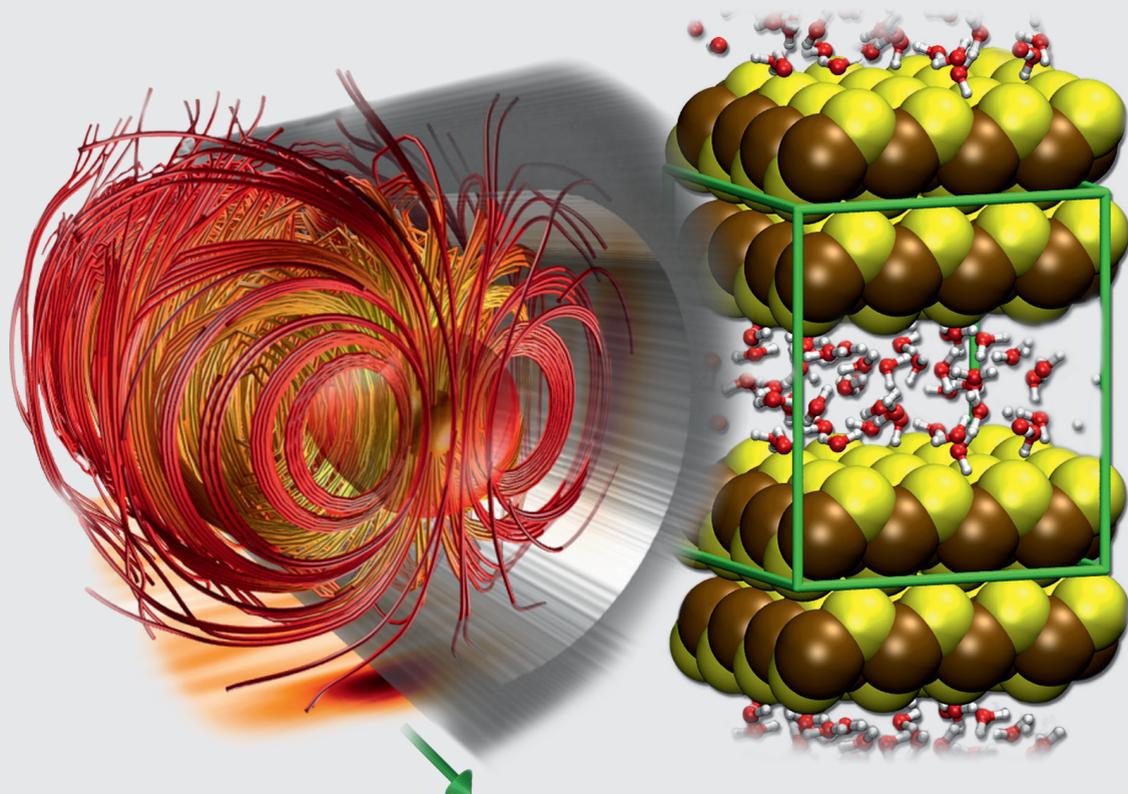


NIC Symposium 2016

11-12 February 2016 | Jülich, Germany

K. Binder, M. Müller, M. Kremer, A. Schnurpfeil (Editors)

Proceedings



Publication Series of the John von Neumann Institute for Computing (NIC)

NIC Series

Volume 48

Forschungszentrum Jülich GmbH
John von Neumann Institute for Computing (NIC)

NIC Symposium 2016

11-12 February 2016 | Jülich, Germany

K. Binder, M. Müller, M. Kremer, A. Schnurpfeil (Editors)

Proceedings

Publication Series of the John von Neumann Institute for Computing (NIC)

NIC Series

Volume 48

ISBN 978-3-95806-109-5

Bibliographic information published by the Deutsche Nationalbibliothek.
The Deutsche Nationalbibliothek lists this publication in the Deutsche
Nationalbibliografie; detailed bibliographic data are available in the
Internet at <http://dnb.d-nb.de>.

Publisher and
Distributor: Forschungszentrum Jülich GmbH
Zentralbibliothek
52425 Jülich
Phone +49 (0) 24 61 61-53 68 · Fax +49 (0) 24 61 61-61 03
e-mail: zb-publikation@fz-juelich.de
Internet: <http://www.fz-juelich.de/zb>

Cover Design: Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH

Printer: Grafische Medien, Forschungszentrum Jülich GmbH

Copyright: Forschungszentrum Jülich 2016

We thank D. Marx and D. Muñoz-Santiburcio (Ruhr-Universität Bochum) as well as E. P. Alves and
L. O. Silva (IST Lisbon) for the images used in the cover design.

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 48

ISBN 978-3-95806-109-5

The complete volume is freely available on the Internet on the Jülicher Open Access Server (JuSER) at
www.fz-juelich.de/zb/openaccess

Neither this book nor any part of it may be reproduced or transmitted in any form or by any
means, electronic or mechanical, including photocopying, microfilming, and recording, or by any
information storage and retrieval system, without permission in writing from the publisher.

Preface

Kurt Binder

Institut für Physik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany
E-mail: kurt.binder@uni-mainz.de

Marcus Müller

Institut für Theoretische Physik, Universität Göttingen, 37077 Göttingen, Germany
E-mail: mmueller@theorie.physik.uni-goettingen.de

Manfred Kremer

John von Neumann Institute for Computing, Jülich Supercomputing Centre,
Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: m.kremer@fz-juelich.de

Alexander Schnurpfeil

John von Neumann Institute for Computing, Jülich Supercomputing Centre,
Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: a.schnurpfeil@fz-juelich.de

The present book accompanies the 8th NIC Symposium, held on February 11 and 12, 2016, at Forschungszentrum Jülich. The John von Neumann Institute for Computing (NIC) has a tradition of biennial meetings, the previous meeting was held on February 12 and 13, 2014. While the first seven NIC Symposia were all held in this century, the NIC builds on the tradition of its predecessor institution, the “Höchstleistungsrechenzentrum (HLRZ)”, which commenced its operation in 1987 at Jülich. Thus, for almost 30 years the Jülich Supercomputing Centre (JSC) has not only served the high performance computing needs of the scientists working at Forschungszentrum Jülich (FZJ), but also provided access to supercomputing resources, as well as the required know-how, to scientists in a broad range of disciplines all over Germany.

The NIC supports research projects from a broad scientific spectrum including topics from Astrophysics, Biology and Biophysics, Chemistry, Elementary Particle Physics, Materials Science, Condensed Matter, Soft Matter Science, Earth and Environment, Computer Science and Numerical Mathematics, Fluid Mechanics, and Plasma Physics. The selection of projects presented in this book can only give a small but representative sample of them. As can be seen from the table of contents, all basic fields of the natural sciences profit from access to supercomputers. One will also note from a look at the presented projects that some of them belong to the field of engineering or have at least a strong motivation from applications.

The High Performance Computing (HPC) infrastructure in Germany is organised in the so-called “HPC pyramid”, which has three layers named tier-1, tier-2, and tier-3, respectively.

Tier-3 comprises computer clusters that locally serve universities and research institutions. Tier-2 comprises regional high performance computers operated for universities and research institutions with special computing needs like for climate research. These regional/topical computing centres are loosely organised in the so-called "Gauss Allianz". Tier-1 is formed by the three national German supercomputing centres, JSC in Jülich, "Leibniz Rechenzentrum" (LRZ) in Garching, and the "Hochleistungsrechenzentrum Stuttgart" (HLRS), which are organised in the "Gauss Centre for Supercomputing" (GCS). At present, the computing power of the supercomputers at these three centres is comparable and amounts to roughly 6 PFlops peak performance at each site, but their architectures follow different designs, to best serve a broad range of requirements. In this spirit, the HLRS supercomputer is most intensively used by projects from the engineering community, whereas the NIC predominately serves the broad natural science community.

The procedures to apply for supercomputing resources and the refereeing process for the proposed projects are closely coordinated by the NIC coordination bureau, in accordance with the GCS governance rules. In particular, the so-called "Gauss large scale projects" (defined, at present, by a request for at least 35 million core hours per year) are handled together by the three GCS sites. For this purpose GCS has established a peer review board. This nucleated from the NIC peer review board, which had been in existence already since the HLRZ was founded in 1987. At present, more than half of the total resources of the three supercomputers (SUPERMUC at LRZ, HORNET/HAZEL HEN at HLRS, and JUQUEEN at JSC) are allocated to these Gauss Large Scale Projects. Due to their extreme computing requirements, it is clear that these projects could not have been carried out at another level of the HPC pyramid. In 2015, the total number of applications for Gauss Large Scale Projects in each granting period (recall that there are two calls for project applications per year, one call for a one-year granting period starting at May 1st and the second for a period starting at November 1st) was about 20. The huge demand clearly demonstrates the urgent need for the available resources.

JSC, as well as HLRS and LRZ, also participated in the "Partnership for Advanced Computing in Europe". PRACE is a consortium of 28 European countries pooling their HPC resources and expertise in order to boost HPC in Europe because this is a key capability for modern technology. All these partners contribute to the management of the infrastructure and services but only the "hosting members" Germany, France, Italy and Spain provide supercomputing resources. The future of PRACE is currently being discussed intensively and is therefore unclear at the time of writing.

In 2014/2015, the situation for high performance computing in Germany has been particularly favourable: both LRZ and HLRS expanded to their full strength, and also at JSC the situation improved, because the JSC had fulfilled its duty to provide one third of its total JUQUEEN resources to PRACE. Therefore, more computing time can be provided to our NIC users.

Another important development is the replacement of the general purpose supercomputer JUROPA by the new system JURECA. After six years of successful and reliable operation, this upgrade increases the performance by about an order of magnitude to 2.2 Pflops. In

view of the great diversity of research fields at NIC, a dual hardware strategy is necessary: top performance is achievable on JUQUEEN, but requires extremely efficient massive parallelisation, which is not achievable for many scientifically very important algorithms and codes. Running these on JUQUEEN would be inefficient and a waste of computing resources. Instead, they can be efficiently run on supercomputers such as JUROPA or JURECA, which imply a more modest level of parallelisation only.

An important activity of NIC is also the establishment and support of NIC research groups, which work on research topics that are of primary interest to the member institutions of NIC. Apart from FZJ these are the Deutsches Elektronen Synchrotron (DESY) in Hamburg and the GSI Helmholtzzentrum für Schwerionenforschung in Darmstadt. The topics of the research groups are: “Computational Materials Science” (headed by Prof. M. Müser, FZJ), “Lattice Gauge Theory” (headed by Prof. R. Sommer, DESY in Zeuthen), “Relativistic Transport Simulation” (headed by Prof. M. Bleicher, GSI) and “Phase Transitions in Lattice Quantum Chromodynamics” (headed by Prof. O. Philipsen, GSI). All NIC research groups have contributed an article to this proceedings volume.

It is a valuable tradition of NIC to award in every year two projects the distinction as “NIC excellence projects”. These projects are identified by the NIC peer review board on the basis of scientifically outstanding results reached in these projects. In 2014 these projects were

“The strong Interaction at Neutron-Rich Extremes” (Prof. Dr. A. Schwenk, Tech. Univ. Darmstadt)

“Numerical Modelling of Lithosphere and Crystal-Scale Deformation on Geological Time Scales” (Prof. Dr. B. Kaus, University of Mainz)

and in 2015

“Structure and Dynamics of Polymers and Lipid Systems” (Prof. M. Müller, Univ. Göttingen)

“Continuous Time Quantum Monte Carlo for Materials” (Prof. A. Lichtenstein, Univ. Hamburg)

Finally, it should be emphasised that much of the excellent scientific output resulting from the use of the NIC supercomputers would not be possible without the very competent and friendly staff of the JSC, who provide very helpful advice to optimise the program codes and the know-how of using the supercomputers very efficiently. It is a real pleasure to thank them sincerely. Also, the administrative staff at NIC always gives competent and friendly advice and help, and similarly deserves the gratitude of the user community.

At this point, it is a sad duty to remember Dr. Walter Nadler, who has been the head of the NIC coordination bureau since 2008, and suddenly passed away on June 9, 2015. In the process of integrating NIC into the GCS structures and improving the application and peer review procedures he has carried out a wonderful job. NIC presently could not function as efficiently as it does without his dedicated and careful and competent work. NIC misses his enthusiasm and his friendly personality and will not forget him.

The editors also want to thank all the authors who have contributed to this book for their valuable effort. We feel that they have succeeded in providing an impressive overview of the wide range of research topics for which supercomputing makes central contributions. We also thank Ms. Martina Kamps who compiled the contributions and produced this high quality book. Finally we want to thank Dr. Florian Janetzko, Ms. Britta Hoßfeld and Ms. Helga Offergeld for their valuable help in organising the 8th NIC Symposium.

Jülich, February 2016

K. Binder

M. Müller

M. Kremer

A. Schnurpfeil

Obituary

Dr. Walter Nadler

† 9 June 2015



On June 9, 2015, Dr. Walter Nadler passed away at the age of 61. Even to those close to him this came as an unexpected shock.

Dr. Walter Nadler was not only an established expert in simulating complex biological systems, he also excelled as a science manager, efficiently and effectively coordinating peer review procedures which had been established to grant Tier-1 and Tier-2 supercomputer resources to different user alliances. He was the driving force in implementing the GCS governance which was agreed upon by the GCS member centres in 2012/13, where he acted as indispensable advisor to the chairmen of numerous compute time granting commissions.

Walter received his PhD in Physics in 1985 at Technische Universität München, supervised by Prof. Klaus Schulten. In the following years he worked at different institutions, including the California Institute of Technology, Wuppertal University and Michigan Technological University. In October 1996 he joined Forschungszentrum Jülich for the first time and became for two years a member of the HLRZ (later NIC) research group Complex Systems, led by Prof. Peter Grassberger. He collaborated scientifically with many groups, among them the Complexity Science Group at Calgary University and the group Cardiac MRT and Biophysics at Würzburg University. Results of the latter collaboration were honoured 2003 with the Helmholtz prize. In 2007 he joined Forschungszentrum Jülich for a second time and became member of the NIC research group “Computational Biology and Biophysics”, headed by Prof. Ulrich Hansmann. Mid 2008 he was appointed head of the newly established NIC coordination office. Since then he focussed on introducing innovative enhancements to the NIC peer review and FZJ-internal allocation procedures. In 2012 an additional regional peer-review process for researchers from Forschungszentrum Jülich and RWTH Aachen University was approved (JARA-HPC partition), an extraordinarily complex undertaking which succeeded largely thanks to the skill and leadership demonstrated by Walter. One year later, the peer-review process for the newly established GCS Large-Scale Projects - defined by the GCS Governance - had to be integrated and implemented, again done by Walter. In the last few months he provided most valuable input to the redesign of JSCs application and peer-review server software.

A role that Walter will also be fondly remembered for both inside and outside JSC is as an enthusiastic, unflappable and fearsomely efficient booth manager: in this capacity

he organized and coordinated the exhibition of the Jülich Supercomputing Centre at the European International Supercomputing Conference (ISC) and the US Supercomputing Conference (SC) every year since 2008. Under his supervision these exhibitions became a cornerstone with respect to the international visibility of the Jülich Supercomputing Centre.

The Jülich Supercomputing Centre and its partner institutions in the Gauss Centre for Supercomputing, the John von Neumann Institute for Computing, the JARA-HPC Vergabegremium and the Vergabekommission für Supercomputer-Ressourcen at Forschungszentrum Jülich will all sorely miss him.

Contents

Scientific Big Data Analytics by HPC <i>T. Lippert, D. Mallmann, M. Riedel</i>	1
The NIC Research Groups	
Relativistic Quantum Molecular Dynamics Simulations of Multi-Strange Particle Production <i>M. Bleicher, J. Steinheimer, H. Petersen</i>	13
Contact Mechanics and Fluid Leakage Near Percolation <i>W. B. Dapp, M. H. Müser</i>	21
The QCD Phase Transition with Two Quark Flavours <i>C. Czaban, F. Cuteri, O. Philipsen, C. Pinke, A. Sciarra</i>	31
Dynamical Simulations of Lattice QCD <i>K. Jansen, S. Schaefer, H. Simma, R. Sommer</i>	39
Astrophysics	
Introduction <i>P. L. Biermann</i>	49
Formation of Star-Forming Clouds from the Magnetised, Diffuse Interstellar Medium <i>R. Banerjee, B. Körtgen</i>	51
The Formation of Planetesimals: Building Bricks for Planetary Systems <i>H. Klahr, A. Schreiber</i>	61
LOFAR: Calibration and Imaging on JURECA <i>M. Hoeft, A. Horneffer, A. Drabent, S. Fröhlich</i>	69
Turbulence and Its Effect on Protostellar Disk Formation <i>D. Seifried, R. Banerjee, R. S. Klessen</i>	77

Computational Biology and Biophysics

Introduction <i>W. Wenzel</i>	85
Structural Predictions of Intrinsically Disordered Proteins with Computational Methods <i>E. Abad, G. Rossetti</i>	89
Towards Restoring Catalytic Activity of Glutamine Synthetase With a Clinically Relevant Mutation <i>B. Frieg, D. Häussinger, H. Gohlke</i>	97
Using Molecular Dynamics to Model the Stacking Behaviour of Perylene Bisimide Derivatives in Aromatic Solvent <i>M. Hollfelder, S. Gekle</i>	105

Computational Chemistry

Introduction <i>C. Peter</i>	115
<i>Ab Initio</i> Molecular Dynamics Simulations of Ionic Liquids <i>M. Thomas, I. Sancho Sanz, O. Hollóczki, B. Kirchner</i>	117
Prebiotic Chemistry in Nanoconfinement <i>D. Muñoz-Santiburcio, D. Marx</i>	125
Simulation of Electron Transfer and Electron Transport in Molecular Systems at Surfaces <i>P. B. Coto, C. Hofmeister, V. Prucker, D. Weckbecker, M. Thoss</i>	133

Elementary Particle Physics

Introduction <i>G. Münster</i>	143
The Mass Difference Between Protons and Neutrons and the Fine Tuning of Physical Constants <i>Sz. Borsanyi, S. Dürr, Z. Fodor, C. Hölbling, S. D. Katz, S. Krieg, L. Lellouch, T. Lippert, A. Portelli, K. K. Szabó, B. C. Tóth</i>	145

The Pseudocritical Line in the QCD Phase Diagram	
<i>H.-T. Ding, P. Hegde, O. Kaczmarek, F. Karsch, E. Laermann, S. Mukherjee, P. Petreczky, H. Sandmeyer, C. Schmidt, P. Steinbrecher, W. Söldner</i>	153
Pushing the Boundaries of Nuclear Physics with Lattice Simulations	
<i>U.-G. Meißner, T. A. Lähde, T. Luu</i>	161
Hadron Structure from Lattice QCD	
<i>A. Schäfer</i>	169
The Strong Interaction at Neutron-Rich Extremes	
<i>A. Gezerlis, K. Hebeler, J. D. Holt, J. Menéndez, A. Schwenk, J. Simonis, I. Tews</i>	179

Materials Science

Introduction	
<i>R. O. Jones</i>	189
Nanostructured Metallic Glasses: Tailoring the Mechanical Properties of Amorphous Metals	
<i>T. Brink, O. Adjaoud, K. Albe</i>	191
Understanding Tribology and Machining Processes through Computationally Intensive Large Scale MD	
<i>P. A. Romero, M. Moseler</i>	199
Memory Effect in Crystallisation of Amorphous $\text{Ge}_2\text{Sb}_2\text{Te}_5$	
<i>R. O. Jones, J. Kalikka, J. Akola</i>	209
Anisotropic Magneto-Thermopower in $M/\text{Co}/M$ ($M = \text{Cu, Pd, Pt}$) Trilayer Systems	
<i>V. Popescu, P. Kratzer</i>	217

Condensed Matter

Introduction	
<i>K. Binder</i>	227
Structures and Phases in (Nano-)Systems in Confined Geometry	
<i>M. Beck, A. Haller, D. Kawatzki, M. Matt, M. Pütz, M. Ring, R. Schmid, K. Scholz, U. Siems, P. Nielaba</i>	229

Quantum Monte Carlo Simulations of Strongly Correlated Electron Systems: The Dimensional Crossover	
<i>M. Raczkowski, M. Bercx, M. Weber, S. Beyl, J. Hofmann, F. Parisen Toldin, M. Hohenadler, F. F. Assaad</i>	241

Optical Spectra of Carbon-Based Nanostructures	
<i>M. Rohlfing</i>	249

Computational Soft Matter Science

Introduction	
<i>K. Kremer</i>	259

Hierarchical Modelling of High-Molecular Weight Polymer Melts: From Soft Blobs to Microscopic Description	
<i>G. Zhang, T. Stuehn, K. Ch. Daoulas, K. Kremer</i>	261

Process-Directed Self-Assembly of Copolymer Materials	
<i>M. Müller</i>	271

Calculation of the Work of Adhesion of Solid-Liquid Interfaces by Molecular Dynamics Simulations	
<i>F. Leroy, F. Müller-Plathe</i>	279

Earth and Environment

Introduction	
<i>U. Hansen</i>	289

Very High Resolution Simulations of African Climate with the Regional Climate Model REMO	
<i>A. Hänsler, N. Koldunov, D. Sein, W. Sauf, D. Jacob</i>	291

Forward and Inverse Modelling of Lithospheric Deformation on Geological Timescales	
<i>B. J. P. Kaus, A. A. Popov, T. S. Baumann, A. E. Püsök, A. Bauville, N. Fernandez, M. Collignon</i>	299

Computational Aspects of High-Resolution Global Gravity Field Determination – Numbering Schemes and Reordering	
<i>J. M. Brockmann, W.-D. Schuh</i>	309

Towards a Better Understanding of Rotating Turbulent Convection in Geo- and Astrophysical Systems	
<i>S. Stellmach, J. Verhoeven, M. Lischper, U. Hansen</i>	319

Computer Science and Numerical Mathematics

Introduction	
<i>D. Kröner</i>	331
Massively Parallel Large Scale Stokes Flow Simulation	
<i>B. Gmeiner, M. Huber, L. John, U. Rüde, C. Waluga, B. Wohlmuth</i>	333
Analysing the Scalability of Climate Codes Using New Features of Scalasca	
<i>M. Harlacher, A. Calotoiu, J. Dennis, F. Wolf</i>	343

Fluid Mechanics

Introduction	
<i>N. A. Adams</i>	353
Numerical Analysis of the FDA Centrifugal Blood Pump	
<i>V. Marinova, I. Kerroumi, A. Lintermann, J. H. Göbbert, C. Moulinec, S. Rible, Y. Fournier, M. Behbahani</i>	355
Direct Numerical Simulations of Impeller Driven Turbulence and Dynamo Action	
<i>S. Kreuzahler, R. Grauer, H. Homann, Y. Ponty</i>	365
Highly-Resolved Numerical Simulation of the Turbulent Combustion Process in Experimental Burners	
<i>F. Proch, M. Rieth, A. Rittler, A. Kempf</i>	373
Supercomputing Studies in Turbulent Rayleigh-Bénard Convection: Challenges and Perspectives	
<i>M. S. Emran, P. Götzfried, A. Kolchinskaya, J. D. Scheel, J. Schumacher</i>	381
On the Parametrisation of Lattice Boltzmann Method in Pore-Scale Flow Simulations	
<i>S. Khirevich, U. Tallarek</i>	389

Computational Plasma Physics

Introduction

K.-H. Spatschek

399

Towards Plasma-Driven Free-Electron Lasers

*I. Dornmair, L. T. Campbell, J. T. Henderson, S. Jalas, O. Karger, M. Kirchen,
A. Knetsch, G. G. Manhan, G. Wittig, B. Hidding, B. W. J. McNeil, A. R. Maier*

401

Collisionless Shocks and Mediating Instabilities

A. Stockem Novo, P. Alves, K. Schoeffler, L. O. Silva, R. Schlickeiser

409

Scientific Big Data Analytics by HPC

Thomas Lippert, Daniel Mallmann, and Morris Riedel

Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany
and JARA-HPC

E-mail: {th.lippert, d.mallmann, m.riedel}@fz-juelich.de

Storing, managing, sharing, curating and especially analysing huge amounts of data face an immense visibility and importance in industry and economy as well as in science and research. Industry and economy exploit “Big Data” for predictive analysis, to increase the efficiency of infrastructures, customer segmentation, and tailored services. In science, Big Data allows for addressing problems with complexities that were impossible to deal with so far. The amounts of data are growing exponentially in many areas and are becoming a drastical challenge for infrastructures, software systems, analysis methods, and support structures, as well as for funding agencies and legislation.

In this contribution, we argue that the Helmholtz Association, with its objective to build and operate large-scale experiments, facilities, and research infrastructures, has a key role in tackling the pressing Scientific Big Data Analytics challenge. DataLabs and SimLabs, sustained on a long-term basis in Helmholtz, can bring research groups together on a synergistic level and can transcend the boundaries between different communities. This allows to translate methods and tools between different domains as well as from fundamental research to applications and industry. We present an SBDA framework concept touching its infrastructure building blocks, the targeted user groups and expected benefits, also concerning industry aspects. Finally, we give a preliminary account on the call for “Expressions of Interest” by the John von Neumann-Institute for Computing concerning Scientific Big Data Analytics by HPC.

1 Introduction

The Helmholtz Association develops and operates large-scale infrastructures and makes them available for scientists, research groups, and communities. The effective usage of these infrastructures is ensured by scientific peer-review. These science-led processes not only guarantee the most beneficial usage of the infrastructures, but also steer their evolution and focus through the involvement of research communities in key areas of science and engineering.

The John von Neumann Institute for Computing (NIC) coordinates the scientific peer-review process for the provision of supercomputing cycles at the Jülich Supercomputing Centre (JSC). Scientists and researchers who apply for computing time on these systems are supported by a continuously growing number of domain-specific Simulation Laboratories (SimLabs) at JSC. The SimLabs offer support on a high level and push forward research and development in co-design together with specific scientific communities that take advantage of High Performance Computing (HPC) techniques in massively parallel applications.

In order to address the challenges of *Scientific Big Data Analytics (SBDA)* it is essential to enhance the central elements (cf. red box in Fig. 1) as research, development of infrastructures, and especially the provisioning through a scientific peer-review process. The user support through SimLabs and DataLabs, who have an integrating role, needs to be expanded within Helmholtz as well as to communities outside Helmholtz.

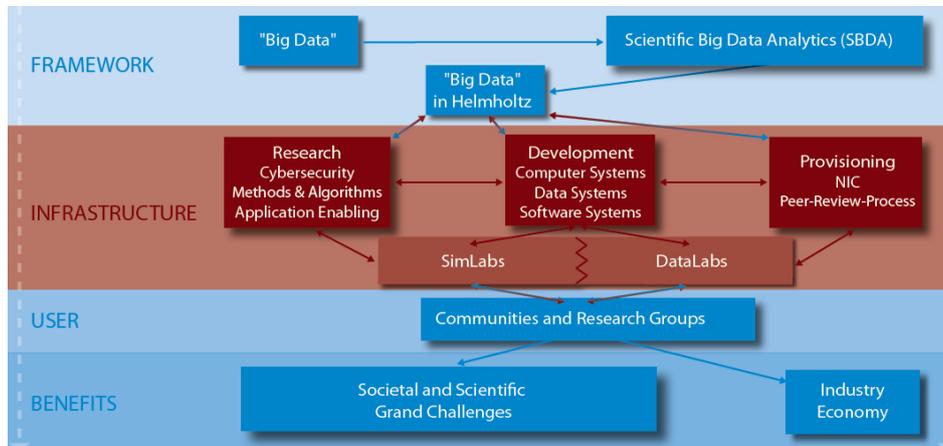


Figure 1. “Scientific Big Data Analytics” and the role of lage-scale research.

The remainder of this paper is structured as follows. After the introduction, the scene is set by clarifying the terms “Big Data” with a focus on “Analytics” as a framework concept for scientific environments in general and Helmholtz in particular. Sect. 3 describes the core building blocks of the infrastructure that are necessary to implement this framework concept, while Sect. 4 provides information on relevant communities and research groups. Sect. 5 then outlines benefits for research and industry when implementing the framework concept. Sect. 6 gives insight into the initial steps already undertaken to realise parts of the framework and its implementation using a couple of case studies as examples.

2 Big Data in Science and Engineering

2.1 On Understanding “Big Data”

Simple definitions of “Big Data” refer to data sets that become so large that it is impossible to process them using traditional methods. This definition leaves many questions unanswered such as “what are traditional methods” or “what exactly means processing in this context”. This is in particular the case when one considers the difference of processing and traditional methods in science and industry. More recently, the term “Big Data” is often used to refer to data challenges with an increasing number of the (in)famous “Vs” such as “Variety”, “Volume”, “Velocity”, “Veracity”, “Variability”, and most notably “Value” that altogether emphasise the fact that the “Volume” alone is not necessarily the only problem.

Economy and industry have a variety of applications for “Big Data”: personalised marketing and product offerings, product development according to trends, infrastructure expansion based on usage statistics, or better logistics are only a few examples. For these purposes the customer behaviour is analysed and used to e.g. personalise advertisements, or to optimise logistics chains of supermarkets, or the infrastructure of wireless carriers. The goal is to deliver the best product or best service literally tailored for a customer at minimal cost. This is often achieved by focusing on “correlation” and related statistical

models that indicate a pattern in datasets. For product development, additional data is used, e.g. material properties, production costs, or measurement data obtained from production machines to help developing innovative products and to remain competitive in the market.

In science and research, Big Data in the context of our framework illustrated in Fig. 1 means, that based on data analysis, complex questions are dealt with, models are refined, or new models are developed. In contrast to just analysing Big Data for correlations and trends (cf. Google Flu Predictions¹) as often done in industrial research questions, science and research need to focus on *causality* in order to prove or disprove a specific hypothesis including also the reproducibility of scientific findings (cf. “lack of Google Flu predictions”²).

In all scientific areas, the amount of data is increasing and the addressed problems become more complex. This is also due to the impact of “open scientific data” providing much more opportunities for data fusion by using multiple datasets in one scientific use case. Unfortunately, current state-of-the-art processing technologies can only handle a few TeraBytes. Beyond these limitations, methods of parallel data management and high performance computing need to be exploited. The largest data sets that have to be analysed in leading projects require to process several PetaBytes. It is almost impossible to handle data sets of this size in universities or research institutes. This task is reserved for large-scale research institutions and research infrastructures.

2.2 Scientific Big Data Analytics (SBDA)

The field of Big Data is strongly influenced by books such as “Big data: A revolution that will transform how we live, work, and think”³ as well as commercial use cases such as “Industry 4.0”, whereby factories of the future are automatically built and optimised by their machines that continuously measure and exchange data or industrial case studies about predictive maintenance, potentially leading to massive reductions in operating costs of large machines. The enormous momentum in media and industry leads to the development of a wide variety of technologies that constantly change and thus tends to impede stable algorithm development (cf. the relatively fast move from map-reduce⁴ solutions based on Apache Hadoop⁵ 1.0, 2.0, and more recently Apache Spark⁶).

Some of our latest findings reveal that many solutions based on these recent technologies cannot offer effective algorithmic solutions required by science. Examples include classification algorithms like support vector machines (SVMs)⁷ as shown in Cavallaro *et al.*⁸ or clustering algorithms like density-based spatial clustering of applications with noise (DBSCAN)⁹ as shown by Goetz *et al.*¹⁰. Due to constant change of the basis technologies and because of the significant lack of yet to be developed or commonly adopted standards for remote and distributed processing, a community agreed algorithm code (e.g. as often seen in the simulation sciences – based on the mature and strong message passing interface – such as numerical weather prediction codes¹¹) is thus hard to establish. The constant changes also lead to resource usage focusing on demonstrating new technology options rather than algorithms supporting deep scientific hypothesis analysis and testing. We strongly believe that a community-led and accepted peer-review in the field of SBDA can bring a major consolidating effect to the development of this technology field in general and could push its technical maturity in particular.

Researchers are constantly distracted from exploring new and partly innovative technologies (with often rather similar functions for a variety of scientific problems) and in too

many cases lack a sound infrastructure for developing solutions beyond small-scale technology islands or simple testbeds. The notion *Scientific Big Data Analytics* comprises the work of researchers and engineers, that is based on the creation and analysis of big data sets, and thus relies on the availability of appropriately sized infrastructures, in order to be competitive in the respective domain. It is necessary to provide large-scale infrastructures to scientists and engineers of universities and research institutes, who perform projects with highest demands on storing, transfer, and processing of data sets. The provisioning should be done similar to the provisioning of HPC resources as done for the simulation sciences since many years. This constitutes a key element in the SBDA building block in our framework shown in Fig. 1.

To guarantee that the data analysis achieves the highest scientific quality, it is necessary to apply the principle of autonomous controlling of resource allocation by science in a competitive peer-review process, like it is common practice for international large-scale infrastructures. In addition, the scientific controlling of resource allocations will allow to focus on problem areas that are highly relevant for science and society. The steering process prevents that science gets lost in the details of this industry-driven topical area, with many technologies that are highly relevant only for industry (e.g. recommender engines, shopping basket association rule mining, etc.). Instead, new approaches will be developed and, subsequently have to be translated to economy and industry. Scientific approaches will mature, leading to community-approved codes to tackle an ever increasing amount of research questions.

2.3 Helmholtz-Specific Elements

The framework as outlined in Fig. 1 puts strong emphasis on the Helmholtz association being in a key position to realise SBDA. Firstly, in its role as operator of infrastructures, like supercomputers and data storages of the highest performance category, secondly, as leading organisation for research and development in this domain, and thirdly, with its internationally acknowledged competence in the scientifically controlled allocation of resources of large-scale instruments, not to forget its experts from a variety of application domains. Given these strong competences, Helmholtz is able to push the field forward and help to achieve new insights for science and society. To give an example, solving a wide variety of inverse problems from the field of Big Data as available in Helmholtz centres can actually lead to better algorithms for simulation sciences that in turn then deliver more accurate models to understand our world. In order to establish this strong productive loop between HPC simulations and more data-intensive applications, a strong foundational infrastructure is required with technologies as mature as seen today in the field of scientific computing.

Fig. 2 illustrates one of the key capabilities of Helmholtz we refer to as “full productive loop for SBDA”. Big Data is already an integral part of many Helmholtz centres and most centres have demands for data storage, management, and data analysis, as it was already outlined in a Helmholtz concept paper 2012 (“Further Development of High-Performance Computing and Big Data Management in the Helmholtz Association”). The user demand is the key driver for the Helmholtz Association as a whole. But in contrast to many other scientific associations, Helmholtz can as well provide the large-scale resources that are required in order to tackle Big Data and its challenges – with sustained superior performance.

The leading organisation in the Helmholtz Association at the forefront of simulation and data intensive computing is the John von Neumann-Institute for Computing (NIC), a long-standing partnership of the research centres Jülich, GSI and DESY.

3 Infrastructure

3.1 Infrastructure Research

The analysis of scientific data sets is characterised by well established methods and algorithms from mathematics and computer science, which evolved over decades, and which now reach their limits with the exponentially growing demands of Big Data. This is especially true for the scaling of algorithms and the throughput in data analysis. Core principles, like parallelisation, can be adapted from other areas, like HPC. Examples include the use of “halo/ghost areas” in smart domain decompositions or the use of parallel and scalable I/O. In this manner, several limitations can be overcome. Of course, new methods have to be developed, and in parallel, the process for scientific and engineering application enabling must be established in the infrastructures. Concrete issues will be solved in co-design with industrial partners. Further challenges in the area of data privacy and security are closely coupled to the research area cybersecurity, and SBDA will be a central use case for this. This kind of infrastructure research is a first core building block of our framework implementation shown in Fig. 1.

3.2 Infrastructure Development

The continuous development of compute and data systems is essential for the continuous advancement of the infrastructure, which simultaneously must offer stable operation and

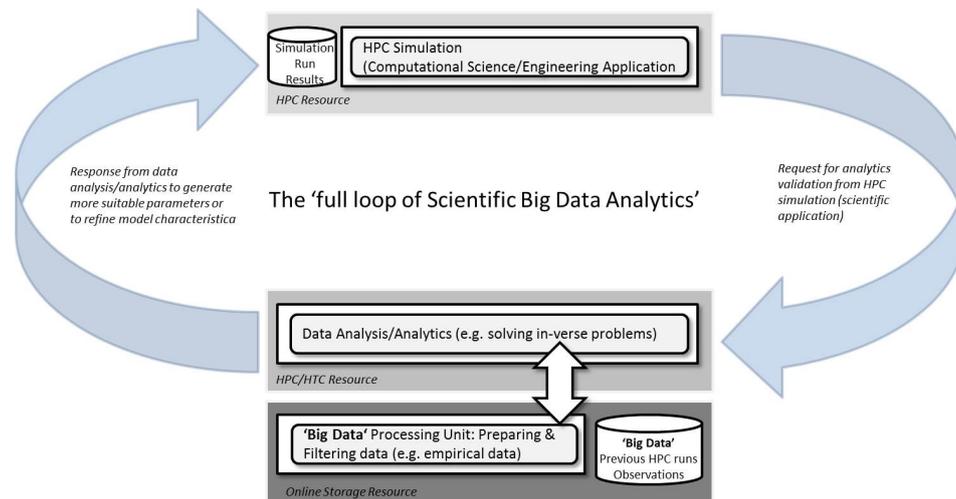


Figure 2. The “full productive loop of SBDA”.

usage, in order to allow researchers and engineers to stay competitive in their domain. Big Data currently is shaped by many new approaches for software systems, which are too disjunct to offer efficient and stable usage. Several of these approaches are driven by technologies originating in the fast moving industrial sectors as outlined above (Hadoop 1.0, Hadoop 2.0, Google Dataflow, Apache Spark, or Hana and Teracotta for in-memory databases). The usability of these software systems for scientific applications needs to be evaluated and compared to well established approaches (e.g. message passing and/or parallel I/O in simulation sciences), and, if suitable, must be further developed. The development of the provided compute, data, and software systems in community and multi-disciplinary compute and data centres will be driven by the demands of the scientific applications.

3.3 Infrastructure Provisioning

In order to reach the highest scientific quality in data analysis and the greatest possible advancement in research and development in SBDA by HPC, it is crucial to establish the scientific competitive peer-review process for the allocation of resources as shown in the third core building block in Fig. 1. A well known institution for scientifically controlled allocation of resources is the Scientific Council of the John von Neumann-Institut for Computing (NIC), who allocates a major part of the computing time of the supercomputers operated by JSC to scientists in Helmholtz and throughout Germany. The NIC allocation principles served as a blue print for the allocation of computing time in the Gauss Centre for Supercomputing (GCS) as well as the Partnership for Advanced Computing in Europe (PRACE). Given NIC's strong experience over many years and its scientific advancement, it is natural to apply the NIC provisioning concept to data infrastructures and analysis resources. This kind of provisioning together with the activities in the Helmholtz programme "Supercomputing & Big Data", will promote research and development for supercomputing and data infrastructures. It is at the heart of truly innovative SBDA.

3.4 Data- and SimLabs

The efficient usage of these infrastructure will be made possible through interlocking the domain-specific SimLabs for scientific computing with domain-specific DataLabs. In this respect, it is important to realise that data analysis, and simulation are not separable. This fact is clearly outlined in one of the recent reports of US Department of Energy that suggests that "computing gets more intertwined with data analysis"¹². Many scientific projects show that there is no border between simulation and data analysis, because the HPC-driven data modelling is mandatory for solving grand challenge problems as shown in Fig. 2 while data analysis is equally important for empirical research and theory. While numerical simulations with HPC, based on physical laws or engineering models are key of bridging experiment and theory, SBDA and its full loop provide the "breakwaters for HPC" to safeguard these bridges from collapsing while faced with an ever increasing number of big data waves.

What is more, DataLabs and SimLabs as shown in Fig. 1 have an important role in establishing SBDA as a central element with international dimension promoted by their members. In order to assemble teams with highest competence, the members need to be

recruited based on international reputation and scientific standing. More notably, they have to be enabled to link with their scientific communities in the best possible way. Ideally, as shown in SimLabs by supporting and driving community codes, key scientific and engineering data-intensive algorithms, mostly developed in international collaborations, must be supported and further developed within such DataLabs.

4 Communities and Research Groups

It is the users of the infrastructures, scientific communities and research groups, that define which methods, algorithms, and approaches, are relevant for their domain. Those methods, algorithms, and approaches are the ones to be further developed and improved and are selected as part of the infrastructure provisioning. The SimLabs and DataLabs as installed by the JSC in recent years are the liaison teams that can coordinate the development and can push the transfer of algorithms and tools into the infrastructure. In order to accomplish this objective, SimLabs and DataLabs must be integrated into the respective research communities. The latter condition is of key importance since many techniques and solutions to Big Data challenges will be developed by geographically dispersed research teams, that often lack proper communication and exchange on the synergistic level.

Analysing JSC's SimLabs activities reveals that several of them are already working on data-intensive aspects. The JSC will strongly foster data science aspects in the existing SimLabs. One can expect to see quite some impact in the following communities served by JSC's Sim/DataLabs in the next years: terrestrial systems, climate sciences, neuroscience, computational biology, molecular systems, to name a few. Examples of fields, that today demonstrate these trends, can be found in the earth sciences (e.g. hyper-spectral earth observation image data) and in the medical sciences (e.g. high-resolution medical image data).

An important driver of SBDA activities will be the availability of publicly shared datasets. Many large funding organisations (e.g. the European Commission) or major community journals are pushing in the direction to publish not only scientific findings in a journal, but also to make associated datasets publicly available. The technologies to enable full reproducibility of paper findings from such datasets are still in their infancy, but are continuously developed. These technologies and the subsequent larger public data access will profoundly influence the activities of communities and research groups in data-intensive computing in the next decades.

What is more, one can foresee an increasing number of scientific cases where the availability of data across different scientific domains will provide so far unknown research opportunities. This demands that SimLabs and DataLabs from different domains should not work in isolation from each other, but may tackle joint Big Data challenges. Examples would be in the area of earth sciences where climate challenges directly or indirectly influence modelling and data-intensive computing of terrestrial systems.

5 Benefits

5.1 Grand Challenges

The framework sketched for large-scale research including infrastructure activities and users as shown in Fig. 1 will allow to tackle societal and scientific grand challenges in

innovative ways (e.g. solving inverse problems jointly with HPC simulations). In an initial survey throughout the Helmholtz Association, 14 domains were identified including one or more Helmholtz centres each, which encounter major challenges in their scientific data analytics. Among others those domains are climate research, epidemiology, environmental diseases, water research, or polar and marine research. Additional domains and research areas will be identified in the context of research groups at universities and research institutes.

It should be emphasised that methodologies, algorithms, tools, and analysis procedures originating in fundamental research very often are the basis to support solutions of grand challenges pursued in the applied sciences. The interlinked SimLab-DataLab framework is a catalyst to realise this transfer. Examples from energy are areas like Smart Grids, electric mobility, and battery research, in health sciences the DataLab Neuroscience supports, for instance, the development of the three-dimensional model of the human brain.

5.2 Benefits for Industry and Economy

Industry and economy in general analyse data in order to solve issues different from scientific research. However, methodologically, both sides often need to use the same algorithms and tools. The activities in research and development will influence the activities in industry and economy through improved algorithms and methods developed in joint research. Small and medium enterprises can benefit, because they can overcome scaling issues they face while expanding without the need to employ their own data science teams. In addition Sim- and DataLabs are supposed to also support industry and economy through targeted contract research for a limited period, coordinated by a so-called industry hub currently put up by the JSC.

Moreover, SBDA activities have been started such as the German initiative “Smart Data Innovation Lab”¹³. Here, data-intensive computing challenges are tackled in the four domain areas energy, industry 4.0, smart cities, and personalised medicine. Research organisations in Germany with quite a few companies formed this collaboration in order to jointly tackle problems in industry and economy. While the infrastructure is just getting started to be used by selected scientific cases, one can already observe many options for improvements taking into account major arguments of this paper. Most notably, the current infrastructure is driven by High Throughput Computing (HTC) needs and technologies while solutions based on HPC are equally important to be considered in order to consistently provide benefits for industry and economy, not to forget the quality assuring principle of peer-review.

6 Scientific Big Data Analytics by HPC – Expression of Interest

In order to gain better insights into the demand by communities and requirements for SBDA, the JSC has performed an initial step towards implementing principles as proposed in this paper. The importance of data analytics, management, sharing, and preservation of very big, often heterogeneous or distributed data sets from experiments, observations and simulations is of increasing significance for science, research and industry. This development has been recognised by many research institutions, among them leading HPC centres. They want to advance their support for researchers and engineers using SBDA by HPC.

At the beginning of 2015, the John von Neumann Institute for Computing (NIC) has, for the first time, invited Expressions of Interest (EoI) for SBDA projects using HPC to identify and analyse the needs of scientific communities. The goal is to extend and optimise the HPC and data infrastructures and services in order to provide optimal methodological support. As described above, a peer-review was performed on the EoI submissions in a similar manner as known from HPC calls. The EoIs have been submitted from various domains. They clearly demonstrate in which areas SBDA by HPC is of major importance:

EoI submissions have been received in the field of Biology where HPC knowledge mining of molecular structure data was considered to be an interesting challenge for SBDA. These contributions already showed how intertwined approaches for SBDA, using machine learning and statistical data mining on the one hand, can be successfully combined with HPC technologies on the other hand. This fact is further underpinned with evidence by two more EoI submissions related to HPC simulations: One is on the statistical analysis of high-Rayleigh number turbulent convection data while the second outlines Big Data challenges and opportunities in the area of turbulence databases from direct numerical simulations.

Two EoI submissions in the area of earth sciences are related to each other. Big Data challenges from the SimLab Climate Science raises the demand for a joint atmospheric data repository and processing unit in order to advance in SBDA while the SimLab Terrestrial Systems outlines challenges for the implementation of a high-performance big data storage and analysis framework for large-scale earth science simulations.

A further EoI submission in the field of neurosciences raised the request for the availability of more parallel and scalable SBDA algorithms such as deep learning and unsupervised clustering for analysis of cellular cortical structures in the human brain. While this contribution was focused on statistical data mining aspects, a further EoI contribution described SBDA challenges to solve inverse problems by big data evaluations for a high complexity energy meteorological *in situ* analytics application. Finally, turning to particle physics, a lattice QCD submission is highlighting several Big Data challenges in the post-processing steps of a large-scale simulation project.

7 Conclusions

The first experiences with SBDA by HPC in the EoI call of the John von Neumann Institute for Computing tell that there is an urgent demand for an SBDA framework concept as provided in this paper. The Helmholtz infrastructure as outlined promises to implement SBDA by HPC in a sustainable manner, with SimLabs and DataLabs as well as autonomous scientific peer-review by community representatives as a cornerstone.

Acknowledgements

We commemorate the late Walter Nadler who was invaluable in forming the EoI call for SBDA by HPC in the John von Neumann Institute for Computing. Walter, we will not forget you!

References

1. J. Ginsburg, M. Mohebbi, R. Patel, L. Brammer, M. Smolinski, L. Brilliant, *Detecting influenza epidemics using search engine query data*, Nature 457, 2009.
2. D. Lazer, R. Kennedy, G. King, and A. Vespignani, *The Parable of Google Flu: Traps in Big Data Analysis*, Science Vol (343), 2014.
3. V. Mayer-Schoenberger and C. Kenneth, *Big data: A revolution that will transform how we live, work, and think*, Houghton Mifflin Harcourt, 2013.
4. J. Dean and G. Sanjay, *MapReduce: simplified data processing on large clusters*, Communications of the ACM 51.1 (2008): 107-113.
5. D. Borthakur, J. Sen Sarma, J. Gray, *Apache Hadoop goes realtime at Facebook*, Proceedings of the 2011 ACM SIGMOD International Conference on Management of data, ACM, 2011.
6. S. Ryza, U. Laserson, S. Owen, J. Wills, *Advanced Analytics with Spark: Patterns for Learning from Data at Scale*, O'Reilly Media, ISBN 1491912766, 2015.
7. C. Cortes and V. Vapnik, *Support-vector networks*, Machine Learning, vol. 20(3), pp. 273-297, 1995.
8. G. Cavallaro, M. Riedel, M. Richerzhagen, J. A. Benediktsson, A. Plaza, *On Understanding Big Data Impacts in Remotely Sensed Image Classification Using Support Vector Machine Methods*, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Issue 99, pp. 1-13, 2015.
9. M. Ester *et al.*, *A density-based algorithm for discovering clusters in large spatial databases with noise*, Kdd. Vol. 96, 1996.
10. M. Goetz, M. Richerzhagen, G. Cavallaro, C. Bodenstein, P. Glock, M. Riedel, J. A. Benediktsson, *On Parallel and Scalable Classification and Clustering Techniques for Earth Science Datasets*, Sixth Workshop on Data Mining in Earth System Science (DMESS 2015), International Conference on Computational Science (ICCS), Reykjavik, 2015.
11. J. Dudhia *et al.*, *The weather research and forecast model: software architecture and performance*, Proceedings of the Eleventh ECMWF Workshop on the Use of High Performance Computing in Meteorology, 2005.
12. US DOE ASCAC Report, *Synergistic Challenges in Data-Intensive Science and Exascale Computing*, 2013.
13. Smart Data Innovation Lab (SDIL), German Initiative,
Online: <http://www.sdil.de>

The NIC Research Groups

Relativistic Quantum Molecular Dynamics Simulations of Multi-Strange Particle Production

Marcus Bleicher^{1,2,3}, Jan Steinheimer¹, and Hannah Petersen^{1,2,3}

¹ Frankfurt Institute for Advanced Studies,
Ruth-Moufang-Straße 1, 60438 Frankfurt am Main, Germany
E-mail: bleicher@fias.uni-frankfurt.de

² GSI Helmholtzzentrum für Schwerionenforschung, Planckstraße 1, 64291 Darmstadt, Germany

³ Institut für Theoretische Physik, Goethe Universität,
Max-von-Laue-Straße 1, Frankfurt am Main, Germany

The Ultra-relativistic Quantum Molecular Dynamics Model (UrQMD) has been updated and applied to explore rare particle probes in nuclear collisions. Such simulation tools allow to explore the dynamics of the strongly interacting matter (i.e. matter described by Quantum Chromodynamics, QCD). The model was supplemented with novel reaction channels to explore the production of multi-strange particle probes in heavy ion collisions. The extended model has been used for simulations of nuclear collisions at beam energies in the 1-2 GeV/nucleon domain to explore the properties and particle production mechanisms in the dense QCD medium created.

1 Introduction

Quantum Chromodynamics (QCD), sometimes also called the strong interaction is the fundamental theory governing the interactions in the atomic nucleus, i.e. between the protons and neutrons. Also the neutrons and protons themselves are build up from strongly interacting elementary particles, called the quarks and the gluons. While lattice QCD allows to obtain *ab initio* results based on the Lagrangian of QCD, this approach is restricted to static equilibrium conditions and does not allow to model dynamical processes. Unfortunately, dense and hot QCD matter is usually created by the collision of heavy ions, e.g. Gold on Gold collisions, producing a transient state that finally decays into “stable” particles that can be detected by the experiment. Therefore it is necessary to connect the lattice QCD results with a dynamical evolution model of the matter to allow for a comparison with the experimental observables. A well established approach to simulate the evolution of QCD matter is the Ultra-relativistic Quantum Molecular. This model combines relativistic Boltzmann dynamics for the initial and final stages of the reaction with a hydrodynamic model for the hot and dense stages of the reaction. At lower energies, however, a decoupling of the hydrodynamic stage and the Boltzmann phase is not possible and a direct Quantum Molecular Dynamics simulation is employed to integrate the equations of motion. Within the last years the simulations were updated^{1,2} to describe rare channels of multi-strange particle production to gain important insights into the creation mechanisms of these states and the properties of the matter created.

In Sec. 3 the UrQMD model will be introduced and the main modifications will be outlined. Sec. 4 introduces the topic of subthreshold particle production. Finally, Sec. 5 will then discuss the results for the multi-strange particle production and provide a comparison to the HADES data. The paper will finally conclude and provide some outlook.

2 Motivation

QCD matter under extreme conditions is created in today's largest accelerator facilities, namely the Large Hadron Collider (LHC) and the Super Proton Synchrotron (SPS) at CERN. At lower energies, the relativistic heavy ion collider RHIC at Brookhaven allows to explore the energy excitation function of observables, while GSI and FAIR at Darmstadt will map out the low energy and high density regime. The systems created in such collision have a typical life time of 10^{-23} seconds and temperatures above a billion degrees, the densities are similar to those expected in neutron stars. The measured particle spectra generally only allow to reconstruct the emission momenta of the particles and their mass after most of the produced particles have already decayed. A major task is to relate these particle spectra to the properties of the matter, e.g. the temperature and density (the Equation of State), the degrees of freedom during the hottest and densest time and other properties, e.g. viscosities and in-medium spectral functions of hadrons.

Recently two kinds of rare probes have gained a lot of attention in the GSI energy regime: I) multi-strange hadron production, i.e. the production of exotic particles consisting of strange quarks like the Φ meson (a strange-anti-strange state, $s\bar{s}$) or the Ξ baryon, a sss state. II) leptonic probes, here specifically pairs of electrons that are emitted from the reaction zone. The HADES experiment at GSI has recently observed unexpectedly high yields of multi-strange Ξ baryons in Ar+KCl reactions at 1.75 GeV beam energy. This observation could not be explained in terms of statistical models or state-of-the-art transport simulations. A similar situation arose in the recent past with the production of the multi-strange Φ meson in the same energy range. The observed strong enhancement of both hadron species can now be explained and will allow to probe multi-step processes and new production channels as will be detailed below.

A better understanding of strange particle production is not only relevant for elementary particles, but also serves as input for simulations of exotic hyper matter states, e.g. $^{\Lambda}$ He and heavier states³⁻⁵.

3 The UrQMD Model

UrQMD is based on the solution of a coupled set of Boltzmann equations for the time evolution of each hadron species. The main inputs for the simulation are the properties of the individual hadrons supplemented by the interaction cross sections. The cross sections are either fitted to available data, calculated from effective models or taken from the additive quark model. Hadron production in the UrQMD transport model^{6,7} proceeds through different channels: The excitation and de-excitation (decay) of hadronic resonances, the excitation and de-excitation of a string and the annihilation of a particle with its anti-particle. Strangeness exchange processes, which can change the flavour content of a hadron are also included¹. The probabilities of the different processes are governed by their reaction cross sections. These cross sections serve as input for the model and are, whenever possible, taken from experimental measurements of elementary (binary) collisions. For example the total and inelastic cross section of binary proton+proton collisions has been measured in many experiments over a wide range of beam energies⁸. The difference between the total and elastic cross section therefore should correspond to the inelastic cross section.

Here again, many different reactions are possible. In UrQMD the inelastic part of the nucleon+nucleon cross section (up to a certain energy) is described by resonance production channels. The possible channels of resonance excitations are divided into several classes:

1. $NN \rightarrow N\Delta_{1232}$
2. $NN \rightarrow NN^*$
3. $NN \rightarrow N\Delta^*$
4. $NN \rightarrow \Delta_{1232}\Delta_{1232}$
5. $NN \rightarrow \Delta_{1232}N^*$
6. $NN \rightarrow \Delta_{1232}\Delta^*$
7. $NN \rightarrow R^*R^*$

Here R^* could be any excited N^* or Δ^* state. Since a large part of the channels are not known, or only measured within a limited energy interval one uses an effective parametrisation of the different cross sections.

4 Sub-Threshold Particle Production

If the centre-of-mass energy of the collision is below the threshold for the production of a certain hadron type, these interaction are called sub-threshold. Sub-threshold collisions are therefore prime testing grounds for multiple interactions. When discussing sub-threshold production of ϕ 's and Ξ 's in nuclear collisions one should note that there are two distinct mechanisms which allow for the production of hadrons with masses, higher than what would be energetically forbidden in elementary reactions:

1. One is the fact that in a nucleus, the nucleons acquire a Fermi momentum due to their bound state. Because of the Fermi momenta, the actual energy of two colliding nucleons will not be exactly the beam energy but a smeared out energy distribution. This allows for collisions of nucleons at energies higher than the actual beam energy.
2. Furthermore energy can be accumulated due to secondary interactions of already excited states, produced earlier in the collision^{9,10}.

As an example for deep sub-threshold production of multi-strange hadrons we will first investigate the production probability of resonance states with sufficiently high mass to produce a ϕ or Ξ in collisions of Ca+Ca (corresponding to the Ar+KCl collisions studied at the HADES experiment). In such collisions, from the Fermi momenta alone, about two percent of all primary $N + N$ collisions will have an invariant mass large enough to produce a ϕ meson, while essentially none has sufficient energy to produce a Ξ . Even though a small fraction of initial collisions has sufficient energy to produce a ϕ , none can be produced at this energy through a string excitation because in the string picture a ϕ can only be created together with a $K + \bar{K}$ pair, increasing the effective threshold for ϕ production by a string fragmentation by an additional 1 GeV.

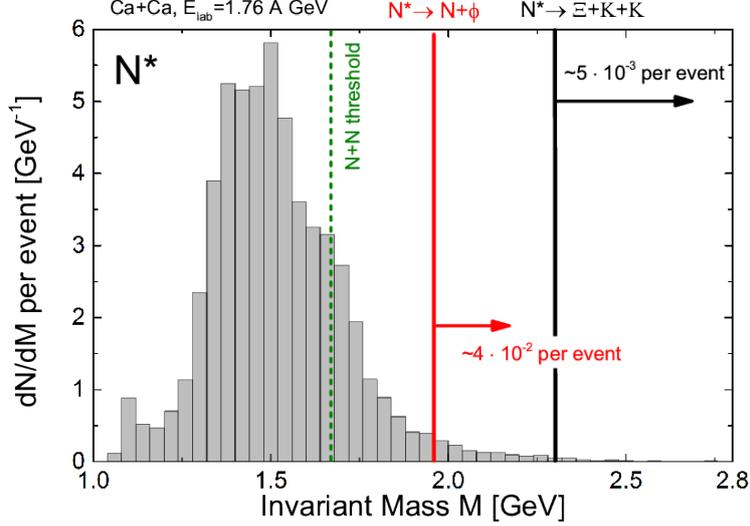


Figure 1. Invariant mass distribution of N^* resonances produced in collisions of Ca+Ca at a fixed target beam energy of $E_{lab} = 1.76$ A GeV. We consider events with an impact parameter smaller than $b < 5$ fm, in accordance with HADES experiment specifications. The vertical green dashed line indicates the maximum mass a N^* can have in an elementary $N + N$ collision at the same beam energy. The vertical red line depicts the ϕ production threshold mass while the black line corresponds to the $\Xi + K + K$ threshold mass².

Fig. 1 shows the calculated invariant mass distributions of N^* baryons produced in Ca+Ca collisions at a fixed target beam energy of $E_{lab} = 1.76$ A GeV and a centrality range of $b < 5$ fm. The vertical green dashed line indicates the maximal N^* mass possible in elementary NN reactions at the same beam energy. It is clear from this figure that a substantial number of N^* resonances with masses larger than the apparent threshold energy is produced in the nuclear collision. We also indicate, as vertical lines, the minimal mass a N^* would need in order to decay into a final state with a ϕ (red line) or even Ξ (black line). From this distribution we conclude that a moderate amount of excited states with sufficiently high mass are available that may produce ϕ mesons as well as Ξ baryons. In the following we explore decays of the most massive N^* resonances implemented in UrQMD, namely the $N^* \rightarrow N + \phi$ and $N^* \rightarrow \Xi + K + K$ channels, in order to describe the production of ϕ and Ξ particles near and below their elementary threshold energies. In particular we will use the $N^*(1990)$, $N^*(2080)$, $N^*(2190)$, $N^*(2220)$ and $N^*(2250)$ states included in the UrQMD model, as their decay channels are experimentally not well constrained and they have a sufficiently large mass.

Using ANKE data¹¹ we find, that a branching fraction of $\frac{\Gamma_{N+\phi}}{\Gamma_{tot}} = 0.2\%$, for all the above mentioned N^* resonances, provides a very good description of the measured ϕ production cross section in proton+proton reactions at ANKE. The new HADES data on Ξ production in p+Nb reactions is used as a proxy for the unavailable elementary collision data to fix the $N^* \rightarrow \Xi + K + K$ branching fraction. To describe the measured production yield of Ξ^- 's we obtain a branching fraction $\frac{\Gamma_{\Xi+K+K}}{\Gamma_{tot}} = 10\%$ for all N^* states mentioned above, that have a sufficiently high mass for this decay.

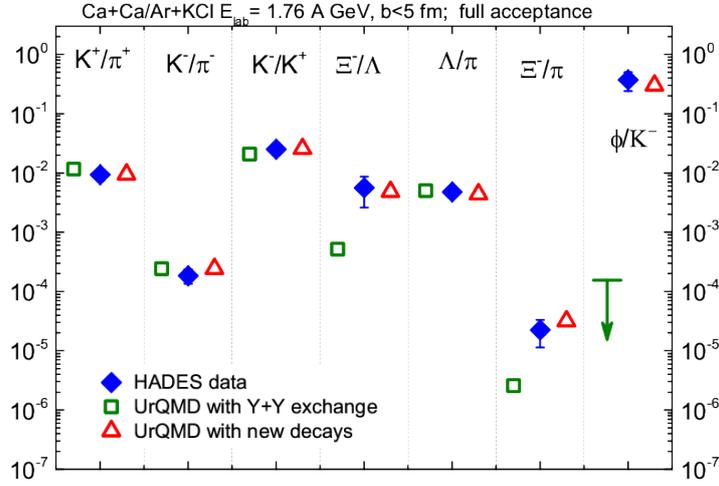


Figure 2. Different strange particle ratios from the UrQMD model in its default settings (green squares), compared to our results including the new N^* decays (red triangles). We compare the simulations for Ca+Ca at $E_{\text{lab}} = 1.76$ A GeV and $b < 5$ fm with published HADES data¹²⁻¹⁵ for Ar+KCl collisions at the same beam energy (blue diamonds)².

5 Comparison to HADES Data

Having constrained the branching fractions, we employ this new mechanism to estimate the production probabilities of ϕ 's and Ξ 's in nuclear collisions, particularly at sub-threshold energies. In Fig. 2 we present results on strange particle ratios, in Ca+Ca collisions at $E_{\text{lab}} = 1.76$ A GeV. The default calculation with the previously released UrQMD version (v3.4) is shown as green squares. Compared to the default calculation we show the new results, including the ϕ and Ξ decay channels of the N^* as red triangles. A considerable increase in the ϕ and Ξ^- production is visible. More importantly when we compare all the obtained strange particle ratios with Ar+KCl data from the HADES experiment (blue diamonds) we observe a very good description of all measured ratios, including the ϕ and Ξ . Such a good description of the full set of data has not been achieved in any previous study. Hence, we conclude that strange particle production in Ar+KCl collisions at the HADES experiment can be explained, and is in fully consistent, with production cross sections obtained in elementary reactions.

In Fig. 3 we present predictions for the same strange particle ratios shown in Fig. 2, in Au+Au reactions at 1.23 A GeV with the new ϕ and Ξ production channels. The red triangles indicate the ratios for Ca+Ca collisions already shown in Fig. 2 for comparison, while the black triangles are the predictions for Au+Au collisions at a beam energy of $E_{\text{lab}} = 1.23$ A GeV and $b < 9.5$ fm. Collisions at this energy have been recently investigated at the HADES experiment. Up to now only preliminary data is available for few particle ratios, shown as blue diamonds. Apparently the preliminary ϕ/K^- ratio seems to indicate a small difference between data and our model study. It will be very interesting to

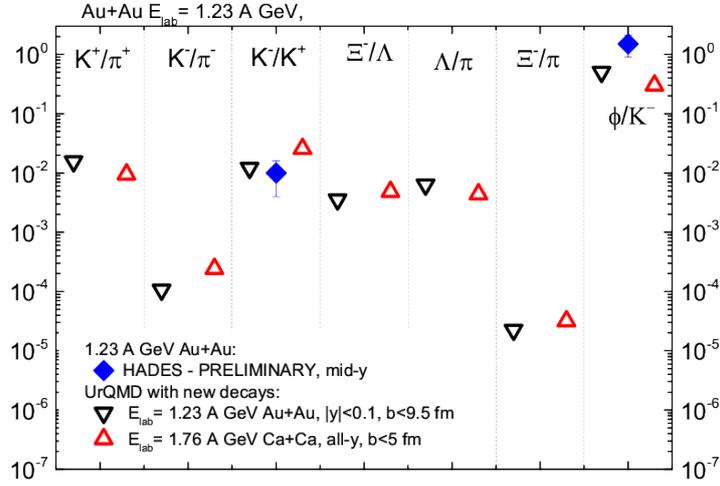


Figure 3. The same particle ratios as in Fig. 2 for Au+Au reactions at $E_{\text{lab}} = 1.23 \text{ A GeV}$, calculated with the UrQMD model, including the new resonance decays. The red triangles correspond to the results for Ca+Ca, already shown in Fig. 2. Preliminary HADES data¹⁶ is shown as blue diamonds².

see whether this holds for the final data and whether a similar difference will also be seen for the Ξ^-/Λ ratio.

6 Summary

We have implemented novel production channels into the Ultra-relativistic Quantum Molecular Dynamics model. These new channels allow to explain both, the Φ and Ξ production in sub-threshold reactions studied at GSI-HADES. As next steps we aim to explore the implications of these novel channels for further reactions.

Acknowledgements

This work was partially supported by the State of Hesse via the LOEWE-Center HIC for FAIR. H.P. was supported by the Helmholtz Association.

References

1. G. Graef, J. Steinheimer, F. Li and M. Bleicher, Phys. Rev. C **90**, no. 6, 064909, 2014.
2. J. Steinheimer and M. Bleicher, arXiv:1503.07305 [nucl-th].
3. J. Steinheimer, M. Mitrovski, T. Schuster, H. Petersen, M. Bleicher and H. Stoecker, Phys. Lett. B **676**, 126 (2009) [arXiv:0811.4077 [hep-ph]].
4. J. Steinheimer, K. Gudima, A. Botvina, I. Mishustin, M. Bleicher and H. Stoecker, Phys. Lett. B **714**, 85 (2012) [arXiv:1203.2547 [nucl-th]].

5. A. S. Botvina, J. Steinheimer, E. Bratkovskaya, M. Bleicher and J. Pochodzalla, Phys. Lett. B **742**, 7 (2015) [arXiv:1412.6665 [nucl-th]].
6. S. A. Bass *et al.*, Prog. Part. Nucl. Phys. **41**, 255 (1998) [Prog. Part. Nucl. Phys. **41**, 225 (1998)] [arXiv:nucl-th/9803035].
7. M. Bleicher *et al.*, J. Phys. **G25**, 1859 (1999) [arXiv:hep-ph/9909407].
8. K. A. Olive *et al.* [Particle Data Group Collaboration], Chin. Phys. C **38**, 090001, 2014.
9. C. Spieles, A. Jahns, H. Sorge, H. Stoecker and W. Greiner, Mod. Phys. Lett. A **8**, 2547, 1993.
10. G. Zeeb, M. Reiter and M. Bleicher, Phys. Lett. B **586**, 297 (2004) [arXiv:nucl-th/0312015].
11. Y. Maeda *et al.* [ANKE Collaboration], Phys. Rev. C **77**, 015204 (2008) [arXiv:0710.1755 [nucl-ex]].
12. G. Agakishiev *et al.* [HADES Collaboration], Eur. Phys. J. A **47**, 21 (2011) [arXiv:1010.1675 [nucl-ex]].
13. P. Tlusty *et al.* [HADES Collaboration], arXiv:0906.2309 [nucl-ex].
14. G. Agakishiev *et al.* [HADES Collaboration], Phys. Rev. C **80**, 025209 (2009) [arXiv:0902.3487 [nucl-ex]].
15. G. Agakishiev *et al.* [HADES collaboration], Phys. Rev. Lett. **103**, 132301 (2009) [arXiv:0907.3582 [nucl-ex]].
16. M. Lorenz [HADES Collaboration], Nucl. Phys. A **931**, 785, 2014.

Contact Mechanics and Fluid Leakage Near Percolation

Wolf B. Dapp and Martin H. Müser

John von Neumann Institute for Computing and Jülich Supercomputing Centre,
Institute of Advanced Simulation, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {w.dapp, m.mueser}@fz-juelich.de

In this work, we present calculations on the leakage of fluid through the gap that forms between a rigid substrate with microscopic roughness and a flat but elastically deformable counter body that is pressed against the substrate. We find that the resistance to flow near the percolation threshold is determined by one, or at most very few constrictions. These constrictions can be described as saddle points and their contact mechanics determine the critical flow rather than the gap topography at large scales.

1 Introduction

The leak rate of simple fluids through an interface formed by a seal and a rigid, randomly rough surface decreases roughly exponentially with intermediate loads L squeezing seal and surface together¹⁻³. At very high loads, the flow falls off even more quickly with L and eventually comes to a complete stop. The accepted picture is that beyond a threshold load, no more percolating open channel exists in which the fluid can pass from one side of the contact to the opposite side⁴⁻⁶. The analysis of the leak-rate problem has gained recent momentum mainly for two reasons: the contact mechanics theory by Persson⁷ allows one to make quantitative predictions – also for the leakage of seals^{2-4,8} – with little numerical effort. Large-scale computer simulations^{5,6,9-11} have likewise progressed due to algorithmic developments and an increase in computing power making it possible to simulate ever more complex systems.

To describe theoretically the leak-rate of (static) seals, Persson and coworkers have pursued two approaches. One² is based on the idea that close to percolation all fluid pressure falls off at the constriction interrupting the last percolating channel at the critical load L_c . The topography of this critical constriction and thus its flow resistance is determined with Persson's contact mechanics theory. The second approach³ is footed on Bruggeman's effective medium theory. It allows one to estimate the resistance to Reynolds flow from the distribution of interfacial separations¹². The latter can be deduced to high accuracy from Persson's contact mechanics theory^{4,13,14}.

Although the leak-rate theory of seals has successfully reproduced both experiments^{2,3} and simulations⁶ in the pressure range where flow is suppressed exponentially with L , it is questionable whether it also describes flow close to the percolation point. Investigations on this issue are presented in this contribution. We note that our problem differs from the random (on/off) resistance networks¹⁵, or simple generalisations thereof, like the Lorentz¹⁶ or Swiss cheese model, which are often seen as generic models for percolation. In those models, conductivity is finite or zero in a given bond or domain. In our percolation problem, the local conductivity evolves smoothly to smaller values as the load increases and continuously approaches zero when the local gap between seal and substrate closes.

2 Methods

To find a realistic channel geometry, we need to simulate contacts in which roughness scales over at least 2 – 3 decades of spatial scale. In order to make predictions that can be compared to continuum theory, the smallest lengthscale has to be sufficiently discretised, which adds another decade (away from the percolation threshold) or even more near the percolation point. Thus, in a brute-force analysis of the problem, the linear length of the system can be easily 10^5 discretisation points long, which boils down to 10^{10} surface grid points in a Fourier-based boundary element method like Green’s function molecular dynamics (GFMD)¹⁷. Modern multi-scale methods that are not based on (fast) Fourier transforms cannot handle the problem more efficiently, due to the complexity of real contact (except at very small contact pressures, when contact is only formed at an isolated mesoscale asperity) and the long-range nature of elastic interactions. Last but not least, solving Reynolds’ thin-film equation on a given domain can entail lengthy computations, whose solution time increases with the complexity of the open channel structure unless efficient solution strategies are pursued. In the following, we summarise the numerical methods we use.

General setup. We consider a system where there is a reservoir of liquid on one side, with a fluid pressure of 1, and on the other there is a sink for said liquid, with a pressure of 0. In the transverse direction the system is periodic. Since the choice of which direction is vertical and which one is horizontal is arbitrary, for each surface we also calculate the flow through the corresponding contact with a transposed surface. In addition, we also compute the flow through a contact where the surface is inverted. We performed these calculations for 6 different surfaces (with different random seeds).

Main aspects pertaining to contact mechanics. The liquid has to pass through channels created by pressing a rigid, randomly rough and a compliant, smooth surface against one another. The surfaces are created by treating the Fourier coefficients of the height spectra as (complex) random numbers whose first moment disappears and whose second moments satisfy

$$\langle \tilde{h}(\mathbf{q}) \tilde{h}^*(\mathbf{q}) \rangle \propto q^{-2(1+H)} \Theta(q_s - q), \quad (1)$$

where H is the Hurst roughness exponent and $q_s = 2\pi/\lambda_s$ with λ_s being the cutoff at short wavelengths. The elastic side is pressed against this substrate but not allowed to penetrate it via a hard-wall constraint. The displacement is considered as a scalar, with all lateral displacements neglected. Details can be found in the literature¹⁸.

Main aspects related to computing the flow. We calculate the fluid flow using the stationary Reynolds thin-film equation. This means that each point in the interface is assigned a conductivity that scales as the third power of the interfacial separation. The `hypre` package¹⁹ together with GMRES (generalised minimal residual) methods²⁰ and multigrid preconditioners is used to solve the sparse discretised Reynolds equation. Our in-house code is MPI-parallelised and uses HDF5 for I/O.

Scaling tests. Both GFMD and our Reynolds solver scale well with the number of processors. This is demonstrated in Fig. 1, where we show the extrapolated time it would take

one MPI rank to make one time step per particle or one iteration per grid point as a function of the requested nodes. In GFMD, this time is around $\gtrsim 1.3 \mu\text{s}$ with a cumulative loss of efficiency of a factor of $\lesssim 2$ when increasing the number of cores. The FFTW library decomposes the domain of a 2-dimensional Fourier transform into stripes. We find that below a stripe width of 16, the efficiency plummets because communication starts to dominate. This, together with the limit of 16 GiB of memory per JUQUEEN node limits the maximum system size we can tackle to $524,288$ grid cells in linear direction, using $65,536$ cores. A different FFT library with a different domain decomposition may remove this limitation, and allow the GFMD code to scale across the entire JUQUEEN. The number of iterations needed to reach convergence does not depend strongly on the system size, and the complexity of each step scales with $\mathcal{O}(N \log N)$. Each grid point requires ≈ 128 Byte of memory.

The Reynolds solver uses `hypre`, which is parallelised using a square domain decomposition. We find it ceases to scale well if each processor works on less than 512×512 grid points. At this workload per processor, compared to larger squares, the efficiency has dropped by only $\lesssim 20\%$. However, the number of steps to reach convergence increases with increasing system size and complexity, and thereby sets limits on the maximum problem size. The memory demands per grid point of the Reynolds solver are greater than that of GFMD by about a factor of 2 to 4, depending on the specific solver used.

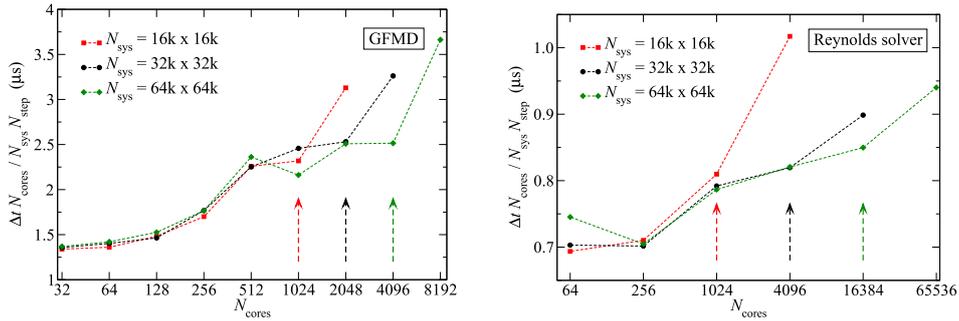


Figure 1. Strong scaling test of GFMD simulations (left) and the Reynolds solver (right) run on JUQUEEN for system sizes of $N_{\text{sys}} = 16, 384^2, 32, 768^2$, and $65, 536^2$. The abscissae correspond to the number of MPI ranks, N_{cores} , while the ordinates show the extrapolated time it would take one core to propagate one grid point by one time step. Here, Δt stands for the wall-clock time and N_{step} for the number of time steps or iterations. The arrows indicate a stripe width of 16 (for GFMD), and 512×512 grid points per process (Reynolds).

3 Results

3.1 Flow through Gaps Between Seals and Self-Affine Rough Surfaces

As the initial step of our simulations, we determine the critical load L_c , which is the external load L squeezing the seal against the self-affine rough surface, above which no more open fluid channel percolates through the system. We find (for a Hurst exponent of $H = 0.8$, which is a typical value for technical and natural surfaces alike) that L_c is the

load needed to induce a relative (critical) contact area of $a_r^* = 0.413 \pm 0.018$. This result is in line with our previous estimate⁶ of $a_r^* = 0.42 \pm 0.02$. While this value may seem remarkably close to the value of ≈ 0.4073 for a square random site-percolation model, there is no *a priori* reason why the result should be the same, because the self-affine roughness induces long-range correlations in the surfaces that are not present in a random-site model. Additionally our contacts live on a continuum, which we represent by resolving the smallest wavelength in our system by many points. Their number increases from 4 far from the percolation threshold to 128 close to it.

Well below the critical load, i.e., at $1 - L/L_c \gtrsim 0.2$, we find that the fluid current drops roughly exponentially with the load. For larger loads, i.e., in the vicinity of the percolation threshold, see also reference²¹, the fluid current j disappears according to

$$j \propto (L_c - L)^\beta \quad (2)$$

with $\beta \approx 7/2$. The exponent is deduced from data such as that shown in Fig. 2.

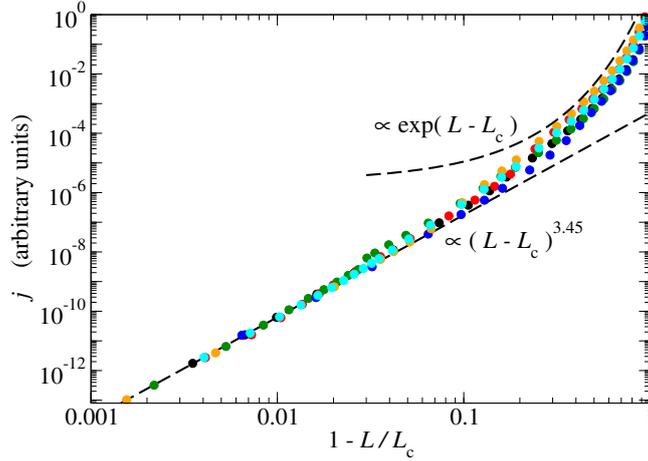


Figure 2. Disorder-averaged fluid flow j as a function of the reduced load $\Delta L = 1 - L/L_c$. Differently coloured symbols represent different random realisations. For clarity not all available data sets are included in the figure, but they all show the same critical behaviour. The data are shifted vertically by a factor of up to $\lesssim 7$ to make them superimpose in the critical region.

To rationalise this behaviour, it is instructive to visualise the fluid flow. This is done in Fig. 3. It shows that the assumption of the critical junction theory appear to be valid, i.e., the fluid pressure drops in quasi-discrete jumps at narrow constrictions. In-between two constrictions, the fluid pressure is essentially constant.

3.2 Critical Flow in Isolated Constrictions

To rationalise the critical flow in the previous section, we now focus on isolated constrictions (see also Ref. 22). These are realised by single-wavelength substrate geometries. The

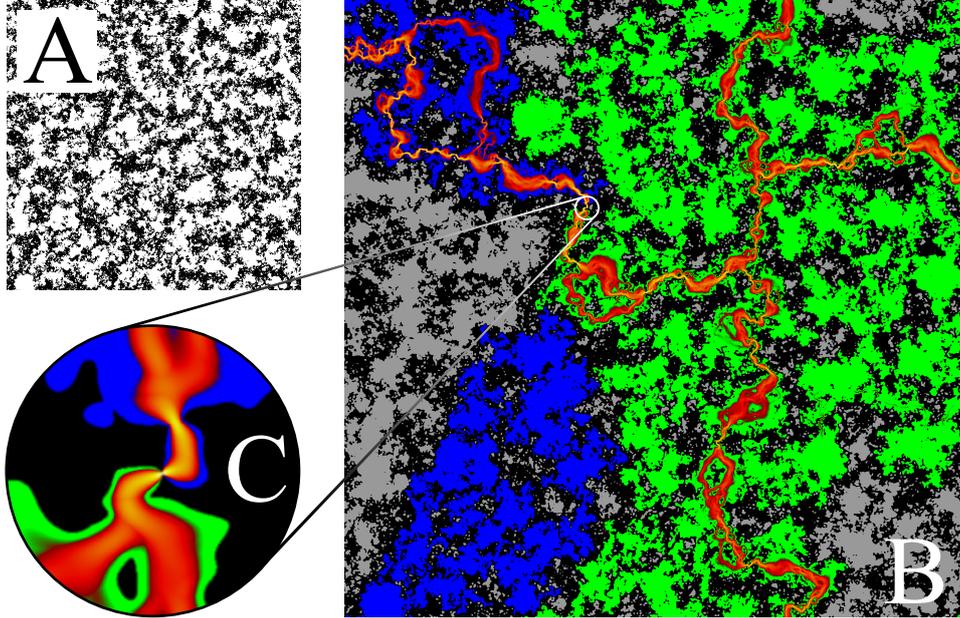


Figure 3. Visualisation of flow passing through the gap between a linearly elastic seal and a self-affine rough substrate (top-down view). **A**: true contact (black) near the percolation threshold. **B**: Flow pattern through the seal. Grey colour marks all non-contact regions that do not belong to the percolating fluid channel. Blue and green colours indicate the fluid pressure, which drops from one (blue) on the left border of the interface to zero (green) on the right border. Red and yellow hues indicate the absolute value of the fluid current density. **C**: Zoom onto the critical constriction.

simplest considered geometry is the square lattice, i.e., a height profile given by

$$\frac{h_{\text{sq}}(x, y)}{h_0} = 2 + \cos\left(\frac{2\pi x}{\lambda}\right) + \cos\left(\frac{2\pi y}{\lambda}\right), \quad (3)$$

for which the simulation cell dimensions along x and y direction are chosen to coincide with the wavelength λ of the height undulation, that is $L_x = L_y = \lambda$. For the realisation of the hexagonal and its dual triangular lattice, we refer to the literature²². Here, we only note that their scaling is similar to that of square lattice, although critical contact areas differ quite substantially, as do the load ratios $r_L \equiv L_c/L_f$, where L_f is the load needed to go into full contact, while L_c is the critical load at which the contact area percolates. We found the following numerical values for the relative contact area needed for percolation $a_c(\text{sq}) = 0.40185(6)$, $a_c(\text{tri}) = 0.17826(11)$, and $a_c(\text{hex}) = 0.67323(1)$. The load ratios were $r_L(\text{sq}) = 4$, $r_L(\text{hex}) = 30$, and $r_L(\text{tri}) = 1.4$.

In Fig. 4, we confirm the critical exponent that we found in the fractal case, and test how well Bruggeman theory describes the flow through an isolated constriction. A mean-field approach cannot reasonably be expected to describe critical behaviour, and indeed, as expected, it fails to predict flow accurately near the percolation threshold.

Since all three investigated saddlepoints have one positive and one negative height curvature, they show universal scaling of width, length, and height of the constriction with

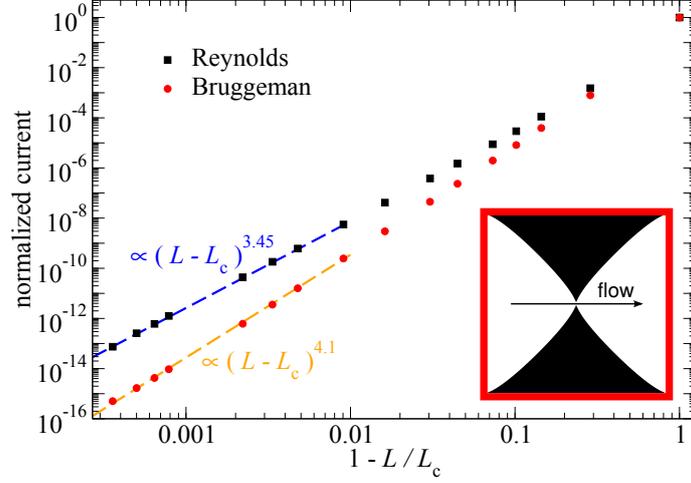


Figure 4. Reynolds flow through a square saddle-point constriction (described by Eq. 3). The inset illustrates the flow, where black is contact and white is a non-zero gap. The power law exponent is the same as for the fractal surface, while the Bruggeman mean field approximation predicts a slightly different scaling.

load

$$g(l, x, y) = |l|^\zeta g^\pm \left(\frac{x}{|l|^\chi}, \frac{y}{|l|^\nu} \right), \quad (4)$$

where $l = (L - L_c)/L_c$ is the reduced load and $g^\pm(\dots)$ are two master functions depending on whether the critical point is approached from high pressures or low pressures. The scaling exponents turn out $\zeta = 6/5$, $\chi = 3/5$, and $\nu = 9/20$, which means that they naturally lead to the observed exponent for the Reynolds flow of $\beta = 3\zeta + \nu - \chi = 69/20$. Only the prefactors and scaling factors depend on the ratio of positive and negative surface curvature.

Fig. 5 demonstrates this scaling relation. It shows the contact line in the x - y -plane (top-down view on the contact) in the left panel, and the shape of the gap in the right panel (x - z -plane, side-view of the contact), very close to the critical point ($l \leq 2\%$). The red and blue solid curves are the master functions $g^\pm(\dots)$ of Eq. 4, which both asymptotically approach the critical green curve. Different symbols show different saddle-point geometries, where we omit the hexagonal case, for clarity, and do not overplot the symbols on all legs of the curves. For all negative reduced loads l (i.e., $L < L_c$) and all geometries the points can be scaled to fall exactly on the blue curve, while for positive reduced loads ($L > L_c$), the red curve is the pertinent curve. Exactly at the critical point, the green curve is traced by the data.

This universality is the reason why the critical behaviour in the fractal case is the same for any of the random realisations we studied, even though the shape of the last constriction might be very different in every case. We note, however, that the precise *magnitude* of the leakage current *does* depend on the local geometry. Hence it remains impossible to predict the prefactor of the flow, just its scaling.

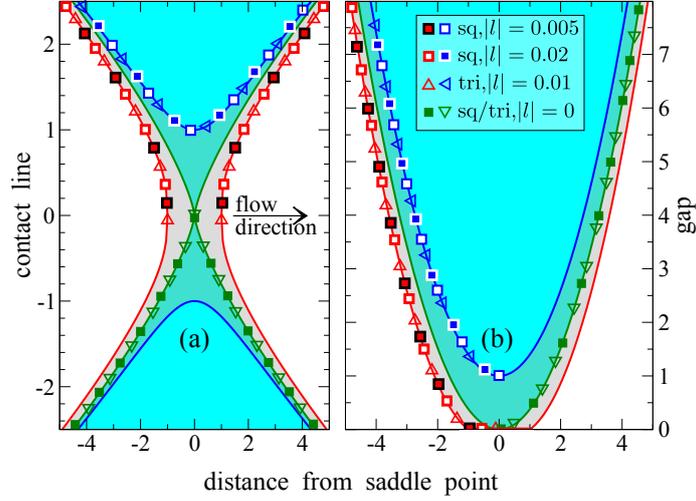


Figure 5. **(a)** Contact line shape for normal loads below (blue), at (green), and, above (red) the critical load. **(b)** Gap on the symmetry axis as a function of the distance from the saddle point. Different symbols represent different saddle-point geometries and different colours different reduced loads l , as indicated in the caption. Solid areas and lines are obtained from analytical approximations to the contact shape. At a given value of l , distances are rescaled according to Eq. 4. Units are chosen such that gap height, length, and width are 1 for $|l| = 0.01$ in the square model. From Ref. 22: Wolf B. Dapp & Martin H. Müser *Contact mechanics of and Reynolds flow through saddle points: On the coalescence of contact patches and the leakage rate through near-critical constrictions*, EPL 109 (2015) 44001. Reprinted with permission.

4 Conclusions

In summary, we carried out numerical simulations of contact between an elastic body with a rigid substrate with self-affine roughness, and computed the fluid flow through the resulting network of narrow channels. We find that, if one stays far enough away from the percolation point, Persson theory in combination with Bruggeman theory accurately predicts the leakage to be exponentially decreasing with external load. As the critical point is approached, at loads of $L \gtrsim 0.9 L_c$, essentially all fluid pressure drops at a single constriction. Then, the dependence of the current on load becomes a powerlaw, as the last constriction determines the resistance of the entire seal. Unsurprisingly, this is not well described by mean-field theory, and detailed calculations are necessary to compute the flow. We found the powerlaw exponent to be $\beta = 69/20$, and confirmed and rationalised this for single-wavelength saddle-point geometries. Independent on the curvature radius of the constriction, we obtain the same shape of the gap, which causes the universal scaling behaviour of the resistance.

However, the precise magnitude of the current does depend on the geometry, as that sets the prefactor of the powerlaw. As such, it is impossible to predict not only at what mechanical load or pressure a seal (like a faucet) stops dripping but also how it stops dripping unless detailed calculations are carried out for a specific (disorder realisation of the) surface. The situation can become even more complex as the universality class can change when additional features are included into the model such as adhesion between the surfaces or more realistic flow boundary conditions.

Acknowledgements

We gratefully acknowledge computing time on JUQUEEN at the Jülich Supercomputing Centre.

References

1. G. Armand, J. Lapujoulade, and J. Paigne, *A theoretical and experimental relationship between the leakage of gases through the interface of two metals in contact and their superficial micro-geometry*, *Vacuum*, **14**, 53–57, 1964.
2. B. Lorenz and B. N. J. Persson, *Leak rate of seals: Comparison of theory with experiment*, *EPL*, **86**, 44006, 2009.
3. B. Lorenz and B. N. J. Persson, *Leak rate of seals: Effective-medium theory and comparison with experiment*, *Eur. Phys. J. E*, **31**, 159, 2010.
4. B. N. J. Persson and C. Yang, *Theory of the leak-rate of seals*, *J. Phys. Condens. Matter*, **20**, 315011, 2008.
5. F. Bottiglione, G. Carbone, L. Mangialardi, and G. Mantriota, *Leakage mechanism in flat seals*, *J. Appl. Phys.*, **106**, 104902, 2009.
6. W. B. Dapp, A. Lücke, B. N. J. Persson, and M. H. Müser, *Self-affine elastic contacts: percolation and leakage*, *Phys. Rev. Lett.*, **108**, 244301, 2012.
7. B. N. J. Persson, *Theory of rubber friction and contact mechanics*, *J. Chem. Phys.*, **115**, 3840–3861, 2001.
8. M. Scaraggi, G. Carbone, B. N. J. Persson, and D. Dini, *Lubrication in soft rough contacts: A novel homogenized approach. Part I - Theory*, *Soft Matter*, **7**, 10395–10406, 2011.
9. F. Sahlin, R. Larsson, P. Marklund, A. Almqvist, and P. M. Lugt, *A mixed lubrication model incorporating measured surface topography. Part 2: roughness treatment, model validation, and simulation*, *P. I. Mech. Eng. J-J Eng.*, **224**, 353, 2009.
10. C. Vallet, D. Lasseux, P. Sainsot, and H. Zahouani, *Real versus synthesized fractal surfaces: Contact mechanics and transport properties*, *Tribol. Int.*, **42**, 250–259, 2009.
11. M. Scaraggi, G. Carbone, and D. Dini, *Lubrication in soft rough contacts: A novel homogenized approach. Part I - Theory*, *Soft Matter*, **7**, 10407–10416, 2011.
12. B. N. J. Persson, *Relation between interfacial separation and load: A general theory of contact mechanics*, *Phys. Rev. Lett.*, **99**, 125502, 2007.
13. C. Campañá, B. N. J. Persson, and M. H. Müser, *Transverse and normal interfacial stiffness of solids with randomly rough surfaces*, *J. Phys.: Condens. Matter*, **23**, 085001, 2011.
14. A. Almqvist, C. Campañá, N. Prodanov, and B. N. J. Persson, *Interfacial separation between elastic solids with randomly rough surfaces: Comparison between theory and numerical techniques*, *J. Mech. Phys. Solids*, **59**, 2355–2369, 2011.
15. I. Webman, J. Jortner, and M. H. Cohen, *Critical exponents for percolation conductivity in resistor networks*, *Phys. Rev. B*, **16**, 2593–2596, 1977.
16. S. K. Schnyder, M. Spanner, F. Höfling, T. Franosch, and J. Horbach, *Rounding of the localization transition in model porous media*, *Soft Matter*, **11**, 701–711, 2015.
17. C. Campañá and M. H. Müser, *Practical Green's function approach to the simulation of elastic semi-infinite solids*, *Phys. Rev. B*, **74**, 075420, 2006.

18. N. Prodanov, W. B. Dapp, and M. H. Müser, *On the contact area and mean gap of rough, elastic contacts: Dimensional analysis, numerical corrections and reference data*, Tribol. Lett., **53**, 433–448, 2014.
19. R.D. Falgout, J.E. Jones, and U.M. Yang, *The Design and Implementation of hypre, a Library of Parallel High Performance Preconditioners*, in: Numerical Solution of Partial Differential Equations on Parallel Computers, 2006.
20. Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., **7**, 856–869, 1986.
21. W. B. Dapp and M. H. Müser, *Fluid leakage near the percolation threshold*, in press, 2015.
22. W. B. Dapp and M. H. Müser, *Contact mechanics of and Reynolds flow through saddle points*, EPL, **109**, 44001, 2015.

The QCD Phase Transition with Two Quark Flavours

Christopher Czaban^{1,2}, Francesca Cuteri¹, Owe Philipsen^{1,2}, Christopher Pinke¹,
and Alessandro Sciarra^{1,2}

¹ Institut für Theoretische Physik, Goethe Universität Frankfurt,
Max-von-Laue-Str. 1, 60438 Frankfurt, Germany

E-mail: {czaban, cuteri, philipsen, pinke, sciarra}@th.physik.uni-frankfurt.de

² John von Neumann Institute for Computing (NIC), GSI, Planckstr. 1, 64291 Darmstadt, Germany

The QCD phase diagram, specifying the form of strongly interacting matter as a function of temperature and density, is important in many disciplines of physics. Finite densities are not amenable to standard Monte Carlo simulations and knowledge of the phase diagram remains scarce. We report from a long term project to determine the phase diagram of QCD with two mass degenerate quark species at zero and imaginary chemical potential, where there is no sign problem, in order to constrain the phase diagram of physical QCD at real chemical potential.

1 Introduction

The fundamental theory of the strong interactions governing the forces between nuclear and subnuclear particles is Quantum Chromodynamics (QCD). Its fundamental degrees of freedom are light u - and d -quarks, a heavier s -quark and gluons. A key feature of the theory is asymptotic freedom with the coupling strength depending on the energy scale of a scattering process. For energies below a few GeV, the coupling is large and quarks and gluons combine into numerous tightly bound states, the hadrons, among them the familiar nucleons proton and neutron. On the other hand, at large temperatures or densities, the average energy per particle is higher and the theory enters a weak coupling regime, where the quarks and gluons form a plasma. The QCD phase diagram determines the form of matter under different conditions as a function of temperature, T , and chemical potential for baryon number, μ_B , as sketched in Fig. 1. Whether and where these regions are separated

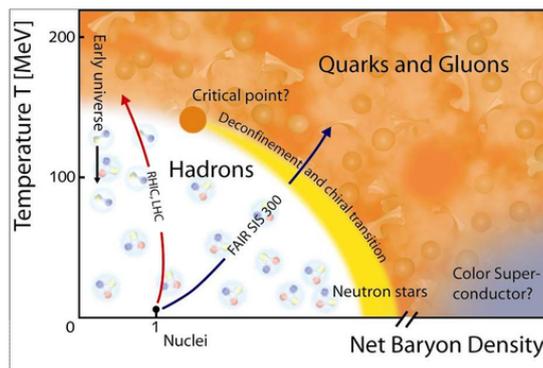


Figure 1. The QCD phase diagram as a function of temperature and chemical potential for baryon density.

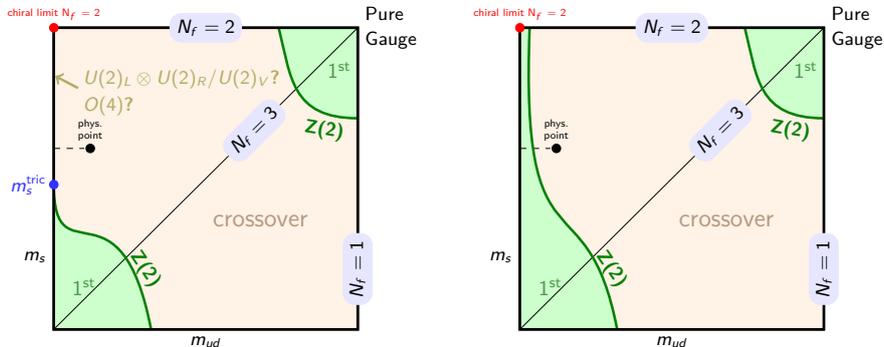


Figure 2. Possible scenarios for the QCD phase diagram at $\mu = 0$ as function of quark mass.

by true phase transitions has to be determined by first principle calculations and experiments. Since QCD is strongly coupled on scales of hadronic matter, a non-perturbative treatment is necessary and Monte Carlo simulations of lattice QCD are the only fully reliable approach. Unfortunately, the so-called sign problem prohibits straightforward simulations at finite baryon density. There are several approximate ways to circumvent this problem, all of them valid for $\mu/T \lesssim 1$ only¹ (with quark chemical potential $\mu = \mu_B/3$). On the other hand, at imaginary chemical potential there is no sign problem. Results from such simulations can be either analytically continued after fitting to power series, or used as a way to constrain the phase diagram at zero and real μ .

The order of the finite temperature phase transition at zero density depends on the quark masses and is schematically shown in Fig. 2 (left), where N_f denotes the number of mass degenerate quark flavours. In the limits of zero and infinite quark masses (lower left and upper right corners), order parameters corresponding to the breaking of a global symmetry can be defined, and for three degenerate quarks one numerically finds first order phase transitions at small and large quark masses at some finite temperatures $T_c(m)$. On the other hand, one observes an analytic crossover at intermediate quark masses, with second order boundary lines separating these regions. Both lines have been shown to belong to the $Z(2)$ universality class of the 3d Ising model²⁻⁴. The critical lines bound the quark mass regions featuring a chiral or deconfinement phase transition, and are called chiral and deconfinement critical lines, respectively. The former has been mapped out on $N_\tau = 4$ lattices⁵ and puts the physical quark mass configuration in the crossover region. The chiral critical line recedes with decreasing lattice spacing^{6,7}: for $N_f = 3$, on the critical point $m_\pi(N_\tau = 4)/m_\pi(N_\tau = 6) \sim 1.8$. Thus, in the continuum the physical point is deeper in the crossover region than on coarse lattices.

An open question to this day remains the order of the transition in the limit of zero light quark masses, called the chiral limit. As explained below, this limit cannot be directly simulated. Consequently, it is still not known whether the chiral phase transition for two quark flavours is of first or second order. Hence an alternative scenario is Fig. 2 (right). Clarifying this question is important because of the proximity of the critical line to the physical point.

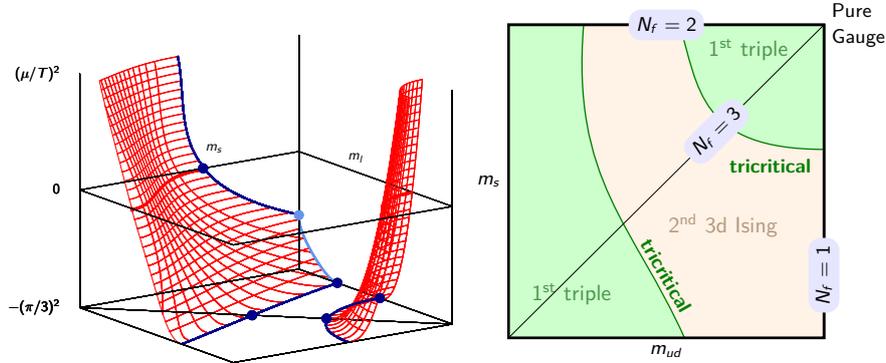


Figure 3. Left: As Fig. 2 with an additional axis for quark chemical potential. Right: Bottom plane of the left.

2 Imaginary Chemical Potential

At imaginary chemical potential $\mu = i\mu_i$ the sign problem is absent and standard simulation algorithms can be applied. Because of an exact symmetry $Z(\mu) = Z(-\mu)$, the partition function is a function of μ^2 . We can then plot Fig. 2 with a third axis, Fig. 3 (left). The chiral and deconfinement critical lines now become critical surfaces as functions of μ^2 , with analytic continuations between real and imaginary μ . The curvature of the chiral critical surface has been computed⁸. In this region of phase space, there is an exact periodic (Roberge-Weiss (RW)) symmetry⁹,

$$Z(\mu) = Z(\mu + 2\pi ik/3), \quad k \in \mathbb{N}, \quad (1)$$

with critical values of $\mu_i^c = (2k + 1)\pi/3$, ($k \in \mathbb{N}$), marking boundaries between adjacent centre sectors of the $SU(3)$ gauge group of QCD. The first boundary constitutes the bottom of Fig. 3 (left) and is shown separately in Fig. 3 (right). First order transition regions for light and heavy quarks are bounded by tricritical lines on which the critical surfaces terminate¹⁰. Note that at this value of imaginary chemical potential, the first order region for two light quark flavours is large enough so that it can actually be seen in simulations on $N_\tau = 4, 6$. In this work, we compute the positions of the tricritical points on the upper $N_f = 2$ line in that diagram on finer $N_\tau = 6$ lattices using Wilson fermions. We also map out the boundary line of the chiral critical surface in the $N_f = 2$ plane on $N_\tau = 4$ lattices.

3 Simulating Thermodynamical Systems in Lattice QCD

The central object in statistical physics is the partition function \mathcal{Z} of a system. For lattice QCD it is expressed as a path integral over the gluon fields U , including a determinant of the Dirac operator for the quarks. An expectation value of some observable A then reads

$$\langle A \rangle = \mathcal{Z}^{-1} \int \mathcal{D}U A \det D[U] \exp \{-S_g[U]\}. \quad (2)$$

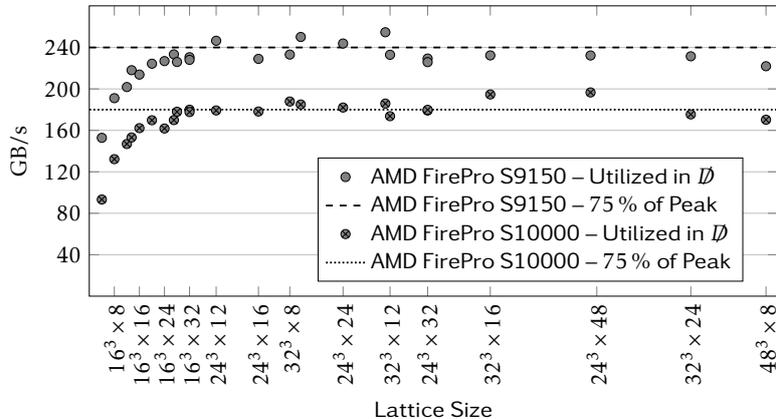


Figure 4. Benchmark of \mathcal{D} for the L-CSC GPUs (taken from Ref. 14).

For our study we used the standard Wilson gauge action $S_g[U]$ as well as the standard Wilson Dirac operator $D[U]$, for detailed expressions see Ref. 11.

The bare parameters are the lattice gauge coupling $\beta = 6/g^2$, and the bare quark mass $m_{u,d} \equiv m$, which is encoded in the hopping parameter κ

$$\kappa = \frac{1}{2(am + 4)} .$$

A finite temperature T is specified by the inverse spatial lattice extent

$$T = (a(\beta)N_\tau)^{-1} . \quad (3)$$

On a lattice with given N_τ , temperature is tuned by changing the lattice spacing a indirectly via the running coupling $\beta(a)$. On the other hand, a continuum limit at fixed temperature implies $a \rightarrow 0$, $N_\tau \rightarrow \infty$, and larger values of N_τ imply smaller lattice spacings.

The determinant of the fermion matrix D is expressed in terms of pseudo fermions ϕ ,

$$\det D[U] \sim \int \mathcal{D}\phi^\dagger \mathcal{D}\phi \exp \{ -\phi^\dagger D^{-1}[U]\phi \} , \quad (4)$$

yielding the effective action $S_{\text{eff}}[U, \phi] = S_{\text{gauge}}[U] + \phi^\dagger D^{-1}[U]\phi$. Importance sampling methods are used to evaluate this high-dimensional integral. Using the Boltzmann-weight $p[U, \phi] = \exp \{ -S_{\text{eff}}[U, \phi] \}$ as probability measure, an ensemble of N gauge configurations $\{U_m\}$ is generated. Then, $\langle A \rangle$ may be approximated by

$$\langle A \rangle \approx \frac{1}{N} \sum_m A[U_m] . \quad (5)$$

The standard simulation algorithm to generate QCD gauge configurations is the Hybrid Monte-Carlo (HMC) algorithm, where the effective action is embedded in a fictitious classical system evolved over a time τ according to the Hamiltonian equations of motion.

Since the fermion matrix D is high-dimensional and sparse, iterative Krylov space methods are used. D^{-1} is calculated indirectly out of equations like

$$D\phi = \psi \Rightarrow \phi = D^{-1}\psi . \quad (6)$$

This inversion is the most cost-intensive part of a simulation. Thus, for performance it is crucial to have a well tuned implementation, in particular of the derivative term \not{D} . The numerical costs for the HMC scales like $V^{5/4}$ and m_π^{-6} . This illustrates that LQCD studies are very cost-intensive, especially when going towards the chiral limit, and efficient simulation programs are needed. We thermalised configurations on JUQUEEN. For production we employed our publicly available^a OpenCL^b-based code CL2QCD¹², which is optimised to run efficiently on the GPUs of the LOEWE-CSC¹³ at Goethe-University Frankfurt and the L-CSC¹⁴ at GSI in Darmstadt. The latter was recently ranked as the most energy efficient computing cluster in the world^c. Fig. 4 shows the performance of the \not{D} on this cluster and underlines the efficient implementation.

4 Identifying Phase Transitions

	Crossover	1 st triple	Tricritical	3D Ising
B_4	3	1.5	2	1.604
ν	–	1/3	1/2	0.6301(4)
γ	–	1	1	1.2372(5)

Table 1. Critical values of ν , γ and $B_4 \equiv B_4(X, \alpha_c)$ for the universality classes needed here¹⁶.

In order to identify the phase transition as a function of the QCD parameters, we use the Binder cumulant¹⁵ defined as

$$B_4(X, \alpha_1, \dots, \alpha_n) \equiv \frac{\langle (X - \langle X \rangle)^4 \rangle}{\langle (X - \langle X \rangle)^2 \rangle^2},$$

where X is a general observable and $\alpha_1, \dots, \alpha_n$ is a set of parameter on which B_4 depends. Our observables are the Polyakov loop and the chiral condensate, and the QCD parameters are $\{\alpha_i\} = \{\beta, \kappa, \mu_i\}$. First we find the location of phase transitions by the vanishing third moment of the fluctuations in the observables, $\langle (X - \langle X \rangle)^3 \rangle \approx 0$, for critical parameter sets α_i^c . The nature of the transition then has to be extracted from finite size scaling on those parameter values. Non-analytic phase transitions only exist in the thermodynamic limit $V \rightarrow \infty$, for which the Binder cumulant takes different values depending on the nature of the phase transition (see Tab. 1). Even though B_4 is a non-analytic step function for $V \rightarrow \infty$, at finite volume it gets smoothed out and its slope increases with the volume. Around the critical coupling β_c , the Binder cumulant is a function of $x \equiv (\beta - \beta_c)N_\sigma^{1/\nu}$ only and can be Taylor-expanded around 0. The first finite-size corrections are

$$B_4(\beta, N_\sigma) = B_4(\beta_c, \infty) + a_1 x + a_2 x^2 + \mathcal{O}(x^3). \quad (7)$$

^aSee github.com/CL2QCD/.

^bSee www.khronos.org/opencl for more information.

^cSee <http://www.green500.org/news/green500-list-november-2014>.

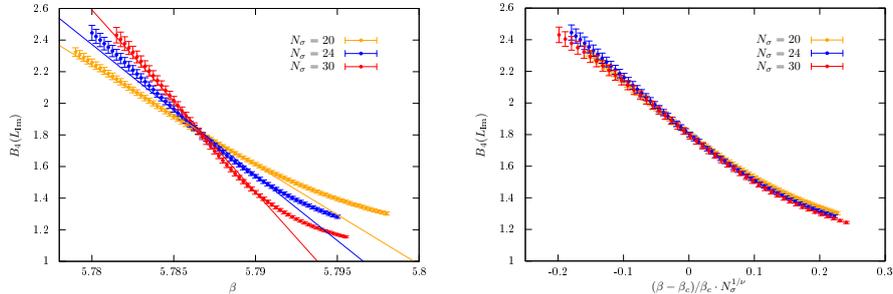


Figure 5. Binder cumulant for $\kappa = 0.13$ and $N_\tau = 6$. Left: Reweighted data including linear fits. Right: Collapse plot with second order exponents.

Close to the thermodynamic limit, the intersection of different curves from different volumes gives β_c and the critical exponent ν takes its value depending on the type of transition. While the value of the Binder cumulant is very sensitive to finite size corrections and in some cases not close to its infinite volume value, the exponents are more stable.

Our strategy to locate the two tricritical values of κ is described in detail in Ref. 17. For each simulated quark mass parametrised by κ , we measured the Binder cumulant around the critical coupling β_c and extracted the value of ν fitting our data according to Eq. 7. Because of the particularly delicate fitting procedure required to extract the critical exponent ν from Eq. 7, we usually produced four different Markov chains for each value of the coupling in order to acquire statistics in a faster manner and also to better understand if the collected statistics was enough. Ferrenberg-Swendsen reweighting¹⁸ was used to produce additional β -points and thus to smooth the data. An example is shown in Fig. 5 (left). We then check the consistency of our fitted exponents by replotting the data with the nearest exact critical exponent. To the extent that the data are correctly described by these exponents, they collapse onto a universal curve in the neighbourhood of the critical coupling as in Fig. 5 (right).

5 Results

Fig. 6 shows the critical exponents ν for QCD with two flavours of Wilson quarks simulated at $\mu_i = \pi T/3$ as a function of hopping parameter and thus quark masses on $N_\tau = 4$ (left)¹⁷ and $N_\tau = 6$ (right) lattices. We clearly observe the change from a first order behaviour to second order values and back, passing through tricritical points. These correspond to the tricritical points on the upper boundary of Fig. 3 for two different lattice spacings. As is the case for $\mu = 0$, the chiral first order region is shrinking with growing N_τ , i.e. as the lattice gets finer. Additional finer lattice spacings are required in future simulations before extrapolations to the continuum can be undertaken and conclusions can be drawn.

In a second set of calculations we have mapped out the critical boundary line in the $N_f = 2$ backplane of Fig. 3 (left)¹⁹, following a recent suggestion to allow for a determination of the nature of the chiral phase transition at $\mu = 0$ ²⁰. The situation and its particular interest are shown in Fig. 7 (left). The boundary line coming from the tricritical point at $\mu_i = \pi T/3$ has to terminate in another tricritical point on the μ^2 -axis. If the line hits the

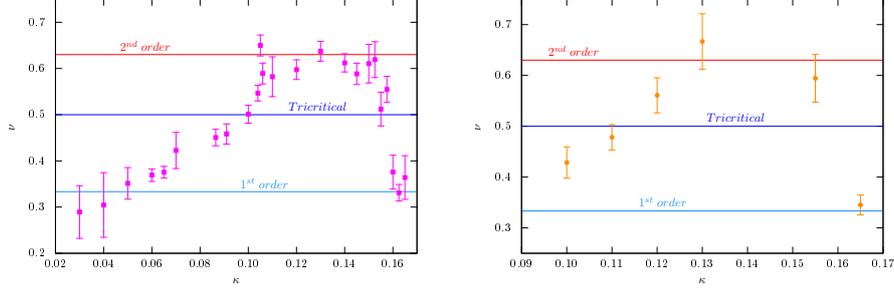


Figure 6. Critical exponents in the RW plane as a function of mass for $N_\tau = 4$ (left) and $N_\tau = 6$ (right).

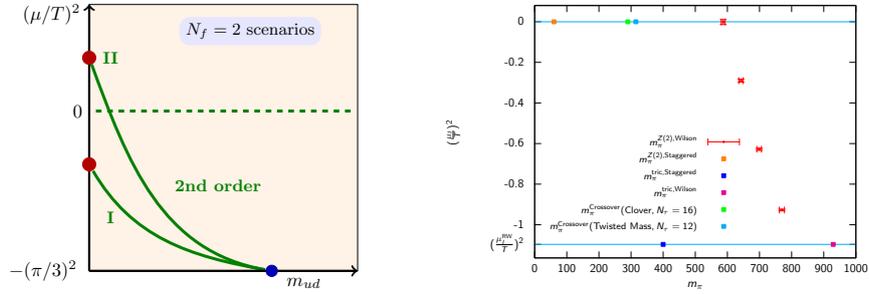


Figure 7. Left: Possibilities for the boundary of the chiral critical surface in the $N_f = 2$ backplane of Fig. 3 (left). To the left of the lines is first order region, to the right crossover. Right: The boundary line for $N_\tau = 4$.

axis at $\mu < 0$, the chiral phase transition at $\mu = 0$ is of second order, whereas if it hits at $\mu > 0$ there is a first order region around the chiral limit. The boundary line is in the 3D Ising universality class and we can again use the Binder cumulant to identify and map it out. Fig. 7 (right) shows the results of our simulations. Note that, in order to be able to compare different lattice spacings as well as different lattice discretisation, the horizontal axis of the figure is the pion mass rather than κ or the quark mass. This is because the quark mass receives renormalisation factors specific to a particular choice of discretisation, whereas the pion mass is a renormalisation group invariant physical quantity, that directly depends on κ and can be compared between different lattice discretisations.

The figure illustrates that Wilson fermions on coarse lattices clearly feature a large region of first order chiral transitions. Note that this region is much larger than in the case of staggered fermions, which are also shown. This difference must be due to discretisation effects, since the physics in the continuum must be independent of the discretisation scheme chosen. One may speculate that the Wilson results have the larger discretisation effects, since there are already simulations with improved Wilson actions at smaller lattice pion masses which are in the crossover region. The huge discrepancies illustrate the need to go to much finer lattices, and hence require significantly more High Performance Computing time, before any conclusions for the continuum can be drawn.

6 Summary

We have started a systematic investigation of the phase structure of QCD with two flavours of standard Wilson fermions at zero and imaginary chemical potentials. On coarse lattices we find clear evidence for a first order chiral phase transition. However, comparison between different lattice spacings and different discretisation schemes shows huge discretisation artefacts. These have to be removed by future simulations on much finer lattices in order to draw conclusions for continuum physics. This highlights the continued and growing need for High Performance Computing in elementary particle and nuclear physics.

Acknowledgements

C.C., C. P. and A.S. were supported by the Helmholtz International Center for FAIR and HGS-Hire within the LOEWE program of the State of Hesse, O.P. was supported by BMBF grant 06FY7100. The authors thank the administrating teams of LOEWE-CSC at GU, L-CSC at GSI and JUQUEEN at NIC.

References

1. P. de Forcrand, PoS **LAT2009** (2009) 010 [arXiv:1005.0539 [hep-lat]].
2. F. Karsch, E. Laermann and C. Schmidt, Phys. Lett. B **520** (2001) 41 [arXiv:hep-lat/0107020].
3. P. de Forcrand and O. Philipsen, Nucl. Phys. B **673** (2003) 170 [arXiv:hep-lat/0307020].
4. S. Kim *et al.*, PoS **LAT2005**, (2006) 166 [arXiv:hep-lat/0510069].
5. P. de Forcrand and O. Philipsen, JHEP **0701** (2007) 077 [hep-lat/0607017].
6. P. de Forcrand, S. Kim and O. Philipsen, PoS **LAT2007** (2007) 178 [arXiv:0711.0262 [hep-lat]].
7. G. Endrödi *et al.*, PoS **LAT2007** (2007) 182 [arXiv:0710.0998 [hep-lat]].
8. P. de Forcrand and O. Philipsen, JHEP **0701** (2007) 077 [hep-lat/0607017].
9. A. Roberge and N. Weiss, Nucl. Phys. B **275** (1986) 734.
10. P. de Forcrand and O. Philipsen, Phys. Rev. Lett. **105** (2010) 152001 [arXiv:1004.3144 [hep-lat]].
11. C. Gattringer and C. B. Lang, Lect. Notes Phys. **788** (2010) 1.
12. M. Bach, V. Lindenstruth, O. Philipsen and C. Pinke, Comput. Phys. Commun. **184** (2013) 2042 [arXiv:1209.5942 [hep-lat]].
13. M. Bach *et al.*, Computer Science - Research and Development **26** (2011).
14. D. Rohr, M. Bach, G. Neskovic, V. Lindenstruth, C. Pinke and O. Philipsen, High Performance Computing LNCS **9137** (2015).
15. K. Binder, Z. Phys. B **43** (1981) 119.
16. A. Pelissetto and E. Vicari, Phys. Rept. **368** (2002) 549 [cond-mat/0012164].
17. O. Philipsen and C. Pinke, Phys. Rev. D **89** (2014) [arXiv:1402.0838 [hep-lat]].
18. A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63** (1989) 1195.
19. O. Philipsen, C. Pinke, PoS **LAT2015** (2015) 149, [arXiv:1508.07725 [hep-lat]].
20. C. Bonati, P. de Forcrand, M. D'Elia, O. Philipsen and F. Sanfilippo, Phys. Rev. D **90** (2014) 7, 074030 [arXiv:1408.5086 [hep-lat]].

Dynamical Simulations of Lattice QCD

Karl Jansen, Stefan Schaefer, Hubert Simma, and Rainer Sommer

John von Neumann Institute for Computing, DESY, Platanenallee 6, 15738 Zeuthen
E-mail: {karl.jansen, stefan.schaefer, hubert.simma, rainer.sommer}@desy.de

1 Introduction

Lattice calculations of Quantum Chromodynamics (QCD) are continuously becoming more realistic. Where Ukawa¹ famously concluded only fourteen years ago that simulations including two physically light sea quarks are basically impossible even with today's computers, algorithmic developments over the last years have changed this situation drastically. Nowadays up and down quark masses light enough to control the chiral extrapolation reliably are standard and also the sea quark effects of strange (and charm) quark are included.

Modern lattice simulations are an intricate interplay between a large variety of numerical methods on one side and the computer hardware on the other side. The main areas of progress have been the solvers used for the Dirac equation, fermion determinant factorisations and better integrators for the molecular dynamics which is at the heart of most algorithms used for QCD simulations.

In lattice QCD simulations the path integral is computed via a Markov Chain Monte Carlo method. In virtually all projects with dynamical fermions a variant of the Hybrid Monte Carlo algorithm is employed to generate the Markov chain, where the fields are updated using molecular dynamics. But there is considerable freedom in how to include the fermion determinant into the simulation. Factorisations^{2,3} of this determinant have been essential in the progress of recent years, being successful in particular together with improved integrators of the molecular dynamics⁴.

The solution of the Dirac equation constitutes the most computer time consuming element of simulations with fermions. The dramatic speedup for small fermion mass due to locally deflated solvers^{5,6} has therefore had a significant impact on what is possible in the simulations. These algorithms have practically eliminated the increase in cost of the solution as the quark mass is lowered.

1.1 Autocorrelations

One of the main difficulties of Markov Chain Monte Carlo is the presence of autocorrelations: subsequent field configurations depend on previous ones and one essential requirement of a correct simulation is that the whole chain is much longer than the modes decorrelating most slowly.

In the study of the autocorrelations it has also become clear that significantly more statistics is required to reach precision results than previously thought. In particular observables constructed from smoothed gauge fields⁷ turned out to be particularly sensitive to these long autocorrelations.

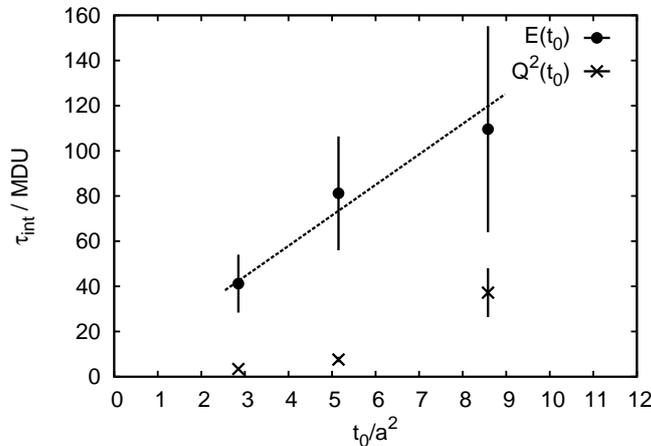


Figure 1. Scaling of the autocorrelation times of the action density E constructed from smoothed links and the topological charge Q with the lattice spacing a . The results are from the CLS 2+1 simulations using open boundary conditions. For E we observe scaling compatible with the theoretical expectation with a dynamical critical exponent of 2. There is no abnormal slowing down of the charge observed.

Finally, in the standard setup with periodic boundary conditions in space and time a dramatic critical slowing down of the algorithms has been observed⁸, i.e. lowering the lattice spacing the number of update steps to reach an independent configuration in the Markov chain needed to be increased drastically. This was due to topological modes which need more and more time to decorrelate. This topological freezing is a property of the continuum theory which exhibits disconnected sectors in field space labelled by an integer topological charge.

A way out is to use open boundary conditions in time⁹. This avoids the sector formation in the continuum and avoids the presence of exceptionally slow modes in the Markov chain. These boundary conditions have therefore been adopted in the CLS simulations described below¹⁰. In Fig. 1 it is demonstrated that in simulations with open boundary conditions no dramatic slowing down is observed, even for quantities known for their sensitivity to slow modes.

2 Status of Dynamical Simulations

By their nature, three principal effects have to be controlled in simulations of lattice QCD: the effect of the finite lattice spacing a , the finite size of the box L and the unphysical quark masses. Furthermore also the consequences of the finite computer time available, and therefore a limited statistics of the Monte Carlo estimates, including the problem of the autocorrelations.

In these simulations, a compromise has to be found between the prominence of these various sources of systematic error. For example at fixed physical volume smaller light quark masses will lead to larger volume effects. Equally smaller lattice spacing will lead to longer autocorrelations – due to the critical slowing down of the algorithms – and therefore

ensemble	β	κ_c	$a\mu_\ell$	L/a	N_τ	N_{conf}	τ	$\tau_{\text{int}}(P)$
<i>cA2.09.48</i>	2.10	0.13729	0.0009	48	6900	2950	1	15(6)
<i>cA2.30.24</i>	2.10	0.13730	0.0030	24	3400	1300	1	3.2(8)
<i>cA2.60.24</i>	2.10	0.13730	0.0060	24	9000	4000	1	3.8(6)
<i>cA2.60.32</i>	2.10	0.13730	0.0060	32	11840	5350	1	2.9(5)

Table 1. The $N_f = 2$ ETMC ensembles, all of which have temporal extent $T = 2L$ with L/a the spatial lattice extend. In addition we give the total number of trajectories N_τ , the number of thermalised configurations N_{conf} and the HMC trajectory length τ and the integrated autocorrelation time of the plaquette $\tau_{\text{int}}(P)$.

observable	<i>cA2.09.48</i>	<i>cA2.30.24</i>	<i>cA2.60.24</i>	<i>cA2.60.32</i>
aM_{π^\pm}	0.06196(09)($^{+12}_{-05}$)	0.1147(7)($^{+4}_{-7}$)	0.15941(38)($^{+15}_{-21}$)	0.15769(26)($^{+15}_{-14}$)
$aM_{\pi^{(0,c)}}$	0.1191(05)($^{+07}_{-10}$)	0.1541(13)($^{+05}_{-05}$)	0.18981(61)($^{+21}_{-25}$)	0.18840(44)($^{+46}_{-29}$)
aM_{π^0}	$-\dagger$	0.09(1)*	0.11(2)*	0.13(2)*

Table 2. The splitting between the charged and the neutral pion is a discretisation effect for twisted-mass fermions. In the new simulations it is much reduced compared to the previously used formulation. \dagger : For *cA2.09.48*, disconnected diagrams have not been computed yet. $*$: For the full neutral pion, a study of systematic effects from excited states was not possible due to the poor signal.

require increased length of the simulations.

In the following we will present two sets of such simulations: one by the ETM collaboration using twisted mass fermions, one by CLS with non-perturbatively improved Wilson fermions. Some of the choices of these simulations are similar, some also different and therefore allow for complementary checks concerning the control of systematic effects.

2.1 ETMC

Along with the $N_f = 2 + 1 + 1$ simulations pursued by the European Twisted Mass collaboration (ETMC) in recent years, there has been a renewed interest in $N_f = 2$ flavour simulations¹¹. The leading discretisation effects of twisted mass fermions at maximal twist come always at $O(a^2)$, but by adding the clover term with a suitably tuned coefficient, their magnitude could be reduced.

The simulation parameters are summarised in Tab. 1. As can be seen from the table, these simulations are so far restricted to one lattice spacing and relatively small volume ($m_\pi L \approx 3$ at the physical point), however, varying the lattice size L allows to study the volume effects and give estimates for these effects.

The main effect of the reduction of discretisation effects can be seen in the difference between the mass of the neutral and the charged pion. Where this has been quite noticeable in previous simulations without the clover term, the effect is now significantly reduced as can be seen from the results in Tab. 2.

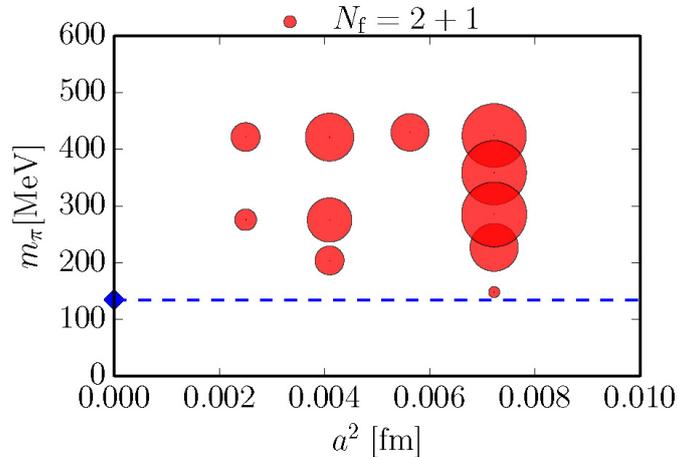


Figure 2. Lattice spacing a and pion mass m_π of the CLS 2+1 ensembles. The size of the circle is proportional to the number of independent gauge field configurations. The blue diamond indicates the physical point.

2.2 CLS

The Coordinated Lattice Simulations (CLS) have a programme to simulate $N_f = 2 + 1$ flavours of non-perturbatively improved Wilson fermions. The project started in 2013 and by now has generated lattices at four different lattice spacings between 0.085 fm and 0.05 fm¹². Lattice spacing and pion mass of these ensembles are visualised in Fig. 2, from which it becomes clear that the control over both, the chiral and the continuum limit is an essential part of this project. Both can bring significant corrections with respect to the lattice results. An example of this is given in Fig. 3 where the product of the pion decay constant with the gluonic scale parameter t_0 defined through the Wilson flow⁷ is displayed. As we can see, the accuracy which can be reached in such a quantity is on the level of 1%.

The extrapolation towards the continuum limit agrees with leading scaling violations of $O(a^2)$, as expected for this non-perturbatively $O(a)$ improved theory. From the plot it is also obvious that only the fine lattices used here can lead to such a 1% accuracy – the points at 0.085 fm and 0.065 fm being $O(5\%)$ away from the continuum result.

Also the chiral corrections are well under control as can be seen from Fig. 4. We show

$$f_{\pi K} = \frac{2}{3}(f_K + \frac{1}{2}f_\pi) \quad (1)$$

along the line of constant sum of quark masses $m_u + m_d + m_s$. In particular in comparison to Fig. 3 it becomes clear that chiral and continuum extrapolations are equally important and to reach per cent level accuracies it is pivotal to control both at a high level of accuracy. It is remarkable that in the full range of quark masses, next-to-leading order chiral perturbation theory describes the effect to the 20% level.

With these lattices the foundation of a large variety of projects has been laid which are currently pursued in a number of European groups.

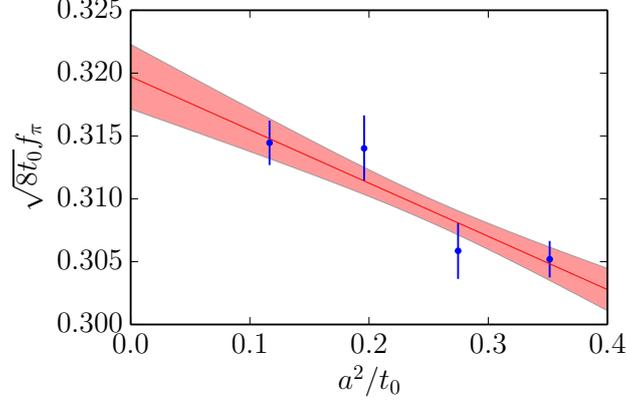


Figure 3. Continuum extrapolation of the dimensionless product $f_\pi \sqrt{8t_0}$ of the pseudoscalar decay constant f_π and the gluonic scale t_0 along the line $m_\pi = m_K \approx 420$ MeV. Fine lattices are needed to reach a per cent level result.

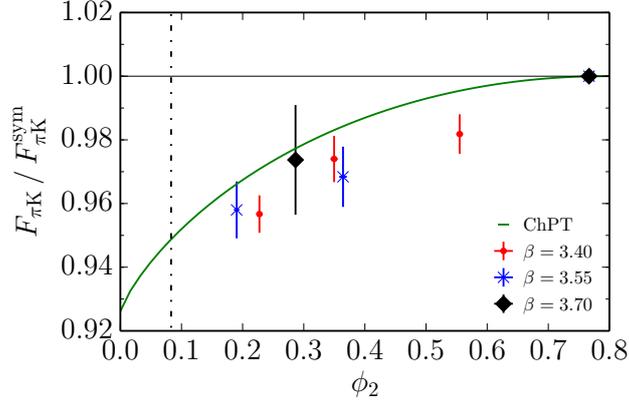


Figure 4. Chiral extrapolation of $f_{\pi K} = \frac{2}{3}(f_K + \frac{1}{2}f_\pi)$ along the line of constant sum of quark masses $m_u + m_d + m_s$.

3 Anomalous Magnetic Moments

One particular result of the ETMC simulations described above are hadronic contributions to anomalous magnetic moments of the leptons, which give the deviation of Dirac's leading order value of the magnetic moment of a fermion $g = 2$. Hadronic contributions are the dominant source of uncertainty in the comparison of the quantum field theory prediction of the quantity to experiments which are accurate to eight digits for the electron and

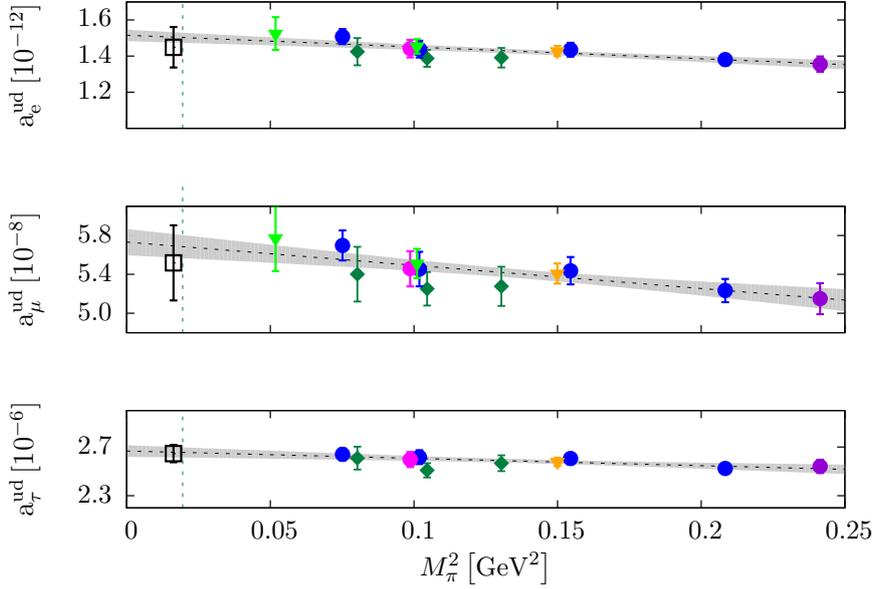


Figure 5. Comparison of the chiral extrapolation of the light quark contributions to the three lepton anomalous magnetic moments obtained from $N_f = 2 + 1 + 1$ simulations to the values at the physical value of the pion mass (black square). The dark green diamonds correspond to $a = 0.086\text{fm}$ and $L = 2.8\text{fm}$ and the circles to $a = 0.078\text{fm}$, the violet one stands for $L = 1.9\text{fm}$, the blue ones for $L = 2.5\text{fm}$, and the pink for $L = 3.7\text{fm}$. The orange triangle shows the value obtained for $a = 0.061\text{fm}$ and $L = 1.9\text{fm}$ and the light green triangle denotes $a = 0.061\text{fm}$ and $L = 2.9\text{fm}$.

seven digits for the muon anomalous magnetic moments. Since the tau lepton has a very short lifetime $O(10^{-13}\text{seconds})$ there is presently no experimental result for its anomalous magnetic moment. Deriving the anomalous moment from field theory, one may, and does, employ an expansion in the electromagnetic and weak couplings but not in the strong coupling. The hadronic contribution is the given by

$$a_1^{\text{hvp}} = \alpha^2 \int_0^\infty \frac{dQ^2}{Q^2} w\left(\frac{Q^2}{m_1^2}\right) \Pi_R(Q^2), \quad (2)$$

with $\Pi_R(Q^2)$ the renormalised hadronic vacuum polarisation function and $w(x)$ a known function.

In Fig. 5 a comparison of the previous results in $N_f = 2 + 1 + 1$ flavours at pion masses above the physical point and the new $N_f = 2$ numbers is shown. The two results are consistent. Note that in this plot there is a significant variation of the number of flavours in the sea, the lattice volume, pion mass and the lattice spacing. A mutual agreement between these results is therefore not trivial and suggests that these systematic errors are under control.

4 Conclusion

The ability to simulate QCD on the lattice has progressed significantly over the last years, to the extent that we can now control all systematic effects coming from the finite volume, the non-physical quark masses and the finite lattice spacing to the level of per cent accuracies for a number of observables.

This is due to improved algorithms, computational strategies and codes which exploit the improved hardware resources. Of course challenges remain. Many observables are still not accessible to lattice calculations with the accuracies required to make an impact in the phenomenological analysis of high-energy physics data. And ever changing hardware will also pose new requirements on the algorithms and the software.

References

1. A. Ukawa, *Computational cost of full QCD simulations experienced by CP-PACS and JLQCD Collaborations*, Nucl. Phys. Proc. Suppl., **106**, 195–196, 2002, [,195(2002)].
2. M. Hasenbusch, *Speeding up the hybrid Monte Carlo algorithm for dynamical fermions*, Phys.Lett., **B519**, 177–182, 2001.
3. M. Hasenbusch and K. Jansen, *Speeding up lattice QCD simulations with clover improved Wilson fermions*, Nucl.Phys., **B659**, 299–320, 2003.
4. I. P. Omelyan, I. M. Mryglod, and R. Folk, *Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations*, Computer Physics Communications, **151**, no. 3, 272 – 314, 2003.
5. M. Lüscher, *Local coherence and deflation of the low quark modes in lattice QCD*, JHEP, **0707**, 081, 2007.
6. A. Frommer, K. Kahl, S. Krieg, B. Leder, and M. Rottmann, *Adaptive Aggregation-Based Domain Decomposition Multigrid for the Lattice Wilson–Dirac Operator*, SIAM Journal on Scientific Computing, **36**, no. 4, A1581–A1608, 2014.
7. M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, JHEP, **1008**, 071, 2010.
8. S. Schaefer, R. Sommer, and F. Virota, *Critical slowing down and error analysis in lattice QCD simulations*, Nucl.Phys., **B845**, 93–119, 2011.
9. M. Lüscher and S. Schaefer, *Lattice QCD without topology barriers*, JHEP, **1107**, 036, 2011.
10. M. Lüscher and S. Schaefer, *Lattice QCD with open boundary conditions and twisted-mass reweighting*, Comput.Phys.Commun., **184**, 519–528, 2013.
11. A. Abdel-Rehim *et al.*, *Simulating QCD at the Physical Point with $N_f = 2$ Wilson Twisted Mass Fermions at Maximal Twist*, 2015.
12. M. Bruno *et al.*, *Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions*, JHEP, **02**, 043, 2015.

Astrophysics

Astrophysics

Peter L. Biermann

MPI for Radioastronomy, Bonn, Germany
E-mail: plbiermann@mpifr-bonn.mpg.de

Dept. Phys. & Astron., Univ. of Alabama, Tuscaloosa, AL, USA

Karlsruhe Inst. for Tech., KIT, Germany

Dept. for Phys. & Astron., Univ. Bonn, Germany

One of the most interesting quests in astrophysics is exploring where we come from, how stars like the Sun and others form and evolve, how the planet system is born and develops. In parallel we obtain every deeper insight from the latest radio telescopes, such as LOFAR, arrays that extend across the entire Netherlands and now much of Europe. LOFAR can observe radio galaxies as well as ultra high energy cosmic rays that traverse the atmosphere, and many other phenomena at low radio frequencies.

Introduction

Observations show clearly that stars have formed already very early in the universe, with the deepest observations now reaching redshift $z \simeq 10$, corresponding to 3 percent of the age of the universe, now known to be 13.8 billion years, so 450 million years after the big bang. When star formation really began, we do not know, but it may have been very much earlier even, with some explorative speculations allowing star formation at redshift $z \simeq 100$ even, one per mille of the present age of the universe. Observations also show that almost all stars have planets around them, usually revolving in a disk like arrangement, due to angular momentum alignment, as already speculated hundreds of years ago, and the first mathematical theory already more than sixty years ago, called “accretion disk”. In dense star clusters such orbits get heavily disturbed. Our own Solar system is only 4.5 billion years old, corresponding to a redshift slightly less than 0.5 only. How do we explain all these phenomena? At the same time our observational methods are getting better and better, with one now famous example the LOFAR radio array, a telescope by computer, since the phased radio signal is transmitted via high-speed internet, and combined in a supercomputer to image the entire sky at once, and very quickly, so that even transient phenomena can be detected, such as ultra high energy particles traversing the atmosphere.

Stars and Planets

Already in ancient times the planets in our own Solar system were observed in the evening and morning skies, and it was noted later that they all follow essentially co-planar orbits. The first mathematical theory to describe the formation of such an accretion disk was formulated in the early 1950s³. Now we have abundant observations of such systems, and go very much deeper into the questions, how stars form out of the interstellar medium, how then planetary systems develop, and do all this in exquisite detail. Three articles here

describe this search: 1) Formation of Star-Forming Clouds from the Magnetised, Diffuse Interstellar Medium, by R. Banerjee & B. Körtgen; 2) The Formation of Planetesimals: Building Bricks for Planetary Systems, by H. Klahr & A. Schreiber, and 3) Turbulence and Its Effect on Protostellar Disk Formation, by D. Seifried, R. Banerjee, R. S. Klessen.

These explorations show the influence of magnetic fields on the evolution of cloud cores, and proto-planetary disks; however, since instabilities and turbulence greatly reduces the angular momentum transport outwards in the disk by magnetic torques, there is still a fair amount of uncertainty about the quantitative description, as noted by the authors. It is interesting to note, that related physics have been invoked to explain why massive stars explode at all¹, and also there the effect of turbulence is a key that is not yet full explored to date.

New Radio Telescopes

Using amplitude and phase information, much of Western Europe has been transformed into a gigantic low frequency radio telescope, LOFAR, centred on the Netherlands: LOFAR: Calibration and Imaging on JURECA, by M. Hoeft, A. Horneffer, A. Drabent, & S. Fröhlich. This telescope can observe the trails of ultra high energy cosmic rays, and obtain their chemical composition, just as well as the ionosphere, and radio galaxies at cosmic distances. The data analysis requires a supercomputer.

This implies that with LOFAR we can observe one class of sources of ultra high energy cosmic ray particles, and also observe these same particles when they hit Earth, from start to finish. Radio galaxies are the most likely sources of ultra high energy cosmic ray particles, as predicted already 1963², via very strong shock waves caused by powerful relativistic jets that in turn also produce the spectacle on the sky, in the case of the radio galaxy Centaurus A 20 times larger than the Moon.

Conclusions

Apart from extreme plasma simulations, nowadays even astronomical data analysis needs supercomputers. In both aspects astronomy follows particle physics. In the case of LOFAR both the acceleration and the final interaction of ultra high energy cosmic ray nuclei on Earth can be observed.

References

1. G. S. Bisnovatyi-Kogan, *The Explosion of a Rotating Star As a Supernova Mechanism*, Astron. Zh. **47**, 813, 1970; transl. Sov. Astron. A.J. **14**, 652, 1971.
2. V. L. Ginzburg & S.I. Syrovatskii, *Cosmic Rays in Metagalactic Space*, Astron. Zh. **40**, 466, 1963; transl. Sov. Astron. A.J. **7**, 357, 1963.
3. R. Lüst, *Die Entwicklung einer um einen Zentralkörper rotierenden Gasmasse I. Lösungen der hydrodynamischen Gleichungen mit turbulenter Reibung*, Z.f. Nat. **7**, 87, 1952.

Formation of Star-Forming Clouds from the Magnetised, Diffuse Interstellar Medium

Robi Banerjee and Bastian Körtgen

Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany
E-mail: banerjee@hs.uni-hamburg.de

Molecular clouds, the birthplaces of stars in galaxies, form dynamically from the diffuse atomic gas of the interstellar medium (ISM). The ISM is also threaded by magnetic fields which have a large impact on its dynamics. In particular, star forming regions must be magnetically *super-critical* in order to accommodate gas clumps which can collapse under their own weight. Based on a parameter study of three dimensional magneto-hydrodynamical (MHD) simulations, we show that the long-standing problem of how such supercritical regions are generated is still an open issue.

1 Introduction

Present day stars form within the densest regions of molecular clouds (MCs) and giant molecular clouds (GMCs), in gravitationally unstable cores and clumps. Our common understanding is that those MCs and GMCs form from the diffuse, atomic (HI) gas within timescales of less than 10 Myr¹. The generation of filaments and substructures within GMCs is primarily controlled by magnetic fields and turbulence⁴. In particular, magnetic fields are an elemental part of the interstellar medium^{5,6} which have a large impact on the dynamics of the ISM on various spatial scales⁷ as the magnetic energy density is comparable to the thermal energy density of the ISM⁸.

One long-standing issue is the formation of supercritical clumps and cores. Similarly to thermal pressure, magnetic fields prevent contraction of otherwise (thermally) self-gravitating gas clumps if the magnetic fields are strong enough. Therefore, gaseous overdensities must be magnetically *supercritical*, quantified by the mass-to-flux ratio, μ , to collapse and to subsequently allow the formation of stars.

Already in 1956, Mestel & Spitzer realised that molecular clouds should be magnetically *subcritical* assuming field strengths that correspond to the equipartition of magnetic and kinetic energy density within the ISM⁹. To generate supercritical cloud cores out of those subcritical conditions, they suggested that the non-perfect coupling between charged particles and neutrals, i.e. the ambipolar diffusion (AD) drift, could locally increase the mass-to-flux ratio which allows the cloud to break up and to form stars. For a long time this was the standard theory of star formation out of the magnetised ISM^{2,10}. In this fairly static “standard model” of magnetically-supported, AD-mediated supercritical cores, low-mass stars would form by the slow gravitational contraction of isolated cores containing a very small fraction of the clouds’ mass. This picture would also account for the very low observed global star formation efficiency (SFE) of giant molecular clouds^{11,12}. The slow contraction results from the typical timescale for ambipolar diffusion, t_{AD} , which is an order of magnitude larger than the free-fall time, t_{ff} , of individual cloud cores (their ratio is about $t_{AD}/t_{ff} \approx 10(x_e/10^{-7})$, where x_e is the ionisation fraction). On the

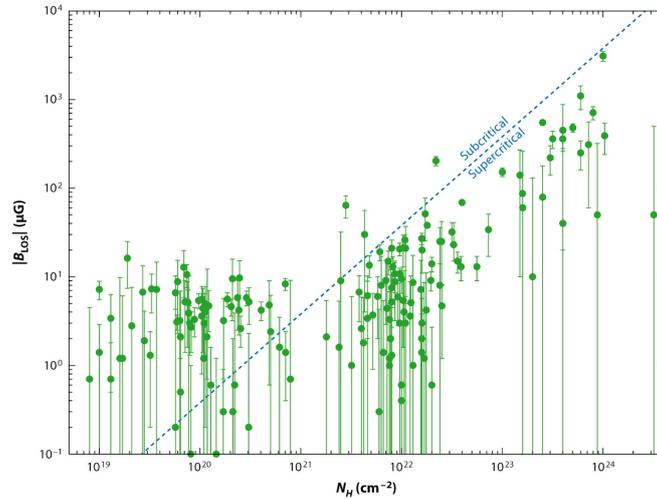


Figure 1. These observational data summarise the main motivation for our proposed study: How do subcritical (HI) clouds become supercritical (H_2) clouds? Our previous studies have shown that it is everything but trivial to build up supercritical clouds out of the magnetised interstellar medium, because the mass-to-flux ratio is fairly well conserved, even in the presence of ambipolar diffusion and enhanced non-ideal MHD setups^{26,27}. [From Crutcher (2012)⁵].

other hand, rather recent observations by Crutcher (2009)¹³ of individual cloud cores including Zeeman measurements to determine their magnetic field distribution indicate that idealised models of ambipolar-diffusion driven star formation are unlikely to be operative. In idealised models with ordered background magnetic fields, efficient ambipolar diffusion would lead to a local increase of the mass-to-flux ratio towards the centre of cloud cores which is not seen in their observed sample¹³ (but see also Bertram *et al.* (2012)¹⁴ on the difficulty to interpret those observations).

However, present-day models of star formation also account for the fact that molecular clouds are also pervaded by supersonic random motions, i.e. turbulence^{15,16}. Eventually, this resulted in a paradigm shift of the theory of star formation where magnetic fields only play a minor role and supersonic, super-Alfvénic turbulence controls the star formation efficiency within molecular clouds^{3,17}. As a consequence, the magnetic fields are expected to be highly disordered rather than being an ordered background field. Hence, idealised models of ambipolar diffusion drift should not apply¹⁸. Additionally, the AD characteristic timescale is expected to decrease in this case^{19,20} and other diffusive effects like turbulent reconnection might be operative²². Indeed, a number of studies have suggested that both MCs²³ and their clumps^{24,25} are close to being magnetically critical, with a moderate preference for being supercritical. Moreover, recent compilations of observational data show that cloud cores and clumps with column densities of $N \gtrsim 2 \times 10^{21} \text{ cm}^{-2}$ are essentially *all* supercritical (see Fig. 1).

Whether those supercritical cloud cores and clumps are the result of ambipolar diffusion together with random motions in the ISM is far from being certain and has to be investigated further. For instance, recently Heitsch & Hartmann (2014)²⁸ argued in their

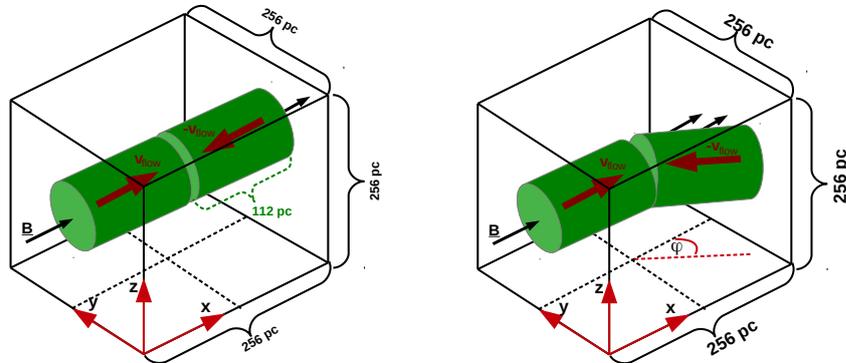


Figure 2. Setup of the initial conditions. *Left*: The well studied head-on collision where both flows are in the direction of the background magnetic field. *Right*: Flows with an inclination angle with respect to the background magnetic field. This setup resembles the impact of external driving of such flows (e.g. driving by SN blast waves). [From Körtgen & Banerjee (2015)²⁷].

parameter study that ambipolar diffusion in concert with turbulence is unlikely to control the formation of supercritical cores and hence star formation. They again propose an alternative scenario where large scale flows are the main driver to generate supercritical cores. This idea, where supercritical clouds could be assembled from large scale flows was already discussed in Mestel & Spitzer (1956)⁹ as an alternative to the AD-mediated scenario and to avoid the “magnetic flux problem”. But only in combination with supersonic turbulence this scenario becomes more feasible because gravitational fragmentation could be suppressed during the assembly of the clouds by those turbulent motions²⁹. This accumulation idea would also support a number of recent observations which show that magnetic fields are dynamically important on all scales in the Milky Way and other spiral galaxies^{30,7}. This is particularly evident from Fig. 1: The low column density HI gas is magnetically *subcritical*, whereas clouds which exceed columns of $N \gtrsim 2 \times 10^{21} \text{ cm}^{-2}$ are magnetically *supercritical*.

In the presented numerical parameter study, we investigated the possibility of diffusion mediated generation of *supercritical* clouds showing that it is unlikely that such unstable clouds can be built up from *subcritical* HI-clouds.

2 Numerical Method and Initial Conditions

For these studies we used the FLASH adaptive mesh refinement (AMR) code³¹. In addition to the basic ideal MHD equations (for which we employ the Bouchut solver^{33,34}) we also used the *ambipolar diffusion* module developed by Duffin & Pudritz (2008)³⁵. Additionally, self-gravity as well as heating and cooling processes were included in those simulations. For the latter, we followed the treatment by Koyama & Inutsuka (2002)³⁶ (an analytic simplification of their detailed calculation in Refs. 37,39). To capture the build-up of self-gravitating cores within the molecular clouds we used *sink particles*⁴⁰ in addition to the Jeans refinement criterion (i.e. the Truelove criterion⁴¹). In particular the detailed

sink particle approach allows us to unambiguously identify supercritical, collapsing regions which are important for our studies quantifying the star formation ability from the magnetised ISM.

Our initial setups for those studies are similar to the ones described in Refs. 38, 39 (see also Fig. 2) where the build-up of molecular clouds is modelled by the collision of cylindrical streams of warm neutral HI gas (WNM). Each flow is $l = 112$ pc long and has a radius of $r = 64$ pc. The bulk flows are slightly supersonic with typical Mach numbers of $\mathcal{M}_f = 2$. On top of those bulk motions, a turbulent velocity field is superimposed which triggers initial instabilities like the non-linear thin-shell instability (NTSI)⁴² and subsequently leads to fragmentation of the cloud. The initially uniform magnetic field has a strength of $B = \{3, 4, 5\}$ μG corresponding to mass-to-flux ratios of $\mu/\mu_{crit} \approx 1, 0.7, 0.6$ if the critical value $\mu_{crit} \approx 0.13/\sqrt{G}$ is applied⁴³.

Furthermore, we also studied the impact of an oblique angle of the flows with respect to the background magnetic field (see right panel of Fig. 2). Those oblique flows are more realistic than the head-on flows and could be generated, for instance, by supernova shock waves and by the gravitational potential of spiral arms. The motion of the flow at an inclination with respect to the magnetic field results in enhanced magnetic diffusivity (by numerical diffusion^{44,45}). Again, those flows resemble streams of the WNM in a thermally bistable configuration⁴⁶. The flows are studied with different oblique angles, which are varied from 10° to 60° , different initial magnetic fields strengths and different strength of the initial turbulence ranging from subsonic to supersonic velocity fluctuations. For details on the numerical setup and our initial conditions see Körtgen & Banerjee (2015)²⁷.

3 Results

As can be seen from Fig. 3, the different initial field strengths have significant implications for the resulting dynamical behaviour of the molecular cloud. The main difference comes about in efficiency to form stars (or not). In the case of a rather weak background field of $3 \mu\text{G}$ supercritical star forming clumps can be generated whereas in the case of a slightly stronger, but more realistic, magnetic field star formation is fully suppressed. Note that, due to the oblique flows with an angle of 60° the effective mass-to-flux ratios are 0.73 in the $3 \mu\text{G}$ case and 0.44 in the $5 \mu\text{G}$ case. That means that both cases are initially subcritical, but only in the cases of the weak magnetic field locally supercritical clumps are assembled due to sufficient flux loss.

An interesting point is also the field morphology. In the weak magnetic field case the field structure in the dense regions is clearly separated from the large scale magnetic field, whereas in the strong field case the field morphology is almost unaffected compared to the initial configuration (see the blue stream lines of Fig. 3). From an observational point of view, the field structure and its dynamical importance within molecular clouds is still debated. On the one hand, some multi-scale polarisation data indicate that magnetic fields in GMCs are essentially just dragged in from larger scales and are dynamically important^{7,30}. On the other hand, Zeeman measurements of individual cloud cores together with analyses of numerical simulations indicate rather weak fields that might not be dynamically important^{13,47}. With our subsequent studies on cloud formation on cloud scales including a more detailed modelling of ambipolar diffusion we hope to clarify this issue.

In Fig. 4 we quantify the main results by means of histograms in the N - B -plane from

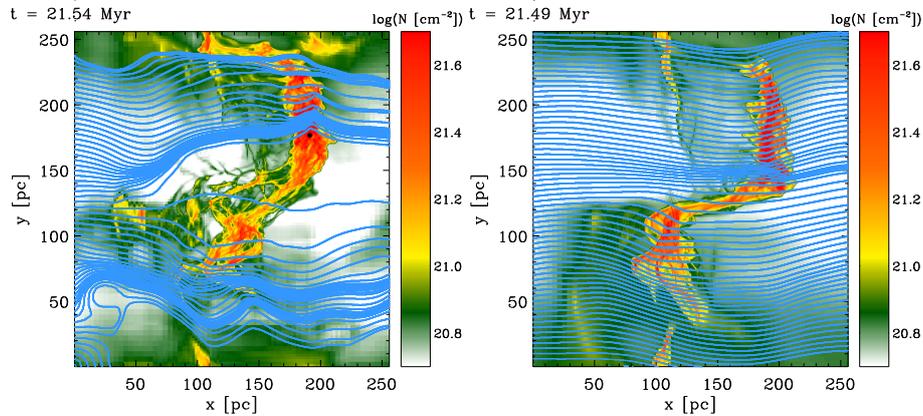


Figure 3. Results from colliding flow simulations investigating the formation of molecular clouds. Here, the flows collide with an oblique angle of 60° . *Left panel*: The weak field case ($3 \mu\text{G}$). The *Right panel* shows the same situations in the case of a stronger background magnetic field of $5 \mu\text{G}$. In the case of a weak magnetic field supercritical cloud cores can form that allow the formation of stars (marked with black dots). Stronger initial magnetic fields prohibit the formation of stars even in the case of large oblique angles of the flows. The blue stream-lines indicate the magnetic field morphology in the projected 2D plane. [From Körtgen & Banerjee (2015)²⁷].

our colliding flow studies for various initial conditions. Only in the *initially* marginally subcritical case ($B = 3 \mu\text{G}$) we observe signs of star formation within supercritical cores^a. For slightly stronger initial magnetic fields ($B \gtrsim 4 \mu\text{G}$) no supercritical cloud cores are generated, hence there is no star formation activity, regardless whether ambipolar diffusion is active or the flows collide with an oblique angle. Nevertheless the results of those simulations show the observed behaviour in the low column regime ($N \lesssim 1 \times 10^{21} \text{ cm}^{-2}$), where gas assembles along field lines without changing the field strength by much (see also the latest analysis from PLANCK observations of individual molecular clouds)³⁰. Only within supercritical, self-gravitating cores the magnetic field gets enhanced by compression due to flux freezing. Furthermore, we observe that star formation is immediately initiated, when the gas becomes supercritical promoting a picture of “rapid” star formation²⁹.

4 Conclusions

Here, we summarise our recent results from MHD simulations of colliding flows with varying initial conditions on the possible formation of supercritical cloud cores from subcritical initial conditions. Although dense clouds are easily formed within colliding flow scenarios due to thermal instability, the generation of supercritical clumps are largely determined by the initial conditions. Furthermore, increasing initial turbulence lead to lower masses of the cores and clumps because the HI streams become less coherent. Otherwise, increasing magnetic field strengths lead to more massive molecular clouds, which nevertheless do not become supercritical. Oblique flows still lead to cloud cores with masses comparable to

^aIf we assume $\mu_{crit} \approx 0.13/\sqrt{G}$ for spherical cores⁴³ we get $\mu/\mu_{crit} = 0.97 (3 \mu\text{G}/B)$ for our head-on colliding flow configurations.

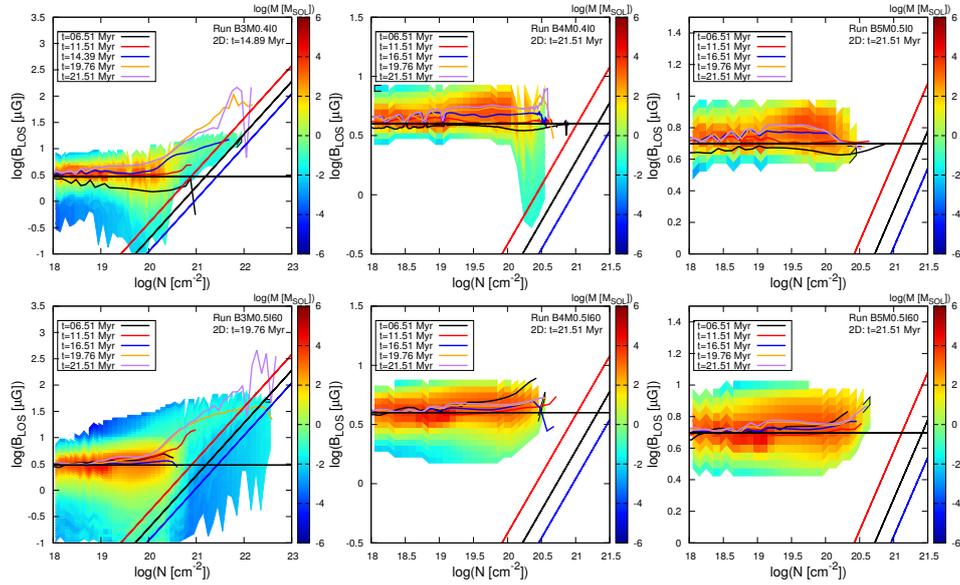


Figure 4. Results from colliding flow simulations with various different initial conditions. Shown are histograms of the line-of-sight field strength B_{LOS} as function of the column density N . From left to right: $B = 3 \mu\text{G}$, $B = 4 \mu\text{G}$ and $B = 5 \mu\text{G}$, respectively. Top: $\Phi = 0^\circ$, Bottom: $\Phi = 60^\circ$. Different line colours denote different times. Also shown are the criticality condition (red line¹⁸), corrected for projection effects (black line⁴⁸), and assuming equipartition of turbulent and magnetic fields (blue line⁴⁹). Colour coded is the mass distribution within this two parameter space. [from Körtgen & Banerjee (2015)²⁷].

what has been observed recently. But starting with subcritical HI flows, in no case the magnetic flux loss is sufficient to allow the build-up of supercritical cloud cores. Generally, increasing inclination of the flows lead to increasing diffusivity of the magnetic field. Again, regardless of the variation of the inclination, no tendency for faster accumulation of gas or faster transition to thermally dominated regions was seen in our simulations.

We therefore stress the role of magnetic fields in the context of molecular cloud and star formation. We point out the complete lack of supercritical regions for realistic initial field strengths. From the observational side, HI clouds may be supercritical as a whole, but their observed, dense subregions be subcritical.

Hence, the question remains, how magnetically *supercritical* cloud cores are formed?

Acknowledgements

RB acknowledges funding by the DFG via the Emmy-Noether grant BA 3706/1-1, the ISM-SPP 1573 grants BA 3706/3-1 and BA 3706/3-2, as well as for the grant BA 3706/4-1. The software used in this work was in part developed by the DOE-supported ASC/Alliance Center for Astrophysical Thermonuclear Flashes at the University of Chicago.

The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN⁵⁰ at Jülich Supercomputing Centre

(JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

References

1. L. Blitz, et al., *Giant Molecular Clouds in Local Group Galaxies*, Protostars and Planets V, 81–96, 2007, arXiv:astro-ph/0602600.
2. F. H. Shu, F. C. Adams, F. C., S. Lizano, *Star formation in molecular clouds - Observation and theory*, ARA&A, **25**, 23–81, 1987.
3. M.-M. Mac Low, R. S. Klessen, *Control of star formation by supersonic turbulence*, Reviews of Modern Physics, **76**, 125–194, 2004.
4. P. André, et al., *From Filamentary Networks to Dense Cores in Molecular Clouds: Toward a New Paradigm for Star Formation*, Protostars and Planets VI, 27–51, 2014, 1312.6232.
5. R. M. Crutcher, *Magnetic Fields in Molecular Clouds*, ARA&A, **50**, 29–63, 2012.
6. H.-B. Li, et al., *The Link Between Magnetic Fields and Cloud/Star Formation*, Protostars and Planets VI, 101–123, 2014, 1404.2024.
7. H.-B. Li, et al., *Self-similar fragmentation regulated by magnetic fields in a region forming massive stars*, Nature, **520**, 518–521, 2015.
8. C. Heiles, T. H. Troland, *The Millennium Arecibo 21 Centimeter Absorption-Line Survey. IV. Statistics of Magnetic Field, Column Density, and Turbulence*, ApJ, **624**, 773–793, 2005, arXiv:astro-ph/0501482.
9. L. Mestel, L. Spitzer Jr., *Star formation in magnetic dust clouds*, MNRAS, **116**, 503, 1956.
10. T. C. Mouschovias, G. E. Ciolek, *Magnetic Fields and Star Formation: A Theory Reaching Adulthood*, In C. J. Lada and N. D. Kylafis (eds.), NATO ASIC Proc. 540: The Origin of Stars and Planetary Systems, 305–+, 1999.
11. P. C. Myers, et al., *Molecular clouds and star formation in the inner galaxy - A comparison of CO, H II, and far-infrared surveys*, ApJ, **301**, 398–422, 1986.
12. N. J. Evans II, et al., *The Spitzer c2d Legacy Results: Star-Formation Rates and Efficiencies; Evolution and Lifetimes*, ApJS, **181**, 321–350, 2009, 0811.1059.
13. R. M. Crutcher, N. Hakobian, T. H. Troland, *Testing Magnetic Star Formation Theory*, ApJ, **692**, 844–855, 2009, 0807.2862.
14. E. Bertram, C. Federrath, R. Banerjee, R. S. Klessen, *Statistical analysis of the mass-to-flux ratio in turbulent cores: effects of magnetic field reversals and dynamo amplification*, MNRAS, **420**, 3163–3173, 2012, 1111.5539.
15. B. Zuckerman, N. J. Evans II, *Models of massive molecular clouds*, ApJ, **192**, L149–L152, 1974.
16. R. B. Larson, *Turbulence and star formation in molecular clouds*, MNRAS, **194**, 809–826, 1981.
17. P. Padoan, et al., *The Star Formation Rate of Molecular Clouds*, Protostars and Planets VI, 77–100, 2014, 1312.5365.

18. R. M. Crutcher, N. Hakobian, T. H. Troland, *Self-consistent analysis of OH Zeeman observations*, MNRAS, **402**, L64–L66, 2010, 0912 . 3024.
19. F. Heitsch, E. G. Zweibel, A. D. Slyz, J. E. G. Devriendt, *Magnetic Flux Transport in the ISM Through Turbulent Ambipolar Diffusion*, Ap&SS, **292**, 45–51, 2004.
20. T. Kudoh, S. Basu, *Formation of Collapsing Cores in Subcritical Magnetic Clouds: Three-dimensional Magnetohydrodynamic Simulations with Ambipolar Diffusion*, ApJ, **728**, 123, 2011, 1012 . 5707.
21. E. T. Vishniac, A. Lazarian, *Reconnection in the Interstellar Medium*, ApJ, **511**, 193–203, 1999.
22. G. Eyink, et al., *Flux-freezing breakdown in high-conductivity magnetohydrodynamic turbulence*, Nature, **497**, 466–469, 2013.
23. C. F. McKee, *Photoionization-regulated star formation and the structure of molecular clouds*, ApJ, **345**, 782–801, 1989.
24. R. M. Crutcher, T. H. Troland, B. Lazareff, G. Paubert, I. Kazès, *Detection of the CN Zeeman Effect in Molecular Clouds*, ApJ, **514**, L121–L124, 1999.
25. T. H. Troland, R. M. Crutcher, *Magnetic Fields in Dark Cloud Cores: Arecibo OH Zeeman Observations*, ApJ, **680**, 457–465, 2008, 0802 . 2253.
26. E. Vázquez-Semadeni, et al., *Molecular cloud evolution - IV. Magnetic fields, ambipolar diffusion and the star formation efficiency*, MNRAS, **414**, 2511–2527, 2011, 1101 . 3384.
27. B. Körtgen, R. Banerjee, *Impact of magnetic fields on molecular cloud formation and evolution*, MNRAS, **451**, 3340–3353, 2015, 1502 . 03306.
28. F. Heitsch, L. Hartmann, *Accretion and diffusion time-scales in sheets and filaments*, MNRAS, **443**, 230–240, 2014, 1406 . 2191.
29. L. Hartmann, J. Ballesteros-Paredes, E. A. Bergin, *Rapid Formation of Molecular Clouds and Stars in the Solar Neighborhood*, ApJ, **562**, 852–868, 2001, arXiv:astro-ph/0108023.
30. Planck Collaboration, et al., *Planck intermediate results. XXXV. Probing the role of the magnetic field in the formation of structure in molecular clouds*, ArXiv e-prints: 1502.04123, 2015, 1502 . 04123.
31. B. Fryxell, et al., *FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes*, ApJS, **131**, 273–334, 2000.
32. A. Dubey, et al., *Challenges of Extreme Computing using the FLASH code*, In N. V. Pogorelov, E. Audit, G. P. Zank (eds.), *Numerical Modeling of Space Plasma Flows*, vol. 385 of Astronomical Society of the Pacific Conference Series, 145–+, 2008.
33. F. Bouchut, C. Klingenberg, K. Waagan, *A multiwave approximate Riemann solver for ideal MHD based on relaxation II - Numerical implementation with 3 and 5 waves*, Numerische Mathematik, **115**, 647–679, 2010.
34. K. Waagan, C. Federrath, C. Klingenberg, *A robust numerical code for compressible magnetohydrodynamics and its application to highly supersonic turbulence*, J. Comp. Phys, submitted 2010.
35. D. F. Duffin, R. E. Pudritz, *Simulating hydromagnetic processes in star formation: introducing ambipolar diffusion into an adaptive mesh refinement code*, MNRAS, **391**, 1659–1673, 2008, 0810 . 0299.

36. H. Koyama, S.-I. Inutsuka, *An Origin of Supersonic Motions in Interstellar Clouds*, *ApJ*, **564**, L97–L100, 2002, arXiv:astro-ph/0112420.
37. H. Koyama, S.-I. Inutsuka, *Molecular Cloud Formation in Shock-compressed Layers*, *ApJ*, **532**, 980–993, 2000, arXiv:astro-ph/9912509.
38. E. Vázquez-Semadeni, et al., *Molecular Cloud Evolution. II. From Cloud Formation to the Early Stages of Star Formation in Decaying Conditions*, *ApJ*, **657**, 870–883, 2007, astro-ph/0608375.
39. R. Banerjee, E. Vázquez-Semadeni, P. Hennebelle, R. S. Klessen, *Clump morphology and evolution in MHD simulations of molecular cloud formation*, *MNRAS*, **398**, 1082–1092, 2009, 0808.0986.
40. C. Federrath, R. Banerjee, P. C. Clark, R. S. Klessen, *Modeling Collapse and Accretion in Turbulent Gas Clouds: Implementation and Comparison of Sink Particles in AMR and SPH*, *ApJ*, **713**, 269–290, 2010, 1001.4456.
41. J. K. Truelove, et al., *The Jeans Condition: A New Constraint on Spatial Resolution in Simulations of Isothermal Self-gravitational Hydrodynamics*, *ApJ*, **489**, L179+, 1997.
42. E. T. Vishniac, *Nonlinear instabilities in shock-bounded slabs*, *ApJ*, **428**, 186–208, 1994, astro-ph/9306025.
43. T. C. Mouschovias, L. Spitzer Jr., *Note on the collapse of magnetic interstellar clouds*, *ApJ*, **210**, 326–+, 1976.
44. F. Heitsch, L. Hartmann, A. Burkert, *Fragmentation of Shocked Flows: Gravity, Turbulence and Cooling*, *ArXiv e-prints*, **805** (2008), 0805.0801.
45. T. Inoue, S.-I. Inutsuka, *Two-Fluid Magnetohydrodynamic Simulations of Converging HI Flows in the Interstellar Medium. I. Methodology and Basic Results*, *ApJ*, **687**, 303–310, 2008, 0801.0486.
46. G. B. Field, *Thermal Instability.*, *ApJ*, **142**, 531–+, 1965.
47. P. Padoan, et al., *MHD Turbulence In Star-Forming Clouds*, In G. Bertin, F. de Luca, G. Lodato, R. Pozzoli, M. Romé (eds.), American Institute of Physics Conference Series, vol. 1242 of American Institute of Physics Conference Series, 219–230, 2010.
48. F. H. Shu, A. Allen, H. Shang, E. C. Ostriker, Z.-Y. Li, *Low-Mass Star Formation: Theory*, In C. J. Lada, N. D. Kylafis (eds.), NATO Advanced Science Institutes (ASI) Series C, vol. 540 of NATO Advanced Science Institutes (ASI) Series C, 193, 1999.
49. C. F. McKee, E. G. Zweibel, A. A. Goodman, C. Heiles, *Magnetic Fields in Star-Forming Regions - Theory*, In E. H. Levy, J. I. Lunine (eds.), *Protostars and Planets III*, 327–+, 1993.
50. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, *Journal of large-scale research facilities*, **1**, A1, 2015.

The Formation of Planetesimals: Building Bricks for Planetary Systems

Hubert Klahr and Andreas Schreiber

Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
E-mail: {klahr, aschreiber}@mpia.de

The asteroids^a and Kuiper Belt objects^b are left overs of building material for our earth and the other planets in our solar systems from 4.567 billion years ago. Those typically 100 km large objects are called planetesimals, built up from icy and dusty grains. In our current paradigm of planet formation it was turbulent flows and metastable flow patterns like zonal flows and vortices that concentrated mm to cm sized grains in sufficient numbers that a streaming instability and gravitational collapse of these particle clumps was triggered. The entire picture is known as gravoturbulent formation of planetesimals. What was missing until recently was a physically motivated prediction on the typical sizes at which planetesimals should form via this process. Our old simulations on JUGENE in Jülich had only shown a correlation between numerical resolution and planetesimal size and thus no answer was possible. But with the latest series of simulations on JUQUEEN¹ covering all the length scales down to the physical size of actual planetesimals we were able to obtain values for the turbulent particle diffusion as a function of the particle load in the gas. Thus we have all necessary data at hand to feed our back of the envelope calculation that predicts the size of planetesimals as result of a competition between gravitational concentration and turbulent diffusion. Using the diffusion values obtained in the numerical simulations on JUQUEEN predicts planetesimal sizes on the order of 100 km, which luckily coincides with the measured data from both asteroids² as well from Kuiper Belt objects³.

1 Introduction

Planet formation is a beneficial side effect of star formation. It is the gas and dust around young stars that does not get accreted directly during the collapse of a cloud core due to angular momentum conservation that will form a planetary system. The goal of our current research activities is to better understand the properties and diversity of planets around distant stars as well as in our own solar system. Starting with numerical simulations of the star formation process we aim to understand the stages of disk formation and thus via a calibration on actual disk observations the initial mass and angular momentum distribution of disks as initial conditions for planet formation. In these disks we study all possible sources of turbulent disk evolution from self-gravity, via magnetic fields to classical hydrodynamical instabilities. Latest developments in the field include deriving various observables for different sources of turbulence, which in conjunction with simulations of planet disk interaction are used to interpret the wealth of observations coming online from interferometers working with many kilometre sized base-length in the submillimetre band (e.g. the Atacama Large Millimeter Array: ALMA⁴) to modern high resolution and high contrast cameras on the largest optical telescopes (e.g. SPHERE⁵). Lastly we follow all evolutionary steps of planets from small dust grains and snow flakes via planetesimals to

^aMore precisely: objects in the main asteroid belt.

^bMore precisely: objects in the cold classical Kuiper Belt.

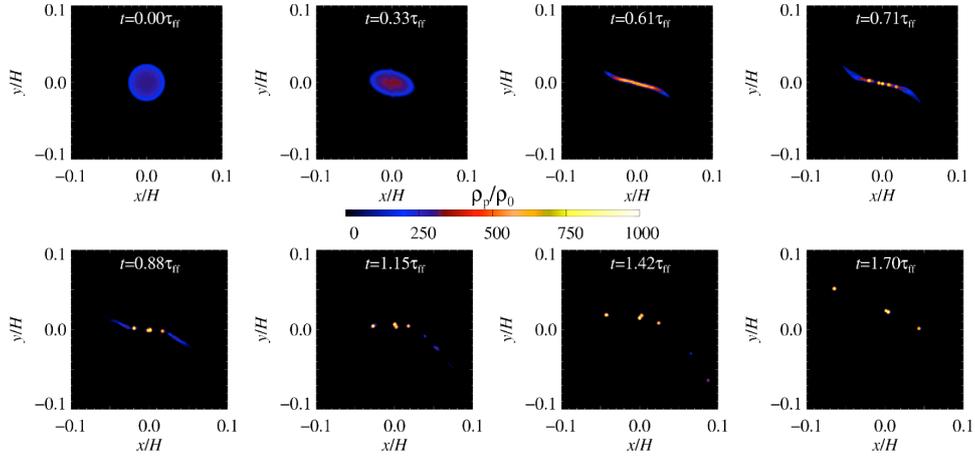


Figure 1. A series of snapshots following the collapse of a particle heap at 10 % of the local Roche density (see text) under its own gravity embedded in an accretion disk around a young star. Colours indicate the particle density. The initial spherical cloud gets sheared apart by Keplerian shear, or in other words by the tidal forces exerted from the star, because its internal density is lower than the local Roche density. Then its local subregions that exceed the Roche density, collapse into a group of many planetesimals around 100 km in size. This 3D Hydro Dynamical Simulation was performed with the Pencil Code on the JUQUEEN cluster (see PhD thesis by Karsten Dittrich¹⁰).

rocky and gaseous planets, partly in order to feed population synthesis with better data, partly to better interpret disk observations, and partly to study the chemical composition of planets - a question of special relevance to understand the formation of life on earth and elsewhere.

In a series of papers⁶⁻⁸ we explained the formation of planetesimals via the process of gravoturbulence. Here it is magnetic fields that trigger the formation of zonal flows, which in turn concentrate particles up to dust to gas ratios of order unity. Then a second instability kicks in, based on the feedback of the dust onto the gas - the streaming instability. The action of this instability is twofold: A: it concentrates the dust even further up to values that locally exceed the Roche density^c and B: destroying local particle concentration via turbulent diffusion on small scales. Exceeding the Roche density is the condition that self gravity is locally stronger than the tidal forces from the star, i.e. object or particle clumps of lower density than the Roche density will not collapse but get tidally sheared apart (see Fig. 1 from the PhD thesis of Karsten Dittrich¹⁰). Also comets are known to break apart once they pass the gravity field of a planet in a way that their internal density is lower than the local Roche density set by the planets mass and the distance from its centre (see the breakup of comet Shoemaker-Levy 9 (SL9)⁹ and its following impact on Jupiter). In the case that gravity wins, e.g. that the local density exceeds the Roche density, a planetesimal can form. In Fig. 1 one finds that even the initial cloud is sheared apart, it forms fragments, which individually can exceed the Roche density and then form planetesimals. Unfortunately the regions that exceed the Roche density are at the resolution limit of our

^cRoche density is the density that a gravity bound body needs to withstand the tidal forces of a second more massive body, for instance a close by planet or here: the sun.

numerical grid. Therefore we started looking for a physical criterion on the smallest possible fragments, that would withstand the tidal forces, while still resolving the turbulence of the gas.

A particle clump at Roche density without any internal pressure will collapse on the free fall time $t_{\text{ff}} = 0.64\Omega^{-1}$ with Ω being the local Keplerian frequency around the central star. Yet as particles feel friction with the gas during their gravitational contraction, this time gets longer proportional to the friction time τ_f^{d} defined via $\dot{v} = -\frac{\delta v}{\tau_f}$. For typical particle sizes in protoplanetary disks the dimensionless friction time is given as the Stokes Number^e $St = \tau_f\Omega$ with values of about $St = 0.1$. Our numerical simulations¹¹ on particle growth and disk evolution have shown that larger Stokes numbers, i.e. larger particles will rarely be available, because they either rain out towards the sun, or destroy each other by collisional fragmentation.

In the case of Stokes numbers smaller than unity the particle clump contraction time (or sedimentation time for the case of self gravity) is on the order of $t_s \approx \frac{t_{\text{ff}}}{St}$ ¹². Yet if the local dust concentration is diffused via the turbulent dust and gas motion with diffusivity D , a clump of size l gets diffused on the typical time scale of

$$t_{\text{d}} = l^2/D. \quad (1)$$

If one compares now the time scale of diffusion versus sedimentation $t_{\text{d}} = t_s$ one can derive a size prediction of clumps to withstand internal diffusion at

$$l \approx \sqrt{\frac{D}{St\Omega}}. \quad (2)$$

In combination with the knowledge that the clump had already Roche density one can determine the mass of the resulting planetesimal as $m_P = \frac{4}{3}\pi l^3$. What remains to be determined is the diffusivity D for typical dust-to-gas-ratios and length-scales at which planetesimal formation is thought to occur.

2 High Resolution Studies of the Streaming Instability

Hydro- and magneto-hydro-dynamical instabilities lead to non-laminar flows within protoplanetary disks^{13,14}. Many of those non-laminar features are known to produce dust overdensities via the local trapping of radially sedimenting particles (sedimentation towards the central star), such as zonal flows, and vortices, resulting in axisymmetric and non-axisymmetric structures¹⁵. Though they are short-lived in terms of viscous disk evolution, they are long-lived in terms of the disk dynamics. In particular, the concentration of dust as well as the collapse of a particle clump to a planetesimal is supposed to occur within a few local orbits. As shown in several papers in the last years particle overdensities are self amplifying over a certain threshold due to the streaming instability (SI)¹⁶, which was numerical investigated at least for the largest possible scales ($\approx 0.1H$)¹⁷.

^dThe friction time describes the friction between a particle of given mass and size with the surrounding gas. If particles are smaller than the mean free path of the gas, the drag formulae for the Epstein regime have to be applied, otherwise the aerodynamic, i.e. Stokes drag is valid.

^eIn astrophysics the Stokes Number is defined on the global dynamical time scale of the system, i.e. the orbital frequency of the accretion disk, rather than via the unknown timescale relevant for the dissipation scale. Particles of $St = 1$ are then the fastest radially sedimenting objects in the radially pressure supported accretion disk.

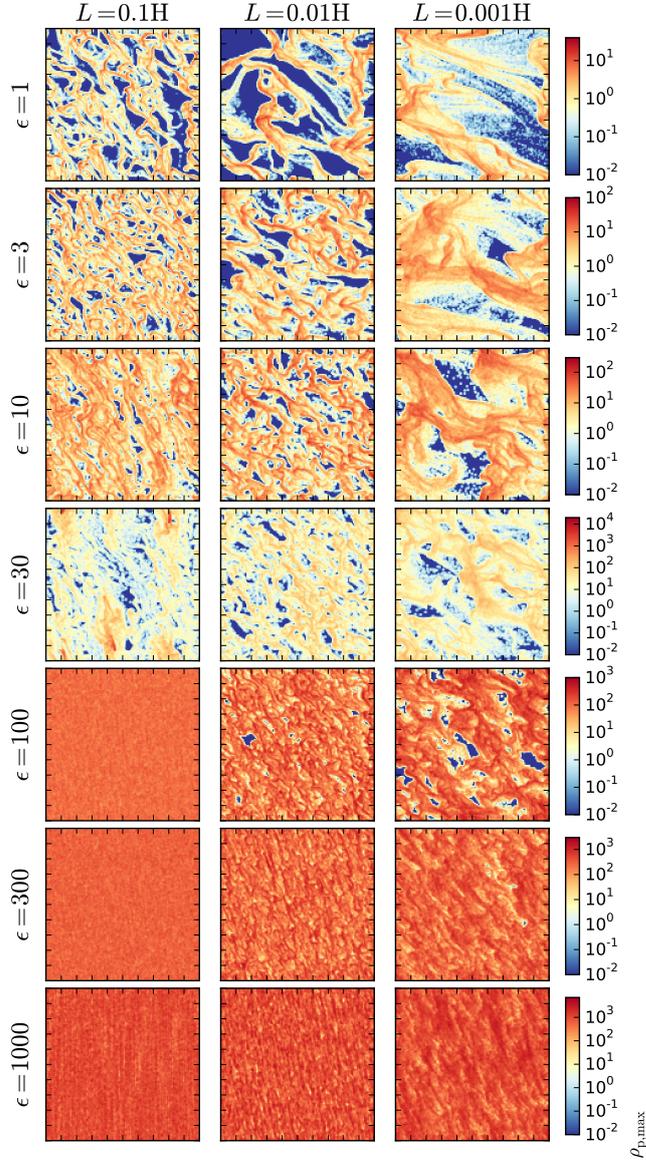


Figure 2. Overview of our parameter study to estimate diffusivities for the streaming instability on small physical length scales $L = 0.1 H$ to $L = 0.001 H$. We thereby investigate high dust to gas density ratios ϵ from 1 up to 1000, to mimic a wide range of phases within a collapsing particle clump. Colour represents the local dust to gas ratio with red high and blue low values.

Following this idea planetesimal collapse must happen in areas where SI is already active and will still be active during the collapsing phase. Here one can expect that for

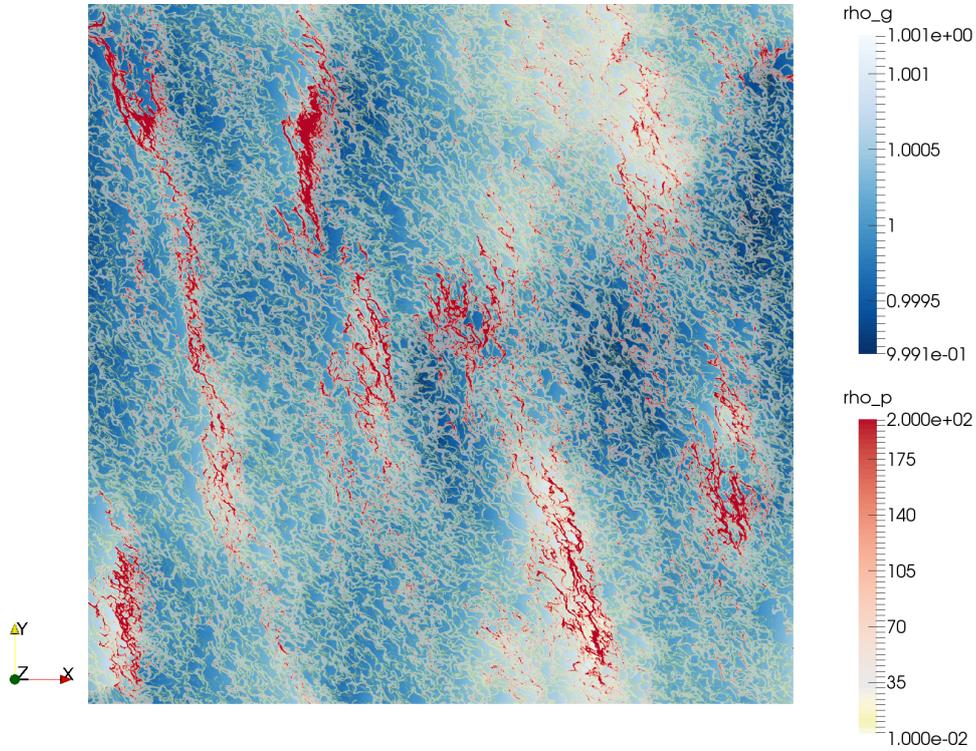


Figure 3. This snapshot shows the largest streaming instability investigation so far, with a resolution of 1250^2 and $1.6e7$ swarm particles. This snapshot shows in blue the underlying gas density field. On top of it in red to white is plotted the particle density. The particles show a typical streaming instability pattern over the whole domain (white) and additionally (red) particle dominated regions which are long lived and highly particle concentrated.

large dust-to-gas-ratios $\epsilon = \rho_p/\rho_g$ the SI should become weaker, because in the extreme case of having only dust, there are no more hydro-instabilities, and this is indeed what we see for numerical simulation at high ϵ -values.

Thus in a forthcoming paper we investigate if the SI can lead to diffusivities D or diffusion times expressed in local orbits that can delay or even prohibit the collapse to planetesimals via internal diffusion of the clump. Since the dust-to-gas-density-ratio ϵ will increase to very high numbers in the collapse phase, we will exceed ϵ -values previously investigated about several orders of magnitude see Fig. 2 for our scanned parameter space. Our boxes range from $L = 0.1H$ down to $L = 0.001H$ and the ϵ values increase from $\epsilon = 1$ to $\epsilon = 1000$. Depending on the local gas density, which is a function of time and space, the Roche density can be reached at values of $\epsilon = 10 - 1000$. Our 3D simulations use typically several Mio. CPU-hours per parameter set, which explains that scanning the necessary parameter space is a very expensive endeavour. We found that for increasing dust-to-gas-ratios the SI was getting indeed a little weaker, but nevertheless did not die out. At the same time the unstable wave-lengths were squeezed to smaller scales, e.g. the scales interesting for the final collapse - sedimentation phase of the planetesimal (see Fig. 4).

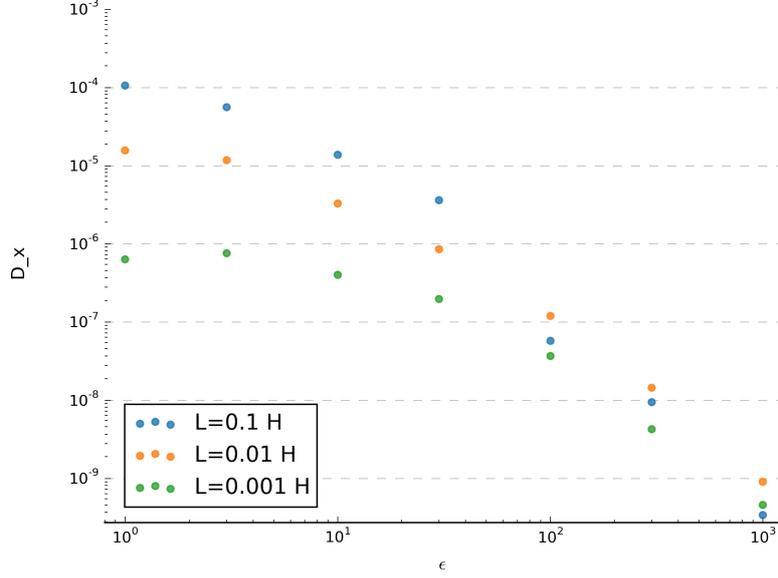


Figure 4. One of our results from our parameter study is a dust to gas ratio ϵ dependence of the diffusivity. This is the input for our planetesimal collapse scenario and will give us the ability to calculate planetesimal properties, such as size and rate.

3 Size Prediction for Planetesimals

The dust-to-gas-ratio that corresponds to the local Roche density depends on the distance from the star as well as on the actual gas content of the disk as a function of the evolution of the disk. As a typical value we choose $\epsilon = 10$ for distances of about 40 AU (= Astronomical Unit = distance Sun-Earth). The measured diffusivities at the medium scale $L = 0.01H$ in dimensionless units are $D = 2.7 \times 10^{-6}$. This value for D can be plugged into our prediction for planetesimal size a :

$$a \approx \frac{H/R}{0.03} \sqrt{\frac{D}{2.7 \times 10^{-6}}} \sqrt{\frac{0.1}{St}} 88\text{km}. \quad (3)$$

This result is remarkably close to the measured sizes of the knee in the observed population of both asteroids and planetesimals and a nice support for our theory of planetesimal formation. An improved prediction of the actual gas content of the disk and the Stokes number of the available dust grains can still change the result, yet not by orders of magnitude.

4 Outlook

Of course, we have to refine our methods in several ways. For instance, we have to study the streaming instability for other particle sizes (e.g. St) as well as for entire particle size distributions. We also have to study the effect of particle-particle collisions in a better way, as they will eventually start to dominate, once we resolve the actual dimensions of the resulting planetesimal.

On the other hand we will surely test our back of the envelope estimate for the criterion of planetesimal formation via straight forward 3D numerical simulations of the streaming instability and gravitational collapse along the parameters as predicted from our non-self-gravitating runs. Then a high resolution case of fully developed SI as already obtained (see Fig. 3) will be the initial state, for which we “simply” have to switch on self-gravity and see A: whether the box size allows for collapse, e.g. the unstable wavelength as determined from diffusivity and self gravity fits into the simulation domain and B: whether the resulting planetesimals obey our size prediction, hopefully independent on further increasing the resolution, i.e. reaching convergence with resolution for our simulations. Then we will be able to say how the planetesimals in our solar system can have formed and explain the size of asteroids and Kuiper belt objects.

After that we can start investigating at what rate they were forming in the solar nebula. When did they form at what distance from the start? Where did their building material come from? What kind of planetesimals did form the earth and does it help us to understand the chemical composition of the earth? Finally this leads to the question of what were the chemical conditions on a young earth, which one day started to harbour life.

Acknowledgements

We are indebted to Anders Johansen, Wladimir Lyra, Hans Baehr and Karsten Dittrich for many fruitful discussions. This research has been supported by the Deutsche Forschungsgemeinschaft Schwerpunktprogramm (DFG SPP) 1385 “The first ten million years of the solar system” under contract KL 1469/4-(1-3) “Gravoturbulente Planetesimal Entstehung im frühen Sonnensystem” and by (DFG SPP) 1833 “Building a Habitable Earth” under contract KL 1469/13-1 “Der Ursprung des Baumaterials der Erde: Woher stammen die Planetesimale und die Pebbles? Numerische Modellierung der Akkretionsphase der Erde.”

The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time for a GCS Large-Scale Project (additional time through the John von Neumann Institute for Computing (NIC)) on the GCS share of the supercomputer JUQUEEN¹ at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

References

1. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities, **1**, A1, 2015, <http://dx.doi.org/10.17815/jlsrf-1-18>.
2. W. F. Bottke, D. D. Durda, D. Nesvorny, R. Jedicke, A. Morbidelli, D. Vokrouhlicky, & H. F. Levison, *Linking the collisional history of the main asteroid belt to its dynamical excitation and depletion*, Icarus **179**(1), 63-94, 2005.

3. D. Nesvorný, D. Vokrouhlický, W. F. Bottke, K. Noll, & H. F. Levison, *Observed Binary Fraction Sets Limits on the Extent of Collisional Grinding in the Kuiper Belt*, *The Astronomical Journal* **141**(5), 159, 2011.
4. ALMA Partnership, C. L. Brogan, L. M. Perez, T. R. Hunter, W. R. F. Dent, A. S. Hales, et al., *The 2014 ALMA Long Baseline Campaign: First Results from High Angular Resolution Observations toward the HL Tau Region*, *The Astrophysical Journal Letters* **808**(1), L3, 2015.
5. M. Benisty, A. Juhasz, A. Boccaletti, H. Avenhaus, J. Milli, C. Thalmann, et al., *Asymmetric features in the protoplanetary disk MWC 758*, *Astronomy & Astrophysics* **578**, L6, 2105.
6. A. Johansen, H. Klahr, T. Henning, *Gravoturbulent Formation of Planetesimals*, *The Astrophysical Journal* **636**, 1121-1134, 2006.
7. A. Johansen, J. S. Oishi, M.-M. Mac Low, H. Klahr, T. Henning, A. Youdin, *Rapid planetesimal formation in turbulent circumstellar disks*, *Nature* **448**, 1022-1025, 2007.
8. A. Johansen, H. Klahr, T. Henning, *High-resolution simulations of planetesimal formation in turbulent protoplanetary discs*, *Astronomy & Astrophysics* **529**, A62, 2011.
9. E. Asphaug, W. Benz, *Size, Density, and Structure of Comet Shoemaker-Levy 9 Inferred from the Physics of Tidal Breakup*, *Icarus* **121**, 225-248, 1996.
10. K. Dittrich, *Numerical Simulations of Planetesimal Formation in Protoplanetary Disks*, Uni Heidelberg, PhD-thesis, <http://www.ub.uni-heidelberg.de/archiv/15664>.
11. T. Birnstiel, H. Klahr, B. Ercolano, *A simple model for the evolution of the dust population in protoplanetary disks*, *Astronomy & Astrophysics* **539**, A148, 2012.
12. K. Shariff, J. N. Cuzzi, *The Spherically Symmetric Gravitational Collapse of a Clump of Solids in a Gas*, *The Astrophysical Journal* **805**, 42, 2015.
13. H. H. Klahr, P. Bodenheimer, *Turbulence in Accretion Disks: Vorticity Generation and Angular Momentum Transport via the Global Baroclinic Instability*, *The Astrophysical Journal* **582**, 869-892, 2003.
14. K. Dittrich, H. Klahr, A. Johansen, *Gravoturbulent Planetesimal Formation: The Positive Effect of Long-lived Zonal Flows*, *The Astrophysical Journal* **763**, 117, 2013.
15. M. Flock, J. P. Ruge, N. Dzyurkevich, T. Henning, H. Klahr, S. Wolf, *Gaps, rings, and non-axisymmetric structures in protoplanetary disks. From simulations to ALMA observations*, *Astronomy & Astrophysics* **574**, A68, 2015.
16. A. Youdin, A. Johansen, *Protoplanetary Disk Turbulence Driven by the Streaming Instability: Linear Evolution and Numerical Methods*, *The Astrophysical Journal* **662**, 613-626, 2007.
17. A. Johansen, A. Youdin, *Protoplanetary Disk Turbulence Driven by the Streaming Instability: Nonlinear Saturation and Particle Concentration*, *The Astrophysical Journal* **662**, 627-641, 2007.

LOFAR: Calibration and Imaging on JURECA

Matthias Hoeft¹, Andreas Horneffer², Alexander Drabent¹, and Stefan Fröhlich³

¹ Thüringer Landessternwarte, Sternwarte 5, 07778 Tautenburg, Germany
E-mail: hoeft@tls-tautenburg.de

² Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany

³ Forschungszentrum Jülich, Jülich Supercomputing Centre, 52425 Jülich, Germany

The Low Frequency Array (LOFAR) is a novel radio telescope, operating at very low radio frequencies. It uses a large frequency bandwidth and many stations comprising a multitude of simple dipole antennas. LOFAR stations are relatively small, giving them a large field-of-view (FoV). Moreover, the stations are distributed over many European countries, which gives LOFAR a superb image resolution. With its large bandwidth, large collecting area, large FoV, and high spatial resolution LOFAR achieves unprecedented sensitivity, resolution and in particular high survey speed at this little explored frequency range. The LOFAR data processing is realised via digital electronics starting from the signals of individual dipoles. Each station produces a tremendous amount of data to be sent to a central correlator via fast internet connections, which again leads to huge amounts of data to be processed further. For LOFAR the analysis is complicated by the ionosphere which can severely disturb the observations. Moreover, the simplicity of the antenna fields introduces distortions that need to be corrected when processing the data. Traditionally, the analysis of radio interferometer data has been carried out interactively on work stations. The enormous amounts of data produced by LOFAR and the complex data reduction schemes demand much larger computing resources and automated processes. The data reduction, thus, needs to be carried out on supercomputers. Our aim is to adapt the LOFAR software for analysing data on JURECA, to develop a framework which allows astronomers to set up a data reduction including observation specifics. Moreover, with selected observations we demonstrate the feasibility of reducing LOFAR data on general-purpose, multi-user supercomputers such as JURECA. The algorithms developed for LOFAR pave the way to a new generation of powerful radio telescopes at low frequencies.

1 The Low Frequency Array

The *Low Frequency Array* (LOFAR) is a novel radio telescope³, mainly developed and operated by the Dutch Institute for Radio Astronomy, ASTRON. LOFAR observes in the largely unexplored frequency range from 10 to 240 MHz. Instead of having large dishes it consists of simple dipole antennas grouped into stations. With several new stations becoming operational in 2015 (one close to Hamburg, three in Poland) LOFAR now comprises more than 100.000 dipoles distributed over Dutch and international stations. Each station operates as a phased array, i.e. the beam and the direction of the beam is set by delays when combining the signals of the individual dipoles (*beam-forming*). LOFAR allows us to observe in many directions at the same time, since the delays are set when processing the data: the processes can run in parallel as often as the processing and data transport resources permit. The beam-formed data are sent from each station to a central correlator via fast internet connections, e.g. international stations are equipped with 10 Gbit/s data links. Via cross-correlations and averaging the correlator combines the incoming data to *visibility* data.

The collecting area and the frequency coverage is maximised to increase the telescope sensitivity. LOFAR samples the signals from each antenna with a rate of 200 MHz, thus

the maximum frequency window allowed by the Nyquist sampling theorem is 100 MHz. The maximum bandwidth a LOFAR station can transmit is 96 MHz, which can be spread over several directions. On the other hand, a very high frequency resolution is necessary to mitigate Radio Frequency Interferences (RFI) and to avoid smearing out signals on long baselines provided by international LOFAR stations. A typical frequency resolution used in LOFAR observations is 3 kHz. The number of frequency channels, N_k , is consequently of the order of 30,000. The collecting area is determined by the number of dipoles used. The diameter of the station determines also the size of the station beam, i.e. the area on the sky visible with one pointing (field-of-view, FoV). The larger the FoV the less pointings are necessary to survey the entire sky. The diameter of each station is chosen as a compromise between survey speed and sensitivity. LOFAR currently uses stations with a diameter of about 50 and 70 metres and 50 stations in total (38 stations in the Netherlands, six stations in Germany, three in Poland, and one in France, UK and Sweden each). The large frequency coverage and the large collecting area allows LOFAR to increase in sensitivity with respect to earlier telescopes by a factor of about 100.

The resolution of the instrument is given by the largest distances between the stations. With stations in UK, France, Sweden and Poland LOFAR has baselines of thousands of kilometres, permitting sub-arcsecond resolution in this frequency regime. Within one observation a FoV of several degrees can be imaged, see Fig. 1, with a resolution even below one arcsecond. The image fidelity also depends on baselines with intermediate distances. Covering baselines from several tens of metres to thousands of kilometres LOFAR allows us to reconstruct the sky brightness distribution with unprecedented imaging capabilities.

The fundamental operation of a radio interferometer is to cross-correlate and average the beam-formed signals of all stations. More precisely, the signals of all pairs of stations are multiplied and averaged by the central correlator. As noted above, the sampling rate for LOFAR is 200 MHz, the typical averaging time is one or two second(s). For each antenna pair, for each frequency channel, and for each averaging time step the correlator computes one complex number (in fact four complex numbers since we have to deal with two polarisations), $\mathcal{V}_{ij}(t_n, \nu_k)$, where ij refers to the baseline between the stations i and j , t_n refers to the time step n , and ν_k refers to the frequency channel k . Initially, a BlueGene/L supercomputer, ranking number six in the TOP 500^a list of the fastest international supercomputers when built, was used as correlator. In 2008 it was replaced with a BlueGene/P. The correlation process itself is excellently suited for Graphics Processing Units (GPUs). So when support from IBM for the BlueGene/P ran out, a NVIDIA Tesla K10 cluster, named COBALT, has taken over the task of correlating the LOFAR station signals. The maximum input rate is 240 Gbit/s and the maximum output rate is 80 Gbit/s. The total number of visibilities produced for a typical 12 hours observation is

$$N_{\text{pol}} \times \frac{N_{\text{st}}(N_{\text{st}} - 1)}{2} \times N_n \times N_k, \quad (1)$$

with N_{pol} the number of polarisations, N_{st} the number of stations, N_n the number of integration time steps, and N_k the number of frequency channels. In total this implies 10^{13} complex visibilities. Data of the order of 100 TByte have to be stored and further processed for one typical LOFAR observation that includes all available stations.

^a<http://www.top500.org/lists/2005/06/>

2 Reconstructing the Sky Brightness Distribution

LOFAR is a very flexible instrument, allowing quite different modes of operation. The instrument is primarily operated in the ‘interferometer mode’ as described above with the aim to reconstruct the sky brightness distribution in the frequency range used by LOFAR. In an ideal case the visibility data, as generated by the correlator, depend on the sky brightness distribution, $I(l, m, n)$, according to:

$$\mathcal{V}_{ij}(u, v, w) = \int \frac{dl dm}{\sqrt{1-l^2-m^2}} \sqrt{A_i(l, m)A_j(l, m)} I(l, m) \times e^{-2\pi i [ul+vm+w(\sqrt{1-l^2-m^2}-1)]}, \quad (2)$$

where l, m gives the sky direction, u, v, w encodes the separation and direction from station i to station j , and $A_i(l, m)$ the station sensitivity into direction l, m , (*station beam*). In case of $l, m \ll 1$ (or $w = 0$) the relation between sky brightness distribution and visibility data becomes very close to a Fourier transform. Therefore, with a planar array, a small FoV, and no distortions we could simply apply a Fourier transform to reconstruct the sky brightness

$$\text{Visibility} \xleftrightarrow{\mathcal{FT}} \text{Sky Brightness} \quad (3)$$

However, there are many aspects which need to be taken into account when reconstructing the sky brightness from LOFAR data.

2.1 Data Preparation

Man-made radio signals may affect or dominate measured visibilities. This RFI typically occurs either at certain times and affect the entire spectrum (broad band RFI), or may be present for long times but only in a small frequency range (narrow band RFI, e.g., transmitters for communication). Often RFI also shows a very complex time-frequency behaviour. It is essential to remove RFI (*flagging*) before further processing the data since it may significantly distort the sky brightness reconstruction. The size of the LOFAR data prohibit a manual inspection, so RFI needs to be identified automatically with using the high frequency and time resolution as produced by the correlator. For LOFAR the special software (AOFlagger), which allows to automatically flag the data affected by RFI⁷, was developed.

One undesired effect of the phased array technique at LOFAR is that the station beams have strong side lobes. Therefore, bright radio sources many degrees away from the actual target may significantly affect the visibilities when located at a side lobe. This causes rapid visibility variation in frequency and time direction. For LOFAR an algorithm called *demixing* has been developed which allows us to subtract those sources before averaging¹⁰.

2.2 Nobody Is Perfect: Calibration

Signal modifications may be caused by the amplifiers and the processing of the analog signal before it gets digitised. Since each antenna has its own electronics which might be affected by the environment, e.g., variations of temperature or humidity, these distortions

are impossible to predict perfectly. In addition to changing the signal as such, the variations between the dipoles in a phased array like a LOFAR station also affect the shape of the station beam.

At the low radio frequencies used by LOFAR the ionosphere has a significant impact. The electron column density along the line of sight determines the “refraction index” of the ionosphere, with two dominant effects on the waves propagating through the ionosphere: introducing a delay and rotating the polarisation direction. Since the column density varies with time and position, the disturbances of the visibilities are extremely difficult to determine. So correction factors have to be derived from the observations, since the ionosphere cannot be measured directly with sufficient accuracy. These correction factors vary from station to station, with time, and across the field of view.

The signal propagation from the source to the telescope can be represented as a series of 2×2 complex matrices⁴. Typical parameters are amplitude and phase for each of the two polarisations and a polarisation rotation angle, all for each station, resulting in hundreds of parameters to solve for in each time and frequency interval. Calibrating on the measured data itself requires a reasonable model of the sky brightness distribution. One calibration step is done by solving for parameters in these matrices that can convert this model to the measured values and then inverting the matrices and multiplying them to the measured data.

Calibration of interferometer data usually is direction independent. It uses only one set of parameters for the full field of view. In contrast to many other radio telescopes LOFAR has a large field of view –which increases the survey speed– but in particular the ionosphere introduces disturbances which vary across the field of view. Dealing with these direction dependent effects is done by solving for solutions in the direction of several bright sources within the field of view and then interpolating between those directions⁶ or imaging only small facets around the bright sources. Correcting for these direction dependent effects is a major task when processing LOFAR data on JURECA, and is a field of ongoing research.

2.3 Imaging and Deconvolution

As any other radio interferometer, LOFAR measures only visibilities at u, v, w coordinates determined by the antenna positions. Hence, the uv -plane is only sparsely sampled. This can be denoted with a function $B(u, v)$, which is 1 for (u, v) pairs for which a visibility is actually measured and 0 otherwise. In other words, the ‘ideal’ visibilities are masked by the uv -coverage $B(u, v)$,

$$\text{Visibility} \times B(u, v) \xleftrightarrow{\mathcal{FT}} \text{Sky Brightness} \otimes \text{PSF}, \quad (4)$$

where the Fourier Transform of the uv -coverage results in the point spread function (PSF), which can be quite complicated. Hence, the brightness distribution obtained after the Fourier Transform can be understood as a convolution of the true sky brightness distribution and the PSF. In order to restore the true sky brightness distribution we ‘deconvolve’ the initially obtained image. A standard *deconvolution* scheme is the *CLEAN* algorithm. In its simple version⁵ the position of the maximum brightness in the image is successively determined, assumed to represent a point of true sky brightness distribution. The PSF is then subtracted from the residual image at this position. A more precise version⁸ interleaves *minor* cycles of simple CLEAN with *major* cycles in which the new sky brightness

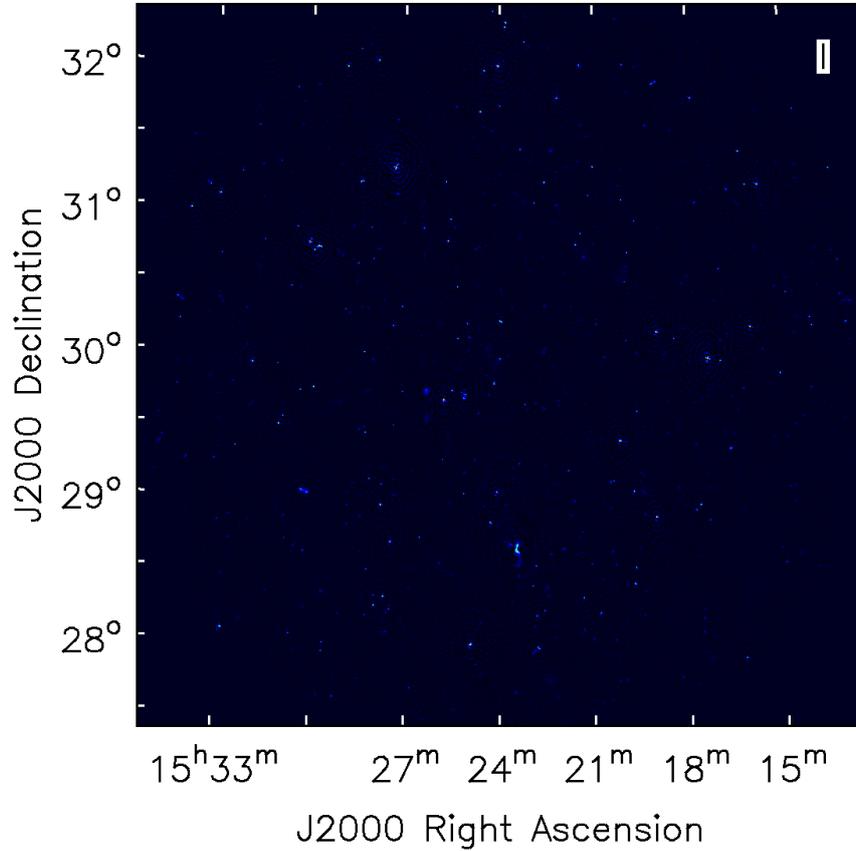


Figure 1. The field of the galaxy cluster Abell 2069 and the Corona Borealis Supercluster observed with LOFAR HBA in May 2014. This map was created by imaging frequency-limited chunks of the whole data and stacking the images afterwards. 100 out of 370 subbands was used. Image resolution is $106''$ *times* $103''$.

components are transformed to Fourier space and subtracted from the visibilities. The new residual image is then generated by transforming the residual visibilities to image space. Hence, the deconvolution is already a complex process in which the image needs to be computed from the visibilities many times.

Since we work with large amounts of data in all imaging steps a Fast Fourier Transform needs to be applied. The field of view of the LOFAR stations is as large as several degrees. On the other hand the resolution, determined by the longest baselines, can be of the order of arcseconds or even below. Even if one is only interested in a comparatively small object the full field of view needs to be imaged in order to be able to deconvolve all sources, which would otherwise generate image artefacts all over the image. The image size must therefore be of the order of several ten thousand pixels in both directions.

The LOFAR stations have a large field of view, which prohibits to ignore the w -term in Eq. 2. Also the beam shape of the stations $A_i(l, m)$ and the polarisation properties of the LOFAR dipoles depend on the pointing direction relative to the station, so they vary

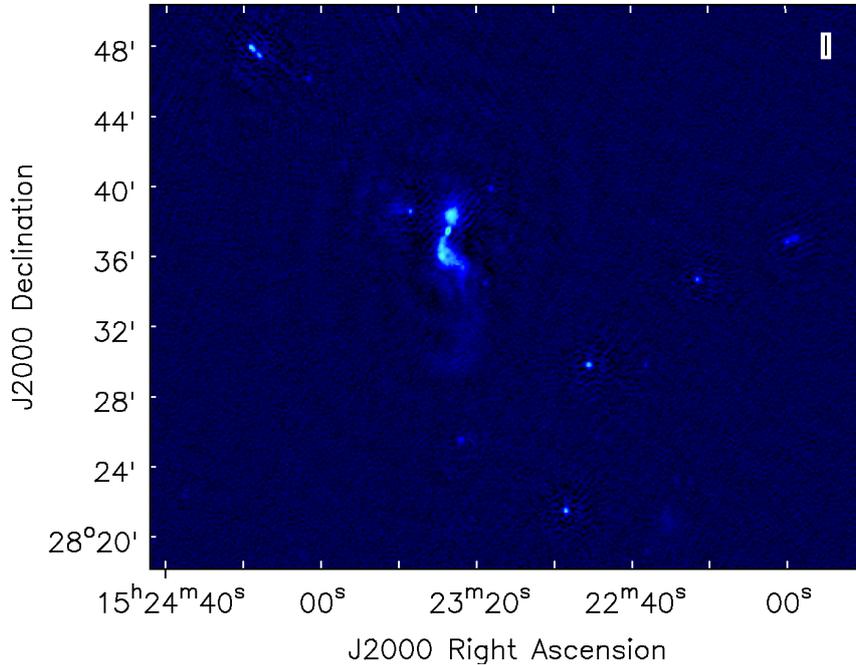


Figure 2. The radio galaxy 4C+28.39 in the field of the galaxy cluster Abell 2069 and the Corona Borealis Supercluster.

with time when tracking a source on the sky. Therefore, the sky brightness for LOFAR cannot be obtained via simply Fourier transforming the visibility data. The visibility data cannot be corrected for direction dependent effects, they have to be corrected in the images, hence a dedicated imager for LOFAR has been developed, the *awimager*⁹. The direction dependent effects are multiplicative in image space but they vary in time and frequency so correcting in image space would require Fourier transforming many time and frequency bins separately. They can be corrected for by a convolution in Fourier space, called the A-projection, and because the effects vary only on large scales in direction this keeps the support of the convolution kernel small, making it more efficient than many Fourier transforms¹. Eq. 2 shows that the w -term can be seen as a multiplicative factor in image space, thus it can also be corrected for by a convolution in Fourier space². Those two corrections together give the aw-projection.

2.4 The Processing Flow: Selfcal

Usually the sky brightness distribution of the target is not fully known. The calibration factors for the instrument can be generated by observing and calibrating on a well known and bright calibrator source, but ionospheric corrections are different for different directions. Fortunately even calibrating on an imperfect model of the sky (e.g. with low resolution) results in an improved image. Therefore several calibration–imaging loops are usually necessary before the sky model converges to a final result. This process is called *selfcal*.

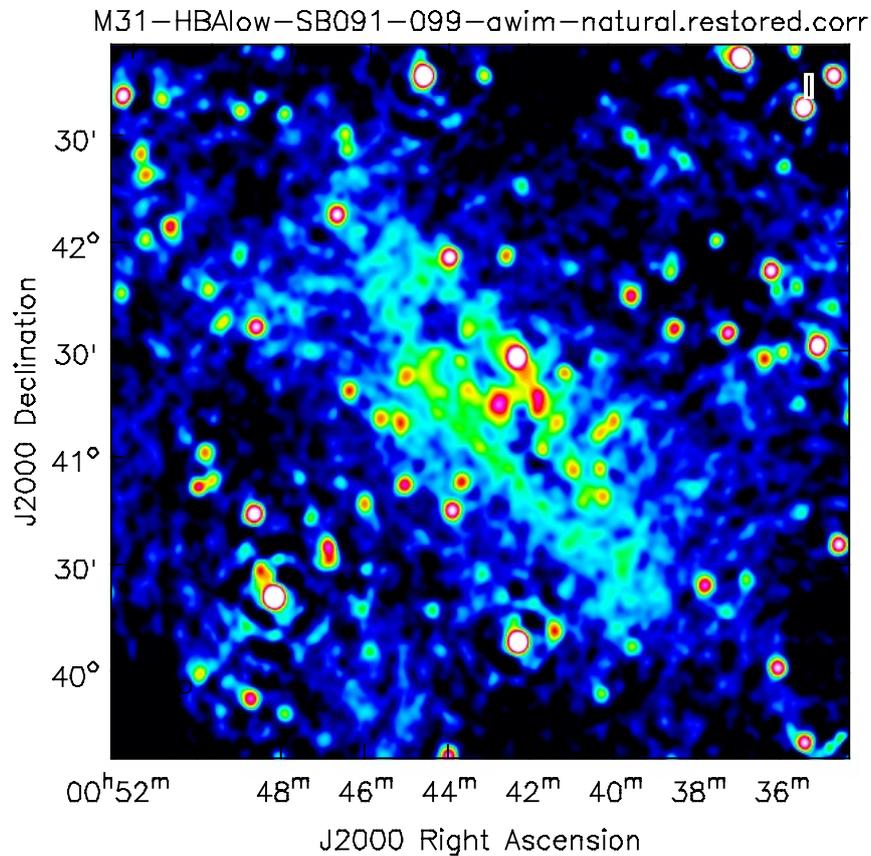


Figure 3. Image of the Andromeda galaxy M31 with LOFAR at 130 MHz. The data for this image was calibrated on JUOPA and imaged on JURECA.

The exact setup of these loops depends on the data and the scientific objective. E.g., at low frequencies the calibration needs to be done on high time resolution in order to track the changes in the ionosphere, while for the imaging the data can be averaged much more before the image is significantly distorted, so it is useful to add an averaging step to speed up processing.

The major steps of creating a sky image out of the measured visibility data are therefore: (i) flagging of RFI, (ii) demixing of bright sources in the vicinity of the target, (iii) averaging the data, (iv) using calibrator sources to calibrate the absolute amplitude of the measured visibilities, (v) image the data and successively determine variations in time, frequency, and across the field of view of the calibration matrices, and (vi) create the final image of the scientific target. For normal users the first three steps are usually done by the observatory right after observation. This minimises the amount of data to be transferred and stored to typically a few ten TByte for each observation. The computing power at JURECA allows us to do those steps on our own, freeing up computing resources at the observatory for more observations and allowing us to perform a more involved processing.

3 Pipeline Framework and Image Examples

In our project we aim for enabling processing of LOFAR data on JURECA. To get good results it is still necessary to adopt the calibration strategy to a specific observation. Many parameters depend on brightness of the sources in the field and the properties of the actual target, e.g., size, morphology and brightness. Figuring out the best sequence of processing steps is still manual work, but once they are known the steps can be added together into a pipeline that runs without human interaction. To run this pipeline a pipeline framework was developed, that executes several consecutive processing steps and takes care of distributing the workload to the processing nodes and generating checkpoints. The original framework is too cumbersome to adopt to different processing strategies and thus was only used for standard processing by the observatory but not by the astronomers. On the other hand simple scripts do not make efficient use of a supercomputer like JURECA. To ease pipeline development we developed an extension to the framework that allows to specify several processing steps in one configuration file in a user-friendly manner. This makes it possible to rapidly specify a processing strategy and have it executed in an efficient manner.

Fig. 3 shows the Andromeda galaxy at 130 MHz, after only direction independent calibration. The direction independent calibration is insufficient, which can be seen in the ring-like artefacts around strong sources. Currently, within the LOFAR community we are developing a strategy which allows to do direction dependent calibration in an automated fashion. This pipeline determines calibration factors in the direction of many bright sources by subtracting all other sources and working on only a small part of the field, the facet calibration. Due to its large size with extended emission of more than 3° adopting this strategy to the observation of the Andromeda galaxy will be a challenge.

Acknowledgements

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC).

References

1. S. Bhatnagar, T. J. Cornwell, K. Golap, and J. M. Uson, *A&A* **487**, 419, 2008.
2. T. J. Cornwell, K. Golap, and S. Bhatnagar, *IEEE J. Select. Top. Signal Proc.* **2**, 647, 2008.
3. M. P. van Haarlem, M. W. Wise, A. W. Gunst, et al., *A&A* **556**, 2, 2013.
4. J. P. Hamaker, J. D. Bregman, and R. J. Sault, *A&AS* **117**, 137, 1996.
5. J. A. Högbom, *A&AS* **15**, 417, 1974.
6. H. T. Intema, arXiv:1402.4889, 2014.
7. A. R. Offringa, J. J. van de Gronde, and J. B. T. M. Roerdink, *A&A* **539**, A95, 2012.
8. F. R. Schwab, *AJ* **89**, 1076, 1984.
9. C. Tasse, S. van der Tol, J. van Zwieten, G. van Diepen, and S. Bhatnagar, *A&A* **553**, 105, 2013.
10. S. van der Tol, B. D. Jeffs, and A.-J. van der Veen, *IEEE Transactions on Signal Processing* **55**, 4497, 2007.

Turbulence and Its Effect on Protostellar Disk Formation

Daniel Seifried¹, Robi Banerjee², and Ralf S. Klessen³

¹ I. Physikalisches Institut, Universität zu Köln, Zùlpicher Str. 77, 50937 Köln, Germany
E-mail: seifried@ph1.uni-koeln.de

² Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany
E-mail: banerjee@hs.uni-hamburg.de

³ Institut für Theoretische Astrophysik, Universität Heidelberg,
Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany
E-mail: klessen@uni-heidelberg.de

We analyse simulations of turbulent, magnetised molecular cloud cores focusing on the formation of Class 0 stage protostellar discs and the physical conditions in their surroundings. We show that for a wide range of initial conditions Keplerian discs are formed in the Class 0 stage already. Furthermore, we show that the accretion of mass and angular momentum in the surroundings of protostellar discs occurs in a highly anisotropic manner, by means of a few narrow accretion channels. The magnetic field structure in the vicinity of the discs is highly disordered, revealing field reversals up to distances of 1000 AU. These findings demonstrate that as soon as even mild turbulent motions are included, the classical disc formation scenario of a coherently rotating environment and a well-ordered magnetic field breaks down.

1 Introduction

In our research we investigate the formation of protostars and their associated protostellar disks, the early precursors of planetary systems. During the last decade simulations of collapsing molecular cloud cores have revealed the so-called catastrophic magnetic braking problem: Magnetic fields are able to transport angular momentum by means of toroidal Alfvén waves. Modelling the collapse of rotating molecular cloud cores, simulations have shown that in the presence of magnetic fields with strengths comparable to observational results, the formation of rotationally supported (Keplerian) protostellar disks is largely suppressed^{1,2}. This is due to the fact that angular momentum is removed very efficiently from the interior of the core by the magnetic field. In previous works we could confirm this effect for the collapse of massive (100 solar masses), molecular cloud cores³. This key result of the suppression of Keplerian disk formation during the earliest stages of star formation is in contrast to recent observational results which state that protostellar disks should be present already in the Class 0 stage^{4,5}.

2 Numerical Methods

The simulations presented here are performed with the hydrodynamics code FLASH, which is written in Fortran 90⁶. The code solves the 3-dimensional, discretised magnetohydrodynamical equations on a Cartesian grid. Making use of the adaptive-mesh-refinement (AMR) technique, only those regions which are of particular interest for us are resolved with the highest possible spatial resolution whereas other regions of minor interest are

resolved more coarsely. This significantly reduces the number of calculations to be performed and hence the computational time required, thus allowing us to perform the simulations over long physical timescales. We also make use of the sink particle routine to model the formation of protostars⁷.

3 Initial Conditions

Observations of the birth places of stars show a wide range of physical quantities, in particular in their initial mass. As we do not simulate a particular region observed by astronomers but rather aim to understand the systematic influence of the initial conditions, we have to perform a number of simulations in our work covering a wide range of masses and turbulence strengths. This allows us to draw conclusions about the effect of the initial conditions on the formation mechanism of stars. We modelled the collapse of molecular cloud cores with masses ranging from about 2 solar masses up to 1000 solar masses. The cores are threaded by a strong magnetic field along the z-axis and have an additional supersonic, turbulent velocity field as indicated by observations.

4 Results

The results of previous simulations described in Sec. 1 show up in case that highly idealised initial conditions are used for the simulations. In particular the lack of turbulent motions - frequently observed in molecular cloud cores - could have a significant effect on the formation of protostellar disks and outflows. For this reason, in our research we here focus on the influence of turbulence on the formation of protostellar disks and outflows. This work has been performed on JUROPA and other supercomputing facilities. Each of the simulations required a computational time of a few 100 000 CPU-hours with a simultaneous use of up to 1000 CPUs per simulation. A few hundreds of files were produced for each simulation requiring a disk space of a few TB in total.

4.1 Turbulence-Induced Disk Formation

We focus our simulations on the question, how turbulence affects the formation of Keplerian disks⁸⁻¹⁰. An example result is shown in Fig. 1 showing the protostellar disk in a representative run with a molecular cloud core of 100 solar masses.

As can be seen, in the presence of turbulence rotationally supported disks are re-obtained again, which is in contrast to the previous simulations with comparable magnetic field strengths but no turbulence³. This suggests that the efficiency of the magnetic braking, which is responsible for removing angular momentum from the midplane, is reduced significantly in the turbulent case. Analysing the surroundings of the disks we can show that this is indeed the case. The turbulent surroundings of the disk exhibit no coherent rotation structure (compare Fig. 1). Therefore, an efficient build-up of a strong toroidal magnetic field responsible for angular momentum extraction is hampered. Moreover, the turbulent motions lead to a strongly disordered magnetic field which further reduces the magnetic braking efficiency. Since simultaneously the angular momentum inwards transport remains high due to the presence of local shear flows in the vicinity of the disks, there is a net inwards angular momentum transport towards the centre of the disk. The combination of

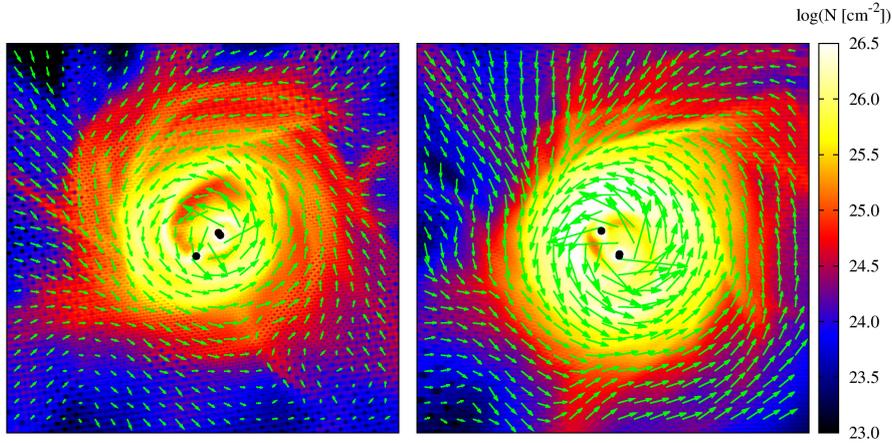


Figure 1. Protostellar disk seen from top-on. Black dots represent protostars, green arrows the velocity field.

these effects results in the observed build-up of Keplerian disks as expected from observations^{4,5}. Varying the core masses (2.6 - 1000 solar masses) and the turbulence strengths does not change our findings. This clearly demonstrates that the turbulence-induced disk formation mechanism works for a wide range of initial conditions. In particular, we could show that the formation of Keplerian disks does not require a uniform rotation of the core – turbulent motions alone lead to the build-up of Keplerian disks. Moreover, we showed that even for subsonic turbulence, which is usually present in low-mass protostellar cores, the turbulence-induced formation mechanism still holds.

In Fig. 2 we show the 3-dimensional structure of the magnetic field and the gas motions around a protostellar disk in one of our simulations. The magnetic field lines clearly reveal a highly complex structure being far off from well-ordered. Several field reversals up to distances of about 1000 AU from the disk centre can be found. This indicates that the recently observed inclination of magnetic fields with respect to the disc axis^{11–13} could simply be the consequence of a spatially unresolved, highly disordered field structure. The anisotropy of the accretion towards the disk is represented by the coloured regions in the figure. Accretion is dominated by a few narrow accretion channels which carry a significant part of the inflowing mass and angular momentum but cover only about 10% of the surface area. We emphasise that we find this anisotropic accretion mode in all of our simulations.

5 Conclusions and Outlook

To summarise, our simulations present a richer picture of the process of disk formation, wherein turbulence, filamentary accretion streams, and magnetic field reversals guarantee that the otherwise overwhelming strength of magnetic braking by ordered fields is significantly degraded, allowing Keplerian discs to form. Non-turbulent collapse simulations, on the other hand, might significantly overestimate the efficiency of magnetic braking and thus underestimate the fraction of Class 0 stage Keplerian discs. We suggest that the anisotropic accretion and disordered magnetic field structure found in the environment of protostellar

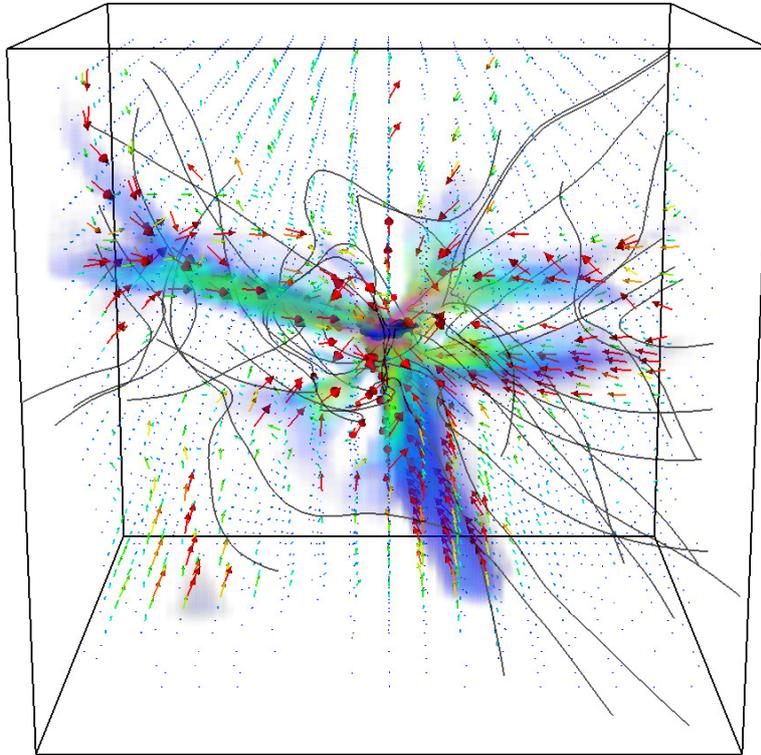


Figure 2. 3-dimensional structure of the magnetic field and gas motions around a Keplerian disk in one of our simulations.

discs might set the stage for other mechanisms contributing to the formation of discs like (pseudo-) disc warping or non-ideal MHD effects^{14,15}

For our future research we intend to study the self-consistent ejection of protostellar outflows from the Keplerian disks formed in our turbulence simulations. In order to reach this goal we will have to redo some of our simulations with increased spatial resolution, which will require further large amounts of computing power in the future. Furthermore, we plan to produce so-called synthetic observations. These synthetic observations will allow us to directly compare our simulation results with real observations. Such comparisons allow us to assess what can be inferred from observations - in particular how reliable parameters obtained from real observations are.

Acknowledgements

DS acknowledges funding by the Bonn-Cologne Graduate School as well as from Deutsche Forschungsgemeinschaft (DFG) via the Sonderforschungsbereich SFB 956. RB acknowl-

edges funding by the DFG via the Emmy-Noether grant BA 3706/1-1, the ISM-SPP 1573 grants BA 3706/3-1 and BA 3706/3-2, as well as for the grant BA 3706/4-1. RSK acknowledges support from the European Research Council under the European Communitys Seventh Framework Programme (FP7/2007-2013) via the ERC Advanced Grant STARLIGHT (project number 339177) and support from the DFG via the Sonderforschungsbereich SFB 881. The FLASH code was developed partly by the DOE-supported Alliances Center for Astrophysical Thermonuclear Flashes (ASC) at the University of Chicago.

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA at Jülich Supercomputing Centre (JSC).

References

1. A. Allen, Z.-Y. Li, and F. H. Shu, *Collapse of Magnetized Singular Isothermal Toroids. II. Rotation and Magnetic Braking*, *ApJ*, **599**, 363–379, Dec. 2003.
2. P. Hennebelle and S. Fromang, *Magnetic processes in a collapsing dense core. I. Accretion and ejection*, *A&A*, **477**, 9–24, Jan. 2008.
3. D. Seifried, R. Banerjee, R. S. Klessen, D. Duffin, and R. E. Pudritz, *Magnetic fields during the early stages of massive star formation - I. Accretion and disc evolution*, *MNRAS*, **417**, 1054–1073, Oct. 2011.
4. J. J. Tobin, L. Hartmann, H.-F. Chiang, D. J. Wilner, L. W. Looney, L. Loinard, N. Calvet, and P. D’Alessio, *A ~ 0.2 -solar-mass protostar with a Keplerian disk in the very young L1527 IRS system*, *Nature*, **492**, 83–85, Dec. 2012.
5. Á. Sánchez-Monge, R. Cesaroni, M. T. Beltrán, M. S. N. Kumar, T. Stanke, H. Zinnecker, S. Etoaka, D. Galli, C. A. Hummel, L. Moscadelli, T. Preibisch, T. Ratzka, F. F. S. van der Tak, S. Vig, C. M. Walmsley, and K.-S. Wang, *A candidate circumbinary Keplerian disk in G35.20-0.74 N: A study with ALMA*, *A&A*, **552**, L10, Apr. 2013.
6. B. Fryxell, K. Olson, P. Ricker, F. X. Timmes, M. Zingale, D. Q. Lamb, P. MacNeice, R. Rosner, J. W. Truran, and H. Tufo, *FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes*, *ApJS*, **131**, 273–334, Nov. 2000.
7. C. Federrath, R. Banerjee, P. C. Clark, and R. S. Klessen, *Modeling Collapse and Accretion in Turbulent Gas Clouds: Implementation and Comparison of Sink Particles in AMR and SPH*, *ApJ*, **713**, 269–290, Apr. 2010.
8. D. Seifried, R. Banerjee, R. E. Pudritz, and R. S. Klessen, *Disc formation in turbulent massive cores: circumventing the magnetic braking catastrophe*, *MNRAS*, **423**, L40–L44, June 2012.
9. D. Seifried, R. Banerjee, R. E. Pudritz, and R. S. Klessen, *Turbulence-induced disc formation in strongly magnetized cloud cores*, *MNRAS*, **432**, 3320–3331, July 2013.
10. D. Seifried, R. Banerjee, R. E. Pudritz, and R. S. Klessen, *Accretion and magnetic field morphology around Class 0 stage protostellar discs*, *MNRAS*, **446**, 2776–2788, Jan. 2015.
11. N. L. Chapman, J. A. Davidson, P. F. Goldsmith, M. Houde, W. Kwon, Z.-Y. Li, L. W. Looney, B. Matthews, T. G. Matthews, G. Novak, R. Peng, J. E. Vaillancourt, and N. H. Volgenau, *Alignment between Flattened Protostellar Infall Envelopes and Ambient Magnetic Fields*, *ApJ*, **770**, 151, June 2013.

12. C. L. H. Hull, R. L. Plambeck, A. D. Bolatto, G. C. Bower, J. M. Carpenter, R. M. Crutcher, J. D. Fiege, E. Franzmann, N. S. Hakobian, C. Heiles, M. Houde, A. M. Hughes, K. Jameson, W. Kwon, J. W. Lamb, L. W. Looney, B. C. Matthews, L. Mundy, T. Pillai, M. W. Pound, I. W. Stephens, J. J. Tobin, J. E. Vaillancourt, N. H. Volgenau, and M. C. H. Wright, *Misalignment of Magnetic Fields and Outflows in Protostellar Cores*, *ApJ*, **768**, 159, May 2013.
13. C. L. H. Hull, R. L. Plambeck, W. Kwon, G. C. Bower, J. M. Carpenter, R. M. Crutcher, J. D. Fiege, E. Franzmann, N. S. Hakobian, C. Heiles, M. Houde, A. M. Hughes, J. W. Lamb, L. W. Looney, D. P. Marrone, B. C. Matthews, T. Pillai, M. W. Pound, N. Rahman, G. Sandell, I. W. Stephens, J. J. Tobin, J. E. Vaillancourt, N. H. Volgenau, and M. C. H. Wright, *TADPOL: A 1.3mm Survey of Dust Polarization in Star-forming Cores and Regions*, *ApJS*, **213**, 13, July 2014.
14. Z.-Y. Li, R. Krasnopolsky, H. Shang, and B. Zhao, *On the Role of Pseudodisk Warping and Reconnection in Protostellar Disk Formation in Turbulent Magnetized Cores*, ArXiv e-prints, Aug. 2014.
15. M. N. Machida, S.-I. Inutsuka, and T. Matsumoto, *Conditions for circumstellar disc formation: effects of initial cloud configuration and sink treatment*, *MNRAS*, **438**, 2278–2306, Mar. 2014.

Computational Biology and Biophysics

Computational Biology and Biophysics: Modelling Molecular Structure with High Performance Computing Approaches

Wolfgang Wenzel

Institute of Nanotechnology, Karlsruhe Institute of Technology, Karlsruhe, Germany
E-mail: wolfgang.wenzel@kit.edu

The continuous development of methods by a large community of scientists over several decades, coupled with the steady advance of the available computational resources, has led to an increasing role of simulation methods in life science¹ and materials science research. In part these achievements were recognised with the 2013 Nobel Prize for chemistry², which honoured the development of multi-scale simulation methods that have been used for decades to elucidate mechanisms of bio-macromolecules. With time, simulation methods have reached an increasing degree of maturity^{3,4}, starting with simulations that were closely tied to the availability of structural data⁵ to eventually fill ever larger gaps, which experimental methods could not yet resolve⁶. Molecular simulation methods have been used to fill the static structures obtained from xray crystallography or NMR techniques with life, generating representative trajectories which visualise, but more importantly enable analysis of the dynamics of the macromolecular machinery of the cell.

It has been a paradigm of molecular life-science research that structure determines function. For this reason the determination of biomolecular structure, often as a function of the environment or the presence of interaction partners, has been a hallmark also of molecular simulations⁷. We now recognise, however, that many biological functions are accomplished by intrinsically disordered proteins, which either exist in an ensemble of many conformations or even attain their conformation only when binding to the respective partners⁸. An example of such a study is the work by Abad and Rossetti in this volume, who discussed methods that enable characterisation of the conformational space of intrinsically disordered proteins. Simulations of these systems are complicated by the large conformational space that these proteins can attain, but offer important insights into structural ensembles that are difficult to characterise with ensemble-level experimental techniques which require either a single or a few dominant structures for successful characterisation.

Another hallmark application of molecular simulation has been the study of receptor-ligand interactions, in particular for pharmaceutically relevant protein receptors⁹. In this field early work addressed the complementarity of specific ligands to the protein receptors¹⁰, which often required suppression of the receptor conformational degrees of freedom for lack of adequate computational resources and simulation techniques. The work by Frieg *et al.* in this volume is another example for the degree of maturity this field of simulation has reached today. Routine sequencing of the human genome, as for instance demonstrated in the thousand genomes project, has led to an unprecedented amount of information of mutations that affect critical metabolic pathways of the cell. As demonstrated in the work reported by Frieg *et al.* simulation methods have now matured to enable explo-

ration of possible treatment of the ensuing disorders through the development of ligands that restore the activity of the damaged enzyme. Further development of such techniques, in combination with modelling of the metabolic pathways and systems biology methods, raises hopes for the development of tools for personalised medicine.

However, applications of molecular simulations are not limited to the life-sciences. The advent of novel two-dimensional materials, such as graphene¹¹, holds great promise for the development of novel materials due to their unique structural and electronic properties. The high demand of such materials, in particular for large-scale industrial application, challenges scientists to develop scalable production techniques of high-quality samples. One of the most promising methods is the development of exfoliation techniques¹², where small organic molecules are used to stabilise flakes of these 2-D materials in solution. The design of appropriate stabilisers, such as the perylene molecules discussed in the study of Hollfelder and Gekle in this volume, requires striking a complex balance between the solubility of the molecules and their affinity to the target substrate, as aggregation competes with exfoliation. The work reported here demonstrates how molecular simulation can elucidate the interaction of complex molecules with their solvent environment in order to enable selection or molecular design of optimal solvents.

References

1. M. Karplus, J. A. McCammon, *Molecular dynamics simulations of biomolecules*, Nat. Struct. Mol. Biol. **9**, 646-652, 2002.
2. M. Karplus, *Nobel Lecture: Development of Multiscale Models for Complex Chemical Systems From H+H2 to Biomolecules*, 2013.
3. D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, W. Wriggers, *Atomic-level characterization of the structural dynamics of proteins*, Science **330**, 341-6, 2010.
4. T. H. Walther, C. Gottselig, S. L. Grage, M. Wolf, A. V. Vargiu, M. J. Klein, S. Vollmer, S. Prock, M. Hartmann, S. Afonin, E. Stockwald, H. Heinzmann, O. V. Nolandt, W. Wenzel, P. Ruggerone, A. S. Ulrich, *Folding and Self-Assembly of the TatA Translocation Pore Based on a Charge Zipper Mechanism*, Cell **152**, 316-326, 2013.
5. A. Miranker, S. E. Radford, M. Karplus, C. M. Dobson, *Demonstration by NMR of folding domains in lysozyme*, Nature **349**, 633-636, 1991.
6. L. V. Bock, C. Blau, G. F. Schröder, I. I. Davydov, N. Fischer, H. Stark, M. V. Rodnina, A. C. Vaiana, H. Grubmüller, *Energy barriers and driving forces in tRNA translocation through the ribosome*, Nat. Struct. Mol. Biol. **20**, 1390-1396, 2013.
7. G. Zhao, J. R. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, P. Zhang, *Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics*, Nature **497**, 643-646, 2013.
8. V. N. Uversky, *Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes*, Chemical Society Reviews **40**, 1623-1634, 2011.
9. D. Kokh, R. Wade, W. Wenzel, *Receptor flexibility in small-molecule docking calculations*, Reviews of Computational Molecular Science **1**, 298-314, 2010.

10. E. C. Meng, K. Shoichet, I. D. Kunz, *Automated Docking with Grid-Based Energy Evaluation*, J. Comp. Chem. **13**, 505, 1992.
11. K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, A. A. Firsov, *Two-dimensional gas of massless Dirac fermions in graphene*, Nature **438**, 197-200, 2005.
12. Y. Hernandez, V. Nicolosi, M. Lotya, F. M. Blighe, Z. Sun, S. De, I. T. McGovern, B. Holland, M. Byrne, Y. K. Gun'Ko, J. J. Boland, P. Niraj, G. Duesberg, S. Krishnamurthy, R. Goodhue, J. Hutchison, V. Scardaci, A. C. Ferrari, J. N. Coleman, *High-yield production of graphene by liquid-phase exfoliation of graphite*, Nat. Nano **3**, 563-568, 2008.

Structural Predictions of Intrinsically Disordered Proteins with Computational Methods

Enrique Abad¹ and Giulia Rossetti^{1,2}

¹ Computational Biomedicine, Institute for Advanced Simulation IAS-5 and Institute of Neuroscience and Medicine INM-9, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: e.abad@grs-sim.de

² Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: g.rossetti@fz-juelich.de

Structural characterisation of intrinsically disordered proteins is highly non-trivial. They exist as dynamic, highly flexible structural ensembles that undergo conformational conversions on a wide range of timescales, spanning from picoseconds to milliseconds. Computational methods may be of great help to characterise these proteins. Here we review recent progress from our lab and other groups to develop and apply *in silico* methods for structural predictions of these highly relevant, challenging systems.

1 Introduction

In the following, we closely follow the more detailed original report in Rossetti, *et al.*¹ (Copyright Elsevier, 2015). Intrinsically Disordered proteins (IDPs) are an important class of functional proteins with high abundance in nature²⁻⁴, specifically in humans, where they represent almost one third of the genome^{5,2-4}. Notably IDPs are extensively associated with human diseases and amyloidosis^{6,7}. Specifically, 79 % of cancer associated proteins, 57 % of the cardiovascular disease associated proteins and 55 % of neurodegenerative disease associated proteins are predicted to contain 30 or more consecutive disordered residues⁸.

Studying the structural determinants of this class of proteins is the key to understand their role for cellular function and dysfunction in both healthy and altered-disease-associated pathways. Unfortunately, traditional computational and experimental approaches have been hampered so far by a variety of challenges. IDPs do not adopt a well-defined native three-dimensional structure⁹ and they lack stable tertiary and/or secondary structures when isolated in solution under near-physiological conditions¹⁰ and exist in an ensemble of states both in solution and when unbound to a ligand *in vivo*¹⁰. This means that an ensemble of inter-converting conformers is required to describe the conformational behaviour of IDPs¹¹.

Apparently IDPs do not simply occur as filler material amongst functional well-structured proteins, instead they are associated with a variety of biological functions². IDPs are indeed enriched in signalling and regulatory functions because disordered segments permit interaction with several proteins and hence the re-use of the same protein in multiple pathways^{12,2,13,14}. Moreover, IDPs can apply different molecular recognition mechanisms and functional modes^{15,14} compared to the globular proteins. The majority of IDPs undergo a disorder-to-order transition upon functioning^{16,17}, a structural transition from a partially disordered state into a more highly ordered conformation in the complex^{17,18}, also called folding-upon-binding mechanism (Fig. 1A).

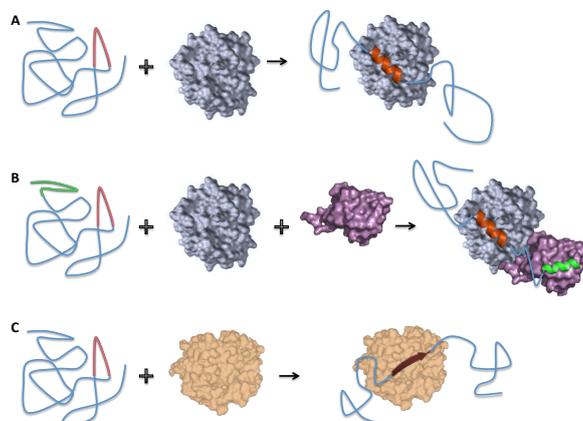


Figure 1. Schematic of IDPs molecular recognition mechanisms. A) disorder-to-order transition¹⁴. B) Binding commonality¹⁹. C) Binding diversity¹⁴. Figure from our original report in Rossetti, *et al.*¹ (Copyright Elsevier, 2015).

The persistence of natively unfolded proteins throughout evolution may reside in advantages of flexible structure during disorder order transitions in comparison with rigid proteins^{10,20,21}. The potential advantages of intrinsic lack of structure and function-related disorder-to-order transitions are:

- i) Decoupling of specificity from binding strength. IDPs are capable to combine high specificity with low affinity¹⁰. This is due to the fact that folding and binding are coupled for IDPs. Therefore the change in enthalpy is compensated by a much larger loss of conformational entropy as compared to globular proteins. This results in a lower absolute value of free-energy, decreasing the stability of the resulting complex^{22,3}.
- ii) Binding commonality in which multiple, distinct sequences fold differently yet each recognises a common binding surface¹⁹ (Fig. 1B). These localised interacting regions allow IDPs to have an increased modularity as different binding regions can be incorporated into the same protein without excessively increasing protein length²³.
- iii) Binding diversity. IDPs may folds differently to recognise differently shaped partners by several structural accommodations at the various binding interfaces^{10,17,2,21} (Fig. 1C). This phenomenon known as one-to-many signalling, illustrate the complexity of the different binding modes of IDPs and enables an exceptional plasticity in cellular responses¹⁴.
- iv) The creation of very large interaction surfaces as the disordered protein wraps-up or surrounds its partner²⁴ making it possible to overcome steric restrictions¹⁰, meaning that these proteins utilise a much larger fraction of their accessible surfaces compared to globular proteins²⁵.
- v) A increased speed of interaction. IDPs display both faster rates of association by reducing dependence on orientation factors and by enlarging target sizes and faster

rates of dissociation¹⁰. The great conformational freedom of IDPs in multidirectional search permits the recognition of distant and/or discontinuous determinants on the target^{6,15}. Moreover, their extended structure enables them to contact their partner(s) over a large binding surface area, which allows the same interaction potential to be realised by shorter proteins overall¹⁵.

All these features collocate IDPs among the major cellular regulators, recognisers, and signal transducers¹¹. Also IDPs reduced life-time in the cell, possibly represents a mechanism of rapid turnover of important regulatory molecules²¹.

The inherent flexibility of IDPs calls upon new experimental and computational strategies for studying these proteins, since describing the ensemble of conformations of IDP at atomistic level remains a considerable challenge.

In this review we will first discuss the computational strategies that have been devised to tackle the conformational plasticity of IDPs, complemented by an application from our lab.

2 Computational Methods for IDPs

Computational methods using physics-based empirical molecular mechanics force fields increasingly release critical contributions in providing general insights into the behaviour of IDPs^{26,11,22}. However, the dynamic and heterogeneous nature of IDPs presents substantial challenges, in terms of force field accuracy and of conformational sampling capability. MD simulations are indeed sensitive to the choice of the protein force field, which are typically parameterised to reproduce the behaviours of folded proteins rather than IDPs, and thus they may fail to capture important aspects of IDP conformational ensembles²⁷. We will therefore describe all the different techniques so far used for IDPs.

2.1 Molecular Dynamics and Monte Carlo Simulations

Molecular dynamics (MD) and Monte Carlo (MC) simulations complement experiments by elucidating chemical details underlying the conformational dynamics of biological macromolecules²⁸. Unfortunately, it is extremely difficult to adequately sample the conformational space accessible to IDPs. In details:

- MD in explicit solvent at room temperature is generally insufficient for achieving convergence in simulated structural ensembles of IDPs, due to their large conformational space and the so-called kinetic trapping, i.e., the system tends to be confined to local energy minima²⁹. Such minima are separated by free-energy barriers, whose heights are often much larger than the thermal energy available to the system²⁹. Therefore MD is not always suitable to sample the dynamical behaviour of IDPs.
- In MC approaches, stochastic conformational searches are used to efficiently sample conformations of the protein chain³⁰. MC surmounts energy barriers by moving through successive discrete local minima in the energy landscape. In this way, MC sample all the minima of conformational without seeing the energy barriers³¹.

2.2 Enhanced Sampling Techniques

These methods achieve a random walk in the potential-energy space, allowing the system to easily overcome the energy barriers that separate local minima. Three well-known approaches for carrying out generalised ensemble MD or MC simulations are the multi-canonical algorithm³², the simulated tempering³³ and the replica exchange method³⁴. They are very briefly summarised here.

- The multi-canonical algorithm (MUCA) method³² assigns to each state with energy E a non-Boltzmann weight that is independent of temperature so that a uniform potential energy distribution is obtained ensuring that all the energy states are sampled with the same likelihood. This approach was applied to the coupled folding and binding of an IDP in order to generate the corresponding free-energy landscape³⁵.
- The simulated tempering (ST) performs a free random walk in temperature space. This random walk, in turn, induces a random walk in the potential energy space and allows the simulation to escape from states of local energy minima. In ST³³, the temperature of the system is randomly switched between several predefined values. ST was applied to study the binding mechanism of two IDPs in combination with classical MD³⁶.
- The replica-exchange method (REM) uses standard Boltzmann weight factors that are known *a priori*³⁴. In this method, a number of non-interacting copies (or replicas) of the original system at different temperatures are simulated in parallel under different conditions³⁴; at given time intervals, the simulation conditions are exchanged with a specific transition probability between replica pairs³⁴. A variation is the replica exchange solute tempering method, REST2³⁷, in which only the protein and the ions (i.e. the solute) are simulated at different effective temperatures by applying an appropriate potential energy function to each replica. REM could also be coupled with Monte-Carlo simulations, (REMC)^{38,39} to explore the conformation space of IDPs (see a recent application from our lab below).

New generalised-ensemble algorithms could be obtained by combining the merits of the above three methods (reviewed in Ref. 40). Another particularly attractive approach to overcome the sampling bottleneck is to combine large numbers of equilibrium and/or generalised ensemble simulations using network methods based on MC algorithms like Markov State Models recently applied also to IDPs^{41,42}.

2.3 Solvent Representation

In all the methods discussed in the previous sections, the solvent can be represented either as a continuum model, or as explicit molecules. Traditional explicit solvent protein force fields arguably provide the most realistic description of solvent, but also significantly increase the system size leading to prohibitive computational cost to sufficiently sample the immense conformational space of IDPs⁴³. Moreover, explicit solvent force fields are known to have a tendency to over-stabilise helices⁴⁴ and overestimate the strength of protein-protein interactions. A substantial reduction in the computational cost could be obtained using implicit solvent models⁴⁵. Recently important advances have been made to greatly

improve the efficiency and achievable accuracy of implicit solvent models based on generalised Born (GB) approximation, however they do not properly describe short-range effects where the detailed interplay of a few non-bulk-like water molecules is important and might be further limited by the specific methodology for calculating the solvation free energy as well as the physical parameters of the solvation model⁴⁶. Despite these caveats, implicit solvent force field has been successfully applied to simulation of regulatory IDPs^{47,48}.

3 Applications from our Lab

A REMC-derived approach based on an implicit-solvent all-atom potential⁴⁹ was recently applied on the disordered N-terminal domain of Prion Protein (PrP)³⁸. We were able to predict the conformational ensemble of the wild type (WT) and mutated mouse PrPC N-terminal domain. Importantly, the work shows how pathogenic mutations (PMs) affect the PrPC binding to functional interactors and/or the translocation³⁸. In Dibenedetto *et al.*⁵⁰, we proposed a computational protocol based on classical MD simulations for investigating how the conformational space of the IDP alpha-synuclein (AS) is affected by the binding of an anti-aggregation drug, dopamine (DOP). Specifically, we analysed the conformational ensemble of AS, alone and in the presence of the drug, with a newly developed tool based on the dihedral angle distributions visited during MD⁵⁰. The latter allows interpreting 2D 1H-15N Heteronuclear multiple-quantum correlation (HMQC) spectra of AS in the presence of the anti-aggregation drug by distinguishing variation of chemical shifts due to direct contacts with the drug from the ones due to conformational changes of the AS induced by long-range effects of the binding. Very recently [under review], we have extended this protocol for the study of the physiological, N-terminally acetylated form of alpha-synuclein (AcAS). We have employed the REST2 method exposed in Sec. 2.2, with the same force field as in the work of Dibenedetto *et al.*⁵⁰. Realistic starting geometries from this previous work underwent 15 ns of REST2 simulations with 32 replicas after manually adding the acetyl group to the N-terminus. The obtained ensemble compares well with experimental measurements of local properties (NMR-chemical shifts) as well as global properties (hydrodynamic radius, circular dichroism spectra).

4 Conclusions

We have presented recent computational investigations regarding proteins of biological relevance, whose structure determination poses challenges to experimental techniques. Our studies of IDPs suggest, as already pointed out, that methods at different resolution might give important insights in their biological function as well as ligand binding, a process which is so far not well understood.

Acknowledgements

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and the JARA-HPC Vergabegremium (VGG) and provided on the supercomputers JUROPA and JURECA at Jülich Supercomputing Centre (JSC).

References

1. G. Rossetti, D. Dibenedetto, V. Calandrini, A. Giorgetti, and P. Carloni, *Structural predictions of neurobiologically relevant G-protein coupled receptors and intrinsically disordered proteins*, Archives of Biochemistry and Biophysics, **582**, 91–100, 2015.
2. H. J. Dyson and P. E. Wright, *Intrinsically unstructured proteins and their functions*, Nat Rev Mol Cell Biol, **6**, no. 3, 197–208, 2005.
3. P. Tompa, *Intrinsically unstructured proteins*, Trends Biochem Sci, **27**, no. 10, 527–33, 2002.
4. V. N. Uversky, *Natively unfolded proteins: a point where biology waits for physics*, Protein Sci, **11**, no. 4, 739–56, 2002.
5. A. K. Dunker, Z. Obradovic, P. Romero, E. C. Garner, and C. J. Brown, *Intrinsic protein disorder in complete genomes*, Genome informatics International Conference on Genome Informatics, **11**, 161–71, 2000.
6. L. M. Iakoucheva, C. J. Brown, J. D. Lawson, Z. Obradović, and A. K. Dunker, *Intrinsic disorder in cell-signaling and cancer-associated proteins*, Journal of Molecular Biology, **323**, no. 3, 573–84, 2002.
7. V. N. Uversky, C. J. Oldfield, and A. K. Dunker, *Intrinsically Disordered Proteins in Human Diseases: Introducing the D²Concept*, Annual Review of Biophysics, **37**, no. 1, 215–246, 2008.
8. V. N. Uversky, C. J. Oldfield, and A. K. Dunker, *Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling*, Journal of Molecular Recognition, **18**, no. 5, 343–384, 2005.
9. V. N. Uversky and D. Eliezer, *Biophysics of Parkinsons Disease: Structure and Aggregation of-Synuclein*, Current Protein and Peptide Science, **10**, no. 5, 483–499, 2009.
10. A. K. Dunker, J. D. Lawson, C. J. Brown, R. M. Williams, P. Romero, J. S. Oh, C. J. Oldfield, A. M. Campen, C. M. Ratliff, K. W. Hipps, J. Ausio, M. S. Nissen, R. Reeves, C. Kang, C. R. Kissinger, R. W. Bailey, M. D. Griswold, W. Chiu, E. C. Garner, and Z. Obradovic, *Intrinsically disordered protein*, Journal of Molecular Graphics and Modelling, **19**, no. 1, 26–59, 2001.
11. A. K. Dunker, I. Silman, V. N. Uversky, and J. L. Sussman, *Function and structure of inherently disordered proteins*, Current opinion in structural biology, **18**, no. 6, 756–764, 2008.
12. A. K. Dunker, E. Garner, S. Guillot, P. Romero, K. Albrecht, J. Hart, Z. Obradovic, C. Kissinger, and J. E. Villafranca, *Protein disorder and the evolution of molecular recognition: theory, predictions and observations*, Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, pp. 473–484, 1998.
13. R. G. Smock and L. M. Gierasch, *Sending Signals Dynamically*, Science, **324**, no. 5924, 198–203, 2009.
14. V. N. Uversky, *Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes*, Chemical Society Reviews, **40**, no. 3, 1623–1634, 2011.
15. P. Tompa, *The interplay between structure and function in intrinsically unstructured proteins*, FEBS Letters, **579**, no. 15, 3346–3354, 2005.

16. A. P. Demchenko, *Recognition between flexible protein molecules: induced and assisted folding*, Journal of molecular recognition, **14**, no. 1, 42–61, 2001.
17. H. J. Dyson and P. E. Wright, *Coupling of folding and binding for unstructured proteins*, Current opinion in structural biology, **12**, no. 1, 54–60, 2002.
18. P. E. Wright and H. J. Dyson, *Linking folding and binding*, Current opinion in structural biology, **19**, no. 1, 31–38, 2009.
19. P. Radivojac, L. M. Iakoucheva, C. J. Oldfield, Z. Obradovic, V. N. Uversky, and A. K. Dunker, *Intrinsic disorder and functional proteomics*, Biophysical journal, **92**, no. 5, 1439–1456, 2007.
20. B. W. Pontius, *Close encounters: why unstructured, polymeric domains can increase rates of specific macromolecular association*, Trends Biochem Sci, **18**, no. 5, 181–6, 1993.
21. P. E. Wright and H. J. Dyson, *Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm*, Journal of Molecular Biology, **293**, no. 2, 321–31, 1999.
22. D. Eliezer, *Biophysical characterization of intrinsically disordered proteins*, Current opinion in structural biology, **19**, no. 1, 23–30, 2009.
23. K. Gunasekaran, C.-J. Tsai, S. Kumar, D. Zanuy, and R. Nussinov, *Extended disordered proteins: targeting function with less scaffold*, Trends in biochemical sciences, **28**, no. 2, 81–85, 2003.
24. Y. Choo and J. W. Schwabe, “All wrapped up”, 00/01 1998, 4.
25. B. Mszros, P. Tompa, I. Simon, and Z. Dosztanyi, *Molecular principles of the interactions of disordered proteins*, Journal of Molecular Biology, **372**, no. 2, 549–561, 2007.
26. J. M. Bourhis, B. Canard, and S. Longhi, *Predicting protein disorder and induced folding: from theoretical principles to practical applications*, Current Protein and Peptide Science, **8**, no. 2, 135–149, 2007.
27. S. Piana, K. Lindorff-Larsen, and D. E. Shaw, *How robust are protein folding simulations with respect to force field parameterization?*, Biophysical journal, **100**, no. 9, L47–L49, 2011.
28. D. Frenkel, B. Smit, and M. A. Ratner, *Understanding molecular simulation: from algorithms to applications*, Physics Today, **50**, 66, 1997.
29. H. A. Scheraga, M. Khalili, and A. Liwo, *Protein-folding dynamics: overview of molecular simulation techniques*, Annu. Rev. Phys. Chem., **58**, 57–83, 2007.
30. A. Vitalis and R. V. Pappu, *Methods for Monte Carlo simulations of biomacromolecules*, Annual reports in computational chemistry, **5**, 49–76, 2009.
31. Z. Li and H. A. Scheraga, *Monte Carlo-minimization approach to the multiple-minima problem in protein folding*, Proceedings of the National Academy of Sciences, **84**, no. 19, 6611–6615, 1987.
32. N. Nakajima, H. Nakamura, and A. Kidera, *Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides*, The Journal of Physical Chemistry B, **101**, no. 5, 817–824, 1997.
33. E. Marinari and G. Parisi, *Simulated tempering: a new Monte Carlo scheme*, EPL (Europhysics Letters), **19**, no. 6, 451, 1992.
34. Y. Sugita and Y. Okamoto, *Replica-exchange molecular dynamics method for protein folding*, Chemical Physics Letters, **314**, no. 1, 141–151, 1999.

35. J. Higo, Y. Nishimura, and H. Nakamura, *A free-energy landscape for coupled folding and binding of an intrinsically disordered protein in explicit solvent from detailed all-atom computations*, *Journal of the American Chemical Society*, **133**, no. 27, 10448–10458, 2011.
36. I. Staneva, Y. Huang, Z. Liu, and S. Wallin, *Binding of two intrinsically disordered peptides to a multi-specific protein: a combined Monte Carlo and molecular dynamics study*, *PLoS computational biology*, **8**, no. 9, e1002682, 2012.
37. L. Wang, R. A. Friesner, and B. J. Berne, *Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2)*, *J. Phys. Chem. B*, **115**, no. 30, 9431–9438, 2011.
38. X. Cong, N. Casiraghi, G. Rossetti, S. Mohanty, G. Giachin, G. Legname, and P. Carloni, *Role of Prion Disease-Linked Mutations in the Intrinsically Disordered N-Terminal Domain of the Prion Protein*, *Journal of Chemical Theory and Computation*, **9**, no. 11, 5158–5167, 2013.
39. R. K. Das and R. V. Pappu, *Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues*, *Proceedings of the National Academy of Sciences*, 2013.
40. Y. Okamoto, *Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations*, *Journal of Molecular Graphics and Modelling*, **22**, no. 5, 425–439, 2004.
41. C. M. Baker and R. B. Best, *Insights into the binding of intrinsically disordered proteins from molecular dynamics simulation*, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2013.
42. Q. Qiao, G. R. Bowman, and X. Huang, *Dynamics of an intrinsically disordered protein reveal metastable conformations that potentially seed aggregation*, *Journal of the American Chemical Society*, **135**, no. 43, 16092–16101, 2013.
43. T. H. Click, D. Ganguly, and J. Chen, *Intrinsically Disordered Proteins in a Physics-Based World*, *Int J Mol Sci*, **11**, no. 12, 5292–5309, 2010.
44. R. B. Best, N.-V. Buchete, and G. Hummer, *Are current molecular dynamics force fields too helical?*, *Biophysical journal*, **95**, no. 1, L07–L09, 2008.
45. M. Feig and C. L. Brooks Iii, *Recent advances in the development and application of implicit solvent models in biomolecule simulations*, *Current opinion in structural biology*, **14**, no. 2, 217–224, 2004.
46. J. Chen and C. L. Brooks, *Critical importance of length-scale dependence in implicit modeling of hydrophobic interactions*, *Journal of the American Chemical Society*, **129**, no. 9, 2444–2445, 2007.
47. J. Chen, *Intrinsically disordered p53 extreme C-terminus binds to S100B ($\beta\beta$) through “fly-casting”*, *Journal of the American Chemical Society*, **131**, no. 6, 2088–2089, 2009.
48. D. Ganguly and J. Chen, *Atomistic details of the disordered states of KID and pKID. Implications in coupled binding and folding*, *Journal of the American Chemical Society*, **131**, no. 14, 5214–5223, 2009.
49. *Folding thermodynamics of peptides*, *Biophys J*, **88**, no. 3, 1560–9, 2005.
50. D. Dibenedetto, G. Rossetti, R. Caliandro, and P. Carloni, *A molecular dynamics simulations-based interpretation of NMR multidimensional heteronuclear spectra of alpha-synuclein/dopamine adducts*, *Biochemistry*, **52**, no. 38, 6672–83, 2013.

Towards Restoring Catalytic Activity of Glutamine Synthetase With a Clinically Relevant Mutation

Benedikt Frieg¹, Dieter Häussinger², and Holger Gohlke¹

¹ Institute for Pharmaceutical and Medicinal Chemistry,
Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
E-mail: {benedikt.frieg, gohlke}@hhu.de

² Clinic for Gastroenterology, Hepatology, and Infectious Diseases,
Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
E-mail: haeussin@hhu.de

Glutamine synthetase (GS) is an essential enzyme in nitrogen metabolism and catalyses the ATP-dependent ligation of glutamate and toxic ammonia to glutamine. So far, three GS mutants with clinically relevant pathologies are known. These mutations lead to hyperammonemia and a global glutamine deficiency. Recently, we investigated the molecular mechanisms underlying these mutations. However, an adequate treatment for a patient carrying the R324S mutation, now eleven years old, remains elusive. Here, we investigate by means of unbiased molecular dynamics simulations if and how the compatible solute betaine, successfully used for the treatment of liver diseases, can restore GS catalytic activity in the case of R324S GS. Upon free ligand diffusion, betaine enters the R324S GS binding site and forms hydrogen bonds with S324 and weak electrostatic interactions with ATP. Thus, betaine could act as a non-covalently bound ATP binding-enhancer in the case of R324S GS.

1 Introduction

The human glutamine synthetase (GS, EC 6.3.1.2) is a homodecamer in which ten subunits form two pentameric rings that stack onto each other (Fig. 1A)¹. GS catalyses the conversion of glutamate and ammonia to glutamine under the use of adenosine triphosphate (ATP)². Thus, GS contributes to glutamate and ammonia detoxification^{3,4}, as well as to glutamine formation, which makes GS essential for the human nitrogen metabolism^{5,6}. Two GS mutations, R324C and R341C (sequence numbering of human GS¹), were linked to a reduced GS activity and led to severe brain malformations and multiorgan failure, which results in neonatal death^{9,10}. Another GS mutation, R324S, was identified in a boy, who is neurologically compromised due to the lack of ammonia detoxification and glutamine synthesis¹¹. The boy had been initially treated with glutamine supplementation¹². However, the treatment was stopped after the patient showed an increased irritability. Thus, no adequate medication is currently available to treat inborn GS deficiency caused by the R324S mutation. Recently, we showed that ATP binding is hampered in the R324S mutant because a salt-bridge interaction between R324 and ATP is lost in the case of the R324S mutant¹³. Additionally, the R324S mutant results in a gap between the S324 sidechain and ATP¹³.

In search of a treatment of this rare disease, we aim at identifying small molecules that can act as ATP binding-enhancing molecules, thereby restoring GS activity in the R324S mutant. Here, we focus on trimethyl glycine (betaine, Fig. 1B) as one such molecule. From a structural point of view, betaine could bridge the gap between S324 and ATP by

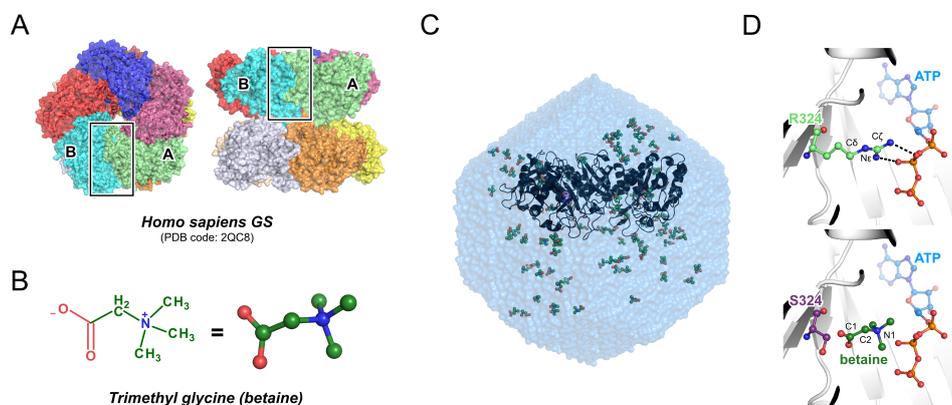


Figure 1. Human glutamine synthetase, betaine, and starting structures for unbiased MD simulations. A: Top view (left panel) and side view (right panel) of the human glutamine synthetase (GS) (PDB entry: 2QC8¹). Each subunit is coloured differently and shown in surface representation. Subunits A and B were used for a reduced dimeric model. The black box depicts the catalytic site. B: Schematic (left panel) and optimised 3D structure (right panel; ball-and-stick model without hydrogens) of the betaine molecule as used for MD simulations. C: Starting structure (“GS_{APO} model”) (black cartoon) of the R324S GS with ~ 150 mM of betaine (green spheres) in a truncated octahedron of TIP3P¹⁵ waters (blue surface representation). D: Close up view of the GS binding site. The upper panel shows wild type R324 (green) with labelled reference atoms and ATP (blue). Salt-bridge interactions are depicted as black dotted lines. The lower panel shows another starting structure (“GS_{ATP} model”) of the R324S GS mutant (magenta) with bound ATP (blue) and betaine (green) with labelled reference atoms.

forming hydrogen bonds and electrostatic interactions with S324 and ATP, respectively. From a clinical point of view, betaine is a safe, well tolerated, and inexpensive drug, and dietary betaine supplementation improved serum levels of liver enzymes in the context of fatty liver diseases¹⁴. We assessed if the GS binding site is accessible for betaine by means of unbiased molecular dynamics (MD) simulations of betaine diffusion in the presence of R324S GS. Moreover, we determined, if betaine simultaneously interacts with S324 and ATP and stabilises ATP binding.

2 Methods

To investigate if the R324S GS binding site is accessible for betaine, we performed unbiased MD simulations of R324S GS in the GS_{APO} state in the presence of ~ 150 mM of betaine (the preparation of the GS_{APO} state is described in Ref. 13). We randomly placed 92 betaine molecules around the R324S GS using *PACKMOL*¹⁶. Furthermore, to test whether a single betaine molecule can stably span the interaction between S324 and the phosphate groups of the GS substrate ATP, we generated an additional model with betaine and ATP bound to R324S GS (further referred to as GS_{ATP} state; for preparation of the GS_{ATP} state see Ref. 13). The coordinates of R324 atoms C δ , N ϵ , and C ζ were assigned to the betaine atoms C1, C2, and N1 such that the betaine’s carboxyl function is oriented towards S324 and the quaternary ammonium function towards ATP (Fig. 1D). After neutralising both the GS_{APO} and GS_{ATP} systems by adding sodium ions, the systems were placed in a truncated octahedron of TIP3P water¹⁵. Following thermalisation, both systems were subjected to

three independent MD simulations of 100 ns length using *Amber14*¹⁷ (simulation parameters are described in Ref. 13). Results are expressed as means. Mean standard error (MSE) is $< 0.02 \text{ \AA}$, unless specified differently. A one-sample Student's t-test was performed using the *R* software¹⁹.

3 Results and Discussion

To test if betaine can enhance ATP binding to the R324S mutant of GS and restore GS activity that way, we performed three unbiased MD simulations of 100 ns length of the GS_{APO} state of the R324S mutant in the presence of $\sim 150 \text{ mM}$ of betaine (Fig. 1C). Initially, we assessed whether the R324S GS_{APO} model is stable in the presence of betaine. The mean backbone root mean square deviation (RMSD) of R324S GS is $< 2.9 \text{ \AA}$, but some conformations showed RMSD values close to 5.0 \AA (Fig. 2A). Visual inspection of the trajectories revealed that those regions of the dimer are highly mobile where adjacent subunits would be located in the decamer¹³. Excluding these regions from the RMSD calculations (resulting in what is termed GS_{core} below) yielded backbone RMSD values that are largely constant and $< 3.0 \text{ \AA}$ with mean RMSD values for the $\text{GS}_{\text{core}} < 2.0 \text{ \AA}$ (Fig. 2B). As the catalytic site is located at the interface between the two subunits of the dimeric model (Fig. 1A) and, hence, in the centre of GS_{core} , we conclude that the dimeric GS_{APO} model is appropriate for investigating putative effects exerted by betaine at the catalytic site. Considering the first 20 ns as an extended equilibration phase, further analyses focused on the 20 - 100 ns interval.

Next, we analysed if betaine can enter the catalytic site of R324S GS because this will be a prerequisite for any direct betaine effect on ATP binding. For this, we computed density maps from the MD trajectories showing the frequency of interactions of betaine on the surface of R324S GS. Regions of high betaine density across all three trajectories are labelled a to f in Fig. 2C. Additional regions are visible in one or two of the MD trajectories only, suggesting that sampling of the unbound state of betaine is not converged after 100 ns of MD simulations. One region was identified where betaine is frequently present within the ATP binding site (Fig. 2D). Monitoring the minimal distance between the two carboxylate oxygens of betaine and the sidechain oxygen of S324 reveals for one MD trajectory that the betaine molecule enters the binding site after $\sim 30 \text{ ns}$ and then remains there for the remainder of the trajectory, reaching a minimal distance to S324 of 6.2 \AA (Fig. 2D). As an explanation, betaine immediately binds to one of the Mg^{2+} ions within the binding site that would interact with ATP; the Mg^{2+} ions are important for GS activity²⁰. In agreement with the slow exchange kinetics of first shell ligands of Mg^{2+} ,²¹ no unbinding of betaine from Mg^{2+} is observed during $\sim 70 \text{ ns}$ of MD simulations. We hypothesise that the bound betaine will be displaced from the Mg^{2+} by subsequent binding of ATP; due to the low likelihood of observing two simultaneous binding events within the accessible time scale of MD simulations, we are at present unable to validate this hypothesis, however. Additional regions of high betaine density close to the glutamate binding site were identified in all three trajectories (Fig. 2E). However, no betaine molecule finally entered the glutamate binding site. This is likely because glutamate binding to GS is highly regulated by an ATP-dependent opening and closing mechanism of the glutamate binding site¹. As we disregarded ATP during MD simulations, the glutamate binding site is shielded by several loop regions (Fig. 2E), preventing betaine from entering it.

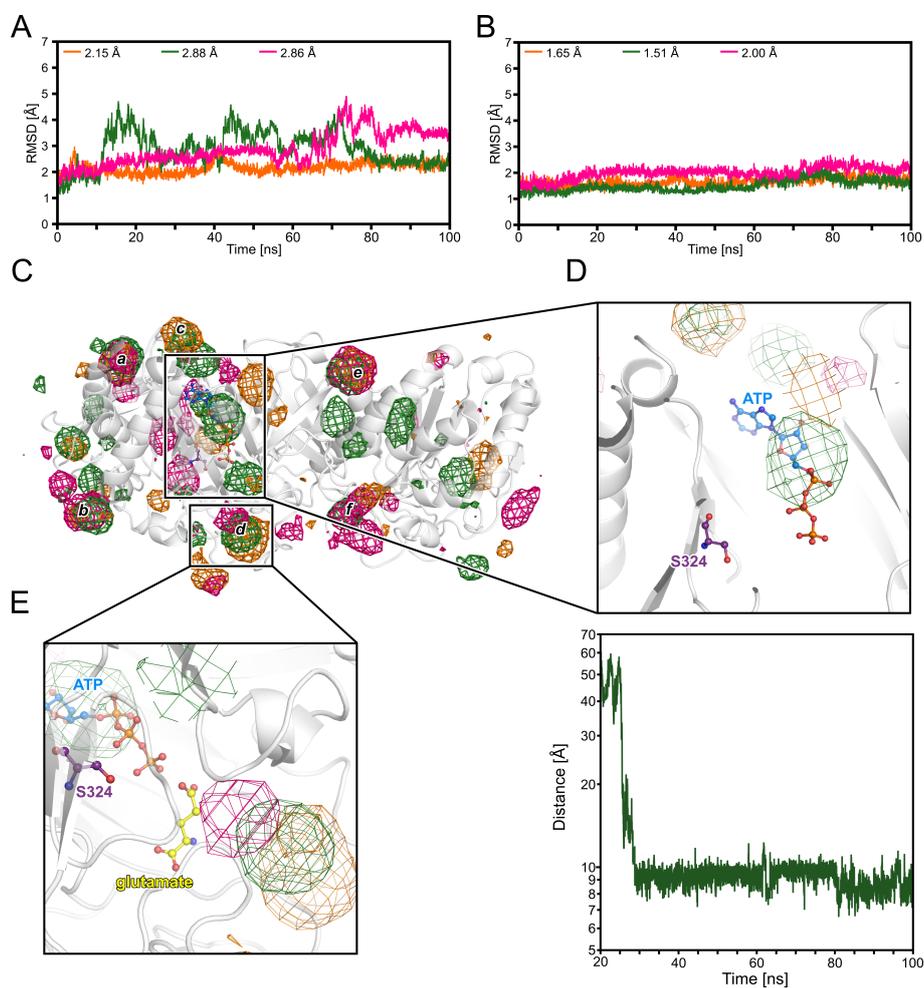


Figure 2. Unbiased MD simulations of betaine diffusion in the presence of R324S GS. A: Backbone RMSD of R324S GS in the GS_{APO} state for three trajectories. Mean values are provided in the legend. B: Backbone RMSD for GS_{core} residues of R324S GS in the GS_{APO} state for three trajectories. Mean values are provided in the legend. C: Overlay of density maps (depicted as isocontour mesh; contour level is 80 sigma) showing the frequency of betaine interactions on the surface of R324S GS_{APO} (white cartoon representation) of trajectory 1 (orange), 2 (green), and 3 (magenta). Labels a to f depict regions of high betaine density coherently found in all trajectories. D: The top panel shows a close up view of the ATP binding site with regions of high betaine density. The lower panel shows the minimal distance between the two carboxylate oxygens of the respective betaine molecule and the sidechain oxygen of S324 for the 20 - 100 ns interval of trajectory 2. E: Close up view of regions of high betaine density close to the glutamate binding site. For panels D and E, ATP (blue ball-stick model) and glutamate (yellow ball-stick model) were added only to visualise their location; both molecules were not present during the MD simulations. Residue S324 is depicted in magenta.

Next, to probe if betaine stably binds between S324 and ATP, thereby bridging the gap between the amino acid's side chain and ATP, we performed three additional MD simulations of 100 ns length in which a single betaine molecule was initially placed in the gap

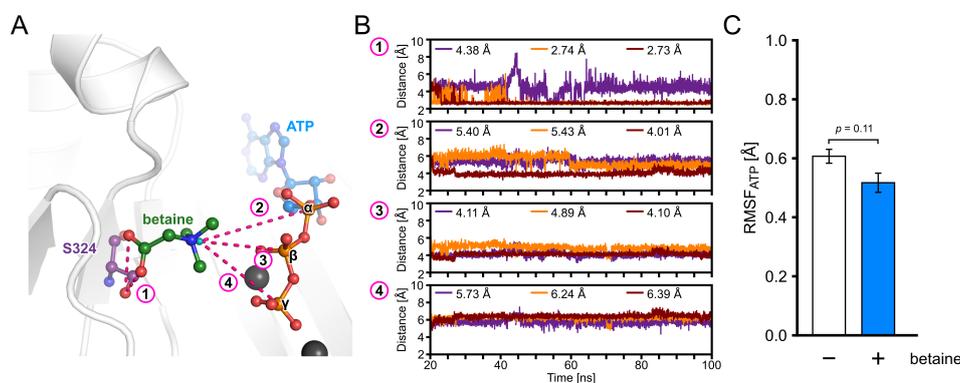


Figure 3. Structural analyses of betaine bound to the R324S glutamine synthetase binding site. A: Close up view of betaine (green) in the S324 GS binding site with residue S324 (purple) and ATP (blue) as blue ball-stick models. The geometric centre (GC) of the betaines trimethyl amine group is depicted as cyan sphere. Distances 1 - 4 (see text for definitions) are shown as dotted magenta lines. B: Distances 1 - 4 over trajectories 1, 2, and 3 (purple, orange, and red lines) with respective mean values in the legend. For distance 1, only the smallest distance to one of the oxygens of the carboxylate group of betaine, and for distance 2 - 4, only the smallest distance to one of the oxygens of the α , β , and γ phosphate groups in ATP are considered, respectively. C: Mean RMSF of ATP over three trajectories (mean standard error depicted as error bars). The white bar shows RMSF values in the absences (-) of betaine (MD simulations were previously described in Ref. 13), the blue bar in the presence (+) of betaine.

(Fig. 1D). To analyse hydrogen bond formation with S324 and formation of an electrostatic interaction between the positively charged trimethylamine function in betaine and the negatively charged phosphate groups of ATP, we measured four distances (Fig. 3A). Distance 1, measured as minimal distance between the two carboxylate oxygens of the betaine molecule and the sidechain oxygen of S324, is rather constant in trajectories 2 and 3 with mean distances of 2.73 Å and 2.74 Å (Fig. 3B). In these trajectories, a hydrogen bond (using a distance of 3.2 Å between the two donor and acceptor atoms and an angle (donor atom, H, acceptor atom) of 120° as cutoff criteria¹⁸) is present between both functional groups in at least 90% of all conformations. The smallest distances between the geometric centre of the quaternary ammonium group in betaine to one of the oxygens of the α , β , and γ phosphate groups in ATP, respectively, show the overall smallest mean values across all three trajectories in the case of the β phosphate group (values between 4.10 Å and 4.89 Å). In the case of trajectory 3, a small mean value of 4.01 Å is found with respect to the α phosphate group, too. Considering that the radial distribution function between nitrogen atoms of the choline cation and oxygen atoms of alanine anions revealed favourable interactions between both groups in a distance range from 4.00 Å to 6.0 Å²², we suggest that such interactions also exist between betaine and at least the β phosphate group of ATP. Together with the hydrogen bond analysis above, our results suggest that betaine can bridge the gap between S324 and ATP and, thus, enhance ATP binding.

To investigate whether betaine has a stabilising effect on ATP binding to R324S GS, we calculated the root mean square fluctuation (RMSF) of ATP in the presence of betaine after superimposition of backbone atoms of residues within 5 Å of ATP in the starting structure (Fig. 3C, blue bar). We compared values to those extracted from MD simula-

tions of R324S GS in the GS_{ATP} state without betaine, which were described in Ref. 13 (Fig. 3C, white bar). The RMSF value decreases from 0.61 Å (\pm 0.03 Å) without betaine to 0.52 Å (\pm 0.04 Å) in the presence of betaine (Fig. 3C); the difference is weakly significant ($p = 0.11$). As the RMSF is a measure of atomic mobility, this indicates that betaine weakly stabilises ATP binding in the R324S GS binding site, in agreement with the above geometric analysis of intermolecular interactions.

4 Conclusion and Outlook

The clinically relevant R324S mutation is suggested to reduce catalytic activity of GS¹², which leads to hyperammonemia and a global glutamine deficiency¹¹. Previously, we found that ATP binding is hampered in the R324S mutant¹³. Here, we investigated by MD simulations of in total 600 ns length if betaine can act as an ATP binding-enhancing molecule. Our results show that betaine can enter the ATP binding site of R324S GS_{APO} state and come close to S324. In the GS_{ATP} state, betaine simultaneously forms hydrogen bonds with S324 and electrostatic interactions with at least the β phosphate group of ATP. In agreement, betaine binding reduces the mobility of ATP in R324S GS. In all, this indicates that betaine weakly stabilises ATP binding in the R324S GS binding site.

From our results, several suggestions arise for how to strengthen an ATP binding-enhancing effect. First, the interactions with the phosphate group may be strengthened if, instead of the bulky trimethylammonium group of betaine, a positively charged group such as an amidino or guanidine moiety is chosen that can form salt bridge interactions with the phosphate oxygens in addition. Glycocyanine (2-guanidino acetate), a metabolite of glycine, or creatine (2-[amidino(methyl)amino]acetate), a methyl derivative of glycocyanine, which are both used as dietary supplements to treat creatine deficiency^{23,24}, contain such guanidine groups. Second, the negatively charged carboxylate group of betaine, which results in disfavoured electrostatic interactions with the phosphate groups, may be replaced by a neutral hydroxymethyl group, resulting in the positively charged molecule choline. Choline is used as a dietary supplement²⁵ and in the treatment of liver diseases²⁶. We intend to investigate these alternatives further by MD simulations and experimental validation in order to aid in the identification or development of ATP binding-enhancing molecules by which the R324S GS mutant can be repaired extrinsically²⁷.

Acknowledgements

We gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA and JURECA at Jülich Supercomputing Centre (JSC) (NIC project HDD13). We are grateful to the “Center for Information and Media Technology” (ZIM) at the Heinrich Heine University of Düsseldorf (Germany) for computational support. This work was supported by the Deutsche Forschungsgemeinschaft through the Collaborative Research Center SFB 974 (Communication and Systems Relevance during Liver Damage and Regeneration, Düsseldorf).

References

1. W. W. Krajewski et al., *Crystal structures of mammalian glutamine synthetases illustrate substrate-induced conformational changes and provide opportunities for drug and herbicide design.*, J. Mol. Biol. **375**, 217-228, 2008.
2. A. Meister, *Glutamine synthetase of mammals.*, The enzymes **10**, 699-754, 1974.
3. D. Häussinger, *Hepatocyte heterogeneity in glutamine and ammonia metabolism and the role of an intercellular glutamine cycle during ureogenesis in perfused-rat-liver.*, Eur. J. Biochem. **133**, 269-275, 1983.
4. N. Qvartskhava et al., *Hyperammonemia in gene-targeted mice lacking functional hepatic glutamine synthetase.*, Proc. Natl. Acad. Sci. U. S. A. **112**, 5521-5526, 2015.
5. D. Häussinger, *Nitrogen metabolism in liver: structural and functional organization and physiological relevance.*, Biochem. J. **267**, 281-290, 1990.
6. D. Häussinger, *Hepatic glutamine transport and metabolism.*, Adv. Enzymol. Relat. Areas. Mol. Biol. **72**, 43-86, 1998.
7. A. Martinez-Hernandez, K. P. Bell, and M. D. Norenberg, *Glutamine synthetase: glial localization in brain.*, Science **195**, 1356-1358, 1977.
8. R. Gebhardt and D. Mecke, *Heterogeneous distribution of glutamine synthetase among rat liver parenchymal cells in situ and in primary culture.*, Embo J. **2**, 567-570, 1983.
9. J. Häberle et al., *Brief report - Congenital glutamine deficiency with glutamine synthetase mutations.*, N. Engl. J. Med. **353**, 1926-1933, 2005.
10. J. Häberle et al., *Inborn error of amino acid synthesis: human glutamine synthetase deficiency.*, J. Inherit. Metab. Dis. **29**, 352-358, 2006.
11. J. Häberle, N. Shahbeck, K. Ibrahim, G. F. Hoffmann, and T. Ben-Omran, *Natural course of glutamine synthetase deficiency in a 3 year old patient.*, Mol. Genet. Metab. **103**, 89-91, 2011.
12. J. Häberle et al., *Glutamine supplementation in a child with inherited GS deficiency improves the clinical status and partially corrects the peripheral and central amino acid imbalance.*, Orphanet J. Rare Dis. **7**, 1-10, 2012.
13. B. Frieg et al., *Molecular mechanisms of glutamine synthetase mutations that lead to clinically relevant pathologies.*, submitted.
14. M. F. Abdelmalek, P. Angulo, R. A. Jorgensen, P. B. Sylvestre, and K. D. Lindor, *Betaine, a promising new agent for patients with nonalcoholic steatohepatitis: results of a pilot study.*, Am. J. Gastroenterol. **96**, 2711-2717, 2001.
15. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *Comparison of simple potential functions for simulating liquid water.*, J. Chem. Phys. **79**, 926-935, 1983.
16. L. Martinez, R. Andrade, E. G. Birgin, and J. M. Martinez, *PACKMOL: a package for building initial configurations for molecular dynamics simulations.*, J. Comput. Chem. **30**, 2157-2164, 2009.
17. D. A. Case et al., *AMBER 14.*, University of California, San Francisco, 2014.
18. G. R. Desiraju, G.R. and T. Steiner, *The weak hydrogen bond.*, Oxford University Press, New York, 2001.
19. R Core Team *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2008.

20. H. B. Pinkofsky, M. R. Maurizi, and A. Ginsburg, *Binding Mn^{2+} or Mg^{2+} to active-sites of glutamine-synthetase from bovine brain.*, Fed. Proc. **44**, 1807-1807, 1985.
21. H. Ohtaki and T. Radnai, *Structure and dynamics of hydrated ions.*, Chem. Rev. **93**, 1157-1204, 1993.
22. M. Campetella et al., *Interaction and dynamics of ionic liquids based on choline and amino acid anions*, J. Chem. Phys. **142**, p234502, 2015.
23. S. M. Ostojic, B. Niess, M. Stojanovic, and M. Obrenovic, *Co-administration of methyl donors along with guanidinoacetic acid reduces the incidence of hyperhomocysteinaemia compared with guanidinoacetic acid administration alone.*, Br. J. Nutr. **110**, 865-870, 2013.
24. E. Hultman, K. Soderlund, J. A. Timmons, G. Cederblad, and P. L. Greenhaff, *Muscle creatine loading in men.*, J. Appl. Physiol. **81**, 232-237, 1996.
25. S. H. Zeisel, *Choline: an essential nutrient for humans.*, Nutrition **16**, 669-671, 2000.
26. A. L. Buchman et al., *Choline deficiency - a cause of hepatic steatosis during parenteral-nutrition that can be reversed with intravenous choline supplementation.*, Hepatology **22**, 1399-1403, 1995.
27. N. Chondrogianni et al., *Protein damage, repair and proteolysis.*, Mol. Aspects Med. **35**, 1-71, 2014.

Using Molecular Dynamics to Model the Stacking Behaviour of Perylene Bisimide Derivatives in Aromatic Solvent

Manuel Hollfelder and Stephan Gekle

Biofluid Simulation and Modeling, University of Bayreuth, 95444 Bayreuth, Germany

E-mail: stephan.gekle@uni-bayreuth.de

Using the power of present-day supercomputers makes it possible to solve Newton's equations of motion numerically, even for large systems. Molecular Dynamics provides a powerful simulation method to get insight in the behaviour of molecules under many conditions when considering atoms as charged solid spheres and hiding the quantum mechanics in effective potentials. We use this method to characterise the physical properties of different systems, e.g. polymers, water, organic and inorganic molecules.

As an example of our work, we present a model of the π - π stacking dynamics of a perylene bisimide molecule when solvated in aromatic solvents. Our calculations show that the transition from the open (unstacked) to the stacked configuration is hindered by a small free energy barrier of approx. $1 k_B T$ in the aromatic solvent toluene. The origin of this barrier is traced back to π - π interactions between perylene and the aromatic solvent which are very similar in nature to those between two PBI monomers. The stacking process proceeds in three phases via two well-defined transition states: (i) in the first phase, the two PBI molecules share part of their respective solvation shells forming the first transition state. Further approach needs to squeeze out the shared solvent layer thus creating the energy barrier. (ii) After removal of the separating solvent the two PBIs form a second transition state with one monomer located at a random position in the other's solvation shell. (iii) Finally, the two PBIs slide on top of each other into their final stacked position.

1 Introduction

Newton's equations of motion cannot be solved analytically for large systems. The term large system is, however, misleading since every system above two particles must be considered of this type. Therefore, modelling atoms as solid and charged spheres in the framework of Molecular Dynamics leads only to new results if one takes the possibility of numerical calculations into account. Nowadays supercomputers allow us to solve the equations of Molecular Dynamics simulations for truly large systems, e.g. number of atoms above 20 000, within finite time.

Using the computation time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA at Jülich Supercomputing Centre (JSC) we model different molecules, e.g. polymers, organic and inorganic molecules, water, and derive models of interaction or calculate physical properties of these molecules. This article gives a brief summary of our work and presents as an example results obtained from numerical calculations of an organic dimer consisting of two covalently linked perylene bisimide molecules (PBM) using highly parallelised supercomputers¹.

Sec. 2 summarises the motivation of our investigations. After introducing briefly the methods in Sec. 3, we use Sec. 4.1 to characterise the structure of the solvation shell around a single PBM when solvated in toluene as a typical aromatic solvent. In Sec. 4.2 we then

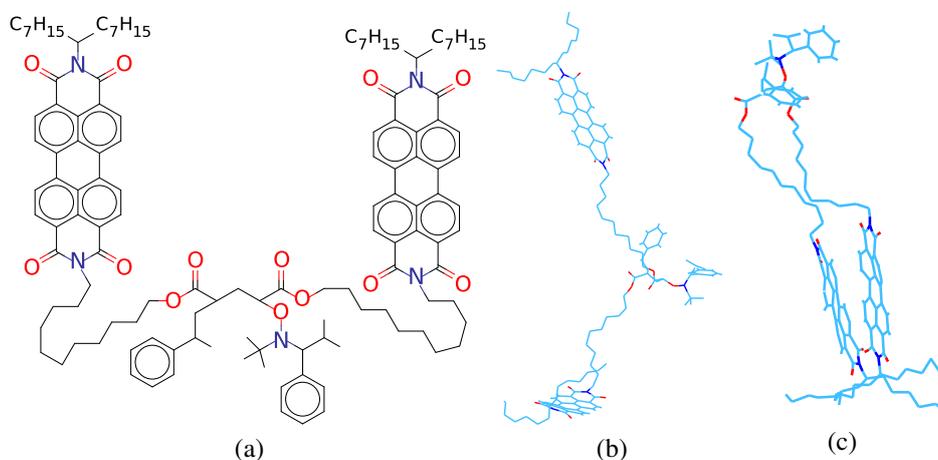


Figure 1. (a) Chemical structure of the investigated perylene bisimide dimer molecule. (b) Snapshot of an unstacked (OPEN) state and (c) snapshot of a stacked (STACK) configuration. The solvent toluene is omitted for clarity.

present the free energy profiles for the dimer solvated in toluene which governs the stacking transition. In Sec. 4.3 we describe the full stacking pathway which proceeds from the open to the stacked configuration via two well-defined transition states.

2 Motivation

Molecules based on derivatives of perylene are a widely used material for many different purposes: they have useful optical properties because their emission colour can be adjusted over a wide range of the visible spectrum, they are used as highly sensitive sensors and show n-type conduction making them suitable for organic transistors and optoelectronic devices in general. Furthermore, molecules based on perylene derivatives are considered a promising candidate for building highly efficient organic solar cells.

Given these numerous applications, perylene derivatives have been intensively investigated both experimentally and computationally in the recent past. One of their salient characteristics is a tendency to form large tower-like aggregates (stacks) of molecules as has been demonstrated by experiments and *ab initio* molecular dynamics simulations. This stacking behaviour can be traced back to the interactions of the π -orbitals and has great influence on the properties of the materials as it dramatically changes, e.g., the efficiency of energy and charge transfer processes. The static structure of these stacks has been well characterised: in the stacked state the perylenes remain mostly planar with the distance between the two planes being around 0.35 nm; the perylene axes are tilted by approx. 45° with respect to each other and the free energy gained from stacking is of the order of 15 kJ/mol depending on the exact chemical structure of the perylene derivative as determined by linear free energy relationships in conjunction with UV/Vis spectroscopy².

A systematic investigation of the dynamic transition pathway from the open (unstacked) to the stacked state, however, has not been conducted so far. Here, we have

investigated the transition dynamics considering a perylene bisimide dimer solvated in toluene by Molecular Dynamics (MD) simulations. The investigated structure consists of two perylene bisimide monomers (PBM) and is shown in Fig. 1 (a). Both monomers are connected by an alkane chain containing furthermore oxygen, nitrogen and two aromatic rings. On the other side of the PBM two C_7H_{15} chains are present. We consider the solvent toluene representing a prototypical aromatic solvent. The toluene-solvated system has recently been investigated by fluorescence spectroscopy^{3,4}. The unstacked (OPEN) and stacked (STACK) configurations are well reproduced by our MD simulations as shown in Fig. 1 (b) and (c). We find the distance of the planes defined by the (almost perfectly) planar PBMs to be 0.36 nm which is in good agreement with values obtained from experiments and earlier MD simulations⁵.

3 Simulation Methods

Classical Molecular Dynamics simulations were run with Gromacs and the Gromos 53a6 force field was used. This means that the parameters of the bonds, harmonic spring constants, Lennard-Jones parameters etc. follow a parameter set of typical bonding values and interactions which is itself consistent. Building new molecules reduces to combining the correct parameters for the bonds between the atoms and assigning charges to the atoms. These force field topologies were automatically calculated for the perylene derivatives using *Automated Force Field Topology Builder (ATB) and Repository*. The force field file for the solvent toluene was calculated using PRODRG. United-atom force field topologies were used for the simulations which means that hydrogen atoms and carbon atoms, e.g. for alkane chains, are combined to super-atoms to reduce the necessary computation time. The visual analysis of the molecular structure files and trajectories was carried out using VMD.

For the simulations the dimer was solvated in a rectangular box with 1000 toluene molecules. After energy minimisation and NVT equilibration, the final runs were simulated as NPT ensembles at 300 K and 1 bar. In order to investigate the general behaviour and stacking of the dimer, simulations were started with different starting configurations of the atoms' positions and velocities. Every simulation was stopped as soon as stacking occurred.

For the free energy calculations in Sec. 4.2 we use umbrella sampling where the distance between the centres of mass of the PBMs was chosen as a reaction coordinate and different states along that coordinate were created using the Gromacs pull code. After equilibration the simulations were simulated as NPT ensembles for each window at 300 K and 1 bar. For the bias potential the *umbrella potential* implemented in the Gromacs package was used and simulations with different spring constants were started. The free energy profile was calculated using the *Weighted Histogram Analysis Method* implemented in Gromacs as *g_wham*.

4 Results and Discussion

4.1 Static Solvation Shell for Toluene

We first analyse the static structure of the toluene solvation shell around a single PBM of the dimer in the OPEN state. We find a clear first solvation shell whose time averaged

density is, however, rather inhomogeneous as shown in Fig. 2 from different perspectives. Since the monomers are fairly stiff they remain almost perfectly planar during the entire simulation time.

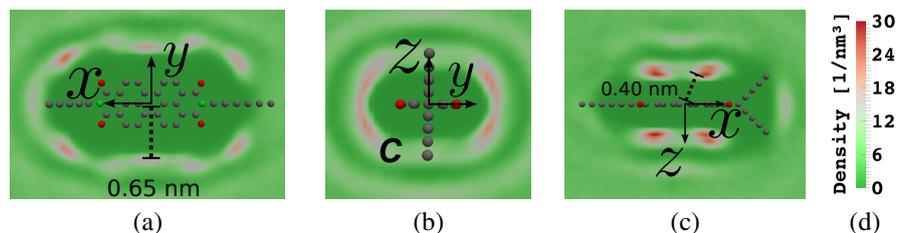


Figure 2. Average density of the solvent toluene around one PBM of the dimer during the OPEN state. The density is illustrated as coloured plots over planar cuts through the PBM centre of mass along the xy (a), yz (b), and xz (c) plane with the legend of the plots in (d). The PBM's carbon, oxygen, and nitrogen atoms are depicted in grey, red, and green, respectively. In (b) C marks a special point which is referred to during the stacking mechanism in Sec. 4.3.

In Fig. 2 (a), which corresponds to a view from the top, alternating regions of high and low densities can be distinguished. The distance from the monomer's centre of mass (COM) to the high density spots on the y -axis is approximately 0.65 nm. Near the partially charged oxygen atoms we observe regions of fairly low solvent density. In Fig. 2 (b), which represents a head-on view, one can see two rather homogeneous ellipsoidal rings formed around the monomer by the solvent. The side view, Fig. 2 (c), illustrates four clear spots of high solvent density right above and below the PBM. The distance between the spots right above and below the PBM and the PBM itself is in z -direction 0.40 nm corresponding to the well-known stacking distance between perylene molecules. This confirms the existence of a π - π -interaction between the perylene and the aromatic solvent very similar in nature to the π - π interaction in perylene stacks. This interaction strongly influences the stacking dynamics as we will show further below.

4.2 Free Energy of Stacking

As a first step towards understanding the stacking pathway of the dimer in aromatic solvent, we present in Fig. 3 the free energy profile as a function of the distance between the COMs of the two monomers for toluene. The profiles are calculated using umbrella sampling as described in the methods section. Any stacking pathway needs to proceed from the plateau region at the right of Fig. 3 corresponding to the OPEN state towards the minimum located at 0.38 nm corresponding to the STACK state at the very left. Note that the distance between the COM considered here is slightly larger than the plane-plane distance of 0.36 nm in the STACK state due to the tilting of the monomers with respect to each other.

There is a clear energy barrier with a height of approx. $1 k_B T$ separating the STACK and OPEN states from each other in the solvent toluene. The barrier is rather broad, starting at 1.40 nm and extending down to 0.8 nm. From the top of the barrier the free energy drops sharply towards its minimum. Around 0.50 nm we observe a turning point marking

a change in the curvature of the free energy curve. Finally, for distances smaller than 0.38 nm the free energy rises steeply which is caused by unfavourable squeezing of the PBMs during the stacked state.

The free energy difference between the minimum identified with the STACK state and the OPEN state with constant free energy for distances beyond 1.40 nm is approximately $5.2 k_B T$ corresponding to 13 kJ/mol. This is in good agreement with values obtained from UV/Vis experiments for the aggregation of perylene bisimide: for a slightly different structure of the PBMs (two additional benzene molecules and no connecting carbon chain) the free energy change was obtained as 15.8 kJ/mol for the solvent toluene².

The density profiles of the toluene solvation shells around the monomers (Fig. 2) allow a physical interpretation of the distinct features observed in the free energy profile in the transition region between 1.4 nm and 0.8 nm. Noting from Fig. 2 (b) and (c) that the solvation shells are located at distances between 0.40 nm in z -direction and 0.65 nm in y -direction, the first contact of the solvation shells is expected to occur for distances between 0.8 nm and 1.3 nm (depending on the relative orientation of the PBMs during approach) which corresponds closely to the width and position of the energy barrier for toluene. After contact, further approach of the PBMs is only possible if (at least) one of the PBMs loses part of its solvation shell and the two PBMs further on “share” solvent molecules. Scraping off the solvation shell clearly requires energy in the case of the solvent toluene thus explaining the observed energy barrier. For distances smaller than about 0.7 nm the PBMs start interacting directly leading to a rapid drop towards the STACK state which represents the absolute minimum of the free energy.

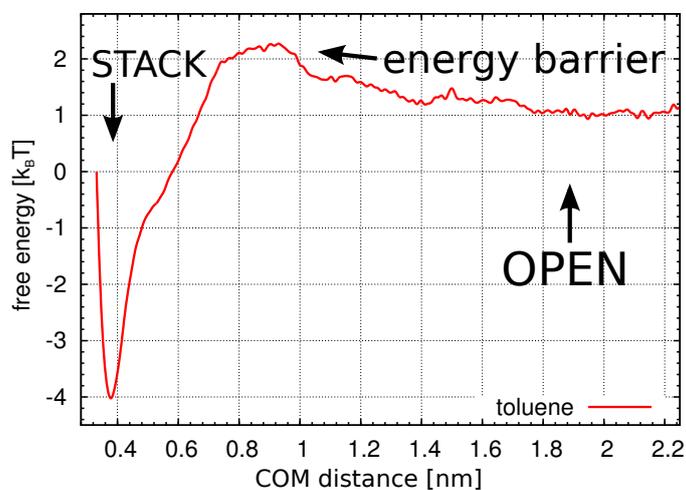


Figure 3. Free energy change as a function of the centre of mass (COM) distance between both PBMs of the dimer for the solvent toluene (red). The states STACK/OPEN and the energy barrier for toluene are marked and the energy profiles are shifted against each other for clarity.

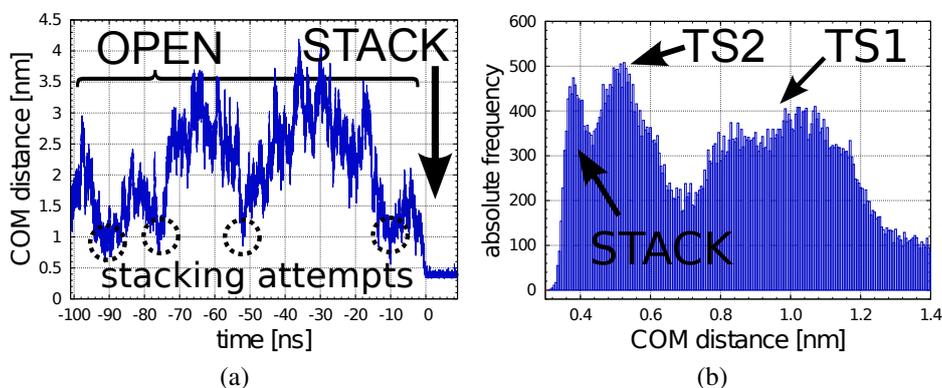


Figure 4. (a) The COM distance between both PBMs with stacking at $t = 0$ ns. The distance decreases several times to values between 0.80 nm and 1.00 nm (marked spots) without causing a change to the stacked state STACK which is characterised by an average COM distance of about 0.38 nm. (b) Histogram over COM distances for the transition from OPEN to STACK in the solvent toluene. Only frames < 2000 ps before stacking are taken into account. The peak around 1.00 nm corresponds to the transition state TS1, the second one around 0.50 nm to TS2 and the third one around 0.38 nm to the state STACK.

4.3 Dynamic Stacking Pathway of Perylene Bisimide in Toluene

We now turn to the actual stacking pathway for toluene as an example of an aromatic solvent. Fig. 4 (a) shows the COM distance of the two monomers as a function of time for a typical trajectory before and during the stacking process. The origin of time at $t = 0$ ns has been located at the stacking transition which clearly separates the STACK from the OPEN configurations. Before the actual transition, however, the distance decreases several times to values between 0.80 nm and 1.00 nm but without continuing to the final stacked state. These “failed attempts” are a consequence of the energy barrier seen in Fig. 3 which can be overcome by thermal fluctuations only in a small number of cases.

We now focus more closely on the actual stacking process. The most notable feature are two regions of fairly constant COM distance during the stacking of the monomers during the simulations. These regions correspond to two well-defined transition states which we shall denominate in the following as TS1 and TS2, respectively. These transition states can be found with variable time duration in all our simulations during the stacking process. To give a full picture, we present in Fig. 4 (b) a histogram taken over all simulations during the last 2000 ps before stacking. The value of 2000 ps was estimated from the average time the stacking process lasts during the different simulations. Beside the expected peak corresponding to the STACK state at 0.38 nm, the histogram clearly shows two further peaks: a broad first peak centred around 1.0 nm and a second sharper one at about 0.50 nm. The COM distances where these peaks occur correspond to the transition states TS1 and TS2 and are evidence of the general nature of the transition states.

The picture that thus emerges is a stacking pathway consisting of three consecutive phases. In the first phase both PBMs approach each other closely until the respective solvation shells first come into contact and subsequently merge to form the first transition state TS1. In this state, which is illustrated in Fig. 5 (a), the two monomers share part of their respective solvation shell (green solvent molecules). Further approach would require the

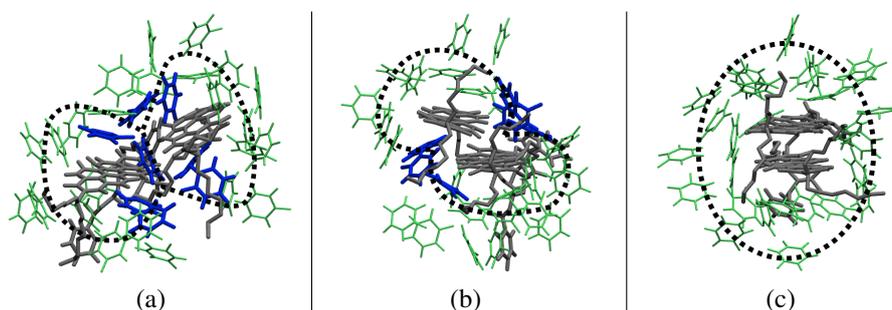


Figure 5. The different states during the stacking process illustrated with the solvation shell (dashed lines) around the PBMs. (a) View of the first transition state TS1 in which the PBMs share part of their solvation shells (shared solvent has blue colour). (b) Second transition state TS2 where the PBMs replace the solvent by themselves and only some solvent remains next to them (blue molecules) which stabilises this state. (c) Final stacked state.

complete removal of the shared solvent molecules which is connected to the energy barrier as observed in Fig. 3. In most cases, the dynamics continues without transcending the barrier and the PBMs separate again (corresponding to the “failed attempts” in Fig. 4). In some cases, however, the energy barrier can be overcome by adequate thermal fluctuations and the dynamics continues to the second phase.

In the second phase, the solvent between the PBMs is expelled, the PBMs approach more closely, and the first monomer becomes part of the second monomer’s solvation shell (and *vice versa*). In TS2 there are still some solvent molecules present as shown in Fig. 5 (b) by the green solvent molecules. These stabilise TS2 by their π - π interactions with the perylene. The existence of TS2 furthermore reflects itself in the slightly flatter part of the free energy profile around 0.50 nm in Fig. 3. Both monomers arrange themselves such that the normal vectors to the PBM planes are parallel. For the position of the two PBMs relative to each other we find the most probable angles around 47° of the connecting COM and the normal vector showing that the monomers approach each other via a well-defined “channel” through the point marked A in Fig. 2 (b). The time span during which TS2 exists strongly varies with the second monomer’s exact location and the stabilising solvent molecules. Once TS2 has been reached, however, the system is already situated on the steep downward slope of the free energy which makes final stacking inevitable.

In the third and final phase, the second monomer revolves around the first one in order to find its final stacked position with the two PBM planes lying right on top of each other, possessing the well-known COM distance of 0.38 nm and a joint solvation shell as shown in Fig. 5 (c).

5 Conclusion

In conclusion, we have elucidated the dynamic stacking pathway of perylene bisimide dimers solvated in toluene using Molecular Dynamics simulations in combination with free energy calculations. In toluene, the stacked and unstacked states are well separated by a free energy barrier of approx. $1 k_B T$ which is due to the fairly rigid solvation shells that the aromatic solvent forms around each perylene bisimide monomer. During the stacking

transition this barrier is overcome in three phases: (i) after the initial approach and the first contact of the solvation shells, the two monomers form a metastable first transition state (TS1) in which they share part of their respective solvation shells. Further approach requires the removal of the shared region of the solvation shell thus explaining the observed energy barrier. (ii) If thermal fluctuations allow the system to overcome this energy barrier, the system forms a second transition state (TS2) in which the interjacent solvent of the shared solvation shell is expelled and substituted by one of the perylene monomers themselves which, loosely speaking, “solvate themselves” in this state. (iii) In the last step the monomers revolve around each other until the two planes lie on top of each other which represents the final stacked configuration.

A central result of our investigations is that the stacking pathway is mainly determined by the ability of the aromatic solvent to form π - π -interactions with the solute very similar to those formed between the perylenes themselves. This leads to an energy barrier which must be overcome before the system can form the familiar stacks.

Acknowledgements

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA at Jülich Supercomputing Centre (JSC).

Financial support comes from the Volkswagen Foundation in the framework of the Lichtenberg program and from the Graduiertenkolleg 1640. Further thanks are owed to Florian Spreitler, Jürgen Köhler and Mukundan Thelakkat for helpful discussions.

References

1. M. Hollfelder and S. Gekle, *Dynamic Stacking Pathway of Perylene Dimers in Aromatic and Nonaromatic Solvents*, *The Journal of Physical Chemistry B*, **119**, no. 32, 10216–10223, 2015, PMID: 26186499.
2. Z. Chen, B. Fimmel, and F. Würthner, *Solvent and Substituent Effects on Aggregation Constants of Perylene Bisimide π -Stacks - a Linear Free Energy Relationship Analysis*, *Org. Biomol. Chem.*, **10**, 5845–5855, 2012.
3. F. Spreitler, M. Sommer, M. Thelakkat, and J. Köhler, *Conformational Dynamics of Di-(Perylene Bisimide Acrylate) and Its Footprints in Steady-State, Time-Resolved, and Fluorescence-Correlation Spectroscopy*, *Phys. Chem. Chem. Phys.*, **14**, 7971–7980, 2012.
4. F. Spreitler, M. Sommer, M. Hollfelder, M. Thelakkat, S. Gekle, and J. Köhler, *Unravelling the Conformations of Di-(Perylene Bisimide Acrylate) by Combining Time-Resolved Fluorescence-Anisotropy Experiments and Molecular Modelling*, *Phys. Chem. Chem. Phys.*, **16**, 25959–25968, Nov. 2014.
5. V. Marcon, J. Kirkpatrick, W. Pisula, and D. Andrienko, *Supramolecular Structure of Perylene Tetracarboxdiimides*, *phys. stat. sol. (b)*, **245**, no. 5, 820–824, May 2008.

Computational Chemistry

Computational Chemistry: *Ab initio* Molecular Dynamics and Electron Transport

Christine Peter

Theoretical and Computational Chemistry, University of Konstanz, 78457 Konstanz, Germany
E-mail: christine.peter@uni-konstanz.de

Due to a combination of methodological advancement, continuous improvement of program packages, and powerful supercomputers such as the ones at the NIC and GCS computing facilities, *Ab initio* Molecular Dynamics (AIMD) simulations and first principles studies of electron dynamics have reached system sizes and time scales that permit the investigation of complex chemical processes in novel materials and fundamental insight into biomolecular systems. The present proceedings of the NIC Symposium 2016 show three beautiful examples of such state-of-the-art simulations in the chemistry world.

In their contribution *Ab Initio Molecular Dynamics Simulations of Ionic Liquids* Martin Thomas, Iris Sancho Sanz, Oldamur Hollóczy, and Barbara Kirchner report on extensive AIMD simulations of two ionic liquids systems, using the CP2K program package. The authors compare the two systems regarding the formation of hydrogen bonds, and they analyse the systems in terms of pair correlation functions and higher order distribution functions where distance-angle correlations are considered. Based on these data the microheterogeneity in the sample is analysed, which results from the microscale segregation of polar and nonpolar regions in the molecular structure.

In *Prebiotic Chemistry in Nanoconfinement* Daniel Muñoz-Santiburcio and Dominik Marx combine AIMD with the metadynamics technique to carry out a mechanistic and energetic study of peptide synthesis reactions with the long-term goal to test the “Iron-Sulfur-World” origin of life hypothesis. The authors have carried out AIMD simulations of nanoconfined water in a slab between two layers of mackinawite, an iron/sulfur mineral that occurs at hydrothermal vents. They find that moderately nanoconfined water in contact with these surfaces has dielectric and proton-conducting properties that make it a good medium for prebiotic peptide synthesis. Following up on these results, the authors have extended their earlier simulations of the prebiotic peptide cycle to nanoconfined hot-pressurised water and come to the conclusion that extreme conditions and nanoconfinement with the resulting steric, entropic and dielectric effects make the investigated environment highly suitable for prebiotic polypeptide synthesis.

In *Simulation of Electron Transfer and Electron Transport in Molecular Systems at Surfaces* Pedro B. Coto, Chriszandro Hofmeister, Veronika Prucker, Dominik Weckbecker, and Michael Thoss show results of combined first-principles electronic structure calculations with quantum dynamics simulations to investigate electron transfer processes in systems of organic molecules on metal surfaces. They have studied, which factors control the electron injection dynamics in nitrile-substituted alkanethiolate self-assembled monolayers on the Au(111) surface. The authors have also simulated electron transport in single molecule junctions of organic molecular bridges. They show that these molecules can be

used as a nanoswitch after a proton transfer reaction is triggered by an external electrostatic field. The results suggest exciting new options for the design of nanoelectronic devices.

These three chemistry articles give an excellent impression of the scales and complexity that can be reached by *ab initio* quantum chemistry and first-principles atomistic calculations. It is now possible to study interactions that govern structure formation or solvent effects on chemical reactions on an electronic structure level, and one can simulate electron transfer in organic electronic devices.

Ab Initio Molecular Dynamics Simulations of Ionic Liquids

Martin Thomas, Iris Sancho Sanz, Oldamur Hollóczy, and Barbara Kirchner

Mulliken Center for Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn,
Beringstraße 4, 53115 Bonn, Germany
E-mail: kirchner@thch.uni-bonn.de

We present *ab initio* molecular dynamics simulations of the ionic liquids 1-butyl-3-methylimidazolium trifluoromethanesulfonate and 1-butyl-1-methylpyrrolidinium trifluoromethanesulfonate. We compare the trajectories regarding the formation of hydrogen bonds and the microheterogeneity due to polar and nonpolar phases.

1 Introduction

In the last decades, ionic liquids (ILs) became an extensively investigated type of substances due to their excellent properties as solvents that allow for a wide range of interesting applications¹⁻³. The structure of ILs is governed by a large variety of interaction forces such as electrostatic attraction and repulsion⁴, dispersion interaction^{5,6}, and directional atom contacts as in hydrogen bonds^{7,8}. A theoretical model has to take into account all these effects to be able to make reasonable predictions about the properties of ILs⁹. On the one hand, static quantum chemical calculations, which explicitly treat the electronic structure of the system on the basis of the Schrödinger equation, can accurately describe the wide range of interaction forces for a particular molecular configuration. On the other hand, molecular dynamics simulations, which rely on empirical force fields to propagate the atoms in the system according to classical mechanics, are able to provide insight into the highly dynamic nature of the interaction network. To benefit from the advantages of both approaches, they can be combined in *ab initio* molecular dynamics (AIMD)^{10,11}, where classical mechanics is used to describe the propagation of the atoms, but a quantum chemical electronic structure method is employed to obtain the forces. Due to its good cost-performance ratio, density functional theory (DFT) is usually applied for this purpose today. Still, such AIMD simulations require very large computational resources and, thus, they are often carried out in supercomputing facilities.

In this article, we present AIMD simulations of two ILs: 1-butyl-3-methylimidazolium trifluoromethanesulfonate ([Bmim][OTf]) and 1-butyl-1-methylpyrrolidinium trifluoromethanesulfonate ([Bmpyr][OTf]). After a description of the computational setup in Sec. 2, we analyse the trajectories in Sec. 3. In particular, we compare these two ILs regarding the formation of hydrogen bonds by investigating radial distribution functions (RDFs) and combined distribution functions (CDFs). Furthermore, we apply our recently developed domain analysis¹² to study the microheterogeneity, which is a known property of ILs due to the segregation of polar and nonpolar phases on a microscopic level¹³.

2 Computational Details

For both ILs, 32 ion pairs were simulated under periodic boundary conditions (see Fig. 1 for representative snapshots of the simulations cells). The AIMD was performed using

	[Bmim][OTf]	[Bmpyrr][OTf]
Composition	32 [Bmim][OTf]	32 [Bmpyrr][OTf]
Cell size (pm)	2281.4	2313.5
Density (g/cm ³)	1.29	1.25
Equilibration time (ps)	5.0	5.0
Physical time (ps)	77.0	74.0

Table 1. Simulation parameters.

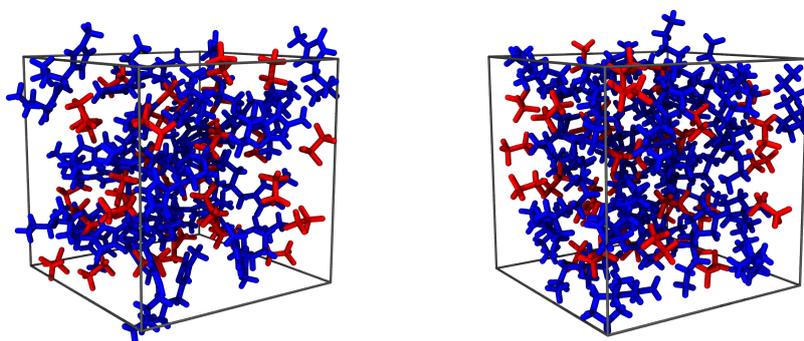


Figure 1. Representative snapshots of the simulation cells; left: [Bmim][OTf], right: [Bmpyrr][OTf], blue: cations, red: anions.

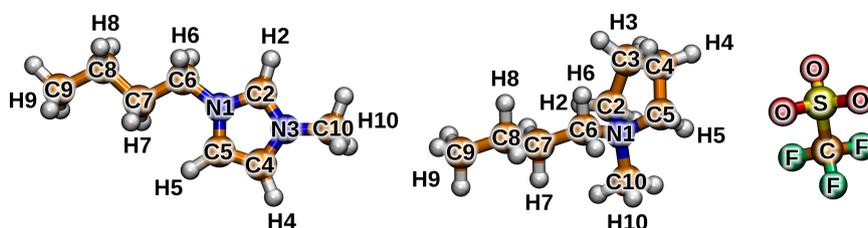


Figure 2. Atom numbering of [Bmim]⁺, [Bmpyrr]⁺, and [OTf]⁻ used throughout this article.

the CP2K program package^{14,15}. DFT was employed as electronic structure method, utilising the BLYP exchange-correlation functional^{16,17} with Grimme's dispersion correction D3¹⁸. The molecularly optimised double-zeta basis set MOLOPT-DZVP-SR-GTH¹⁹ was applied to all atoms together with the corresponding Goedecker–Teter–Hutter pseudopotentials^{20–22} and a plane wave cutoff of 280 Ry. A timestep of 0.5 fs was chosen, and the temperature was adjusted to 350 K by a Nosé–Hoover chain thermostat^{23–25}. The sizes of the cubic simulation cells, the resulting densities, and the simulation times are shown in Tab. 1. The trajectories were analysed using our trajectory analysis software package TRAVIS^{26,12,27}. The atom numbering to present the results is given in Fig. 2.

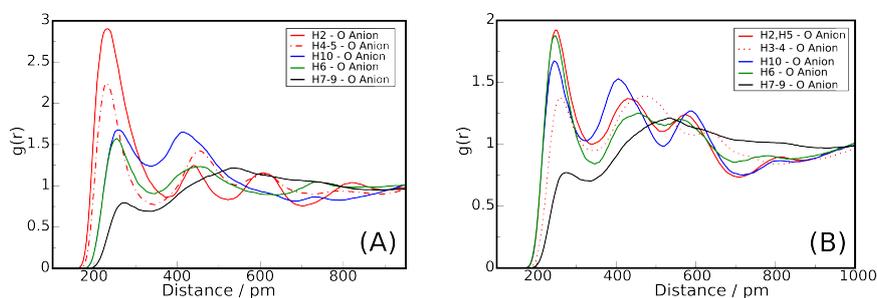


Figure 3. Radial distribution functions between hydrogen atoms of the cation and oxygen atoms of the anion for [Bmim][OTf] (A) and [Bmpyrr][OTf] (B).

The starting structures for the AIMD simulations were created by classical molecular dynamics carried out using LAMMPS²⁸. Force constants for bonds, angles, dihedral angles, and improper torsions, as well as Lennard-Jones parameters were taken from the general AMBER force field²⁹ and from the force field for imidazolium-based ILs by Liu *et al.*³⁰ Atomic partial charges for electrostatic interactions were derived from a restrained electrostatic potential fit³¹. The AIMD simulations were equilibrated by massive thermostating with a coupling time constant of 10 fs for about 2.5 ps and running with a single thermostat chain for additional 2.5 ps. During the second equilibration phase and in the production run, the thermostat coupling time constant was set to 100 fs.

3 Results and Discussion

On the basis of the RDFs, for the [Bmim][OTf] IL, the most intense interaction is coordinated through the ring hydrogen atoms (see Fig. 3 (A)). The most acidic one, placed between the two nitrogen atoms (H2), is the biggest, presenting an intensity of 2.9 for a H–O bond distance of 230.6 pm. This is followed by the other two hydrogen atoms of the ring (H4–5) which present a similar peak with an intensity of 2.2 at a H–O distance of 228.4 pm. The hydrogen atoms of the methyl residue (H10), as well as those from the first carbon atom of the butyl chain (H6) also show sizable peaks with intensities of 1.7 and 1.6, respectively, and longer H–O distances of 257.1 pm for H10 and 256.0 pm for H6. Thus, applying the criteria established by G. A. Jeffrey³², the two first hydrogen bonds from H2 and H4–5 could be considered of medium strength, with an interaction type mostly electrostatic, and the ones where H10 and H6 are the donors are weak hydrogen bonds with an electrostatic interaction type. When looking into the RDFs of [Bmpyrr][OTf], the most intense interaction is also observed through the ring hydrogen atoms close to the nitrogen atom (H2, H5) with a H–O distance of 248.8 pm (see Fig. 3 (B)), following, as above, the rule “Better acid – Better donor”³³, although its intensity of only 1.9 is one unit less than the one depicted by the H2 of [Bmim]⁺. The hydrogen atoms of the first position in the butyl chain (H6) present almost the same peak with an intensity of 1.9 and a distance of 246.3 pm for the hydrogen bond. Finally, the methyl hydrogen atoms (H10) (see Fig. 3 (B)) are better donors than the other ring hydrogen atoms (H3, H4) with a H–O distance 15 pm shorter (244 pm vs. 259 pm) and 0.3 more intense (1.7 vs. 1.4). The four interac-

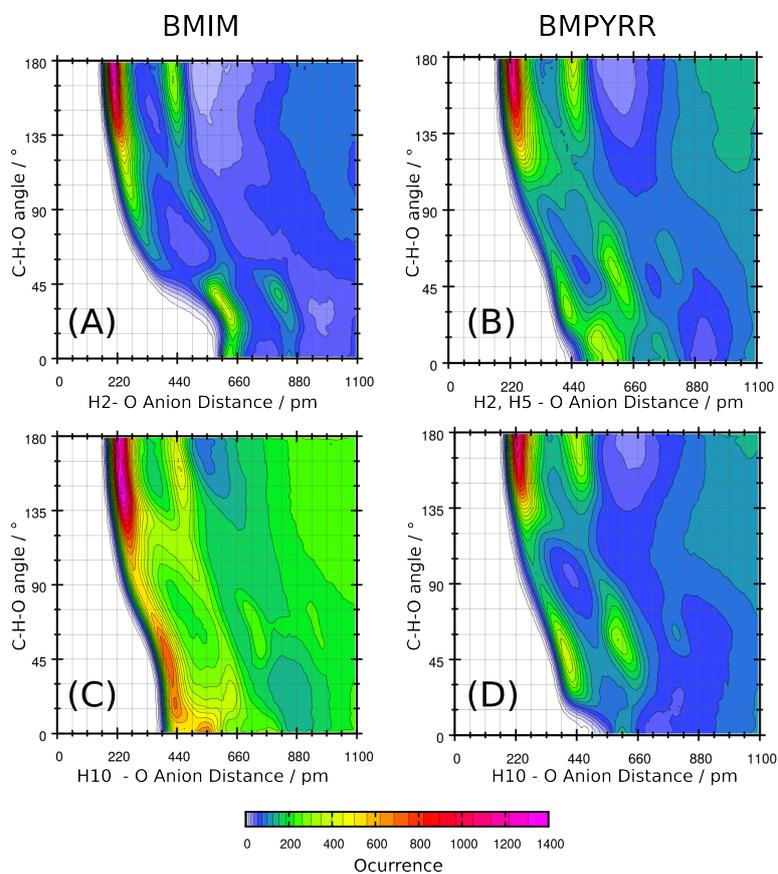


Figure 4. Combined distribution functions showing C–H–O angles against the corresponding H–O distances for [Bmim][OTf] (A, C) and [Bmpyrr][OTf] (B, D).

tions considered here are of an electrostatic nature, leading to weak hydrogen bonding³². The acidity of the hydrogen bond donor can be related to the pK_a of the donor molecule. It is known that electron-withdrawing groups enhance the donor strength, as it opens the possibility of a cooperation through double bonds (namely, π cooperativity) as is shown by *o*-nitrophenol or formamide, just to mention some³³. Having these two concepts in mind, it is easy to understand our results: in the [Bmpyrr]⁺ cation, which has no aromaticity, the hydrogen atoms H2, H5, H10, and H6 are two bonds away of the electronegative nitrogen atom while H3 and H4 are three bonds away. In the case of the [Bmim]⁺ cation, the ring aromaticity allows the π cooperativity, enhancing the hydrogen bond ability of the ring hydrogen atoms.

According to the CDFs, the oxygen atoms of the triflate anion are oriented almost 180° towards the most acidic ring hydrogen atoms, both in [Bmim][OTf] (H2) and in [Bmpyrr][OTf] (H2, H5) (see Fig. 4 (A, B)). In the CDF corresponding to the hydrogen bond between H2 in [Bmim]⁺ and the oxygen atoms in the anion (Fig. 4 (A)) we see

the highest occurrence at a distance of 215 pm and a linear angle corresponding with the oxygen pointing towards H2. This peak has a neighbour with a H2–O distance of 438 pm and a hydrogen bond angle of 180° too which corresponds to the other two oxygen atoms in the triflate anion. With a bond distance of 620 pm we encounter another peak showing C2–H2–O angles between 0° and 45°, these ones correspond to distances and angles between H2 and oxygen atoms that are coordinated to H4 and H5, almost coplanar with H2. At a distance of around 825 pm we can see the neighbour of this peak, with almost the same range of angles and less occurrence. H2 and H5 in [Bmpyr⁺] (Fig. 4 (B)) present a similar CDF but with less definite peaks. The one with the highest intensity can be seen at a H–O distance of 230 pm and an angle of 180°. The other two oxygen atoms of the triflate anion are reflected in the neighbour peak at 440 pm, which presents even less intensity. Another interaction with angles between 0° and 45° can also be observed in this graph with irregular peaks that blend together due to the variety of hydrogen atoms H3–4. These peaks, as seen above, clearly show the distances and angles of an oxygen atom coordinated to H3 or H4. The CDFs plotting the hydrogen bonds formed with H10 are shown in Fig. 4 (C, D). The peaks in [Bmim][OTf] are less distinct and this can be related to the flexibility of C10 and H10 and the weaker hydrogen bonding. The most intense peak appears at a H10–O distance of around 220 pm and with a C10–H10–O angle between 135° and 180°. The second peak in occurrence, which has almost the half of the first one, appears at around 430 pm in the H10–O distance and with an angle between 75° and 0° corresponding to the other hydrogen atoms of the same methyl group. In the [Bmpyr⁺] plot (D), the peaks are much more distinct. The highest peak is shown at a H10–O distance of around 220 pm and angles between 140° and 180° with a secondary peak due to the other two oxygen atoms with the same angle and at around 440 pm. The second peak in intensity can be seen at 430 pm with an angle between 90° and 15°, corresponding to the other hydrogen atoms of the same methyl group again.

The results of the domain analysis are shown in Tab. 2. In both ILs, the nonpolar domains are defined as the butyl chains without the methylene group connected to the ring, so they contain the carbon atoms C7, C8, and C9 with the attached hydrogen atoms, while the polar domains consist of all other atoms. Both ILs show a similar behaviour: the polar moieties form one single domain during the whole simulations, but the nonpolar side chains are sometimes split into several distinct domains, leading to average domain counts of 1.33 and 1.16, respectively. This indicates that the polar groups form an extended network that ranges through the whole simulation cells while the butyl chains tend to form aggregates in between. It has to be noted that a system size of 32 ion pairs is too small to fully observe this aggregation of the nonpolar side chains. The interpretation is supported by the neighbour matrices in Tab. 3 and Tab. 4, which show the average contact area and the average neighbour count between different groups. For this purpose, the polar domains are split into the anions and the head groups of the cations (“ring”). It is clearly apparent that the cations’ rings are primarily surrounded by anions and *vice versa*, supporting the picture that they form an extended network which is stabilised by hydrogen bonding. The self contacts of the side chains seem to be quite low, but it has to be kept in mind that a certain contact area to the rings and the anions the latter interacts with is simply enforced due to the covalent bond.

Domain	IL	n	$V / \text{\AA}^3$	$A / \text{\AA}^2$
Polar	[Bmim][OTf]	1.00	8813	3259
	[Bmpyrr][OTf]	1.00	9434	3155
Nonpolar	[Bmim][OTf]	1.33	2568	2734
	[Bmpyrr][OTf]	1.16	2718	2907

Table 2. Domain analysis in [Bmim][OTf] and [Bmpyrr][OTf]: average number of domains n , average domain volume V , and average domain surface area A .

	Anion	Ring	Nonpolar
Anion	18.6 (4.1)	89.8 (6.1)	48.2 (5.4)
Ring	89.8 (6.1)	31.5 (5.2)	53.7 (5.1)
Nonpolar	48.2 (5.4)	53.7 (5.1)	25.4 (3.2)

Table 3. Neighbour matrix in [Bmim][OTf]: average contact areas in \AA^2 , neighbour counts in parentheses.

	Anion	Ring	Nonpolar
Anion	18.3 (3.5)	92.7 (6.1)	43.7 (5.0)
Ring	92.7 (6.1)	40.5 (5.7)	54.9 (5.0)
Nonpolar	43.7 (5.0)	54.9 (5.0)	24.6 (3.1)

Table 4. Neighbour matrix in [Bmpyrr][OTf]: average contact areas in \AA^2 , neighbour counts in parentheses.

4 Conclusion

We have presented a study of the ILs [Bmim][OTf] and [Bmpyrr][OTf] in which we show how the cations' rings are surrounded by anions and *vice versa*. The strongest hydrogen bond is formed between H2 and O in [Bmim][OTf]. Every other polar hydrogen in the cation coordinates to the oxygen with less intensity and larger distances. Due to the π cooperativity the hydrogen bonds formed in [Bmim]⁺ are more intense than those formed in the analogous [Bmpyrr]⁺ IL which presents no aromaticity in its heterocyclic ring. The domain analysis indicates a similar degree of microheterogeneity in both ILs due to the aggregation of the butyl chains, but larger system sizes would be needed to observe this effect more clearly.

Acknowledgements

We gratefully acknowledge the allocation of computer time by the Jülich Supercomputing Centre. B.K. would like to thank the Deutsche Forschungsgemeinschaft (DFG) for support under the SPP 1708 project KI768/12-1.

References

1. T. Welton, *Room-Temperature Ionic Liquids. Solvents for Synthesis and Catalysis*, Chem. Rev., **99**, no. 8, 2071–2084, 1999.
2. N. V. Plechkova and K. R. Seddon, *Applications of ionic liquids in the chemical industry*, Chem. Soc. Rev., **37**, 123–150, 2008.
3. J. P. Hallett and T. Welton, *Room-Temperature Ionic Liquids: Solvents for Synthesis and Catalysis. 2*, Chem. Rev., **111**, no. 5, 3508–3576, 2011.
4. S. Tsuzuki, H. Tokuda, and M. Mikami, *Theoretical analysis of the hydrogen bond of imidazolium C2-H with anions*, Phys. Chem. Chem. Phys., **9**, 4780–4784, 2007.
5. S. Zahn, F. Uhlig, J. Thar, C. Spickermann, and B. Kirchner, *Intermolecular Forces in an Ionic Liquid ([Mmim][Cl]) versus Those in a Typical Salt (NaCl)*, Angew. Chem. Int. Ed., **47**, no. 19, 3639–3641, 2008.
6. S. Grimme, W. Hujo, and B. Kirchner, *Performance of dispersion-corrected density functional theory for the interactions in ionic liquids*, Phys. Chem. Chem. Phys., **14**, 4875–4883, 2012.
7. S. B. C. Lehmann, M. Roatsch, M. Schöppke, and B. Kirchner, *On the physical origin of the cation-anion intermediate bond in ionic liquids Part I. Placing a (weak) hydrogen bond between two charges*, Phys. Chem. Chem. Phys., **12**, 7473–7486, 2010.
8. M. Brehm, H. Weber, A. S. Pensado, A. Stark, and B. Kirchner, *Proton transfer and polarity changes in ionic liquid-water mixtures: a perspective on hydrogen bonds from ab initio molecular dynamics at the example of 1-ethyl-3-methylimidazolium acetate-water mixtures-Part I*, Phys. Chem. Chem. Phys., **14**, 5030–5044, 2012.
9. B. Kirchner, O. Hollóczki, J. N. Canongia Lopes, and A. A. H. Pádua, *Multiresolution calculation of ionic liquids*, WIREs Comput. Mol. Sci., **5**, no. 2, 202–214, 2015.
10. D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, 2009.
11. B. Kirchner, P. J. Dio, and J. Hutter, “Real-world predictions from ab initio molecular dynamics simulations”, in: *Multiscale Molecular Methods in Applied Chemistry*, Barbara Kirchner and Jadran Vrabec, (Eds.), vol. 307 of *Top. Curr. Chem.*, pp. 109–153. Springer Berlin Heidelberg, 2012.
12. M. Brehm, H. Weber, M. Thomas, O. Hollóczki, and B. Kirchner, *Domain Analysis in Nanostructured Liquids: A Post-Molecular Dynamics Study at the Example of Ionic Liquids*, ChemPhysChem, 2015.
13. R. Hayes, G. G. Warr, and R. Atkin, *Structure and Nanostructure in Ionic Liquids*, Chem. Rev., **115**, no. 13, 6357–6426, 2015.
14. J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *QUICKSTEP: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach*, Comput. Phys. Commun., **167**, no. 2, 103–128, 2005.
15. J. Hutter, M. Iannuzzi, F. Schiffrmann, and J. VandeVondele, *CP2K: atomistic simulations of condensed matter systems*, WIREs Comput. Mol. Sci., **4**, no. 1, 15–25, 2014.
16. A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A, **38**, 3098–3100, Sep 1988.

17. C. Lee, W. Yang, and R. G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B, **37**, 785–789, Jan 1988.
18. S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu*, J. Chem. Phys., **132**, no. 15, 154104, 2010.
19. J. VandeVondele and J. Hutter, *Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases*, J. Chem. Phys., **127**, no. 11, 114105, 2007.
20. S. Goedecker, M. Teter, and J. Hutter, *Separable dual-space Gaussian pseudopotentials*, Phys. Rev. B, **54**, 1703–1710, Jul 1996.
21. C. Hartwigsen, S. Goedecker, and J. Hutter, *Relativistic separable dual-space Gaussian pseudopotentials from H to Rn*, Phys. Rev. B, **58**, 3641–3662, Aug 1998.
22. M. Krack, *Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals*, Theor. Chem. Acc., **114**, 145–152, 2005.
23. S. Nosé, *A unified formulation of the constant temperature molecular dynamics methods*, J. Chem. Phys., **81**, no. 1, 511–519, 1984.
24. S. Nosé, *A molecular dynamics method for simulations in the canonical ensemble*, Mol. Phys., **52**, no. 2, 255–268, 1984.
25. G. J. Martyna, M. L. Klein, and M. Tuckerman, *Nosé–Hoover chains: The canonical ensemble via continuous dynamics*, J. Chem. Phys., **97**, no. 4, 2635–2643, 1992.
26. M. Brehm and B. Kirchner, *TRAVIS – A Free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories*, J. Chem. Inf. Model., **51**, no. 8, 2007–2023, 2011.
27. “TRAVIS – Trajectory Analyzer and Visualizer”,
<http://www.travis-analyzer.de>.
28. S. Plimpton, *Fast Parallel Algorithms for Short-Range Molecular Dynamics*, J. Comput. Phys., **117**, no. 1, 1–19, 1995.
29. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, *Development and testing of a general Amber force field*, J. Comput. Chem., **25**, no. 9, 1157–1174, 2004.
30. Z. Liu, S. Huang, and W. Wang, *A Refined Force Field for Molecular Simulation of Imidazolium-Based Ionic Liquids*, J. Phys. Chem. B, **108**, no. 34, 12978–12989, 2004.
31. U. C. Singh and P. A. Kollman, *An approach to computing electrostatic charges for molecules*, J. Comput. Chem., **5**, no. 2, 129–145, 1984.
32. G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, 1997.
33. P. Gilli, L. Pretto, V. Bertolasi, and G. Gilli, *Predicting Hydrogen-Bond Strengths from Acid–Base Molecular Properties*, Acc. Chem. Res., **42**, no. 1, 33–44, 2009.

Prebiotic Chemistry in Nanoconfinement

Daniel Muñoz-Santiburcio and Dominik Marx

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

E-mail: {daniel.munoz, dominik.marx}@theochem.rub.de

Water presents extremely different properties depending not only on conditions like temperature and pressure but also on the local environment (bulk vs. interfacial water), which in turn profoundly affects chemical reactions in solution. Here, we present the results of extensive *ab initio* molecular dynamics and metadynamics simulations in which we study some properties of nanoconfined water in mackinawite and a whole pathway for prebiotic peptide synthesis *via* glycine and COS condensation. It is shown that the special chemistry observed in this system is the result of the interplay between the steric and entropic effects intrinsic to nanoconfinement and the different dielectric properties of nanoconfined water compared to the bulk. The results show that prebiotic peptide bond formation is enhanced in nanoconfined water, while peptide hydrolysis is hindered, thus making this environment highly suitable for prebiotic polypeptide synthesis.

1 Introduction

The emergence of life is of special interest for mankind. Even though great advances in science and technology have been achieved, the answer to this fundamental question remains to be found. Various ideas have been developed and innovative experiments have been performed to explain how organic molecules might have been synthesised in a prebiotic “inorganic” environment, and how they organised themselves to form biomolecules of contemporary complexity^{1,2}. A two-dimensional scenario like the mineral surface is now considered to be a possible prebiotic stage where such reactions had taken place^{3,4}. One of the most acclaimed hypotheses in this respect is that suggested by Wächtershäuser^{5,6}.

Wächtershäuser proposes a chemo-autotrophic origin of life^{5,6}, which has been called the “Iron–Sulfur–World” (ISW) scenario. The main assumption in the ISW scenario is that prebiotic reactions like condensations of amino acids are catalyzed by iron–sulfur minerals located near deep sea hydrothermal vents^{7,8}, i.e. in a high temperature and high pressure environment. Several experimental studies provide support for a possible ISW scenario^{9–12} as well as offer mechanistic insight into the peptide polymerisation¹³. Up to now, our computations on JUQUEEN (and previously on JUGENE and even on JUBL, being the first Blue Gene platform in Germany)¹⁴ have identified very important mechanistic and energetic details of peptide synthesis reactions and thus the underlying “Iron–Sulfur–World” origin of life scenario. We have already established that extreme thermodynamic conditions have a dramatic impact on both mechanisms and energetics for prebiotic peptide synthesis^{15–17}, due to the change of the dielectric properties of water upon rising temperature and pressure, which makes it essentially a different solvent when compared to water at ambient conditions. Now, we determine the impact of nanoconfinement on chemical reactions in solution, which is of utmost importance for prebiotic chemistry. It has been found that an important iron/sulfur precipitate, namely mackinawite, can be produced at hydrothermal vents under ISW conditions^{7,18}. This mineral presents a layered structure, and its loosely bonded FeS sheets can be intercalated by water thus obtaining a nanocon-

finned water lamella between FeS plates. Taking into account that the dielectric properties of interfacial water are remarkably different from those of bulk water¹⁹, nanoconfined water (being purely interfacial water) reveals itself as a new and different medium for chemical synthesis as we will show in the following.

2 Nanoconfined Water in Mackinawite

We designed a model system consisting of two Fe₃₂S₃₂ mackinawite layers that are intercalated by either 32 or 49 H₂O molecules, being these our setups for “extreme” and “moderately” nanoconfined water (Fig. 1), respectively.

Our *ab initio* simulations of nanoconfined water²⁰ showed that, while extremely nanoconfined water in mackinawite is “arrested water” and the diffusivity is too low to consider it a suitable medium for chemical synthesis, moderately nanoconfined water presents liquid-like properties and is thus a good medium for prebiotic peptide synthesis. Further *ab initio* simulations in which we studied the excess proton in nanoconfined water²¹ revealed most efficient proton transfer. As in bulk water, the excess proton can be considered as a fluxional charge defect even at extreme confinement, being fully incorporated into the hydrogen bond network as a topological defect. The transfer of this excess proton along the hydrogen bond network occurs through the well-known Grotthuss mechanism (Fig. 2) and it is an almost barrierless process, which implies most efficient proton migration even at extreme nanoconfinement conditions.

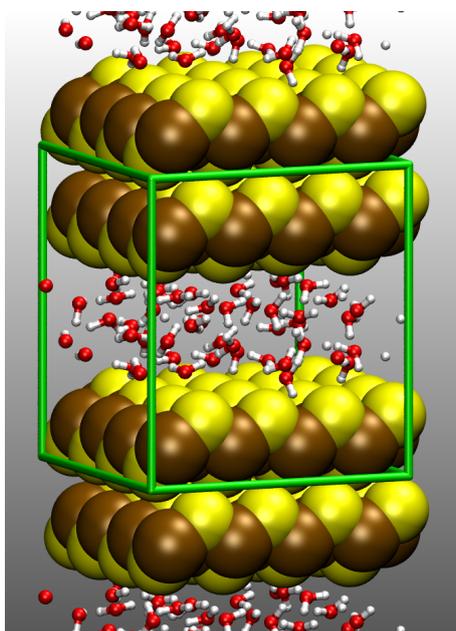


Figure 1. Model system for “moderately nanoconfined” water in mackinawite. Fe atoms are shown as brown spheres, S as yellow spheres, and H₂O molecules as balls-and-sticks. The edges of the simulation box are shown in green, and the periodic images along the *z* axis are also displayed.

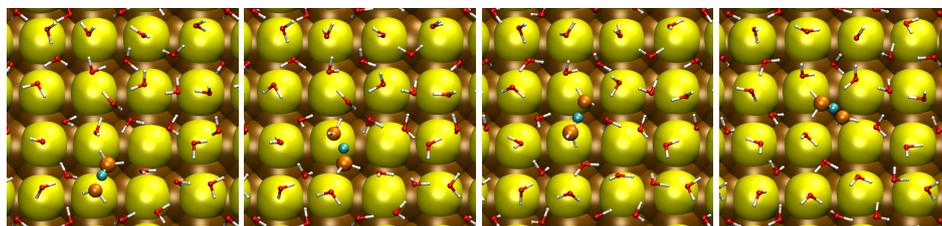


Figure 2. Transfer of an excess proton in “extremely nanoconfined” water by the Grotthuss mechanism. The excess proton is depicted in cyan and the oxygen atoms between which the transfer takes place are depicted in orange.

Our findings not only suggest that, indeed, thin water films confined by such mineral sheets are efficient nanoreactors for prebiotic reactions in Nature, but they might also lead to the design of improved setups for fuel cell membranes.

3 Peptide Synthesis in Nanoconfined Water

In previous stages of the long-term project “*In Silico* Exploration of Possible Routes to Prebiotic Peptide Synthesis by *Ab Initio* Metadynamics”, we simulated full peptide synthesis *via* N-carboxyanhydride (NCA) activation at three different thermodynamic conditions (see Fig. 3): ambient bulk water at 300 K and 0.1 MPa (ABW), hot-pressurised bulk water at 500 K and 20 MPa (HPW), and water at an iron-sulfur mineral interface at 500 K

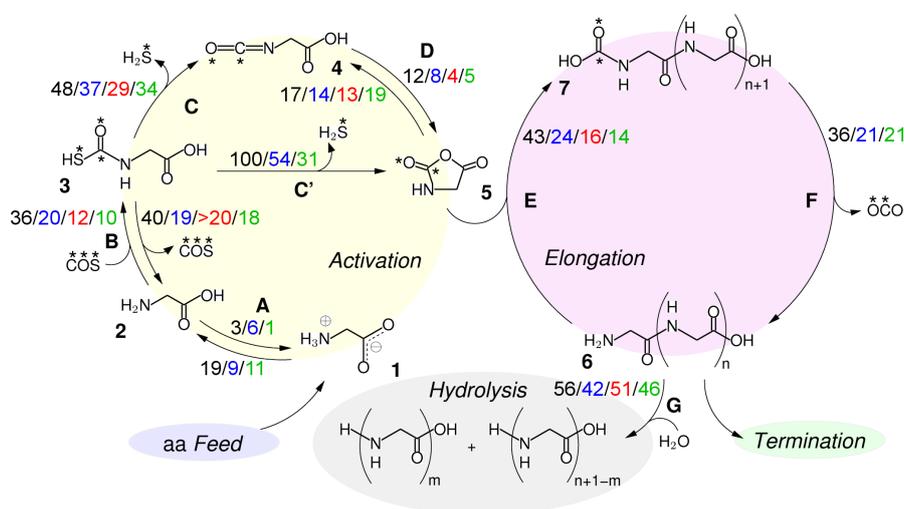


Figure 3. The peptide synthesis route comprising activation, elongation and hydrolysis. The calculated free energy barriers for individual steps of the mechanism are colour-coded; black: ambient bulk water, blue: hot-pressurised bulk water, red: pyrite-water interface at extreme conditions, green: nanoconfined hot-pressurised water in mackinawite; energies are given in $k_B T$ units.

and 20 MPa. Based on free energy calculations and dissecting the detailed reaction mechanisms, our previous studies have concluded the importance of extreme conditions and mineral surfaces for the primordial peptide synthesis^{15–17,22}. Moreover, these calculations have identified activation pathways of an amino acid.

The situation of a growing pyrite surface was simulated by surface defects at the interface to the bulk water. Hydrothermally-formed iron sulphide chimneys that are nowadays found at volcanic vents consist of mostly pyrite (FeS₂). However, it is assumed that pyrite is a reaction product in the sense that it precipitated as the monosulphide (FeS) mackinawite^{7,18}. As exposed before, being a layered structure it is conceivable that water can enter easily at high pressure and temperature conditions thus producing water/FeS interfaces in a local “lamella geometry” resulting into confinement. Incidentally, there are other layered minerals capable of hosting such water lamellae in their interlayer space, some of which like fougèrite (or “green rust”) may be of paramount importance for prebiotic chemistry as they have been proposed as part of a primordial pyro-phosphate synthetase nanoengine²³. Taking up these thoughts, we carried out the investigation of the prebiotic peptide cycle in nanoconfined hot–pressurised water (500 K and 20 MPa, NCW). Our simulations reveal remarkable differences in energetics and mechanisms compared to bulk water. These can be traced back to a unique combination of factors, namely the different dielectric properties of interfacial water w.r.t. the bulk¹⁹ as well as entropic and steric factors intrinsic to nanoconfinement which make nanoconfined water a whole new medium for chemical synthesis.

The way in which extreme conditions influences the reactions comprising the peptide cycle had been successfully studied by us in previous periods of this project^{15–17}. It is well established that the dielectric constant of water is remarkably reduced²⁴ when raising temperature and pressure from ambient to extreme conditions. This reduction is by more than a factor of two in our previously studied cases (from $\epsilon \approx 80$ at AMB to $\epsilon \approx 32$ in HPW)^{24,25}, thus essentially making HPW another solvent than at ambient conditions. In consequence, charged species are much favoured in ambient conditions, while extreme conditions discourage the formation of these and stabilise the neutral form of reactants and intermediates. In stark contrast to what is observed in the limit of bulk solvation, our recent results show that charged species are again stabilised in NCW. This can be explained with an increase in the dielectric constant of water in nanoconfinement w.r.t. the bulk. It has been shown that the parallel component of the dielectric tensor in interfacial water is higher in the interfacial region than in the bulk-like region.^{26,19} This higher dielectric constant of interfacial compared to bulk water is responsible for the lower relative free energy of charged species and transition states in NCW compared to HPW. This is clearly realised in the zwitterionic equilibrium of glycine (reaction **A** in Fig. 3): while the neutral form was stabilised in HPW compared to AMB conditions, in NCW the equilibrium is again displaced towards the zwitterionic form (Fig. 4). In some other reactions of the peptide cycle, this stabilisation of charged species results in changes of the reaction mechanism: from concerted in HPW to stepwise in NCW. Such is the case in reaction steps **C'** (Fig. 5) or **D** (Fig. 6), which in NCW proceed *via* the charged intermediates **3.2** and **4.1**, respectively, while these intermediates are only transition states in HPW.

In addition to this change in the dielectric properties of water, other factors determine the chemistry observed in the nanoconfined water lamella. These are mainly steric and entropic in nature: while addition (like **B** or **E**) or cyclisation reactions (like **C'** or **D**) are

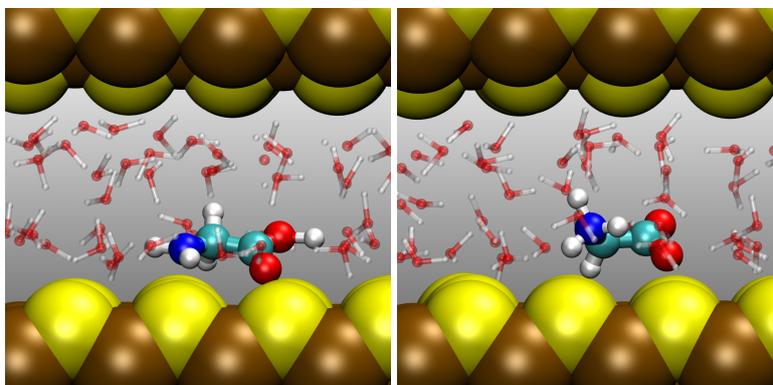


Figure 4. Snapshots for the reaction step A (protonation equilibrium of glycine). In nanoconfined hot–pressurised water, the neutral form of glycine (left) is much less stable than the zwitterionic form (right), just like in bulk, ambient water; while in bulk hot–pressurised water the neutral form is stabilised compared to those two conditions.

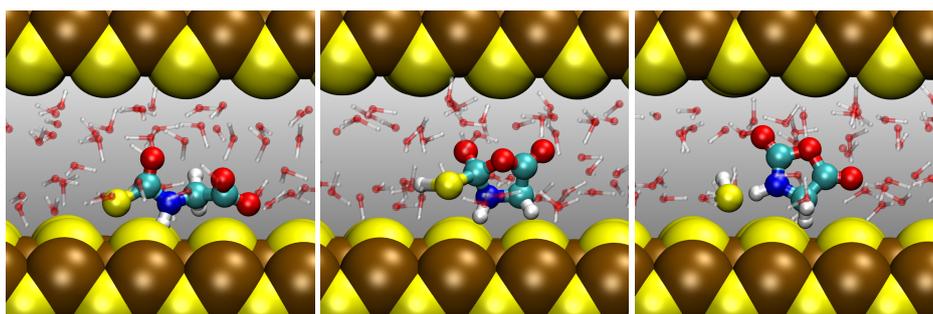


Figure 5. Snapshots for the reaction step C. From left to right: Carbamate **3**, charged intermediate **3.2** and SH^- + isocyanate **4**.

quite facilitated compared to HPW, elimination reactions are not quite affected. For addition reactions, the reduction in accessible space intrinsic to nanoconfinement limits the diffusive modes, forcing the reactants to have more reactive encounters; while for cyclisations it discourages stretched conformations and favours compact foldings which results in intramolecular bond formation. On the contrary, elimination reactions do not benefit from this reduced size of the reaction chamber, thus being less affected by nanoconfinement.

In the special case of the peptide hydrolysis reaction, the peptide bond is protected against water attack thanks to the sterical hindrance provided by the confining mackinawite sheets and the side chains of the dipeptide. This renders peptide hydrolysis unfavoured in NCW when compared to AMB or HPW conditions, while peptide elongation is not penalised as it involves the reaction of the terminal amino group of the dipeptide with a compact NCA molecule, which should not suffer as much sterical hindrance as the water attack on the less accessible peptide bond. Because of this, polypeptide formation should be facilitated in nanoconfinement in comparison with bulk water. The implications of this

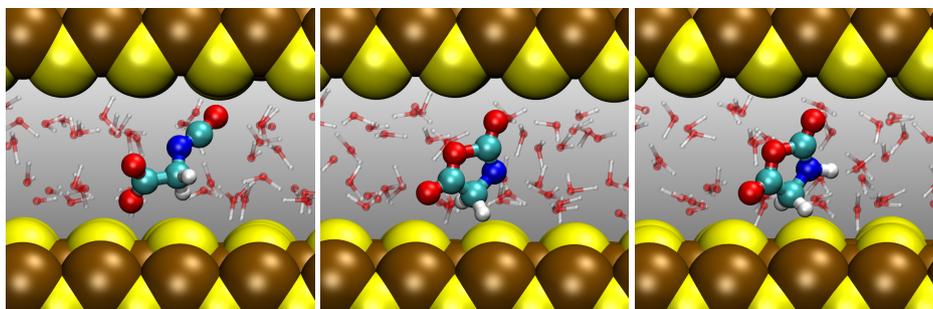


Figure 6. Snapshots for the reaction step **D**. From left to right: Isocyanate **4**, charged intermediate **4.1** and N-carboxyanhydride **5**.

result can be quite deep as it supports some hypotheses: it has been proposed that layered minerals like mackinawite could have acted as inorganic templates for the formation of peptide α -sheets or amyloids, which after incorporating metal atoms or inorganic clusters could have become protoenzymes²³.

4 Outlook

As we have exposed, the subtle interplay between the peculiar properties of nanoconfined water and the steric and entropic factors intrinsic to nanoconfinement determines the very special chemistry observed in such systems. From this starting point, new paths are now open for future lines of research: first, the nature of the reactants may have a great influence on the reactions, since for instance the hydrophobic/hydrophilic character of the aminoacid residues will impose different arrangements of the solvating waters, which is very likely to have a deep impact on both mechanisms and energetics. Second, these findings can be generalised to other systems capable of hosting such nanoconfined water lamellae like green rusts²³, and it would be of great interest to observe how the specific confining surface affects the reactions taking place therein. Finally, nanoconfined water is ubiquitous in biological systems, as it can be found in the crowded regions inside cells or in water pockets and tunnels in biological macromolecules²⁷. Hence, we expect that the study of the distinct properties of nanoconfined water will result not only in unravelling several yet poorly understood phenomena, but also in new and exciting applications for chemistry, biology and materials science.

Acknowledgements

This work has been supported by the German Research Foundation (DFG) *via* the Cluster of Excellence EXC 1069 “RESOLV”. We also gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time for a GCS Large Scale Project on the IBM Blue Gene/Q system JUQUEEN at Jülich Supercomputing Centre (JSC).

References

1. I. Fry, *The Emergence of Life on Earth: A Historical and Scientific Overview*, Rutgers University Press, New Jersey, U.S.A., 2000.
2. P. L. Luisi, *The Emergence of Life from Chemical Origins to Synthetic Biology*, Cambridge University Press, Cambridge, UK, 2006.
3. M. Schoonen, A. Smirnov, and C. Cohn, *A Perspective on the Role of Minerals in Prebiotic Synthesis*, *Ambio*, **33**, 539, 2004.
4. S. Leach, I. W. M. Smith, and C. S. Cockell, *Introduction: conditions for the emergence of life on the early Earth (Royal Society Discussion Meeting Issue)*, *Phil. Trans. R. Soc. B*, **361**, 1675–1679, 2006.
5. G. Wächtershäuser, *Before enzymes and templates: theory of surface metabolism*, *Microbiol. Rev.*, **52**, 452, 1988.
6. G. Wächtershäuser, *Groundworks for an evolutionary biochemistry: the iron-sulphur world*, *Prog. Biophys. Mol. Biol.*, **58**, 85–201, 1992.
7. W. Martin and M. J. Russell, *On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes*, *Phil. Trans. R. Soc. Lond. B.*, **358**, 59–85, 2003.
8. S. C. Singh, W. C. Crawford, H. Carton, T. Seher, V. Combier, M. Cannatand, J. P. Canales, D. Düsünür, J. Escartin, and J. M. Miranda, *Discovery of a magma chamber and faults beneath a Mid-Atlantic Ridge hydrothermal field*, *Nature*, **442**, 1029–1032, 2006.
9. C. Huber and G. Wächtershäuser, *Activated Acetic Acid by Carbon Fixation on (Fe,Ni)S Under Primordial Conditions*, *Science*, **276**, 245, 1997.
10. C. Huber and G. Wächtershäuser, *Peptides by Activation of Amino Acids with CO on (Ni,Fe)S Surfaces: Implications for the Origin of Life*, *Science*, **281**, 670–672, 1998.
11. C. Huber, W. Eisenreich, S. Hecht, and G. Wächtershäuser, *A Possible Primordial Peptide Cycle*, *Science*, **301**, 938–940, 2003.
12. L. Leman, L. Orgel, and M. R. Ghadiri, *Carbonyl Sulfide-Mediated Prebiotic Formation of Peptides*, *Science*, **306**, 283–286, 2004.
13. H. R. Kricheldorf, *Polypeptides and 100 Years of Chemistry of α -Amino Acid N-Carboxyanhydrides*, *Angew. Chem. Int. Ed.*, **45**, 5752–5784, 2006.
14. N. N. Nair, E. Schreiner, and D. Marx, *Prebiotic Peptide Synthesis on Blue Gene Platforms at “Iron-Sulfur-World” Conditions*, *inSiDE*, **6**, 30–35, 2008, See: http://inside.hlr.de/_old/htm/Edition.02.08/article_09.html (accessed on August 25, 2015).
15. E. Schreiner, N. N. Nair, and D. Marx, *Influence of extreme thermodynamic conditions and pyrite surfaces on peptide synthesis in aqueous media*, *J. Am. Chem. Soc.*, **130**, no. 9, 2768–2770, 2008.
16. N. N. Nair, E. Schreiner, and D. Marx, *Peptide synthesis in aqueous environments: The role of extreme conditions on amino acid activation*, *J. Am. Chem. Soc.*, **130**, no. 43, 14148–14160, 2008.
17. E. Schreiner, N. N. Nair, and D. Marx, *Peptide Synthesis in Aqueous Environments: The Role of Extreme Conditions on Peptide Bond Formation and Peptide Hydrolysis*, *J. Am. Chem. Soc.*, **131**, no. 38, 13668–13675, 2009.

18. M. J. Russell and A. J. Hall, *The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front*, *J. Geolog. Soc.*, **154**, 377–402, 1997.
19. D. J. Bonthuis, S. Gekle, and R. R. Netz, *Dielectric profile of interfacial water and its effect on double-layer capacitance*, *Phys. Rev. Lett.*, **107**, no. 16, 166102, 2011.
20. C. Wittekindt and D. Marx, *Water confined between sheets of mackinawite FeS minerals*, *J. Chem. Phys.*, **137**, no. 5, 054710, 2012.
21. D. Muñoz-Santiburcio, C. Wittekindt, and D. Marx, *Nanoconfinement effects on hydrated excess protons in layered materials*, *Nat. Commun.*, **4**, 2349, Aug 2013.
22. E. Schreiner, N. N. Nair, C. Wittekindt, and D. Marx, *Peptide synthesis in aqueous environments: the role of extreme conditions and pyrite mineral surfaces on formation and hydrolysis of peptides*, *J. Am. Chem. Soc.*, **133**, no. 21, 8216–8226, 2011.
23. M. J. Russell, L. M. Barge, R. Bhartia, D. Bocanegra, P. J. Bracher, E. Branscomb, R. Kidd, S. McGlynn, D. H. Meier, W. Nitschke, et al., *The drive to life on wet and icy worlds*, *Astrobiology*, **14**, no. 4, 308–343, 2014.
24. H. Weingärtner and E. U. Franck, *Supercritical water as a solvent*, *Angew. Chem., Int. Ed.*, **44**, no. 18, 2672–2692, 2005.
25. M. Uematsu and E. U. Frank, *Static dielectric constant of water and steam*, *J. Phys. Chem. Ref. Data*, **9**, no. 4, 1291–1306, 1980.
26. V. Ballenegger and J.-P. Hansen, *Dielectric permittivity profiles of confined polar fluids*, *J. Chem. Phys.*, **122**, no. 11, 114711, 2005.
27. D. Lucent, C. D. Snow, C. Aitken Echeverría, and V. S. Pande, *Non-bulk-like solvent behavior in the ribosome exit tunnel*, *PLoS Comput. Biol.*, **6**, no. 10, e1000963, 2010.

Simulation of Electron Transfer and Electron Transport in Molecular Systems at Surfaces

**Pedro B. Coto, Chriszandro Hofmeister, Veronika Prucker,
Dominik Weckbecker, and Michael Thoss**

Institut für Theoretische Physik and Interdisziplinäres Zentrum für Molekulare Materialien,
Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstraße 7/B2, 91058 Erlangen, Germany
E-mail: michael.thoss@physik.uni-erlangen.de

We have investigated electron transfer and transport processes in several molecular systems adsorbed at metal surfaces using a methodology that combines first-principles electronic structure methods with quantum dynamics and transport approaches. Specifically, we have analysed the molecular factors that control electron transfer in a series of nitrile-substituted alkanethiolate self-assembled monolayer models adsorbed at the Au(111) surface that differ in the size of the aliphatic spacer chain. In addition, we have analysed the possibility of using a proton transfer reaction triggered by an external electrostatic field as a novel mechanism for switching a molecular junction. To demonstrate the feasibility of the process, we have investigated electron transport in a junction containing a molecular bridge that can exist in two tautomeric forms, [2,5-(4-hydroxypyridine)] and 2,5-[4(1H)-pyridone], that exhibit very different conductance properties.

1 Introduction

The transfer or transport of electrons is a key step in many processes in physics, chemistry, biology, and technology¹. Examples range from intramolecular charge transfer in donor-acceptor complexes in solution² over ET in sensory proteins or photosynthesis³ to charge transport processes in single-molecule junctions⁴. From the point of view of the design of nanoelectronic devices or new materials with improved characteristics, understanding the fundamental mechanisms of ET at the molecular level is all important. In this respect, theory and simulation can assist the interpretation of experimental results and provide additional information on the underlying ET mechanisms that is not straightforwardly obtained from experiment.

In the following, we discuss the theoretical methods employed and the results obtained in the simulation of ET processes involving organic molecules adsorbed at metal surfaces. Specifically, we report the results obtained in the investigation of the mechanism of ET in a series of nitrile-substituted alkanethiolate self-assembled monolayer (SAM) models adsorbed at the Au(111) surface that have aliphatic spacer chains of different size. In addition, we detail the results obtained in the characterisation of the transport properties of a single molecule junction with an organic molecular bridge that can exist in two tautomeric forms, [2,5-(4-hydroxypyridine)] and 2,5-[4(1H)-pyridone], characterised by very different conductance properties and discuss the possibility of using this device as a molecular switch.

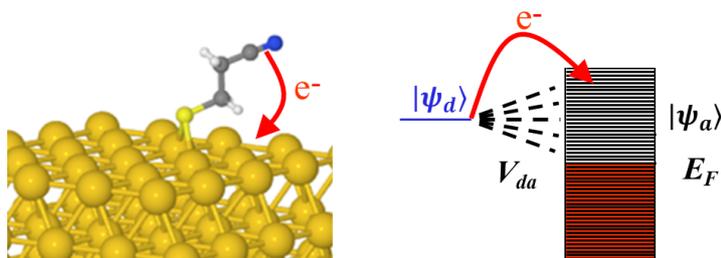


Figure 1. Schematic representation of the ET process. Left: Prototypical system investigated in this work consisting of an organic nitrile-substituted alkanethiolate molecule adsorbed at the Au(111) surface. Right: Energy level diagram for the ET process. The donor state, $|\psi_d\rangle$, is coupled to a quasi-continuum set of empty acceptor states, $|\psi_a\rangle$, through V_{da} , the donor-acceptor electronic coupling terms. E_F is the Fermi energy of the system.

where the diagonal blocks of H contain the energies of the i ($E_{d,i}$) and j ($E_{a,j}$) states of the donor and acceptor moieties, respectively. The eigenvectors of these blocks are identified with the localised states of the donor ($|\psi_d\rangle$) and acceptor ($|\psi_a\rangle$) fragments and the matrix elements of the off-diagonal blocks, V_{da} , are identified with the donor-acceptor electronic coupling terms. To incorporate the effects of the extended nature of the Au surface in the model ET Hamiltonian, we used a model inspired by the surface Greens function approach¹⁵. In this model, the self-energy (which accounts for the effects of the part of the infinite surface not included in the cluster models) is approximated by a complex absorbing potential (CAP)¹⁶ that is added to H and that has the form

$$V_{\text{CAP}}(R) = \begin{cases} -i\alpha(R - 5)^\delta & \text{if } R > 5 \text{ bohr} \\ 0 & \text{if } R \leq 5 \text{ bohr} \end{cases} \quad (2)$$

where R is the distance between the considered Au atom and the sulphur atom of the nitrile-substituted alkanethiolate moiety (see Fig. 1) and α and δ are constants with values 1.0×10^7 hartree/bohr⁴ and 4, respectively, that were obtained after testing.

The aforementioned theoretical protocol provides the necessary information to obtain the time evolution of the population of the donor state, $P_d(t)$ given by

$$P_d(t) = |\langle \psi_d | e^{-i\hat{H}t} | \psi_d \rangle|^2, \quad (3)$$

which we have used for the characterisation of the electron injection dynamics.

2.2 Application to Electron Transfer at Molecule-Metal Interfaces

The methodology detailed above has been applied to investigate the electron injection dynamics in a series of nitrile-substituted alkanethiolate SAM models (**C2**, **C3**, **C4**, and **C8**) having aliphatic spacer chains with an increasing number of methylene bridge groups ($-(\text{CH}_2)_2-$, $-(\text{CH}_2)_3-$, $-(\text{CH}_2)_4-$, and $-(\text{CH}_2)_8-$, respectively). We investigated the ET dynamics of the donor states π_1^* and π_2^* (see Fig. 2) corresponding to the two π^* resonances of the CN group whose electron injection dynamics has been studied using core-clock hole spectroscopy^{17,18}. The motivation underlying the selection of these systems was twofold. On the one hand, they provide a way to investigate the role that the symmetry of the donor state has in the ET process. On the other hand, the use of systems with

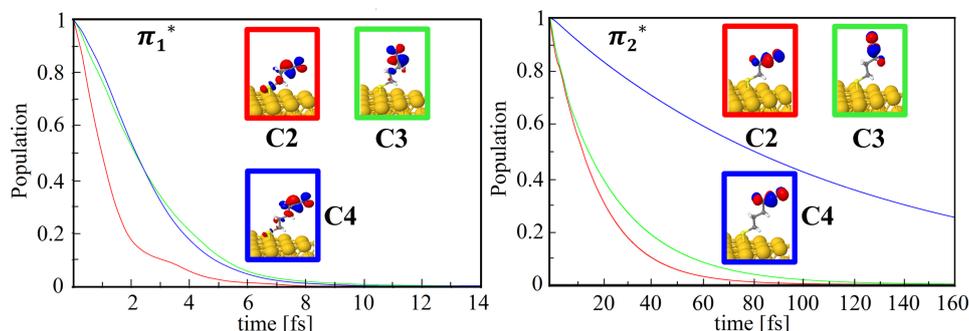


Figure 2. Population decay of π_1^* (left) and π_2^* (right) donor states for the nitrile-substituted alkanethiolate SAM models **C2** (red), **C3** (green), and **C4** (blue). The insets show the localised donor states $|\psi_d\rangle$.

aliphatic linkers of different size provides information about the roles that their electronic and chemical structures have in the ET mechanism.

The results obtained show that for all systems investigated, the most stable adsorption geometry was the bridge-fcc like position with the adsorbed molecule tilted against the surface normal with values ranging from 68.8° in **C4**, over 70.7° in **C2**, to 73.6° in **C8**. For **C3**, the only system having an odd number of methylene bridge groups, the tilting angle obtained was 10.3° (see Fig. 2). Regarding the electron injection dynamics, for all systems investigated the donor states π_1^* and π_2^* exhibit different ET times. In particular, electron injection times from the π_1^* donor state of **C2**, **C3**, and **C4** were found to be ultrafast, with characteristic times of $\tau_{1/2} = 1.0$ fs, $\tau_{1/2} = 2.1$ fs, and $\tau_{1/2} = 2.2$ fs, respectively. On the other hand, the electron injection times from the π_2^* donor state were found to be systematically slower [$\tau_{1/2} = 12.3$ fs (**C2**), $\tau_{1/2} = 15.1$ fs (**C3**), and $\tau_{1/2} = 80.5$ fs (**C4**)]. For **C8**, the electron injection time scales obtained are too long to be correctly described with the CAP method employed in the simulations. The results obtained can be rationalised on the basis of the symmetry of the donor state and the molecular structure of the spacers. Specifically, in all systems investigated the π_1^* donor states are more delocalised than the corresponding π_2^* donor states (see Fig. 2). This leads to larger molecule-substrate coupling V_{da} and therefore, faster electron injection dynamics, a result in qualitative agreement with the experimental data⁸. On the other hand, there is a systematic slowdown of the ET process with the increase in size of the aliphatic spacer chain in all systems investigated for both donor states. The reason for this behaviour is the increase of the distance between the donor and acceptor moieties, which leads to a decrease in the magnitude of V_{da} ¹⁹ and therefore to slower electron injection dynamics.

3 Electron Transport in Single Molecule Junctions

In this section, we discuss the results obtained in the simulation of ET in single molecule junctions. After a brief outline of the theoretical methods employed in the simulation of electron transport in molecular junctions, we report the results obtained in the characterisation of the transport properties of a junction with an organic molecular bridge that can exist in two tautomeric forms, [2,5-(4-hydroxypyridine)] and 2,5-[4(1H)-pyridone], and the possibility of using this device as a molecular switch.

3.1 Theoretical Methods

In the simulation of electron transport in single molecule junctions we have used the model Hamiltonian

$$H = \begin{pmatrix} H_L & H_{LM} & 0 \\ H_{ML} & H_M & H_{MR} \\ 0 & H_{RM} & H_R \end{pmatrix}, \quad (4)$$

that describes a single molecule junction consisting of a molecular bridge (M) and two leads (L and R) where the bridge is coupled to both leads and no direct coupling between the latter exists ($H_{RL} = H_{LR} = 0$). The parameters needed for the construction of this model Hamiltonian have been obtained employing a partitioning technique similar to that described in Sec. 2.1. Specifically, the overall system was divided into two parts. The molecule and part of the leads were described explicitly using first-principles electronic structure methods. For the systems investigated, we used DFT employing the B3LYP hybrid exchange-correlation functional²⁰⁻²², and a SV(P) basis set¹¹ (88 Au atoms were described using the ECP-60-MWB pseudopotential¹²). The remaining part of the leads was described using self-energies¹⁵. After adding the corresponding self-energies to the left and right leads the current-voltage characteristic can be obtained using the Landauer formula²³

$$I(V) = \frac{2e}{h} \int_{-\infty}^{\infty} T(E)[f_L(E) - f_R(E)]dE \quad (5)$$

where $f_L(E)$ and $f_R(E)$ are the Fermi distributions of the left and right leads, respectively, and $T(E)$ is the transmission function that can be expressed as²⁴

$$T(E) = \text{tr}_M[\Gamma_L(E)G_M^\dagger(E)\Gamma_R(E)G_M(E)] \quad (6)$$

where $G_M(E)$ is the molecular Green's function, $\Gamma_{L/R}(E)$ the broadening of the molecular states that results from the coupling to the continuum of states of the Au leads, and the trace is taken over the electronic states of the molecule (M)²⁵.

3.2 Switching the Conductance of a Molecular Junction using a Proton Transfer Reaction

A single molecule junction may be used as a nanoswitch if the molecular bridge has at least two stable states with different conductance that can be reversibly transformed into each other by means of a mechanism that can be externally controlled. We have analysed the possibility of using a proton transfer reaction triggered by an external electrostatic field as a novel mechanism for switching a molecular junction. As an example, we have investigated this possibility using single molecule junctions containing tris{2,5-[4(1H)-pyridone]} (T2PY) and tris[2,5-(4-hydroxypyridine)] (T2HP) as bridges (see Fig. 3). Both tautomers were functionalised for binding to the Au leads with 3,4-dimercapto-(1H)-pyridine and 3,4-dimercaptobenzenone, and 3,4-dimercaptopyridine and 3,4-dimercaptophenol, respectively. To demonstrate the feasibility of the process we have investigated the transport properties of these junctions. Fig. 3 depicts the current-voltage characteristics of T2PY and T2HP. The results show that T2PY exhibits a rather low current for the range of V investigated whereas T2HP exhibits larger current, in particular for bias voltages $|V| > 0.5$ V. This different behaviour originates from the differences in

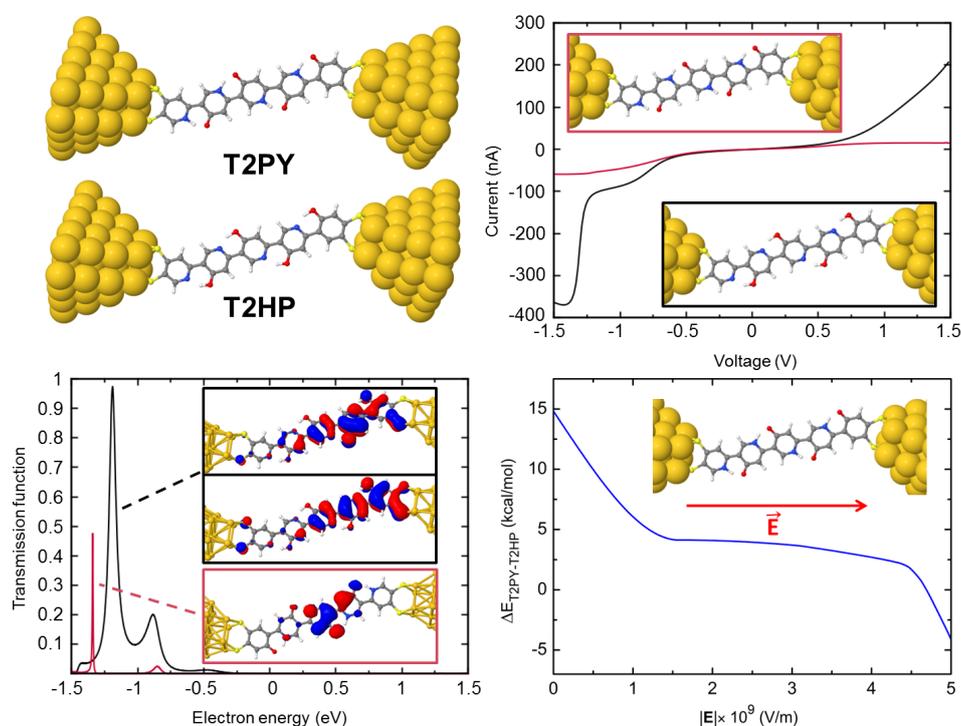


Figure 3. Top, left: Molecular junctions investigated in this work. Right: Current-voltage characteristics. The black and red lines correspond to T2HP and T2PY, respectively. Bottom, left: Transmission function for T2HP (black) and T2PY (red). The molecular orbitals associated to the most significant peaks of $T(E)$ are also shown. Right: Energy difference between T2PY and T2HP for different values of the applied external electrostatic field. The inset shows the direction of the field²⁶.

the electronic structures of both tautomers, specifically the degree of delocalisation of the π bond network²⁶. As shown in Fig. 3, the molecular orbitals associated with the most relevant structure of the transmission function of T2HP, the peak at $E = 1.25$ eV, are π -like orbitals that extend over the molecule while in the case of T2PY the most relevant feature, the peak at $E = -1.4$ eV, is associated with a π -like orbital that has most of its density localised in the centre of the molecule. Therefore, for a given bias voltage the two tautomers show different conductance and can be associated with the “on” (T2HP) and “off” (T2PY) states of a molecular switch. The second condition necessary a functional switch is the existence of a mechanism for the reversible transformation between the stable minima. In the case investigated the tautomers can be interconverted via a proton transfer. To explore this possibility, we have studied the relative stability of both T2HP and T2PY in the presence of a constant external electrostatic field oriented in a direction along the axis of the junction (see Fig. 3). The results obtained show that it is possible to obtain an inversion of the relative stabilities of T2HP and T2PY by increasing the strength of the external electrostatic field applied, therefore providing a proof of principle that a reversible interconversion between T2HP and T2PY using this mechanism is possible.

4 Concluding Remarks

We have described a theoretical methodology for the investigation of electron transfer and transport processes in molecular systems adsorbed at surfaces that combines first-principles electronic structure methods with quantum dynamics and electron transfer approaches. Using this methodology we have investigated the factors controlling the electron injection dynamics in a series of nitrile-substituted alkanethiolate SAM models adsorbed at the Au(111) surface. In addition, we have explored the possibility of using a proton transfer reaction triggered by an external electrostatic field as a novel mechanism for switching a molecular junction, providing a proof of principle of the feasibility of the process. The results obtained may pave the way for the design of new nanoelectronic devices.

Acknowledgements

This work has been supported by the German-Israeli Foundation for Scientific Development (GIF), the Deutsche Forschungsgemeinschaft (DFG) through the Cluster of Excellence “Engineering of Advanced Materials” (EAM), SFB 953, and a research grant. Generous allocation of computing time at the Jülich Supercomputing Centre (JSC) is greatly acknowledged. We thank Prof. A. L. Sobolewski for helpful discussions.

References

1. A. M. Kuznetsov and J. Ulstrup, *Electron Transfer in Chemistry and Biology: An Introduction to the Theory*, Wiley, Chichester, 1999.
2. R. A. Marcus, *Electron-Transfer Reactions in Chemistry: Theory and Experiment (Nobel lecture)*, *Angew. Chem., Int. Ed.*, **32**, 1111–1121, 1993.
3. M. E. Michel-Beyerle, *The Reaction Center of Photosynthetic Bacteria: Structure and Dynamics*, Springer, 1996.
4. J. C. Cuevas and E. Scheer, *Molecular Electronics: An Introduction to Theory and Experiment*, World Scientific Publishing Co. Pte. Ltd., Singapore, 2010.
5. J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, *Phys. Rev. Lett.*, **77**, 3865–3868, 1996.
6. P. E. Blöchl, *Projector Augmented-Wave Method*, *Phys. Rev. B*, **50**, 17953–17979, 1994.
7. “Viena Ab Initio Simulation Package (VASP), V5.2.11, 2006” available from <http://cms.mpi.univie.ac.at/cmsspage/main/>.
8. F. Blobner, P. B. Coto, F. Allegretti, M. Bockstedte, O. Rubio-Pons, H. Wang, D. L. Allara, M. Zharnikov, M. Thoss, and P. Feulner, *Orbital-Symmetry-Dependent Electron Transfer through Molecules Assembled on Metal Substrates*, *J. Phys. Chem. Lett.*, **3**, 436–440, 2012.
9. V. Prucker, O. Rubio-Pons, M. Bockstedte, H. Wang, P. B. Coto, and M. Thoss, *Dynamical Simulation of Electron Transfer Processes in Alkanethiolate Self-Assembled Monolayers at the Au(111) Surface*, *J. Phys. Chem. C*, **117**, 25334–25342, 2013.
10. “TURBOMOLE V6.3 2011”, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.

11. A. Schäfer, H. Horn, and R. Ahlrichs, *Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr*, J. Chem. Phys., **97**, no. 4, 2571–2577, 1992.
12. D. Andrae, U. Häußermann, M. Dolg, H. Stoll, and H. Preuß, *Energy-Adjusted Ab Initio Pseudopotentials for the Second and Third Row Transition Elements*, Theor. Chim. Acta, **77**, 123–141, 1990.
13. H. Feshbach, *Unified Theory of Nuclear Reactions II*, Ann. Phys., **19**, 287–313, 1962.
14. W. Domcke, *Theory of Resonance and Threshold Effects in Electron-Molecule Collisions: The Projection-Operator Approach*, Phys. Rep., **208**, 97–188, 1991.
15. Y. Xue and M. A. Ratner, *Microscopic Study of Electrical Transport through Individual Molecules with Metallic Contacts. I. Band Lineup, Voltage Drop, and High-Field Transport*, Phys. Rev. B, **68**, 115406, 2003.
16. U. V. Riss and H.-D. Meyer, *Calculation of Resonance Energies and Widths Using the Complex Absorbing Potential Method*, J. Phys. B: At., Mol. Opt. Phys. **26**, 4503–4535, 1993.
17. S. Neppl, U. Bauer, D. Menzel, P. Feulner, A. Shaporenko, M. Zharnikov, P. Kao, and D. L. Allara, *Charge Transfer Dynamics in Self-Assembled Monomolecular Films*, Chem. Phys. Lett., **447**, 227–231, 2007.
18. P. Kao, S. Neppl, P. Feulner, D. L. Allara, and M. Zharnikov, *Charge Transfer Time in Alkanethiolate Self-Assembled Monolayers via Resonant Auger Electron Spectroscopy*, J. Phys. Chem. C, **114**, 13766–13773, 2010.
19. M. N. Paddon-Row, *Orbital Interactions and Long-Range Electron Transfer*, Adv. Phys. Org. Chem., **38**, 1–85, 2003.
20. A. D. Becke, *Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior*, Phys. Rev. A, **38**, 3098–3100, 1988.
21. C. Lee, W. Yang, and R. G. Parr, *Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density*, Phys. Rev. B, **37**, 785–789, 1988.
22. A. D. Becke, *Density-Functional Thermochemistry. III. The Role of Exact Exchange*, J. Chem. Phys., **98**, 5648–5652, 1993.
23. D. S. Fisher and P. A. Lee, *Relation Between Conductivity and Transmission Matrix*, Phys. Rev. B, **23**, 6851–6854, 1981.
24. Y. Meir and N. S. Wingreen, *Landauer Formula for the Current Through an Interacting Electron Region*, Phys. Rev. Lett., **68**, 2512–2515, 1992.
25. C. Benesch, M. Čížek, J. Klimeš, I. Kondov, M. Thoss, and W. Domcke, *Vibronic Effects in Single Molecule Conductance: First-Principles Description and Application to Benzenealkanethiolates between Gold Electrodes*, J. Phys. Chem. C, **112**, 9880–9890, 2008.
26. C. Hofmeister, R. Härtle, O. Rubio-Pons, P. B. Coto, A. L. Sobolewski, and M. Thoss, *Switching the Conductance of a Molecular Junction using a Proton Transfer Reaction*, J. Mol. Model. **20**, 2163, 2014.

Elementary Particle Physics

Elementary Particle Physics

Gernot Münster

Institut für Theoretische Physik, Universität Münster, 48149 Münster, Germany
E-mail: munsteg@uni-muenster.de

The projects in the field of elementary particle physics are aimed at questions in fundamental physics, concerning the constituents of matter and their interactions. For example, one of these questions is about the mass difference between protons and neutrons. These particles are the constituents of atomic nuclei. The small difference between their masses plays an important role for the formation of nuclei after the big bang and for the abundance of elements in the universe. If the mass difference were only slightly different from its value in nature, the formation of stars and planets and the conditions for the possibility of life would drastically differ from the actual situation. The mass difference between protons and neutrons depends on fundamental constants of nature, in particular on the masses of quarks and on the strength of the electromagnetic forces. This dependence can be investigated in the framework of the theory of strong interactions of elementary particles, Quantum Chromodynamics (QCD), combined with Quantum Electrodynamics (QED). An *ab initio* calculation, avoiding approximations of unknown accuracy, requires numerical simulations of QCD plus QED, regularised on a space-time lattice. This has been achieved by Borsanyi *et al.* in their project. Careful studies of systematic and statistical errors were necessary to obtain reliable results on the delicate parameter dependence of the proton-neutron mass difference. Their results show that small variations of the parameters in the percentage range would lead to significant changes in the mass difference.

In general, the goals of numerical investigations in the area of lattice field theory have shifted in recent years. Whereas in earlier years issues concerning the foundations and basic properties of models and their numerical simulations have been in the focus, present projects concentrate on detailed questions of experimental and phenomenological relevance. Also, problems of nuclear physics, that previously were difficult to approach, have become feasible with present methods and computer resources. Direct *ab initio* calculations of many-nucleon systems in lattice QCD still exceed the possibilities. However, using an intermediate approach, in which effective field theories, including complicated interaction terms, are simulated on space-time lattices, it has become possible to attack demanding problems of theoretical nuclear physics. In this way Meißner *et al.* have investigated production rates of carbon and oxygen, which are essential for the formation of life, in their dependence on fundamental constants of nature. In addition they have studied interesting issues concerning the structure of nuclei. Related methods have been applied by Gezerlis *et al.* to study problems in nuclear physics. Their investigations reveal the importance of three-body forces for the structure of nuclei. Taking the detailed form of these interactions into account, and employing methods from renormalisation group theory, they are able to obtain insights into the nature of exotic nuclei and neutron matter, that are of astrophysical relevance.

Back to the realm of elementary particles, the inner structure of hadrons like protons or neutrons is studied with lattice QCD in the project of Schäfer *et al.*. They calculate probability distributions characteristic for the dynamics of quarks inside hadrons, which can be related to experimental results from particle accelerators. Employing appropriate boundary conditions they are able to avoid obstacles resulting from the topology of gauge fields.

The behaviour of hadronic matter at high temperatures and densities is of great interest for astrophysics and cosmology. Costly experiments at heavy ion colliders at CERN and Brookhaven are designed for its study. From the theoretical side the thermodynamical phase diagram of hadronic matter can be studied in lattice QCD. Combining theoretical methods from statistical physics with numerical methods, Karsch *et al.* obtain results on the phase transition from hadrons to the quark-gluon plasma, in particular about freeze-out parameters that can be compared to experimental results.

The Mass Difference Between Protons and Neutrons and the Fine Tuning of Physical Constants

Szabolcs Borsanyi¹, Stephan Dürr^{1,2}, Zoltan Fodor^{1,2,3}, Christian Hölbling¹,
Sándor D. Katz³, Stefan Krieg², Laurent Lellouch⁴, Thomas Lippert^{1,2},
Antonin Portelli⁵, Kálmán K. Szabó^{1,2}, and Bálint C. Tóth¹

¹ Department of Physics, University of Wuppertal, D-42119 Wuppertal, Germany
E-mail: fodor@theorie.physik.uni-wuppertal.de

² Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52428 Jülich, Germany

³ Institute for Theoretical Physics, Eötvös University, H-1117 Budapest, Hungary

⁴ Aix-Marseille Université, CPT, UMR 7332, F-13288, Marseille, France

⁵ School of Physics & Astronomy, University of Southampton, SO17 1BJ, UK

We report on a large scale calculation of the origin of the mass difference between the proton and the neutron that was carried out to a large extent on the JUQUEEN computer.

1 Introduction

To our current knowledge, the strong nuclear force is described at a fundamental level by quantum chromodynamics (QCD)¹, a quantum field theory of quarks and gluons that is strongly coupled at low energies^{2,3}. The strong coupling, which gives the theory its very rich dynamical content is at the same time responsible for the failure of conventional perturbative approaches that are the usual tool for solving quantum field theories: A small perturbation on top of the unperturbed “free” theory of quarks and gluons will never be able to even reproduce the hadrons, i.e. the particles connected to the strong nuclear force that we observe in experiment like protons and neutrons.

A direct numerical solution of the theory is possible if one discretises it on a space-time lattice⁴. This lattice approach has matured over the last decade into a reliable tool to compute otherwise unattainable predictions of QCD⁵. A few years ago, it has eventually been demonstrated⁶, using lattice techniques, that QCD indeed does give the correct masses to the most common hadrons within a few percent accuracy (see Fig. 1).

With the validity of the approach thus confirmed, one can start to ask more subtle questions. One of these questions is the origin of the mass difference between proton and neutron. Although this difference is relatively tiny – the neutron is only 0.14% more massive than the proton⁷ – it does have profound consequences for the evolution and existence of the universe as we know it. As the more massive particles, neutrons can decay into protons (plus an electron and a neutrino), but since the mass difference is tiny, the process is relatively slow: Free neutrons have a half life of about 15 minutes. These properties are rather crucial, as a more massive proton in a hydrogen atom would decay into a neutron, obviously leading to an entirely different universe. When the nuclei that predominantly populate our universe today formed during the first twenty minutes after the big bang, the value of the neutron half-life and thus the proton neutron mass difference also determined

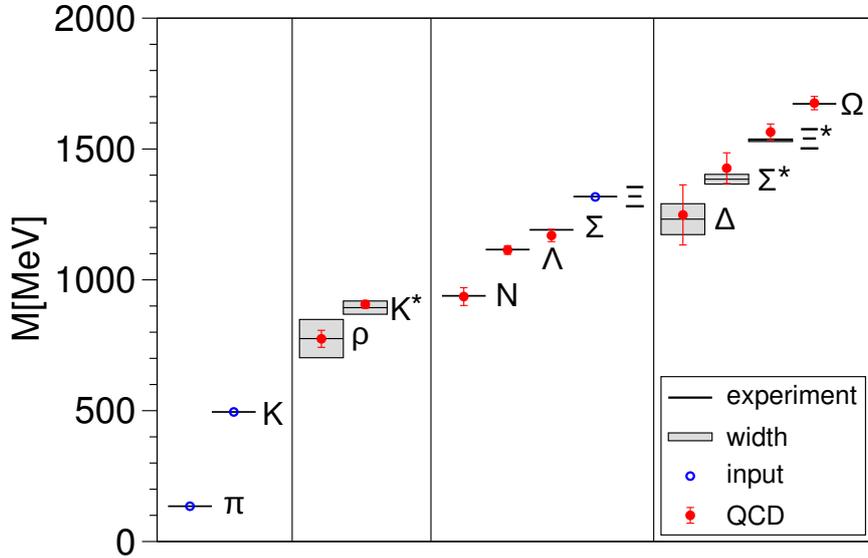


Figure 1. The light hadron spectrum⁶.

what fraction of the primordial protons would fuse to helium nuclei. A smaller mass difference and thus longer half-life would have left little hydrogen as fuel for stars⁸.

It is therefore very interesting to note, that the mass difference between proton and neutron actually results from a cancellation of two competing effects. On the one hand, the proton as a charged particle has a higher electromagnetic self energy and thus mass than the neutron. This effect is however overcompensated by the neutron consisting of one additional “down” quark as compared to the protons “up” quark. We know that the “down” quark is slightly more massive than the “up” quark and we also know that this is an exceptional situation, as the second and third generation “down” type quarks (the “strange” and the “bottom”) are actually lighter than their respective “up” type counterparts (the “charm” and “top”)⁷.

While these qualitative features are well understood, the complex dynamics of QCD makes any quantitative understanding of the proton-neutron mass difference in terms of the fundamental theory extremely difficult. We report here on a recent calculation that computed the proton-neutron mass splitting and related quantities *ab initio* utilising a lattice formulation of QCD and the quantum field theory of the electromagnetic interaction (QED)⁹ which was mostly carried out on the JUQUEEN supercomputer of Forschungszentrum Jülich.

2 Computational Challenges

In order to compute the mass difference between proton and neutron, one has to eliminate two approximations that are usually done in a lattice calculation: The equal mass of up and down quark (usually called the isospin limit) and the neglect of electromagnetic effects.

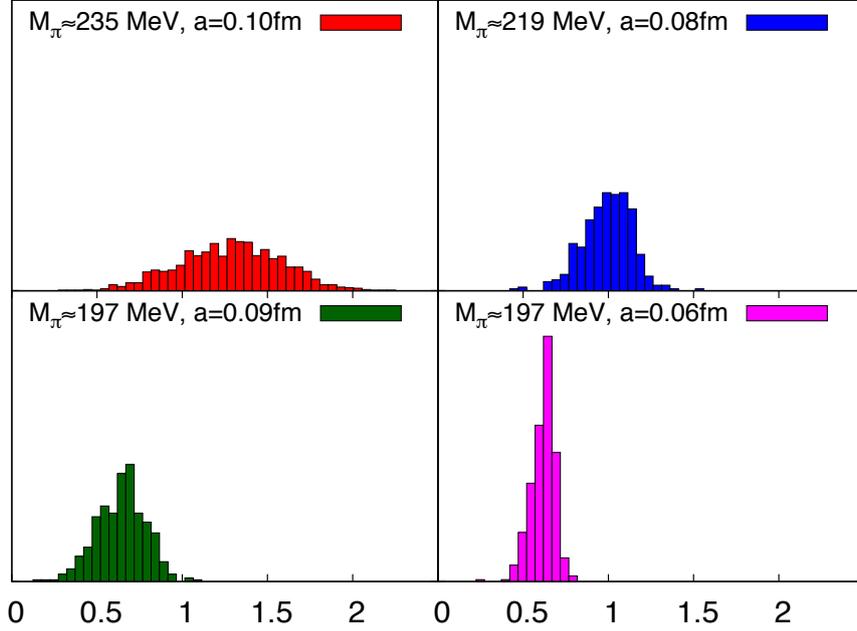


Figure 2. Inverse iteration number for each ensembles with the smallest quark mass at each lattice spacing⁹.

2.1 Nondegenerate Up and Down Quarks

In typical lattice computations of QCD, the mass of the up and down quark are usually set to be equal. Although the up quark mass is known to be less than half that of the down quark¹⁰

$$\frac{m_u}{m_d} = 0.46(2)(2) \quad (1)$$

the approximation is well justified by the fact that both up and down quark masses are substantially smaller than the dynamical scale of the theory

$$m_u, m_d \ll \Lambda_{\text{QCD}} \sim 300 \text{ MeV} \quad (2)$$

For the *ab initio* calculation of the mass splittings, however, we had to lift this degeneracy. In principle this is straightforward to achieve by using different masses in the update algorithm. In practice however, the smallness of the up quark mass poses a technical challenge. At the core of the update algorithms are inversions of large, sparse matrices whose condition number is inversely proportional to the quark mass minus a fluctuating offset (see e.g. Ref. 11). In the worst case, the condition number could diverge leading to a so-called exceptional configuration. Because the size of the fluctuations of this offset are related to the quality of the lattice discretisation (more precisely to the chiral symmetry of the fermion operator, see e.g. Ref. 12) an efficient lattice discretisation is essential. We have chosen a highly efficient discretisation that specifically suppresses the couplings of long range physical modes to short range lattice artefacts^{13,14}. As demonstrated in Fig. 2, the

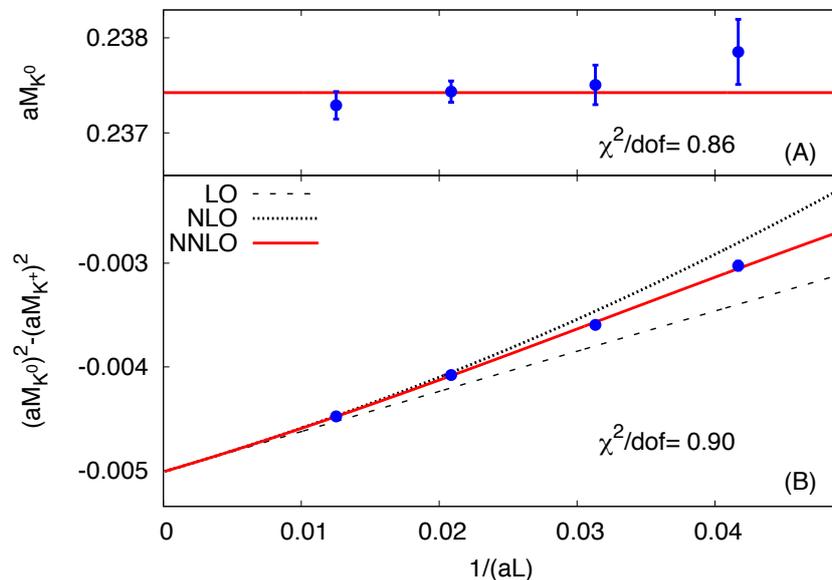


Figure 3. The measured mass of the neutral kaon (top) and the mass square difference between the neutral and charged kaon (bottom) plotted as a function of inverse lattice size⁹. Blue points are lattice data while the lines represent our analytical predictions in leading (LO), next to leading (NLO) and third order (NNLO) in $1/L$.

inverse iteration number remains safely above 0 even for our ensembles with the smallest quark masses and we are thus clear of exceptional configurations.

2.2 QED on the Lattice

The second major challenge one faces when computing the proton neutron mass splitting is the inclusion of electromagnetic effects into the lattice calculation. At first sight this difficulty might seem odd: After all, the quantum field theory of the electromagnetic interaction, quantum electrodynamics (QED) is much better understood and typically simpler than the theory of the strong interaction, QCD. On a fundamental level however, QED poses some conceptual problems that QCD successfully avoids. First of all, the flip side of the strong QCD coupling at low energies is a vanishing coupling at high energies, also known as asymptotic freedom^{2,3}. In QCD this property guarantees that at short distances no additional dynamics appears and thus a finer and finer lattice discretisation will ultimately reproduce the continuum theory. In QED the opposite is true: at smaller distance the coupling diverges and no continuum limit exists in principle. We are dealing with an effective theory that is only valid down to a certain distance scale. This scale fortunately is very small so that it can be neglected in our calculations, but it does pose a limit on the electromagnetic coupling we were able to simulate.

A second, potentially more severe problem results from the massless photon present in QED. While the strong interaction is also carried by a massless particle – the gluon – in QCD, these particles can never be observed outside bound states, which are always mas-

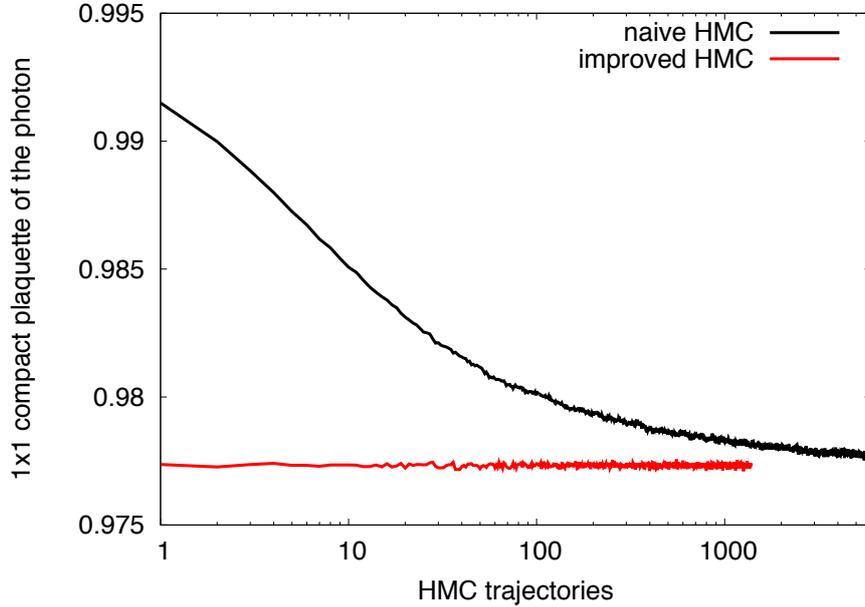


Figure 4. Evolution of the plaquette with simulation time for a standard HMC and the improved momentum space HMC described in the text⁹.

sive. Since the propagation of a particle over a long distance is exponentially suppressed by its mass¹⁵, a massive particle usually does not feel the finite extent of its spacetime lattice very strongly. The massless photon on the other hand induces severe self interactions of particles winding around the entire lattice. The most severe of them with vanishing momentum $\vec{p} = 0$ have to be removed in order to render the theory consistent¹⁶. We demonstrated analytically and numerically that the resulting theory is well defined and computed the remaining finite volume effects to third order in the inverse box length⁹ $1/L$ unearthing on the way some subtleties of a nonrelativistic effective description¹⁷ (see Fig. 3).

The long range correlations induced by the massless photons also have a negative effect on the update algorithm. In a standard molecular dynamics based hybrid Monte Carlo (HMC) algorithm¹⁸, long range modes have huge autocorrelations. To counter these effects, we designed a momentum space version of this algorithm that drives modes with different momenta k with a force proportional to $1/k^2$. The resulting speedup is documented in Fig. 4.

There is a further subtlety involved when trying to answer the question of QED contributions to a certain effect like the neutron-proton mass difference. The general strategy to compute QED contributions is to simply switch off the electromagnetic interaction and check for the difference. Although it is easy enough to simply turn off the electromagnetic interaction in our simulations, there is no easy answer to the question of which quantities should be kept constant while doing so to reproduce “nature without QED”. Nature can not guide us here since we can not simply turn off electromagnetic interactions between elementary particles. What we can do however is taking a sensible quantity that we do

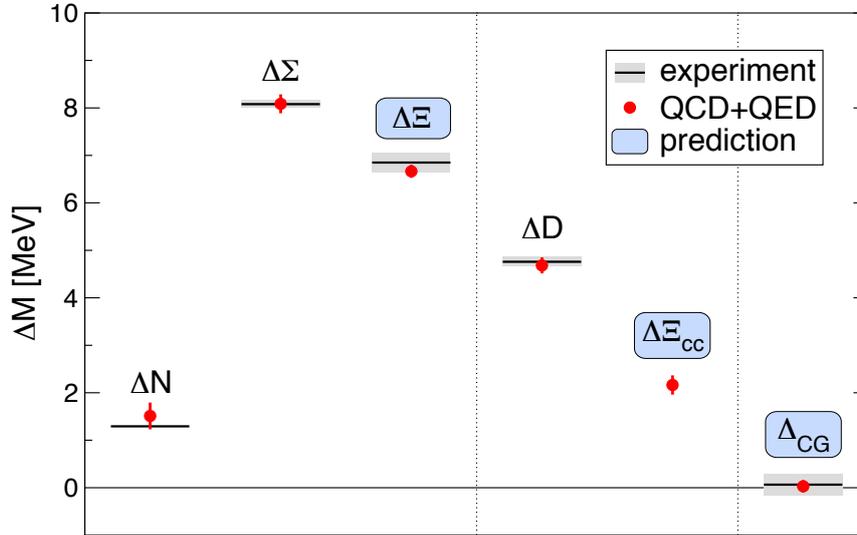


Figure 5. A comparison of the small mass differences between some related hadrons as found in experiment (black lines) and an *ab initio* calculation in QCD+QED (red dots). The grey bands on experimental values show current experimental uncertainties. The leftmost point, ΔN , is the mass difference between neutron and proton. For further details see Ref. 9.

expect to vary very little while changing the electromagnetic coupling and simply define that it stays precisely constant during that process. We could demonstrate⁹ that the mass difference between a positively and negatively charged Σ particle is a suitable quantity that, within our statistical accuracy, is not expected to be affected by changing the electromagnetic coupling and consequently used it to define our “nature without QED” point.

3 Data Sets and Analysis

We generated a total of 41 ensembles of lattice discretised QCD+QED gauge fields at different parameter values with four non-degenerate dynamical quarks. Ground state masses were extracted and a statistical analysis of the pertaining fit quality ensured the absence of excited state contaminations. The parameters themselves were chosen such that a safe interpolation and extrapolation to the physical point was possible and the total error of the procedure was minimised. Interpolations to the physical point were carried out in the electromagnetic coupling constant and the strange and charm quark masses. Extrapolations were performed in the lattice spacing, the box size and the light quark mass. All interpolations and extrapolations were performed with a variety of ansätze that covered the possible influence of effects that our data were not accurate enough to discern and thus give a reliable estimate of the possible systematic errors of our analysis procedure. The spread of the results thus obtained, properly weighted by the quality and estimated information content of the respective fit, was used as an estimate of the systematic error.

For further details we refer the interested reader to the supplementary material provided in Ref. 9.

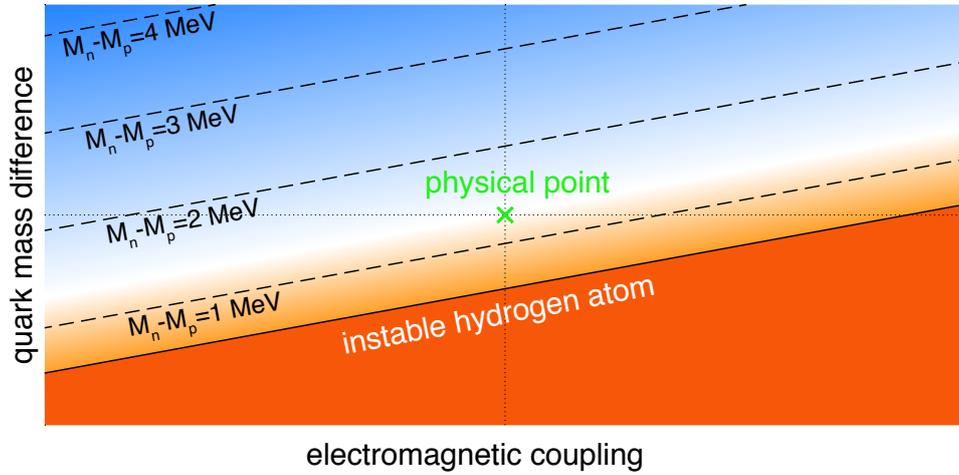


Figure 6. The figure shows a contour plot of the proton neutron mass difference as one varies the electromagnetic coupling (x-axis) and the difference between the “down” and the “up” quark mass from 0 to twice their physical values. The physical point, designating the values actually realised in nature, is marked by a green cross. The red region corresponds to a mass difference between neutron and proton that is less than the mass of the electron. In such a world, the hydrogen atom would not be stable as it would be favourable for the proton in its core to capture the electron and transform into a neutron.

4 Results

Fig. 5 displays the main result of our computation, the small mass differences between some closely related hadrons compared to experimentally known values where available. The leftmost point, ΔN is the mass difference between neutron and proton, which we could confirm to be correctly described by QCD+QED. Our calculations also give a prediction of the mass splittings between the doubly charmed baryons Ξ_{cc}^{++} and Ξ_{cc}^+ , which have not clearly been observed in experiment yet⁷.

Coming back to the original question of quantifying the contribution to the neutron-proton mass difference of electromagnetic effects versus the mass difference of “up” and “down” quarks, Fig. 6 displays a contour plot of the neutron-proton mass difference when hypothetically varying both the electromagnetic coupling and the difference of “down” and “up” quark mass between zero and twice their physical values respectively. This figure illustrates how much wiggle room nature had in choosing the fundamental parameters before our universe would be dramatically different.

Acknowledgements

The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN¹⁹ at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF)

and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

References

1. H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett., **B47**, 365–368, 1973.
2. H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett., **30**, 1346–1349, 1973.
3. D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett., **30**, 1343–1346, 1973.
4. K. G. Wilson, *Confinement of Quarks*, Phys. Rev., **D10**, 2445–2459, 1974.
5. Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabo, *The Order of the quantum chromodynamics transition predicted by the standard model of particle physics*, Nature, **443**, 675–678, 2006.
6. S. Durr et al., *Ab-Initio Determination of Light Hadron Masses*, Science, **322**, 1224–1227, 2008.
7. K. A. Olive et al., *Review of Particle Physics*, Chin. Phys., **C38**, 090001, 2014.
8. J. Bernstein, L. S. Brown, and G. Feinberg, *COSMOLOGICAL HELIUM PRODUCTION SIMPLIFIED*, Rev. Mod. Phys., **61**, 25, 1989.
9. Sz. Borsanyi et al., *Ab initio calculation of the neutron-proton mass difference*, Science, **347**, 1452–1455, 2015.
10. S. Aoki et al., *Review of lattice results concerning low-energy particle physics*, Eur. Phys. J., **C74**, 2890, 2014.
11. T. DeGrand and C. E. Detar, *Lattice methods for quantum chromodynamics*, 2006.
12. C. Hoelbling, *Lattice QCD: concepts, techniques and some results*, Acta Phys. Polon., **B45**, no. 12, 2143, 2014.
13. S. Capitani, S. Durr, and C. Hoelbling, *Rationale for UV-filtered clover fermions*, JHEP, **11**, 028, 2006.
14. S. Durr et al., *Scaling study of dynamical smeared-link clover fermions*, Phys. Rev., **D79**, 014501, 2009.
15. H. Yukawa, *On the Interaction of Elementary Particles I*, Proc. Phys. Math. Soc. Jap., **17**, 48–57, 1935, [Prog. Theor. Phys. Suppl.1,1(1935)].
16. M. Hayakawa and S. Uno, *QED in finite volume and finite size scaling effect on electromagnetic properties of hadrons*, Prog. Theor. Phys., **120**, 413–441, 2008.
17. Z. Fodor, C. Hoelbling, S. D. Katz, L. Lellouch, A. Portelli, K. K. Szabo, and B. C. Toth, *Quantum electrodynamics in finite volume and nonrelativistic effective field theories*, 2015.
18. S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Hybrid Monte Carlo*, Phys. Lett., **B195**, 216–222, 1987.
19. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities, **1**, A1, 2015, <http://dx.doi.org/10.17815/jlsrf-1-18>.

The Pseudocritical Line in the QCD Phase Diagram

Heng-Tong Ding¹, Prasad Hegde¹, Olaf Kaczmarek², Frithjof Karsch^{2,3},
Edwin Laermann², Swagato Mukherjee³, Peter Petreczky³, Hauke Sandmeyer²,
Christian Schmidt², Patrick Steinbrecher^{2,3}, and Wolfgang Söldner⁴

¹ Institute of Particle Physics, Central China Normal University, Wuhan 430079, China

² Fakultät für Physik, Universität Bielefeld, 33501 Bielefeld, Germany
E-mail: edwin@physik.uni-bielefeld.de

³ Physics Department, Brookhaven National Laboratory, Upton, NY 11973, USA

⁴ Fakultät für Physik, Universität Regensburg, 93053 Regensburg, Germany

We report about ongoing work to determine the chiral transition line in 2+1 flavour QCD for small values of the light quark chemical potential. The curvature of this line can be derived from a scaling analysis of the chiral condensate and its derivative with respect to the chemical potential. We outline the method and show results as available.

1 Introduction

One of the central goals of the physics program of the present and future heavy ion colliders is the exploration of the phase diagram of Quantum Chromo Dynamics (QCD) the theory of the strong interactions of quarks and gluons. At high temperatures and/or baryon densities the theory undergoes a transition from a confined, hadron phase where chiral symmetries are spontaneously broken to the so-called quark gluon plasma which is characterised by deconfinement and restored chiral symmetries. The general picture of the phase diagram is such that at small densities the two regions are presumably separated by a second order phase transition line in the chiral limit of vanishing quark masses. For small quark masses the transition happens along a pseudocritical line and is a rapid crossover. At sufficiently large densities or, equivalently, chemical potentials the transition could turn into a first order phase transition. It is the aim of the work presented here to shed more light on the location and the properties of the transition lines and perhaps provide some means to study the Critical End Point separating the first order line from second order/crossover.

In the experimental studies of the QCD phase diagram mentioned above the analysis of conserved charge fluctuations, the correlation between fluctuations of different charges and their higher moments, play an increasingly important role; not the least because experiments start to be able to measure them with rising reliability. In particular higher order cumulants of net baryon number, net electric charge and net strangeness carry information on the phase structure of strong interaction matter as function of the corresponding chemical potentials. At present there are thus major on-going efforts undertaken to measure observables related to these higher-order non-Gaussian fluctuations in various heavy-ion collision experiments¹⁻⁵.

Cumulants of conserved charge fluctuations characterise the thermal conditions in a heavy ion collision at the time of freeze-out, i.e. at the time when partons have recombined again and formed hadrons and when inelastic scatterings have ceased to change the particle

decomposition of the expanded fireball. Large deviations of these fluctuation observables from a non-critical baseline model, e.g. the hadron resonance gas (HRG) model, will occur in the vicinity of a phase transition. The hope is that experiments may detect a non-monotonic behaviour of cumulants of charge fluctuations at zero or non-zero values of the chemical potentials which, on the one hand, would be indicative⁶ for critical behaviour in the vicinity of the chiral phase transition at small μ_B and, on the other hand, would provide evidence for the existence of the QCD Critical End Point (CEP)⁷⁻¹⁰ at non-zero chemical potential. While the former may be studied at the highest energies at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory as well as at the Large Hadron Collider (LHC) at CERN in its lead-lead collision mode, the latter may be found with the Beam Energy Scan (BES) program at RHIC.

This critical point, if it exists, is likely to be connected to the chiral phase transition at zero chemical potential for two degenerate light (l) and a strange (s) quark flavour. In fact, in the chiral limit the CEP will be a tri-critical point at which the line of second order chiral transitions turns into a line of first order transitions. An important aspect of the theoretical analysis of the chiral transition is the quantitative determination of the chiral transition line at moderate, non-zero values of the chemical potential. In this context, it is an important issue to determine the relation between the QCD transition temperature T_c at non-zero quark chemical potential μ_q , and the experimentally determined freeze-out curve because the experimental success may depend crucially on the relative position of these two lines.

The determination of the QCD phase diagram clearly is a non-perturbative problem. As such it invites to be tackled by a numerical evaluation of the QCD path integral on space time lattices i.e. lattice QCD. However, at non-vanishing baryo-chemical potential the QCD action becomes complex such that standard Monte Carlo methods to compute the path integrals stochastically don't work anymore. This is the famous sign problem. Various ways have been invented to by-pass this problem at least at small densities. Here, in order to determine the crossover transition line at non-zero chemical potential, we follow an approach that is based on Taylor expansion around $\mu_q = 0$ and a scaling analysis relating a pseudocritical line (crossover) at non-zero quark masses to the true critical line in the chiral limit.

2 The Method

A theoretically sound and attractive approach to determine the pseudocritical (crossover) line in the QCD phase diagram at nonvanishing quark masses is based on a scaling analysis which relates this line to a true critical line in the chiral limit. We have worked out this method¹⁹ and sketch it in the following.

At leading order in the quark chemical potential μ_q the variation of the transition temperature with chemical potential is parametrised in terms of the constant κ_q

$$\frac{T_c(\mu_q)}{T_c} = 1 - \kappa_q \left(\frac{\mu_q}{T}\right)^2 + \mathcal{O}\left(\left(\frac{\mu_q}{T}\right)^4\right). \quad (1)$$

Here T_c is the phase transition temperature in the chiral limit. This constant κ_q can be determined by analysing the dependence of the light quark, chiral condensate $\langle\bar{\psi}\psi\rangle_l$ on the quark chemical potential. Of course, at vanishing light quark mass one would simply

determine the temperature at which $\langle \bar{\psi}\psi \rangle_l$ vanishes. At non-zero but small values of the quark mass this information is encoded in scaling functions.

Scaling functions arise from the partition function or, equivalently, the free energy density f . In the vicinity of a critical point regular contributions f_r to the partition functions become negligible in higher order derivatives and the singular behaviour of response functions will generally be dominated by contributions coming from the singular part f_s of the free energy density^a

$$f(T, m_l, m_s, \mu_q) = f_s(T, m_l, m_s, \mu_q) + f_r(T, m_l, m_s, \mu_q). \quad (2)$$

The singular part of the free energy density depends on the parameters of the QCD Lagrangian, e.g. the quark masses, and the external control parameters, temperature and chemical potentials, enter only through two relevant couplings. These scaling variables, t and h , control deviations from criticality, $(t, h) = (0, 0)$, along the two relevant directions, which in the case of QCD characterise fluctuations of the energy and chiral condensate, respectively. To leading order the scaling variable h depends only on parameters that break chiral symmetry in the light quark sector, while t depends on all other couplings. In particular, t will depend on the quark chemical potential

$$\begin{aligned} t &\equiv \frac{1}{t_0} \left(\frac{T - T_c}{T_c} + \kappa_q \left(\frac{\mu_q}{T} \right)^2 \right), \\ h &\equiv \frac{1}{h_0} \frac{m_l}{m_s}, \end{aligned} \quad (3)$$

The curvature parameter κ_q appearing in the above expression for t is the same as in Eq. 1 as the (pseudo) critical line $T_c(\mu_q)$ is determined by $t = 0$. Note that the strange quark mass has been chosen to render the magnetic field like parameter h dimensionless, and t_0 , h_0 are non-universal scale parameters.

Since the singular part of the free energy, f_s , is a homogeneous function of its arguments it can be rewritten in terms of the scaling variable $z = t/h^{1/\beta\delta}$ as

$$f_s(t, h) = h^{1+1/\delta} f_s(z, 1) \equiv h^{1+1/\delta} f_s(z). \quad (4)$$

where β , δ are known critical exponents of the three-dimensional $O(N)$ universality class^{21,22}, $\beta = 0.349$ and $\delta = 4.780$ for three-dimensional $O(2)$ models and $\beta = 0.380$ and $\delta = 4.824$ for $O(4)$, respectively. All parameters entering the definition of t and h , i.e. t_0 , h_0 and T_c may depend on the strange quark mass, but are otherwise unique in the continuum limit of (2+1)-flavour QCD. Just like the transition temperature T_c , however, also t_0 and h_0 are cut-off dependent and will need to be extrapolated to the continuum limit.

The universal critical behaviour of the order parameter, $M \sim \partial f / \partial m_l$, which in QCD is the chiral condensate, is controlled by a scaling function $f_G(z)$ that arises from the singular part of the free energy density after taking a derivative with respect to the light quark mass,

$$M(t, h) = h^{1/\delta} f_G(z). \quad (5)$$

^aFor systems belonging to the 3-dimensional $O(2)$ or $O(4)$ universality classes this does not hold for the thermal response function (specific heat) as the relevant critical exponent α is negative in these cases.

Also the scaling function $f_G(z)$ is well-known for the $O(2)$ and $O(4)$ universality classes through studies of three-dimensional spin models²¹.

To extract information about the curvature κ_q it suffices to consider the leading order Taylor expansion coefficient of the chiral condensate,

$$\frac{\langle \bar{\psi}\psi \rangle_l}{T^3} = \left(\frac{\langle \bar{\psi}\psi \rangle_l}{T^3} \right)_{\mu_q=0} + \frac{\chi_{m,q}}{2T} \left(\frac{\mu_q}{T} \right)^2 + \mathcal{O}((\mu_q/T)^4), \quad (6)$$

the mixed susceptibility $\chi_{m,q}$

$$\frac{\chi_{m,q}}{T} = \frac{\partial^2 \langle \bar{\psi}\psi \rangle_l / T^3}{\partial(\mu_q/T)^2} = \frac{\partial \chi_q / T^2}{\partial m_l / T} \quad (7)$$

which may also be viewed as the quark mass derivative of the light quark number susceptibility (χ_q). By means of the scaling relations the mixed susceptibility is immediately proportional to κ_q ,

$$\frac{\chi_{m,q}}{T} = \frac{2\kappa_q T}{t_0 m_s} h^{-(1-\beta)/\beta\delta} f'_G(z), \quad (8)$$

where $f'_G(z) \equiv df_G(z)/dz$ is easily obtained from $f_G(z)$ by using the parametrisations for the latter given in Refs. 21, 23. The way to proceed thus is to compute $\chi_{m,q}$ as a function of T and m_l numerically and extract κ_q , exploiting the knowledge on β, δ etc..

3 Some Technical Remarks

As pointed out in the previous section the first task is to determine the non-universal parameters t_0, h_0 and T_c at a given value for the lattice spacing, a , from the scaling properties of the chiral order parameter. For QCD this order parameter is (related to) the chiral quark condensate. However, owing to QCD being a Quantum Field Theory, the order parameter receives multiplicative as well as, at non-vanishing quark masses, additive renormalisations. The multiplicative renormalisation can be taken care of by using the renormalisation group invariant quantity

$$M_b = m_s \langle \bar{\psi}\psi \rangle_l. \quad (9)$$

The quadratically divergent additive correction $\sim m_l$ can in addition be subtracted according to

$$M = m_s (\langle \bar{\psi}\psi \rangle_l - m_l/m_s \langle \bar{\psi}\psi \rangle_s). \quad (10)$$

Multiplication by N_τ^4 , where N_τ is the lattice extent in the temporal direction related to temperature by $aT = 1/N_\tau$, renders both quantities dimensionless. The difference between those two definitions becomes irrelevant in the chiral limit. Alternatively it can be accounted for by scaling correction terms.

On the lattice the quark condensate is obtained from the inverse of the Dirac matrix M_f for a given quark flavour f ,

$$\langle \bar{\psi}\psi \rangle_f = \frac{1}{4} \frac{1}{N_\sigma^3 N_\tau} \langle \text{Tr} M_f^{-1} \rangle \quad (11)$$

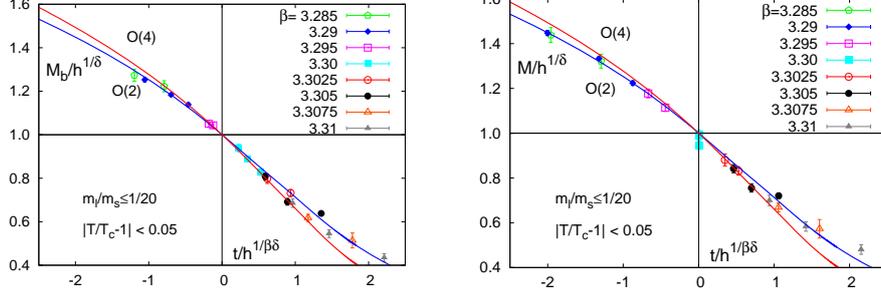


Figure 1. Fit of the $O(2)$ scaling function to numerical results for the light quark condensate. This analysis has been performed within the p4 discretisation with light quark masses $m_l/m_s \leq 1/20$ on $N_\tau = 4$ lattices. From Ref. 18.

by taking the trace Tr over Dirac and space-time indices. Similarly, the mixed susceptibility is obtained from traces over expressions involving M_l ,

$$\frac{\chi_{m,q}}{T} = \frac{1}{N_\sigma^3} (\langle \mathcal{C}_2 \rangle + 2 \langle \mathcal{C}_1 \mathcal{D}_1 \rangle + \langle \mathcal{C}_0 \mathcal{D}_2 \rangle + \langle \mathcal{C}_0 \mathcal{D}_1^2 \rangle - \langle \mathcal{C}_0 \rangle (\langle \mathcal{D}_2 \rangle + \langle \mathcal{D}_1^2 \rangle)) \quad (12)$$

where

$$\mathcal{C}_n = \frac{1}{4} \left. \frac{\partial^n \text{tr} M_l^{-1}}{\partial \hat{\mu}^n} \right|_{\hat{\mu}=0}, \quad \mathcal{D}_n = \frac{1}{2} \left. \frac{\partial^n \ln \det M_l}{\partial \hat{\mu}^n} \right|_{\hat{\mu}=0}, \quad (13)$$

with the shorthand notation $\hat{\mu} = \mu_q a$ for the chemical potential expressed in units of the lattice spacing. The different traces are obtained from stochastic estimators: given vectors of random numbers satisfying $\langle r_i^* r_j \rangle = \delta_{ij}$ in the average over the random vectors, one repeatedly solves $MX = r$ to obtain e.g. the quark condensate as $\langle r_j^* X_j \rangle$. Recycling the solutions X and applying them as part of the right hand side in e.g. $MY = M'X$ where $M' = \partial M / \partial \hat{\mu}$ delivers $\mathcal{C}_1 \sim \text{Tr} M^{-1} M' M^{-1}$ as $\langle r_j^* Y_j \rangle$. Thus, quite a few right hand sides are needed to obtain all requested traces so that deflating the Krylov space of the Conjugate Gradient inverter²⁵ from $\mathcal{O}(100)$ pre-calculated low eigenmodes of M is applied.

4 Results

The analysis path sketched in Sec. 2 has been suggested and worked out by us in Ref. 19. As a proof of principle of the method it was applied there to lattice regularised QCD in the p4 discretisation scheme. As a prerequisite, in Ref. 18 we had studied the scaling behaviour on $N_\tau = 4$ lattices down to a light to strange quark mass ratio of $m_l/m_s = 1/80$. With the strange quark mass at its physical value this corresponds to a Goldstone pion as low as 75 MeV. As can be read off Fig. 1 the data is in agreement with $O(N)$ scaling in the chiral order parameter. Here we show results for both definitions of the parameter, Eqs. 9 and 10, for light quark masses smaller than 1/20 of the strange quark mass. The figure further reveals that on the basis of the condensate $O(2)$ is not distinguishable from $O(4)$.

It turned out that the scaling region was entered already at $m_l/m_s = 1/10$. At this and larger light quark mass values besides the singular part two regular terms were needed to obtain a good fit. As a byproduct this study provides evidence for the Goldstone effect in QCD at high temperature which consists of a $1/\sqrt{m_l}$ divergence of the disconnected susceptibility $\chi_m \sim \partial\langle\bar{\psi}\psi\rangle_l/\partial m_l$ in the chirally broken phase, giving further support that the scaling machinery works consistently. In Ref. 19 we added the analysis of $N_\tau = 8$ configurations, yet only light quark masses down to $m_l/m_s = 1/20$ were available. We found again agreement with scaling also at this finer lattice. A separate scaling fit for only the lightest quark masses was beyond the scope of this paper, and the non-universal scaling parameters were determined from a combined fit to all data in this case.

While these results are based on the p4 discretisation, in Ref. 20 we applied successfully the $O(N)$ scaling relation to HISQ configurations, an action with the smallest taste violations of all known staggered type actions, and obtained very good descriptions of the data on chiral observables. These were used subsequently in chiral interpolations of the quark condensate in order to determine the pseudocritical temperature T_c .

Whereas in all these studies agreement with $O(N)$ critical behaviour was observed, we want to stress that the nature of the chiral phase transition is not a question settled yet; other scenarios e.g. suggest a first order transition, see Ref.²⁶ for a recent paper. Thus, in the ongoing project we are carrying out a dedicated scaling analysis within the HISQ discretisation, with quark masses reaching down to values that lead to smaller than physical pion masses. The scaling parameters are then input to the combined analysis of the transition line.

Once the results on the non-universal scaling parameters t_0, h_0 are at hand, the curvature constant κ_q can be extracted from the mixed susceptibility, Eq. 8. This is shown in Fig. 2 taken from our pilot study. Note that κ_q is the only free parameter and the fit is completely fixed otherwise. As far as the value for κ_q is concerned a phenomenologically motivated approach¹² has obtained a result very close to our value from that study while a recent analysis based on simulations at imaginary chemical potential²⁷ suggests a larger curvature. It is the aim of the present project to reduce the systematic errors and greatly improve the prediction of that curvature at real chemical potentials.

5 Concluding Remarks

We have sketched a procedure to determine the curvature of the chiral transition line of QCD in the temperature - chemical potential plane at small baryon densities by means of numerical simulations of lattice QCD. This method is based on the analysis of the scaling behaviour of QCD in the vicinity of the chiral transition. We have demonstrated the feasibility of this procedure by providing quantitative results for the curvature in a pilot study. The present project is aiming at establishing these results at decreased systematic uncertainties.

Being of high interest on its own, the curvature result obtained so far has been used to compare the chiral transition line in the $T - \mu$ QCD phase diagram with the freeze-out curve, see Fig. 3. Here, for the first time the freeze-out parameters T^f and μ_B^f were obtained by comparing experimental results from heavy ion collisions on fluctuations and correlations of conserved quantum numbers with lattice computations of the same quantities¹⁴. While in such comparisons care must be taken that effects of conservation laws due to finite system sizes, acceptance cuts etc. do not invalidate grand canonical statistics, an

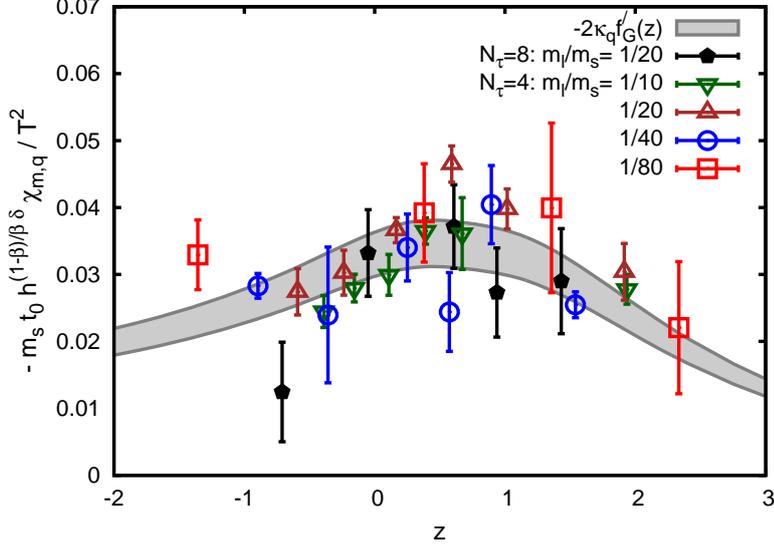


Figure 2. The scaled mixed susceptibility as function of the scaling variable $z = t/h^{1/\beta\delta}$. The data from simulations with the p4 action are compared to the $O(2)$ scaling curve. The band shows a 10% error on this curve which arises from statistical errors on the calculated observables as well as from the errors on the scaling parameters t_0 and z_0 . From Ref. 19.

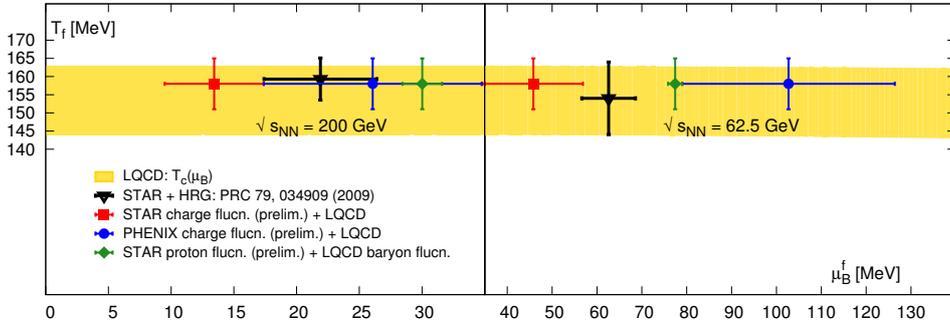


Figure 3. Freeze-out temperatures T^f and baryon chemical potentials μ_B^f obtained through direct comparisons between lattice calculations and the preliminary STAR and PHENIX data for cumulants of net charge and net proton fluctuations at two different beam energies $\sqrt{s_{NN}} = 200$ GeV and 62.5 GeV. The shaded region indicates the present lattice results^{19,15} for the pseudocritical temperature T_c as a function of μ_B . From Ref. 24.

approach which does not rely on model assumptions certainly deserves further work. At present it is very interesting that the results for the freeze-out line $T^f(\mu_B^f)$ are close to the chiral/deconfinement crossover, raising the hope that the collision experiments may signal the presence of criticality in the phase diagram.

Acknowledgements

This work has been partially supported through the Bundesministerium für Bildung und Forschung (BMBF) under grant no. 05P15PBCAA and the U.S. Department of Energy (DoE) under Contract no. DE-SC0012704.

The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN²⁸ at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

References

1. M. M. Aggarwal et al. (STAR Coll.), Phys. Rev. Lett. **105**, 022302, 2010.
2. L. Kumar (STAR Coll.), Nucl. Phys. **A862-863**, 125, 2011.
3. S. Jena (ALICE Coll.), PoS WPCF2011, 046, 2011.
4. L. Adamczyk et al. [STAR Coll.], Phys. Rev. Lett. **113**, 092301, 2014.
5. N. M. Abdelwahab et al. [STAR Coll.], Phys. Rev. **C92**, 021901, 2015.
6. K. Morita, B. Friman, K. Redlich and V. Skokov, Eur. Phys. J. **C74**, 2706, 2014; Phys. Rev. **C88**, 034903, 2013.
7. M. A. Stephanov, Phys. Rev. Lett. **102**, 032301, 2009.
8. M. A. Stephanov, Phys. Rev. Lett. **107**, 052301, 2011.
9. M. Asakawa, S. Ejiri and M. Kitazawa, Phys. Rev. Lett. **103**, 262301, 2009.
10. B. Friman, F. Karsch, K. Redlich and V. Skokov, Eur. Phys. J. **C71**, 1694, 2011.
11. P. de Forcrand, S. Kim, O. Philipsen, PoS LAT 2007, 175, 2007.
12. G. Endrodi et al., JHEP **1104**, 001, 2011.
13. A. Bazavov et al. (HotQCD Coll.), Phys. Rev. **D86**, 034509, 2012.
14. A. Bazavov et al., Phys. Rev. Lett. **109**, 192302, 2012.
15. A. Bazavov et al., Phys. Rev. Lett. **111**, 082301, 2013.
16. R. Bellwied et al., Phys. Rev. Lett. **111**, 202302, 2013.
17. E. Follana et al. (HPQCD and UKQCD Coll.), Phys. Rev. **D75**, 054502, 2007.
18. S. Ejiri et al., Phys. Rev. **D80**, 094505, 2009.
19. O. Kaczmarek et al., Phys. Rev. **D83**, 014504, 2011.
20. A. Bazavov et al. (HotQCD Coll.), Phys. Rev. **D85**, 054503, 2012.
21. J. Engels et al., Phys. Lett. **B514**, 299, 2001.
22. J. Engels, L. Fromme and M. Seniuch, Nucl. Phys. **B675**, 533, 2003.
23. J. Engels and F. Karsch, Phys. Rev. **D85**, 094506, 2012.
24. S. Mukherjee and M. Wagner, PoS CPOD2013, 039.
25. K. Wu and H. Simon, SIAM J. Matrix Anal. Appl. **22**, 602, 2000.
26. C. Bonati et al., Phys. Ref. D90, 074030, 2014.
27. C. Bonati et al., Phys. Ref. D90, 114025, 2014.
28. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities, **1**, A1, 2015, <http://dx.doi.org/10.17815/jlsrf-1-18>.

Pushing the Boundaries of Nuclear Physics with Lattice Simulations

Ulf-G. Meißner^{1,2}, Timo A. Lähde², and Thomas Luu²
for the NLEFT collaboration

¹ Helmholtz-Institut für Strahlen- und Kernphysik and Bethe Center for Theoretical Physics,
Universität Bonn, D-53115 Bonn, Germany
E-mail: meissner@hiskp.uni-bonn.de

² Institute for Advanced Simulation (IAS-4), Institut für Kernphysik (IKP-3) and Jülich Center for
Hadron Physics, Forschungszentrum Jülich, 52425 Jülich, Germany

We review the recent highlights of nuclear lattice effective field theory, which is continuing to push the boundaries of *ab initio* nuclear many-body calculations, both in terms of nuclear structure and nuclear reactions. Significantly, this provides a deeper understanding of key nuclear processes that produce the elements necessary for life as we know it, such as carbon and oxygen. This remarkable progress has been made possible by recent dramatic increases in HPC resources, as well as advances in computational methods and algorithmic improvements.

1 Introduction

Recent advances in high-performance computing (HPC) has enabled nuclear physics to enter a new and exciting era. Calculations of nuclear structure and reactions that were once considered nearly impossible are now being readily performed. The research performed by the NLEFT (Nuclear Lattice Effective Field Theory) collaboration is at the forefront of this development. Such calculations are *ab initio* in the sense that they use nuclear forces derived from the chiral effective Lagrangian of Quantum Chromodynamics (QCD), which is the underlying theory that describes the interactions of quarks and gluons. For few-nucleon systems, the chiral effective field theory (EFT) for the forces between two, three and four nucleons have been worked out to high orders in the chiral power counting. This force consists of long-ranged exchanges of one or more pions, and shorter ranged multi-nucleon contact interactions. By combining these EFT forces with Monte Carlo methods developed by the lattice QCD community, the NLEFT collaboration has successfully studied the properties of *p*-shell nuclei (such as ¹²C and ¹⁶O). These nuclei have formed the calculational boundary of more traditional nuclear many-body techniques, such as Green's function Monte Carlo.

With recent advances in the methods and algorithms of NLEFT, this boundary has been pushed further by recent *ab initio* calculations of nuclei in the *sd*-shell. Furthermore, the NLEFT formalism has been developed and adapted to include the treatment of nuclear reactions. This ongoing line of research is now rapidly addressing key questions related to the formation of elements, including those that enable life as we know it. In the following, we briefly review the motivation and methodology behind the NLEFT formalism. We also present recent highlights of our research and conclude with an outlook on future progress.

2 Theoretical Background of NLEFT Simulations

Nuclei are self-bound systems of nucleons (protons and neutrons). As the nucleons themselves consist of quarks and gluons, and hence are not fundamental degrees of freedom, the forces between nucleons are not completely given in terms of two-body interactions, but include three-body and higher terms. Computing the properties of multi-nucleon systems presents a very difficult challenge. The complicated structure of the interaction coupled with the quantum mechanical nature of such systems leads to an exponential growth in the computational effort as a function of the number of nucleons A . For $A \leq 4$, bound state energies and scattering phase shifts have traditionally been calculated by exact (numerical) solution of the Lippmann-Schwinger or Faddeev-Yakubovsky equations. For $A \geq 5$, well established many-body techniques have been developed, such as the no-core shell model and coupled-cluster methods. These ultimately rely on the direct diagonalisation of a large matrix M in order to solve a problem of the form $Mx = b$. As the size of M increases exponentially, the memory and processing power of currently available HPC systems are quickly exhausted, which confines such methods to systems with $A \leq 12$. In order to push beyond $A = 12$, simplifications to the interaction between nucleons as well as other *ad hoc* assumptions become necessary.

In the context of QCD, systems of quarks and gluons also exhibit exponential scaling in the number of degrees of freedom, but instead of relying on direct diagonalisation in order to calculate observables, methods have been developed to stochastically *estimate* observables. The quarks and gluons are placed on a discrete space-time lattice, and Monte Carlo sampling of the propagation of the particles is performed in order to capture the most relevant contributions to a given observable. Such “lattice QCD” calculations provide a much reduced calculational complexity. Moreover, lattice QCD calculations are fully non-perturbative and provide the only known rigorous way to compute the properties of QCD in the non-perturbative (low-energy) regime. Still, it should be kept in mind that the stochastic nature of lattice QCD induces an associated uncertainty in each calculated observable, in addition to possible issues arising from numerical sign oscillations (the “sign problem”) or from an unfavourable signal-to-noise ratio.

While a formalism similar to lattice QCD is used in NLEFT calculations, in the latter case the nucleons form the degrees of freedom that propagate on the space-time lattice, such that the interactions between nucleons are provided by chiral EFT. The stochastic nature of the Monte Carlo importance sampling of the nucleons’ trajectories provides a softer scaling of computational complexity with A . This, in turn, is what allows NLEFT to push the boundaries of *ab initio* calculations beyond those reached by more traditional methods.

3 Nuclear Physics on a Space-Time Lattice

In NLEFT simulations, Euclidean space-time is discretised on a torus of volume $L_s^3 \times L_t$, where L_s is the side length of the (cubic) spatial dimension, and L_t denotes the extent of the Euclidean time dimension. The lattice spacing in the spatial dimensions is denoted a , analogously to a_t in the temporal dimension. The maximal momentum on the lattice is thus $p_{\max} \equiv \pi/a$, which serves as the UV regulator of the theory. Nucleons exist as pointlike particles on the lattice sites, and the interactions between nucleons (pion exchanges and

contact terms) are treated as insertions on the nucleon world lines via auxiliary-field representations. The nuclear forces have an approximate spin-isospin SU(4) symmetry (Wigner symmetry) that is of fundamental importance in suppressing numerical sign oscillations that plague any Monte Carlo simulation of strongly interacting fermions at finite density. This is in contrast to lattice QCD, where any finite baryon chemical potential renders the Monte Carlo simulation unfeasible.

We compute the properties of multi-nucleon systems by means of the transfer matrix projection Monte Carlo method. There, each nucleon is treated as a single particle propagating in a fluctuating background of pion and auxiliary fields, the latter representing the multi-nucleon contact interactions. Due to the very strong binding between four nucleons occupying the same lattice site, we find that the convergence of the chiral EFT expansion can be greatly accelerated by means of smeared LO contact interactions. We start the Euclidean time projection from a Slater determinant Ψ_A of single-nucleon standing waves for Z protons and N neutrons (with $A = Z + N$) in a periodic cube. We then use a Wigner SU(4) symmetric Hamiltonian as a computationally inexpensive filter for the first few Euclidean time steps, which also suppresses sign oscillations dramatically. Finally, we apply the full LO chiral EFT Hamiltonian and calculate the ground state energy and other properties from the correlation function

$$Z(t) \equiv \langle \Psi_A | \exp(-tH) | \Psi_A \rangle, \quad (1)$$

in the limit of large Euclidean projection time t . Higher-order contributions, such as the Coulomb repulsion between protons and other isospin-breaking effects (due to the light quark mass difference), are computed as perturbative corrections to the LO amplitude. The properties of excited states are obtained from a multi-channel projection Monte Carlo method. In our LO lattice action, the nucleon kinetic energy and momentum-dependent smearing factors of the contact interactions are treated using $\mathcal{O}(a^4)$ improvement. Moreover, all lattice operators are included up to $\mathcal{O}(Q^3)$, where Q denotes the momentum transfer between pions and nucleons. This includes operators related to the breaking of rotational symmetry on the lattice. The strengths of such operators can be tuned to eliminate unphysical effects, such as the mixing of the 3D_3 partial wave into the 3S_1 - 3D_1 channel.

4 Recent NLEFT Results

We shall now discuss the highlights of selected recent NLEFT calculations, which demonstrate both the strengths of the NLEFT approach as well as recent algorithmic developments. These include the sensitivity of the triple alpha process to the variation of the fundamental physical parameters, and the *ab initio* demonstration of alpha clustering in ${}^{12}\text{C}$ and ${}^{16}\text{O}$. On the algorithmic side, we report recent results on the restoration of rotation symmetry on the lattice, as well as a new method for ameliorating sign oscillations in Monte Carlo simulations of NLEFT.

4.1 Variation of the Fundamental Physical Parameters

A significant advantage of an *ab initio* approach to nuclear physics is that we can determine the sensitivity of key nuclear processes to changes in the fundamental physical parameters, such as the light quark mass m_q (or, equivalently, the pion mass M_π) and the

electromagnetic fine-structure constant α_{em} ¹. The only additional input required for such a calculation is knowledge of the M_π -dependence of the coefficients of the contact terms in the chiral EFT interaction between nuclei. Such knowledge enables us to determine the level of “fine-tuning” that occurs in nuclear physics, and in particular address “anthropic” questions related to the formation of life in the Universe. The production rate of carbon and oxygen via the triple-alpha process and alpha capture on carbon are of particular anthropic significance, since these elements appear to be essential to life as we know it. The only known environment where the required density, temperature, and timescale for the production of such elements exists is within the cores of red giant stars. Our NLEFT simulations have revealed that the production rates of carbon and oxygen are likely to remain stable under perturbations smaller than $\simeq 2 - 3\%$ in m_q or α_{em} . This result is shown in Fig. 1, together with current knowledge of the M_π -dependence of the chiral EFT interaction².

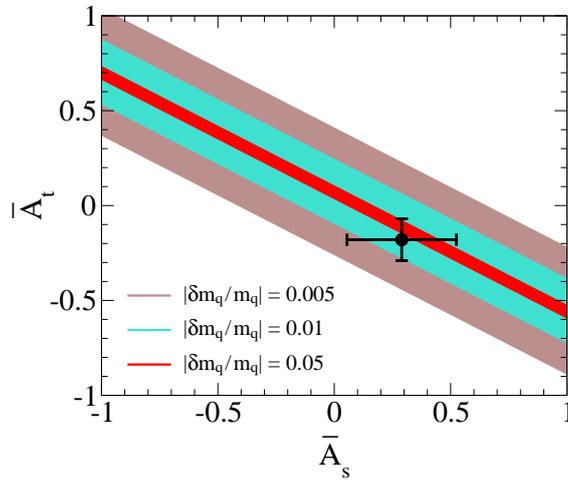


Figure 1. “Survivability bands” for carbon-oxygen based life from NLEFT simulations, due to 0.5% (broad outer band), 1% (medium band) and 5% (narrow inner band) changes in the light quark mass m_q , in terms of the slopes of the inverse NN S -wave scattering lengths a_s^{-1} and a_t^{-1} , where $A_i \equiv \partial a_i^{-1} / \partial M_\pi$ at the physical value of the pion mass M_π . The data point with error bars represents the most recent NNLO determination of the pion mass dependence of the NN scattering lengths in chiral EFT.

4.2 Alpha Clustering in Light and Medium-Mass Nuclei

Alpha clustering within light and medium mass nuclei is an important physical phenomenon that is captured in our NLEFT simulations. Our NLEFT simulations of ^{12}C and ^{16}O , for example, provide insight into the exact configuration of alpha clustering. In the case of ^{12}C , the 0_1^+ ground state corresponds to a configuration of alpha clusters in the shape of an equilateral triangle (on the lattice an isosceles right triangle), while the 0_2^+ “Hoyle state” appears to closely resemble an obtuse triangle similar to the ozone molecule. In Fig. 2, we show the alpha clustering realised in the analogous states of ^{16}O . The alpha clustering picture which emerges from NLEFT simulations is also strongly supported

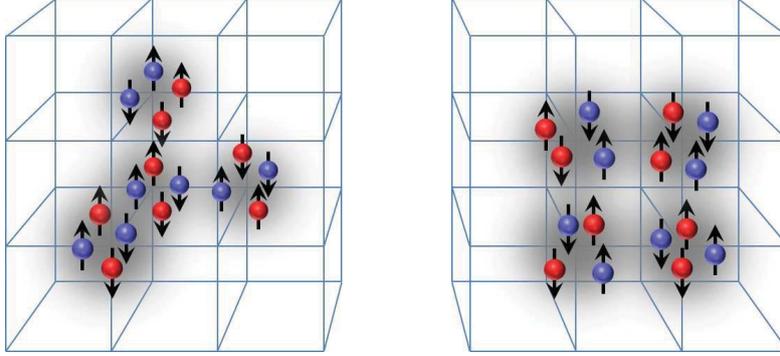


Figure 2. Illustration of alpha cluster states with tetrahedral and square configurations. The tetrahedral configuration is found to have large overlap with the 0_1^+ ground state of ^{16}O , while the square configuration corresponds closely to the excited 0_2^+ state.

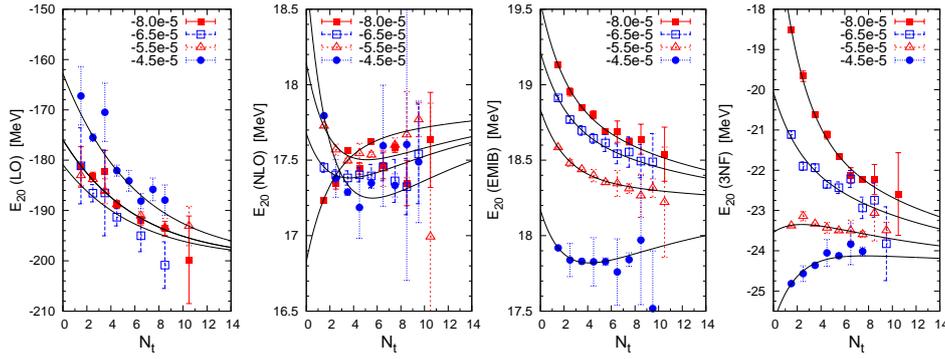


Figure 3. NLEFT simulation of the ground-state energy of ^{20}Ne to NNLO in the EFT expansion. From left to right, the contributions are the LO energy, the isospin-symmetric NLO correction, the electromagnetic and isospin-breaking (EMIB) corrections, and the three-nucleon force (3NF) correction. Contributions of NLO and higher order are treated perturbatively. The curves show a correlated fit to four different choices of the Wigner $\text{SU}(4)$ symmetric Hamiltonian, in terms of a spectral density given by a sum of three energy delta functions. The physical results are found by extrapolation $t = a_t N_t \rightarrow \infty$.

by our results for the electromagnetic transitions between the states in question, as these are sensitive probes of the structure of the underlying wave functions. The same NLEFT simulations have also enabled the *ab initio* study of nuclei within the *sd*-shell, including the medium-mass nuclei ^{20}Ne , ^{24}Mg and $^{28}\text{Si}^4$. Our NLEFT simulation results for ^{20}Ne are shown in Fig. 3, including the full structure of the chiral EFT force up to NNLO in the EFT expansion. The extent of the Monte Carlo data in Euclidean time is limited by sign oscillations, which necessitates an extrapolation in order to obtain the final result for each observable.

4.3 Restoration of Rotational Symmetry on the Lattice

On the lattice, rotational invariance is broken from the full $SO(3)$ rotational group to the cubic group. Hence, observables computed on the lattice will in general be affected by rotational symmetry breaking effects. In particular, the unambiguous identification of excited states and the computation of transition amplitudes may suffer significantly due to the relatively large lattice spacings of $a \simeq 2$ fm in present NLEFT simulations. Hence, it makes sense to carefully determine the sources of rotational symmetry breaking in actual NLEFT simulations, and search for methods that minimise their impact on physical observables. We have therefore used a simplified alpha cluster model to study the lattice matrix elements of irreducible tensor operators as a function of $a^{5,6}$.

In order to minimise the effects of rotational symmetry breaking, we have introduced the ‘‘isotropic average’’ which consists of a linear combination of the components of a given matrix element, such that each component is weighted according to the Clebsch-Gordan coefficient with the associated quantum numbers. This method, which is equivalent to averaging over all lattice orientations, enables the unambiguous computation of matrix elements even at large lattice spacings. In Fig. 4, we illustrate the effect of isotropic averaging on the mean square radius of ${}^8\text{Be}$ within the alpha cluster model.

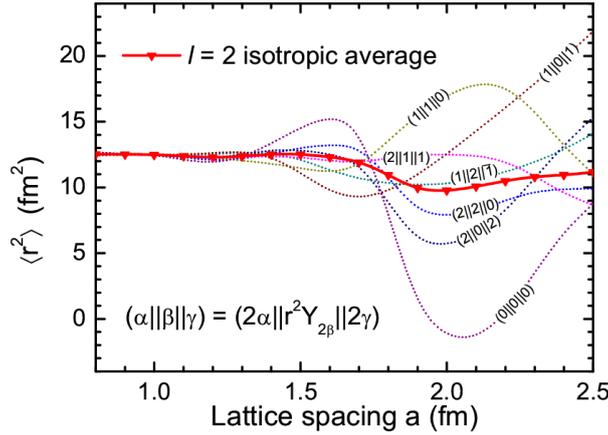


Figure 4. Mean square radii $\langle r^2 \rangle$ for the lowest 2^+ multiplet of ${}^8\text{Be}$ states within a simplified alpha-cluster model calculation. The reduced lattice matrix elements all merge in the limit $a \rightarrow 0$, while at finite a the matrix elements depend on the quantum number α , β and γ , which is indicative of rotational symmetry breaking. Such effects are nearly eliminated in the isotropic average, especially when $a \leq 1.7$ fm.

4.4 Combating the Sign Problem in NLEFT Simulations

While the Monte Carlo approach to NLEFT exhibits relatively mild scaling of the computational effort as a function of A , the applicability of NLEFT simulations is potentially limited by the fermionic sign oscillations inherent to the chiral EFT interaction, especially when $N \neq Z$. As already noted, the approximate Wigner $SU(4)$ symmetry of our trial wave functions greatly alleviates the sign oscillations, which are also found to be minimised for

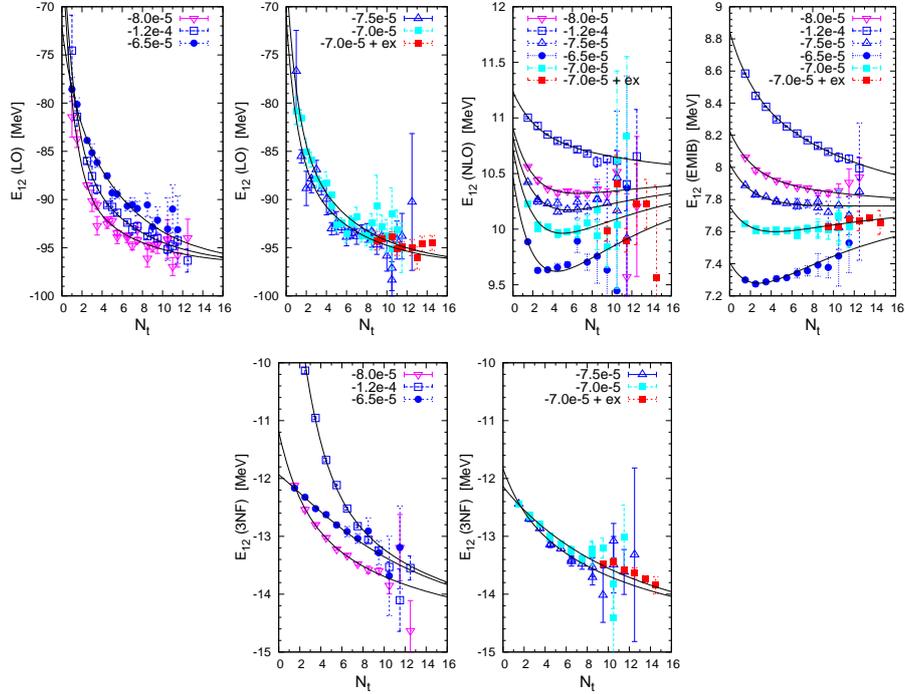


Figure 5. Updated extrapolation in Euclidean projection time for the ground state of ^{12}C , including data points (filled red squares) obtained by means of symmetry-sign extrapolation. Notation and conventions are similar to Fig. 3. The other sets of Monte Carlo data have been obtained directly at $d_h = 1$, for which simulations at $N_t = 14$ and $a_t = 1.32$ fm would be out of the question due to severe sign oscillations.

the alpha nuclei. Likewise, the momentum-dependent smearing of the LO interactions and the perturbative treatment of NLO and higher order corrections allow the sign problem to be circumvented further. Nevertheless, substantial extrapolation in Euclidean time remains necessary, which introduces additional uncertainties into the results. In order to reliably access larger nuclei as well as neutron-rich halo nuclei, we have developed a systematic approach called ‘‘Symmetry-sign extrapolation’’ (SSE) which allows us to perform NLEFT simulations for nuclei for which the sign problem is more severe⁷.

In the SSE method, Monte Carlo simulations are performed for the Hamiltonian H_{SSE} , given by

$$H_{\text{SSE}} \equiv d_h H_{\text{phys}} + (1 - d_h) H_{\text{SU}(4)}, \quad (2)$$

where H_{phys} is the full NLEFT Hamiltonian, $H_{\text{SU}(4)}$ is an SU(4) symmetric Hamiltonian which (for most cases of interest) is free of a sign problem, and $0 \leq d_h \leq 1$ is an adjustable parameter. By taking $d_h < 1$, we can always decrease the sign oscillations to a tolerable level, at the price of introducing an extrapolation $d_h \rightarrow 1$ in addition to $t \rightarrow \infty$. An additional constraint is provided by the requirement that the limit $d_h \rightarrow 1$ be independent of the choice of $H_{\text{SU}(4)}$. In Fig. 5, we show how the SSE method allows us to extend the

region in Euclidean projection time accessible to Monte Carlo simulations. Clearly, SSE provides us with a valuable tool for extending the boundaries of NLEFT simulations to the regime of neutron-rich halo nuclei and to nuclei near drip lines.

5 Future of NLEFT Simulations

At this time, NLEFT has matured into a well-established *ab initio* framework at the forefront of modern theory of nuclear structure and reactions. The ongoing development of both theory and algorithms is expected to provide further insight into a number of key problems, and to further extend the applicability of NLEFT to heavier nuclei. For example, NLEFT is well placed to address the issue of possible *P*-wave pairing in neutron matter, which would modify the nuclear equation of state and hence the cooling rate of neutron stars. Significant progress has also been achieved in applying NLEFT to nuclear reactions using the “adiabatic projection” method to reduce a complicated many-fermion problem to a computationally simpler cluster-cluster scattering problem⁸. These investigations pave the way for *ab initio* studies of processes relevant for stellar astrophysics, such as alpha-alpha scattering, alpha-carbon scattering and radiative capture. In particular, a model-independent calculation of the “holy grail” of nuclear astrophysics, *i.e.* the $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ reaction at stellar energies, appears to be within reach. In short, NLEFT appears headed toward an exciting time of progress and discovery.

Acknowledgements

We thank our NLEFT colleagues for their contributions to the results presented here. This work is supported in part by DFG, HGF, BMBF and CAS. The computational resources were provided by the Jülich Supercomputing Centre at Forschungszentrum Jülich and by RWTH Aachen.

References

1. U.-G. Meißner, Sci. Bull. (2015) **60**(1):43-54 [arXiv:1409.2959 [hep-th]].
2. E. Epelbaum, H. Krebs, T. A. Lähde, D. Lee and U.-G. Meißner, Eur. Phys. J. A **49** (2013) 82 [arXiv:1303.4856 [nucl-th]].
3. E. Epelbaum, H. Krebs, T. A. Lähde, D. Lee, U.-G. Meißner and G. Rupak, Phys. Rev. Lett. **112** (2014) 10, 102501 [arXiv:1312.7703 [nucl-th]].
4. T. A. Lähde, E. Epelbaum, H. Krebs, D. Lee, U.-G. Meißner and G. Rupak, Phys. Lett. B **732** (2014) 110 [arXiv:1311.0477 [nucl-th]].
5. B. N. Lu, T. A. Lähde, D. Lee and U.-G. Meißner, Phys. Rev. D **90** (2014) 3, 034507 [arXiv:1403.8056 [nucl-th]].
6. B. N. Lu, T. A. Lähde, D. Lee and U.-G. Meißner, Phys. Rev. D **92** (2015) 1, 014506 [arXiv:1504.01685 [nucl-th]].
7. T. A. Lähde, T. Luu, D. Lee, U.-G. Meißner, E. Epelbaum, H. Krebs and G. Rupak, Eur. Phys. J. A **51** (2015) 7, 92 [arXiv:1502.06787 [nucl-th]].
8. M. Pine, D. Lee and G. Rupak, Eur. Phys. J. A **49** (2013) 151 [arXiv:1309.2616 [nucl-th]].

Hadron Structure from Lattice QCD

Andreas Schäfer for the RQCD collaboration

Institute for Theoretical Physics, Regensburg University, 93040 Regensburg, Germany
E-mail: andreas.schaefer@physik.uni-r.de

We discuss the goals and techniques of our lattice QCD calculations and explain the criteria for declaring success in this field. Based on these general considerations we have adopted a new strategy to simulate with open boundary conditions as part of the CLS collaboration. Our NIC projects form part of this large scale international effort. We illustrate our approach by showing and shortly discussing three examples, one addressing parton distribution functions (PDFs) one addressing Generalised Parton Distributions (GPDs) and one addressing Distribution Amplitudes (DAs). The upshot is that while a very detailed picture of hadron structure is emerging, reaching the required level of theoretical control will still require a long term coordinated effort.

1 Introduction

Particle physics is presently in a somewhat peculiar situation: On the one hand the Standard Model of the electro-weak and strong interactions is extremely successful, passing successfully one high precision test after the other, and one has all reason to be highly satisfied with its achievements. On the other hand, however, we know that there must exist Beyond the Standard Model physics (BSM). Standard quantum field theory (QFT) breaks down for energies above the Planck scale so it is clearly an incomplete theory. On the experimental side one knows from astrophysics and cosmology that dark energy and dark matter do exist, but they do not fit into the Standard Model. Seen from this perspective every successfully passed test implies that it will be still harder to learn about BSM physics. As there are limits to the energy of affordable particle accelerators (the “energy frontier”) making progress along the “precision frontier” becomes ever more important. Progress in this direction requires primarily a reduction of QCD uncertainties. In some cases this can be achieved by performing perturbative QCD calculations of ever higher order but in other cases more precise lattice simulations are needed. In fact, progress in both directions is equally important as it is always the largest individual uncertainty which determines the total. Because more and more lattice results for non-perturbative quantities cannot be tested experimentally, the complete control of the theoretical uncertainties is indispensable and actually defines success. We contribute to the continuous extension in scope and increase in precision by our lattice simulations.

Unfortunately, the level of technical sophistication in QCD has reached a point where it has become very difficult to explain to a non-expert precisely how any given example fits into the grand picture just sketched. Therefore, in this report we will focus on just three topics for which this is still relatively easy and explain these in some detail. These topics are: 1.) parton distribution functions, 2.) generalised parton distributions and 3.) distribution amplitudes.

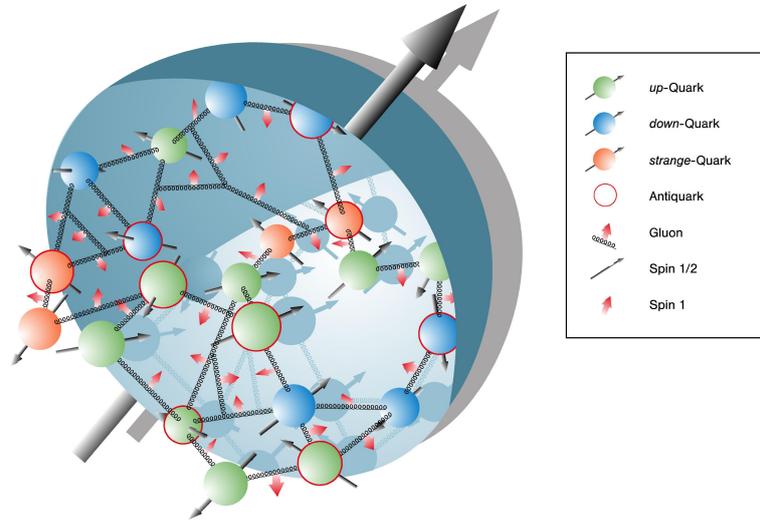


Figure 1. An illustration used by DESY to provide a feeling for the highly dynamical, fluctuating nature of the proton quantum ground state.

2 Fundamentals of Lattice QCD

While the hydrogen atom can be visualised quite well as being built from an electron and a proton with a mass which differs from the sum of their masses by only a fraction of 10^{-8} , the sum of the valence quark masses is nearly negligible compared to the mass of, e.g. the proton, i.e. the proton mass is generated nearly exclusively by quantum mechanical interactions. Attempts to illustrate its completely quantum, highly dynamical and relativistic nature will always be a bit misleading, see Fig. 1. In reality the proton can only be described by an incredibly complicated many particle quantum wave function. Lattice QCD allows to single out and determine very specific pieces of information which are relevant for very specific experimentally observable reactions.

The first step to achieve this is to perform an analytic continuation to imaginary time. It is one of the fundamental and quite fascinating properties of QFT that this mathematically rigorously and uniquely defined operation, translates many QFT problems into purely statistical ones which can be solved numerically by Monte-Carlo techniques. Part of progress in lattice QCD is to extend the classes of objects for which this is achieved. This is highly non-trivial and requires progress in continuum QFT as well as lattice QCD and progress is, therefore, painfully slow. On the other hand any new result on fundamental interactions with a truly reliable error margin (reliable error estimates are actually the crucial distinction between theory and model building) will remain valid (within this error estimate) as long as our universe will exist. With this time perspective in mind progress in the last decades was explosive, although the impression one gets on the time scale of a PhD thesis can be somewhat different.

Fig. 2 taken from the EIC White Paper¹ gives a more technical illustration of what was said above. Lattice input has acquired by now a similar importance as direct experimental

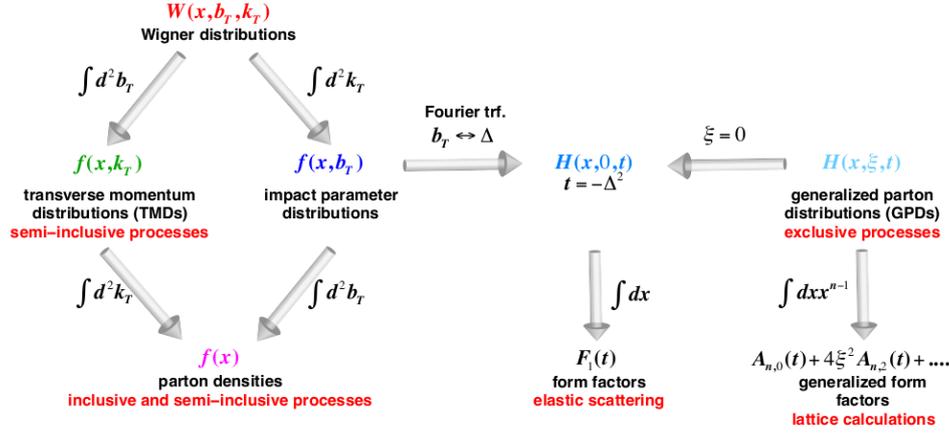


Figure 2. Part of the information characterising a hadron wave function can be expressed by Wigner distributions $W(x, b_T, k_T)$. Their information content can be reduced, e.g., by integrating out some variables. Experimentally, one was able to pin down several twist-2 parton densities $f(x)$, or rather $f(x, Q^2)$, and several form factors $F(t)$. Presently TMDs and GPDs have moved into the focus of attention. It became already clear that the purely experimental determination of all of them will hardly be possible without massive lattice QCD input to supplement experimental results. (A nucleon is, e.g., characterised by three independent types of twist-2 parton densities but eight GPDs.)

data and this might justify to invest for lattice QCD a noticeable fraction of the resources needed for such experiments.

The structure of lattice QCD calculations consists of two parts. First one generates ensembles of quark-gluon configurations with the correct probability distributions of QCD. These ensembles can be used, in principle, to obtain all information on all hadrons. The task is then in a second step to extract very specific pieces of information, e.g., information which is related to a given experiment or a given theoretical consideration. In practice, this is limited by statistical and systematic errors. Consequently, one combines information of all ensembles at ones disposal, e.g. ensembles with various quark masses (simulations with heavier than physical masses are much cheaper) various lattice spacings and lattice volumes. Due to this constant reuse of all generated data one cannot cite all results including all future results from an individual lattice QCD project but rather one has to pick some examples as illustrations, as we do in this report. Let us add that BG computers like JUQUEEN are perfectly suited for the generation of ensembles, while, e.g., JURECA is perfectly suitable for analysis. This is why quite often we submit joint computer time proposals. It should be obvious that the needed continuous progress in hadron physics requires a continuous increase in available computer resources.

3 The Epic Story of $\langle x_{u-d} \rangle$

The decade long story of $\langle x_{u-d} \rangle$ is possible best suited to illustrate the general situation described above. Twist-2 parton distributions have the great advantage to allow for a sim-

ple probabilistic interpretation while in QFT one usually has to discuss everything on the level of (interfering) probability amplitudes. For example $u(x, Q^2)$ gives the probability that one finds an up quark with a longitudinal momentum xP_N^μ , where P_N^μ is the (large) 4-momentum of a nucleon, when probed, e.g., by deep-inelastic electron scattering, at a spacial resolution scale of $1/Q$. Such scatterings were investigated, e.g., at HERA at DESY, Hamburg. For reasons we cannot explain here the difference between the expectation value of x for up and down quarks in a proton is partially protected from higher QCD corrections and thus an especially suitable quantity for a comparison with lattice calculations. Also it is known rather precisely from experiment. Therefore, a long existing strategy is to use the uncertainties of lattice QCD calculations for $\langle x_{u-d} \rangle$ as error estimate for all similar lattice QCD calculations. Obviously, this strategy works only, if the discrepancy between experimental and lattice QCD results is compatible with the error estimate.

As all earlier lattice calculations of $\langle x_{u-d} \rangle$ still had substantial uncertainties, we staged a large effort to achieve unprecedented accuracy in checking this agreement. The result is shown in Fig. 3. What one can not see from this figure is that we did our very best to reduce all systematic errors and to estimate the remaining uncertainties reliably. The remaining 30% discrepancy is far too large to be ignored and in our opinion points to a fundamental problem with some present day lattice QCD calculations for hadron structure. While the cause of this discrepancy is open to debate we strongly suspect discretisation errors, i.e. artefacts due to the finite lattice spacing used. The latter is typically larger than 0.05 fm which has to be compared to a typical hadron radius of 0.7 fm. Simulations on finer lattices are not only much more expensive but also face the problem of diverging topological auto-correlation times. This problem is linked to a fundamental property of QCD. In QCD the vacuum is not an empty state, but the state of lowest energy, which has, in fact, a highly complex structure. Part of this structure is that there exist infinitely many, distinct, degenerate vacuum states differing in their topological structure. A lattice QCD calculation gives correct results if all sectors are sampled democratically. The size of artefacts if this choice is biased or if the topological sector is even fixed are unclear and heavily debated. We decided, therefore, to change our whole strategy and to switch from simulations with periodic boundary conditions to simulations with open boundary conditions, which offer the only known way to avoid the sketched topology problem. Consequently we joined the CLS collaboration which has devoted its efforts to exploring this approach. As a consequence of this decision we had to start completely from scratch generating ensembles and analysing physical observables. Meanwhile a substantial number of first preliminary results exist but they are not yet finally released. Therefore, we will rather show results obtained for the same ensembles as used for Fig. 3, which have for the stated reason not well known discretisation errors, but illustrate nicely the type of investigations we perform.

4 The Total Angular Momentum of Quarks in the Nucleon

The quantities displayed in Fig. 2 and others allow to answer many questions on hadron structure, although often the formal relations are rather involved and non-intuitive. (The techniques allowing to derive them are usually subsumed under the heading Operator Product Expansion (OPE).) For example a calculation which is similar to that leading to $\langle x_{u-d} \rangle$

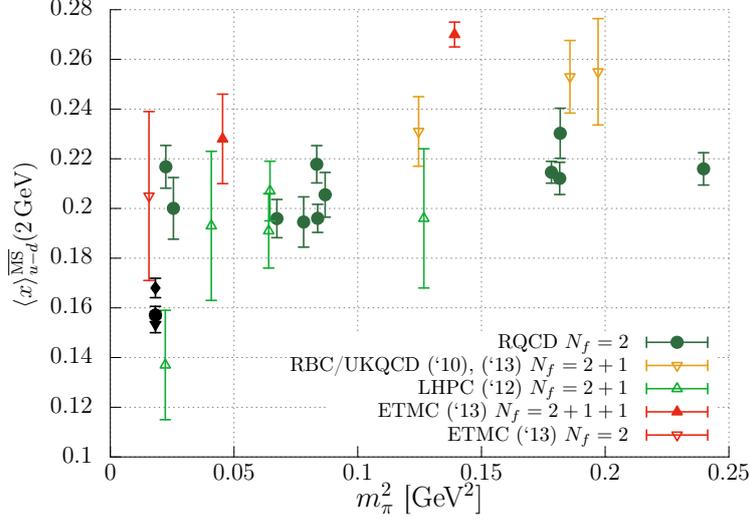


Figure 3. Results for $\langle x_{u-d} \rangle$ from various lattice collaborations as function of the pion mass squared, which is a measure of the up and down quark mass. Physical quark masses correspond to the position of the black symbols, which show the results of phenomenological fits to experimental data. Our results are the dark green filled circles. Our lowest mass point was obtained for our largest volume.

leads to a quantity which can be identified with the total angular momentum carried by individual quark species in a proton J_q , fulfilling

$$\begin{aligned}
 J_G + \sum_q J_q &= \frac{1}{2} \\
 J^q &= \frac{1}{2} (A_{20}^q(0) + B_{20}^q(0)) \\
 &= \frac{1}{2} \int_{-1}^1 dx x [H(x, \xi, 0) + E(x, \xi, 0)] \quad (1)
 \end{aligned}$$

where $H(x, \xi, 0)$ and $E(x, \xi, 0)$ are generalised parton distributions (GPDs). Fig. 4 shows our result for for the generalised form factor $A_{20}^q(Q^2)$ for the isovector combination of valence quarks ($B_{20}^q(Q^2)$ is not shown) and the resulting total angular momentum $J_{u-d} = J_u - J_d$. To obtain the error estimate, we varied every part of the analysis, fitting ranges, parametrisations as suggested by various levels of Chiral Perturbation Theory (ChPT), smearing strategies, This resulted in thousands of fits which we then histogrammed to read of a combined error estimate. Let us stress that moments ($\int_0^1 dx \dots$) are very difficult to estimate from experimental data alone, because any experiment can only probe a limited x range.

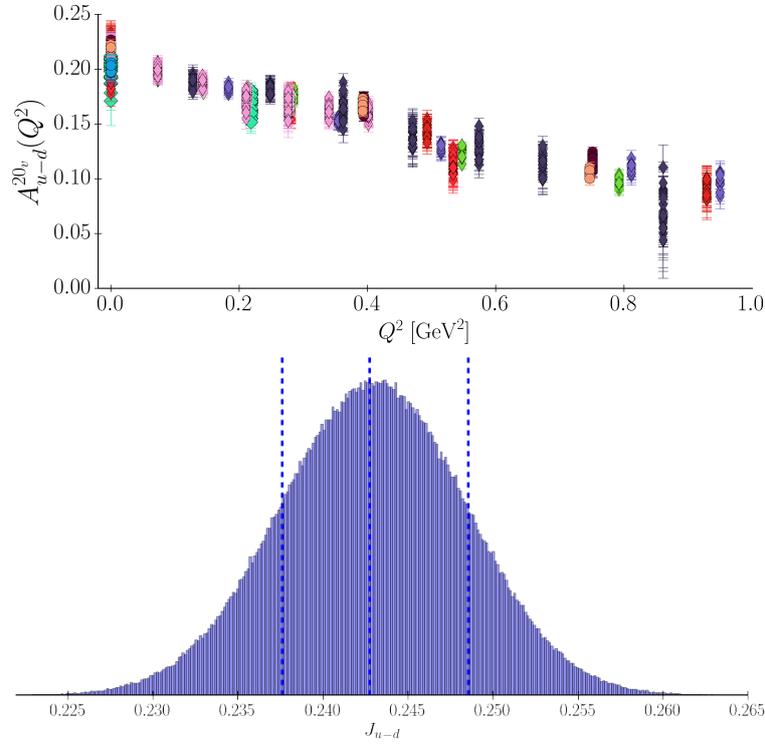


Figure 4. The generalised form factor A_{20}^{u-d} [top] and the histogram of fits for the resulting J_{u-d} [bottom]. The multiple symbols in the upper figure show the results of different fits to the primary lattice data. Very many ChPT fits to these and other similar generalised form factors resulted in the histogram in the lower figure from which we can extract a well determined fitting error.

5 Distribution Amplitudes

Fig. 2 comes not even close to displaying all well established elements of hadron structure. Distribution Amplitudes (DAs) contain independent information and play a pivotal role for the description of “exclusive” reactions (while distribution functions are central for “inclusive reactions”). In exclusive reactions, the 4-momenta of all participating particles are measured. In inclusive reactions this is not the case. Sufficiently hard exclusive reactions single out the properties of the leading Fock-state of a hadron multi-particle wave function. This is well known since decades, but with modern high-luminosity accelerators their precise investigation became feasible and in many respects even necessary to make full use of these machines. As the theory behind DAs is rather advanced we just show as illustration plots for the 3-quark component of the wave functions of the nucleon, the $N^*(1535)$ and $N^*(1650)$. Note that the latter wave functions contain nodes. As the momentum fractions x_1 , x_2 and x_3 of all three quarks have to add up to 1 these wave functions can be plotted best in the manner used in that figure. We are presently preparing a paper with the DAs of the full baryon octet based on our new CLS $N_f = 2 + 1$ ensembles for publication.

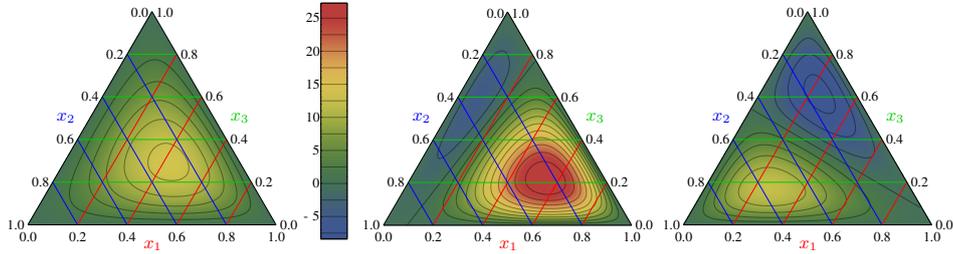


Figure 5. Barycentric plots of the nucleon [left], $N^*(1650)$ [centre] and $N^*(1535)$ [right] wave functions. Only the first moments of the distribution amplitude have been used to create these plots.

6 Algorithms

Obtaining precision results from lattice QCD poses specific challenges to software design and special algorithms are needed to achieve the needed statistical accuracy in reasonable time. The problem of topological freezing has been solved by the introduction of open boundary conditions in the time direction⁴. Relinquishing translational invariance in time, the method allows the topological charge to flow freely in and out of the lattice which prevents the topological charge from freezing and allows to go to finer and finer lattice spacings. The reduction in quark masses can be achieved by the use of highly efficient and parallel solvers for the inversion of the Wilson-Clover Dirac operator, employing various state-of-the-art techniques, e.g., domain decomposition and deflation. For the generation of the gauge ensembles within the CLS effort, we use a software package called “open-QCD”. It incorporates the above mentioned features like open boundary conditions and an efficient deflated solver. There are many other features that make this software very efficient, including twisted-mass Hasenbusch frequency splitting that allows for a nested hierarchical integration of the molecular dynamics at different time scales, decoupling the quickly changing but cheaper forces of the action from the more expensive low frequency part of the fermion determinant. The analysis is done with the publicly available lattice QCD software “chroma”⁵, steadily being extended by our group and others. E.g., recently an adaptive, aggregation based multigrid solver⁶ has been made available, enabling us to perform inversions of the Dirac operator also at very light quark masses at low cost, see Fig. 6.

The Hierarchical Data Format (HDF5) is used to handle parallel I/O and the management of our big amounts of data. Using Data-Grid technology established by the experimental groups at the LHC we are also able to move large files in short times. By making available such technology and know-how, computer centres contribute significantly to the success of large-scale numerical efforts.

7 Summary

We have presented a few example for how specific information can be isolated from the complete multi-particle wave function of a hadron and related to specific experimental

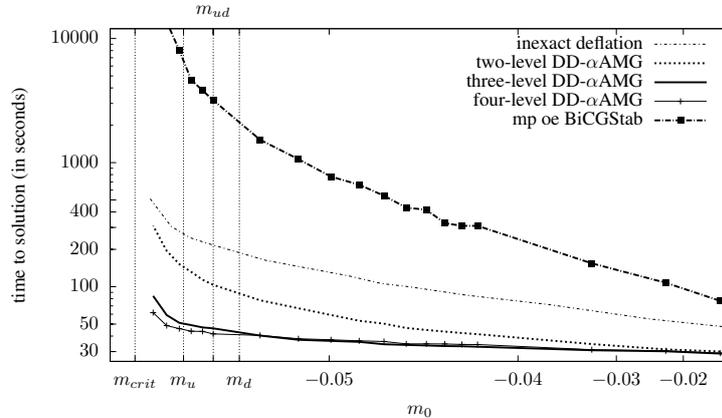


Figure 6. Time to solution for various different solvers of the Wilson-Clover Dirac operator as a function of the bare quark mass m_0 . Three versions of our new algorithm resulted in the three lowest curves.

observables. We argued that to do so with really controlled errors requires a qualitative change in how lattice simulations are done, e.g., by the adaption of open boundary conditions. As part of CLS we contribute to this extensive long-term effort, using the computer time we have been granted by NIC.

Acknowledgements

We acknowledge support by DFG (SFB/TRR-55).

The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN⁷ at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

References

1. A. Accardi *et al.*, *Electron Ion Collider: The Next QCD Frontier - Understanding the glue that binds us all*, arXiv:1212.1701 [nucl-ex].
2. G. S. Bali *et al.*, *The moment $\langle x \rangle_{u-d}$ of the nucleon from $N_f = 2$ lattice QCD down to nearly physical quark masses*, [arXiv:1408.6850 [hep-lat]]. Phys. Rev. **D90**, 074510, 2014.
3. V. M. Braun *et al.*, *Light-cone Distribution Amplitudes of the Nucleon and Negative Parity Nucleon Resonances from Lattice QCD*, [arXiv:1403.4189 [hep-lat]]. Phys. Rev. **D89**, 094511, 2014.

4. M. Lüscher and S. Schaefer, *Lattice QCD without topology barriers*, [arXiv:1105.4749 [hep-lat]]. JHEP **1107**, 036, 2011.
5. R. G. Edwards *et al.* [SciDAC and LHPC and UKQCD Collaborations], *The Chroma software system for lattice QCD*, [hep-lat/0409003]. Nucl. Phys. Proc. Suppl. **140**, 832, 2005.
6. A. Frommer, K. Kahl, S. Krieg, B. Leder and M. Rottmann, *Adaptive Aggregation Based Domain Decomposition Multigrid for the Lattice Wilson Dirac Operator*, [arXiv:1303.1377 [hep-lat]]. SIAM J. Sci. Comput. **36**, A1581, 2014.
7. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities, **1**, A1, 2015, <http://dx.doi.org/10.17815/jlsrf-1-18>.

The Strong Interaction at Neutron-Rich Extremes

Alexandros Gezerlis¹, Kai Hebeler^{2,3}, Jason D. Holt⁴, Javier Menéndez⁵,
Achim Schwenk^{2,3}, Johannes Simonis^{2,3}, and Ingo Tews^{2,3}

¹ Department of Physics, University of Guelph, Guelph, Ontario N1G 2W1, Canada

² Institut für Kernphysik, Technische Universität Darmstadt, 64289 Darmstadt, Germany
E-mail: schwenk@physik.tu-darmstadt.de

³ ExtreMe Matter Institute EMMI, GSI Helmholtzzentrum für Schwerionenforschung GmbH,
64291 Darmstadt, Germany

⁴ TRIUMF, 4004 Wesbrook Mall, Vancouver, British Columbia, V6T 2A3 Canada

⁵ Department of Physics, University of Tokyo, Hongo, Tokyo 113-0033, Japan

In our JUROPA/JURECA project, we have applied powerful many-body methods to strongly interacting nuclear systems. We have developed quantum Monte Carlo calculations with chiral effective field theory interactions to study light nuclei and neutron matter. Moreover, we have systematically implemented chiral three-nucleon forces in many-body calculations of medium-mass nuclei, using renormalisation group methods in momentum space and in-medium for calculations of valence-shell interactions. This enabled us to make predictions of exotic nuclei studied at rare isotope beam facilities and of extreme neutron-rich matter in astrophysics.

1 Introduction

The microscopic understanding of atomic nuclei and of high-density matter is a very challenging task. Powerful many-body simulations are required to connect the observations made in the laboratory to the underlying strong interactions between neutrons and protons, which govern the properties of nuclei and of strongly interacting matter in the universe. Renewed interest in the physics of nuclei is driven by discoveries at rare isotope beam facilities worldwide, which open the way to new regions of exotic, neutron-rich nuclei, and by astrophysical observations and simulations of neutron stars and supernovae, which require controlled constraints on the equation of state of high-density matter. Fig. 1 shows the substantial region of exotic nuclei that will be explored at the future FAIR facility in Darmstadt.

The nuclear many-body problem involves two major challenges. The first one concerns the derivation of the strong interaction between nucleons, which is the starting point of few- and many-body *ab initio* calculations. Since nucleons are not elementary particles, but composed of quarks and gluons, the strong interaction has a very complex structure. Although it is becoming possible to study systems of few nucleons directly based on quarks and gluons¹, the fundamental degrees of freedom of quantum chromodynamics (QCD), high-precision calculations of nuclei based on quarks and gluons will not be feasible in the foreseeable future. As a systematic approach, chiral effective field theory (EFT) allows to derive nuclear forces in terms of low-energy degrees of freedom, nucleons and pions, based on the symmetries of QCD². The chiral EFT framework provides a systematically improvable Hamiltonian and explains the hierarchy of two-, three-, and higher-body forces. The presence of such many-body forces is an immediate consequence of strong interactions³.

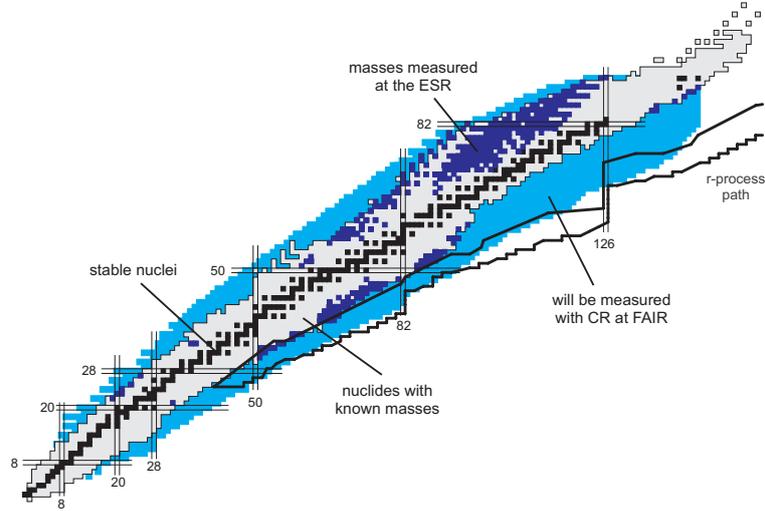


Figure 1. Chart of atomic nuclei. The black points mark the stable nuclei, while the grey region denotes all nuclei known to date. Of the latter, the points in purple have been studied with mass measurements at the GSI Helmholtzzentrum für Schwerionenforschung in Darmstadt. The blue points mark the exotic nuclei that will be explored at the future FAIR facility. Figure courtesy Y. Litvinov (GSI).

In particular, the computation and inclusion of three-nucleon (3N) forces in many-body calculations is one of the current frontiers⁴.

The second challenge concerns the practical solution of the many-body problem based on nuclear forces. Since the computational complexity grows significantly with the number of particles, up to about 10 years ago the scope of *ab initio* calculations was limited to light nuclei up to around carbon (with nucleon number $A = 12$). Due to advances on several fronts and also due to rapidly increasing computing power, this limitation has nowadays been pushed to much heavier systems (see, e.g., the recent work of Ref. 5). One key step was the development of powerful renormalisation group (RG) methods that allow to systematically change the resolution scale of nuclear forces^{6,7}. Such RG transformations lead to less correlated wave functions at low resolution and the many-body problem becomes more perturbative and tractable.

Our work focuses on the derivation of RG evolved interactions and electroweak operators, the inclusion of chiral 3N forces in many-body calculations of extreme neutron-rich nuclei, and on the development of quantum Monte Carlo simulations with chiral EFT interactions, which open up nonperturbative benchmarks for nuclei and high-density matter. These calculations enable us to explore the formation of structure in exotic nuclei, the properties of neutron-rich nuclei and matter that play a key role in the synthesis of heavy elements in the universe, as well as the nuclear physics involved in applications to fundamental symmetries, e.g., for the nuclear matrix elements of neutrinoless double-beta decay that probes the nature and mass scale of the neutrinos.

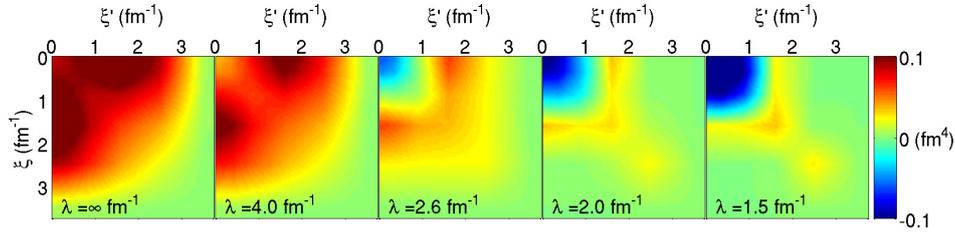


Figure 2. Contour plots of SRG-evolved 3N interactions as a function of three-body hypermomenta (see Ref. 8 for details). Shown is the systematic decoupling of low- and high-momentum parts in the interaction as the SRG resolution scale λ is lowered from left to right. (Figure taken from Ref. 8).

2 Renormalisation Group Evolution of Nuclear Interactions

The convergence behaviour and the required computational resources of many-body calculations for a given nucleus are governed by the properties of the employed nuclear forces. It is convenient to visualise nuclear interactions as a function in momentum space, where low momenta correspond to large interparticle distances and high momenta to short-range correlations. In general, a strong coupling of low- and high-momentum parts in nuclear interactions induce strong virtual excitations of particles and imply a poor perturbative convergence and large required basis spaces for the solution of the many-body Schrödinger equation. The similarity renormalisation group (SRG) allows to systematically decouple low-momentum physics from high-momentum details via a continuous sequence of unitary transformations that suppress off-diagonal matrix elements, driving the Hamiltonian towards a band-diagonal form⁷. This decoupling is illustrated in Fig. 2 on the basis of a representative chiral 3N interaction.

Computationally, the SRG evolution of two-nucleon (NN) interactions is straightforward and can be performed on a local computer. However, when evolving nuclear interactions to lower resolution, it is inevitable that higher-body interactions are induced even if initially absent. This might be considered unnatural, if nuclei could be accurately calculated based on only NN interactions. However, chiral EFT reveals the natural scale and hierarchy of many-body forces, which dictates their inclusion in calculations of nuclei and nuclear matter. In fact, the importance of 3N interactions has been demonstrated in many different calculations^{3,4}. The RG evolution of 3N forces is computationally challenging since typical dimensions of interaction matrices in a momentum-space partial-wave representation can reach about $10^4 - 10^5$. This means that the required memory for storing a single interaction matrix in double precision can reach about 40 GB. For the solution of the RG flow equations it is necessary to evaluate efficiently matrix products of this dimension. Since numerical solvers of differential equations need several copies of the solution vector for a stable and efficient evolution, a distributed storage of all matrices and vectors is mandatory. For an efficient evaluation of large matrix products we have employed a hybrid OpenMP/MPI strategy for our implementation.

3 Three-Nucleon Forces and Neutron-Rich Nuclei

Nuclei with a certain number of protons and neutrons are observed to be particularly well-bound. These closed-shell or “magic” nuclei form the basis of the nuclear shell model, which is a key computational method in nuclear physics. Exploring the formation of shell structure and how these magic configurations evolve with nucleon number towards the limits of the nuclear chart is a frontier in the physics of nuclei, and the microscopic understanding from nuclear forces represents a major challenge. The theoretical shortcomings in predicting shell structure are particularly evident in the calcium isotopes. While microscopic calculations with well-established NN forces reproduce the standard magic numbers $N = 2, 8, 20$, they do not predict ^{48}Ca as a doubly-magic nucleus when neutrons are added to ^{40}Ca , see Ref. 4. As a result, phenomenological forces have been adjusted to render ^{48}Ca doubly magic, and it has been argued that the need for these phenomenological adjustments may be largely due to neglected 3N forces. In recent works^{9,10}, we have shown that 3N forces play a decisive role in medium-mass nuclei and are crucial for the magic number $N = 28$. For the calcium isotopes, the predicted behaviour of the two-neutron separation energy S_{2n} up to ^{54}Ca is in remarkable agreement with precision mass measurements at TITAN/TRIUMF¹¹ and of the ISOLTRAP collaboration at ISOLDE/CERN using a new multi-reflection time-of-flight mass spectrometer, as shown in Fig. 3. The new $^{53,54}\text{Ca}$ masses are in excellent agreement with our NN+3N predictions and establish $N = 32$ as a shell closure. This work was published with the ISOLTRAP collaboration in Nature¹². The same valence-shell interactions have been successfully applied to study the spectroscopy of calcium isotopes and their electromagnetic moments.

Very recently we studied ground- and excited-state properties of all *sd*-shell nuclei with neutron and proton numbers $8 \leq N, Z \leq 20$, based on a set of low-resolution NN+3N interactions that predict realistic saturation properties of nuclear matter¹⁵. We focused on estimating the theoretical uncertainties due to variation of the resolution scale, the low-energy couplings, as well as from the many-body method. In Fig. 4 we compare theoretical and experimental two-neutron separation energies S_{2n} for all isotopic chains from oxygen to calcium ($Z = 8 - 20$). The theoretical calculations describe the overall experimental trends well, but in general our uncertainty bands underestimate the empirical values. The dominant uncertainties arise from the different Hamiltonians, with smaller differences between second- and third-order MBPT results. Typically the uncertainty range for S_{2n} is ~ 5 MeV. The exception are $N < Z$ isotopes, more visible in heavier elements, where the differences between second- and third-order results is comparable to the uncertainty between input Hamiltonians adding up to a total uncertainty of ~ 10 MeV.

Since it is not possible to solve the many-body problem exactly for general medium-mass nuclei, valence-space methods utilise a factorisation of nuclei into a core and valence nucleons that occupy a truncated single-particle space above the core. The interactions of particles in this valence-space are computed microscopically in many-body perturbation theory (MBPT), whereas the primary computation lies in the self-consistent evaluation of a large number of one- and two-body diagrams. The resulting effective Hamiltonian can then be diagonalised exactly, and within certain limits reproduces the exact eigenvalues. Very recently, we have achieved the first nonperturbative derivation of valence-space Hamiltonians using the in-medium SRG, with excellent results for the spectra of oxygen isotopes¹⁷. This enables a first *ab initio* connection of nuclear forces to the shell model.

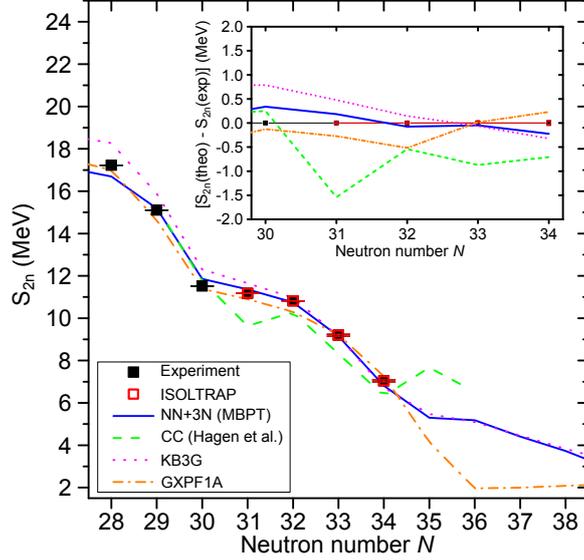


Figure 3. Two-neutron separation energy S_{2n} of the neutron-rich calcium isotopes as a function of neutron number N (see Ref. 12 for details). Shown are the new experimental results from the ISOLTRAP collaboration (red squares) in comparison with our predictions based on chiral NN+3N interactions (blue line), as well as shell-model (KB3G and GXPF1A) and coupled cluster (CC) calculations. (Figure taken from Ref. 12).

4 Quantum Monte Carlo Simulations with Chiral EFT Interactions

Quantum Monte Carlo (QMC) methods have been proven to be a very powerful tool for studying light nuclei and neutron matter¹⁸. In Refs. 19–21, we have presented first QMC calculations based on chiral NN interactions. This was not possible before due to nonlocalities in chiral EFT interactions. However, it is possible to remove all sources of nonlocality in nuclear forces up to next-to-next-to-leading order (N^2LO) in the chiral expansion. This enables us to perform auxiliary-field diffusion Monte Carlo (AFDMC) calculations for the neutron matter equation of state up to nuclear saturation density based on local leading-order (LO), next-to-leading order (NLO), and N^2LO NN interactions, as shown in Fig. 5. Our results exhibit a systematic order-by-order convergence in chiral EFT and provide nonperturbative benchmarks with theoretical uncertainties. For the softer interactions, perturbative MBPT calculations are in excellent agreement with the AFDMC results, as shown in Fig. 6.

These advances also opened up first Greens Function Monte Carlo calculations of light nuclei based on chiral NN interactions²¹. Recently, we implemented the leading 3N forces in QMC simulations of neutron matter and light nuclei^{22,23}. This paves the way for QMC calculations with systematic chiral EFT interactions for nuclei and nuclear matter, for testing the perturbativeness of different orders, and also allows for matching to lattice QCD results in a finite volume.

The QMC methods we use in our calculations treat the Schrödinger equation as a diffusion equation in imaginary time and project out the ground-state wave function from a

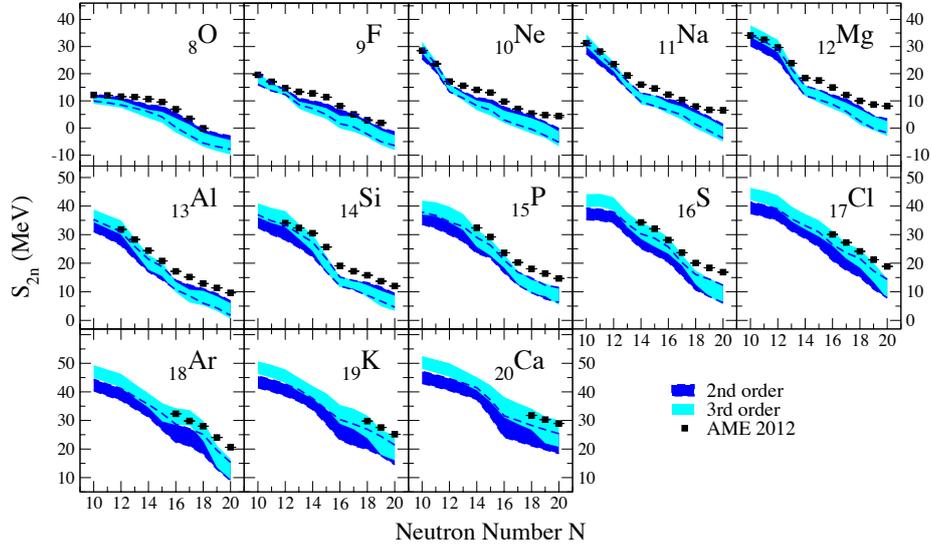


Figure 4. Uncertainty estimates for the two-neutron separation energies S_{2n} of sd -shell isotopic chains at second (blue, darker band) and third order (cyan, lighter band) in MBPT (see Ref. 15 for details) and compared to the Atomic Mass Evaluation (AME 2012)¹⁶. (Figure taken from Ref. 15).

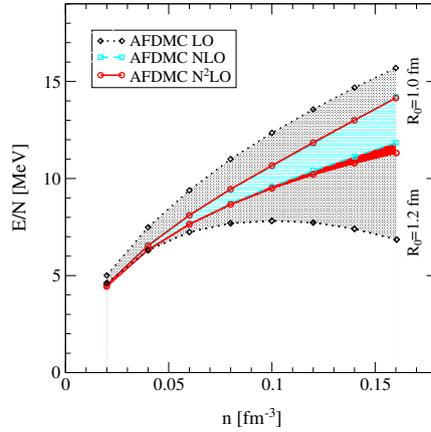


Figure 5. Neutron matter energy per particle E/N as a function of density n using AFDMC with the local chiral NN potentials at LO, NLO, and N^2 LO. The bands are obtained by varying the cutoff $R_0 = 1.0 - 1.2$ fm and the SFR cutoff $\tilde{\Lambda} = 1000 - 1400$ MeV.

trial wave function by evolving to large imaginary times. GFMC performs, in addition to a stochastic integration over the particle coordinates, explicit summations in spin-isospin space, and is thus very accurate but computationally very costly, so that one can only access particle numbers with $A \leq 12$. In contrast, AFDMC also stochastically evaluates summations in spin-isospin space and shows a better scaling behaviour at the cost of less

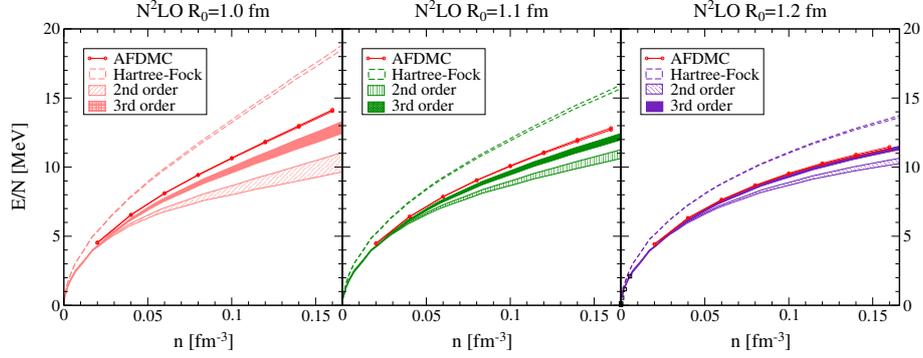


Figure 6. Neutron matter energy per particle E/N as a function of density n based on AFDMC and MBPT calculations (see Ref. 20 for details). The different panels show results using local N^2LO NN interactions with different resolution scale $R_0 = 1.0 - 1.2$ fm (the latter being the softest interaction). For the MBPT results, we show the Hartree-Fock energies as well as the energy at second and third order. The width of the bands provide a measure of the theoretical uncertainty. For the softer interactions (right panel), the perturbative calculations are in excellent agreement with the AFDMC results. (Figure taken from Ref. 20).

accuracy. We can thus simulate 66 fermions in our neutron matter calculations. For our QMC simulations we typically average over 5 – 10k walkers for several thousand time steps. Since we use independent walkers, the code is easily parallelisable and shows an excellent and almost linear scaling behaviour with the number of cores. Typically, we use 200 – 400 cores. In contrast to SRG transformations, we have only moderate memory requirements of typically 1 GB per core.

Acknowledgements

Our results could not have been achieved without an allocation of computing resources at the Jülich Supercomputing Centre. We are grateful to the John von Neumann Institute for Computing for this and for selecting the present project as 2014 Excellence Project. We also thank our collaborators S. Bogner, J. Carlson, E. Epelbaum, R. Furnstahl, S. Gandolfi, H. Hergert, J. Lynn, and A. Nogga on these projects. This work was supported in part by the ERC Grant No. 307986 STRONGINT.

References

1. S. R. Beane, W. Detmold, K. Orginos, and M. J. Savage, *Nuclear physics from lattice QCD*, Prog. Part. Nucl. Phys. **66**, 1, 2011.
2. E. Epelbaum, H.-W. Hammer, and U.-G. Meißner, *Modern theory of nuclear forces*, Rev. Mod. Phys. **81**, 1773, 2009.
3. H.-W. Hammer, A. Nogga, and A. Schwenk, *Three-body forces: From cold atoms to nuclei*, Rev. Mod. Phys. **85**, 197, 2013.
4. K. Hebeler, J. D. Holt, J. Menéndez, and A. Schwenk, *Nuclear forces and their impact on neutron-rich nuclei and neutron-rich matter*, Annu. Rev. Nucl. Part. Sci. **65**, 457, 2015.

5. S. Binder, J. Langhammer, A. Calci, and R. Roth, *Ab initio path to heavy nuclei*, Phys. Lett. B **736**, 119, 2014.
6. S. K. Bogner, R. J. Furnstahl, and A. Schwenk, *From low-momentum interactions to nuclear structure*, Prog. Part. Nucl. Phys. **65**, 94, 2010.
7. R. J. Furnstahl and K. Hebeler, *New applications of renormalization group methods in nuclear physics*, Rept. Prog. Phys. **76**, 126301, 2013.
8. K. Hebeler, *Momentum space evolution of chiral $3N$ interactions*, Phys. Rev. C **85**, 021002, 2012.
9. J. D. Holt, T. Otsuka, A. Schwenk, and T. Suzuki, *Three-body forces and shell structure in calcium isotopes*, J. Phys. G **39**, 085111, 2012.
10. J. D. Holt, J. Menéndez, and A. Schwenk, *The role of three-nucleon forces and many-body processes in nuclear pairing*, J. Phys. G **40**, 075105, 2013.
11. A. T. Gallant et al., *Breakdown of the isobaric multiplet mass equation for the $A = 20$ and 21 multiplets*, Phys. Rev. Lett. **113**, 082501, 2014.
12. F. Wienholtz et al., *Masses of exotic calcium isotopes pin down nuclear forces*, Nature **498**, 346, 2013.
13. J. D. Holt, J. Menéndez, J. Simonis, and A. Schwenk, *Three-nucleon forces and spectroscopy of neutron-rich calcium isotopes*, Phys. Rev. C **90**, 024312, 2014.
14. R. F. Garcia Ruiz et al., *Ground-state electromagnetic moments of calcium isotopes*, Phys. Rev. C **91**, 041304, 2015.
15. J. Simonis, K. Hebeler, J. D. Holt, J. Menéndez, and A. Schwenk, *Exploring sd -shell nuclei from two- and three-nucleon interactions with realistic saturation properties*, arXiv:1508.05040.
16. M. Wang, G. Audi, A. H. Wapstra, F. G. Kondev, M. MacCormick, X. Xu, and B. Pfeiffer, *The AME2012 atomic mass evaluation*, Chin. Phys. C **36**, 1603, 2012.
17. S. K. Bogner, H. Hergert, J. D. Holt, A. Schwenk, S. Binder, A. Calci, J. Langhammer, and R. Roth., *Nonperturbative shell-model interactions from the in-medium similarity renormalization group*, Phys. Rev. Lett. **113**, 142501, 2014.
18. J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. E. Schmidt, and R. B. Wiringa, *Quantum Monte Carlo methods for nuclear physics*, Rev. Mod. Phys. **87**, 1067, 2015.
19. A. Gezerlis, I. Tews, E. Epelbaum, S. Gandolfi, K. Hebeler, A. Nogga, and A. Schwenk, *Quantum Monte Carlo calculations with chiral effective field theory interactions*, Phys. Rev. Lett. **111**, 032501, 2013.
20. A. Gezerlis, I. Tews, E. Epelbaum, M. Freunek, S. Gandolfi, K. Hebeler, A. Nogga, and A. Schwenk, *Local chiral effective field theory interactions and quantum Monte Carlo applications*, Phys. Rev. C **90**, 054323, 2014.
21. J. E. Lynn, J. Carlson, E. Epelbaum, S. Gandolfi, A. Gezerlis, and A. Schwenk, *Quantum Monte Carlo calculations of light nuclei using chiral potentials*, Phys. Rev. Lett. **113**, 192501, 2014.
22. I. Tews, S. Gandolfi, A. Gezerlis, and A. Schwenk, *Quantum Monte Carlo calculations of neutron matter with chiral three-body forces*, arXiv:1507.05561.
23. J. E. Lynn, I. Tews, J. Carlson, S. Gandolfi, A. Gezerlis, K. E. Schmidt, and A. Schwenk, *Chiral three-nucleon interactions in light nuclei, neutron- α scattering, and neutron matter*, arXiv:1509.03470.

Materials Science

Materials Science

Robert O. Jones

Peter Grünberg Institut PGI-1, Forschungszentrum Jülich, 52425 Jülich, Germany

E-mail: r.jones@fz-juelich.de

“Exascale” computing (10^{18} floating point operations per second) is commonly seen as the next milestone in the high performance computing world. If this goal can be reached by the year 2020 or shortly thereafter, it will continue the remarkably consistent improvement in computing capability over many years (approximately three orders of magnitude per decade). This astonishing growth has resulted in dramatic changes in all fields in the scale of problem that can be attacked, and materials science is no exception.

Computational materials science focuses on the properties of families of materials, usually with potential technological applications. Computer simulations have obvious advantages when we deal with materials that are toxic and/or radioactive, but they have other decisive benefits. Francis Crick¹ noted “if you want to study function, study structure,” and calculations of the change of the total energy of the system as a function of the atomic coordinates – the “energy surface” – allow us to determine the geometrical arrangement of the atoms. The lowest energy determines the most stable structure, and the form of the surface determines, for example, reaction paths and energy barriers. This structural information can be found *directly* from the simulation, in contrast to experimental methods of structure determination, such as x-ray and neutron diffraction, or nuclear magnetic resonance. In 30 years of density functional calculations, for example, we have moved from mapping out parts of the energy surface of single molecules as small as ozone² to performing millions of self-consistent calculations of energies and forces in systems with hundreds of atoms (see below).

I wrote not long ago³ that “materials science is one of the great beneficiaries of changes in the landscape of scientific computing,” but this proved to be an over-optimistic snapshot of a changing scene. Problems in condensed matter physics in general and materials science in particular currently play minor and diminishing roles in present-day allocations of resources of JUQUEEN compared, in particular, with elementary particle physics. Every physicist knows that lattice quantum chromodynamics (QCD) calculations are crucial to our understanding of the basic structure of matter, and its practitioners very often emphasize this “fundamental” aspect. Nevertheless, one of the most eminent of their number, Steven Weinberg, noted how far elementary particle physics is from applications of any sort⁴, in sharp contrast to materials science with its many direct technological uses. It is true that some of the programs in popular use do not scale well with increasing core number, and this remains a challenge to workers in our field.

The four applications described in this section provide good examples of the materials simulations that are possible today and represent two directions in such simulations: large-scale molecular dynamics (MD) simulations using classical representations of the forces, and density functional calculations without adjustable parameters but many fewer atoms.

Metallic glasses form a class of metastable alloys that have been studied by Brink *et al.* using molecular dynamics (MD) simulations, with particular focus on the mechanical properties, such as brittle failure. The length scale of the phenomena involved (grain boundary formation, crystalline precipitates, nucleation at interfaces, ...) demands very large simulation samples (up to 18 million atoms), and the forces in the Cu-Zr metallic glasses were described by the interatomic force field of Finnis and Sinclair. A comparable number of atoms (~ 15 million) was used by Romero and Moseler in their MD studies of friction and wear occurring when two model surfaces slide against each other or an asperity is drawn across a metal surface. The forces are approximated by an extended atom potential. Both of these classical MD simulations scale well on JUQUEEN.

Simulations on much shorter length scales, but incorporating the full electronic structure and containing no adjustable parameters, have been performed in two contexts related to future developments in computer electronics. In the first, we studied the crystallisation of an amorphous alloy of Ge, Sb, and Te (GST). The rapid and reversible transition between the amorphous and crystalline phases of chalcogenide alloys, such as GST, is the basis of “phase change memories” that are already in use in mobile phones and are favoured candidates for non-volatile random access memory (RAM) in the future. These simulations (460 atoms, up to 6 nanoseconds) are among the most extensive density functional calculations ever performed. The relativistic, spin-polarised electronic structure calculations of Popescu and Kratzer focus on the thermoelectric power in metallic trilayers of transition metal elements. This is a challenging step in the continuing development of spintronic materials, which began with the work of Albert Fert and Peter Grünberg (Nobel Prize for Physics, 2007). There is no doubt that computer simulations will continue to play an essential role in these four fields and in other areas of materials science.

References

1. Francis Crick, “*What Mad Pursuit*”, Penguin, p. 150, 1990.
2. R. O. Jones, Phys. Rev. Lett. **52**, 2002-2005, 1984.
3. R. O. Jones, Proceedings, NIC Symposium 2008, NIC Series Vol. **39**, 151-152, 2008.
4. “*Although surprises are always possible, my own main research area, elementary particle physics, has no direct applications that one can foresee.*” S. Weinberg, New York Review of Books, **56**, Number 16, 19-22, 2009.

Nanostructured Metallic Glasses: Tailoring the Mechanical Properties of Amorphous Metals

Tobias Brink, Omar Adjaoud, and Karsten Albe

Fachgebiet Materialmodellierung, Institut für Materialwissenschaft,
Technische Universität Darmstadt, Jovanka-Bontschits-Str. 2, 64287 Darmstadt, Germany
E-mail: {brink, adjaoud, albe}@mm.tu-darmstadt.de

The mechanical properties of metallic glasses cannot only be influenced by their chemical composition, but also by their nanostructure: Secondary phases in the form of precipitates, as well as a nanocrystalline-like structure in the glass are viable options to increase the plasticity of the material. We performed molecular dynamics simulations on Cu-Zr based metallic glass systems to investigate the influence of these nanostructures on the mechanical deformation.

1 Introduction

Metallic glasses (MG) are metastable materials that were first synthesised in 1960 by rapidly quenching an Au-Si alloy from the melt¹. In 1988, bulk metallic glasses (BMG) with sample thicknesses above 1 mm were realised and have enabled the use of MGs in engineering applications². These materials are usually highly alloyed to improve the glass forming ability, thereby reducing the critical cooling rate and increasing the sample thickness. They have advantageous mechanical properties, foremost their high yield strength and large elastic strain limit. One obstacle for widespread engineering applications is the brittle failure, especially under tension, of MGs at room temperature^{2,3}. The origin of this behaviour, shown in Fig. 1(a), is the localisation of the deformation in a single, dominant shear band. Unlike dislocations, which are line defects, shear bands are two-dimensional defects which are softer than the surrounding matrix, leading to shear softening and the aforementioned strain localisation. Fig. 1(b) compares dislocations with shear bands. Similar to crystalline materials, carefully engineered microstructures may offer possibilities to tailor the mechanical properties. If we take inspiration from these materials, we can imagine two main approaches for influencing shear band nucleation and propagation in the glass: (1) Secondary phases in the form of precipitates and (2) a “grain” structure in the glass. The idea behind composites consisting of a MG matrix with crystalline precipitates is illustrated in Fig. 1(c): The crystal phase can either act as an obstacle to shear band propagation or partake in the plastic deformation. It was found experimentally that a large volume fraction of crystalline phases can lead to pronounced plasticity even under tension, for example in Cu-Zr-based MGs^{4,5} (for a complete discussion of the relevant literature see Ref. 6). Despite a growing interest in crystal–glass composite materials, our understanding of the interaction between shear bands and precipitates is still incomplete. Recent simulation studies found that crystal–glass interfaces serve as nucleation sites for shear bands⁷, but the influence of crystalline precipitates on the propagation of pre-existing shear bands is not understood in detail. To further advance our understanding of composite materials, we investigated this scenario in Cu-Zr based MGs. Contrary to precipitates, which can be grown easily by thermal annealing, a grain structure in amorphous materials

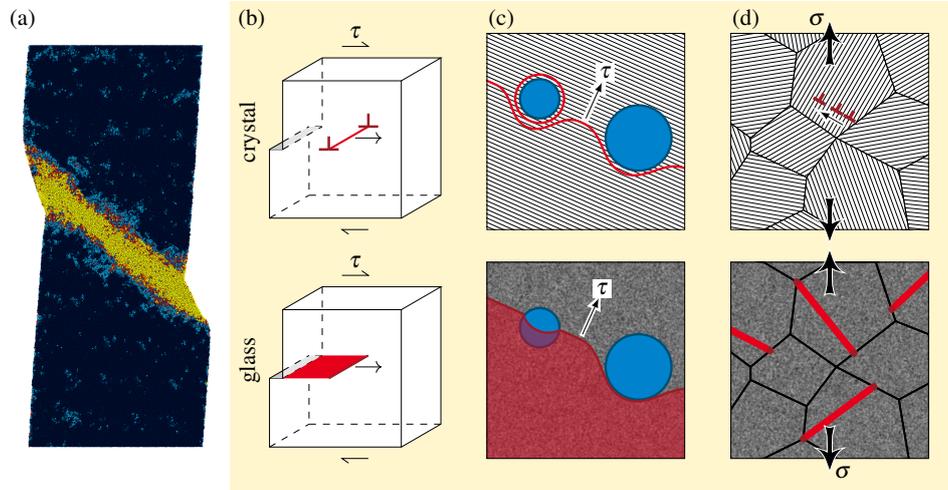


Figure 1. (a) A shear band in a $\text{Cu}_{64}\text{Zr}_{36}$ MG nanowire. The atoms are coloured according to their atomic strain. (b) Comparison between dislocations in crystals, which are line defects, and shear bands in metallic glasses, which are planar defects. Propagation of the shear band front leaves a mechanically softened region behind (red plane in the bottom picture). (c) Similar to precipitates in crystalline metals, precipitates in MGs may block or influence shear band propagation. (d) Inspired by nanocrystalline metals, nanoglasses consist of glassy grains and exhibit modified mechanical properties.

at first seems implausible due to the lack of a crystal lattice mismatch at the grain boundaries. Still, amorphous nanostructures similar to nanocrystalline samples can be obtained [cf. Fig. 1(d)]. These so-called metallic nanoglasses (NG) consist of amorphous grains connected with glass–glass interfaces and can be produced by cold compaction of glassy nanospheres which are prepared by inert-gas condensation⁸. The resulting interfaces exhibit changes in density and/or composition compared to the bulk⁸, as well as a reduced short-range order⁹. $\text{Sc}_{75}\text{Fe}_{25}$ NGs have been shown to exhibit an excellent plasticity under uniaxial tension relative to BMGs with the same chemical composition¹⁰. Molecular dynamics simulations demonstrate that the deformation behaviour of NGs can be modified by varying the volume fraction of the glass–glass interfaces^{7,11}. Nevertheless, previous structural models obtained with computer simulations are not able to correctly reproduce the density and composition changes observed experimentally. The goal is therefore to find a proper structural model for Cu-Zr based NGs and to investigate the deformation mechanisms of NGs.

2 Computational Methods

To investigate the underlying mechanisms of mechanical deformation, a computational method with atomic resolution, high performance, and scaling to millions of atoms is required. In the current work, we therefore employed molecular dynamics simulations using the software LAMMPS¹². The interatomic forces were modelled using the Finnis-Sinclair type potential for Cu-Zr by Mendeleev *et al.* which provides a realistic description of the

short range order of the MG¹³. Deformations used a strain rate of 4×10^7 /s. To stay in a regime of localised deformation despite the high strain rates, the simulations were carried out at a temperature of 50 K⁷. The biggest work packages contained up to 18 million atoms and were run on JUQUEEN using a hybrid OpenMP/MPI scheme with 8 threads \times 4096 processes. We found that LAMMPS efficiently scales to this problem size with only minor losses. The resulting data were analysed and visualised using OVITO¹⁴.

3 Metallic Glasses with Crystalline Nanoprecipitates

We investigated the interaction between an approaching shear band and a crystalline precipitate in order to understand the plasticity in crystal-glass composites⁶. To that end we

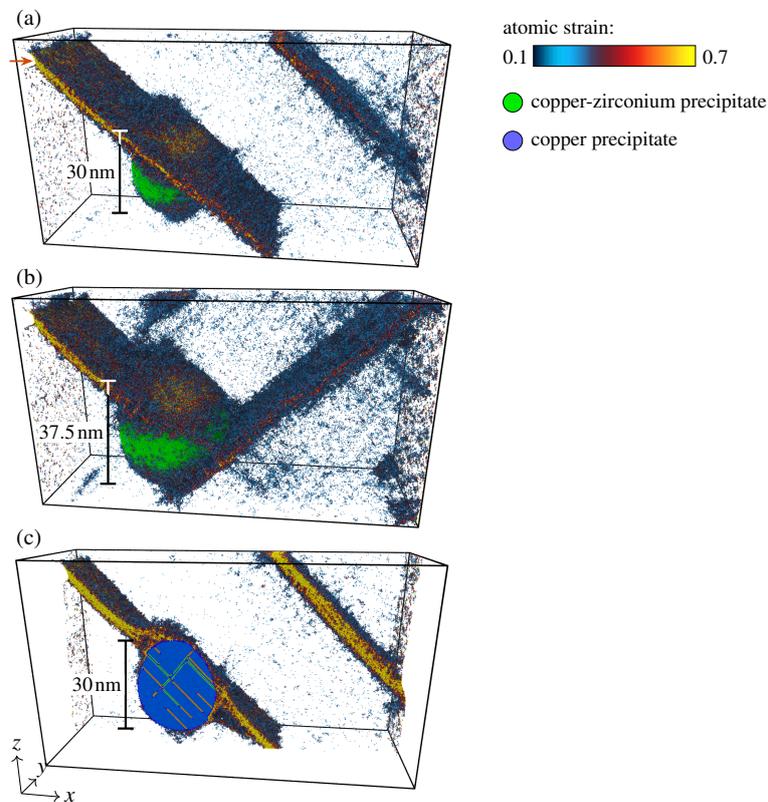


Figure 2. Interaction of shear bands with nanocrystalline precipitates. (a) A shear band originating from the notch on the left (arrow) hits a CuZr precipitate with a diameter of 30 nm. The precipitate does not deform plastically and the shear band simply wraps around the obstacle. (b) If the particle becomes bigger, the wrapping becomes impossible and the shear band is stopped. Due to the remaining stress in the system a new shear band is nucleated immediately perpendicular to the blocked one. (c) By exchanging the crystal phase for a softer one, in this case copper, we can observe slip transfer into precipitate. The snapshot shows a slice through the middle of the precipitate, the orange and green atoms represent stacking faults.

prepared glass samples and inserted precipitates of varying size and with varying distances between them. We used periodic boundary conditions in y and z direction and free surfaces in x direction. To control the origin of the shear band, which would otherwise start from the crystal-glass interface⁷ or a random point on the surface, we introduced a notch on the surface. The notch is positioned such that a shear band nucleating from it under tensile deformation in z direction will hit the precipitate. First, we inserted CuZr precipitates, using the experimentally observed B2 crystal structure¹⁵ and performed tensile tests. We observed that these precipitates only deform elastically and therefore simply present an obstacle for the shear band. This may be an artefact of using the Mendelev potential but presents a good model for precipitates with yield stress far above the glass matrix. Precipitates with a size below 25 nm to 35 nm (the exact diameter depends on the distance between the precipitates) do not stop the shear band propagation. Instead, the shear band simply wraps around the obstacle as depicted in Fig. 2(a). In contrast to the case in a crystalline matrix, this change of direction is made possible by the lack of a crystallographic lattice. Increasing the precipitate size above the critical diameter leads to behaviour as shown in Fig. 2(b): The shear band is blocked by the precipitate. The resulting build-up of stress leads to the nucleation of a second shear band at the crystal-glass interface opposite the initial shear band. Usually, shear bands move on planes of highest resolved shear stress. When bypassing an obstacle, the shear band must leave this favourable plane and temporarily move in a plane of lower resolved shear stress. At a critical precipitate size, the resolved shear stress becomes so low that it is more favourable to nucleate a new shear band. In this case shear bands may be temporarily stopped or deflected, but tensile plasticity cannot be achieved: In the end, a dominant shear band will appear and lead to failure shortly after the elastic limit.

Another case that needs to be taken into account, is a crystal phase that allows a slip transfer from the shear band in the matrix into the crystal. As a model, we chose copper precipitates. Due to greatly overestimated unstable stacking fault energies in the Mendelev potential, we switched to a different Cu-Zr potential, which cannot correctly describe CuZr in the B2 structure but reproduces the copper fcc crystal^{16,17}. The results of this simulation are presented in Fig. 2(c). Instead of blocking the shear band, the precipitate participates in the plastic deformation: Slip transfer through the particle can be observed. In composites like this, the blocking of the shear band is replaced by the participation of the crystalline phase in the plastic deformation. If the crystalline volume fraction is large enough, the macroscopic mechanical response will be a mixture of both phases. These systems hold the promise of a mix of favourable properties of crystalline metals and MGs.

4 Nanoglasses: Metallic Glasses with Microstructure

4.1 Synthesis and Microstructure

NGs are obtained by compacting a powder of amorphous nanospheres produced by inert-gas condensation⁸. In the inert-gas condensation process, gasses of the constituting elements condense into glassy nanospheres in an inert-gas atmosphere. The spheres are initially of a high temperature, an aspect of the synthesis not taken into account by previous atomistic modelling. To simulate the relaxation of the spheres at high temperature, we first cut a glassy nanosphere with a diameter of 7 nm from a $\text{Cu}_{64}\text{Zr}_{36}$ BMG. The glassy

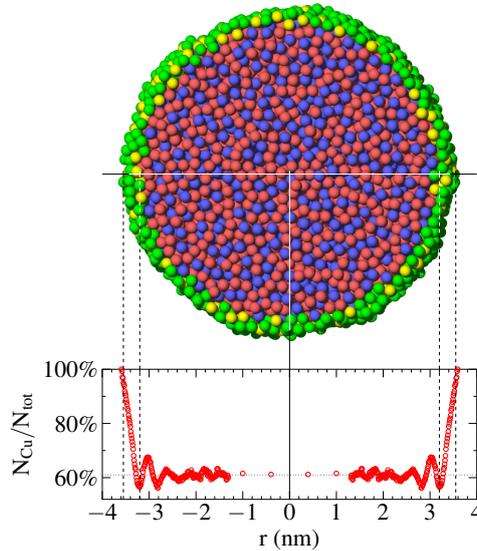


Figure 3. A nanosphere of $\text{Cu}_{64}\text{Zr}_{36}$ glass. To simulate the conditions of inert-gas condensation, the sphere was annealed slightly above T_g . This leads to a segregation of copper to the surface, where a surface shell with different composition arises. The graph shows a radial scan through the particle, plotting the percentage of copper atoms in the current shell ($N_{\text{Cu}}/N_{\text{tot}}$). The sphere has a radius of 3.5 nm. The surface layer has a width of around 0.3 nm. Red atoms are copper, blue atoms zirconium. In the surface shell, copper is depicted in green and zirconium in yellow.

nanosphere was heated above the glass transition temperature (T_g), held there until equilibrated, and then cooled down to 50 K¹⁸. We determined the composition variation across the sphere, plotted as a radial profile in Fig. 3. A Cu-rich surface layer with a width of about 0.3 nm is formed. This layer has an average composition of about 72 at.-% copper. Due to the small size of the sphere, the core is slightly copper-depleted, resulting in a composition of $\text{Cu}_{61}\text{Zr}_{39}$.

We repeated the procedure described above for several glassy nanospheres with diameters ranging from 6 nm to 8 nm to mimic the size variation in the experiment. Fig. 4 portrays the production of a NG from the powder: The differently sized nanospheres are compacted by applying a hydrostatic pressure about 5 GPa¹⁹, comparable to the experimental conditions⁸. As can be seen in Fig. 4(b), the surfaces of the nanospheres turn into interfaces and do not diffuse into the grain interior. The glass–glass interfaces retain their composition of $\text{Cu}_{72}\text{Zr}_{28}$. These effects, which are due to the surface segregation in the nanospheres, can explain the experimental observation that the interfaces in $\text{Sc}_{75}\text{Fe}_{25}$ NGs have a different composition than the grain (they are iron-depleted)⁸, a phenomenon that was not accounted for in earlier atomistic modelling^{7,9,11}.

4.2 Plastic Deformation

Using the previously described NG model, we investigated the plastic deformation mechanisms under tension²⁰. The engineering stress-strain curve is presented in Fig. 5(b). A

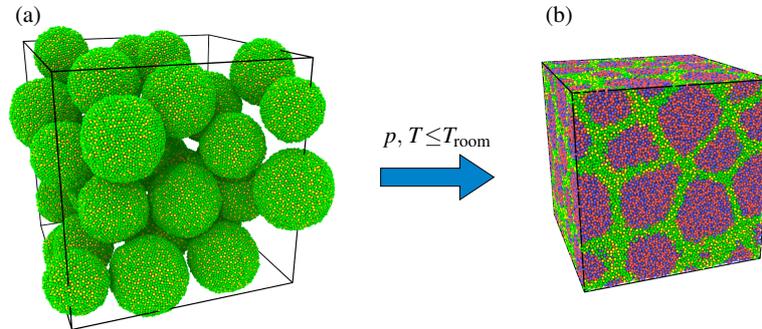


Figure 4. The glass spheres are cold compacted to produce a nanoglass. The surface shell becomes the intergranular interface as evidenced by the colouring of the atoms (same as Fig. 3). This interface retains the changed composition and exhibits a different density compared to the interior of the grain.

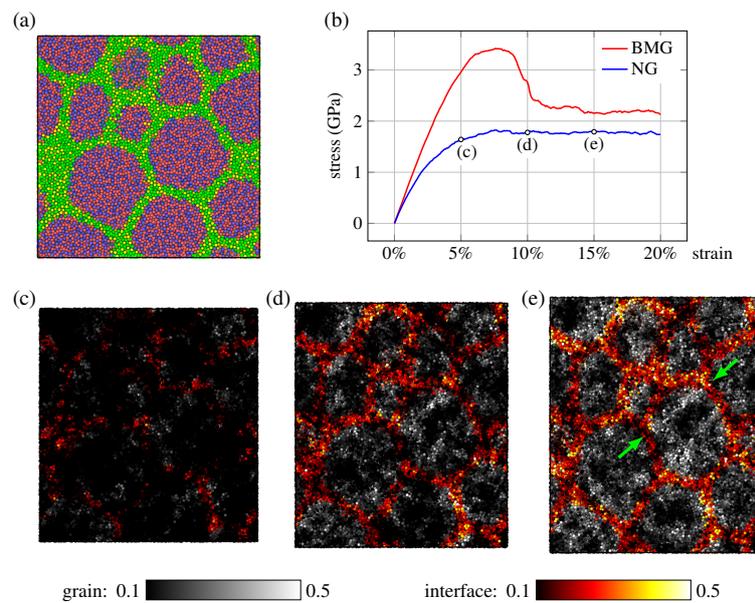


Figure 5. Tensile test of an NG. (a) A slice of the glass before deformation. The stress-strain curve is plotted in (b) together with a stress-strain curve for a bulk metallic glass of the same composition. (c–e) Atomic strain in the slice at 5%, 10%, and 15% macroscopic strain. The deformation almost immediately starts in the interfaces. The grains deform by small shear bands going through them, as indicated by the green arrows. No single shear band has moved through the complete system and become critical, thereby allowing for the finite plasticity of the material.

stress-strain curve of a BMG with the same nominal composition is also shown for comparison. In the case of the BMG, there is a stress drop which is characteristic of the formation of a single shear band, while this stress drop is not observed in the NG. The steady-state

stress remains constant until the end of the simulation at 20% total strain, which indicates that no shear softening is observable. To further get insights into the deformation behaviour at an atomic level, we inspected the atomic strain, which is displayed in Figs. 5(c–e) at 5%, 10%, and 15% macroscopic strain. The deformation starts in the interfaces, which are characterised by a reduced short-range order and therefore reduced strength. At larger strains, shear bands start appearing in the grain interiors due to local stress concentrations induced by grain movement, marked with green arrows in Fig. 5(e). The weakened interfaces also serve as nucleation sources for shear bands. Due to the random 3D arrangement of grains no single shear band becomes dominant, leading to significant plasticity, although at the cost of reduced yield strength. Other simulations already indicate that the grain size also plays a role¹¹.

5 Conclusion

The mechanical properties of MGs can be tailored via their nanostructure: Both the introduction of secondary phases and of glass–glass interfaces lead to modified shear band nucleation and propagation which influences the plastic deformation as a result. Shear bands can get blocked by precipitates only if the crystal does not deform plastically and is greater than 25 nm to 35 nm. This may lead to the nucleation of new shear bands in lieu of continued propagation of the blocked ones, but cannot prevent the formation of critical shear bands. If the precipitates deform plastically, slip transfer from the shear band into the crystal phase can take place. In these composites, the macroscopic mechanical response is a mixture of both the glass and the crystal phase. NGs are rather new materials that avoid critical shear banding much longer than homogeneous MGs due to their inherent nanostructure. The atomic strain is mostly distributed in the glass–glass interfaces, leading to a more homogeneous deformation. At later stages of the deformation, shear bands appear inside the grains, where they stay localised. These shear bands do not propagate through the sample and therefore do not become critical shear bands.

Acknowledgements

The authors gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) through project grants nos. AL 578/13-1 and AL 578/6-2, as well as the sub-project *Nanoglasses* of SPP 1594 (grant no. AL 578/15). Computing time was granted within the scope of project HDA22 by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA at Jülich Supercomputing Centre (JSC), as well as by the Gauss Centre for Supercomputing (GCS) through the NIC on the GCS share of the supercomputer JUQUEEN at JSC.

References

1. W. Klement, R. H. Willens, and P. Duwez, *Non-crystalline structure in solidified gold-silicon alloys*, *Nature*, **187**, 869–870, 1960.
2. A. Inoue, *Stabilization of metallic supercooled liquid and bulk amorphous alloys*, *Acta Mater.*, **48**, 279–306, 2000.

3. M. F. Ashby and A. L. Greer, *Metallic glasses as structural materials*, *Scr. Mater.*, **54**, 321–326, 2006.
4. Z. Liu, R. Li, G. Liu, K. Song, S. Pauly, T. Zhang, and J. Eckert, *Pronounced ductility in CuZrAl ternary bulk metallic glass composites with optimized microstructure through melt adjustment*, *AIP Adv.*, **2**, 032176, 2012.
5. F.-F. Wu, K. C. Chan, S.-T. Li, and G. Wang, *Stabilized shear banding of ZrCu-based metallic glass composites under tensile loading*, *J. Mater. Sci.*, **49**, 2164–2170, 2014.
6. T. Brink, M. Peterlechner, H. Rösner, K. Albe, and G. Wilde, *Influence of Crystalline Nanoprecipitates on Shear Band Propagation in Cu-Zr Based Metallic Glasses*, *Phys. Rev. Appl.*, submitted, 2015, arXiv:1505.00723 [cond-mat.mtrl-sci].
7. K. Albe, Y. Ritter, and D. Şopu, *Enhancing the plasticity of metallic glasses: Shear band formation, nanocomposites and nanoglasses investigated by molecular dynamics simulations*, *Mech. Mater.*, **67**, 94–103, 2013.
8. J. X. Fang, U. Vainio, W. Puff, R. Würschum, X. L. Wang, D. Wang, M. Ghafari, F. Jiang, J. Sun, H. Hahn, and H. Gleiter, *Atomic structure and structural stability of $Sc_{75}Fe_{25}$ nanoglasses*, *Nano Lett.*, **12**, 458–463, 2012.
9. Y. Ritter, D. Şopu, H. Gleiter, and K. Albe, *Structure, stability and mechanical properties of internal interfaces in $Cu_{64}Zr_{36}$ nanoglasses studied by MD simulations*, *Acta Mater.*, **59**, 6588–6593, 2011.
10. X. L. Wang, F. Jiang, H. Hahn, J. Li, H. Gleiter, J. Sun, and J. X. Fang, *Plasticity of a scandium-based nanoglass*, *Scr. Mater.*, **98**, 40–43, 2015.
11. D. Şopu and K. Albe, *Influence of grain size and composition, topology and excess free volume on the deformation behavior of Cu–Zr nanoglasses*, *Beilstein J. Nanotechnol.*, **6**, 537–545, 2015.
12. S. Plimpton, *Fast parallel algorithms for short-range molecular dynamics*, *J. Comp. Phys.*, **117**, 1–19, 1995, <http://lammps.sandia.gov/>.
13. M. I. Mendeleev, M. J. Kramer, R. T. Ott, D. J. Sordelet, D. Yagodin, and P. Popel, *Development of suitable interatomic potentials for simulation of liquid and amorphous Cu–Zr alloys*, *Philos. Mag.*, **89**, 967–987, 2009.
14. A. Stukowski, *Visualization and analysis of atomistic simulation data with OVITO – the Open Visualization Tool*, *Model. Simul. Mater. Sci. Eng.*, **18**, 015012, 2010, <http://ovito.org/>.
15. S. Pauly, J. Das, C. Duhamel, and J. Eckert, *Martensite formation in a ductile $Cu_{47.5}Zr_{47.5}Al_5$ bulk metallic glass composite*, *Adv. Eng. Mater.*, **9**, 487–491, 2007.
16. X. W. Zhou, R. A. Johnson, and H. N. G. Wadley, *Misfit-energy-increasing dislocations in vapor-deposited CoFe/NiFe multilayers*, *Phys. Rev. B*, **69**, 144113, 2004.
17. L. Ward, A. Agrawal, K. M. Flores, and W. Windl, *Rapid production of accurate embedded-atom method potentials for metal alloys*, arXiv:1209.0619 [cond-mat.mtrl-sci], 2012.
18. O. Adjaoud and K. Albe, *On the origin of compositional gradients in metallic nanoglasses*, in preparation, 2015.
19. O. Adjaoud and K. Albe, *Molecular dynamics simulations of sintering processes of nanometer-sized metallic glassy spheres*, in preparation, 2015.
20. O. Adjaoud and K. Albe, *On the origin of plasticity in metallic nanoglasses*, in preparation, 2015.

Understanding Tribology and Machining Processes through Computationally Intensive Large Scale MD

Pedro A. Romero^{1,2} and Michael Moseler^{1,3,4}

¹ Fraunhofer Institute for Mechanics of Materials IWM,
Wöhlerstraße 11, 79108 Freiburg, Germany
E-mail: michael.moseler@iwmfraunhofer.de

² Karlsruhe Institute of Technology, IAM-CMS, Kaiserstraße 12, 76131 Karlsruhe, Germany

³ University of Freiburg, Physics Department,
Hermann-Herder-Straße 3, 79104 Freiburg, Germany

⁴ Freiburg Materials Research Center, Stefan-Meier-Straße 21, 79104 Freiburg, Germany

We show in two examples how massive molecular dynamics simulations can provide a more fundamental understanding of tribology and machining processes of nanocrystalline metal surfaces. First, we show how a rigid indenter sliding over a nanocrystalline metal surface can cause surface folding during unconstrained plastic surface flow. In this case, the evolution of surface plastic flow and the subsequent folding of the initial surface are dictated by the initial orientation of the surface grains along with the lattice defects extending to the free surface. The folded surface grows into a rough chip with stratified lamellae that are identified as the precursors of wear debris.

The behaviour is very different when the counter surfaces are of similar hardness. Tribological shearing of polycrystalline metals typically leads to grain refinement at the sliding interface. Here, however, we show that nanocrystalline metals exhibit qualitatively different behaviour. Specifically, large time and space scale atomistic simulations demonstrate, that during sliding, contact interface nanocrystalline grains self-organise through extensive grain coarsening and lattice rotation until the optimal plastic slip orientation is established. Afterwards, plastic deformation is frequently confined to localised shear bands aligned with the shearing direction and emanating from voids and other defects in the vicinity of the sliding interface. These findings demonstrate the importance of surface texture and grain structure engineering to achieve ultralow wear in metals.

1 Introduction

The fundamental atomistic mechanisms responsible for friction and wear on polycrystalline metal surfaces are not accessible with conventional experimental setups and/or typical continuum level modelling frameworks like finite elements (FE) modelling or computational fluid dynamics (CFD). In order to establish a fundamental understanding of shear and wear mechanisms on metal surfaces during machining processes, large scale computationally intensive atomistic simulations of polycrystalline metal substrates are necessary. Such simulations require efficient parallel molecular dynamics (MD) algorithms¹ that can run on high performance computers². Typically, metal surfaces under tribological loads are in contact only at certain asperity peaks where most of the initial surface deformation takes place^{3,4}. The size of these asperity contacts can be of the order of tens or hundreds of nanometres, which is accessible to large scale classical MD. Therefore the mechanisms responsible for local plasticity and grain structure evolution as well as the different processes

mitigating the sliding motion can to a large extent be studied using large scale classical MD simulations. In tribology and machining processes involving surfaces with different hardness, understanding the unconstrained plastic flow (UPF) of polycrystalline metallic surfaces in sliding contact with hard asperities is crucial for the control of wear in metal-based tribological systems³⁻⁵ and the generation of surfaces in metal-working processes⁶⁻⁸. The wear resistance of metallic machine parts can be influenced by preconditioning – either by the final manufacturing steps^{9,10} (such as lapping, honing, and grinding) or by suitable running-in procedures during the initial life of a tribological contact¹¹. Traditionally, models of UPF have been based on smooth laminar material displacement⁵ – an assumption whose general validity has recently been challenged by experimental *in situ* observations and FE simulations of the disruption of laminar flow via folding on a microcrystalline copper surface sliding against a hard steel wedge¹². Unfortunately, FE lacks the description of the atomistic and crystallographic details underlying the deformation mechanisms responsible for the folding process. Consequently, the microscopic origins of UPF-induced folds in polycrystalline metallic surfaces are still not well understood.

In a tribological process involving surfaces with similar hardness, it is unclear how the initial grain structure at the shear interface will evolve and how the evolving grain structure will accommodate the shear deformation. There exists clear experimental evidence of grain refinement during sliding of conventional polycrystalline metals¹³⁻¹⁵. In contrast, grain growth has been observed in various nanocrystalline metals (Cu¹⁶, Al¹⁷, Ni¹⁸, Ni-Fe and Ni-W¹⁹). Therefore, it is still unclear whether nanocrystalline metals undergo grain refinement or coarsening. And even though tribo-induced coarsening might be expected, due to the tendency of contacting surfaces to develop highly deformable layers at the contact interface^{20,21} and the lower ductility of nanocrystalline metals relative to their microcrystalline counterparts, the exact mechanisms governing this grain growth are hardly understood.

In this communication, we use massive (millions of atoms) atomistic simulations to shed light on the basic mechanisms underlying fold formation during UPF between a hard asperity and a nanocrystalline metal surface. We demonstrate in detail that bulges in front of a moving indenter with a -45° rake angle are dictated by grains with slip systems that are favourably oriented with respect to the indenters cutting plane. For two nanocrystalline metal surfaces with similar hardness, large-scale atomistic simulations reveal local grain coarsening at the sliding interface followed by significant twinning and the formation of localised nano-shear bands (NSB).

2 Methods

2.1 Hard Asperity on Metal Sliding

The plowing of a nanocrystalline (nc) copper surface by a rigid tapered indenter (asperity) is studied by massively parallel MD¹ employing an embedded atom method (EAM) potential²⁵. The surface is modelled as a $160 \times 40 \times 30 \text{ nm}^3$ block (see Fig. 1). To construct the copper surface, a Voronoi algorithm is used to create a fully periodic system with 1584 randomly oriented grains (see top-right of Fig. 1) with grain diameters ranging from 4 to 12 nm. The constructed substrate contains ~ 15.41 million atoms. The initial Voronoi construction is first optimised with the fast inertial relaxation engine (FIRE)²⁶ before annealing

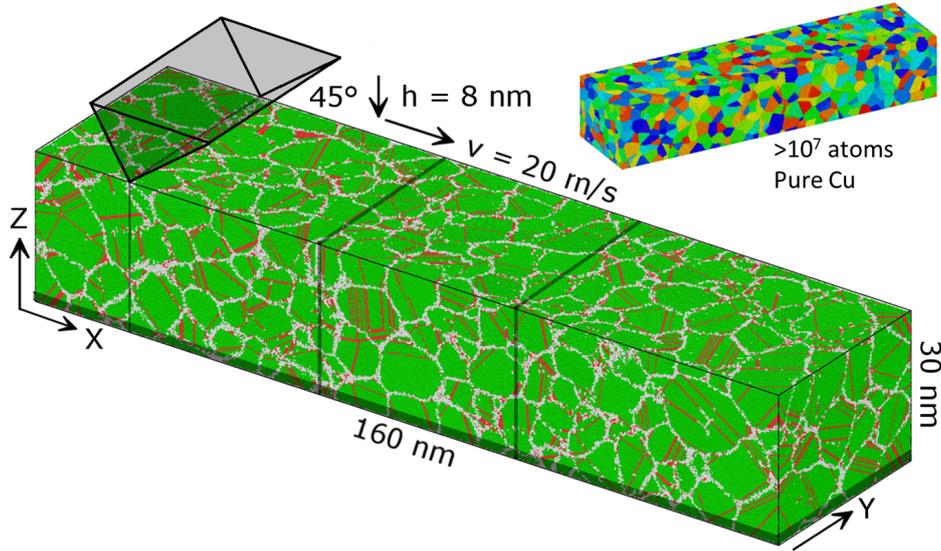


Figure 1. Molecular dynamics simulation of unconstrained surface plastic flow. (a) Copper nanocrystalline substrate and rigid indenter model. Atoms are coloured according to their local environment using a common neighbour analysis (CNA)²⁰. Green spheres represent atoms with an fcc environment; red spheres indicate stacking fault and twin boundary atoms; and gray spheres locate GBs and other defect sites. The grains are randomly oriented and contain defects characteristic of annealed copper^{21–24}.

the sample by heating and holding it at 80% of the melting temperature ($T_m = 1152$ K) for 500 ps followed by cooling it to 300 K within 50 ps and finally holding the sample at 300 K for 50 ps in order to create a system with relaxed grains and grain boundaries. The temperature changes are conducted within 10 ps respectively with a Berendsen thermostat²⁷ while holding the pressure constant at 0 GPa with a Berendsen barostat. During annealing, the system dimensions shrink to $158.8 \times 39.5 \times 29.6$ nm³. After annealing, the system acquired an approximate log normal distribution of grains ranging in size from 4 to 18 nm.

During the plowing simulations, periodic boundary conditions are employed in the lateral directions and a 0.5 nm layer of fixed atoms at the bottom of the block mimics the anchoring of the Cu nano-film to a large sample. The indenter is modelled as a nonreactive rigid wedge [with a 90° opening angle; see Fig. 1] interacting with the Cu surface via a harmonically repulsive potential with a stiffness of 32 Nm^{-1} per atom and a 0.5 nm cutoff. A 300 K dissipative particle dynamics (DPD) thermostat²⁹ with a dissipation constant of $0.1 \text{ eV ps } \text{\AA}^{-2}$ and a cutoff of 0.45 nm is applied to all atoms in the system in order to reflect ambient temperature and to reduce thermally induced grain growth^{21,28} in nc-Cu.

2.2 Metal on Metal Sliding

For sliding contact between metal surfaces with similar hardness, we employed interatomic potentials based on the embedded atom method^{30,25} that provide a reliable and efficient description of the atomic interactions for the investigated metals within the temperature

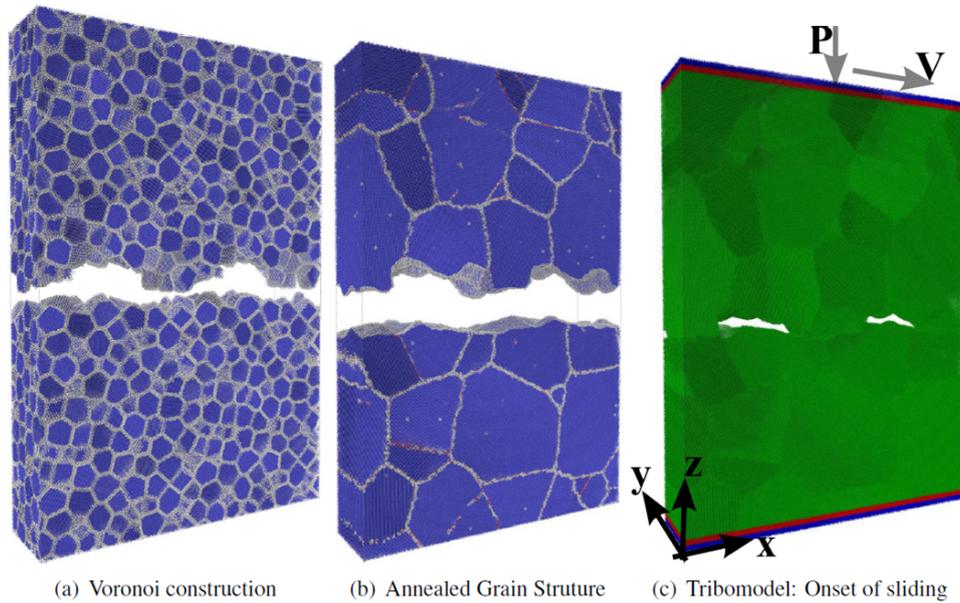


Figure 2. (a) Initial Voronoi construction. (b) Final annealed relaxed nanocrystalline counter bodies where blue/grey/red stand for bcc/nonbcc/twin iron atoms base on CNA analysis²⁰. (c) Depiction of the tribosystem at the onset of sliding where green/red/blue represent Newtonian/thermostated/rigid atoms.

range of our loading conditions³¹. Here, we present the results for a system consisting of nanocrystalline iron counter surfaces, which corresponds to a controlled volume around a pair of interacting asperities between mated nanocrystalline metal surfaces. To construct this system a Voronoi algorithm was used to generate pure nanocrystalline fully periodic iron counter bodies with randomly oriented grains having an average grain size of ~ 5 nm. Next, a region with approximately two semi-circles is cut out of the initial Voronoi construction in order to roughen the contacting surfaces (see Fig. 2(a)). The generated system has dimensions of $59.4 \times 14.9 \times 89.1$ nm³ and contains ~ 5.6 million atoms. In order to prepare the system, the initial construction is annealed by first heating the sample to $\sim 0.85T_m$ or 1500 K (for iron) for the employed interatomic potential and holding the system at this temperature for ~ 2.0 ns. Subsequently, the system is quenched to 300 K within a time of ~ 1.0 ns and then held at 300 K for an additional ~ 1.0 ns in order to ensure and verify that the annealed grains and grain boundaries are fully relaxed and stable at 300 K (see Fig. 2(b)). After applying and equilibrating a 1 GPa pressure, the system is sheared at a velocity (V) of 5 m/s for ~ 58 ns under average constant pressure (~ 1 GPa) by slightly damping (0.451 (Ev/Å)/(Å/ps)) the vertical motion of the top rigid zone³².

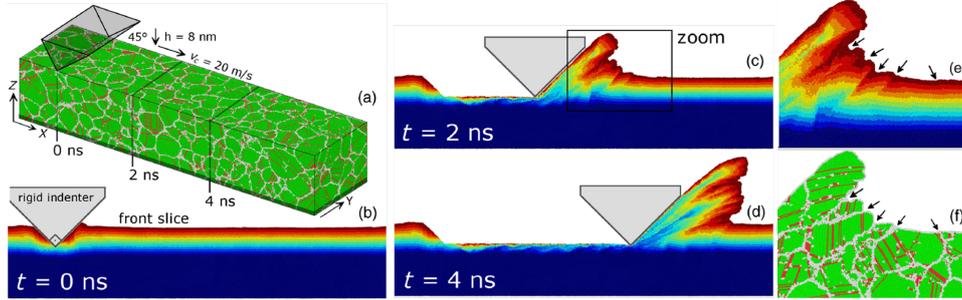


Figure 3. Molecular dynamics simulation of unconstrained surface plastic flow. (a) CNA view of Cu nanocrystalline substrate and indenter. (b)-(d) Sequence of snapshots [at the times and indenter positions indicated in (a)] showing an xz cross section of the Cu substrate. In (b)-(d) the atoms have been coloured according to their initial vertical position revealing the initiation and evolution of folding patterns during plowing. (e) Close-up of the chip shown in (c). (f) Same as (e) but coloured according to CNA.

3 Results and Discussion

3.1 Hard Asperity on Metal Sliding

In order to generate UPF on a metal nanocrystalline (nc) surface with a rigid asperity (see Fig. 3(a)), the surface is first indented to a depth of $h = 8$ nm with a velocity of $v_i = 20$ m/s. This already results in significant material pileup (see cross-sectional view in Fig. 3(b) where atoms have been coloured according to their initial vertical position). Subsequently, the nc copper is plowed at constant height (h) and a velocity of $v_c = 20$ m/s. The buildup of pronounced fold patterns is clearly observed during the evolution of UPF. Fig. 3(c) shows a snapshot of the system at $t = 2$ ns. The folding leads to the formation of a chip with a rough front face [Fig. 3(d)]. Similarities to the micron-scale folds of Ref. 12 are apparent.

The mechanisms governing the formation of surface folds can be understood by inspecting the microstructure of the copper film³⁴. A close-up of the chip in Fig. 3(c) (marked by a black frame) is shown in Fig. 3(e) along with additional arrows that mark the notches between the surface protrusions. The observed notches occur consistently at lattice defects (i.e. grain boundaries (GB)) as exhibited by the common neighbour analysis in Fig. 3(f) for the same atoms in Fig. 3(e). Clearly, the arrows coincide with gray zones in Fig. 3(f) representing GBs that separate crystalline areas (green zones).

Fig. 3(e) also shows that the grains at the surface bulge outwards towards the free surface to different degrees. The geometric form of the indenter with a -4° rake angle suggests that an optimal crystal orientation for outward flow contains a slip system whose slip direction coincides with the rake face and motion of the indenter. In order to quantify the ability of a grain to support a certain shear deformation, an atomic stress projection factor (ASPF) is defined as follows. First, for each atom with an fcc environment the projections $m_s = \cos \varphi_s \cdot \cos \lambda_s$ are calculated for the $s = 1, \dots, 12$ fcc slip systems. Here, λ_s are the angles between the normal of the indenter's rake face and the normals of the $\{111\}$ slip planes of the atomic fcc environment. φ_s denote the angles between chip flow direction (reflecting the upwards material transport along the -45° rake angle) and the $\langle 110 \rangle$ slip directions of the atomic fcc environment of the grain. For each grain, the ASPF

is defined as the maximum m_s of the 12 fcc slip systems similar to the Schmid factor for uniaxial deformation. Indeed, surface grains with higher ASPF tend to exhibit higher degrees of bulging towards the surface³⁴.

At some arrows in Fig. 3(e), discontinuities in the strata of the initial-z colour scheme can be discerned at GBs which may indicate GB sliding³⁵ as one mechanism that plays a role in the initiation of bulging and fold formation in nc metals. As shown in Fig. 3(f), the generated chip consists of coarsened larger grains, which form and merge from the start of the plowing. The underlying mechanisms are lattice reorientation and grain boundary migration. Both processes, bulging and coarsening, are evidently linked with the motion of dislocations and their interaction with the GBs. While dislocation-assisted GB migration may generally occur in deforming materials³⁶ and contribute to the deformation of polycrystalline materials by shear coupling³⁷, it can also contribute by clearing the path for plastic slip towards the surface in favourably oriented grains and by straightening the GBs parallel to suitably oriented slip planes.

3.2 Metal on Metal Sliding

For surfaces with similar hardness, the interaction between the counter metals is very different relative to hard asperities sliding on metal surfaces. The key stages of deformation for metal on metal sliding contact are presented in Fig. 4. Initially, the relative motion is primarily accommodated via surface grain boundary sliding between the grains forming the initial contact interface [see the bottom panel in Fig. 4(a)]. Accordingly, the starting grain structure is not significantly altered during the initial 4 ns of sliding.

The corresponding evolution of the encountered shear resistance $\tau(t)$ and of the ratio of non-bcc to bcc atoms $R(t)$ are presented in Fig. 5(a). During the first stage, the shear stress $\tau(t)$ sharply increases [see Fig. 5(a) between arrows a and b] with only a brief relaxation event (at 0.85 GPa between 1 and 2 ns) resulting from the flattening of small surface obstacles. At about 9 ns, $\tau(t)$ reaches a maximum of 1.75 GPa due to asperity interlocking, which is accompanied by moderate grain refinement near the initial sliding interface. The formation of new grains and grain boundaries is detected by the steady increase in the non-bcc to bcc atom ratio $R(t)$ in Fig. 5(a) between 2 and 9 ns. This grain refinement and the accompanying shear resistance enhancement is characteristic of a Hall-Petch-type mechanism.

Continued asperity interlocking from 4 to 9 ns [between Figs. 4(a) and 4(b)] leads to grain rotation and plastic deformation which in turn causes the counter bodies to cold weld [see Fig. 4(b)]. In the next stage of sliding (10 to 19 ns), the grains in the vicinity of the sliding zone gradually coarsen and merge, eventually forming a large grain that extends over the whole horizontal length of the simulation box [compare the top panels of Figs. 4(b) and 4(c)]. In a Hall-Petch fashion, this coarsening period is characterised by a marked decrease in $\tau(t)$ [see Fig. 5(a) between arrows b and c].

The sliding mechanism changes drastically between 19 and 38 ns, during which the sliding motion becomes confined within narrow shear zones spanning the entire horizontal distance of the system [see the bottom panels of Figs. 4(d) and 4(e)]. These NSBs are partly composed of twin boundaries [red zones indicated by the arrow in the middle panels of Fig. 4(d) and 4(e)], which propagate vertically within the coarsened grain. The evolution of this twinning phase is captured by the increase and decrease in $R(t)$ displayed in

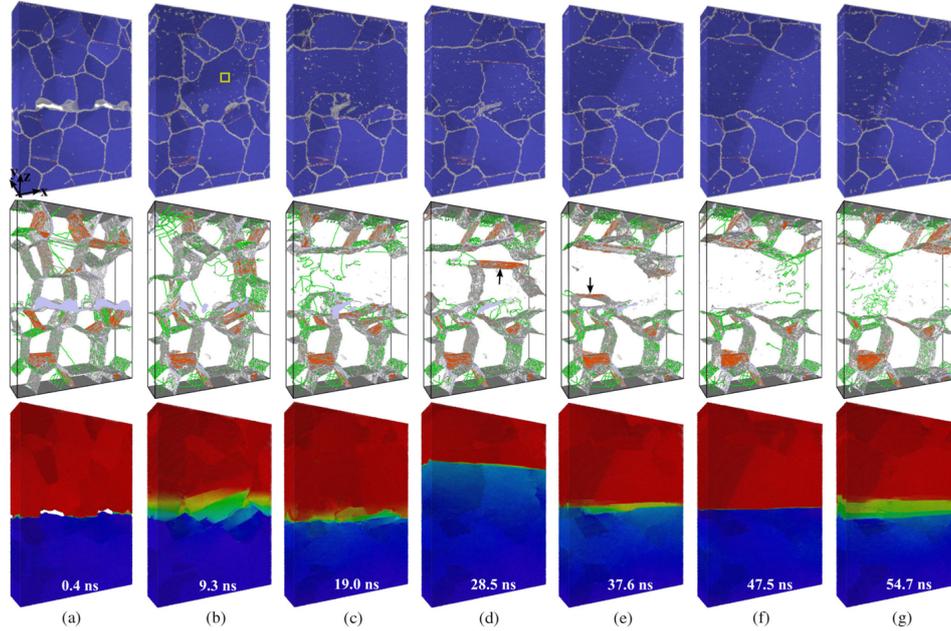


Figure 4. Snapshots of the evolution of sliding contact between two nanocrystalline iron tribo-partners. Each column presents the instantaneous state of the grain structure (top panel), the grain boundaries and dislocations (middle panel), and the corresponding sliding interface (bottom panel) at selected sliding times. The grain structure in the top row is exposed by performing a CNA²⁰ using a 3.46 cutoff radius (with blue, white, and red standing for bcc, non-bcc, and twin boundary atoms). The middle row shows the simultaneous state of the extracted dislocations^{39,40} and twin boundaries (with green and gray lines representing dislocations with $1/2\langle 111 \rangle$ and non- $1/2\langle 111 \rangle$ Burgers vectors and with gray and red zones marking the location of voids and twin boundaries). In order to correlate the structural dynamics with the way the system distributes the sliding induced velocity gradient over the system, the strain rates are visualised in the bottom row by colouring the atoms according to their velocity along the sliding x-direction after every 0.2 ns on a scale from 0 m/s (blue) to 5 m/s (red). As sliding progresses, the counter bodies cold weld and establish a coarsened grain layer where plasticity concentrates and the sliding interface localises.

Fig. 5(a) between arrows c and e. Additionally, this interval contains the lowest values of the shear resistance $\tau(t)$ [see Fig. 5(a) between $22 < t < 30$ ns] which is consistent with a low Peierls stress of twinning dislocations.

Interestingly, the first minimum in $\tau(t)$ at $t \sim 22$ ns is accompanied by a realignment of the bcc $\langle 111 \rangle$ close-packed direction within the coarse-grained layer along the direction of sliding. This is quantified in Fig. 5(b) by the change in the angle formed between the sliding x-direction and the $\langle 111 \rangle$ lattice direction within the coarsening grain inside the region marked with a yellow rectangle in the top panel of Fig. 4(b). The starting misalignment of about 35° at $t \sim 8$ ns is progressively reduced until virtually a complete alignment with the direction of sliding is reached at ~ 22 ns after which the orientation is maintained for the rest of the sliding time. These alignment changes can also be observed in the insets of Fig. 5(b) which show the atomic structure in the coarsened grain region for an observer looking along the direction of sliding.

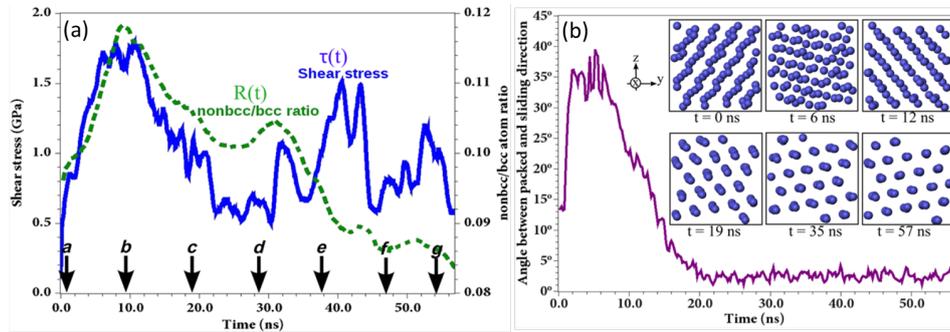


Figure 5. (a) Evolution of the shear stress [$\tau(t)$] showing the materials resistance along the direction of sliding and the progression of the ratio of non-bcc to bcc atoms [$R(t)$] indicating the general reduction in the number of grains and grain boundary volume. The arrows recall the instances corresponding to the snapshots in Fig. 4. (b) Evolution of the bcc packed direction within the coarsened centre grain. The graph shows the angle between the bcc $\langle 111 \rangle$ packed direction and the sliding x-direction. The close-up insets show states of the crystal structure within the coarsened grain while looking in the direction of sliding. Clearly, the coarsened grain region strongly aligns its bcc packed direction along the direction of sliding.

4 Concluding Remarks

In conclusion, the large-scale MD simulations presented here demonstrate how computationally intensive atomistic simulations using high performance computing facilities can lead to fundamental understanding in tribology and machining processes. For instance our massive MD simulations for a hard asperity sliding on a metal surface shows that differently oriented slip systems between neighbouring surface grains along with lattice defects (i.e. GB) dictate the degree of material bulging towards the surface and the shape and size of the resulting surface folds. The chosen simulation conditions favour dislocation-mediated processes that are by no means restricted to nanoscale grains but are rather well documented for texture formation in conventional polycrystalline metals⁴¹. Therefore, plasticity-induced bulging around surface extending lattice defects (e.g., GBs) can be regarded as the elementary process underlying fold formation in polycrystalline metals, which represents an important mechanism for the generation of lamellar wear debris (as suggested by nano-wear experiments).

As shown by the large-scale spatial (~ 5.6 million atoms) and temporal (58 ns) atomistic simulations for metal on metal sliding contact, pure nanocrystalline metallic surfaces tend to undergo grain coarsening at the contact interface during extensive sliding plastic shear. Once the interface grains cold weld, they self-organise by coarsening through grain boundary migration and by simultaneously reorienting the lattice of the coarsening grains until the optimal plastic slip direction is aligned with the sliding direction. Subsequently, the relative sliding motion is accommodated through heavily localised NSBs and twin boundary planes located within the coarsened textured interface grains. This grain coarsening and realignment in conjunction with the reduction of frictional resistance exemplifies the tendency of tribologically driven systems to self-organise during a running-in phase in order to create more easily deformable material layers at the sliding interface^{42,43}.

Acknowledgements

We acknowledge funding support from the Deutsche Forschungsgemeinschaft (DFG), the Robert Bosch GmbH, and the European Commission under Contract No. NMP.2010.2.5-1.263335 (MultiHy). Simulations were conducted on the supercomputer JUROPA and JUQUEEN at the Jülich Supercomputing Centre (JSC). Visualisation was mainly carried out with OVITO^{39,40}.

References

1. S. J. Plimpton, *Fast parallel algorithms for short-range molecular dynamics*, J. Comp. Phys. **117**, 1–19, 1995.
2. <http://www.fz-juelich.de>
3. F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids*, Clarendon Press, Oxford, 1950.
4. F. P. Bowden, A. J.W. Moore, and D. Tabor, *The ploughing and adhesion of sliding metals*, J. Appl. Phys **14**, 80, 1943.
5. K. L. Johnson, *Contact mechanics and the wear of metals*, Wear **190**, 162, 1995.
6. P. L. B. Oxley, *The Mechanics of Machining: An Analytical Approach to Assessing Machinability*, Halsted Press, New York, 1989.
7. P. A. Romero, G. Anciaux, A. Molinari, and J.-F. Molinari, *Friction at the toolchip interface during orthogonal nanometric machining*, Model. Simul. Mater. Sci. Eng **20**, 055007, 2012.
8. P. A. Romero, G. Anciaux, A. Molinari, and J.-F. Molinari, *Insights into the thermo-mechanics of orthogonal nanometric machining*, Comput. Mater. Sci. **72**, 116, 2013.
9. P. Berlet, M. Dienwiebel, and M. Scherge, *The effect of sample finishing on the tribology of metal/metal lubricated contacts*, Wear **268**, 1518, 2010.
10. P. Iglesias, M. D. Bermdez, W. Moscoso, and S. Chandrasekar, *Influence of processing parameters on wear resistance of nanostructured OFHC copper manufactured by large strain extrusion machining*, Wear **268**, 178, 2010.
11. Peter J. Blau, *On the nature of running-in*, Tribol. Int. **38**, 1007, 2005.
12. N. K. Sundaram, Y. Guo, and S. Chandrasekar, *Mesoscale folding, instability, and disruption of laminar flow in metal surfaces*, Phys. Rev. Lett. **109**, 106001, 2012.
13. A. Emge, S. Karthikeyan, and D. Rigney, *Wea* **267**, 562, 2009.
14. P. Stoyanov, P. A. Romero, T. T. Järvi, L. Pastewka, M. Scherge, P. Stemmer, A. Fischer, M. Dienwiebel, and M. Moseler, Tribol. Lett. **50**, 67, 2013.
15. A. Zhilyaev, S. Swaminathan, A. Pshenichnyuk, T. Langdon, and T. McNelley, J. Mater. Sci. **48**, 4626, 2013.
16. K. Zhang, J. R. Weertman, and J. A. Eastman, Appl. Phys. Lett. **87**, 061921, 2005.
17. M. Jin, A. Minor, E. Stach, and J. W. Morris Jr., Acta Mater. **52**, 5381, 2004.
18. X. Z. Liao, A. R. Kilmametov, R. Z. Valiev, H. Gao, X. Li, A. K. Mukherjee, J. F. Bingert, and Y. T. Zhu, Appl. Phys. Lett. **88**, 021909, 2006.
19. B. Boyce and H. A. Padilla II, Metall. Mater. Trans. A **42**, 1793, 2011.
20. J. D. Honeycutt, and H. C. Andersen, *Molecular dynamics study of melting and freezing of small Lennard-Jones clusters*, J. Phys. Chem. **91**, 4950, 1987.

21. P. A. Romero, T. T. Järvi, N. Beckmann, M. Mrovec, and M. Moseler, *Coarse graining and localized plasticity between sliding nanocrystalline metals*, Phys. Rev. Lett. **113**, 036101, 2014.
22. M. Chandross and E. A. Holm, *Measuring grain junction angles in discretized microstructures*, Metall. Mater. Trans. A **41**, 3018, 2010.
23. D. P. Field, L. T. Bradford, M. M. Nowell, and T. M. Lillo, *The role of annealing twins during recrystallization of Cu*, Acta Mater. **55**, 4233, 2007.
24. D. Farkas, E. Bringa, and A. Caro, *Annealing twins in nanocrystalline fcc metals: A molecular dynamics simulation*, Phys. Rev. B **75**, 184111, 2007.
25. Y. Mishin, M. J. Mehl, D. A. Papaconstantopoulos, A. F. Voter, and J. D. Kress, *Structural stability and lattice defects in copper: Ab initio, tight-binding, and embedded-atom calculations*, Phys. Rev. B **63**, 224106, 2001.
26. E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch, *Structural relaxation made simple*, Phys. Rev. Lett. **97**, 170201, 2006.
27. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *Molecular dynamics with coupling to an external bath*, J. Chem. Phys. **81**, 3684, 1984.
28. E. A. Holm, and S. M. Foiles, *How grain growth stops: A mechanism for grain-growth stagnation in pure materials*, Science **328**, 1138, 2010.
29. R. D. Groot, and P. B. Warren, *Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation*, J. Chem. Phys. **107**, 4423, 1997.
30. G. Simonelli, R. Pasianot, and E. Savino, Mater. Res. Soc. Symp. Proc. **291**, 567, 1993.
31. M. Müller, P. Erhart, and K. Albe, J. Phys. Condens. Matter **19**, 326220, 2007.
32. L. Pastewka, S. Moser, M. Moseler, B. Blug, S. Meier, T. Hollstein, and P. Gumbsch, Int. J. Mat. Res. **99**, 1136, 2008.
33. T. Soddemann, B. Dünweg, and K. Kremer, Phys. Rev. E **68**, 046702, 2003.
34. N. Beckmann, P. A. Romero, D. Linsler, M. Dienwiebel, U. Stolz, M. Moseler, P. Gumbsch, *Origins of Folding Instabilities on Polycrystalline Metal Surfaces*, Phys. Rev. Applied **2(6)**, 06400, 2014.
35. D. Farkas, A. Frseth, and H. van Swygenhoven, *Grain boundary migration during room temperature deformation of nanocrystalline Ni*, Scr. Mater. **55**, 695, 2006.
36. Y. Cheng, D. Weygand, and P. Gumbsch, *Simulation of small-angle tilt grain boundaries and their response to stress*, Comput. Mater. Sci. **45**, 783, 2009.
37. J. W. Cahn, Y. Mishin, and A. Suzuki, *Coupling grain boundary motion to shear deformation*, Acta Mater. **54**, 4953, 2006.
38. J. Schiotz, F. D. D. Tolla, and K. W. Jacobsen, Nature **London**, 391, 5611998.
39. A. Stukowski, Model. Simul. Mater. Sci. Eng. **18**, 015012, 2010.
40. A. Stukowski, V. V. Bulatov, and A. Arsenlis, Model. Simul. Mater. Sci. Eng. **20**, 085007, 2012.
41. D. Helm, A. Butz, D. Raabe, and P. Gumbsch, *Microstructure-based description of the deformation of metals: Theory and application*, JOM **63**, 26, 2011.
42. M. Godet, Wear **100**, 437, 1984.
43. Y. Berthier, M. Godet, and M. Brendle, Tribol. Trans. **32**, 490, 1989.

Memory Effect in Crystallisation of Amorphous $\text{Ge}_2\text{Sb}_2\text{Te}_5$

Robert O. Jones^{1,2}, Janne Kalikka³, and Jaakko Akola^{4,5}

¹ Peter Grünberg Institut PGI-1 and JARA-HPC, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: r.jones@fz-juelich.de

² German Research School for Simulation Sciences, Forschungszentrum Jülich and RWTH Aachen University, 52425 Jülich, Germany

³ Singapore University of Technology and Design, Singapore 487372
E-mail: janne.kalikka@gmail.com

⁴ Department of Physics, Tampere University of Technology, FI-33101, Tampere, Finland
E-mail: jaakko.akola@tut.fi

⁵ COMP Centre of Excellence, Department of Applied Physics, Aalto University, FI-00076 Aalto, Finland

The rate-limiting process in phase change (PC) optical memories is the extremely rapid (nanosecond time scale) crystallisation of nanosized amorphous “marks” in a polycrystalline layer. Our knowledge of the amorphous and ordered structures of Ge/Sb/Te and Ag/In/Sb/Te alloys has improved significantly in recent years and has led to plausible pictures for the transition between them. Nevertheless, the simulation of the actual crystallisation process is complicated by the need to study large numbers of atoms over time scales that are difficult to attain, even with modern supercomputers. We have performed a series of density functional/ molecular dynamics (DF/MD) simulations on a sample of the prototype PC material $\text{Ge}_2\text{Sb}_2\text{Te}_5$ (GST-225) with 460 atoms. Simulations at 500 K, 600 K, and 700 K have been performed for up to 600 picoseconds in samples where crystallisation was promoted by fixing the structure of a crystalline “seed”, and for over 4 ns in four samples at 600 K without constraints. A comparison of the last four simulations shows a striking memory effect where crystallisation is favoured in an amorphous sample that had previously been crystallised.

1 Introduction

Phase change (PC) materials are chalcogenide alloys (usually Se, Te) that switch very rapidly between the amorphous (a-) and crystalline (c-) phases. They are used extensively in rewritable high-density data storage, especially in optical recording [Digital Versatile Disc (DVD), Blu-ray Disc]¹. Information is stored as rows of nanosized amorphous marks in a polycrystalline layer and accessed via the different optical or electrical properties of the two phases. The most common materials are GeTe-Sb₂Te₃ pseudobinary compounds and Sb-Te binary compounds with small amounts of In, Ag, and/or Ge. Recrystallisation in the two groups is strikingly different, as shown in Fig. 1: In the first [Group 1, Fig. 1(b)] it proceeds mainly via nucleation inside the marks, in the second [Group 2, Fig. 1(c)] via crystal growth from the rim¹. In this report we restrict ourselves to the former.

Close collaboration between theory (density functional/ molecular dynamics simulations) and experiment (particularly x-ray diffraction and EXAFS) has shown that the amorphous structure of $\text{Ge}_2\text{Sb}_2\text{Te}_5$ (GST-225) and other alloys of Ge, Sb, and Te can be

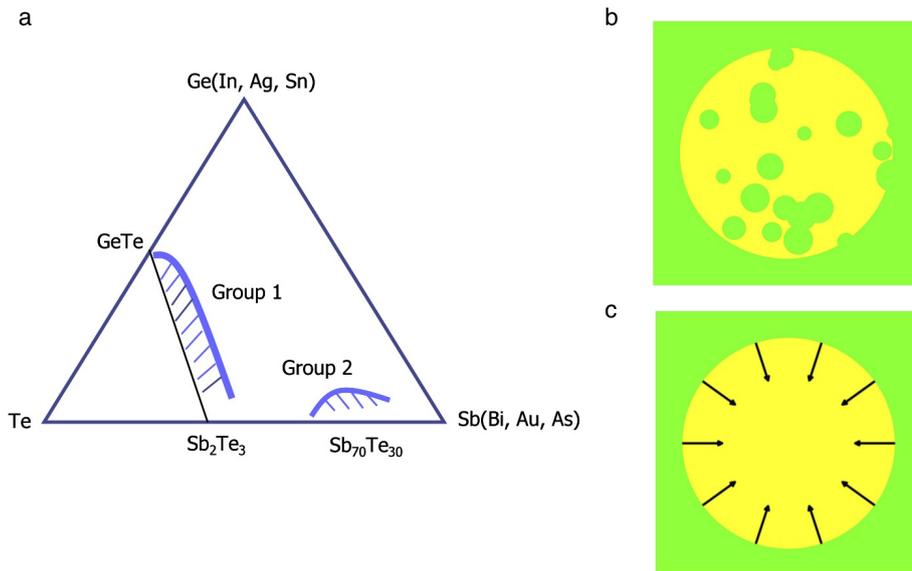


Figure 1. PC materials and their crystallisation patterns. (a) Commonly used materials for optical recording, (b) group 1: Nucleation-dominated (GST), and (c) group 2: growth-dominated recrystallisation (AIST).

characterised by “*ABAB* alternation” (*A*: Ge,Sb, *B*: Te) with four-membered *ABAB* rings being a dominant motif^{2,3}. Since this pattern is prevalent in the metastable (rock salt) crystalline structure, it is plausible that the rapid amorphous-to-crystalline transition be viewed as a re-orientation (nucleation) of disordered *ABAB* squares supported by the space provided by cavities. Knowledge of the amorphous and crystalline structures of an Ag/In/Sb/Te alloy allows us to develop a picture of crystallisation in this case as well⁴. Simulations of crystallisation in GST-225 have been carried out on samples with less than 200 atoms with interesting results^{5,6}, but small samples leave many questions unanswered. We describe here DF/MD simulations of a sample of 460 atoms. In the first simulations, crystallisation is promoted by fixing the structure of a crystalline “seed” (58 atoms, 6 vacancies) throughout. The most recent calculations, referred to below as *run0*–*run3*, there was no seed. In all cases, the densities were adjusted during the simulation to allow for the difference between the amorphous and crystalline forms.

2 Computational Methods

The combined DF/MD calculations were performed on GST-225 with the CPMD program package⁷, using the approximation of Perdew et al. (PBEsol)⁸ for the exchange-correlation energy and scalar-relativistic Troullier-Martins pseudopotentials⁹ with a plane wave cutoff energy of 20 Ry. Periodic boundary conditions were used, with a single point ($k=0$) in the Brillouin zone of the cubic unit cell, and the temperature was controlled by a Nosè-Hoover thermostat.

The initial structures of all simulations were based on the amorphous GST-225 structure of Ref. 3, which agreed well with experimental XRD and XPS measurements. Seven simulations were performed: three with a fixed crystalline seed in the centre of the simulation cell (“fixed-seed runs”)¹⁰, one with an unconstrained seed (*run0*), and three without a seed (*run1*, *run2*, *run3*). The “seeded” starting structure incorporated a $4 \times 4 \times 4$ crystallite (13 Ge, 13 Sb, 32 Te atoms, 6 vacancies, rock salt structure with lattice constant 3.0 Å) into the amorphous structure³. The crystallite had the rock salt structure that is often assumed for GST-225, with Te atoms on one sublattice, and randomly distributed Ge, Sb, and vacancies on the other¹¹. The overlapping atoms were removed, and the structure was relaxed with the atoms of the seed fixed throughout.

The initial structure for *run1* was the original amorphous structure, and those for *run2* and *run3* were derived from this by running 500 MD steps with velocity scaling for *run2*, and another 500 steps for *run3*. The fixed seed runs were performed at 500, 600, and 700 K¹⁰, while all other simulations were at 600 K. The simulation box was adjusted to the change in density ($\sim 7\%$) between the amorphous and crystalline densities by reducing the starting box size (24.629 Å) in five steps of 0.114 Å to the final size (24.06 Å), following the number of crystalline atoms.

The order in the sample was studied using several measures: First, “crystalline” atoms were identified with the aid of the order parameter of Steinhardt, Nelson, and Ronchetti¹², which has proved to be of value in discussing bond orientation order in disordered systems:

$$\bar{Q}_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^l |\bar{Q}_{lm}(i)|^2}, \quad (1)$$

where

$$\bar{Q}_{lm}(i) = \frac{1}{N_b(i)} \sum_{k=0}^{N_b(i)} Q_{lm}(k), \quad \text{and} \quad Q_{lm}(i) = \frac{1}{N(i)} \sum_{j=1}^{N(i)} Y_{lm}(\vec{r}_{ij}). \quad (2)$$

$N(i)$ is the number of neighbours for atom i , $N_b(i)$ includes the atom i and its neighbours, and $Y_{lm}(\vec{r}_{ij})$ are spherical harmonics. The first non-zero value of \bar{Q}_ℓ for cubic structures is for $\ell = 4$, and we define “crystalline” atom to be those for which $\bar{Q}_4 \geq 0.6$.

Second, the clustering of such atoms allows us to analyse the individual structures for “percolation”. If we assume a maximum bond length 3.2 Å, a cluster “percolates” if there is a path connecting an atom to its replica in the neighbouring unit cell. We have also studied the changes that occur in the (partial) pair distribution functions, the numbers of “wrong bonds” and *ABAB* squares, the electronic density of states, and the mean square displacement of the atoms.

3 Results and Discussion

The fixed seed simulations at 600 and 700 K crystallised within 600 ps¹⁰, and *run0* crystallised fully within 1.2 ns. The fixed seed run at 500 K showed signs of crystal growth but crystallisation was not complete within 600 ps. Simulations *runs1-3* showed nucleation, and varying degrees of crystal growth within 4 ns with the degree of crystallinity being more than 50% in *run1* and *run2*, and less than 50% in *run3* at that time. More details are provided in Ref. 13.

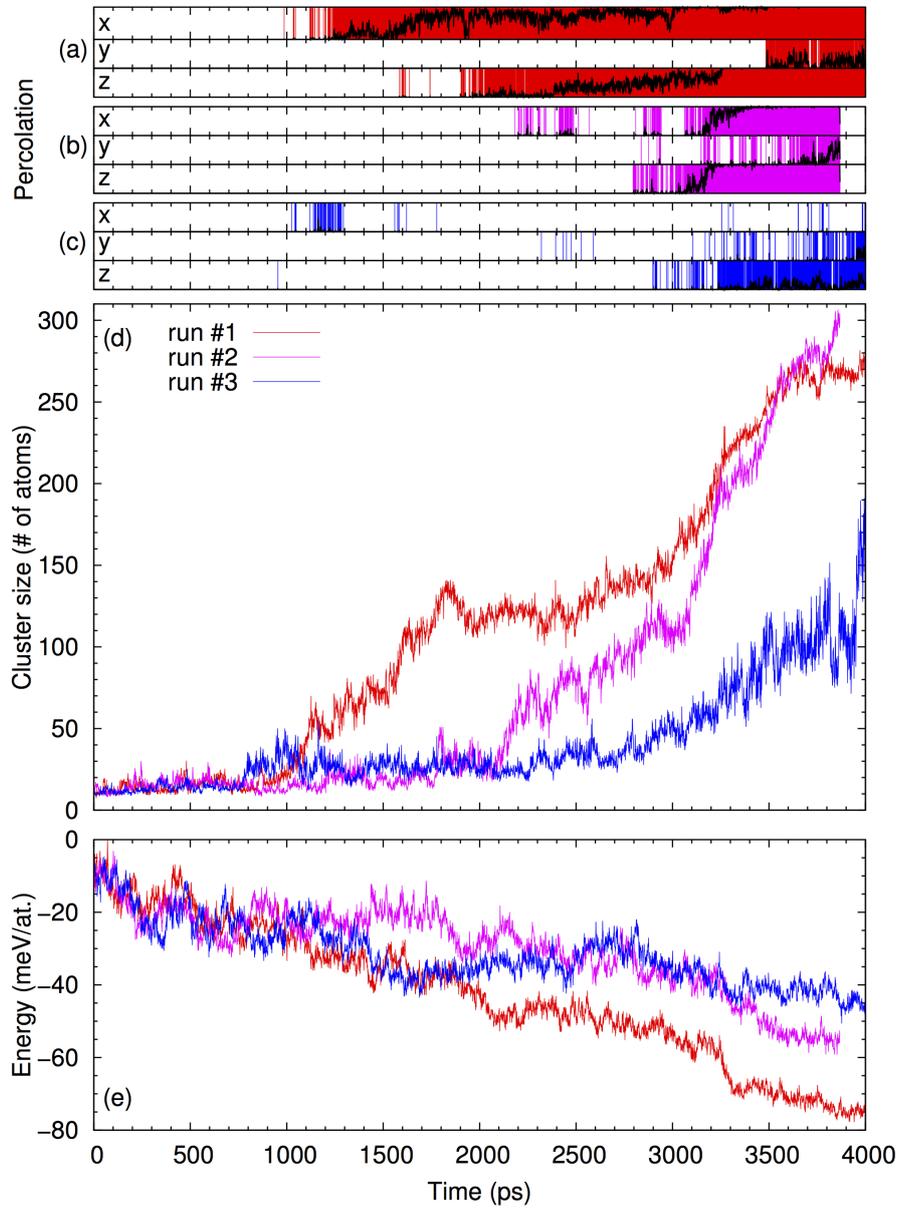


Figure 2. Three simulations (*run1* – *run3*) starting from the amorphous structure of Ref. 3. (a-c) Percolation in *x*-, *y*-, and *z*-directions. Black: fraction of percolating frames in 1 ps windows; coloured background: percolating frames; (d) size of largest cluster; and (e) total energy (normalised for box size). Red: *run1*, purple: *run2*, blue: *run3*.

The early nuclei in *runs0-3* differ significantly in size and shape. The initial nucleus in *run0* grows steadily until crystallisation is complete, while *runs1-3* have subcritical nuclei that fluctuate between 10-50 atoms. The shapes of the nuclei are often stringlike chains of crystalline atoms and isolated ABAB-squares, but more spherical fused blocks of ABAB-cubes (complete or incomplete) are also common. The shapes of the nuclei can differ from the classical nucleation theory picture, and the motifs common in them are AB-alternation and bond angles close to 90 degrees.

In Fig. 2 we plot the crystallisation and percolation properties, as well as the total energies, of *runs1-3*. Percolation is found in *runs0-3* well before the rapid crystallisation phase and can occur with as little as 20% crystallinity. In all these simulations, percolation is initially intermittent in one direction. Sometimes it proceeds along another direction, but in all cases it is more constant later, and extends to two and three dimensions. The rapid crystallisation phase happens during three-dimensional constant percolation in *run0*, and

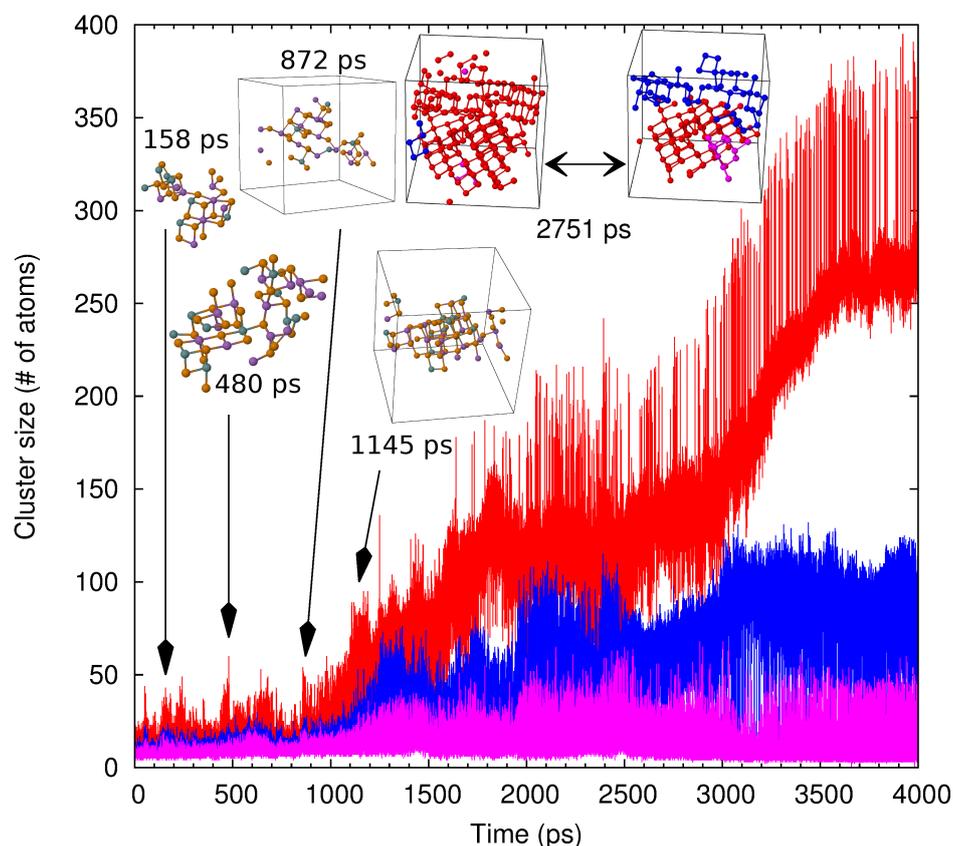


Figure 3. Evolution of the three largest crystalline nuclei of *run1* as a function of time. Red: largest cluster; blue: second largest; purple: third largest. Also shown are the largest crystalline nuclei of *run1* when the size of the largest cluster had a maximum. After ~ 2 ns, the two largest clusters (red, blue) have a fluctuating interface that affects the cluster labelling.

during two-dimensional constant percolation in *run1* and *run2*. In *run3* there is constant percolation in one direction after 2.5 ns, and in two directions just before 4 ns.

The evolution of the size of the crystalline nuclei in *run1* is shown in Fig. 3. The crystal growth rate is low until the critical nucleus size is reached, and then increases until the fraction of crystalline atoms reaches 30-50%. The crystallisation rate in *run2* during the percolating phase is of the order of 1 m/s, which is less than recently measured with differential scanning calorimetry (DSC) at 600 K (2.5-3.0 m/s)¹⁴. After this stage, the structure collapses rapidly until all atoms are crystalline. There is one main crystallite in *run2*, and crystallisation proceeds to completion. In the case of multiple crystallites where none is dominant in size (as in *run1*), crystallisation is incomplete.

Crystallisation takes place within a few nanoseconds. In *run0* there appears to be memory of the ordered seed, since the distribution of bond directions in the simulation cell was higher along the simulation cell axes. The time required for the onset of crystallisation is also lower than in *runs1-3*, and the rate of crystallisation after onset is higher. Full crystallinity occurred after 1.2 ns in *run0*. After this time, *run1* was at 25%, *run2* at 15%, and *run3* at 20% crystallinity and there was no percolation in any of the three. In all simulations that crystallised fully, the final structures have “wrong” bonds that do not exist in the undistorted Yamada model¹¹. The number of wrong bonds is highest in the amorphous structures, and decreases during crystallisation. Nevertheless, the final structures have 0.1 – 0.3 wrong bonds/atom.

4 Concluding Remarks

Crystallisation occurs in most disordered systems on time scales that are far beyond the scope of density functional based calculations. Phase change materials used as optical storage and computer memory are exceptional in that the time scale (some nanoseconds) is accessible with modern supercomputers, and the results could provide insight into crystallisation in general. Nevertheless, DF/MD simulations on a nanosecond time scale with adequate sample sizes remain a great computational challenge. Here we have studied the early stages of crystallisation in GST-225 using samples with 460 atoms over simulation times of over 4 ns. In some cases, crystallisation has been promoted by inserting a small crystallite in the cell, and we have been able to follow the process in detail in all cases. The crystalline structures contain many “wrong bonds”, and Te atoms are located on both rock salt sublattices. The commonly accepted picture of a structure with a perfect Te sublattice and random occupancy of Ge, Sb, and vacancies on the other is a substantial oversimplification.

The memory of the order in *run0* remained after the seed was no longer evident, and the acceleration of crystallisation in an amorphous bit with a crystalline history should be a consideration in the design of future memory cells. The speed of crystallisation found in *run0* and the alignment of the final structure along the axes of the simulation cell were not found in simulations *runs1-3*. The starting configuration of the latter in each case was the amorphous structure found in Ref. 3, the only differences being in the initial velocity distributions at 600 K. These apparently small differences lead to large differences in the crystallisation process. Nevertheless, they involve many more atoms and much longer times than all previous DF studies of the process, and they raise questions about the findings of most, in particular those of a recent study⁶ that indicated that all cavities segregate

to the amorphous-crystalline boundary, leaving a cavity-free crystal. There is no evidence for this mechanism in any of our simulations.

The simulation trajectories show crystallisation directions that are unrelated to the axes of the simulation cell, and they show subcritical phases with ordered clusters of 10-50 atoms prior to the onset of crystallisation. Although the onsets differed, the speed of crystallisation from the subcritical phases is similar in all cases, and two of the simulations show multiple clusters and “polycrystalline” final structures. The final structure in all cases show the existence of low-frequency, localised vibrational modes that are not present in the original amorphous structures. Percolation initiates the rapid phase of crystallisation and is coupled to the directional p-type bonding in metastable GST-225. The apparent acceleration of crystallisation in a sample with an ordered history could lead to improved optical storage media.

Each simulation of 4 ns with a time step of ~ 3 fs requires over 1.3 million self-consistent density functional calculations of energies and forces for a sample of 460 atoms. Even using a computer with the performance of JUQUEEN, this is a task that demands both great computing *and* human resources.

Acknowledgements

We acknowledge gratefully computer time provided by the JARA-HPC Vergabegremium on the JARA-HPC partition of the supercomputer JUQUEEN at Forschungszentrum Jülich and for time granted on the supercomputer JUROPA at the Forschungszentrum Jülich.

References

1. M. Wuttig and N. Yamada, *Phase-change materials for rewriteable data storage*, Nature Mater. **6**, 824, 2007 and references therein.
2. J. Akola and R. O. Jones, *Binary alloys of Ge and Te: order, voids, and the eutectic composition*, Phys. Rev. Lett. **100**, 205502, 2008; *Density functional study of amorphous, liquid and crystalline $Ge_2Sb_2Te_5$: Homopolar bonds and/or AB alternation?* J. Phys.: Condens. Matter **20**, 465103, 2008.
3. J. Akola, R. O. Jones, S. Kohara, S. Kimura, K. Kobayashi, M. Takata, T. Matsunaga, R. Kojima, and N. Yamada, *Experimentally constrained density-functional calculations of the amorphous structure of the prototypical phase-change material $Ge_2Sb_2Te_5$* , Phys. Rev. B **80**, 020201(R), 2009, and references therein.
4. T. Matsunaga, J. Akola, S. Kohara, T. Honma, K. Kobayashi, E. Ikenaga, R. O. Jones, N. Yamada, M. Takata, and R. Kojima, *From local structure to nanosecond recrystallization dynamics in $AgInSbTe$ phase change materials*, Nature Mater. **10**, 129, 2011.
5. J. Hegedüs and S. R. Elliott, *Microscopic origin of the fast crystallization ability of $Ge-Sb-Te$ phase-change memory materials*, Nature Mater. **7**, 399, 2008.
6. T. H. Lee and S. R. Elliott, *Structural role of vacancies in the phase transition of $Ge_2Sb_2Te_5$ memory materials*, Phys. Rev. B **84**, 094124, 2011; *Ab initio computer simulation of the early stages of crystallization: Application to $Ge_2Sb_2Te_5$ phase-change materials*, Phys. Rev. Lett. **107**, 145702, 2011.

7. CPMD version 3.15. © IBM Corporation (1990-2012), MPI für Festkörperforschung, Stuttgart (1997-2001) (<http://www.cpmd.org/>).
8. J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhao, and K. Burke, *Restoring the density-gradient expansion for exchange in solids and surfaces*, Phys. Rev. Lett. **100**, 136406, 2008.
9. N. Troullier and J. L. Martins, *Efficient pseudopotentials for plane-wave calculations*, Phys. Rev. B **43**, 2006, 1991.
10. J. Kalikka, J. Akola, J. Larrucea, and R. O. Jones, *Nucleus-driven crystallization of amorphous $Ge_2Sb_2Te_5$: A density functional study*, Phys. Rev. B **86**, 144113, 2012.
11. N. Yamada, *Erasable phase-change optical materials*, MRS Bulletin **21**, 48, 1996.
12. P. J. Steinhardt, D. R. Nelson, and M. Rochetti, *Bond-orientational order in liquids and glasses*, Phys. Rev. B **28**, 784, 1983.
13. J. Kalikka, J. Akola, and R. O. Jones, *Simulation of crystallization in $Ge_2Sb_2Te_5$: A memory effect in the canonical phase-change material*, Phys. Rev. B **90**, 184109, 2014.
14. J. Orava, A. L. Greer, B. Gholipour, D. W. Hewak, and C. E. Smith, *Characterization of supercooled liquid $Ge_2Sb_2Te_5$ and its crystallization by ultrafast-heating calorimetry*, Nature Mater. **11**, 279, 2012.

Anisotropic Magneto-Thermopower in $M/\text{Co}/M$ ($M = \text{Cu}, \text{Pd}, \text{Pt}$) Trilayer Systems

Voicu Popescu and Peter Kratzer

Faculty of Physics and CENIDE, University Duisburg-Essen,
Lotharstr. 1, 47048 Duisburg, Germany
E-mail: {voicu.popescu, peter.kratzer}@uni-due.de

Ab initio investigations performed within the framework of the spin-polarised relativistic Korringa-Kohn-Rostoker Green's function method show that the longitudinal thermopower in the metallic $M/\text{Co}/M$ systems (with M a non-magnetic transition metal) depends in a significant manner on the relative orientation (in-plane versus out-of-plane) of the sample magnetisation. As thermoelectric analogue of the conventional anisotropic magneto-resistance (AMR), the amplitude of this magneto-thermopower (MTP) signal is related to the asymmetry of the AMR around the Fermi energy. This asymmetry can be sizable, and thus the MTP accordingly large, even if the AMR itself is small. An enhancement of the MTP based on this understanding opens the possibility of implementing efficient spin read-out thermoelectric devices based on a single ferromagnetic layer. Our calculations reveal a rather non-intuitive behaviour of the MTP in the investigated systems, with the Pd/Co/Pd trilayer exhibiting an extremely small valued MTP, not necessarily expected from the monotonous increase of the spin-orbit coupling strength for the different heterostructure partners $M = \text{Cu}, \text{Pd}, \text{Pt}$.

1 Introduction

Electron spin injection, manipulation and detection are the primary targets¹ of any application within the rich field of spintronics². Exploiting the spin dependence of thermoelectric phenomena triggered, in addition, the fast development of a new branch, termed spin caloritronics^{3,4}. Successful reports on thermally driven spin injection⁵ and on detecting a magnetic response of the longitudinal thermopower in multilayered metallic nanowires⁶ and tunnelling junctions⁷ suggested the possibility of a practical magneto-thermoelectric device implementation.

While at the core of spintronics lies the electron spin degree of freedom, this, in turn, is coupled to the translational degree of freedom by virtue of the spin-orbit interaction. This ever present phenomenon is at the origin of a magnetisation orientation dependence of various ground state, excited state or transport properties. For the latter case, the most common example is the conventional anisotropic magneto-resistance (AMR)^{8,9}: The resistivity (or, equivalently, the conductivity) of a ferromagnetic sample depends on the relative orientation between its magnetisation direction and that of the current. In the general terms of the electronic transmission probability $\mathcal{T}_{\vec{M}}(E)$ at energy E , this quantity will depend on the magnetisation \vec{M} of the sample and the difference

$$\Delta\mathcal{T}(E) = \mathcal{T}_{\vec{M}_1}(E) - \mathcal{T}_{\vec{M}_2}(E) \quad (1)$$

between two different magnetic configurations \vec{M}_1 and \vec{M}_2 will be a direct measure of the AMR. In the case of a sizeable $\Delta\mathcal{T}$, it becomes obvious that one can use this effect as a magnetic read-out device. While in normal transition metal alloys and heterostructures this

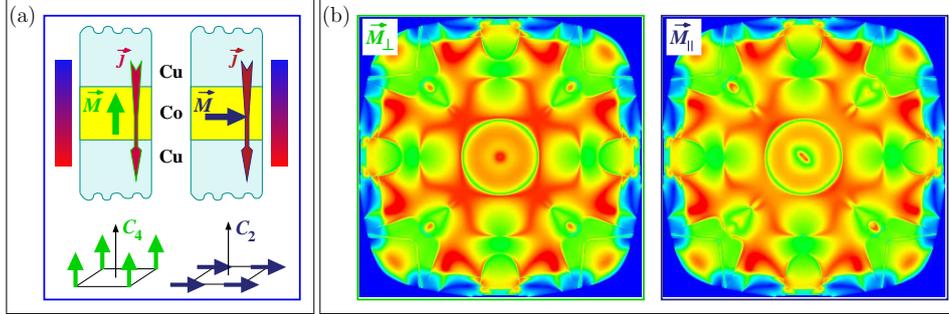


Figure 1. Illustration of spin-orbit coupling induced magnetic anisotropy manifestation in a thin Co slab embedded in Cu: (a) sketch of the geometry of the setup considered for the longitudinal anisotropic magnetothermopower. The current density vector and the temperature gradient are oriented perpendicular to the interface, while the magnetisation is either perpendicular (green) or parallel (dark blue) to it. (b) Calculated electronic transmission probability $\mathcal{T}(\vec{k}_{\parallel}, E)$, Eq. 4, for the two different orientations of the magnetisation, \vec{M}_{\perp} and \vec{M}_{\parallel} , in a 4 monolayer thick Co slab embedded in Cu. The transmission channels \vec{k}_{\parallel} cover the full 2D Brillouin zone and the energy argument E is fixed. Note that the lowering of the symmetry illustrated in panel (a), from \vec{M}_{\perp} configuration to \vec{M}_{\parallel} is accordingly reflected in the two transmission profiles of panel (b).

effect is usually very small, such a setup has been successfully used to generate a typical low/high resistance spin-valve signal similar to that of the giant magneto-resistance (GMR) systems in tunnelling junctions¹⁰. These so-called Dirac devices have the major technological advantage of providing a spin resolution through a single ferromagnetic layer.

As an alternative, the read-out of such a logical element could also be accomplished by making use of the magneto-thermopower (MTP) effect that has been measured for various systems^{11–14}. This effect, the thermoelectric equivalent of the AMR, can be described completely analogously: under a temperature gradient $\vec{\nabla}T$, the generated longitudinal thermopower $S(T)$ (Seebeck effect) takes on different values, depending on the angle between the sample magnetisation and $\vec{\nabla}T$:

$$\Delta S(T) = S_{\vec{M}_1}(T) - S_{\vec{M}_2}(T) \neq 0, \quad (2)$$

where \vec{M}_1 and \vec{M}_2 have the same meaning as above.

An illustration of such a setup is provided in Fig. 1(a). The longitudinal MTP signal is generated by a single magnetic layer (here, without restricting the generality, taken to be Co) sandwiched between two non-magnetic leads (here Cu). The magnetic anisotropy occurring in this system is caused by the symmetry break-off at the interface, as illustrated, for the case of a quadratic two-dimensional (2D) lattice in Fig. 1(a). Upon this junction one can apply either an electric field or a thermal gradient *perpendicular* to the Co/Cu interface and detect an AMR or an MTP response by flipping the magnetisation direction from perpendicular (\vec{M}_{\perp}) to parallel (\vec{M}_{\parallel}) to the interface. We have primarily focused on the correlation between the MTP and the conventional AMR, establishing a practical way to maximise the former.

Our first results for Cu/Co/Cu(001) trilayers¹⁵, obtained by *ab initio* spin-polarised relativistic calculations that account directly, in a parameter free way, for the simultaneous manifestation of spin polarisation and spin-orbit coupling¹⁶, have shown that both AMR

and MTP have a common origin: the spin-orbit coupling induced anisotropy in the electronic transmission probability. The calculated MTP was found to have a large value over a wide range of temperatures and Co thickness values. We could establish, in addition, that the MTP is maximised by an enhanced asymmetry in the AMR energy dependence. Indeed, rather than mapping the Fermi-surface alone, the thermoelectric phenomena depend on transmission channels extending over a finite interval around the Fermi energy.

These investigations are hereby extended to the $M/\text{Co}/M$ (111)-oriented trilayers ($M = \text{Cu}, \text{Pd}, \text{Pt}$). Our results reveal that, while Pt/Co/Pt and Cu/Co/Cu systems exhibit a sizeable MTP, the Pd-based trilayers are characterised by a very small anisotropic thermoelectric response, in spite of having a rather large AMR, comparable to that of Pt/Co/Pt. The outcome of our study is two-fold: The MTP-AMR relationship is generally applicable, however, a large MTP signal is indeed not conditioned by the size but rather by a stark energy dependence of the AMR.

2 Theoretical Aspects and Computational Details

2.1 Geometry Setup and Self-Consistent Potentials

Our investigations on the trilayer systems involve two steps. First, we perform self-consistent electronic structure calculations for both magnetic configurations \vec{M}_\perp and \vec{M}_\parallel using a spin-polarised relativistic implementation¹⁶ of the screened Korringa-Kohn-Rostoker Green's function method (KKR-GF)¹⁷⁻¹⁹, combined with the decimation technique²⁰ for 2D-periodic systems.

A prototypical trilayer system consists of two half-infinite M leads with an interaction region inserted in-between, all sharing the same in-plane 2D periodic lattice. The interaction region contains the n -monolayer (ML) thick Co slab and up to 10 MLs of the heterostructure partner M on both sides of the Co slab. We have considered one of the metals adopting the in-plane lattice constant of its partner, thus being subject to epitaxial strain and manifesting a tetragonal distortion. With the in-plane lattice fixed, the vertical inter-layer separation was determined using continuum elasticity. In our calculations we use spherical potentials in the atomic sphere approximation (ASA) determined within the local spin-density approximation²¹. An angular momentum cut-off of $l_{max} = 3$ was taken for the Green's function expansion and $2l_{max} = 6$ for the charge density.

2.2 Transport Properties

In a second step, the self-consistent potentials are used as input for a calculation scheme that relies on a relativistic implementation²² of the Landauer-Büttiker formula within the KKR-GF method²³. This provides the electron transmission probability $\mathcal{T}_{\vec{M}}(E)$ at energy E for the magnetic configuration \vec{M} :

$$\mathcal{T}_{\vec{M}}(E) = \frac{1}{A_{2D-BZ}} \int_{2D-BZ} d^2\vec{k}_\parallel \mathcal{T}_{\vec{M}}(\vec{k}_\parallel, E) , \quad (3)$$

obtained as an integral over the 2D Brillouin zone (2D-BZ) of the single-channel, ballistic transmission probability $\mathcal{T}_{\vec{M}}(\vec{k}_\parallel, E)$ for a 2D vector \vec{k}_\parallel ²³. This latter quantity can be determined from the z -component of the current operator \underline{J} and the Green's function matrices

$\underline{G}^{22,23}$ (omitting the configuration index \vec{M}):

$$\mathcal{T}(\vec{k}_{\parallel}, E) = \sum_{n \in L} \sum_{n' \in R} \text{Tr} \left[\underline{J}_n^T(E) \underline{G}_{nn'}(\vec{k}_{\parallel}, E) \underline{J}_{n'}(E) \underline{G}_{nn'}^{\dagger}(\vec{k}_{\parallel}, E) \right] \quad (4)$$

for atoms n (n') belonging to the left (right) lead, where the underscored symbols designate matrices in the relativistic representation (j, m_j).

The energy dependent transmission given in Eq. 3 is used to determine the transport coefficients²⁴:

$$L_{\vec{M}}^{(\alpha)}(T) = \int dE \left[\left(-\frac{\partial f_0}{\partial E} \right) \mathcal{T}_{\vec{M}}(E) (E - \mu)^{\alpha} \right], \quad (5)$$

with $f_0(E, T, \mu)$ the Fermi-Dirac distribution function at energy E , temperature T and chemical potential μ . Knowledge of these quantities allows one to calculate the Seebeck coefficient $S_{\vec{M}}(T)$ as²⁴:

$$S_{\vec{M}}(T) = -\frac{1}{eT} \frac{L_{\vec{M}}^{(1)}}{L_{\vec{M}}^{(0)}}. \quad (6)$$

The formalism sketched here only considers the electronic temperature, neglecting the electron-phonon or electron-magnon scattering. As these effects are less important at low temperatures, we have restricted our investigations to a temperature range below 350 K and focused on phenomena solely related to the changes in the electronic structure induced by modifying the magnetic configuration of the system. We finally note that the dependence of the self-consistent potentials and the transport properties on the magnetisation direction is implicitly taken into account by solving the Dirac equation in the local frame of reference at each atomic site and then applying unitary rotations to transform into the global frame of reference with the quantisation axis parallel to the z -axis¹⁶.

2.3 Relationship between the MTP and AMR

The way in which a transmission probability profile $\mathcal{T}_{\vec{M}}(E)$ influences the sign and size of the Seebeck coefficient at finite temperatures can be understood on the basis of Eq. 6. In this equation, a temperature increase effectively extends the integration range, by increasing the non-zero width of $\mathcal{T}_{\vec{M}}(E)(\partial f_0/\partial E)$. Because of the $(E - E_F)$ term, the numerator can be seen as a centre of mass of $\mathcal{T}_{\vec{M}}(E)(\partial f_0/\partial E)$ ²⁵. Consequently, both sign and value of $S_{\vec{M}}(T)$ will be sensitive even to small changes in the numerator's integrand below or above E_F . This behaviour illustrates the simple path towards maximising the Seebeck coefficient of a sample, by achieving a strongly asymmetric $\mathcal{T}_{\vec{M}}(E)$ around E_F .

The AMR in typical transition metal systems, alloys or heterostructures, is rather small, not exceeding few percent. We show in Fig. 1(b) the full transmission profiles $\mathcal{T}_{\vec{M}}(\vec{k}_{\parallel}, E)$ calculated for $\vec{M} \equiv \vec{M}_{\perp}$ and $\vec{M} \equiv \vec{M}_{\parallel}$ configurations over the whole 2D-BZ in a Cu/Co₄/Cu(001) trilayer. The energy argument corresponds to $E = E_F - 0.14$ eV, chosen in such a way that $\Delta\mathcal{T}(E)$ attains a local maximum. We make two important observations: (i) the $\mathcal{T}_{\vec{M}}(\vec{k}_{\parallel}, E)$ contours clearly reflect the lowering of the symmetry when switching from \vec{M}_{\perp} to \vec{M}_{\parallel} configuration (from four-fold to two-fold rotation axis); and

(ii) it is rather difficult to note quantitative differences between the two quantities. In fact, explicit evaluation of Eq. 3 led to a $\Delta\mathcal{T}(E)$ difference in the range of 4 %.

The latter is an important observation in establishing a direct connection between the magnitude of the MTP signal and the asymmetry of the AMR around the Fermi energy. On its basis, one can make the reasonable approximation that the zero-order transport coefficient, Eq. 5, has a negligible dependence on \vec{M} , such that $L_{\vec{M}_\perp}^{(0)}(T) \simeq L_{\vec{M}_\parallel}^{(0)}(T)$. It follows, from the definition of the Seebeck coefficient, Eq. 6, that the MTP introduced in Eq. 2 can be expressed as

$$\Delta S(T) \propto \int dE (\partial_E f_0)(E - E_F) \Delta\mathcal{T}(E) , \quad (7)$$

with $\Delta\mathcal{T}(E)$ given by Eq. 1. From this equation it can be seen that, since $(E - E_F)$ is anti-symmetric about E_F , a $\Delta\mathcal{T}(E)$ of odd parity about E_F is needed to maximise the MTP. We emphasise here on the analogy with the condition for $S(T)$, which is maximised by an asymmetry of $\mathcal{T}(E)$.

2.4 Use of Computer Resources

As outlined above, the central quantity to be calculated for the electronic transport investigations is the \vec{k}_\parallel - and E -resolved transmission probability $\mathcal{T}_{\vec{M}}(\vec{k}_\parallel, E)$ given by Eq. 4. Closer inspection of this equation reveals that, amongst its two ingredients – the current density operator \underline{J}_n and the structural Green’s function matrix $\underline{G}_{nn'}(\vec{k}_\parallel, E)$ – only the latter depends on the transmission channel \vec{k}_\parallel . Since the evaluation of Eq. 4 essentially implies a trace of a matrix product, one can easily separate and optimise the integration loops required in this formula. This leads to an efficient parallelisation over the energy and \vec{k}_\parallel variables.

In spite of this convenient form and the additional advantages brought about by the screened KKR method (essentially an order- N method), the calculation of the transmission probability and of the Seebeck coefficient remains a quite tedious and computationally demanding task: (i) The integration range in Eq. 5 is, in principle, infinite; in practice, however, it can be restricted to a domain in which the derivative of the Fermi-Dirac is greater than a user-defined “zero”, a quantity which needs to be tested basically for every system; (ii) The \vec{k}_\parallel -resolved transmission $\mathcal{T}_{\vec{M}}(\vec{k}_\parallel, E)$, Eq. 4, may be a fast-varying, non-monotonous function of \vec{k}_\parallel , as seen in Fig. 1(b). Convergence tests have shown that the number of \vec{k}_\parallel -points needed to ensure the convergence of $\mathcal{T}_{\vec{M}}(E)$ at any energy in the proximity of μ has to be in the order of 10^6 .

3 MTP versus AMR in $M/\text{Co}/M$ ($M = \text{Cu}, \text{Pd}, \text{Pt}$)

Having established the link between the MTP and the AMR via Eq. 7, we have focused on the following issues: (i) is the MTP-maximisation recipe provided by Eq. 7 generally valid? (ii) does the inclusion of a material, such as Pt and Pd, with a higher spin-orbit coupling (SOC) and hence a higher AMR necessarily lead to an increase of the MTP?

We illustrate our results for the case of an $M/\text{Co}_6/M$ (111)-oriented heterostructure ($M = \text{Cu}, \text{Pd}$ and Pt) with a fixed thickness of the ferromagnetic Co layer. Fig. 2(a)

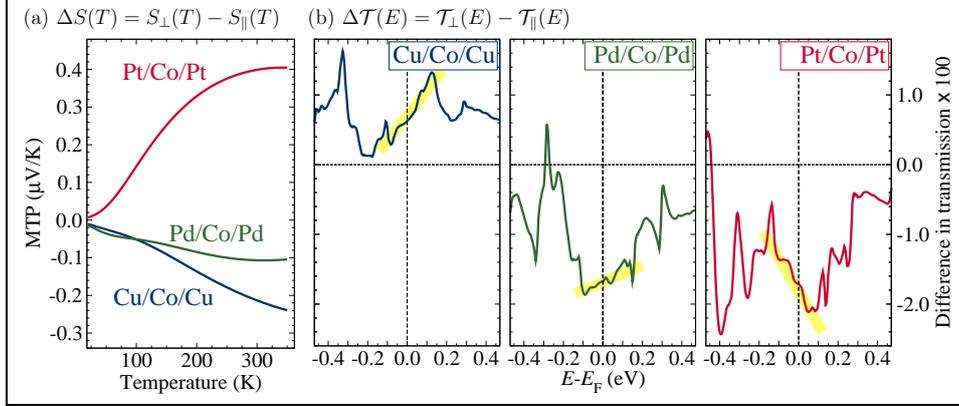


Figure 2. (a) The magneto-thermopower in the $M/\text{Co}_6/M$ (111)-oriented trilayers ($M = \text{Cu, Pd, Pt}$) and (b) the difference in the electronic transmission probability for the magnetisation oriented perpendicular and parallel to the M/Co interface in a current-perpendicular-to-the-plane (CPP) setup for the same systems, from left to right. The different slopes around the Fermi energy (taken as reference) are highlighted as guide for the eyes.

displays the calculated MTP, $\Delta S = S_\perp - S_\parallel$, in the three systems [Eq. 2 and lhs of Eq. 7], while panel (b) of the figure shows, from left to right for each M , the calculated energy-dependent AMR, $\Delta\mathcal{T}(E) = \mathcal{T}_\perp(E) - \mathcal{T}_\parallel(E)$ [Eq. 1 and the integrand of the rhs of Eq. 7]. In addition, the different slopes of $\Delta\mathcal{T}(E)$ around the Fermi energy (taken as reference) are highlighted as guide for the eyes in Fig. 2(b).

As seen in Fig. 2(a), Pt/Co/Pt and Cu/Co/Cu systems are characterised by rather large MTP values, in the order of few tenths of $\mu\text{V/K}$, negative for Cu and positive for Pt. These values of $\Delta S(T)$ need to be compared with the calculated Seebeck coefficients for the respective systems (not shown here), which lie in the range of $-5 \mu\text{V/K}$ (in Cu/Co/Cu) and $\pm 1 \mu\text{V/K}$ (in Pd/Co/Pd and Pt/Co/Pt). This effectively corresponds to an MTP response as high as $\simeq 35\%$ for the Pt/Co/Pt system. In contrast, the highest calculated AMR, found for the same heterostructure, does not exceed 2% [note the multiplicative factor 100 in Fig. 2(b)], at typical values for $\mathcal{T}(E)$ in the order of unity. In other words, using the MTP effect as a basis for a spin read-out Dirac device might be more efficient than through a conventional AMR element.

The Pd-based heterostructure stands out through its extremely small anisotropic MTP effect, as compared to the other two non-magnetic metals. This behaviour might appear counter-intuitive at a first glance, since the SOC increases as going downwards in the periodic table, $3d \rightarrow 4d \rightarrow 5d$. One notes, on the other hand, that the AMR *does* follow the expected trend, with $\Delta\mathcal{T}(E)$ increasing along the series Cu-Pd-Pt.

This apparent disagreement with what an educated guess might suggest can be relatively easily understood on the basis of the established link between the MTP and the AMR expressed by Eq. 7. It becomes clear that both sign *and* magnitude of the calculated MTP correlate with the *slope* of $\Delta\mathcal{T}(E)$ around the Fermi energy [highlighted in Fig. 2(b)], which is nothing else but the expected result from Eq. 7: since the integrand contains the product $(\partial_E f_0)(E - E_F) \Delta\mathcal{T}(E)$ with $\partial_E f_0$ of even and $(E - E_F)$ of odd symmetry with respect to E_F , the integral is zero for $\Delta\mathcal{T}(E)$ even, being maximised by

an as asymmetric as possible AMR. This analysis answers the two issues raised above: (i) numerical results confirm that the MTP-AMR connection expressed by Eq. 7 is quite general; and (ii) a large AMR is not *alone* a sufficient ingredient to ensure a maximum MTP signal. Indeed, as evidenced by the comparison of the $\Delta\mathcal{T}(E)$ profiles for Pd/Co/Pd and Pt/Co/Pt in Fig. 2(b), the two systems differ less in the magnitude of the AMR as in its *energy dependence* around the Fermi energy E_F . In conclusion, we have shown, for a series of heterostructures containing a ferromagnetic layer in a trilayer configuration with non-magnetic leads, a general recipe to enhance the MTP. This may in principle enable the implementation of efficient spin read-out thermoelectric devices based on a single ferromagnetic layer, with significant technological advantages over the, often very complex, multilayer spin valves.

Acknowledgements

This work was supported by the German Research Foundation within the Priority Program 1538 “Spin Caloric Transport (SpinCaT)”. The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA at Jülich Supercomputing Centre (JSC).

References

1. I. Žutić, J. Fabian, and S. Das Sarma, *Spintronics: Fundamentals and applications*, Rev. Mod. Phys., **76**, 323–410, 2004.
2. E. Y. Tsymbal and I. Žutić, (Eds.), *Handbook of Spin Transport and Magnetism*, Chapman and Hall/CRC Press, Boca Raton, FL, USA, 2011.
3. G. E. W. Bauer, E. Saitoh, and B. J. van Wees, *Spin caloritronics*, Nature Materials, **11**, 391–399, 2012.
4. B. Scharf, A. Matos-Abiague, I. Žutić, and J. Fabian, *Theory of thermal spin-charge coupling in electronic systems*, Phys. Rev. B, **85**, 085208, Feb 2012.
5. A. Slachter, F. L. Bakker, J. P. Adam, and B. J. van Wees, *Thermally driven spin injection from a ferromagnet into a non-magnetic metal*, Nature Physics, **6**, 879–882, 2010.
6. L. Gravier, S. Serrano-Guisan, F. Reuse, and J.-P. Ansermet, *Thermodynamic description of heat and spin transport in magnetic nanostructures*, Phys. Rev. B, **73**, 024419, 2006.
7. M. Walter, J. Walowski, V. Zbarsky, M. Münzenberg, M. Schäfers, D. Ebke, G. Reiss, A. Thomas, P. Peretzki, M. Seibt, J. S. Moodera, M. Czerner, M. Bachmann, and C. Heiliger, *Seebeck effect in magnetic tunnel junctions*, Nature Materials, **10**, 742–746, 2011.
8. J. Smit, *Magnetoresistance of ferromagnetic metals and alloys at low temperatures*, Physica, **17**, 612 – 627, 1951.
9. I. A. Campbell, A. Fert, and O. Jaoul, *The spontaneous resistivity anisotropy in Ni-based alloys*, Journal of Physics C: Solid State Physics, **3**, S95, 1970.

10. C. Gould, C. Rüster, T. Jungwirth, E. Girgis, G. M. Schott, R. Giraud, K. Brunner, G. Schmidt, and L. W. Molenkamp, *Tunneling Anisotropic Magnetoresistance: A Spin-Valve-Like Tunnel Magnetoresistance Using a Single Magnetic Layer*, Phys. Rev. Lett., **93**, 117203, 2004.
11. Y. Pu, E. Johnston-Halperin, D. D. Awschalom, and J. Shi, *Anisotropic Thermopower and Planar Nernst Effect in $\text{Ga}_{1-x}\text{Mn}_x\text{As}$ Ferromagnetic Semiconductors*, Phys. Rev. Lett., **97**, 036601, 2006.
12. P. Wiśniewski, *Giant anisotropic magnetoresistance and magnetothermopower in cubic 3:4 uranium pnictides*, Applied Physics Letters, **90**, 192106, 2007.
13. A. D. Avery, M. R. Pufall, and B. L. Zink, *Observation of the Planar Nernst Effect in Permalloy and Nickel Thin Films with In-Plane Thermal Gradients*, Phys. Rev. Lett., **109**, 196602, 2012.
14. A. D. Avery, M. R. Pufall, and B. L. Zink, *Determining the planar Nernst effect from magnetic-field-dependent thermopower and resistance in nickel and permalloy thin films*, Phys. Rev. B, **86**, 184408, 2012.
15. V. Popescu and P. Kratzer, *Large Seebeck magnetic anisotropy in thin Co films embedded in Cu determined by ab initio investigations*, Phys. Rev. B, **88**, 104425, 2013.
16. H. Ebert, D. Ködderitzsch, and J. Minár, *Calculating condensed matter properties using the KKR-Green's function method - recent developments and applications*, Rep. Prog. Phys., **74**, 096501, 2011.
17. L. Szunyogh, B. Újfalussy, P. Weinberger, and J. Kollár, *Self-consistent localized KKR scheme for surfaces and interfaces*, Phys. Rev. B, **49**, 2721–2729, 1994.
18. R. Zeller, P. H. Dederichs, B. Újfalussy, L. Szunyogh, and P. Weinberger, *Theory and convergence properties of the screened Korringa-Kohn-Rostoker method*, Phys. Rev. B, **52**, 8807, 1995.
19. K. Wildberger, R. Zeller, and P. H. Dederichs, *Screened KKR-Green's-function method for layered systems*, Phys. Rev. B, **55**, 10074–10080, 1997.
20. I. Turek, V. Drchal, J. Kudrnovský, M. Šob, and P. Weinberger, *Electronic Structure of Disordered Alloys, Surfaces and Interfaces*, Kluwer Academic Publishers, Boston, 1997.
21. S. H. Vosko, L. Wilk, and M. Nusair, *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, Canadian Journal of Physics, **58**, 1200, 1980.
22. V. Popescu, H. Ebert, N. Papanikolaou, R. Zeller, and P. H. Dederichs, *Influence of spin-orbit coupling on the transport properties of magnetic tunnel junctions*, Phys. Rev. B, **72**, 184427, 2005.
23. Ph. Mavropoulos, N. Papanikolaou, and P. H. Dederichs, *Korringa-Kohn-Rostoker Green-function formalism for ballistic transport*, Phys. Rev. B, **69**, no. 15, 125104, 2004.
24. U. Sivan and Y. Imry, *Multichannel Landauer formula for thermoelectric transport with application to thermopower near the mobility edge*, Phys. Rev. B, **33**, 551–558, 1986.
25. M. Czerner, M. Bachmann, and C. Heiliger, *Spin caloritronics in magnetic tunnel junctions: Ab initio studies*, Phys. Rev. B, **83**, 132405, 2011.

Condensed Matter

Condensed Matter

Kurt Binder

Institut für Physik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany
E-mail: kurt.binder@uni-mainz.de

In principle, “condensed matter” includes fluids as well as crystalline or amorphous solids, from “hard matter” (e.g. metals, semiconductors, insulators such as oxides and silicates) to “soft matter” (e.g. colloids, polymers, cell membranes, etc.) The latter systems are mostly dealt with in other chapters of this book (see the chapters on Computational Soft Matter Science as well as Computational Biology and Biophysics). Condensed Matter also provides the foundation of Materials Science, and studies focusing on applications of condensed matter systems typically can be found in the chapter on Materials Science, although there is certainly no sharp dividing line between these fields; e.g., the original results on atomic size contacts between aluminium nanotips (Fig. 1 of the article by M. Beck *et al.*) can be found in the journal Nature Nanotechnology and hence can be considered as “Materials Science” as well.

In the present chapter the focus is on condensed matter systems in restricted geometry and in nanoscopic confinement. The article by M. Beck *et al.* deals mostly with transport phenomena in nanosystems, although structure formation processes such as spinodal decomposition of liquids also is studied by hydrodynamic simulations. Related methods are even applied to magnetohydrodynamic phenomena in astrophysics. Thus the scope of this article ranges from the scale of single-atom wide “wires” (a problem studied in close collaboration with experimentalists) to the scale of objects in the universe. Complementary to the wide range of physics problems, also a great variety of different simulations methods and codes is used.

M. Raczkowski *et al.* describe a study of a system in between one and two spatial dimensions, namely weakly coupled Hubbard chains, applying Quantum Monte Carlo methods. For the understanding of strong electronic correlations, the Hubbard Hamiltonian is the “fruitfly model” of condensed matter physics: it describes the competition between the hopping of electrons (localised at lattice sites) from site to site and Coulomb on-site repulsion. Binding of particle hole pairs from neighbouring chains near half-filling (one electron per site) generates a finite interchain magnetic interaction, and may induce the onset of a symmetry-broken phase. One very interesting effect, demonstrated in the present study, is the confinement of magnetic quasiparticles, the so-called “spinons”, in the zero temperature limit. However, at nonzero temperature a crossover takes place, charge fluctuations take over, which in principle show up in optical spectra.

Optical spectra of nanoscopic systems are also treated in the article by M. Rohlfing, but the approach taken is completely different. Two touching carbon nanotubes are studied by many-body perturbation theory methods, i.e. the local density approximation for electronic structure calculations is extended in terms of a quasiparticle Hamiltonian for the exciton (electron + hole quasiparticle). But the resulting Bethe-Salpeter equation is approximated in terms of a perturbative treatment. A crucial aspect for the spectral shifts is that the

full screened Coulomb interaction (being inhomogeneous, anisotropic, and non-local!) is accounted for. By these methods, it is shown that the shift in the optical spectra of two aligned touching carbon nanotubes is very sensitive to their spatial shift, and since the latter is hard to control experimentally, this shift of optical spectra due to the admixture of charge-transfer configurations is probably irrelevant in practice, unlike the case of graphite, where the spectrum involves a large number of electronic states.

Thus, this chapter provides beautiful examples for the wide range of topics in condensed matter physics that today can be addressed with elaborate simulation methods on modern supercomputers, and shows that impressive progress in the understanding of the properties of these systems is gained.

Structures and Phases in (Nano-)Systems in Confined Geometry

**Marcus Beck, Andreas Haller, Daniel Kawetzki, Manuel Matt, Martin Pütz,
Markus Ring, Ralf Schmid, Kristian Scholz, Ullrich Siems, and Peter Nielaba**

Department of Physics, University of Konstanz, 78457 Konstanz, Germany

E-mail: {marcus.beck, andreas.haller, daniel.kawetzki, manuel.matt, martin.puetz, markus.ring,
ralf.schmid, kristian.scholz, ullrich.siems, peter.nielaba}@uni-konstanz.de

Computer simulations by members of the NIC-project HMZ07 are reviewed. Structural and transport properties have been investigated in atomic wires by Molecular Dynamics and electronic transport methods and in model colloids by Brownian Dynamics. Nano-alloys and nano-resonators have been analysed by Molecular Dynamics and path integral Monte Carlo. Spinodal decomposition and domain growth of liquid-vapour systems and galaxy magnetism have been studied by hydrodynamics simulations, and nucleation phenomena in lattice systems by rare event methods.

1 Introduction

Nanostructures and colloidal systems in reduced geometry and external fields have become an interesting research domain in the last years, and in this field, computer simulation studies have contributed many important results. Here we review results obtained in our NIC project HMZ07 and present recent results.

In Sec. 2, we report on the computations of electronic and structural properties of atomic wires. In Sec. 3, the effects of external periodic light fields and driving forces on the structural and dynamical properties of two dimensional colloidal systems, confined in a channel geometry, are analysed by Brownian Dynamics simulations as well as the effect of counterflow on segregation phenomena and the ordering on non-spherical particles in channel geometry. In Sec. 4, the properties of nano-alloys and of nano-resonators in the beam geometry are discussed. Effects of the system size and stretching on the resonator frequencies and damping constants have been computed by Molecular Dynamics, and quantum effects at low temperatures by path integral Monte Carlo. In Sec. 5, hydrodynamics simulations for spinodal decomposition phenomena and galaxy magnetism are presented, and in Sec. 6, results for nucleation phenomena in lattice systems, obtained by rare event methods.

2 Control of Properties of Atomic-Sized Contacts

The potential to construct electronic devices with functional building blocks of atomic size is a major driving force of nanotechnology. When wires are stretched to structures, where only few atoms connect the two ends of the wire, interesting current quantisation phenomena occur. In collaborative work, involving J. C. Cuevas, F. Pauly, E. Scheer, and others, in recent studies¹⁻⁴ we investigated the structural and electronic properties of atomic wires. For the Molecular Dynamics (MD) simulations we used the open source program package

LAMMPS⁵. Within LAMMPS, we employed the Embedded Atom Method (EAM) with the semi-empirical potentials from Sheng *et al.*⁶ to model the interactions between the atoms. To calculate the transport properties of the geometries obtained from the MD simulations, the Landauer-Büttiker formalism was used. We expressed it in terms of Green's function techniques, as explained in Refs. 7, 4. According to the Landauer formula, the

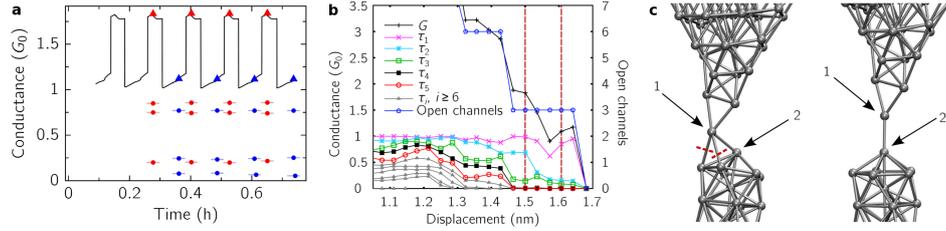


Figure 1. (a) Experimental results from a switching process. (b) A conductance trace G with its channel contributions τ_i as function of the displacement of the two leads from its ideal structure. (c) Connected structures which would explain the experimental findings³.

transmission eigenchannels τ_i are related to the conductance via $G = G_0 \sum_i \tau_i$, where $G_0 = 2e^2/h$ is the quantum of conductance. Conductance histograms for aluminium wires have been compared with experimental data³. Further on we used the information of the channels of the individual traces to connect the experimental results with information about atomic configurations. An example is given in Fig. 1, where in (a) one can find the experimental results from the switching process, in (b) a conductance trace with its channel contributions as function of the displacement of the two leads, and in (c) the connected structures which would explain the experimental findings, are shown. With this formalism we have also access to other properties like shot noise, Fano factor² and thermopower¹. With statistics of 100 curves each we compared conductance histograms, Fano factor, shot noise and thermopower data with experimental results and supported them further with channel distribution data and atomic configurations. With this, we were able to explain² a peak in the variance of the shot noise experiments of Au at about $0.5G_0$. This peak indicates the contribution of more than one channel. Our studies showed low probability appearances of more than one wire in parallel contributing with one channel less than one G_0 . In Ref. 1 we showed, in agreement with experiments, that both Au and Pt statistically have a non vanishing thermopower while going from big contacts to single atom contacts, but with different sign. We could explain the different sign with the help of energy dependent transmission calculations $\tau(E)$ and the calculation of the local density of states (DOS) in the central region of the contacts.

3 Colloidal Systems

Colloidal particles in two- and three-dimensional micro-channels have been studied by Brownian Dynamics (BD) simulations⁸⁻¹², where the particle interactions were modelled by dipolar or Yukawa interactions and hard or soft core potentials with particle diameter σ

for the treatment of the particles excluded volume, and the particle-wall interactions by several methods^{13,14}. Interesting layering phenomena as function of the channel width were detected in two dimensional systems⁸, in good agreement with experimental studies⁸, and are predicted for three dimensional systems⁹. An analysis of the particles diffusion properties in channel direction revealed a non-monotonic crossover from single-file to regular diffusion¹⁰, in close agreement with experimental results¹⁰.

In a recent study¹¹ we examined the effect of a sinusoidal washboard potential on the structure and transport properties of a colloidal system in channel geometry, in order to contribute to the analysis of recent related experimental studies¹⁵. In addition to the driving force F , an external force, periodic in x direction, was applied: $F_x = F + F_p \sin\left(\frac{2\pi}{\lambda}x\right)$. Here, F_p is the amplitude of the periodic force and λ the period length in x -direction.

Without a substrate potential, the particle velocity would equal the applied force. With a substrate potential, the mean velocity is clearly reduced compared to the case of a free moving particle. For small forces the motion of the particles is similar to that in a pinned case. At a certain threshold force the particles start to move, and for high forces the velocity slowly merges into the free velocity. For small period lengths, the particles order more in a triangular structure and for large periods $\lambda > 1.5$, a rectangular order can be identified, see Fig. 2. The layering is only suppressed around $\lambda = 1.5$. For small as well as for large period lengths, four layers are well pronounced, for $\lambda = 0.9$, each second minimum at the boundary is occupied by a particle. For $\lambda = 2.1$, a very strong increase in the mobility at low forces occurs. This is due to very mobile boundary particles that are located between two minima. These topological defects are localised on a few particles around the extra particle and are the so-called “kink” solitons, as known from the Frenkel-Kontorova model.

The effect of counterflow in model colloids has been investigated as well by BD methods: (a) In a recent study¹², segregation phenomena of particles driven in opposite directions have been studied. The particles interacted via a repulsive potential and were confined in three-dimensional hard-walled pipes with quadratic cross sections. In a systematic finite-size study, the pipe length was varied. Ordering on finite length scales was found, but global segregation seems to vanish for infinite channel length, as it was recently found for lane formation in two dimensions. As an additional effect of the finite hard-walled boundary conditions, interface vibrations and longitudinal demixing were found. (b) In another study¹⁶, we considered channels with radial symmetry. The particles were confined in a hollow cylinder with inner radius R_{min} , outer radius R_{max} and height H with quadratic cross section (Fig. 2(b) $R_{max} - R_{min} = H = 9.5$). The two particles types were driven against each other by a radial force $F = F_{max}(R/R_{max})$, acting on each particle in such a way, that the particles moved with a constant angular velocity. At a force of about $F_{max} = 60$ the particles segregated into an inner ring and an outer ring of one particle type, moving in opposite directions (see Fig. 2(b)).

In a recent study, we investigated the layering behaviour of hard spherocylinders in channel geometry¹⁷. The spherocylinders interacted by the Kihara potential¹⁸ and were confined by soft walls. A strong effect of the channel width on the particles ordering (see Fig. 2(c)) and their dynamics was found, similar to previous findings for spherical particles. Detailed further investigations of structural and dynamical properties of mixtures with spherical particles and of finite site effects are planned by a variation of aspect ratios and channel widths, in order to achieve a better understanding of the ordering effect of the confinement on structural and dynamical properties of such systems.

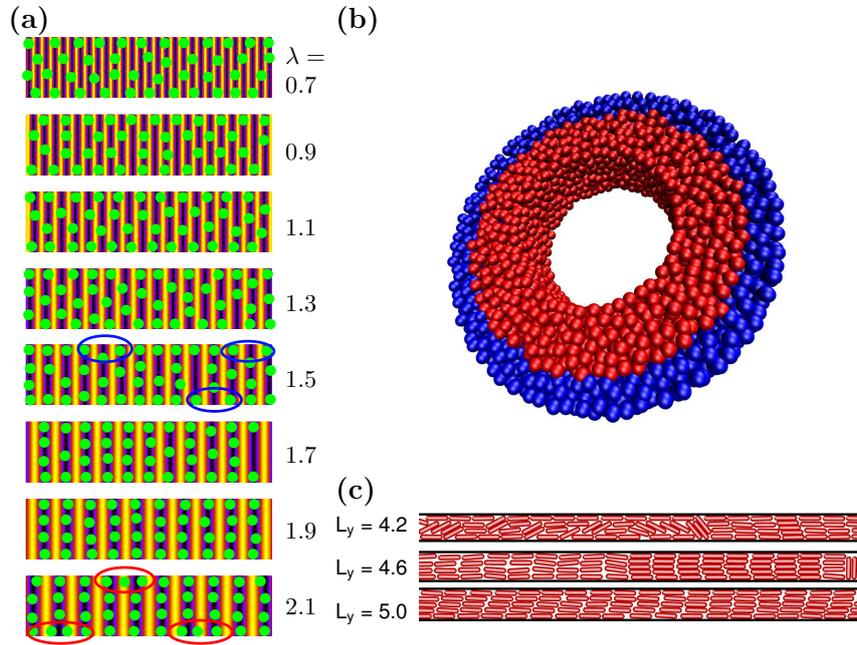


Figure 2. (a) Snapshots of the equilibrium positions for different period lengths ($\lambda=0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9, 2.1$ from top to bottom) on the colour-coded substrate potential (bright, maximum; dark, minimum). Red ellipses, kinks; blue ellipses, antikinks. (b) Snapshot of ordered oppositely driven particle on a ring geometry. (c) Typical configurations for spherocylinders with aspect ratio 5 for various channel wall distances L_y .

4 Nano-Alloys and Nano-Resonators

Due to their remarkable mechanical properties upon loading followed by heating [“shape memory effect” (SME)] and vice versa [“superelasticity” (SE)], shape memory alloys (SMAs) are widely applied these days in diverse macroscopic engineering devices¹⁹. Microscopically, these effects result from the martensitic phase transformation between crystal structures of different symmetries, and from the formation of differently oriented variants of the low-temperature ground state phase (“martensite”).

In our project, atomistic simulation methods (Molecular Dynamics) have been developed and utilised to compute the phase diagram of NiTi²⁰ and FeNi alloys, the effect of the system size on the transition temperature in nano-scale systems²¹, and a full thermodynamic cycle of a temperature driven nano-motor²². The phase transition temperatures for Ni₅₀Ti₅₀ obtained in our studies were in excellent agreement with the experimental data.

The properties of nano-beam resonators have attained much attention recently²³, due to their experimental realisation and the investigation of their properties at sub micro-metre length scales²⁴ accompanied by analytical approximations and computer simulations^{25,26}.

The experimentally studied systems typically show a temporal decay of the oscillation amplitude. Here we report on results of systematic atomistic MD simulations of the dynamical properties of oscillating resonators of various geometries at the nanometre length scale

for various temperatures, under the influence of external stresses, and with local vacancies, using the program SEEVN, developed in our project²⁷, or by the package LAMMPS⁵. Since we are interested in general properties of beam resonator systems, we studied and compared the properties of three systems: Si, SiN, and NiTi, accompanied by studies of systems with simple Lennard-Jones interactions. Our systems are modelled by well tested interaction potentials^{28,29,25}.

In our studies²⁷, we have found interesting structural properties, like a 2x1 dimer surface- and an edge-reconstruction in Si nano-beams, as well as interesting dynamical properties, which are qualitatively similar in all systems studied, but may vary in magnitude due to the different particle interactions. Typical beam sizes were 10-30 nm in length and 2-4 nm in width corresponding to a few thousand up to several ten thousand atoms. After the initial preparation of the structures, the beams were deflected by applying a force until the maximum deflection was obtained, followed by a free oscillation of the beams²⁷. Whenever thermalisation was required in order to heat the beams to a specific temperature, either an Andersen or a Langevin thermostat was used. The end points of the beam were kept fixed. From the time evolution of the centre of mass coordinates we obtained the oscillation frequency f and damping coefficient γ either by Fourier transformation or by directly fitting the resulting curve to an exponentially damped oscillation according to $A \exp(-\gamma t) \sin(2\pi f t - \varphi) + c$.

Fig. 3 shows the high and low frequency parts of the kinetic energy for a NiTi beam for its first and 150th period of oscillation. The energy of the collective motion decreases significantly and dissipates into the fluctuations of the individual atoms around their local lattice positions of the constituting atoms. Fig. 3 (c) shows histograms of the high fre-

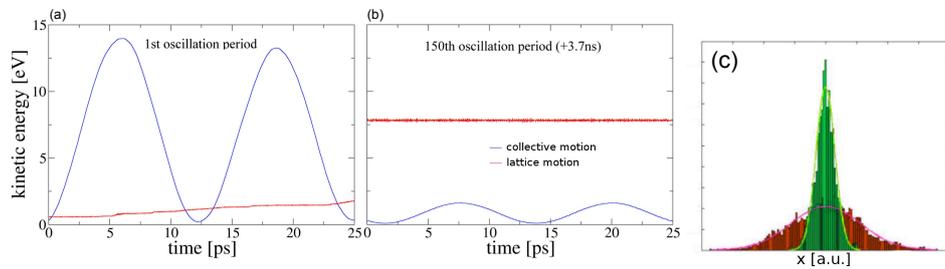


Figure 3. (a) and (b) High (red) and low (blue) frequency parts of the kinetic energy of a NiTi beam summed over all atoms. (a) first oscillation period. (b) 150th oscillation period (after 3.7 ns of oscillation). (c) Normalised histograms of the high frequency part of the x coordinate of an atom in the centre of a NiTi beam for the first (green) and 150th (orange) period of oscillation.

quency part of the x coordinate of an atom in the centre of a NiTi beam for the same two periods of oscillation. At the start of the oscillation, the distribution is sharp due to the collective motion. During the oscillations, the distribution broadens, see Fig. 3(c), due to the local energy flow from the collective mode into the fluctuations of the individual atoms around their local lattice positions, which results in an increase of occupied phase space Γ of the single atoms, and therefore an increase in entropy according to $S \propto \ln \Gamma$. For the collective mode this mechanism results in a damping.

A systematic variation of the system sizes, temperatures and external stresses revealed a strong frequency increase by external stress, e.g. an increase from about 45 GHz to about 62 GHz in Si-beams (with 4x4x60 unit cells) by a 12 % stretching of the wire, a decrease of f with increasing temperature, an increase of the damping coefficient with the temperature and a decrease of the damping with increasing stress. The frequency depends on the beam length L nearly as $1/L^2$, in agreement with the Euler-Bernoulli beam theory.

Quantum effects

A method for the computation of low temperature materials properties is the path integral

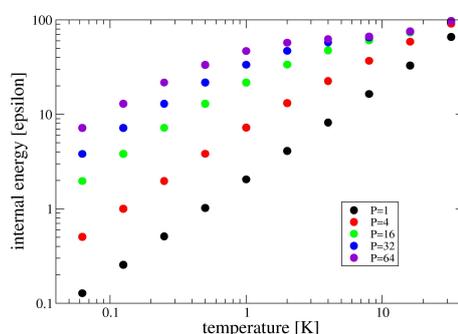


Figure 4. Internal energy of a 100 atom chain with fixed end points for different Trotter numbers P . With increasing Trotter number, the data scale to the quantum limit, resulting in proper low temperature data²⁷.

Monte Carlo method, which has been successfully applied to study the low temperature properties of silicates, the effect of external light fields on the phase diagram of quantum disks, and related systems³¹⁻³⁴. This method permits a systematic approach to the quantum limit, and thus - by comparison with purely classical simulations - the determination of the Debye temperature in nano-beams and -membranes, which is a goal of our project. Fig. 4 shows the internal energy of a 100 atom long Lennard-Jones chain with fixed end points. At low temperature, the classical linear temperature dependence (for Trotter number $P=1$) results in a wrong low temperature heat capacity, contradicting the third law of thermodynamics. With increasing Trotter number P , the results approach the correct low temperature quantum limit, which can be obtained by extrapolation for large P . By this method, such qualitative defects of classical computations can be cured. Our results also indicate that the average width of the chain shows characteristic differences.

5 Hydrodynamics Simulations

One goal of our NIC project is to develop a new method for hydrodynamics computations in soft matter systems using a smoothed particle approach as implemented in astrophysical codes (GADGET³⁵). In colloidal dispersions, the solvent often is charged, resulting in

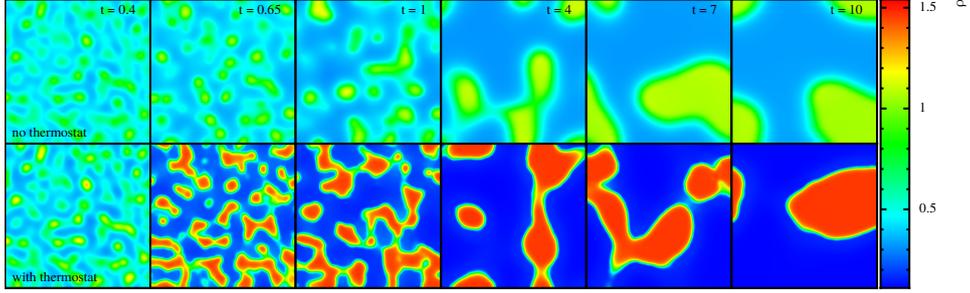


Figure 5. Comparison of density rendered cross section slices from simulation snapshots at the times $t = 0.4, 0.65, 1.00, 4.00, 7.00, 10.00$. The pure thermal simulations with no thermostat (top row) and with thermostat (bottom row), with a side length of the box size $L = 1$. Note that the actual shapes of the three dimensional structures can not be determined in the two dimensional representation.

long range forces. In addition, colloidal particles often are (para-) magnetic systems. The treatment of long range forces in continuum simulations has been optimised in codes like GADGET, and we plan to develop these methods further for systems with long range interactions (magnetic fields in galaxies) and to transfer these methods to condensed matter phenomena in systems with long range interactions at a later stage.

Spinodal decomposition with hydrodynamics

We successfully applied a modification of the massively parallelised astrophysical Smoothed Particle Hydrodynamics (SPH) code GADGET2 to simulations of spinodal decomposition of liquid-vapour systems³⁶. The results are in excellent agreement with theoretical predictions for all expected time regimes from the initial growth of “homophase fluctuations” up to the inertial hydrodynamics regime. The numerical approach follows a modern formulation of the SPH method^{37,38} with a van der Waals equation of state³⁹ and thermal conduction⁴⁰. The dynamics and thermal evolution of instantaneously temperature-quenched systems are investigated. Therefore, we used a simple scaling thermostat that allows thermal fluctuations at a constant predicted mean temperature. Therefore the scaling approach is implemented in the way, that the new particle temperatures \tilde{T}_i are given by $\tilde{T}_i = T_i \cdot (T_0/\bar{T})$, where \bar{T} and T_0 are the mean and initial temperature, respectively.

We found that the initial stage spinodal decomposition is strongly affected by the temperature field. The separated phases react on density changes with a change in temperature. Although, the thermal conduction acts very slowly, thermal deviations are eventually compensated. The domain growth in the late stage of demixing is found to be rather unaffected by thermal fluctuations. We observe a transition from the Lifshitz-Slyozov growth rate with $1/3$ exponent to the inertial hydrodynamics regime with a rate of $2/3$, only excepted from simulations near the critical point where the liquid droplets are observed to nucleate directly in a spherical shape. A comparison of cross section slices of simulations with and without thermostat is shown in Fig. 5. The transition between the growth regimes when the thermostat is applied has been studied by simulations with several initial temperatures and is found to occur earlier for higher initial temperatures. We explain this time dependency with the phase interfaces that become more diffuse and

overlap with approaching the critical point. A prolonging behaviour of the demixing process is observed and also expected to depend on temperature. It is further found that the observations can excellently explain the growth behaviour for pure nonisothermal simulations that are performed without thermostat.

Magnetohydrodynamics simulations

Magnetic fields are an important property of virtually every astrophysical system. Thereby cooperative work between radio observations and theoretical work based on computer simulation plays an important role and is needed to constrain e.g. galactic magnetic fields and the physical processes responsible for its creation. We successfully applied the simulation code GADGET on a variety of astrophysical scenarios (e.g. star formation⁴¹, protostellar outflow⁴¹ and galaxy dynamics and interaction^{42–44}). Many details concerning the influence of magnetic fields are still left to be explored due to the intrinsic complexity associated with the treatment of magnetohydrodynamics in numerical simulations. Detailed knowledge of the structure of our Galaxy from modelling and simulation is crucial for a variety of astrophysical questions, e.g. Ref. 45. For our simulations of the (gas) dynamics and their interplay with magnetic fields in galaxies we use the development version of GADGET³⁵, GADGET-3. Thereby, a state-of-the-art SPH formulation with improved accuracy for simulations of galaxies and the large-scale structure is used⁴⁶. In order to investigate magnetic fields the smoothed particle magnetohydrodynamics implementation of Ref. 47 is used, and in a recent study⁴⁸ we investigated the small scale magnetic field of our own Galaxy, the Milky Way.

6 Nucleation Studies

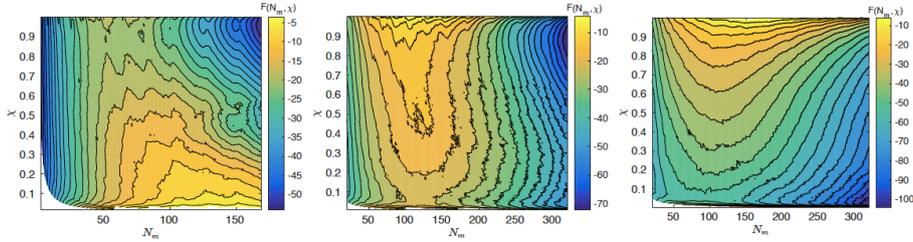


Figure 6. Free energy as function of the two component Potts model with cluster size N_m and crystallinity ratio χ for three temperatures: $k_B T = 0.6$ (left), $k_B T = 0.7$ (middle), $k_B T = 0.8$ (right)⁵⁵.

Based on Refs. 49, 51, 52, metadynamics⁵³, transition path sampling⁵⁰ and transition interface sampling⁵⁴ techniques have been used in a recent study⁵⁵, to compute structures of critical nuclei and nucleation rates in Ising and Potts models. As described in Ref. 49, at a temperature below the roughening transition, the critical nuclei in the Ising system have a rough and non-spherical structure. The novel application of metadynamics to the nucleation scenario in the Ising model⁵⁵ gave easy access to the free energy. Further application of the transition path ensemble⁵⁰ and transition interface sampling⁵⁴ methods were used to

compute nucleation rates⁵⁵. The methods developed for the nucleation studies in the Ising model are transferred to the Potts model in our project. Nucleation phenomena based on a Potts model description have been studied in Refs. 51, 52. In a two component system with solvent and solute particles, lattice sites are occupied by one particle type with one of 24 possible orientations in a three dimensional cubic system. The Hamiltonian is given by Ref. 51: $H = -\sum_{k=1}^2 \left[\sum_{\langle i,j \rangle} [\delta_{m(i),k} \delta_{m(j),k} (G_k + \delta_{s(i)s(j)} A_k)] \right]$. Here, $\langle i, j \rangle$ is the nearest neighbour summation convention, $m(i)$ is the particle type (1 or 2) and $s(i)$ the orientation of a particle at lattice site i , and A_k and G_k control the freezing points and solubilities of both components. In order to compare our results with Ref. 51, the values are chosen as $G_1 = G_2 = 1, A_1 = 1, A_2 = 0$. In a recent study⁵⁵, we could show that metadynamics can be easily applied to the nucleation scenario. The resulting free energy as a function of the particle number N_m of the largest cluster of the same component and of the number N_s of particles with the same orientation s in this cluster, or the crystallinity ratio $\chi = N_s/N_m$, shows interesting features for different temperatures, see Fig. 6. For high temperatures, the nucleation can be described by an amorphous cluster growth (Fig. 6 (middle and right)), whereas at low temperature (Fig. 6(left)), first the cluster crystallises and then grows. These two different scenarios beautifully mirror the different experimentally observed nucleation and growth pathways, namely on the one hand via amorphous precursors versus on the other hand classical critical nuclei⁵⁶.

Acknowledgements

Discussions and collaborations are gratefully acknowledged to J. C. Cuevas, F. Pauly, E. Scheer, C. Schirm, N. Agraït, R. Chen, C. Evangeli, D. Natelson, L. Rincón-García, G. Rubio-Bollinger on the electronic transport studies, to B. Heinze, N. Schwierz, S. Sengupta, A. Erbe, P. Leiderer, C. Kreuter on the colloidal transport studies in channels, to D. Mutter on nano-alloys simulations, to A. Beck and K. Dolag on hydrodynamics simulations, and to C. Peter on nucleation studies. We thank the NIC for computer time on JUROPA / JURECA for the project HMZ07 and the SFBs 767, TR6 and 1214 for support.

References

1. C. Evangeli, M. Matt, L. Rincon-Garcia, F. Pauly, P. Nielaba, G. Rubio-Bollinger, J. C. Cuevas, N. Agraït, *Nano Lett.* **15**, 1006, 2015.
2. R. Chen, M. Matt, F. Pauly, P. Nielaba, J. C. Cuevas, D. Natelson, *J. Phys. Cond. Mat.* **26**, 474204, 2014.
3. C. Schirm, M. Matt, F. Pauly, J. C. Cuevas, P. Nielaba, E. Scheer, *Nature Nanotech.* **8**, 645, 2013.
4. F. Pauly, J. K. Viljas, M. Bürkle, M. Dreher, P. Nielaba, J. C. Cuevas, *Phys. Rev. B* **84**, 195420, 2011.
5. S. Plimpton, *J. Comp. Phys.* **117**, 1, 1995.
6. H. W. Sheng, M. J. Kramer, A. Cadien, T. Fujita, M. W. Chen, *Phys. Rev. B* **83**, 134118, 2011.
7. M. Dreher, F. Pauly, J. Heurich, J. C. Cuevas, E. Scheer, P. Nielaba, *Phys. Rev. B* **72**, 075435, 2005.

8. M. Köppl, P. Henseler, A. Erbe, P. Nielaba, P. Leiderer, Phys. Rev. Lett. **97**, 208302, 2006.
9. N. Schwierz, P. Nielaba, Phys. Rev. E **82**, 031401, 2010.
10. U. Siems, C. Kreuter, A. Erbe, N. Schwierz, S. Sengupta, P. Leiderer, P. Nielaba, Sci. Rep. (Nature) **2** 1015, 2012.
11. U. Siems, P. Nielaba, Phys. Rev. E **91**, 022313, 2015.
12. B. Heinze, U. Siems, P. Nielaba, Phys. Rev. E **92**, 012323, 2015.
13. D. M. Heyes, J. R. Melrose, J. Non-Newtonian Fluid Mechanics **46**, 1, 1993.
14. H. Behringer, R. Eichhorn, Phys. Rev. E **83** 065701(R), 2011.
15. T. Bohlein, J. Mikhael, C. Bechinger, Nat. Mater. **11**, 126, 2011.
16. D. Kawetzki, Bachelor Thesis, University of Konstanz, 2014.
17. A. Haller, Bachelor Thesis, University of Konstanz, 2014.
18. T. Kihara, J. Phys. Soc. Japan **289**, 6, 1951.
19. D. C. Lagoudas(Ed.), *Shape Memory Alloys: Modeling and Engineering Applications*, Springer, 2010.
20. D. Mütter, P. Nielaba, Phys. Rev. B **82**, 224201, 2010.
21. D. Mütter, P. Nielaba, Eur. Phys. J. B **84**, 109, 2011.
22. D. Mütter, P. Nielaba, J. Alloys Compd. **577S**, S83, 2013.
23. A. N. Cleland, *Foundations of Nanomechanics*, Springer, 2002.
24. J. Rieger, A. Isacsson, M. J. Seitner, J. P. Kotthaus, E. M. Weig, Nature Comm. **5**, 3345, 2014.
25. P. Vashishta, R. K. Kalia, A. Nakano, in *Handbook of Materials Modeling*, Springer, 2005.
26. M. Chu, R. E. Rudd, M. P. Blencowe, arXiv 0705:0015v1, 2007.
27. K. Scholz, Dissertation, University of Konstanz (in preparation).
28. F. H. Stillinger, T. A. Weber, Phys. Rev. B **31**, 5262, 1985.
29. J. Tersoff, Phys. Rev. B **37**, 6991, 1988.
30. M. Ring, Master Thesis, University of Konstanz, 2013.
31. M. H. Müser, P. Nielaba, K. Binder, Phys. Rev. B **51**, 2723, 1995.
32. C. Rickardt, P. Nielaba, M. H. Müser, K. Binder, Phys. Rev. B **63**, 045204, 2001.
33. J. Hoffmann, P. Nielaba, Phys. Rev. E **67**, 036115, 2003.
34. P. Nielaba, W. Strepp, in: *Path Integrals New Trends and Perspectives*, ed. by W. Janke and A. Prestel, World Scientific, Singapore, 321, 2008.
35. V. Springel, Mon. Not. R. Astron. Soc. **364**, 1105, 2005.
36. M. Pütz, P. Nielaba, Phys. Rev. E **91**, 032303, 2015.
37. V. Springel, Annu. Rev. Astron. Astrophys. **48**, 391, 2010.
38. D. J. Price, J. Comput. Phys. **231**, 759, 2012.
39. S. Nugent, and H. A. Posch, Phys. Rev. E **62**, 4968–4975, 2000.
40. M. Jubelgas, V. Springel, K. Dolag, Mon. Not. R. Astron. Soc. **351**, 423, 2004.
41. F. Bürzle, P. C. Clark, F. Stasyszyn, T. Greif, K. Dolag, R. S. Klessen, P. Nielaba, Mon. Not. R. Astron. Soc. **412**, 171, 2011.
42. A. Geng, H. Kotarba, F. Bürzle, K. Dolag, F. Stasyszyn, A. Beck, P. Nielaba, Mon. Not. R. Astron. Soc. **419**, 3571, 2012.
43. A. Geng, A. M. Beck, K. Dolag, F. Bürzle, M. C. Beck, H. Kotarba, P. Nielaba, Mon. Not. R. Astron. Soc. **426**, 3160, 2012.

44. A. Beck, H. Lesch, K. Dolag, H. Kotarba, A. Geng, F. Stasyszyn, *Mon. Not. R. Astron. Soc.* **422**, 2152, 2012.
45. K. Dolag, B. M. Gaensler, A. M. Beck, M. C. Beck, *Mon. Not. R. Astron. Soc.* **451**, 4277, 2015.
46. A. M. Beck, G. Murante, A. Arth, R.-S. Remus, A. F. Teklu, J. M. F. Donnert, S. Planelles, M. C. Beck, P. Förster, M. Imgrund, K. Dolag, S. Borgani, *Mon. Not. R. Astron. Soc.* **455**, 2110, 2016.
47. K. Dolag, F. Stasyszyn, *Mon. Not. R. Astron. Soc.* **398**, 1678, 2009.
48. M. C. Beck, A. M. Beck, R. Beck, K. Dolag, A. W. Strong, P. Nielaba, preprint (arXiv:1409.5120).
49. A. C. Pan, D. Chandler, *J. Phys. Chem. B* **108**, 19681, 2004.
50. P. G. Bolhuis, C. Dellago, D. Chandler, P. Geissler, *Annu. Rev. Phys. Chem.* **59**, 291, 2002.
51. N. Duff, B. Peters, *J. Chem. Phys.* **131**, 184101, 2009.
52. V. Agarwal, B. Peters, *J. Chem. Phys.* **140**, 084111, 2014.
53. A. Laio, M. Parrinello, *Proc. Nat. Acad. Sci. U.S.A.*, **99**, 12562, 2002.
54. T. S. van Erp, D. Morioni, P. G. Bolhuis, *J. Chem. Phys.* **118**, 7762, 2003.
55. R. Schmid, Master Thesis, University of Konstanz, 2015.
56. D. Gebauer, H. Cölfen, *Nano Today* **6**, 564, 2011.

Quantum Monte Carlo Simulations of Strongly Correlated Electron Systems: The Dimensional Crossover

Marcin Raczkowski, Martin Bercx, Manuel Weber, Stefan Beyl, Johannes Hofmann, Francesco Parisen Toldin, Martin Hohenadler, and Fakher F. Assaad

Institut für Theoretische Physik und Astrophysik, Universität Würzburg,
Am Hubland, 97074 Würzburg, Germany
E-mail: assaad@physik.uni-wuerzburg.de

The research carried out in this project aims to exactly solve models of correlated electron systems with large-scale fermion quantum Monte Carlo simulations. The goal is to understand the emergent many-body phenomena and critical behaviour. We will first provide an overview of themes and then concentrate on the so-called dimensional crossover phenomena. In this domain, we were able to provide dynamical information about the confinement of fractionalised spinon excitations in weakly coupled Hubbard chains. Such calculations are relevant for the understanding of neutron scattering experiments on KCuF_3 . From the technical point of view we are using and developing a number of algorithms, including continuous-time interaction expansion and determinantal quantum Monte Carlo methods.

1 Introduction

Computational approaches to quantum matter play an important role for our understanding of experiments and theoretical models. In recent years, we have witnessed sustained progress both at the algorithmic level as well as in computational power. In some domains, quantum simulations have achieved such a degree of sophistication that they can calibrate experiments. When electronic correlations play an important role, this is possible only for those systems where simple models can account for the experimental setup. Experiments studying transport through a Kondo dot coupled to superconducting leads can be understood to amazing precision with state-of-the-art continuous-time quantum Monte Carlo (QMC) impurity solvers¹. Advances in QMC simulations of fermions now allow to include long-range Coulomb interactions in model calculations. In the absence of screening – as appropriate for graphene – this is an important step towards realistic modelling². QMC simulations of simplified models are equally important for our understanding of fundamental phenomena such as the electron-phonon interaction³⁻⁵, metal-insulator transitions of Dirac fermions^{6,7}, as well as the interplay between topology and interactions^{8,9}. These topics are being actively pursued in our group. In this article we will concentrate on the so-called dimensional crossover, which plays an important role in the understanding of many experiments.

The interplay between interactions and spatial dimensionality is the key to the essential physics of a large number of materials, ranging from high-temperature superconductors¹⁰ to low-dimensional charge-transfer organic salts¹¹. Dimensionality also plays a decisive role in phase transitions and critical phenomena in quantum magnets¹². In three-dimensional (semiclassical) antiferromagnets, the spontaneous breaking of a continuous symmetry below a finite critical temperature is accompanied by gapless Goldstone modes associated with long-wavelength fluctuations of the order parameter. In contrast, quantum

antiferromagnets such as layered or one-dimensional systems do not develop long-range order at finite temperatures. Numerical simulations aimed at understanding how the semi-classical dynamics characteristic of an ordered phase crosses over to the critical dynamics of lower-dimensional quantum-disordered magnets are thus necessary to explain complex excitation spectra of real quasi-one-dimensional magnetic materials.

In the realm of the solid state, controlling dimensionality implies that the thermal energy (or frequency) is larger than the effective coupling which triggers a dimensional crossover. Under those circumstances, there is no coherence between the lower-dimensional units, so that they effectively decouple, and one can thus expect drastic changes in the physical properties of the system. The dimensional crossover is particularly interesting when elementary excitations fractionalise in the lower-dimensional limit¹³. For example, neutron scattering experiments on weakly coupled spin ladders of CaCu_2O_3 show the two-spinon continuum at frequencies larger than the interchain exchange and their confinement in the higher-dimensional ladder system emerging at lower energies¹⁴. Another experimental realisation is provided by KCuF_3 which consists of weakly coupled spin-1/2 chains¹⁵. At high frequencies one observes the two-spinon continuum, signalling free spinons. In the low-frequency limit, pairs of spinons bind to form the Goldstone mode (spin-waves) of the broken-symmetry phase. Intermediate dimensions and short-range quantum magnetism became recently accessible in ultracold atomic gases by tuning optical lattices comprising weakly coupled one-dimensional chains¹⁶. A variable strength of the interchain coupling allows one to study the impact of a dimensional crossover on antiferromagnetic spin correlations and stimulated a renewed interest in low-dimensional quantum many-body physics¹⁷⁻¹⁹.

Other systems to explore the interplay between low-dimensional quantum dynamics and electron correlations are quasi-one-dimensional organic Bechgaard-Fabre salts. A rich variety of phenomena in their global temperature-pressure phase diagram has been ascribed to an increase in electronic dimensionality with applied pressure which triggers a metal-insulator transition¹¹. Whether the nature of this higher-dimensional metallic phase and its low-energy excitations can be accounted for by a Fermi liquid theory is the key question in the physics of these compounds.

2 The Hubbard Model and the Quantum Monte Carlo Method

A simple model of various correlated electron systems is the Hubbard model. It captures the essence of nontrivial phenomena which are due to the interplay between the kinetic and potential energy. In one dimension, the Bethe ansatz solution and the bosonization approach have lead to a thorough understanding of the low-energy physics¹³. In higher dimensions, however, rigorous results are scarce. Thus, it is natural to ask the question how the crossover from one to two or higher dimensions takes place. Such a dimensional crossover can be studied by smoothly increasing the hopping amplitude between the individual one-dimensional Hubbard chains. To this end, we consider the Hubbard model on a square lattice with an anisotropic hopping at half-filling (one electron per site),

$$H - \mu N = - \sum_{ij,\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_{\mathbf{i}} n_{\mathbf{i}\uparrow} n_{\mathbf{i}\downarrow} - \mu \sum_{\mathbf{i},\sigma} n_{\mathbf{i}\sigma}, \quad (1)$$

with a local Coulomb repulsion U and electron hopping amplitudes: $t_{ij} = t$ on the intrachain bonds, $t_{ij} = t_{\perp}$ on the interchain bonds, and $t_{ij} = -t_{\perp}/4$ between next-nearest-neighbour sites on two adjacent chains. The latter is expected to suppress at least partially the tendency towards the onset of low-temperature symmetry-broken states and opens a possibility to study a transition from the one-dimensional Mott insulator to a higher-dimensional metal. To ensure a half-filled band away from the one-dimensional limit, we adjust the chemical potential μ .

The proper treatment of the dimensional crossover requires non-perturbative techniques to deal with the interchain hopping. This becomes clear by considering a half-filled Hubbard chain with the Mott insulating ground state: if one starts with the Mott fixed point, a metal-insulator transition upon increasing the interchain kinetic energy is certainly a non-perturbative phenomenon since the interchain hopping is an irrelevant perturbation (in the sense of Wilson's renormalisation group) of the Mott fixed point.

To handle the full complexity of the problem, we use a finite-temperature implementation of the auxiliary-field QMC algorithm (see Ref. 20 and references therein) which allows one to compute the expectation value of an observable O in the grand-canonical ensemble:

$$\langle O \rangle = \frac{\text{Tr}[e^{-\beta(H-\mu N)} O]}{\text{Tr}[e^{-\beta(H-\mu N)}]}. \quad (2)$$

It is based on a path integral formulation of the partition function which maps a quantum system in d spatial dimensions onto a $d + 1$ -dimensional classical problem with an additional imaginary-time dimension $\beta = 1/T$. The essence of the QMC method is to separate the one-body kinetic term H_t and the two-body Hubbard term H_U with the help of the Trotter decomposition,

$$e^{-\Delta\tau(H_U+H_t)} \simeq e^{-\Delta\tau H_U} e^{-\Delta\tau H_t}. \quad (3)$$

A fixed small discretisation of the temporal axis $\Delta\tau$ introduces an overall controlled systematic error of order $(\Delta\tau)^2$. The simulations reviewed in Sec. 3 were performed for lattice sizes up to 20×20 in the presence of weak frustration and close ($t_{\perp}/t \lesssim 0.3$) to the one-dimensional limit. Here, the limiting factor is the onset of the negative sign problem which ultimately leads to an exponential scaling of computer time as a function of system size and inverse temperature $\beta = 1/T$.

3 Selected Results

In this section, we summarise our present understanding of dimensional crossover phenomena in strongly anisotropic correlated electron systems such as weakly coupled one-dimensional Hubbard chains, as it emerges from large-scale QMC simulations^{17,19}. From the theoretical point of view, the complexity of a dimensional crossover in coupled Hubbard chains comes from single- and two-particle processes generated by the interchain coupling²¹. On the one hand, the crossover in metallic chains is easily induced by the interchain *one-particle* hopping process thus replacing the Luttinger liquid behaviour with a conventional Fermi liquid metal²². On the other hand, the Mott gap in a half-filled band makes the problem more difficult due to the enhanced relevance of *two-particle* fluctuations: binding of particle-hole pairs from neighbouring chains generates a finite interchain

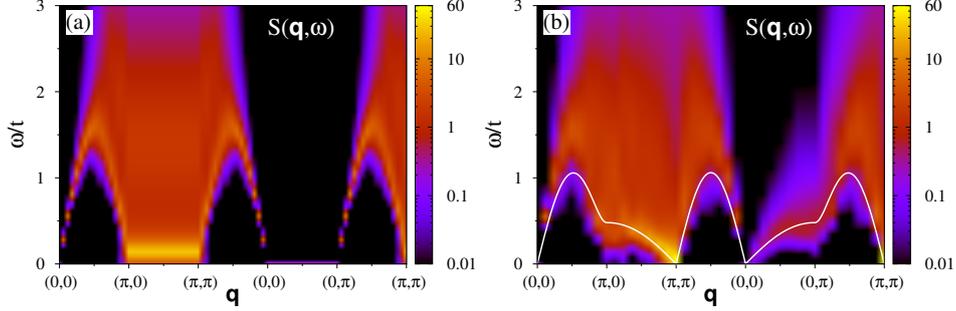


Figure 1. Dimensional-crossover-driven evolution of the magnetic excitation spectrum $S(\mathbf{q}, \omega)$ in the effective zero-temperature limit: (a) 32-site Hubbard chain and (b) 16×16 lattice with strong spatial anisotropy $t_{\perp}/t = 0.3$. The solid line in panel (b) depicts the magnon spectrum obtained within a linear spin-wave theory with the antiferromagnetically ordered ground state. Results taken from Ref. 17.

magnetic interaction J_{\perp} and may induce an onset of the broken-symmetry phase²³. In contrast, when the single-particle tunnelling t_{\perp} overcomes the magnetic coupling J_{\perp} , it drives a metal-insulator transition^{24–26}.

3.1 Spinon Confinement: Dynamics of Weakly Coupled Hubbard Chains

We begin with the two-particle crossover and a dynamical confinement, a phenomenon mostly known from elementary particle physics. Due to the strong interaction, quarks in a hadron are asymptotically free at a shorter distance and never exist as individual particles. Their condensed-matter counterparts, magnetic quasiparticles (spinons), exhibit similar confinement behaviour. Indeed, in the one-dimensional limit, the magnon excitation which one produces by flipping a spin decomposes into two spinons corresponding to domain walls in the antiferromagnetic background. The spinons can move independently along the chain by subsequent spin flips without cost of energy. Since there is a continuum of ways to distribute the momentum \mathbf{q} of the spin-flip process among the two spinons, the corresponding magnetic excitation spectrum is continuous (two-spinon continuum). In the presence of weak interchain interactions, two spinons bind together because their separation would frustrate the interchain bonds. Thus, in analogy with quantum chromodynamics, the interchain interaction plays the role of an attractive potential which grows with the distance between the spinons and confines them back into magnons.

As shown in Fig. 1, lattice QMC simulations allow us to track the evolution of elementary spin excitations upon increasing the coupling between the chains. Our numerical results in the effective zero-temperature limit for the dynamical spin structure factor $S(\mathbf{q}, \omega)$ resolve the frequency dependence of the transition from deconfined to confined spinons. In particular, in Fig. 1(b) we identify essentially two frequency regimes dominated by magnetic excitations of different nature: (i) low-frequency magnons, i.e., the Goldstone mode of the broken continuous $SU(2)$ symmetry group, and (ii) intermediate-frequency two-spinon excitations. We interpret this energy-scale separation as follows. Coupling the chains triggers binding of spinons into spin-waves, thereby replacing the low-energy part of the two-spinon continuum depicted in Fig. 1(a). As shown in Fig. 1(b), a simple linear

spin-wave theory with the antiferromagnetically ordered ground state provides quite a good description of the low-energy part of the magnon spectrum along the $(\pi, \pi) \rightarrow (\pi, 0)$ and $(0, 0) \rightarrow (0, \pi)$ paths. However, deconfinement of spinons still occurs above a threshold energy set up by the strength of attractive potential between the spinons. In proximity to the one-dimensional regime, this potential might be easily overcome, explaining the transfer of spectral weight out of the magnon peak into the two-spinon continuum observed at higher frequencies. The simultaneous observation of low-energy *magnons* and high-energy *spinons* is a fingerprint of a magnetically ordered phase coexisting with strong quantum fluctuations brought by reduced dimensionality. Such a dual nature of magnetic excitation spectra has been resolved in the inelastic neutron scattering data on weakly coupled spin-1/2 chains of KCuF_3 ¹⁵.

3.2 Spin and Charge Dynamics of a Quasi-One-Dimensional Antiferromagnetic Metal

In Sec. 3.1 we have presented numerical evidence of the change in the nature of spin excitations on coupling one-dimensional Hubbard chains in the effective zero-temperature limit, where charge fluctuations can be ignored. Here, we discuss a complementary study at finite temperatures where charge fluctuations become progressively more important as a function of the interchain coupling. Consequently, we find the crossover from the one-dimensional Mott phase – which exhibits spin-charge separation – to a higher-dimensional antiferromagnetic metal¹⁹.

A response function particularly suitable for investigating transport properties of anisotropic systems such as weakly coupled Hubbard chains is the frequency- and polarisation-dependent optical conductivity $\sigma_\alpha(\omega)$: it allows one to resolve a distinct behaviour of the charge dynamics along $[\sigma_\parallel(\omega)]$ or perpendicular $[\sigma_\perp(\omega)]$ to the chains. Since the dimensionality is experimentally controlled not only by the physical or chemical pressure, which changes the ratio of interchain to intrachain transfer, but also by the energy scale used in the measurement, the knowledge of $\sigma_\alpha(\omega)$ offers a possibility to track the evolution of remnant aspects of the one-dimensional physics in the high-energy part of the spectrum.

A dimensional-crossover-driven reconstruction of electronic states as evinced by frequency-dependent intrachain $[\sigma_\parallel(\omega)]$ and interchain $[\sigma_\perp(\omega)]$ optical conductivities is shown in Figs. 2(a,b). At our smallest interchain coupling $t_\perp/t = 0.05$, both optical conductivities display solely a finite-frequency feature, as expected for the one-dimensional Mott insulating phase with charge confined to the individual chains. Around $t_\perp/t = 0.15$ a zero-frequency Drude peak in $\sigma_\parallel(\omega)$ develops accompanied by a tiny Drude-like weight in $\sigma_\perp(\omega)$. Given a nearly t_\perp -independent position of the finite-frequency absorption and a strongly reduced zero-frequency weight in $\sigma_\perp(\omega)$, the system continues to exhibit a substantial tendency to confine charge carriers. In contrast, a pronounced Drude-like feature in $\sigma_\perp(\omega)$ at larger coupling $t_\perp/t = 0.2$ signals the onset of a higher-dimensional metallic phase with electronic quasiparticles replacing fractionalised excitations characteristic of the one-dimensional regime. Interestingly, most of the optical weight does not form a coherent Drude peak as in a usual Fermi liquid metal, but accumulates at finite frequency thus signalling unconventional charge dynamics.

In order to elucidate the origin of this unconventional behaviour we examine the magnetic excitation spectrum $S(\mathbf{q}, \omega)$ at our largest interchain coupling $t_\perp/t = 0.3$ shown in

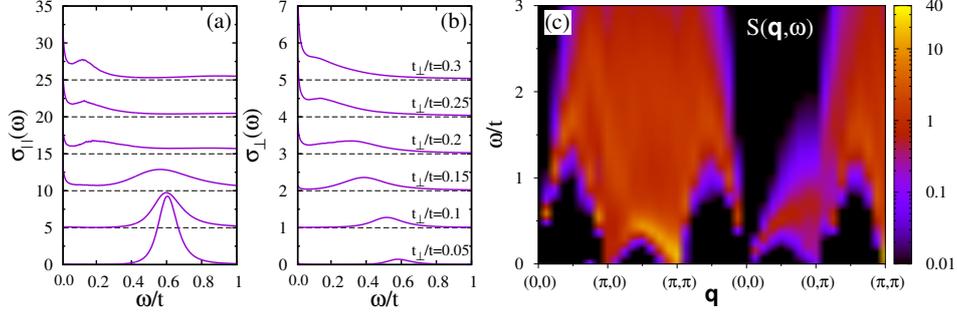


Figure 2. Metal-insulator crossover in weakly coupled one-dimensional Hubbard chains upon increasing the interchain hopping t_{\perp} : (a) intrachain [$\sigma_{\parallel}(\omega)$] and (b) interchain [$\sigma_{\perp}(\omega)$] optical spectra for a 16×16 lattice at temperature $T = t/20$. From bottom to top: $t_{\perp}/t = 0.05, 0.1, 0.15, 0.2, 0.25$, and 0.3 . Panel (c) shows the magnetic excitation spectrum $S(\mathbf{q}, \omega)$ in the metallic phase with short-range antiferromagnetic spin fluctuations at $t_{\perp}/t = 0.3$. Results taken from Ref. 19.

Fig. 2(c). The most striking difference with respect to the one-dimensional regime is a broad dispersive feature along the $(\pi, \pi) \rightarrow (\pi, 0)$ direction. It is a signature of damped antiferromagnetic spin fluctuations (paramagnons) which are not strong enough to gap out the Fermi surface and develop long-range antiferromagnetic order but nevertheless can propagate an appreciable distance. As indicated by their broad spectral width, these paramagnons have a short lifetime due to scattering off mobile charge carriers and merge into a Fermi-liquid-like particle-hole continuum on moving away from $\mathbf{q} = (\pi, \pi)$.

4 Conclusions

We presented a large-scale QMC study of a dimensional crossover in weakly coupled one-dimensional Hubbard chains at half-filling. The simulations allow us to study how the exotic one-dimensional physics relates to a more familiar semiclassical picture valid in higher-dimensional systems. In particular, the evolution of the dynamical spin structure factor upon increasing the interchain coupling in the effective zero-temperature limit indicates a change in the nature of elementary spin excitations from a two-spinon continuum to low-energy spin-waves. Furthermore, finite-temperature results provide evidence of the crossover from the one-dimensional Mott phase to a higher-dimensional antiferromagnetic metal. We have placed emphasis on clarifying the nature of this metallic phase and its low-energy excitations. Our simulations provide a deeper understanding of the intricate spin and charge dynamics of low-dimensional correlated electron systems.

Acknowledgements

Part of this work has been done in collaboration with L. Pollet (LMU Munich). Funding from the DFG grants AS120/8-2 (FOR 1346), AS120/10-1 (FOR 1807), AS120/6-2 (FOR 1162), and HO 4489/3-1 (FOR 1807) is acknowledged. This work was also supported by the FP7/ERC Starting Grant number 306897.

The computations were performed with a grant of CPU time provided by the John von Neumann Institute for Computing (NIC) on the supercomputers JUROPA and JURECA at Jülich Supercomputing Centre (JSC).

References

1. D. J. Luitz, F. F. Assaad, T. Novotný, C. Karrasch, and V. Meden, *Understanding the Josephson Current through a Kondo-Correlated Quantum Dot*, Phys. Rev. Lett., **108**, 227001, 2012.
2. M. Hohenadler, F. Parisen Toldin, I. F. Herbut, and F. F. Assaad, *Phase diagram of the Kane-Mele-Coulomb model*, Phys. Rev. B, **90**, 085146, 2014.
3. F. F. Assaad and T. C. Lang, *Diagrammatic determinantal quantum Monte Carlo methods: Projective schemes and applications to the Hubbard-Holstein model*, Phys. Rev. B, **76**, 035116, 2007.
4. M. Weber, F. F. Assaad, and M. Hohenadler, *Excitation spectra and correlation functions of quantum Su-Schrieffer-Heeger models*, Phys. Rev. B, **91**, 245147, 2015.
5. M. Weber, F. F. Assaad, and M. Hohenadler, *Excitation spectra and correlation functions of quantum Su-Schrieffer-Heeger models*, Phys. Rev. B, **91**, 245147, 2015.
6. F. F. Assaad and I. F. Herbut, *Pinning the Order: The Nature of Quantum Criticality in the Hubbard Model on Honeycomb Lattice*, Phys. Rev. X, **3**, 031010, 2013.
7. F. Parisen Toldin, M. Hohenadler, F. F. Assaad, and I. F. Herbut, *Fermionic quantum criticality in honeycomb and π -flux Hubbard models: Finite-size scaling of renormalization-group-invariant observables from quantum Monte Carlo*, Phys. Rev. B, **91**, 165108, 2015.
8. M. Hohenadler and F. F. Assaad, *Rashba coupling and magnetic order in correlated helical liquids*, Phys. Rev. B, **90**, 245148, 2014.
9. M. Bercx, M. Hohenadler, and F. F. Assaad, *Kane-Mele-Hubbard model on the π -flux honeycomb lattice*, Phys. Rev. B, **90**, 075140, 2014.
10. E. W. Carlson, D. Orgad, S. A. Kivelson, and V. J. Emery, *Dimensional crossover in quasi-one-dimensional and high- T_c superconductors*, Phys. Rev. B, **62**, 3422, 2000.
11. V. Vescoli, L. Degiorgi, W. Henderson, G. Grüner, K. Starkey, and L. K. Montgomery, *Dimensionality-Driven Insulator-to-Metal Transition in the Bechgaard Salts.*, Science, **281**, 1181, 1998.
12. A. Auerbach, *Interacting Electrons and Quantum Magnetism*, Springer-Verlag, New York, 1994.
13. T. Giamarchi, *Quantum Physics in One Dimension*, Oxford University, Oxford, 2004.
14. B. Lake, A. M. Tsvelik, S. Notbohm, D. A. Tennant, T. G. Perring, M. Reehuis, C. Sekar, G. Krabbes, and B. Büchner, *Confinement of fractional quantum number particles in a condensed-matter system*, Nat. Phys., **6**, 50, 2010.
15. B. Lake, D. A. Tennant, C. D. Frost, and S. E. Nagler, *Quantum criticality and universal scaling of a quantum antiferromagnet*, Nat. Mater., **4**, 329, 2005.
16. D. Greif, T. Uehlinger, G. Jotzu, L. Tarruell, and T. Esslinger, *Short-Range Quantum Magnetism of Ultracold Fermions in an Optical Lattice*, Science, **340**, 1307, 2013.
17. M. Raczkowski and F. F. Assaad, *Spinon confinement: Dynamics of weakly coupled Hubbard chains*, Phys. Rev. B, **88**, 085120, 2013.

18. J. Imriška, M. Iazzi, L. Wang, E. Gull, D. Greif, T. Uehlinger, G. Jotzu, L. Tarruell, T. Esslinger, and M. Troyer, *Thermodynamics and Magnetic Properties of the Anisotropic 3D Hubbard Model*, Phys. Rev. Lett., **112**, 115301, 2014.
19. M. Raczkowski, F. F. Assaad, and L. Pollet, *Spin and charge dynamics of a quasi-one-dimensional antiferromagnetic metal*, Phys. Rev. B, **91**, 045137, 2015.
20. F. Assaad and H. Evertz, “World-line and determinantal quantum monte carlo methods for spins, phonons and electrons”, in: Computational Many-Particle Physics, H. Fehske, R. Schneider, and A. Weiße, (Eds.), vol. 739 of *Lecture Notes in Physics*, pp. 277–356. Springer, Berlin Heidelberg, 2008.
21. A. O. Gogolin, A. A. Nersesyan, and A. M. Tsvelik, *Bosonization and Strongly Correlated Systems*, Cambridge University Press, Cambridge, 2004.
22. C. Castellani, C. Di Castro, and W. Metzner, *Dimensional crossover from Fermi to Luttinger liquid*, Phys. Rev. Lett., **72**, 316, 1994.
23. I. Affleck and B. I. Halperin, *On a renormalization group approach to dimensional crossover*, J. Phys. A: Math. Gen., **29**, 2627, 1996.
24. S. Biermann, A. Georges, A. Lichtenstein, and T. Giamarchi, *Deconfinement Transition and Luttinger to Fermi Liquid Crossover in Quasi-One-Dimensional Systems*, Phys. Rev. Lett., **87**, 276405, 2001.
25. F. H. L. Essler and A. M. Tsvelik, *Weakly coupled one-dimensional Mott insulators*, Phys. Rev. B, **65**, 115117, 2002.
26. M. Raczkowski and F. F. Assaad, *Dimensional-Crossover-Driven Mott Transition in the Frustrated Hubbard Model*, Phys. Rev. Lett., **109**, 126404, 2012.

Optical Spectra of Carbon-Based Nanostructures

Michael Rohlfing

Institut für Festkörpertheorie, Universität Münster, Germany
E-mail: Michael.Rohlfing@wwu.de

Carbon exists in many forms, including zero-dimensional fullerenes, one-dimensional nanotubes, two-dimensional graphene and three-dimensional graphite and diamond. Electronic and optical spectroscopy are important tools to analyse these structures and their properties. Here we present optical spectra from *ab initio* many-body perturbation theory for nanotubes and graphite. The data allow to understand details of excited states in these materials. This is of great significance for the interpretation of experimental spectroscopy and for the future manipulation and tuning of optical properties of materials.

1 Introduction

Optical spectroscopy of carbon-based materials is an active field of research, both for fundamental reasons and for practical purposes. On the side of fundamental research, comparison of theoretical and computational results with data from experimental spectroscopy allows to identify general mechanisms of electronic-structure properties inside the material, and may allow to identify materials, defects, dopants, and other details of nanostructured systems. On the side of practical applications, computational spectroscopy might guide experiment in the preselection of materials and systems, to search for new materials and geometries, and to exclude useless systems, before going into the demanding process of sample preparation. Furthermore, theory might demonstrate and elucidate novel mechanisms in the design and control of optical excitations, like e.g. the red-shift tuning of transitions inside a carbon nanotube by touching it with other material (see below).

Internally, optical transitions involve excitons, i.e. coupled excitations of electrons and holes in the material's band structure. Excitons occur everywhere in semiconducting and insulating materials (crystals and molecules) in any dimension, and their interrelation with photons is at the heart of all optics and optoelectronics, including photovoltaics, photocatalysis and more.

Excitons in carbon nanotubes (CNT) and in graphite and graphene have become a highly active research field, providing deep insight in light-matter interaction in carbon-based materials¹⁻⁶. In addition to the optical spectra of single CNT or a single sheet of graphene, their modification by interaction with the environment constitutes an interesting field of research. In the case of nanotubes, characteristic measurements were performed, e.g., on CNT in nitrogen atmosphere³ and on pairs of CNT, with two CNT running along each other⁴. In both cases, a red-shift of the optical transitions to lower excitation energy was observed. In the case of graphene, the excitations inside one sheet start to interact with each other when graphene sheets are stacked to form graphite, followed by spectral changes. Here we take these observations as a motivation for a theoretical study to elucidate the physical mechanisms of spectral shifts. There are two mechanisms involved. On the one hand, the incorporation of additional polarisability (e.g., from a neighbouring nanotube or from the additional graphene sheets inside graphite) cause redshifts of the

optical excitations. On the other hand, there are exciplex contributions that do not occur in a single nanotube or graphene sheet. An exciplex (or charge-transfer) configuration consists of an electron and a hole on different components of the system, i.e. on the two nanotubes in a nanotube pair or on neighbouring sheets in graphite. The admixture of these configurations also lowers the excitation energy, but requires perfect coherence of the quantum-mechanical degrees of freedom. This is in fact given in graphite, but difficult to achieve in two adjacent nanotubes that might be slightly rotated or shifted relative to each other and may not even have the same chirality. Therefore we consider this second mechanism as relevant for graphite, but not for a pair of nanotubes in which the polarisability effect is the only significant effect of spectral shifts. We investigate all issues within many-body perturbation theory (MBPT)⁷, notably by employing the GW approximation (GWA) and the Bethe-Salpeter equation (BSE)⁹, which has become the standard approach for describing CNT excitons^{10–13} and has also been employed for graphene and graphite.

2 Theory

In this section we briefly discuss the computational method used in this work. For a more extended discussion we refer the reader to Ref. 14.

Ab initio quasiparticle (QP) band structures result from the electron self-energy operator $\Sigma(E)$. The state-of-the-art approach to Σ is given by Hedin's GW approximation⁷, which is usually evaluated and employed on top of an underlying density-functional theory (DFT) calculation. The typical procedure employs DFT data to generate the single-particle Green function G_1 and the screened interaction W (usually within the random-phase approximation). Thereafter, the resulting self-energy operator $\Sigma = iG_1W$ replaces the DFT exchange-correlation potential, V_{xc} , arriving at a QP Hamiltonian of

$$\hat{H}^{QP} := \hat{H}^{DFT} + iG_1W - V_{xc} \quad . \quad (1)$$

Eq. 1 yields quasiparticle (QP) states $|m, \mathbf{k}\rangle$ and related band-structure energies $E_{m, \mathbf{k}}$.

Based on the QP states and energies correlated electron-hole states

$$|S\rangle = \sum_{\mathbf{k}} \sum_v^{hole} \sum_c^{elec} A_{v\mathbf{c}\mathbf{k}}^S |v, \mathbf{k}\rangle |c, \mathbf{k} + \mathbf{Q}\rangle \quad (2)$$

are considered as linear combinations of interband transitions between valence band v and conduction band c at wave vector \mathbf{k} . In here, \mathbf{Q} is the total momentum of the electron-hole state which, in optical processes, corresponds to the momentum of the involved photon.

The ansatz of Eq. 2 leads to the Bethe-Salpeter equation (BSE)

$$(E_{c, \mathbf{k} + \mathbf{Q}}^{QP} - E_{v, \mathbf{k}}^{QP}) A_{v\mathbf{c}\mathbf{k}}^S + \sum_{\mathbf{k}'} \sum_{v'}^{hole} \sum_{c'}^{elec} \langle v\mathbf{c}\mathbf{k} | K^{eh} | v'c'\mathbf{k}' \rangle A_{v'c'\mathbf{k}'}^S = \Omega_S A_{v\mathbf{c}\mathbf{k}}^S \quad , \quad (3)$$

with $E_{m, \mathbf{k}}$ being the QP energies from Eq. 1 and K^{eh} being the electron-hole interaction. The Bethe-Salpeter equation and the nature of the resulting states and spectra has been discussed extensively in the literature.

GW/BSE calculations commonly employ the random-phase approximation for evaluating dielectric screening properties (i.e., the W in the self-energy operator and the corresponding electron-hole interaction kernel). This procedure is very time consuming. For the

issues addressed in this work, a simplified, perturbative “LDA+ GdW ” version of MBPT is equally appropriate and much more efficient to evaluate^{14,15}. While being somewhat less accurate (on an absolute energy scale) than a full GW/BSE calculation with RPA dielectric screening, LDA+ GdW still fully incorporates all relevant aspects of the screening (atomic resolution, local-field effects, and non-locality). Our reference calculations within the conventional $GW/BSE/RPA$ approach confirms the applicability of LDA+ GdW .

Note that the screened interaction W depends strongly on the environmental conditions due to the long-ranged nature and non-locality of Coulomb-interaction effects. Simply speaking, a charge at position \mathbf{r} causes an electric field at position \mathbf{r}' . If there is material at position \mathbf{r}' , its electrical polarizability yields an induced charge, which in turn will generate a change of the electric fields at position \mathbf{r} and therefore change the properties of the screened Coulomb interaction W . Via the self energy operator in Eq. 1, $\Sigma = iG_1W$, this long-range mechanism affects the single-particle energy levels at \mathbf{r} . Prominent examples include image-potential effects of molecules on metallic substrates.

A red-shift of the (optical) gap due to the spatial vicinity to a polarisable object has often be interpreted as resulting from a weakening of the (GW) self-energy operator: an increase of dielectric screening weakens W , and the gap closes. For a molecule on a metal, this would be just the image-potential effect. Similarly, a blue-shift of excitons due to the spatial vicinity to a polarisable object is sometimes interpreted as resulting from the weakening of W , as well: the attractive electron-hole interaction becomes smaller, and the exciton binding energy is reduced. It is worth to note that both effects (reduction of the fundamental gap and reduction of the electron-hole binding energy) are real, but are (to first order) exactly opposite in size, thus cancelling each other, provided that the additional polarisability is homogeneous (e.g., a simple dielectric background completely given by a dielectric constant).

Non-zero spectral shifts of excitons require that the additional polarisability be inhomogeneous. This is in fact given for many systems, e.g. the additional polarisability from a neighbouring nanotube or from an adjacent graphene sheet in graphite. It should also be noted that in such situations, model approaches like solvent models that are common in quantum chemistry to describe molecules in solution might not be applicable due to the non-locality, inhomogeneity and anisotropy of the additional polarisability. Our MBPT approach, on the other hand, fully accounts for all these effects automatically, without further effort or modelling, since the full W (being inhomogeneous, anisotropic and non-local) is a key ingredient for the QP energies as well as for the BSE.

3 Results for Nanotubes and for Graphite

For illustration we briefly discuss prototypical results for the optical spectrum of two examples of carbon-based materials, i.e. a semiconducting carbon nanotube and graphite. Carbon nanotubes are formed from a single sheet of graphene (i.e., one monolayer of graphite material) which is rolled into a tube. Depending on the geometrical details (in particular, the chirality), the tubes are metallic or semiconducting. In particular the semiconducting carbon nanotubes show one-dimensional semiconductor physics, i.e. a one-dimensional band structure with a fundamental gap and the formation of excitons across the gap. The corresponding optical transitions start at energies of around 1 eV and above, with a trend to shift to lower energies for tubes of larger diameter⁸.

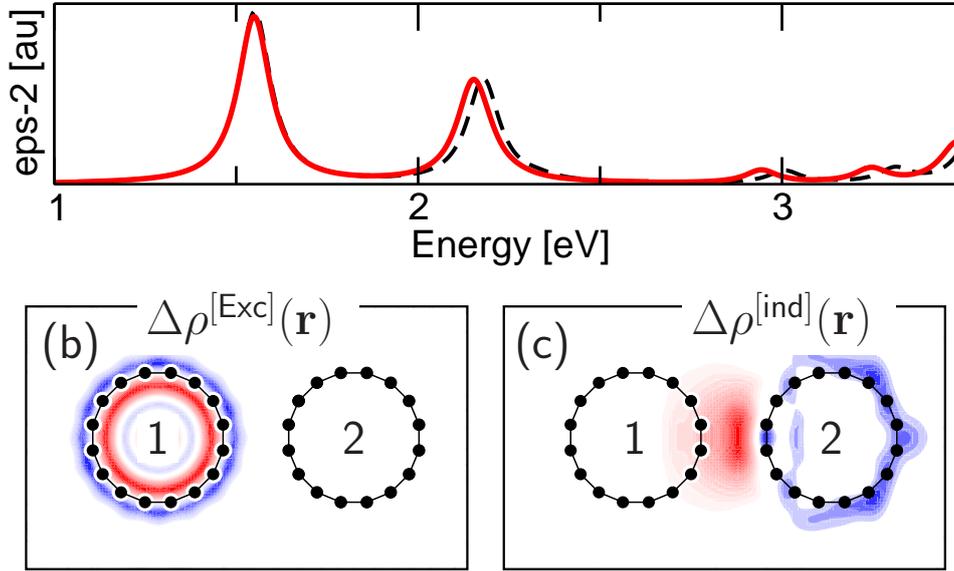


Figure 1. (a) Optical spectrum of a single (8,0) carbon nanotube. The dashed line shows the spectrum of the tube in vacuum. The solid line shows the spectrum in the vicinity of a second nanotube at van der Waals distance. The second nanotube runs along the first one and contributes environment polarisability, only. In all cases the orientation of the electric field vector is along the nanotube. (b) Charge distribution $[\Delta\rho^{[Exc]}(\mathbf{r}) := \rho_v(\mathbf{r}) - \rho_c(\mathbf{r})]$ of exciton D from panel a (blue: negative charge, red: positive charge). (c) Induced charge distribution on the other CNT (blue: negative charge, red: positive charge).

Optical spectra for an individual (8,0) CNT are shown in Fig. 1 a (upper panel). The first four optically active excitations are found at 1.55 eV, 2.18 eV, 2.33 eV, and at 3.01 eV. These results, that were obtained from the simplified LDA+ GdW approach, differ slightly from our full GW/BSE/RPA reference calculation, which yields 1.60 eV, 2.05 eV, 2.42 eV and 3.16 eV for the four peaks. A previous GW/BSE/RPA calculation¹⁰ yielded 1.55 eV and 1.80 eV in comparison with experimental data of 1.60 eV and 1.88 eV^{1,2}. The slight deviations of our LDA+ GdW data result from the approximations involved and from the employment of a model screening.

Starting from the dashed-line spectra of Fig. 1, we now include in the screening the polarisability of another nanotube. In experimental situations the two nanotubes stick to each other due to attractive van der Waals interaction, which makes them lie side by side. If more tubes are involved this would finally result in a bundle. Here we focus on just two nanotubes, both of which are supposed to be (8,0) tubes. The solid line in Fig. 1 shows the effect on a (8,0) CNT when another (8,0) tube is attached to it (at a distance of 3.15 Å). All peaks are redshifted to lower excitation energy. Note that the redshifts are significantly smaller than the reduction of the fundamental gap and of the exciton binding energy (both ~ 0.3 eV). Both effects largely cancel each other (provided that they are described on equal footing, as in our present realisation of MBPT), yielding only a small net effect of a few meV. Our full GW/BSE/RPA reference calculation yields the same redshifts to within 10 meV.

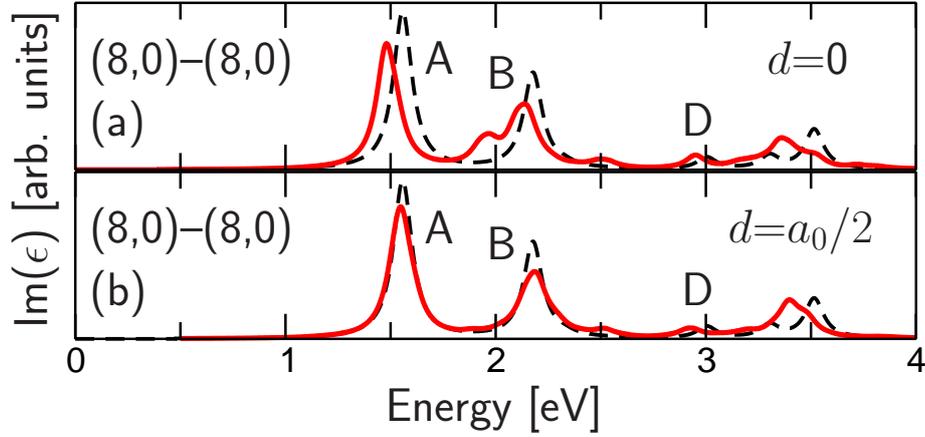


Figure 2. Effect of *electronic coupling* between two (8,0) CNT on their spectra. (a) The two CNT run along each other, with no spatial shift between their unit cells. (b) The two CNT are shifted relative to each other (along their axis) by 2.1 Å (i.e., one half of their lattice constant). In both panels the dashed curve indicates the spectrum of a single CNT in vacuum (cf. Fig. 1).

The redshift results from the polarisation of CNT 2 when an exciton on CNT 1 is excited¹⁶. As illustration, Fig. 1 b shows the change of electronic charge [$\Delta\rho^{[Exc]}(\mathbf{r}) := \rho_v(\mathbf{r}) - \rho_c(\mathbf{r})$] when exciton D is excited on CNT 1. Since the conduction (c) states are closer to the vacuum level than the valence (v) states, the former extend farther into the vacuum, causing $\Delta\rho(\mathbf{r})$ to be slightly positive inside CNT 1 and slightly negative outside. This slight inhomogeneous charge distribution of the exciton leads to a polarisation of the material nearby (here: CNT 2), as shown in Fig. 1 c [induced charge density $\Delta\rho^{[ind]}(\mathbf{r})$]. The interaction between $\Delta\rho^{[Exc]}(\mathbf{r})$ and $\Delta\rho^{[ind]}(\mathbf{r})$ finally redshifts the excitation.

Note that such effects are particularly important if $\Delta\rho^{[Exc]}(\mathbf{r})$ is non-zero at such positions \mathbf{r} where system 2 has high charge susceptibility (caused by its own electronic structure) and inhomogeneity. This is mostly the case at distances of about 1-3 Å from the nuclei of system 2. Here system 2 can be polarised by $\Delta\rho^{[Exc]}(\mathbf{r})$ even if it carries no dipole. For any exciton, $\Delta\rho^{[Exc]}(\mathbf{r})$ must be non-zero somewhere (if not simply for the above-mentioned argument that electrons extend farther into vacuum than holes). The effect described here should thus be of widespread relevance.

In addition to the influence of environmental polarisability, as discussed above, another effect occurs between two touching CNT: the admixture of exciplex (or charge-transfer) configurations to the excitons. For a single tube in vacuum, the electron and the hole have to reside on just the one CNT. For two touching CNT, there are configurations in which the electron is on one CNT and the hole on the other (and vice versa). Simple considerations from second-order perturbation theory indicate that this extension of the configurational space yields an energetic downshift of all excited states, i.e. red-shift trends. This is confirmed by our results shown in Fig. 2 which exhibits the spectrum of a pair of CNT (solid line) in comparison to the spectrum of a single CNT (dashed line). However, this effect of exciplex admixture depends very sensitively on geometric details of the interface. For example, a sliding shift of 2.1 Å (i.e., half a lattice constant) of one tube relative to the

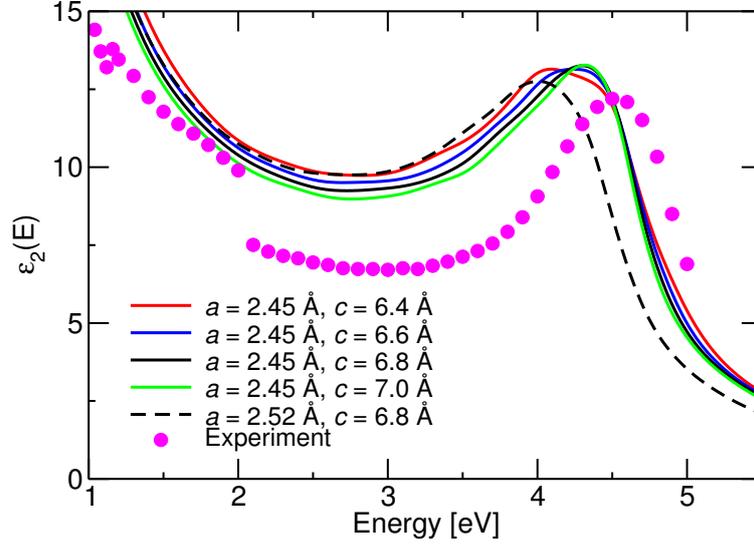


Figure 3. Imaginary part of the dielectric function of graphite, calculated for various values for a and c . Experimental data from Ref. 17.

other completely changes the redshifts (while still being small), as shown by the difference between the two spectra in Fig. 2 a and b. Similar sensitivity was observed for rotations of the CNT around its axes, even for smallest angles⁸. Apparently, imperfect coherence between the electronic (or hole) orbitals of the two components causes uncontrollable scattering of the redshift (while, however, always being negative). Further “chaotic” behaviour of these effects can be expected if the two CNT have different chirality, as in experimental situations. Therefore we conclude that the exciplex admixture is not relevant for CNT, while the environmental polarisability effects is found to be very stable against geometrical details⁸.

The situation in graphite, while being composed of the same graphene sheets from which the nanotubes are formed, is nonetheless significantly different. In particular, the graphene sheets are flat instead of rolled up, and the stacking of the sheets adds three-dimensional character to the material. As a consequence of the different structure, graphite has no fundamental band gap. The optical spectrum can, however, again fully be described by GW-BSE^{18–23}, with self-energy effects and electron-hole attraction similar to semiconductors, as demonstrated by Yang *et al.*²³. Within LDA+ GdW we obtain a spectrum (see Fig. 3) in good agreement with the GW-BSE study by Yang *et al.*, with a maximum at 4.30 eV (4.50 eV in GW-BSE²³) which is 0.32 eV lower than in the free interband spectrum (0.27 eV in GW-BSE²³) due to electron-hole attraction. Here we focus on the dependence of the spectrum and of the contributing excitons on the lattice constants a and c around the experimental equilibrium of $a_0 = 2.45$ Å and $c_0 = 6.71$ Å. Fig. 3 shows the LDA+ GdW optical spectrum for various combinations of a and c . For increasing c the peak near 4.3 eV shifts to higher energies (by about 0.3 eV/Å). The spectrum (including the peak at 4.3 eV) is formed from a large number of resonant rather than bound excitonic states. Changes

in the spectrum do not only result from changes in the excitation energy of each exciton, but also from changes in their optical dipole strength. Between 1 eV and 2 eV, for instance, the spectrum seems to be shifted towards *lower* energy for increasing c . A closer analysis of our data, however, shows that each exciton is rather shifted towards *higher* energy for increasing c . Note that for the graphite case, both the environmental polarisability effect and the admixture of exciplex configurations contribute to the above mentioned spectral shifts. Different from the case of two touching CNT, neighbouring sheets in graphite form a perfect match and allow for wave-function coherence over large distances, thus forming the perfect phase-matching conditions for electrons and holes which is necessary for redshifts from charge-transfer configurations.

4 Summary

In conclusion, we have shown that electronic polarisability of neighbouring systems can redshift exciton states of carbon nanotubes. Here the exciton's charge-density distribution induces charge density in the neighbouring system. This mechanism is particularly effective when the excited system is very close to the neighbouring system, e.g. at physisorption distance. This should be relevant not only for carbon nanotubes (which were taken as an example in the present study), but also for other molecules, polymers, etc. in contact with chemically inert systems. In addition to the polarisability effect, electronic coupling between the systems can significantly enhance the redshifts. However, very precise control of the contact structure would be required for electronic coupling, since it depends very sensitively on the atom positions of the two components relative to each other. For two touching CNT, this condition is not given. In contrast, graphite shows such perfect matching between the sheets that electronic coupling is equally relevant. Both systems demonstrate the delicate relationship between structure and spectra of nanoscale systems.

Acknowledgements

The author gratefully acknowledges the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputers JUROPA and JURECA at Jülich Supercomputing Centre (JSC).

References

1. S. M. Bachilo *et al.*, Science 298, 2361 (c022002).
2. R. B. Weisman and S. M. Bachilo, Nano Lett. 3, 1235, 2003.
3. P. Finnie *et al.*, Phys. Rev. Lett. **94**, 247401, 2005.
4. F. Wang *et al.*, Phys. Rev. Lett. **96**, 167401, 2006.
5. F. Carbone, P. Baum, P. Rudolf, and A. H. Zewail, Phys. Rev. Lett. **100**, 035501, 2008.
6. S. Ghosh, S. M. Bachilo, R. A. Simonette, K. M. Beckingham, and R. B. Weisman, Science **330**, 1656, 2010.
7. G. Onida *et al.*, Rev. Mod. Phys. **74**, 601, 2002.
8. M. Rohlfing, Phys. Rev. Lett. **108**, 087402, 2012.

9. M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927, 2000.
10. C. D. Spataru *et al.*, Phys. Rev. Lett. **92**, 077402, 2004.
11. E. Chang *et al.*, Phys. Rev. Lett. **92**, 196401, 2004.
12. J. Maultzsch *et al.*, Phys. Rev. B **72**, 241402, 2005.
13. C. D. Spataru and F. Leonard, Phys. Rev. Lett. **104**, 177402, 2010.
14. M. Rohlfing, Phys. Rev. B **82**, 205127, 2010.
15. A. Greuling *et al.*, Phys. Rev. B **84**, 125413, 2011.
16. J. M. Garcia-Lastra and K. S. Thygesen, Phys. Rev. Lett. **106**, 187402, 2011.
17. A. Borghesi and G. Guizzetti, in *Handbook of optical constants of solids II*, edited by E.D. Palik, Academic Press, Boston, 1991.
18. V. N. Strocov, P. Blaha, H. I. Starnberg, M. Rohlfing, R. Claessen, J.-M. Debever, and J.-M. Themlin, Phys. Rev. B **61**, 4994, 2000.
19. V. N. Strocov, A. Charrier, J.-M. Themlin, M. Rohlfing, R. Claessen, N. Barrett, J. Avila, J. Sanchez, and M.-C. Asensio, Phys. Rev. B **64**, 075105, 2001.
20. A. Grüneis *et al.*, Phys. Rev. Lett. **100**, 037601, 2008.
21. A. Grüneis, C. Attaccalite, L. Wirtz, H. Shiozawa, R. Saito, T. Pichler, and A. Rubio, Phys. Rev. B **78**, 205425, 2008.
22. P. E. Trevisanutto, C. Giorgetti, L. Reining, M. Ladisa, and V. Olevano, Phys. Rev. Lett. **101**, 226405, 2008.
23. L. Yang, J. Deslippe, C.-H. Park, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **103**, 186802, 2009.

Computational Soft Matter Science

Computational Soft Matter Science

Kurt Kremer

Max-Planck Institut für Polymerforschung, Ackermannweg 10, 55128 Mainz, Deutschland
E-mail: kremer@mpip-mainz.mpg.de

Soft Matter stands for a huge class of liquids and solids which neither can be described as simple liquids nor as hard solids like typical metals or other inorganic materials. *Soft Matter* is pervasive and affects our daily life in many respects. As such it includes biological as well as synthetic materials and systems ranging from man made oil and colloidal liquids to biological fluids or from tissue to everyday plastics. It can be made of relatively small molecules interacting with diverse non covalent interactions such as dispersion and electrostatic interactions or hydrogen bonds or of huge macromolecules such as polymers or proteins, where intra-molecular degrees of freedom are decisive for their properties. The dominance of interactions, which are by far weaker than covalent bonds means that the thermal energy $k_B T$, which is about $4.1 \times 10^{-21} J$ and 1/80 of a covalent bond at room temperature, is the relevant energy scale for these systems, i.e. free energies rather than plain energies are to be considered. As a consequence only the combination of chemically detailed studies combined with analytical theory and simulations of generic, simplified models provide the urgently needed comprehensive understanding of systems and processes. It has been a good tradition at NIC, that both classes of questions and projects are supported. Along this line three contributions are presented in the present volume of the NIC Proceedings.

The first contribution by Guojie Zhang *et al.* describes a new and powerful approach to generate well equilibrated “computer samples” of dense and semidilute polymer systems. The simulation of large, long chain polymer systems faces a number of obstacles, which even nowadays easily make simulations not only very expensive but often hardly possible. In order to study long time dynamics or phase transitions of polymers the preparation of a good, i.e. well equilibrated starting state is a precondition. However, since the typical relaxation time is proportional to roughly $N^{3.4}$, N being the number of beads per chain, equilibration of standard melts often requires computational times of the order of the whole simulation time itself. Thus it is very desirable to prepare well equilibrated systems from scratch, without being bound to the slow physical dynamics of the melts. Zhang *et al.* report a new hierarchical method, which makes systematic use of the universality concept and de Gennes’ blob picture. Consequently the total CPU cost for equilibration is only proportional to the total system size. By describing the chains as soft, strongly overlapping blobs it is possible to setup a structural hierarchy of blobs, which eventually allows to introduce the classical bead models. On each level of description only local relaxation is needed. By resorting to the concept of generalised polymerisation index \bar{N} , which is based on characteristic chain-chain overlaps, it is possible to generate rather different polymer systems within one equilibration scheme, providing a database for a whole class of systems.

In the second contribution Marcus Müller focuses on nonequilibrium processes in copolymer systems. While equilibrium phases of diblock copolymers have been studied extensively by theory and simulations, a detailed comparison to experiment suffers from the fact that experimental systems often do not reach thermal equilibrium, but are rather trapped in a metastable, long lived state. The state of the system depends on the history of the sample. Exactly this problem is addressed in the contribution by M. Müller. How can one understand development of a specific morphology, to what extent does the resulting metastable morphology depend on the starting state and how can one control or guide the process towards a desired result. For this a highly optimised approach based on the so called soft model interactions is employed. As before polymers can easily cross each other, which significantly speeds up the simulations. As long as the chains are short enough, i.e. entanglements do not play a role, such methods still provide the qualitatively correct dynamics. Two different cases are studied. In the first a block length ratio of $f = 1/4$ is chosen with a bcc like initial state, where the segregated minority part forms the “centre of the particles”. By choosing a compressible version of the model, different compressibilities can be assigned to different blocks. Upon rapid expansion large pressure differences occur leading to a fast system response, which is governed by the competition of relatively slow mass transport and fast conformational distortions. The ratio of both significantly influences the resulting metastable structure. In the second example the initial state is disordered. Then the system is elongated at isochoric conditions and then deeply quenched into the segregation regime for the two blocks ($\chi N = 20$). This leads to a prealignment of the blocks and a preferential orientation of the resulting lamellae. A detailed analysis of the emergent morphologies illustrates the huge potential of steered self assembly.

While the previous two contributions do not explicitly include soft matter solid material interfaces, this is the topic of the studies of Frederic Leroy and Florian Müller-Plathe. For many scientific and technical problems the work of adhesion of the liquid with a solid surface plays an important role. When it comes to coatings or other technological processes such as spin casting the wetting properties of the liquid itself or the solvent are crucial. Usually the related work of adhesion is determined by a balance of surface tensions between the liquid and the surface and the two surface tensions of the solid and the liquid surface. This balance can experimentally be determined by a macroscopic measurement of the contact angle between a droplet and the surface. This, however, is numerically often a rather cumbersome task. Thus the authors develop and exploit a complementary ansatz based on thermodynamic integration. Two versions of that approach are presented, the so called phantom wall method, where the droplet and the surface are slowly separated and the dry-surface method where the attractive liquid surface interaction is slowly turned into a repulsive one. In both cases the thermodynamic work of adhesion is determined by the corresponding thermodynamic integration. The applicability of this direct approach has been shown for the example of water on graphene, where the density of defects in the graphene surface is shown to reduce the surface tension with the water.

Though very different in methodology and topic the three different contributions reveal that the importance of high level computer simulations for soft matter problems and demonstrate that only the combined view on generic and chemistry specific information will provide the necessary comprehensive understanding for further materials developments.

Hierarchical Modelling of High-Molecular Weight Polymer Melts: From Soft Blobs to Microscopic Description

Guojie Zhang¹, Torsten Stuehn², Kostas Ch. Daoulas², and Kurt Kremer²

¹ Institute for Theoretical Physics, Georg-August University, 37077 Göttingen, Germany
E-mail: guojie.zhang@theorie.physik.uni-goettingen.de

² Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany
E-mail: {stuehn, daoulas, kremer}@mpip-mainz.mpg.de

We review recent progress in developing a hierarchical strategy for modelling homopolymer melts described with microscopic detail. The samples are gradually equilibrated at all length scales, proceeding from long to short wavelengths through sequential backmapping within a hierarchy of coarse-grained (CG) models. The models stem from the same generic concept, mapping macromolecules on chains of soft spheres (blobs). Each of them represents a microscopic subchain with N_b monomers and their bonded and non-bonded interactions are captured by simple force-fields. The hierarchy of CG models is obtained varying N_b . Melts described with the lowest resolution soft-sphere model are generated by Monte Carlo simulations and are gradually fine-grained into representations of higher resolution. Once N_b becomes sufficiently small, millions of monomers underlying the soft-sphere melt can be reinserted through short Molecular Dynamics simulations, enabled by parallel computing. Chemically different melts can be grouped into classes according to the degree of overlap of their chains (quantified by the invariant degree of polymerisation). In spirit of Renormalisation Group procedures, we demonstrate that long-wavelength conformational and structural properties of an entire class can be captured through a single generic blob-based description. This “blueprint” has to be established only once and can be hierarchically backmapped to describe different materials within the same class. Focusing on method development, we employ for the microscopic description a generic bead-spring model, which retains key qualitative features of chemistry-specific models. Based on this description, it is demonstrated that the new techniques can straightforwardly equilibrate samples of unprecedented size, containing thousands of chains with degrees of polymerisation that are two orders of magnitude larger than the entanglement length. This achievement opens the way for future studies of molecular-scale mechanisms underlying polymer rheology.

1 Introduction

Understanding the connection between structure, processing, and properties of polymers with computer simulations, often requires their description with microscopic detail. This presents, however, a major challenge due to the broad spectrum^{1–3} of characteristic length- and time-scales that are involved. For instance, polymer molecules are mesoscale objects with characteristic dimensions (e.g. radius of gyration) between 10 and 100 nm, formed by covalently linked monomers with sub-nanometre size. The time-scales associated with chain dynamics range from femtoseconds (bond vibrations) up to milliseconds or seconds (conformational relaxation of entire chains). To bridge the extremities of the scale-spectrum hierarchical modelling strategies are particularly attractive. The studied material is considered on several different levels of resolution, each described by an appropriate model. Above the atomistic level, coarse-grained (CG) models are implemented, representing a group of microscopic degrees of freedom by a single effective interaction centre.

The amount of decimated microscopic particles sets the resolution of the CG model. After the system is equilibrated with the coarser resolution, the degrees of freedom of the next level are reinserted and the configuration is re-equilibrated. The final reinsertion of chemical details may introduce millions of particles. Nevertheless computations remain tractable when parallelised, since fine-graining requires only local sampling.

Due to local relaxation of reinserted degrees of freedom, hierarchical approaches implicitly assume that the long-wavelength structure of the material is accurately reproduced on the crudest level of description. For many polymeric systems, this assumption can be rationalised by considering³ hierarchical modelling strategies from the standpoint of renormalisation group (RG) method. For macromolecules, the RG approach was pioneered by de Gennes⁴, who treated polymer chains on the basis of sequentially enlarged blocks of microscopic monomers. As the observation scale increases, these renormalisation transformations converge – irrespective of the underlying microscopic structure – to a generic description depending on a few characteristic parameters, also known as “invariants”. In this scope, the universal long-wavelength behaviour can be captured by any model provided that it is projected on the same set of invariants.

Homopolymer melts are one of the simplest examples of polymeric materials. They are liquids of macromolecules derived from one species of monomer⁵, presenting significant interest for industry (e.g. commodity polymers are processed in molten state) and basic polymer physics. For example, melts (alongside with solutions) provided a framework for the first, simple, mean-field-like descriptions of macromolecular behaviour⁶. These arguments lead to the astonishing conclusion (Flory theorem) that in melts excluded volume intra- and inter-molecular interactions compensate each other. Hence, on large scales chain conformations follow the statistics of an ideal random walk. The scaling law $R_e^2 = Nb_e^2$, is one of the hallmarks of this statistics, where R_e^2 is the mean-square end-to-end distance of the molecule, N the polymerisation degree, and b_e^2 a chemistry-specific coefficient. Hence the number of polymer molecules threading through the volume spanned by a test chain is proportional to $\sqrt{\bar{N}} \equiv \rho R_e^3 / N \sim \sqrt{\bar{N}}$, where ρ is the monomer number density. \bar{N} is the invariant degree of polymerisation, playing a central role in polymer theory. Corrections to Flory’s theorem^{7,8}, demonstrate that the behaviour of simple polymeric systems can be surprisingly complex even for static properties.

The interpretation of dynamics and rheology of molten homopolymers is more challenging. Many of the underlying processes are understood only in the limit of linear response, on the basis of tube theories^{9–13} presenting a mean-field-like approach to polymer dynamics. The collective dynamics of polymer molecules is reduced to the problem of single-chain motion in a transient tube, representing topological constraints (entanglements) imposed on chain dynamics by surrounding chains. *Provided* that large, equilibrated samples of long homopolymer melts described with atomistic resolution are available, tube theories offer a framework for predicting macroscopic linear rheological behaviour from the underlying microscopic picture. For example, the plateau modulus can be determined⁹ as $G_N^0 = (4/5)\rho k_B T / N_e$. The average number of monomers, N_e , between two consecutive topological constraints (entanglement length) can be obtained via, e.g. primitive path analysis¹⁴. For nonlinear viscoelasticity, extensions of the tube model^{9,13,15} can account for some experimental observations. The agreement, however, is by far not perfect^{16,17} and understanding nonlinear viscoelasticity remains an open field of research.

Here, we summarise a novel hierarchical strategy^{18,19} which enables the study of struc-

tural, dynamical, and rheological properties by equilibrating melts of unprecedented size and chain-length, described with microscopic detail. The approach represents homopolymer chains as strings of soft spheres^{20,21} (blobs), each of them corresponding to a subchain of N_b monomers. The quantity of microscopic monomers “lumped” into each sphere sets the resolution at the given level of the hierarchy. In addition, it determines the degree of sphere/sphere overlap (the “softness” of CG interactions) since the amount of interdigitating N_b -monomer subchains scales as $\sqrt{N_b}$. Fine-graining is performed within the sequence of these soft models and microscopic details are efficiently reinserted using high-performance parallel computing, once the resolution becomes sufficiently high. Notably¹⁹, for different homopolymers with the same \bar{N} the hierarchy of CG models at crudest resolutions converges to a universal representation, allowing the generation of their equilibrated configurations from a single long-wavelength “blueprint”.

The manuscript is organised as follows. Sec. 2 presents a generic microscopic model for homopolymer melts. Sec. 3 introduces the soft-sphere model. Sec. 4 presents the backmapping strategy placed in Sec. 5 in context of universal representations of homopolymer melts. Sec. 6 provides a short summary.

2 Generic Microscopic Model

Focusing on method development, we employ for the microscopic description of polymer melts a generic bead-spring model²², incorporating microscopic features that are crucial for reproducing qualitatively the behaviour of real-life molten polymers, especially regarding dynamics and rheology. These features include hard excluded volume, strong covalent bonds, and high density.

Homopolymer melts are modelled as ensembles of $n_{(\gamma)}$ linear chains, composed of $N_{(\gamma)}$ monomers linked by finitely extensible nonlinear elastic (FENE) springs. Chain stiffness is controlled through a standard angular potential²³. Non-bonded interactions are captured through the repulsive Weeks-Chandler-Andersen (WCA) potential. By varying the number density of the monomers, $\rho_{(\gamma)}$, and the strength of the angular potential, $\kappa_{\theta(\gamma)}$, the model can mimic chemical diversity. The subscript γ denotes homopolymer “type”, i.e. two homopolymers are of different types if their molecules are unlike in some respect. A standard parameterisation of the model is employed²², while all lengths and energies are expressed in units of the characteristic WCA length-scale and thermal energy, $k_B T$.

Moderately-sized melts are equilibrated with more conventional techniques to provide reference data for parameterising the soft-sphere model and validating hierarchical fine-graining. We implement a configuration-assembly procedure^{23,24}, operating directly on the microscopic scale. Chains with conformations drawn from the distribution expected in the melt are treated as rigid bodies and arranged relaxing excluded volume constraints. Random placement of molecules generates high density fluctuations, leading^{23,24} to significant conformational distortions when recovering excluded volume. Thus the molecular arrangement is optimised through a Monte Carlo (MC) scheme^{23,24}. Then excluded volume can be reintroduced through a slow “push-off”, employing²³ force-capped WCA potentials. During the “push-off”, realised using Molecular Dynamics (MD), chain-deformation is minimised adjusting force capping through a feedback loop²⁴.

Recent studies demonstrated²⁴ that the computational time of the MC optimisation, increases significantly with chain length. Thus, although state-of-art configuration-assembly

procedures can relax large melts of entangled polymers²⁴, in practice they are more suitable for samples with medium-sized chains, $N_{(\gamma)} \sim 500$. Postulating conformational properties constitutes a more fundamental limitation, since their estimation is challenging in melts with non-linear molecules and inhomogeneous systems.

3 Drastically Coarse-Grained Blob-Based Models

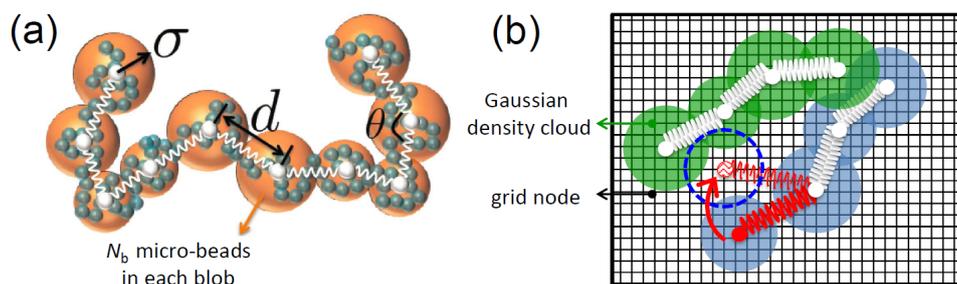


Figure 1. a) Polymer molecule described with microscopic detail (blue) represented^{20,21} by a chain of spheres (yellow) with fluctuating radii. b) Two-dimensional cartoon illustrating the three-dimensional method²¹, assigning the Gaussian density clouds of the spheres to the nodes of a lattice. For clarity only two chains are shown with different colours. Total densities at each node are obtained summing the contributions from all density clouds and employed to calculate the total energy due to non-bonded interactions. During a MC move (red arrow), one needs to calculate changes in local densities only at nodes affected by the move, without neighbour lists.

A CG representation with adjustable resolution is obtained by mapping microscopic polymer molecules on chains of soft spheres, as illustrated in Fig. 1a. The radius σ and the coordinates of the centre of a sphere, match the instantaneous gyration radius and centre-of-mass (COM) of the underlying N_b -monomer subchain. Each CG chain contains N_{CG} blobs linked by simple harmonic bond and angular potentials^{20,21,25}. Fluctuations of σ are controlled^{20,21} by a combination of potentials proposed by Lhuillier²⁶ and Flory²⁷. Non-bonded interactions between two spheres are given by a Gaussian potential, corresponding to an overlap of two Gaussian density clouds. Each of them approximates the distribution in space of N_b monomers underlying the two spheres.

The CG model can be treated with standard MC or MD techniques. However, computational efficiency can be hampered by the large number of interacting neighbours, increasing as $\sim \sqrt{N_b}$. To avoid neighbour-lists, the non-bonded energy can be *equivalently rewritten*²¹ as a functional $\mathcal{H}_{nb} \sim \int d\mathbf{r}(\phi^2(\mathbf{r}) - \psi^2(\mathbf{r}))$. The collective variables $\phi(\mathbf{r})$ and $\psi(\mathbf{r})$ are expressed through the Gaussian density clouds of the spheres. The functional is discretised introducing a fine mesh. As illustrated in Fig. 1b, at each grid node the collective variables are calculated from the contributions of all clouds. Thus an MC scheme²¹ can be employed, where energy differences entering the Metropolis criterion are calculated from the changes induced by a MC move to the collective variables on the “lattice background”, avoiding an explicit consideration of pairwise interactions.

To parameterise the soft-sphere model, in reference samples N_b -monomer subchains are identified to calculate quantifiers of *local* conformations and liquid structure. Typi-

cal examples are¹⁸ distributions of gyration radii of subchains, distance between COMs of sequential subchains, and angles between vectors connecting COMs of sequential subchains. The parameters then follow from standard structure-based coarse-graining¹⁻³ with these quantifiers as an input. After parameterisation, the soft-sphere model can accurately reproduce¹⁸ remaining structural and conformational properties.

4 Backmapping Strategy

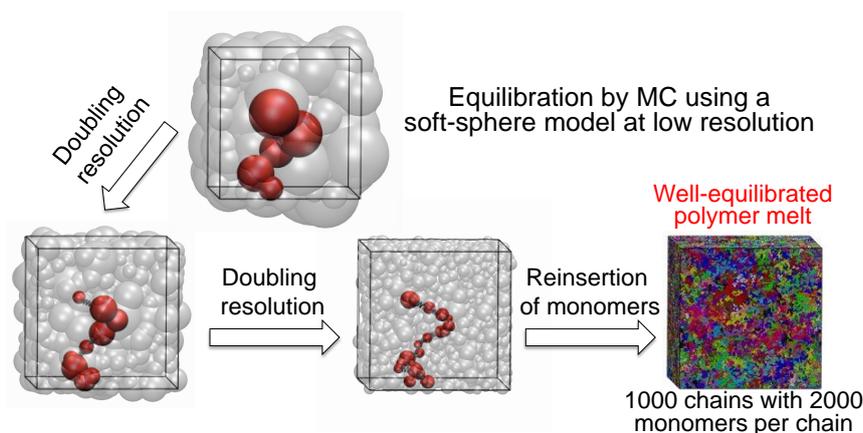


Figure 2. Sketch of the hierarchical backmapping strategy. The last snapshot on the right presents an equilibrated microscopic melt (colours are randomly chosen to improve visibility of different chains). Adapted from Ref. 18.

The sequential backmapping strategy¹⁸ for creating large samples of high molecular weight polymers is summarised in Fig. 2. Initially a CG melt is equilibrated, each sphere representing a large number of microscopic beads. These configurations are sequentially fine-grained, doubling at every step the resolution. The procedure involves only motion of dumbbells, characterised by short relaxation time. The configurations after the last fine-graining step are employed to recover the full microscopic description. Each soft-sphere polymer is replaced by a microscopic bead-spring chain in a matching conformation without non-bonded interactions. This configuration, is subjected to a “push-off” procedure^{23,24} and short re-equilibration. Provided that the melt of the soft-sphere chains, into which the microscopic details are reinserted, corresponds to $N_b < N_e$ the “push-off” and relaxation require only short MD simulations¹⁸.

Most steps of hierarchical backmapping are inexpensive computationally. Parallel computations are involved only during “push-off” and re-equilibration. Typical CPU-time demands are illustrated considering $N_{(\gamma)} = 2000$ melts with thousand chains equilibrated in less than three days, when employing 32 processors during reinsertion of microscopic details¹⁸. Equilibrating the same systems with configuration-assembly on 32 processors lasts more than ten days, even when utilising the efficient ESPReso++ package²⁸.

Fig. 3 illustrates the equilibration of backmapped samples, comparing structural and conformational properties with their counterparts in a reference melt. It follows from

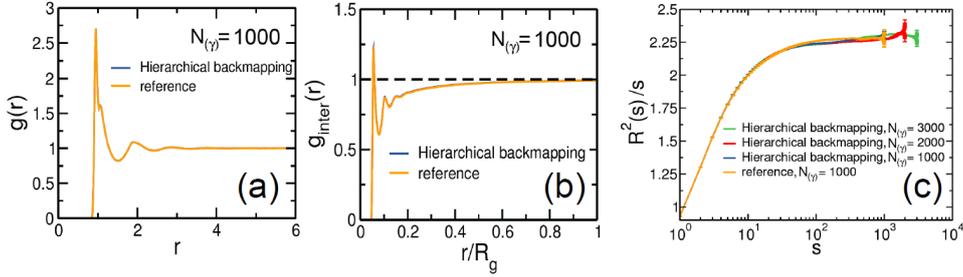


Figure 3. a) Pair-distribution function, $g(r)$, calculated in backmapped (blue line) and reference (orange line) melts ($\kappa_{\theta(\gamma)} = 1$ and $\rho_{(\gamma)} = 0.85$) with $N_{(\gamma)} = 1000$. b) Same as left panel but for the intermolecular part of the pair-distribution function, $g_{\text{inter}}(r)$. c) Internal distance plot, $R^2(s)/s$, calculated in backmapped melts for three different chain lengths $N_{(\gamma)} = 1000, 2000$, and 3000 (blue, red, and green lines) and reference systems (orange line) with $N_{(\gamma)} = 1000$. Similar comparison has been reported in Ref. 18.

Fig. 3a that in backmapped melts the pair-distribution function, $g(r)$, is indistinguishable from the reference data, demonstrating the correct description of microscopic liquid structure. Fig. 3b compares the intermolecular part of the pair-distribution function, $g_{\text{inter}}(r)$, in the same melts against reference data. The match of the curves confirms that the packing of the polymer chains (correlation hole) is also reproduced. The agreement between polymer conformations in backmapped and reference melts is demonstrated in Fig. 3c by the internal distance plot, $R^2(s)/s$. Here $R^2(s)$ is the mean square distance of intramolecular monomers and s is the difference of their ranking numbers along chain backbone. Within the noise of the data, the curves match very well, which is a significant achievement due to the sensitivity of the internal distance plot to potential distortions of chain shape^{18,23}.

5 Universal Long-Wavelength Description of Homopolymer Melts

Coarse-graining polymer melts into ensembles of chains of soft-spheres is similar to a RG procedure. For homopolymer melts with the same $\bar{N}_{(\gamma)}$ (denoted as “class”) this process converges¹⁹, increasing N_b , to a universal long-wavelength representation – an analog of a “fixed point”. For each melt in the same class, this universal long-wavelength representation can be defined introducing a coarse-graining length scale²⁹, $\Delta L_{(\gamma)}$, renormalising the microscopic coordinate space as $\tilde{\mathbf{r}} = \mathbf{r}/\Delta L_{(\gamma)}$. The shortest subchains that are resolved after this coarse-graining operation contain an amount of monomers, $N_{b(\gamma)}$, such that their characteristic size equals $\Delta L_{(\gamma)}$. Thus, (cf. Fig. 4) in the renormalised space all homopolymers are represented as melts of chains of blobs, with the same average diameter equal to unity. In the renormalised blob representation the statistics of the a) distance between the COMs of two sequential blobs and b) angle between two vectors joining the COM of a blob with the COMs of the preceding and the succeeding blob, are universal^{19,25}. Therefore the conformations of CG chains in all homopolymers will be the same, provided that they are represented by the same number of blobs N_{CG} . The last requirement can be fulfilled by matching¹⁹ their coarse-graining scales. In addition, the melts will have the same liquid structure in renormalised space. This follows, first observing¹⁹ that for same N_{CG} the melts will have an identical degree of polymerisation, $\bar{N}_{b(\gamma)}$, also on the level of

subchains. The connection between mesoscale liquid structure of polymers and invariant degree of polymerisation was predicted by de Gennes⁴. Considering the packing of polymer molecules he demonstrated that the chain density within the volume of a test chain is reduced by a universal factor $1 - \bar{N}_{(\gamma)}^{-1/2}$ (Fig. 3b presents this correlation hole). Similarly, liquid packing of subchains depends in a universal way³⁰ on $\bar{N}_{b(\gamma)}$.

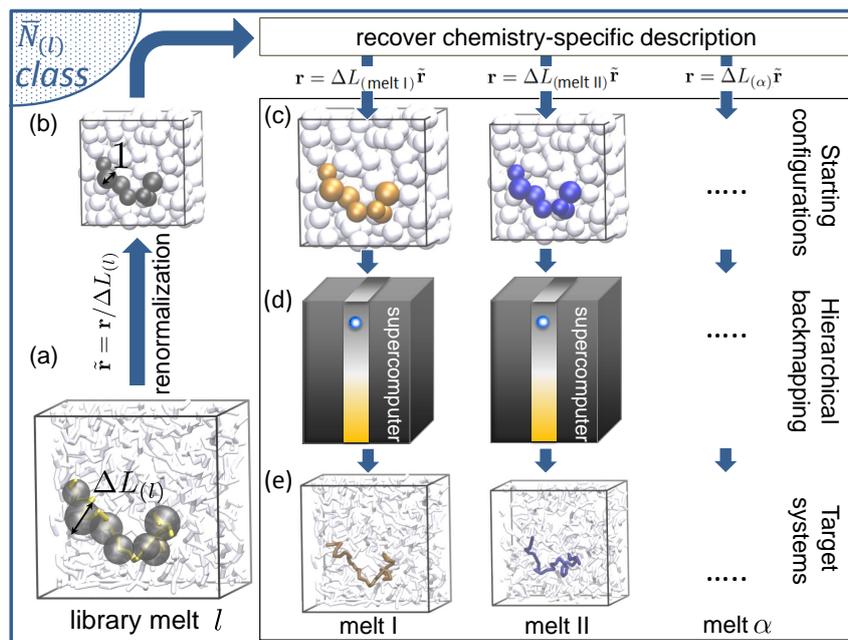


Figure 4. Hierarchical modelling scheme for homopolymer melts with the same invariant degree of polymerisation, $\bar{N}_{(l)}$. a) A library sample described with microscopic detail is subjected to coarse-graining (the average blob size is $\Delta L_{(l)}$) and renormalisation of coordinate-space, $\tilde{\mathbf{r}} = \mathbf{r}/\Delta L_{(l)}$, to obtain b) a universal blob-based blueprint of long-wavelength structure. c) The blueprint is projected on the coordinate space of any other target α -type melt of the same class, back-transforming as $\mathbf{r} = \Delta L_{(\alpha)} \tilde{\mathbf{r}}$. The scales $\Delta L_{(l)}$ and $\Delta L_{(\alpha)}$ are properly matched to ensure the same amount of blobs. d) The blob-based description serves as an input for hierarchical backmapping, delivering e) equilibrated α -type melts described with microscopic detail. Adapted from Ref. 19.

These arguments rely on ideal random walk statistics of polymer conformations predicted by the Flory theorem. The latter was corrected^{7,8}, demonstrating that the volume spanned by a chain exhibits a hierarchy of nested correlation holes of all possible subchains. This introduces weak self-avoidance leading to deviations from the ideal random walk statistics. These deviations, however, diminish increasing N_b and for our practical purposes can be neglected after a certain threshold. The latter can be established monitoring $R_{b(\gamma)}^2/N_{b(\gamma)}b_{e(\gamma)}^2$ as a function of $N_{b(\gamma)}$ in reference samples.

The above suggest that homopolymer melts within the same class can be interconverted via the common blob-based description, simplifying their hierarchical equilibration. Namely, one needs to generate the universal CG representation only once, preferably on the basis of the simplest microscopic model with the desired $\bar{N}_{(\gamma)}$. For example, “materials

genomic” libraries can be developed, containing for each invariant degree of polymerisation a few morphologies which can be easily coarse-grained to define the blob-based “blueprint”. As outlined in Fig. 4, microscopic representations of any class-member can be recovered reinserting chemical details into this common description via the hierarchical backmapping approach¹⁸.

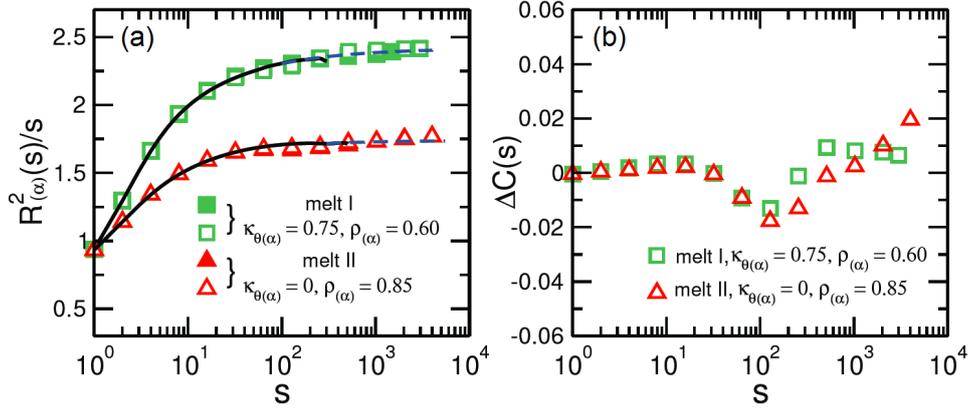


Figure 5. a) Symbols present the internal distance plot, $R_{(\alpha)}^2(s)/s$, for melts I and II equilibrated by backmapping the universal long-wavelength description. Solid lines present $R_{(\alpha)}^2(s)/s$ calculated from reference simulations of shorter melts, extrapolated to larger s (blue dashed lines) on the basis of theoretical predictions^{7,8}. b) Relative deviation, $\Delta C(s)$, of internal distance plots in backmapped melts I and II for $\bar{N}_{(\alpha)} = 15 \times 10^3$ from their counterparts in reference samples. Adapted from Ref. 19.

The feasibility of this approach is illustrated¹⁹ considering library samples characterised by $\kappa_{\theta(l)} = 1.5$ and $\rho_{(l)} = 0.85$, covering a broad range of invariant degrees of polymerisation $7.5 \times 10^3 \leq \bar{N}_{(l)} \leq 15 \times 10^3$ which are representative of values encountered in experiments. The library melts were equilibrated through the hierarchical backmapping procedure¹⁸ of Sec. 4. Configurations of all other systems can be obtained by hierarchically backmapping the blob-based blueprint generated from one of these library melts with the same $\bar{N}_{(l)}$. As an example, we discuss two chemically different systems with $\kappa_{\theta(\alpha)} = 0.75$, $\rho_{(\alpha)} = 0.60$ (melt I) and $\kappa_{\theta(\alpha)} = 0$, $\rho_{(\alpha)} = 0.85$ (melt II). For $N_{(\text{melt I})} = 1500$ and $N_{(\text{melt II})} = 2000$ both melts map on $\bar{N}_{(l_1)} = 7.5 \times 10^3$ while for $N_{(\text{melt I})} = 3000$ and $N_{(\text{melt II})} = 4000$ they correspond to $\bar{N}_{(l_2)} = 15 \times 10^3$.

Fig. 5a demonstrates equilibration of melts I and II backmapped from the universal blueprint, presenting their internal distance plots, $R_{(\alpha)}^2(s)/s$, which follow closely their counterparts in *independent* reference simulations. Fig. 5b quantifies the deviation of the plots for $\bar{N}_{(\gamma)} = 15 \times 10^3$, through their relative difference $\Delta C(s) = [C_{(\alpha)} - C_{(\text{ref})}]/C_{(\text{ref})}$. Here $C_{(\alpha)}$ are the internal distance plots in melt I and II while $C_{(\text{ref})}$ is their counterpart in reference samples. It can be observed that the relative difference is at most 2%. Comparing pair correlation functions in backmapped and reference melts confirms¹⁹ that liquid structure is also correctly reproduced.

6 Concluding Remarks and Outlook

We demonstrated that melts of high molecular weight homopolymers can be efficiently equilibrated employing a novel hierarchical strategy. The method combines low resolution blob-based models capturing accurately long-wavelength properties and fine-graining techniques, allowing for gradual reinsertion of decimated degrees of freedom. Parallel computing is essential for the last stage of fine-graining where *short-time* sampling of configurational space is required by millions of reinserted microscopic particles. Above certain scales, the blob-based models provide a universal description of an entire class of melts, corresponding to homopolymers with the same invariant degree of polymerisation. This blueprint can be hierarchically backmapped to equilibrate any other class-member. Although a generic bead-spring model was chosen for the microscopic description, we expect that similar backmapping strategies can be developed for chemistry-specific models.

Acknowledgements

It is a pleasure to thank Livia Moreira for fruitful discussions and collaborations. The computing time granted by the John von Neumann Institute for Computing (NIC) on JUROPA at Jülich Supercomputing Centre (JSC) is gratefully acknowledged.

References

1. K. Kremer, *Simulation studies of soft matter: generic statistical properties and chemical details*, Eur. Phys. J B **64**, 525, 2008.
2. K. Kremer and F. Müller-Plathe, *Multiscale simulation in polymer science*, Mol. Simulat. **28**, 729, 2002.
3. C. Peter and K. Kremer, *Multiscale simulation of soft matter systems*, Faraday Discuss. **144**, 9, 2010.
4. P. G. de Gennes, *Scaling concepts in polymer physics* (Cornell University Press, Ithaca, New York, 1979).
5. IUPAC, *Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")* (Compiled by A. D. McNaught and A. Wilkinson, Blackwell Scientific Publications, Oxford, 1997).
6. P. Flory, *The configuration of real polymer chains*, J. Chem. Phys. **17**, 303, 1949.
7. J. P. Wittmer, P. Beckrich, H. Meyer, A. Cavallo, A. Johner, and J. Baschnagel, *Intramolecular long-range correlations in polymer melts: the segmental size distribution and its moments*, Phys. Rev. E **76**, 011803, 2007.
8. J. Glaser, J. Qin, P. Medapuram, and D. C. Morse, *Collective and single-chain correlations in disordered melts of symmetric diblock copolymers: quantitative comparison of simulations and theory*, Macromolecules **47**, 851, 2014.
9. M. Doi and S. F. Edwards, *The theory of polymer dynamics* (Oxford University Press: New York, 1986).
10. S. F. Edwards, *The statistical mechanics of polymerized material*, Proc. Phys. Soc. London **92**, 9, 1967.
11. P. G. de Gennes, *Reptation of a polymer chain in presence of fixed obstacles*, J. Chem. Phys. **55**, 572, 1971.

12. M. Daoud and P. G. de Gennes, *Some remarks on the dynamics of polymer melts*, J. Polym. Sci. **17**, 1971, 1979.
13. T. C. B. McLeish, *Tube theory of entangled polymer dynamics*, Adv. Phys. **51**, 1379, 2002.
14. R. Everaers, S. K. Sukumaran, G. S. Grest, C. Svaneborg, A. Sivasubramanian, and K. Kremer, *Rheology and microscopic topology of entangled polymeric liquids*, Science **303**, 823, 2004.
15. M. Doi, *Stress-relaxation of polymeric liquids after double-step strain*, J. Polym. Sci. **18**, 1891, 1980.
16. L. A. Archer, *Separability criteria for entangled polymer liquids*, J. Rheol. **43**, 1555, 1999.
17. S.-Q. Wang, Y. Y. Wang, S. W. Cheng, X. Li, X. Y. Zhu, and H. Sun, *New experiments for improved theoretical description of nonlinear rheology of entangled polymers*, Macromolecules **46**, 3147, 2013.
18. G. Zhang, L. A. Moreira, T. Stuehn, K. Ch. Daoulas, and K. Kremer, *Equilibration of high molecular weight polymer melts: a hierarchical strategy*, ACS Macro Lett. **3**, 198, 2014.
19. G. Zhang, T. Stuehn, K. Ch. Daoulas, and K. Kremer, *One size fits all: equilibrating chemically different polymer liquids through universal long-wavelength description*, J. Chem. Phys. **142**, 221102, 2015.
20. T. Vettorel, G. Besold, and K. Kremer, *Fluctuating soft-sphere approach to coarse-graining of polymer models*, Soft Matter **6**, 2282, 2010.
21. G. Zhang, K. Ch. Daoulas, and K. Kremer, *A new coarse grained particle-to-mesh scheme for modeling soft matter*, Macromol. Chem. Phys. **214**, 214, 2013.
22. K. Kremer and G. Grest, *Dynamics of entangled linear polymer melts: a molecular dynamics simulation*, J. Chem. Phys. **92**, 5057, 1990.
23. R. Auhl, R. Everaers, G. S. Grest, K. Kremer, and S. J. Plimpton, *Equilibration of long chain polymer melts in computer simulations*, J. Chem. Phys. **119**, 12718, 2003.
24. L. A. Moreira, G. Zhang, F. Müller, T. Stuehn, and K. Kremer, *Direct equilibration and characterization of polymer melts for computer simulations*, Macromol. Theory Simul., 2015, DOI:10.1002/mats.201500013.
25. M. Laso, H. C. Öttinger, and U. W. Suter, *Bond-length and bond-angle distributions in coarse-grained polymer chains*, J. Chem. Phys. **95**, 2178, 1991.
26. D. Lhuillier, *A simple model for polymeric fractals in a good solvent and an improved version of the Flory approximation*, J. Phys. France **49**, 705, 1988.
27. P. J. Flory, *The configuration of real polymer chains*, J. Chem. Phys. **17**, 303, 1949.
28. J. D. Halverson, T. Brandes, O. Lenz, A. Arnold, S. Bevc, V. Starchenko, K. Kremer, T. Stuehn, and D. Reith, *ESPResSo++: A modern multiscale simulation package for soft matter systems*, Comput. Phys. Commun. **184**, 1129, 2013.
29. K. F. Freed, *Renormalization group theory of macromolecules* (Wiley, New York, 1987).
30. A. J. Clark, J. McCarty, and M. G. Guenza, *Effective potentials for representing polymers in melts as chains of interacting soft particles*, J. Chem. Phys. **139**, 124906, 2013.

Process-Directed Self-Assembly of Copolymer Materials

Marcus Müller

Institute for Theoretical Physics, Georg-August University,
Friedrich-Hund-Platz 1, 37077 Göttingen, Germany
E-mail: mmueller@theorie.physik.uni-goettingen.de

Using computer simulations of a soft, coarse-grained particle model and numerical self-consistent field calculations, we study the kinetics of self-assembly of lamella-forming block copolymers. The role of non-equilibrium single-chain conformations, conserved densities and dependence of the complex free-energy landscape on the thermodynamic state are highlighted.

1 Introduction

Diblock copolymers are comprised of two, incompatible, flexible polymer blocks that are covalently bonded into linear amphiphilic molecules. The two blocks cannot macroscopically phase separate but, instead, self-assemble into spatially periodic structures whose length scale is dictated by the molecular extension, R_{e0} . The equilibrium structure results from an interplay between the free-energy cost of internal AB interfaces and the entropy loss as the molecules stretch to uniformly fill space. Different morphologies are observed as a function of the volume fraction, f , of one of the blocks and the incompatibility, $\chi N^{1,2}$.

In experiments, however, the self-assembled structures often do not correspond to the thermodynamic equilibrium state but the kinetics of structure formation becomes trapped in a metastable state – e.g., the self-assembly without external guiding fields results in a finger-print like structure, and protracted annealing procedures using solvents or elevated temperature are required to improve long-range order³. Different metastable states can be categorised by their dimensionality: (a) metastable three-dimensional structures like the hexagonally perforated phase, (b) two-dimensional interfaces between grains of structures with the same symmetry but different orientations, or (c) localised defects like edge dislocations in a lamellar (smectic) structure or 5-7 defects in hexagonal structures⁴.

Each of these structures corresponds to a local minimum of the free-energy landscape of the copolymer system, which has been likened to that of glass-forming systems⁵. In order to avoid defect formation or, alternatively, reproducibly direct the self-assembly into a specific metastable morphology, it is important to understand and control the kinetics of structure formation. To this end, one has to explore the complex, rugged free-energy landscape and devise process protocols that reproducibly steer the kinetics of self-assembly. Using two examples I discuss the following questions:

- What are the relevant slow collective variables that describe structure formation?
- What is the free-energy landscape as a functional of these slow variables and how to compute it?
- How can one control the evolution on the free-energy landscape?

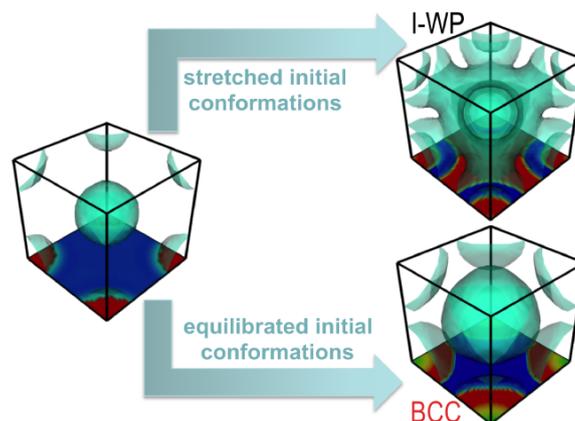


Figure 1. Illustration of the metastable morphologies obtained by kinetic Single-Chain-in-Mean-Field (SCMF) simulations of process-directed self-assembly by rapid expansion of a BCC morphology. The density-contour plots of the average morphology are obtained after back-folding the 27 unit cells of our large simulation box into a single unit cell. Adapted from Ref. 6.

2 Rapid-Quench Structure Formation

In the first example, we investigate the structure formation of a compressible AB block copolymer⁶. In the initial, high-pressure state, both segment types, A and B , are characterised by the same segmental volume and compressibility. The matrix component B is hardly compressible at all densities but the minority component exhibits a super-critical equation of state at lower densities. The fraction of A segments is $f = 1/4$, and a body-centred array of A -spheres is formed in equilibrium at high pressure. Then, the pressure is rapidly reduced, i.e., the system is affinely expanded in the computer simulation of our soft, coarse-grained particle model^{7,8}. Right after the expansion, the density in the A -spheres and the B matrix is identical, but the pressure inside the A -spheres is much larger. Moreover the molecular conformations are stretched, in particular the long B block, which is comprised of the smaller segments. The system has two possibilities to establish mechanical equilibrium: (1) The pressurised A -spheres expand against the low-pressure B matrix, or (2) A segments can evaporate from the spherical A domains. The latter mechanism is facilitated by the tension in the B block that rips the short A fragments out of the A domain into the B matrix. In the simulations we observe that the A -spheres not only expand but that the density of A inside the matrix increases, this A excess in the matrix condenses and forms an additional bicontinuous A -rich network structure inside the B matrix. Thus a new, metastable morphology is formed: Schoen's I-WP phase⁹ as illustrated in Fig. 1⁶.

In order to assess, how important the stretching of the molecular conformations in the initial state is, we equilibrated the conformations in the expanded state at fixed density distribution by applying a field-theoretic umbrella potential⁸. This procedure corresponds to the basic assumption of self-consistent field theory, which asserts that the molecular conformations are in equilibrium with the instantaneous, non-equilibrium density distribution, i.e. the density is the only slow, collective variable. After removing the field-theoretic umbrella potential, the A -spheres only expand, but no additional A network in-

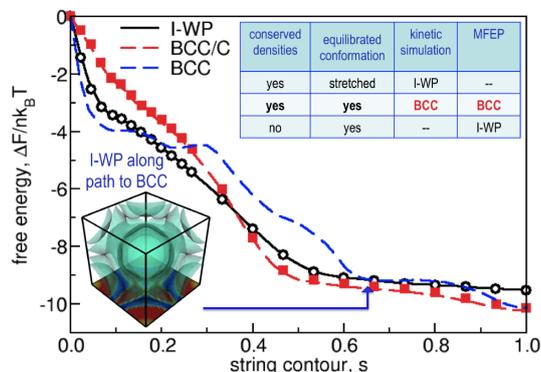


Figure 2. Minimum free-energy path (MFEP) from an expanded BCC morphology to an I-WP morphology or a BCC morphology ignoring the local conservation of the order parameter. Along the path to the BCC structure a metastable I-WP morphology forms at $s \approx 0.67$. Conserving locally the order-parameter, BCC/C, the MFEP directly passes from the BCC to the expanded BCC structure without additional metastable morphology along the path. The table summarises the finding of the SCMF simulation and MFEP calculations. Adapted from Ref. 6.

side the matrix is formed as shown in Fig. 1. Thus, these simulations demonstrate that non-equilibrium molecular conformations are important for the rapid, spinodal structure formation and can qualitatively alter the final structure. In principle, the importance of non-equilibrium molecular conformations must be generally expected because the time scale of spinodal structure formation, $\tau = R_{e0}^2/D$ (with D being the single-chain self-diffusion coefficient) is of the same order of magnitude as the Rouse time $\tau_R = \tau/3\pi^2$ that controls the relaxation of the molecular conformations in a non-entangled melt¹⁰.

We also attempt to describe the structure formation by the minimum free-energy path of the free energy, \mathcal{F} , as a functional of the slow collective A density, $\phi_A(\mathbf{r})$, alone. The minimum free-energy path (MFEP)¹¹ is a sequence of structures, $\phi_A(\mathbf{r}, s)$, that depends on the continuous contour parameter $0 \leq s \leq 1$ and that is commonly defined by the condition that the thermodynamic force (i.e., gradient of the chemical potential) perpendicular to the MFEP vanishes, i.e., $\left(\frac{\delta \mathcal{F}}{\delta \phi_A(\mathbf{r}, s)}\right)^\perp = 0$. Using large-scale computer simulations in conjunction with the improved string method¹², we have computed the MFEP from the affinely deformed BCC structure, and the results are presented in Fig. 2. We observe that the first minimum encountered is the I-WP structure even if we initialise the end-state of the string with the expanded BCC structure. This finding is in contrast to the dynamic SCMF simulation of the kinetics because the free-energy functional does not incorporate information about the non-equilibrium, stretched molecular conformations, and therefore should predict the expanded BCC structure as observed in the simulations with equilibrated chain conformations in the initial state.

This difference can be rationalised as follows: In the kinetic simulations we observed the I-WP structure because the stretched B -blocks pull A fragments into the matrix, whereas in the MFEP calculations the A density in the matrix increases along the path because the local conservation of density is not enforced and the evolution along the MFEP rather corresponds to Allen-Cahn instead of Cahn-Hilliard dynamics¹³. Instead, if we require that the composition current perpendicular to the path vanishes¹⁴ – a requirement that

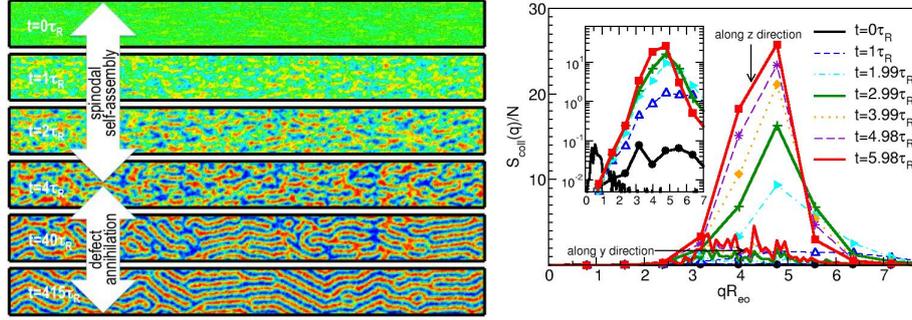


Figure 3. Left: Kinetics of structure formation after a quench of an uniaxial stretched disordered system, $\chi N = 0$, into the lamellar phase $\chi N = 20$. Top views of composition fluctuations are presented where the horizontal axis denotes the stretch direction and the vertical axis the undeformed one. Times, in units of the Rouse time τ_R , are indicated in the key. Right: Time evolution of the structure factor of composition fluctuations, $S_{\text{coll}}(\mathbf{q})$, along the horizontal stretch directions, y , and perpendicular to it, z . The inset depicts the SCMF simulation data $S_{\text{coll}}(qz)$ on a log-scale to illustrate the exponential growth in the spinodal regime. $S_{\text{coll}}(qy, t = 0)$ is also shown as a line without symbols to demonstrate that the height of the peak of the structure factor in the initial anisotropic Gaussian state does not depend on the direction. Adapted from Ref. 22.

can be formally justified by including the Jacobian of the transformation from the particle coordinates, $\{\mathbf{r}\}$, to the collective density, $\phi_A(\mathbf{r})$ ^{11,6,15} – the MFEP calculation BCC/C in Fig. 2 predicts the formation of an expanded BCC structure (in accord with the dynamic SCMF simulations).

3 Directing the Morphology by a Step-Like Elongation

The interplay between structure formation of the collective density, $\phi_A(\mathbf{r}, t)$, and the relaxation of the molecular conformation from a highly stretched, non-equilibrium state can be exploited to direct the large-scale orientation of the self-assembled morphology. This strategy is used in technical process like roll-casting^{16–18} or melt-drawing¹⁹ where experiments observe that the normals of the lamellar stripe structure are preferentially aligned perpendicular to the direction of an uniaxial stretch.

In our dynamic SCMF simulations we mimic this process by (1) equilibrating a polymer film in the disordered state, $\chi N = 0$, (2) affinely elongate the film by a factor $\lambda = 10$ along the horizontal y direction while compressing the film thickness so as to conserve the volume, and then (3) quench the system below the order-disorder transition, $\chi N = 20$. The last step corresponds to the evaporation of solvent after the step-wise elongation or the heating of the specimen above the glass transition temperature. The subsequent kinetics of structure formation is depicted in Fig. 3. During an initial spinodal regime, $t \leq 2\tau_R$, composition fluctuations exponentially grow in time. We observe that they will increase much faster if the wavevector, \mathbf{q} , is aligned perpendicular to the stretch direction (lines with symbols) that in the parallel direction (lines without symbols). These spinodal fluctuations template the structure that, later, is merely improved by defect annihilation. In accord with experiment, we observe a preferential orientation of the lamella normals perpendicular to the stretch direction.

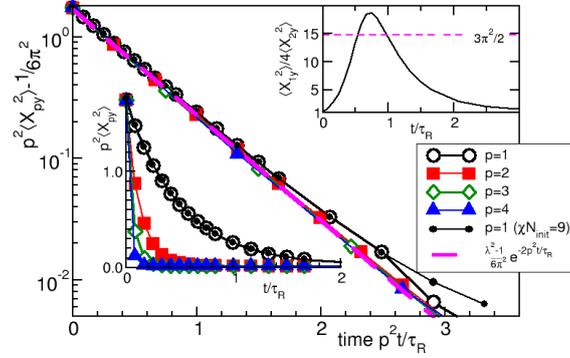


Figure 4. Time evolution of the variance of the first 4 Rouse modes $\langle X_{py}^2 \rangle$ along the stretch direction, y . The main panel presents the single-exponential relaxation with time scale $\tau_R/2p^2$, whereas the left inset shows the unscaled data. The right inset depicts the time evolution of the ratio $\langle X_{1y}^2 \rangle/4\langle X_{2y}^2 \rangle$, which adopts the value 1 for Gaussian chains. Adapted from Ref. 22.

At $t = 0$, the molecules are affinely deformed Gaussian chains with anisotropic statistical segment lengths in the different Cartesian directions. The anisotropy of the statistical segments is directly related to the anisotropy of the (single-chain) stress, $\Sigma_{\alpha\beta} = \frac{3k_B T(N-1)}{V R e_0^2} \sum_{\text{bonds}, i} b_{i\alpha} b_{i\beta}$ where the Greek indices denote the Cartesian directions, x , y and z . If we use these anisotropic Gaussian chain conformations to compute the free energy of the lamellar phase within the self-consistent field theory or the structure factor of composition fluctuations, $S_{\text{coll}}(\mathbf{q})$, according to the Random-Phase-Approximation¹, we will obtain the same result for the two directions – along the stretch direction, y , or perpendicular to it z – provided that we scale the length scale of the lamellar structure or the wavevector by the statistical segment length in that direction. Therefore, the thermodynamics of anisotropic Gaussian chain does not give rise to an orientational preference.

Whereas the initial, affinely deformed chain conformations and the final, equilibrated conformations obey (anisotropic) Gaussian statistics, the time evolution results in temporarily non-Gaussian conformations because the molecular structure on the large length scale retains the anisotropy longer whereas the small-scale structure quickly relaxes towards the isotropic state. During the spinodal stage, $t < 2\tau_R$, the structure formation does not severely interfere with the conformational dynamics of the macromolecules because the bonded forces are significantly stronger than the non-bonded forces that give rise to structure formation. Thus, the conformational dynamics of our soft, coarse-grained particle model, which does not enforce the non-crossability of the chain molecules, obeys Rouse-like dynamics. Fig. 4 demonstrates that the individual Rouse modes exhibit a single-exponential relaxation behaviour during structure formation and that different modes, p , are characterised by the time scale, τ_R/p^2 , like in a disordered melt. The non-Gaussianity of the molecular conformations can be quantified by the ratio of the first two Rouse modes, which is depicted in the inset of Fig. 4. We clearly observe strong deviations from the anisotropic Gauss behaviour during the spinodal time regime where composition fluctua-

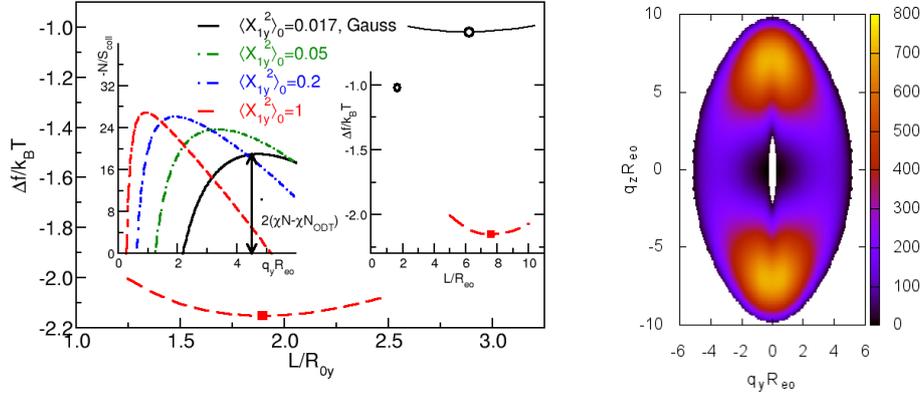


Figure 5. Left: Excess free energy per chain of the lamellar structure and negative, inverse structure factor in the disordered state at $\chi N = 20$ for non-Gaussian chains characterised by $\langle X_{1y}^2 \rangle = 1$. Right: Growth rate, $-q^2/S_{coll}(\mathbf{q})$, of composition fluctuations in the disordered state at $\chi N = 20$. Adapted from Ref. 22.

tions exponentially grow, $t \approx \tau_R$.

Since the stress is comprised of all Rouse modes, however, it is not characterised by a simple, i.e., single-exponential relaxation behaviour. Moreover, if we restrained the diagonal components of the stress, the molecular conformations would remain anisotropic Gaussian^{20,21}. Thus, it is more convenient to consider the variance of the slowest, first Rouse mode (instead of the virial stress), as slow collective variable that characterises the non-equilibrium molecular conformations.

In Fig. 5 (left) we study the thermodynamics of an ensemble of chains whose variance of the first Rouse mode along the stretch direction is fixed to $\langle X_{1y}^2 \rangle = 1$, whereas all other Rouse modes have relaxed to their equilibrium values. Since the Rouse mode is non-local along the chain, the conformations cannot be described by the Markovian propagator of self-consistent field theory, and we use a partial enumeration scheme with $n = 99\,840\,000$ single-chain conformations to compute the single-chain partition function or structure factor. Stretching or increasing $\langle X_{1y}^2 \rangle$ “polarises” the chains by increasing the distance between the centres of the A and B blocks. This effect gives rise to a thermodynamic preference of the ordering with lamella normals along the stretch direction. As shown in Fig. 5 (left), both the excess free energy per chain, Δf , as well as the thermodynamic driving force, $-1/S_{coll}(\mathbf{q})$, for composition fluctuations in the homogeneous disordered state suggest that lamella normals should be aligned along the stretch direction – in marked contrast to the result of the dynamic SCMF simulations and experiments.

Thus, we conclude that the orientation must be a kinetic effect²². Our SCMF simulations indicate that the orientation is decided during the regime of spinodal self-assembly. In Fig. 4 we demonstrated that the single-chain relaxation, which is governed by the strong bonded forces, is hardly affected by the spinodal structure formation. In turn, we expect that a spatially homogeneous deformation does not give rise to a current. Therefore the kinetics of ϕ_A is given by model-B dynamics¹³ and we plot in Fig. 5 (right) the kinetic growth

rate of composition fluctuations, $-\mathbf{q}^2/S_{\text{coll}}(\mathbf{q})$. Here the additional factor, q^2 , arises from diffusive dynamics of the locally conserved composition. Indeed, Fig. 5 (right) reveals that the growth rate of fluctuations with \mathbf{q} perpendicular to the stretch direction is significantly larger than for those with orientation y . Although these fluctuations are thermodynamically less favourable, their growth only requires transport over a shorter distance.

4 Concluding Remarks

The two examples – structure formation after a rapid pressure quench and after a step-wise elongation – illustrate that the structure formation of block copolymers is not only determined by thermodynamic equilibrium considerations of the density but that (1) the non-equilibrium molecular conformations during the early, spinodal stages of self-assembly and (2) the diffusive kinetics due to the local conservation of composition can qualitatively alter the predictions.

An appropriate theoretical description of these self-assembly processes from a highly non-equilibrium state cannot solely be based on the free-energy landscape as a functional of the collective composition but our simulations suggest that the variance of the first, slowest Rouse mode should be included as an additional variable. The qualitative properties of such a free-energy functional, $\mathcal{F}[\phi_A(\mathbf{r}), \mathbf{X}_1^2(\mathbf{r})]$ are largely unexplored. A complete description also has to complement the Cahn-Hilliard equation for the time evolution of the composition by a dynamic equation for the collective variable, $\mathbf{X}_1^2(\mathbf{r})$, that characterises the molecular conformations. In the present case, it simply followed Rouse dynamics of the disordered system but in general the coupling between conformations and morphology and also entanglement effects shall be considered. Such a theoretical framework will provide a general strategy to devise processes (e.g., temporal protocols for varying thermodynamic control parameters like temperature, pressure, mechanical strain or light-induced chemical transformations) that reproducibly direct the kinetics of self-assembly into long-lived metastable morphologies.

Acknowledgements

It is a great pleasure to thank De-Wen Sun and Jiuzhou Tang for enjoyable and fruitful collaboration. Financial support has been provided by the DFG Mu1674/14, SFB937 TP A04, and the EU-FP7 project CoLiSA.MMP. We gratefully acknowledge a generous grant of computing time at the John von Neumann Institute for Computing, Jülich, Germany.

References

1. L. Leibler, *Theory of microphase separation in block copolymers*, *Macromolecules*, **13**, 1602–1617, 1980.
2. F. S. Bates and G. H. Fredrickson, *Block Copolymers - Designer Soft Materials*, *Physics Today*, **52**, 32–38, 1999.
3. M. Müller and J. J. de Pablo, *Computational approaches for the dynamics of structure formation in self-assembling polymeric materials*, *Annu. Rev. Mater. Sci.*, **43**, 1–34, 2013.

4. W. H. Li and M. Müller, *Defects in the self-assembly of block copolymers and their relevance for directed self-assembly*, *Annu. Rev. Chem. Biomol. Eng.*, **6**, 187–216, 2015.
5. C.-Z. Zhang and Z.-G. Wang, *Random isotropic structures and possible glass transitions in diblock copolymer melts*, *Phys. Rev. E*, **73**, 031804, 2006.
6. M. Müller and D.-W. Sun, *Directing the self-assembly of block copolymers into a metastable complex network phase via a deep and rapid quench*, *Phys. Rev. Lett.*, **111**, 267801, 2013.
7. K. C. Daoulas and M. Müller, *Single Chain in Mean Field simulations: Quasi-instantaneous field approximation and quantitative comparison with Monte Carlo simulations*, *J. Chem. Phys.*, **125**, 184904, 2006.
8. M. Müller, *Studying amphiphilic self-assembly with soft coarse-grained models*, *J. Stat. Phys.*, **145**, 967–1016, 2011.
9. A. H. Schoen, *Infinite periodic minimal surfaces without self-intersections*, NASA Technical Note TN D-5541i.
10. M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, Oxford University Press, New York, 1994.
11. L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, *String method in collective variables: Minimum free energy paths and isocommittor surfaces*, *J. Chem. Phys.*, **125**, 024106, 2006.
12. W. E. W. Ren, and E. Vanden-Eijnden, *Simplified and improved string method for computing the minimum energy paths in barrier-crossing events*, *J. Chem. Phys.*, **126**, 164103, 2007.
13. P. C. Hohenberg and B. I. Halperin, *Theory of dynamic critical phenomena*, *Rev. Mod. Phys.*, **49**, 435–479, 1977.
14. W. Zhang, T. J. Li, and P. W. Zhang, *Numerical study for the nucleation of the one-dimensional stochastic Cahn-Hilliard dynamics*, *Commun. Math. Sci.*, **10**, 1105–1132, 2012.
15. M. Müller and D. W. Sun, *Process-directed self-assembly of block copolymers: A computer simulation study*, *J. Phys.: Condens Matter*, **27**, 194101, 2015.
16. R. J. Albalak and E. L. Thomas, *Roll-casting of block copolymers and of block copolymer-homopolymer blends*, *J. Polym. Sci. B: Polym. Phys.*, **32**, 341–350, 1994.
17. M. A. Villara, D. R. Rueda, F. Ania, and E. L. Thomas, *Study of oriented block copolymers films obtained by roll-casting*, *Polymer*, **43**, 5139–5145, 2002.
18. T. Xu, J. T. Goldbach, and T. P. Russell, *Sequential, orthogonal fields: A path to long-range, 3-D order in block copolymer thin films*, *Macromolecules*, **36**, 7296–7300, 2003.
19. T. Keller, C. Semmler, and K. D. Jandt, *Strain-induced phase morphology in melt drawn ultrathin highly oriented block copolymer films*, *Macromol. Rapid Commun.*, **29**, 876–884, 2008.
20. G. H. Fredrickson, *Dynamics and rheology of inhomogeneous polymeric fluids: a complex Langevin approach*, *J. Chem. Phys.*, **117**, 6810–6820, 2002.
21. T. Shima, H. Kuni, Y. Okabe, M. Doi, X. F. Yuan, and T. Kawakatsu, *Self-Consistent-Field Theory of viscoelastic behavior of inhomogeneous dense polymer systems*, *Macromolecules*, **36**, 9199–9204, 2003.
22. M. Müller and J. Tang, *Alignment of copolymer morphology by planar step elongation during spinodal self-assembly*, *Phys. Rev. Lett.*, **115**, 228301, 2015.

Calculation of the Work of Adhesion of Solid-Liquid Interfaces by Molecular Dynamics Simulations

Frédéric Leroy and Florian Müller-Plathe

Eduard-Zintl-Institut für Anorganische und Physikalische Chemie and
Research Cluster Center of Smart Interfaces, Technische Universität Darmstadt,
Alarich-Weiss-Straße 4, 64287 Darmstadt, Germany
E-mail: f.leroy@theo.chemie.tu-darmstadt.de

Colloids, polymer-nanocomposites or polymers and liquids in contact with extended surfaces are examples of systems in which interfaces play a crucial role. The stability of such materials is dominated by the thermodynamic properties of these interfaces. Further advances in the development of those materials require a better understanding of the connection between the interfacial intermolecular interactions and these thermodynamic properties. In our project, we develop algorithms to quantify thermodynamic quantities at the solid-liquid interface through molecular dynamics simulations and relate them to the intermolecular interactions and to the interfacial structure of the liquids. We illustrate our approach through the example of the water-graphene system which has intensively been discussed in the last few years. We also discuss the perspectives our approach opens in multiscale modelling of interfacial soft-matter systems.

1 Introduction

Adhesion of paint on a metal surface, the roll-off motion of a water drop on a plant leaf or liquid flow on a chip for medical applications are typical examples of phenomena where the interaction between a polymer or a liquid and a surface plays a crucial role. The stability and the thermal, mechanical, optical or electrical properties of materials such as colloidal suspensions, nanofluids or polymer nanocomposites to mention a few also strongly depend on the intermolecular interactions at the interface between a liquid or a polymer matrix and solid particles. While experiments to probe structures on distances of a few angstroms in the vicinity of solid surfaces and relate the corresponding observations to thermodynamic information require further development, classical molecular dynamics (MD) simulations and quantum calculations are methods of choice to characterise interfaces at the molecular scale. These approaches show that liquids and polymers adopt a layer structure in the vicinity of surfaces whose magnitude depends on the strength of the interaction. This is illustrated in Fig. 1 where we compare the mass distribution of water on a gold surface and on a graphene monolayer as obtained by MD simulations.

Through the combination of statistical thermodynamics and MD simulations performed on JUROPA and now on JURECA we address the question of how such a peculiar interfacial structure influences the wetting and adhesion properties of liquids and polymers on surfaces. The principles of the methodologies we have developed to reach this goal are briefly discussed. We illustrate how they have been employed to contribute to understand better the wetting properties of graphene. In fact, this research topic of considerable importance has strongly benefited from molecular simulations carried out by several groups to understand experiments.

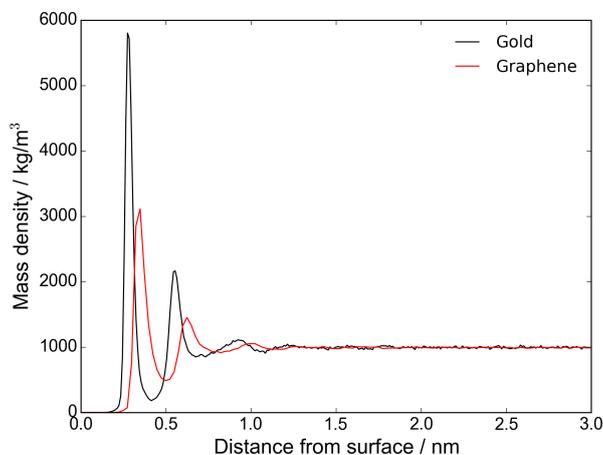


Figure 1. Mass density distribution of water perpendicular to the (111) gold surface (black line) and to a single layer of carbon atoms (graphene monolayer, red line). The results were obtained by classical MD simulations using the interaction parameters of Ref. 1 and 2, respectively. Note that a similar structure of water on graphene was recently obtained through *ab initio* calculations³.

2 Algorithms to Calculate Solid-Liquid Work of Adhesion

In this section we define the solid-liquid work of adhesion and present two algorithms derived by us to calculate it. The solid-liquid work of adhesion W_{SL} is defined as the reversible work to separate a solid and a liquid initially at contact and to bring them to a distance where they no longer interact. It is expressed in terms of surface and interfacial tensions following:

$$W_{SL} = \gamma_S + \gamma_L - \gamma_{SL} \quad (1)$$

where γ_S is the solid surface tension and γ_L is the liquid surface tension, whereas γ_{SL} is the solid-liquid interfacial tension. In Eq. 1, it is assumed that liquid has a vapour pressure low enough such that the excess amount of vapour adsorbed on the surface is negligible⁴. Moreover, solids are considered to be stiff enough, such that there is negligible plastic deformation of the surface upon separating solid and liquid. Under these conditions, the solid-liquid interfacial tension (force per unit distance) is equal to the solid-liquid interfacial excess free energy (energy per unit area). We are interested in fluids with low vapour pressure such as water at ambient condition and ionic liquids as well as polymers on substrates like gold or graphitic surfaces. Therefore, those conditions are met in the study of the systems we aim to consider. Since W_{SL} is defined as the free energy change per unit area, it may be accessed through free energy calculations, as discussed in more detail below. Such calculations represent a central question in the characterisation of the thermodynamic properties of liquid and of soft matter systems. They have thus benefited from numerous methodological studies and are well established.

The thermodynamic quantities that define W_{SL} also control the shape of a droplet of a given liquid on a given surface. This shape is measured through the contact angle θ which

relates to W_{SL} through the equation of Young-Dupré below:

$$W_{SL} = \gamma_L(1 + \cos \theta) \quad (2)$$

Eq. 1 and Eq. 2 suggest several numerical approaches that may be employed to quantify W_{SL} . A first possible approach consists in determining γ_{SL} , γ_L and γ_S independently and combine them following Eq. 1. The calculation of γ_L has benefited from several methodological developments⁵, and the estimation of γ_S for stiff and ordered materials is in principle accessible provided that a model to describe the interactions within the solid exist. In contrast much fewer methodologies lead to the determination of γ_{SL} . γ_L is often determined through the spatial integration of the stress tensor anisotropy⁵. This method proves very demanding in terms of computation time for realistic systems. Due to their mathematical nature (second order quantities), the pressure tensor components are characterised by slow convergence, such that several tens or hundreds of nanoseconds are generally required to determine γ_L with statistical uncertainty less than 10 % for polymers or fluids with electrostatic interactions. γ_{SL} may also be determined through the stress tensor analysis mentioned above with the same slow convergence. Alternatively, one may determine γ_L on the one hand, and θ through the simulation of droplets on the other hand, finally to combine them following Eq. 2. However, θ may be affected by size effects in a way that remains poorly understood (see Ref. 1 and references therein). In fact, the dependence of θ on the droplet size is still an open question in surface science⁶.

Rather than calculating interfacial tensions separately and combining them in Eq. 1, we determine W_{SL} directly. We treat this quantity as a free energy change per unit area, and determine it following the formalism of the well established thermodynamic integration approach⁷. This approach leads to faster convergence, such that computation time can be employed to sample extensively the parameter space and gain detailed knowledge that would not be accessible otherwise. In this context, we derived two methods that will be detailed below. These two methods are based on a common observation about the behaviour of water on repulsive surfaces. MD simulations showed that water tends to avoid such surfaces and forms a liquid-vapour like interface in their vicinity. This result had already been obtained in the case of water in contact with spherical repulsive particles whose radius is larger than a few nanometres⁸. We found that it also applies to planar surfaces¹ and exploited it to derive two algorithms, namely the phantom-wall methodology⁹ (Sec. 2.1) and the dry-surface method¹ (Sec. 2.2). In these two approaches, the interface of interest is turned into an effectively repulsive interface. As mentioned above, thermodynamic integration is employed to calculate the Gibbs free energy change ΔG upon turning the actual interface into the effectively repulsive one. This approach can be understood from the following generic formula:

$$\Delta G = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial \mathcal{U}_{SL}}{\partial \lambda} \right\rangle d\lambda \quad (3)$$

where λ is a parameter that quantifies the reversible path along which the transformation from the actual interface (λ_A) to the repulsive interface (λ_B) is carried out. \mathcal{U}_{SL} is the total solid-liquid interaction energy and the brackets denote an average in an appropriate statistical ensemble. Eq. 3 shows the connection between ΔG which is directly proportional to the macroscopic quantity W_{SL} and the microscopic quantity \mathcal{U}_{SL} which directly depends on the solid-liquid intermolecular interactions. Both the phantom-wall and the dry-surface

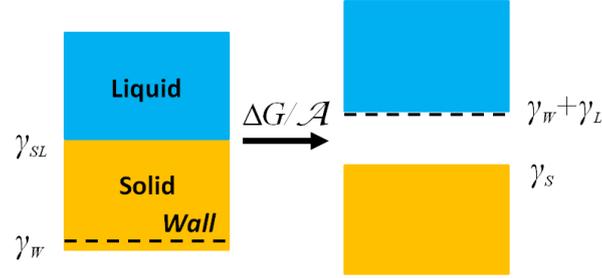


Figure 2. Principle of the phantom-wall method to obtain the solid-liquid work of adhesion. Only the initial step where liquid only interacts with solid and the final step where liquid only interacts with the phantom wall are represented.

methods are implemented in such a way that the free energy change per unit area of this process is equal to $\gamma_s + \gamma_L - \gamma_{SL}$, i.e. the quantity that defines W_{SL} (Eq. 1).

2.1 The Phantom-Wall Algorithm

In the phantom-wall approach⁹, the solid-liquid interface of interest is turned into a repulsive interface by the action of a wall. This wall is initially present within the solid surface (see Fig. 2), but has no interaction with liquid. The wall is then reversibly shifted perpendicular to the surface such that it interacts with liquid. At the end of the process, the wall is located at a distance such that the liquid only interacts with it and no longer with the solid. The Gibbs free energy change per unit area associated with the transformation described above is:

$$\frac{\Delta G}{\mathcal{A}} = \gamma_s + \gamma_L - \gamma_{SL} + P_N \Delta V \quad (4)$$

where \mathcal{A} is the cross-sectional area of the interface, P_N is the pressure component perpendicular to the surface and ΔV is the volume change of the system that arises from the wall's displacement. The direct outputs of MD simulations are ΔG and $P_N \Delta V$, whose determination directly leads to W_{SL} .

The phantom-wall approach proved to be particularly adapted for the study of rough surfaces^{10,11}, as will be demonstrated in a concrete example below.

2.2 The Dry-Surface Approach

Following the dry-surface approach¹, a given solid-liquid interface of interest is turned into an effective repulsive interface. This is realised by turning the actual interfacial interaction potential into an effectively repulsive potential, while maintaining the solid-solid and liquid-liquid interactions unchanged. Similarly to the phantom-wall approach, water avoids the surface as it is turned repulsive. Consequently, the free energy change per unit area associated with this process is equal to $\gamma_s + \gamma_L - \gamma_{SL}$, which is the quantity that defines W_{SL} (see Eq. 1).

The dry-surface method is well adapted to smooth interfaces and may be used to optimise force-field parameters based on experimental quantities such as contact angles, as will be discussed below. It represents a valuable tool to rapidly sample the parameter space. We have recently generalised its application to fluids other than water and to coarse-grained models in which select degrees of freedom are removed¹², such that phenomena occurring at time and length scales larger than accessible through atomistic simulations may be studied. We briefly discuss this extension in Sec. 4.

3 The Interaction between Water and Graphene

The question of the strength of the interaction between a given compound and graphene is of crucial importance in materials where graphene flakes or even carbon nanotubes are embedded in a polymer matrix or dispersed in a fluid, as well as when graphene is used to coat a surface and is in direct contact with liquid droplets. In a series of experimental works supported by MD simulations, measurements of the contact angle of water droplets on graphene layers adsorbed on various substrates were performed¹³⁻¹⁵. It was found that coating a substrate may have different effect on the wetting behaviour of the coated surface. The contact angle on hydrophilic surfaces like gold is almost not affected by the presence of graphene¹³. In contrast, the contact angle on coated hydrophobic surfaces was found to be influenced by the presence of graphene¹⁴. Hydrophilic surfaces which form hydrogen bonds with water are turned into much less hydrophilic substrates when coated with graphene¹³. In fact, the presence of graphene prevents the possibility to form hydrogen bonds between water and the underlying substrate, interactions which are at the origin of the strong interaction between water and the hydrophilic bare substrate.

The experiments mentioned above were conducted with surfaces coated with graphene, such that inferring the intrinsic wetting behaviour of graphene is not straightforward. In this context we performed MD simulations of water on isolated graphene surfaces, i.e. with no substrate, and calculated W_{SL} through phantom-wall calculations depending on the number of graphene layers¹⁶. We showed that and explained why the contact angle of water on a single graphene layer differs from its value on graphite (quasi-infinite stack of graphene layers) only by a few degrees. Our conclusion was confirmed by experiments on super-hydrophobic surfaces¹⁴. We found that the interaction between water and graphene is short-ranged, such that the interaction with the first carbon layer of graphite dominates. The second and third layers have a weaker influence, and layers beyond the third one play a negligible role. Thus, if one removes all carbon layers but the last one of a stack, the interaction between water and the surface is only weakly modified. At the macroscopic scale, a small increase in the contact angle would be observed when turning graphite into mono-layer graphene¹⁶.

Further experiments have shown that the contact angle of water on graphite surfaces is very sensitive to the adsorption of volatile organic compounds^{17,18}. These compounds are hydrophobic molecules like short alkanes and alkenes which make the substrates on which they adsorb appear effectively more hydrophobic than they are. MD simulations have confirmed that the adsorption of alkanes on graphite increases the contact angle of water drops, but no molecular mechanism has been proposed¹⁸. In a work performed on JUROPA, we studied the effect of superficial roughness on the work of adhesion of water on graphite through phantom-wall calculations¹⁰. We employed graphite model surfaces

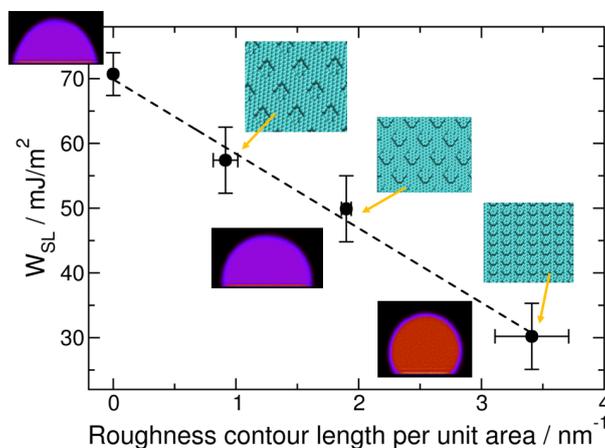


Figure 3. Work of adhesion for water on graphite with superficial defects of various contour length. The defect patterns are represented on the plot next to the respective value of W_{SL} they yield. Additionally, the shape that a nanometre sized droplet takes on these surfaces is represented in purple and red. As expected from Eq. 2, the contact angle increases when W_{SL} decreases, i.e. when graphite becomes more hydrophobic.

where atoms of the top layer were selectively removed to form specific patterns whose contour length was varied (see Fig. 3). Surfaces with such superficial defects can serve as model surfaces to study the effect of the adsorption of hydrocarbons on the wetting behaviour of a surface. Indeed, alkane molecules have interaction with water which is close to the interaction between water and graphitic carbon atoms. Moreover, surface defects may be seen as isolated molecules or small clusters of molecules adsorbed on the surface. Similarly to the effect of alkane adsorption, we found that the work of adhesion of water on rough surfaces decreases compared with smooth defect-free surfaces. In other words, the surface becomes more hydrophobic, when superficial roughness is implemented. We also showed (as illustrated in Fig. 3) that W_{SL} is proportional to the contour length per unit area of the defects. This behaviour was explained by the fact that the upper corners of the defects represent high energy sites for water due to the missing carbon atoms. In order to minimise the perturbation of its hydrogen-bonding network induced by the presence of the surface, water is forced to experience these high energy sites, and the energy of the system increases along these defect corners¹⁰. Hence the linear decrease of W_{SL} reported in Fig. 3. It can thus be observed that molecular simulations represent a valuable tool to both confirm experimental trends and propose models to interpret them. In this regard, it is interesting to note that all the experimental works reported above were published with MD simulation results.

The fact that adsorbed volatile organic compounds have a non-negligible effect on the contact angle of water on graphite implies that the intrinsic interaction between water and graphitic surfaces is stronger than previously thought^{17,18}. Precise contact angle measurements are of crucial importance because this macroscopic quantity may be used to optimise classical force-field parameters. For example, MD simulations of nanometre sized droplets were used to parameterise the water-carbon interaction used in numerous sim-

ulation works¹⁹. Quantum calculations may also be employed to obtain parameters for classical force-fields². It is important to note that contact angle and work of adhesion depend on liquid-liquid, solid-solid but also on solid-liquid interactions. Thus these quantities are observables which inherently depend on the collective behaviour of water molecules. There is *a priori* no direct connection between the single water molecule energies delivered by quantum calculations and the macroscopic contact angle and W_{SL} values measured in experiments. Dry-surface calculations of W_{SL} performed in our group on JUROPA have achieved an important step towards this missing link¹⁹. We show how to connect W_{SL} obtained from experiments and MD simulations to the binding energy of water molecules on graphene surfaces obtained by quantum calculations.

4 Concluding Remarks

We have presented and discussed the capabilities of two methodologies derived by us to calculate the solid-liquid work of adhesion. This quantity characterises the thermodynamics of solid-liquid interfaces which play a major role in systems whose behaviour is dominated by such interfaces. Results concerning the interface between water and graphene layers were presented. The dry-surface method was recently extended to coarse-grained (CG) models. In such models, the number of atoms is reduced such that phenomena like the assembly of nanoparticles occurring at a more mesoscopic scale can be tackled. When the number of degrees of freedom of a given system is decreased, its potential energy surface is also affected. We then addressed the question of how the interfacial thermodynamics is modified by the process of coarse-graining, a topic where little progress has been reported so far. We are interested in CG models which retain the chemical reality of the system of interest, i.e. models beyond generic behaviour. As a first step, we characterised the work of adhesion of *n*-hexane on graphene layers and formulated general recommendations on how to design CG potentials to preserve the value of W_{SL} upon coarse-graining. This work is expected to enhance the understanding of CG potentials for heterogeneous interfacial systems, systems we will address in future works.

Acknowledgements

We thank the John von Neumann Institute for Computing at the Jülich Supercomputing Centre for allocating computer time on JUROPA and JURECA. We are grateful to the German Research Foundation (DFG) for financial support within the Cluster of Excellence 259 “Smart Interfaces: Understanding and Designing Fluid Boundaries” and the Collaborative Research Center TRR 146 “Multiscale Simulation Methods for Soft Matter Systems”. FL acknowledges the group of Nico van der Vegt for fruitful collaborations.

References

1. F. Leroy, and F. Müller-Plathe, *Dry-surface simulation method for the determination of the work of adhesion of solid-liquid interfaces*, *Langmuir* **31**, 8335–8345, 2015.
2. Y. Wu, and N. R. Aluru, *Graphitic carbon-water nonbonded interaction parameters*, *J. Phys. Chem. B* **117**, 8802–8813, 2013.

3. G. Tocci, L. Joly, and A. Michaelides, *Friction of water on graphene and hexagonal boron nitride from ab initio methods: very different slippage despite very similar interface structures*, *Nano Lett.* **14**, 6872–6877, 2014.
4. L. J. M. Schlangen, L. K. Koopal, M. A. Cohen Stuart, and J. Lyklema, *Wettability: thermodynamic relationships between vapour adsorption and wetting*, *Coll. Surf. A: Physicochem. Eng. Aspects* **89**, 157–167, 1994.
5. G. J. Gloor, G. Jackson, F. J. Blas, and E. de Miguel, *Test-area simulation method for the direct determination of the interfacial tension of systems with continuous or discontinuous potentials*, *J. Chem. Phys.* **123**, 134703, 2005.
6. R. David, and A. W. Neumann, *Empirical equation to account for the length dependence of line tension*, *Langmuir* **23**, 11999–12002, 2007.
7. J. G. Kirkwood, and F. P. Buff, *The statistical mechanical theory of solutions*, *J. Chem. Phys.* **19**, 774–777, 1951.
8. D. M. Huang, P. L. Geissler, and D. Chandler, *The statistical mechanical theory of solutions*, *J. Phys. Chem. B* **105**, 6704, 2001.
9. F. Leroy, and F. Müller-Plathe, *Solid-liquid surface free energy of Lennard-Jones liquid on smooth and rough surfaces computed by molecular dynamics using the phantom-wall method*, *J. Chem. Phys.* **133**, 044110, 2010.
10. F. Leroy, and F. Müller-Plathe, *Rationalization of the behavior of solid-liquid surface free energy of water in Cassie and Wenzel wetting states on rugged solid surfaces at the nanometer scale*, *Langmuir* **27**, 637–645, 2011.
11. F. Leroy, and F. Müller-Plathe, *Can continuum thermodynamics characterize Wenzel wetting states of water at the nanometer scale?*, *J. Chem. Theo. Comp.* **8**, 3724–3732, 2012.
12. V. R. Ardham, G. Deichmann, N. F. A. van der Vegt, and F. Leroy, *Solid-liquid work of adhesion of coarse-grained models of n-hexane on graphene layers derived from the conditional reversible work method*, *J. Chem. Phys.* **143**, 243135, 2015. submitted.
13. J. Rafiee, X. Mi, H. Gullapalli, A. V. Thomas, F. Yavari, Y. Shi, P. M. Ajayan, and N. A. Koratkar, *Wetting transparency of graphene*, *Nat. Mater.* **11**, 217–222, 2012.
14. C. J. Shih, Q. H. Wang, S. C. Lin, K. C. Park, Z. Jin, M. S. Strano, and D. Blankschtein, *Breakdown in the wetting transparency of graphene*, *Phys. Rev. Lett.* **109**, 176101, 2012.
15. C. J. Shih, M. S. Strano, and D. Blankschtein, *Wetting translucency of graphene*, *Nat. Mater.* **12**, 866–869, 2013.
16. F. Taherian, V. Marcon, N. F. A. van der Vegt, and F. Leroy, *What is the contact angle of water on graphene?*, *Langmuir* **29**, 1457–1465, 2013.
17. Z. Li, Y. Wang, A. Kozbial, G. Shenoy, F. Zhou, R. McGinley, P. Ireland, B. Morganstein, A. Kunkel, S.P. Surwade, L. Li, and H. Liu, *Effect of airborne contaminants on the wettability of supported graphene and graphite*, *Nat. Mater.* **12**, 925–931, 2013.
18. A. Ashraf, Y. Wu, M. C. Wang, N. R. Aluru, S. A. Dastgheib, and S. W. Nam, *Spectroscopic investigation of the wettability of multilayer graphene using highly ordered pyrolytic graphite as a model material*, *Langmuir* **30**, 12827–12836, 2014.
19. F. Leroy, S. Liu, and J. Zhang, *Parametrizing nonbonded interactions from wetting experiments via the work of adhesion: Example of water on graphene surfaces*, *J. Phys. Chem. C*, DOI: 10.1021/acs.jpcc.5b10267, in press.

Earth and Environment

Earth and Environment

Ulrich Hansen

Institut für Geophysik, Westfälische Wilhelms-Universität Münster,
Corrensstraße 24, 48149 Münster, Germany
E-mail: hansen@earth.uni-muenster.de

The phenomenon of climate change is momentarily at the centre of research efforts in the field of environmental sciences and almost as a matter of course, the first contribution to this year's issue of the NIC Proceedings in the *Earth and Environment* section examines climatic changes on the African continent. We have discovered that the climate is changing and that in fact it seems to change faster than ever. The reasons behind the climatic variations are not yet clear and are still debated. Especially the impact of man on the climate, or on system Earth, in a broader sense, is subject of controversial discussions. While it seems clear to one part of the community that the footprint of mankind upon Earth is too deep, others favour the view, that the variations, as momentarily witnessed, are simply due to natural fluctuations in a complex system. These rather opposing views reflect that our understanding of system Earth is still very patchy. One reason for this is the immense range of spatial and temporal scales on which relevant processes are occurring. Further, we begin to understand that the different geospheres, i.e. hydrosphere, atmosphere and the interior of the Earth are intimately coupled. In 2010 the eruption of the volcano Eyjafjallajökull on Iceland had a significant impact on our daily life. Air traffic even had to be shut down. Volcano eruptions in the past had tremendous influences on the climatic conditions. Vice versa atmospheric conditions can influence even the interior of our planet. According to our knowledge, the presence of an early atmosphere, shortly after formation of the Earth, would have influenced the thermal and chemical evolution of our planet. The picture of a peaceful Earth, if not deranged by mankind, is sometimes unhinged by extreme events, originating from the dynamics of the Earth itself. Especially the large Sumatra earthquake from the 26th of December 2004 and its subsequent tsunami, causing more than 230000 fatalities, reminded us how powerful natural fluctuations can be. While there is little hope to influence these forces or even to predict with high precision their consequences like earthquakes, it seems worthwhile to better understand the inner working of the planet. The second contribution of this issue is concerned with such questions. How do geodynamical processes evolve on global and regional scale? Dense and detailed observational data are of key importance for the understanding of the structure and the dynamics of Earth and Environment. The third contribution describes how an improved picture of the gravity field of the Earth can be obtained. The data can be used to infer the structure of the interior, but also to map mass redistribution at the surface. Clearly some models strive for a picture, as realistic as possible, in order to provide estimates and predictions of satisfactory quality. A different model approach aims at more insight into principal physical mechanisms of particular phenomena. Personally, I prefer to name this approaches numerical experiments, rather than numerical simulations or modelling. Often these efforts are driven by the perception that it is virtually impossible to obtain scientific understanding of a real system in its entity and minutest detail. The standard procedure in physical sciences is to understand

the behaviour through a model which is obtained after simplifying the physics to the fundamental essentials. The fourth contribution treads this type of path. Its subject is the role of rotation in turbulent convection in Geo- and Astrophysical systems.

Let's have a closer look at the individual contributions:

In their article "Very High Resolution Simulations of African Climate with the regional Model REMO", Hänslér *et al.*, report on their experiences, gained while running a high resolution model for the climate on the African continent. Not too surprising in my view, they find that detailed data are required in order to determine mitigation measures on local scales. The authors also report on limitations of the JUQUEEN system with respect to their simulations, due to the relative small amount of memory as available for each individual CPU. They consider this as a typical problem, arising when geophysical models are run on High Performance Computer Systems.

The next contribution by Kaus *et al.* reports likewise on a geophysical modelling effort, however it aims at phenomena operating on scales which are vastly different from the previous one. In their article "Forward and Inverse Modelling of Lithospheric Deformation on Geological Timescales" the authors describe the properties of a new code LaMEM which is able to efficiently handle the elliptic system of equations, as arising when the incompressible Stokes equation is tackled. The Stokes equation, in this case including highly variable material properties describes the flow of rocks and is such of central importance in geodynamical modelling. The authors describe model results, as obtained on crustal and on mantle scale. Interestingly they combine their forward models with inverse calculations, aiming at a better estimate of the viscosity of the material. Obviously, the combination of forward and inverse modelling seems a promising strategy for future research.

The contribution "Computational Aspects of High-Resolution Global Gravity Field Determination" by Brockmann and Schuh, describes a new method allowing for an efficient determination of high order coefficients of the spherical harmonic expansion of the gravitational potential of the Earth. Through satellite missions a detailed analysis of the gravitational field of the Earth is possible. Computer procedures to do so are very demanding. In their present contribution Brockmann and Schuh introduce a reordering strategy for the normal equations, allowing for an improved treatment of the described potential expansion.

Last but not least, the contribution of Stellmach *et al.* focuses on the influence of rotation on turbulent convection. Convective motions are virtually ubiquitous in geophysical and astrophysical systems. They determine the energy transport in the sun from its interior to the surface, the generation of the Earth's magnetic field is intimately coupled to convection in the Earth's outer core and furthermore the dynamics of the oceans and atmospheres is largely driven by convection. Many geo- and astrophysical systems are huge in extent, such that rotational forces play a dominant role. Fundamental aspect of rotating convective flows, for example the heat transport efficiency, are poorly understood. In this contribution the authors employ a powerful numerical model in order to study various effects of rotation. Many fascinating phenomena appear in these flows. Certainly worthwhile to underline is the upscale transport of energy and the resulting appearance of large scale coherent structures. This may possibly explain the large scale structure in the Earth's magnetic field. Furthermore, alternating winds can result in rotating convection, once compressibility of the material is taken into account. The beautiful clouds in Jupiter's atmosphere and its banded structure, as shown in the contribution of Stellmach *et al.*, can possibly be explained by the interplay of rotation, convection and compressibility effects.

Very High Resolution Simulations of African Climate with the Regional Climate Model REMO

Andreas Hänsler¹, Nikolay Koldunov¹, Dmitry Sein²,
Walter Sauf³, and Daniela Jacob¹

¹ Climate Service Center Germany (GERICS), Helmholtz-Zentrum Geesthacht,
20095 Hamburg, Germany
E-mail: {andreas.haensler, nikolay.koldunov, daniela.jacob}@hzg.de

² Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung,
Am Handelshafen 12, 27570 Bremerhaven
E-mail: dmitry.sein@awi.de

³ meteoGRID, Tieloh 10, 22307 Hamburg
E-mail: walter.sauf@meteogrid.de

The regional climate model REMO is used to perform high resolution simulations of African climate. Simulated seasonal precipitation is analysed and compared to a lower resolution run and observational data. We show that the high resolution model is able to represent small scale features of the precipitation distribution and therefore demonstrates considerable improvement compared to the lower resolution version. We also assessed the performance of REMO on the JUQUEEN supercomputer (Jülich Supercomputing Centre, JSC) and found, that it is not the optimal system for conducting regional climate model simulations. Computers with architecture similar to JUROPA or JURECA (JSC) are better suited for this type of applications.

1 Introduction and Motivation

Findings of the 5th IPCC Assessment report (IPCC AR5) released in the year 2013 indicate that the African continent is a hotspot of future climate change. In order to be able to develop options and strategies to adapt to climate change spatially, very detailed climate change information and subsequent climate change impact assessments are required. Therefore the current generation of climate models has to be continuously developed and improved to be able to provide reliable high-resolution climate change information.

The latest generation of global circulation models used in the fifth Coupled Climate Intercomparison Project (CMIP5¹), which is the basis for IPCC AR5 Africa chapter², still only delivers rather coarse scale information on future climate change. Therefore dynamical and statistical downscaling of future climate change projections is required to deliver information on a scale to be used for regional to local climate impact assessments. Hence, the regional climate modelling community launched the CORDEX (Coordinated Regional Climate Downscaling Experiment), which delivers a multi-model multi-scenario ensemble of downscaled climate change information at a spatial resolution of 0.5° (all continents) and 0.11° (Europe only) for almost all inhabited regions of the world³.

Nevertheless, for defining local scale adaptation strategies even CORDEX does not deliver data on a scale that is able to represent local climate features. Hence a further step to provide very high resolved regional climate change projections has to be made. However, running regional climate models on a very high spatial resolution requires a

large amount of computational resources. Therefore simulations are either conducted for rather small regions or for only a very short time period.

In this paper we describe the findings of an initial sensitivity experiment with the regional climate model REMO⁴ at a spatial resolution of 0.11° covering the whole African continent. To conduct this simulation, REMO has been implemented at the JUQUEEN supercomputer in the Jülich Supercomputing Centre (JSC). REMO is a well-known regional climate model which has been used and analysed several times over Africa. Within the CORDEX initiative several simulations have been conducted with REMO for current and future conditions⁵⁻⁷. Already earlier the added value of using REMO at a relatively high horizontal resolution of approximately 18 km could be demonstrated over south-western Africa⁸.

Due to the large amount of computational resources required we conducted only a one-year sensitivity experiment so far. Results of this experiment with respect to the simulation of main atmospheric conditions over Africa will be compared to the REMO CORDEX-Africa hindcast simulations. On top of this analysis also the performance of REMO on the JUQUEEN high performance computing system is analysed. Based on the findings of these assessments it will be decided if an extension of the simulation to longer time periods and maybe also for downscaling of the future climate is worthwhile.

2 REMO Model and Experiment Setup

2.1 Model Description

The sensitivity simulation described in this paper is conducted with the three-dimensional hydrostatic limited-area atmospheric model REMO (version 2009). The hydrostatic version of the regional climate model REMO has been described in several publications before in detail^{9,4} and is credible to be used down to a horizontal resolution of approximately 10 km. In the following, we briefly mention the main characteristics of REMO.

The dynamical core of REMO as well as its discretisation in space and time are based on the Europa-Model of the German Weather service¹⁰. In the original version of REMO the physical parameterisations were taken from the global climate model ECHAM version 4¹¹. The prognostic variables are surface pressure, horizontal wind components, temperature, water vapour, liquid water, and cloud ice. More detailed description of the model can be found in Ref. 4, 6, 12.

2.2 Experiment Setup

REMO was run at a horizontal resolution of 0.11° over the whole African continent (see Fig. 1). The domain was chosen in a way that it reflects the boundaries of the CORDEX Africa simulation domain (see Nikulin⁵ for details). To span the domain 901×901 gridboxes are required in north-south and east-west direction, respectively. With a total of 811891 gridboxes this domain is about 17 times larger than the standard CORDEX Africa domain. To resolve the vertical component 31 vertical layers were used.

With this very large domain only a short simulation period could be calculated for this first sensitivity study. Hence, existing and already in equilibrium state soil moisture and soil temperatures were taken from the REMO CORDEX-Africa hindcast experiment⁵ and

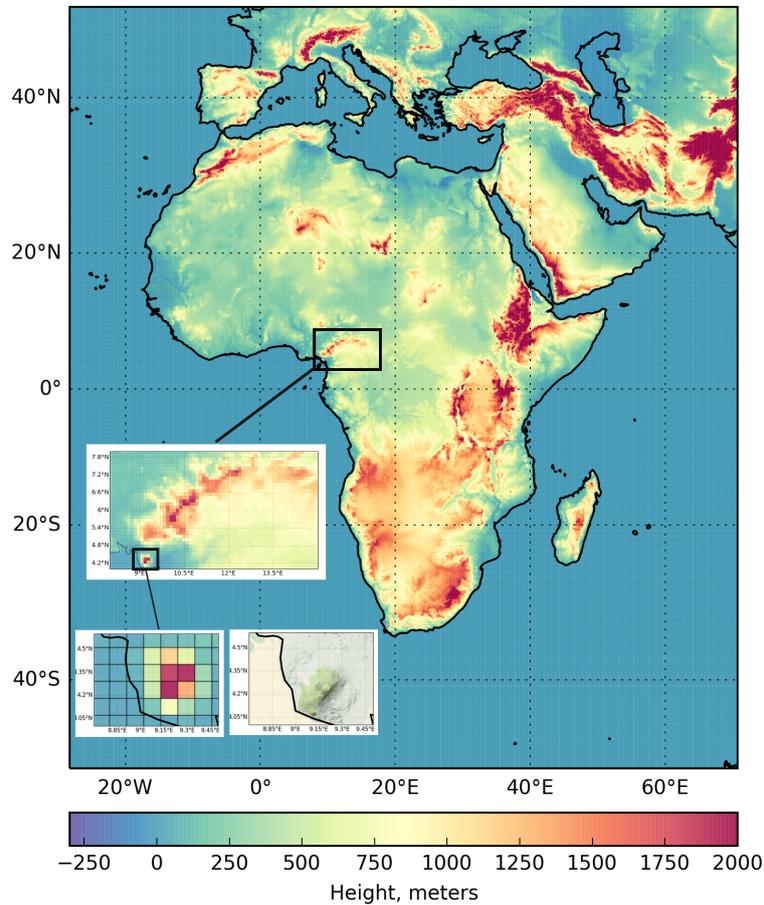


Figure 1. Orography. Zoom to mount Cameroun.

remapped onto this domain in order to reduce the spin-up time to six months. After spin-up, REMO was run for the full year 1979. Lateral boundary conditions for this sensitivity simulation were taken from ERA-Interim reanalysis.

2.3 Parallelisation and Performance

REMO is parallelised using MPI. The model domain is split up into several subdomains which are distributed among CPU cores. Each subdomain has overlapping gridboxes with neighbouring subdomains (halos), so that the information between subdomains can be transferred. This information exchange has to be performed every time step so considerable amount of interprocess communication is required. From our previous experience on another high performance computing systems, the optimal number of grid boxes for a subdomain is about 400 per core.

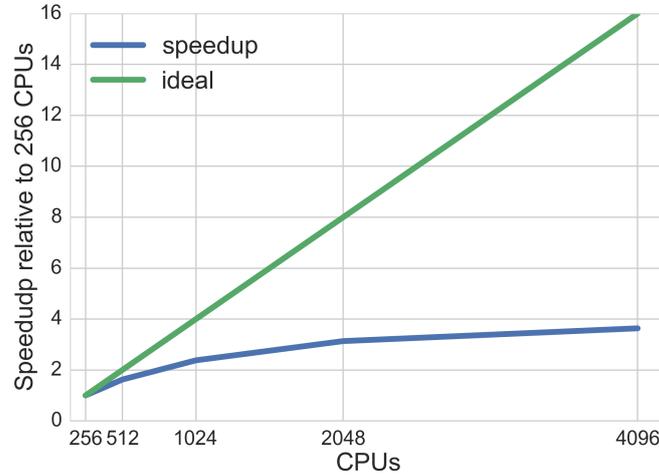


Figure 2. Scaling behaviour of REMO on JUQUEEN at JSC. This data was obtained with a domain size of 901 x 901 grid-boxes.

The amount of memory required by each CPU depends on the size of the model sub-domains and usually is about 4 GB per core. This memory is used for storing fields of the model’s prognostic and diagnostic variables as well as the boundary condition fields. The IO of REMO is not parallel, so all reading and writing is done by one process. The snapshots of two dimensional fields are usually stored every model hour, while snapshots of three dimensional fields are stored every six hours. The amount of data generated by the model depends on the size of the model domain, frequency of the snapshots and number of variables. In the present experiment one model month generates about 430 GB of output data, that makes it 5 TB for one year.

2.3.1 Performance on JUQUEEN

We perform our experiment on the JUQUEEN supercomputer, that has 28,672 nodes equipped with IBM PowerPC® A2, 1.6 GHz CPU cores, 16 cores per node. The amount of memory per node is 16 GB, so there is only 1 GB per core available. As mentioned before this amount of memory is not enough for REMO, so we had to use a reduced number of CPU cores per node (2 CPUs instead of 16 for some configurations). Since CPU time is accounted on the basis of the actual node usage, huge amounts of CPU time were charged, but only a small fraction of cores was actually computing. This has to be taken in to account when one is planing to use geophysical models on JUQUEEN.

As can be seen on the Fig. 1, the model’s performance does not scale very well with an increase of CPU cores used and reaches the plateau relatively fast. The limiting factor here is probably the substantial amount of inter-process communications which is consistent with our previous experience on other high performance computing systems.

Based on our findings, we conclude that the JUQUEEN is more suitable for pure dy-

namical applications, e.g. turbulence. However when the physics starts to play a big role, and therefore the payload of each individual processor significantly increases, JUQUEEN, with its relatively slow processors becomes inefficient. Large amount of physics related computations are characteristics for most of the ocean and atmospheric models and in REMO it is ca. 80% of the processor time. On the contrary, our experience with running REMO on JUROPA and its successor JURECA with their larger amount of available memory and faster processors is very positive and we advise to use supercomputers that use a similar architecture for geophysical applications.

3 Model Results

In order to estimate the quality of the sensitivity simulations various output fields of REMO have been validated against observations. Here we only present the analyses of the seasonal cycle of total precipitation for the REMO 0.11° sensitivity simulation (REMO₁₁) compared to the 50 km REMO CORDEX-Africa hindcast simulation (REMO₅₀) as well as to the CRU gridded observation set¹³. Although the various datasets included into the analysis cover different time periods (only the year 1979 for REMO₁₁ versus the five year period from 1979 to 1984 for REMO₅₀ and CRU), the analysis still allows to identify, if the major precipitation features are skillfully reproduced in the high-resolution simulation.

Generally, seasonal rainfall characteristics over most parts of Africa are strongly linked to the position of the ITCZ, leading to a pronounced annual cycle with an extended rainy and dry season, respectively, over large parts of the continent. Only in the tropical regions closer to the equator two rainy seasons are present. On top of the strong rainfall seasonality, the huge precipitation gradient between the tropics (more than 2000 mm/yr in the Congo basin region) and the desert regions in the north and south makes it difficult for climate models to simulate the precipitation amounts in the right magnitude.

On a regional scale, both the temporal and spatial precipitation patterns represented by the REMO model simulations are skillful. In general, there seems to be a rather small difference in simulated seasonal mean precipitation in the REMO₁₁ and REMO₅₀ simulations (Fig. 3) with a tendency towards wetter conditions in the spatially higher resolved REMO₁₁ simulation. When compared to the CRU dataset, a general dry bias of about 60 to 70 % persists for the whole year in both REMO simulations in the region around the Lake Victoria. Additionally a wet bias is present over south-eastern parts of southern Africa during the rainy season from September to March (up to about 60 to 80% during peak rain season). This wet bias has been reported before and seems to be related to an overestimation of the seasonal heat low conditions in REMO⁸.

Nevertheless while interpreting the results of the sensitivity simulation, two things have to be kept in mind. First, in the high-resolution simulation the seasonal mean of only one year is compared to the average seasonal mean of five years. Hence also the year-to-year variation of precipitation, which is known to be quite substantial over large parts of Africa (e.g. Ref 14, 15), could lead to enhanced biases. Second, also the CRU gridded observation set is not perfect as it is compiled on the basis of a limited number of station records and therefore includes artefacts due to interpolation over large areas¹³.

The differences between the two REMO simulations might also occur due to a set of parametrisations which might not fully represent the spatial scale of the simulation. Especially the parametrisation of cloud processes is linked to the horizontal resolution and

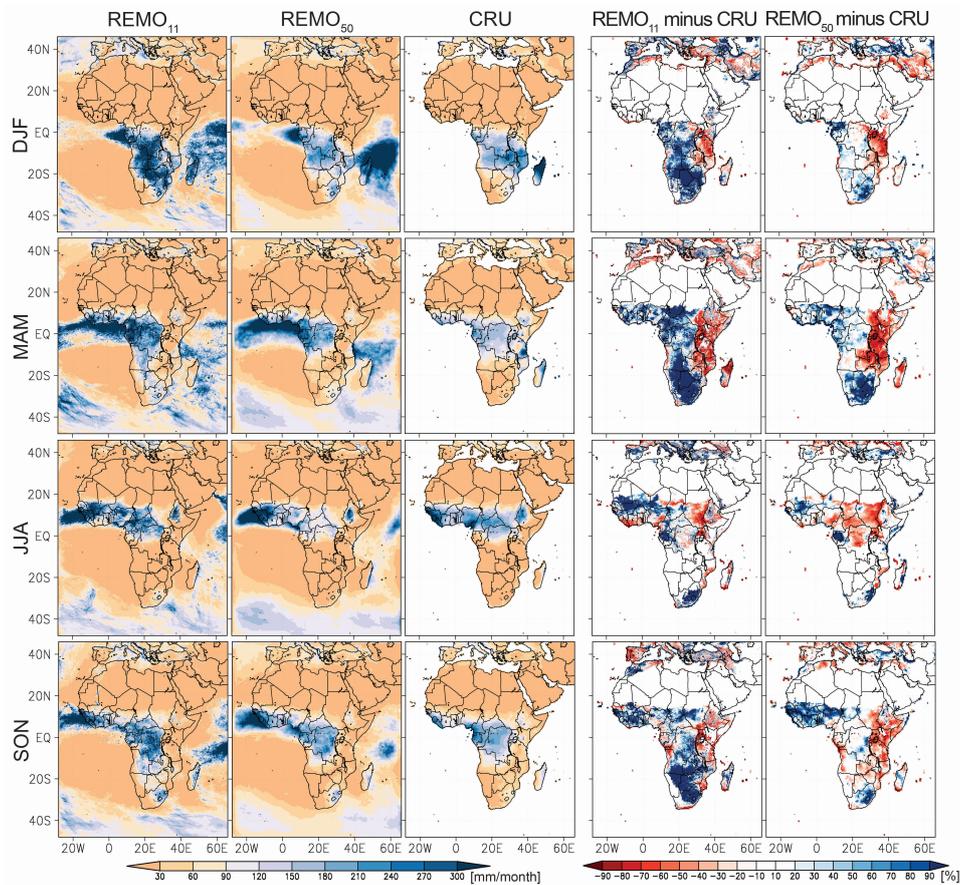


Figure 3. Precipitation in DJF and JJA for REMO₁₁, REMO₅₀ and CRU data.

potentially need to be adapted to the scale, particularly when simulating over tropical regions. However, as a starting point, we used the identical parametrisation for this sensitivity simulation as was used in the REMO₅₀ simulation.

More interesting than the larger regional scale features is the simulation of more complex small scale features. A good example in Africa is the Mount Cameroon region, at the coast of Cameroon. This is the region that receives more than 11000 mm of precipitation a year, which is the maximum annual rainfall amount over the whole of Africa. Fig. 4 shows a zoom of simulated annual precipitation sum over that region. Here the improved representation of complex topographic features in the higher resolved REMO₁₁ simulations leads to a more realistic representation of annual precipitation sums compared to lower resolved simulations and also compared to the gridded observation datasets.

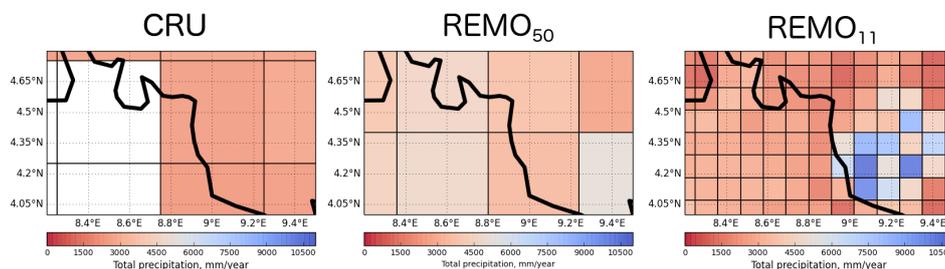


Figure 4. Simulation of small scale features in REMO₁₁ compared to REMO₅₀ and CRU data for the Mount Cameroon region. For orography of the region see Fig. 1.

4 Concluding Remarks

We have described the results obtained by a high resolution simulation with the regional climate model REMO over the African continent and our experience in using the JUQUEEN supercomputer at JSC for this computational very demanding run.

We identified that the model realistically reproduces the major spatial and temporal patterns of seasonal mean precipitation. Remaining biases might be linked to imperfect model parametrisations for this high spatial resolution; however more sensitivity tests have to be done to be able to better judge on this. Additionally, we showed that a better representation of very complex terrain in the high-resolution simulation leads to a better representation of the precipitation amounts along these features, when compared to lower resolved simulations. Hence this underlines the need for spatially detailed climate change information in order to be able to design and shape future adaptation and mitigation measures on the local scale.

Regarding the REMO performance on the JUQUEEN, we found that the performance is limited by the relatively small amount of memory available for one CPU core. Other limiting factors are the relatively low speed of the individual processors and inter process communications. The later is not unique for the JUQUEEN system, but is a common problem when running geophysical models on high performance computer systems.

Acknowledgements

The computations were performed with a grant of computer time provided by the Forschungszentrum Jülich. NK is supported by the funding from the Federal Ministry of Education and Research, Germany, for GLACINDIA project (contract 033L164).

References

1. K. E. Taylor, R. J. Stouffer, and G. A. Meehl, *An Overview of CMIP5 and the Experiment Design*, Bull. Amer. Meteor. Soc., **93**, no. 4, 485–498, Oct. 2011.

2. I. Niang, O. C. Ruppel, M. A. Abdrabo, A. Essel, C. Lennard, J. Padgham, and P. Urquhart, "Africa", in: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, 2014.
3. F. Giorgi, C. Jones, G. R. Asrar, et al., *Addressing climate information needs at the regional level: the CORDEX framework*, World Meteorological Organization (WMO) Bulletin, **58**, no. 3, 175, 2009.
4. D. Jacob, *A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin*, Meteorology and Atmospheric Physics, **77**, no. 1-4, 61–73, 2001.
5. G. Nikulin, C. Jones, F. Giorgi, G. Asrar, M. Büchner, R. Cerezo-Mota, O. Bøssing Christensen, M. Déqué, J. Fernandez, A. Hänsler, et al., *Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations*, Journal of Climate, **25**, no. 18, 6057–6078, 2012.
6. A. Hänsler, F. Saeed, and D. Jacob, *Assessing the robustness of projected precipitation changes over central Africa on the basis of a multitude of global and regional climate projections*, Climatic Change, **121**, no. 2, 349–363, 2013.
7. M. L. Mbaye, A. Haensler, S. Hagemann, A. T. Gaye, C. Moseley, and A. Afouda, *Impact of statistical bias correction on the projected climate change signals of the regional climate model REMO over the Senegal River Basin*, International Journal of Climatology, 2015.
8. A. Haensler, S. Hagemann, and D. Jacob, *The role of the simulation setup in a long-term high-resolution climate change projection for the southern African region*, Theoretical and applied climatology, **106**, no. 1-2, 153–169, 2011.
9. D. Jacob and R. Podzun, *Sensitivity studies with the regional climate model REMO*, Meteorology and Atmospheric Physics, **63**, no. 1-2, 119–129, 1997.
10. D. Majewski, *The Europa-Modell of the Deutscher Wetterdienst*, in: ECMWF Seminar on numerical methods in atmospheric models, vol. 2, pp. 147–191, 1991.
11. E. Roeckner, K. Arpe, M. Bengtsson, M. Christoph, M. Claussen, L. Duemenil, M. Esch, M. Giorgetta, U. Schlese, and U. Schulzweida, *The Atmospheric General Circulation Model ECHAM4: Model Description and Simulation of PresentDay Climate*, Max Planck Institute for Meteorology, Technical report, **218**, 1996.
12. D. V. Sein, U. Mikolajewicz, M. Gröger, I. Fast, W. Cabos, J. G. Pinto, S. Hagemann, T. Semmler, A. Izquierdo, and D. Jacob, *Regionally coupled atmosphere-ocean-sea ice-marine biogeochemistry model ROM: 1. Description and validation*, Journal of Advances in Modeling Earth Systems, **7**, no. 1, 268–304, 2015.
13. M. New, D. Lister, M. Hulme, and I. Makin, *A high-resolution data set of surface climate over global land areas*, Climate research, **21**, no. 1, 1–25, 2002.
14. N. Balas, S. E. Nicholson, and D. Klotter, *The relationship of rainfall variability in West Central Africa to sea-surface temperature fluctuations*, International journal of climatology, **27**, no. 10, 1335–1349, 2007.
15. S. E. Nicholson, *Climatic and environmental change in Africa during the last two centuries*, Climate Research, **17**, no. 2, 123–144, 2001.

Forward and Inverse Modelling of Lithospheric Deformation on Geological Timescales

Boris J. P. Kaus^{1,2,3}, Anton A. Popov¹, Tobias S. Baumann¹, Adina E. Püsk¹, Arthur Bauville¹, Naiara Fernandez^{1,4}, and Marine Collignon⁵

¹ Institute of Geosciences, Johannes Gutenberg University Mainz,
Joh. Joachim Becherweg 21, 55128 Mainz, Germany
E-mail: {kaus, popov, baumann, pusesoek, abauvill}@uni-mainz.de

² Center for Computational Sciences, JGU Mainz

³ Center for Volcanoes and Atmosphere in Magmatic Open Systems (VAMOS), JGU Mainz

⁴ now at: Bureau of Economic Geology, University of Texas, Austin, USA

⁵ Centre for Earth Evolution and Dynamics, University of Oslo, Oslo, Norway

Geological processes such as mountain belt formation, subduction of tectonic plates and the development of sedimentary basins occur on a million-year timescale and involve rocks that have nonlinear visco-elasto-plastic material properties and experienced very large deformations. In order to simulate such processes in 3D, we developed a scalable parallel code, LaMEM, that employs a staggered finite difference discretisation combined with a marker and cell approach. Here, we describe the numerical approach and discuss some case studies in which we employed the code (i) to study the physics of crustal scale folding and faulting, (ii) to understand how continental collision might result in mountain belt and plateau formation, (iii) how it can be combined with an inversion strategy to constrain the rheology of the crust and lithosphere.

1 Introduction

Computational geodynamics uses numerical modelling to understand fundamental geoscientific questions such as: Why do we have plate tectonics on Earth and not on other planets? How did mountain belts such as the Alps and the Himalaya form and what is the role of erosion in this? At the same time, computational geodynamics also addresses more practical questions related to the evolution of so-called fold-and-thrust-belts and salt structures, which are closely linked with the majority of the world's oil reservoirs. Numerically, the problems are challenging as the rheology of rocks varies from elasto-plastic at low temperatures (close to the Earth surface), to viscous at higher temperatures (deeper in the lithosphere and mantle). In addition, the codes should be able to handle very large strains (even after plastic material failure has occurred). As geological processes are slow, inertial terms are negligible and one has to solve the (incompressible) Stokes equations but with visco-plastic or visco-elasto-plastic rheologies, which results in an elliptic system of equations with strongly varying coefficients (that can be 6 orders of magnitude or more over several grid cells). Over the last few years, we have developed a new code (LaMEM - Lithosphere and Mantle Evolution Model), which fulfils these requirements and has a range of multigrid preconditioners combined with Newton iterations for nonlinearities to solve the resulting equations. In addition, we have coupled the code with a Monte-Carlo inversion approach, in order to better constrain the mechanical structure of active mountain belts.

We will describe the mathematical background and numerical implementation of the forward and inverse modelling approach in Sec. 2, and discuss some modelling results in Sec. 3.

2 Numerical Approach

2.1 Mathematical Approach

We solve the coupled system of momentum, mass, and energy conservation equations, respectively, with velocity (v_i), pressure (p) and temperature (T) as primary unknowns:

$$\frac{\partial \tau_{ij}}{\partial x_j} - \frac{\partial p}{\partial x_i} + \rho g_i = 0 \quad (1)$$

$$\frac{1}{K} \frac{Dp}{Dt} - \alpha \frac{DT}{Dt} + \frac{\partial v_i}{\partial x_i} = 0 \quad (2)$$

$$\rho C_p \frac{DT}{Dt} = \frac{\partial}{\partial x_i} \left(\lambda \frac{\partial T}{\partial x_i} \right) + H. \quad (3)$$

Here x_i ($i = 1, 2, 3$) denotes Cartesian coordinates, $\tau_{ij} = \sigma_{ij} + p\delta_{ij}$ is the Cauchy stress deviator, ρ density, g_i the gravity acceleration vector, K the bulk modulus, α the thermal expansion coefficient, C_p the specific heat, λ thermal conductivity, H volumetric heat source, and D/Dt stands for the material time derivative, respectively. The visco-elasto-plastic constitutive equation for the deviatoric stress is given by:

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}_{ij}^{el} + \dot{\varepsilon}_{ij}^{vs} + \dot{\varepsilon}_{ij}^{pl} = \frac{\overset{\diamond}{\tau}_{ij}}{2G} + \dot{\varepsilon}_{II}^{vs} \frac{\tau_{ij}}{\tau_{II}} + \dot{\varepsilon}_{II}^{pl} \frac{\tau_{ij}}{\tau_{II}}, \quad (4)$$

where $\dot{\varepsilon}_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) - \frac{1}{3} \frac{\partial v_k}{\partial x_k} \delta_{ij}$ is the deviatoric strain rate tensor, $\dot{\varepsilon}_{ij}^{el}$, $\dot{\varepsilon}_{ij}^{vs}$, $\dot{\varepsilon}_{ij}^{pl}$ are the elastic, viscous and plastic components, respectively, $\overset{\diamond}{\tau}_{ij} = \frac{\partial \tau_{ij}}{\partial t} + \tau_{ik} \omega_{kj} - \omega_{ik} \tau_{kj}$ is the Jaumann objective stress rate, $\omega_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} \right)$ is the spin tensor, G is the elastic shear modulus, and the subscript II denotes the square root of the second invariant of a corresponding tensor, e.g. for the deviatoric stress $\tau_{II} = \left(\frac{1}{2} \tau_{ij} \tau_{ij} \right)^{1/2}$. The magnitude of the viscous creep strain rate is subdivided into diffusion ($\dot{\varepsilon}_l$), and dislocation ($\dot{\varepsilon}_n$) components:

$$\dot{\varepsilon}_{II}^{vs} = \dot{\varepsilon}_l + \dot{\varepsilon}_n = A_l \tau_{II} + A_n (\tau_{II})^n, \quad (5)$$

where n is the stress exponent of the dislocation creep, and the pre-exponential factor (A) of each creep mechanism is defined by:

$$A_l = B_l \exp \left[- \frac{E_l + pV_l}{RT} \right], \quad A_n = B_n \exp \left[- \frac{E_n + pV_n}{RT} \right]. \quad (6)$$

Here B , E , and V denote the creep constant, activation energy, and activation volume, respectively, of the corresponding creep mechanism, and R is the gas constant. The magnitude of the plastic strain rate ($\dot{\varepsilon}_{II}^{pl}$) is determined by enforcing the Drucker-Prager yield criterion:

$$\tau_{II} \leq \tau_Y = \sin(\phi) p + \cos(\phi) c, \quad (7)$$

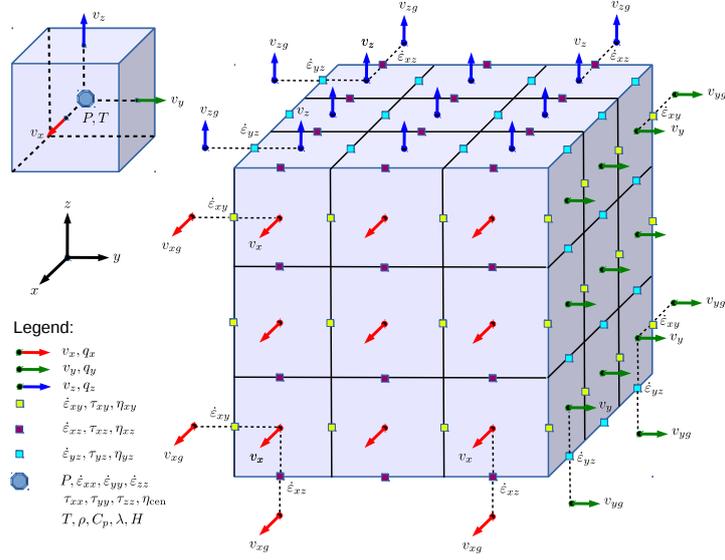


Figure 1. Staggered grid finite difference spatial discretisation, illustrating how the variables are ordered.

where ϕ is the friction angle, and c is the cohesion. The volumetric heat source includes shear heating, controlled by efficiency parameter $0 \leq \chi \leq 1$, and the radiogenic heat (A):

$$H = \chi \tau_{ij} (\dot{\epsilon}_{ij} - \dot{\epsilon}_{ij}^{el}) + \rho A. \quad (8)$$

2.2 Numerical Formulation

We discretise the conservation Eqs. 1–3 in space using staggered grid finite differences¹ as it is a low-order but stable discretisation for (nearly) incompressible fluid flow (see Fig. 1 for grid layout). To achieve scalability on massively parallel machines we use the distributed arrays (DMDA) and iterative solvers (KSP, SNES) from the PETSc library². The free surface is implemented using a so-called sticky air approach, which assigns a relatively low but nonzero viscosity to the air phase, together with an appropriate stabilisation method to allow for sufficiently large time steps^{3,4}. The topography of the free surface is explicitly tracked by an internal 2D grid that covers the entire domain.

We employ a Marker And Cell (MAC) method¹ to track material properties and implement material advection in an Eulerian kinematical framework. To prevent spurious clustering of the material particles (markers) we use a combination of a 4th-order Runge-Kutta method with a conservative velocity interpolation scheme⁵. During the advection, the elastic history stresses from previous time step (τ_{ij}^n) are corrected on the markers to account for the rigid-body rotation, and then interpolated on the edge and cell control volumes (Fig. 1) using the distance-based averaging⁴ to obtain the effective strain rates:

$$\dot{\epsilon}_{ij}^* = \dot{\epsilon}_{ij} + \frac{\tau_{ij}^*}{2G\Delta t}, \quad \tau_{ij}^* = \tau_{ij}^n + \Delta t (w_{ik}\tau_{kj}^n - \tau_{ik}^n w_{kj}). \quad (9)$$

The second invariant of the effective strain rate is computed by cross-interpolation and averaging of the missing data (squares of the corresponding components) between all the control volumes.

The effective viscosity (η^*) and the updated deviatoric stresses (τ_{ij}) are computed from the effective strain rates, using the standard quasi-viscous expression:

$$\tau_{ij} = 2\eta^* \dot{\epsilon}_{ij}^*, \quad \eta^* = \min \left[\left(\frac{1}{G\Delta t} + \frac{1}{\eta_l} + \frac{1}{\eta_n} \right)^{-1}, \frac{\tau_Y}{2\dot{\epsilon}_{II}^*} \right]. \quad (10)$$

Here, the individual creep viscosities are defined by

$$\eta_l = \frac{1}{2} (A_l)^{-1}, \quad \eta_n = \frac{1}{2} (A_n)^{-\frac{1}{n}} (\dot{\epsilon}_{II}^*)^{\frac{1}{n}-1}. \quad (11)$$

The discretised coupled system of nonlinear algebraic equations is solved at each time step using the preconditioned Jacobian-Free Newton-Krylov (JFNK) method with line-search as implemented in the PETSc SNES nonlinear solver framework²:

$$\mathbf{A}^{-1} \mathbf{J}(\mathbf{x}_k) \delta \mathbf{x}_k = -\mathbf{A}^{-1} \mathbf{r}(\mathbf{x}_k), \quad \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \delta \mathbf{x}_k, \quad (12)$$

where \mathbf{r} and \mathbf{x} are the coupled residual and solution vectors, respectively, $\delta \mathbf{x}$ is the iterative correction vector, k is the iteration index, and α the line-search step length. The Jacobian (\mathbf{J}) is defined implicitly by a matrix-vector product approximated by finite differencing. The preconditioning matrix (\mathbf{A}) is obtained by discretising the conservation Eqs. 1–3 using the current effective viscosity and by ignoring the coupling terms between the Stokes block and the energy equation. \mathbf{J} and \mathbf{A} are given by:

$$\mathbf{J} \mathbf{y} \approx \frac{\mathbf{r}(\mathbf{x} + h \mathbf{y}) - \mathbf{r}(\mathbf{x})}{h}, \quad \mathbf{A} = \begin{pmatrix} \mathbf{K} & \mathbf{G} & 0 \\ \mathbf{D} & \mathbf{C} & 0 \\ 0 & 0 & \mathbf{E} \end{pmatrix}, \quad (13)$$

where h is the perturbation parameter, \mathbf{y} is an arbitrary vector to be multiplied with the Jacobian, \mathbf{K} , \mathbf{C} , and \mathbf{E} denote the stiffness matrices of the velocity, pressure, and temperature blocks, respectively, \mathbf{G} is the pressure gradient matrix and \mathbf{D} is the velocity divergence matrix.

To achieve optimal scalability of the linear solver we employ a multigrid method to approximately invert the Stokes block in the preconditioning matrix. The coarse grid operators for the k -th level are obtained algebraically via Galerkin coarsening: process $\mathbf{A}_k = \mathbf{R}_{k+1}^k \mathbf{A}_{k+1} \mathbf{P}_k^{k+1}$. We have incorporated custom restriction (\mathbf{R}) and prolongation (\mathbf{P}) operators suitable for the staggered grid discretisation⁶ into the PETSc multigrid framework. The multigrid preconditioner is implemented in either a coupled form, using simultaneous coarsening of the velocity and pressure blocks, or in a block triangular form, in which coarsening is applied only to the velocity matrix. Accordingly, the resulting preconditioners are referred to as coupled (\mathbf{A}_c) or uncoupled (\mathbf{A}_u)

$$\mathbf{A}_c = \begin{pmatrix} \mathbf{K} & \mathbf{G} \\ \mathbf{D} & -\frac{1}{\eta^*} \mathbf{I} \end{pmatrix}, \quad \mathbf{A}_u = \begin{pmatrix} \mathbf{K} & \mathbf{G} \\ 0 & -\frac{1}{\eta^*} \mathbf{I} \end{pmatrix}. \quad (14)$$

In both cases, we approximate pressure Schur complement by the inverse viscosity matrix.

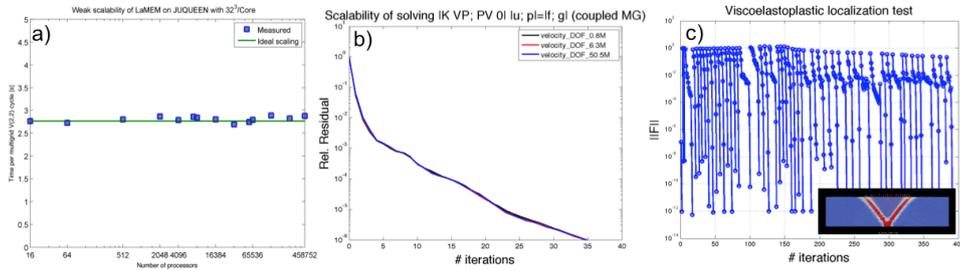


Figure 2. a) Weak scalability test of LaMEM for a single V-cycle with 32^3 gridpoints/processor. The largest simulation has > 15 billion grid points. b) Convergence of a solver step for a variable viscosity falling block test with viscosity contrast 10^3 , demonstrating that our coupled multigrid preconditioner results in resolution independent convergence (KSP: GMRES, with GMG and Chebyshev/Jacobi(3,3) smoothers & GMRES/GAMG as coarse grid solver). c) convergence of the nonlinear residual for various time steps for the viscoelastoplastic shear localisation setup shown in the inset.

2.3 Scalability Tests

We performed a number of tests to verify the scalability of LaMEM on JUQUEEN. A weak scalability test shows nearly perfect scalability for a single multigrid step up to the whole machine size (Fig. 2a). Our coupled multigrid preconditioner gives convergence behaviour that is resolution independent for a variable viscosity setup with a viscosity contrast of 1000 and 10 falling spheres, which was demonstrated to be a realistic test setup for geodynamic problems (Fig. 2b)⁷. A typical example of convergence during a visco-elasto-plastic shear localisation test during several time steps shows that the Newton solver results in rapid convergence once the initial residual has been reduced sufficiently (Fig. 2c).

3 Application Examples

3.1 Crustal-Scale Folding and Faulting

Under compression, crustal rocks can either fault or deform through folding. Folding results in quasi-regular structures, but it was incompletely understood how such structures grow, particularly in 3D. We therefore performed a systematic study and could demonstrate, using a combination of scaling laws and 3D simulations, that the wavelength of folds is mainly controlled by the effective viscosity structure of the crust that is deformed above a weaker salt layer. Lateral growth of folds results in linking of individual segments in a way that is quite similar to those observed in the Zagros (Iran)⁸. Yet, in some locations in the Zagros, salt crops out at the surface. Geological arguments suggest that these salt outcrops are salt diapirs sourced at a deep salt level, but that they were shallowly buried or exposed at the surface before the onset of collision. The formation of such salt structures is controlled by the speed with which sediments are deposited on top of the salt⁹. As such pre-existing salt structures form large-scale heterogeneities, they might affect the folding process. In order to understand that, we performed simulations in which we added the observed salt structure spacing in the Zagros to the initial model setup. Results show that this indeed localises deformation and result in folds with larger amplitudes, even though

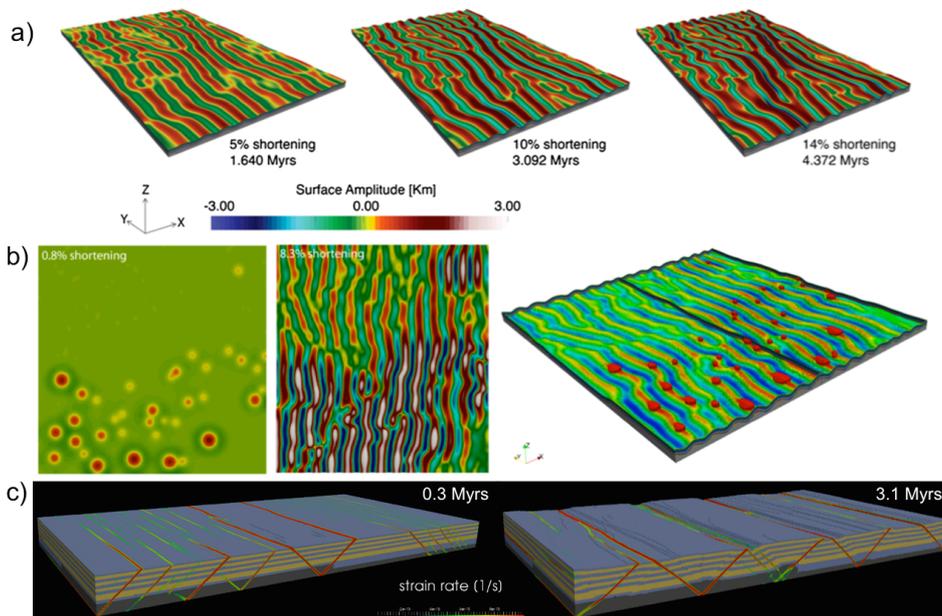


Figure 3. Examples of fold-and-thrust belt simulations: a) 3D fold pattern formation that develops once a system of multilayered crustal rocks above a weaker salt layer is compressed. Random heterogeneities are added to initiate the instability. Yet, fold spacing is mainly controlled by the material properties of the rocks. b) If pre-existing salt diapirs are present before the onset of collision, they affect the fold amplitudes but not their spacing. c) example of brittle fault formation and interaction during compression.

the fold spacing is unaffected (Fig. 3b)¹⁰. In additional work, we studied the effect of erosion on folding patterns and could demonstrate that erosion by itself does not affect fold spacing¹¹ dramatically, even though it might result in a faster growth of the structures and affect the manner in which folds link laterally¹².

Yet, folding is not the only response to compression. If crustal rocks are relatively cold, or have no or only few mechanically weak layers, fault zones will develop. In order to understand how this process works, we performed simulations of a brittle crust, which show that there is significant lateral interaction between fault zones, which ultimately develop larger scale structures (Fig. 3c).

3.2 Continental Collision and Plateau Formation

The collision of India with Asia resulted in the largest mountain belt on Earth, the Himalaya, but also in the Tibetan plateau, which has an average height (~ 5 km) that is more than the highest mountain in Europe. A similar plateau exists in the Andes (Altiplano), whereas the Alps does not have one. Why does a plateau form in some cases and not in others? In order to understand this, we performed systematic simulations using a setup in which an oceanic plate subducts underneath an overriding continental plate, followed by continent-continent collision. The results show that in order to form a plateau, we need to

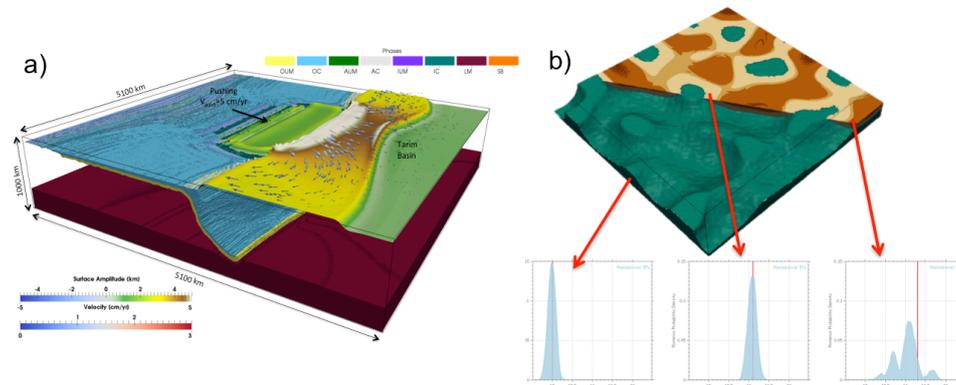


Figure 4. a) Simulation to study how subduction followed by continental collision results in the formation of a mountain belt and a high-elevation plateau (for example the Tibetan plateau)¹³. b) Example of geodynamic inverse modelling of salt-tectonics. During the inversion process, we assume that the geometry is known and we vary the density and viscosity of each of the layers. The modelled surface velocity and gravity anomalies are compared with the “observed” ones (in this case taken from a synthetic setup with known parameters). Results give a probability density function for each of the inversion parameters.

have sufficiently large convergence velocities, strong parts in the overriding lithosphere and a sufficiently large viscosity of the lithosphere (Fig. 4). We could demonstrate that there are 4 different types of plateaus that can form, which are controlled by two non-dimensional numbers, and which can be used to explain the observations¹³.

3.3 Constraining the Rheology of the Lithosphere through Geodynamic Inverse Modelling

The largest uncertainty in performing geodynamic models comes from our imprecise knowledge of the effective viscosity of rocks, which is typically measured in the laboratory by deforming small rocks samples. Rather than relying on these data, we developed a new approach that couples lithospheric-scale geodynamic models with geophysical observations such as the GPS velocities with which plates deform, the topography of the lithosphere and measured gravity anomalies. During the inversion process, we automatically adopt the input parameters until a low misfit is obtained. As it is unclear whether a single global minimum exists, we employ a Monte Carlo based method in combination with geometrical constraints, which is able to deal with multiple local minima¹⁴. We could demonstrate with an analytical solution that this geodynamic inversion gives unique results for a rising sphere example, which is a significant improvement over earlier gravity-only inversions that has non-unique results. In a next step, we performed synthetic tests using linear viscous 3D models and showed that the method is able to retrieve the material parameters of layers that contribute to the large-scale dynamics of the model¹⁴. More recently, we could demonstrate that it also works for fully non-linear and temperature dependent rheologies that include plastic yielding¹⁵, which implies that it can be employed to constrain the rheology of the lithosphere. A first application to the India-Asia collision zone shows that the viscosity of the Indian lithosphere must be rather large, but also that the

viscosity beneath Asia can be less well constrained given the current data¹⁵. Overall, our results suggest that the geodynamics inversion is a very promising new research direction that will give new insights in our understanding of the physics of the lithosphere.

Acknowledgements

This work was sponsored by the European Research Council under Starting Grant 258830. The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN¹⁶ at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

References

1. F. Harlow, J. Welsh, *Numerical calculation of time-dependant viscous incompressible flow of fluid with free surface*, Phys. Fluids **8**, 2182–2189, 1965.
2. S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, K. Rupp, B. F. Smith, S. Zampini, H. Zhang, *PETSc users manual*, Technical Report ANL-95/11–Revision 3.6, Argonne National Laboratory, 2015.
3. B. J. P. Kaus, H. Mühlhaus, D. A. May. *A stabilization algorithm for geodynamic numerical simulations with a free surface*. Phys Earth Planet In **181**, 12–20, 2010.
4. T. Duretz, D. A. May, T. V. Gerya, P. J. Tackley, *Discretization errors and free surface stabilization in the finite difference and marker-in-cell method for applied geodynamics: A numerical study*, Geochem. Geophys. Geosyst. **12**, Q07004, doi:10.1029/2011GC003567, 2011.
5. H. Wang, R. Agrusta, J. van Hunen, *Advantages of a conservative velocity interpolation (CVI) scheme for particle-in-cell methods with application in geodynamic modeling*, Geochem. Geophys. Geosyst. **16**, 2015–2023, 2015.
6. M. Cai, A. Nonaka, J. B. Bell, B. E. Griffith, A. Donev, *Efficient Variable-Coefficient Finite-Volume Stokes Solvers*, Commun. Comput. Phys. **16**, 1263–1297, 2014.
7. D. A. May, J. Brown, L. Le Pourhiet, *pTatin3D: High-Performance Methods for Long-Term Lithospheric Dynamics*, SC14 , 274–284, 2014.
8. N. Fernandez, B. J. P. Kaus, *Fold interaction and wavelength selection in 3D models of multilayer detachment folding*, Tectonophysics **632**, 199–217, 2014.
9. N. Fernandez, B. J. P. Kaus, *Pattern formation in 3-D numerical models of down-built diapirs initiated by a Rayleigh-Taylor instability*, Geophys. J. Int. **202**, 1253–1270, 2015.
10. N. Fernandez, B. J. P. Kaus, *Influence of pre-existing salt diapirs on 3D folding patterns*, Tectonophysics **637**, 354–369, 2014.

11. M. Collignon, B. J. P. Kaus, D. A. May, N. Fernandez, *Influences of surface processes on fold growth during 3-D detachment folding*, *Geochem. Geophys. Geosyst.* **15**, 3281–3303, 2014.
12. M. Collignon, N. Fernandez, B. J. P. Kaus, *Influence of surface processes and initial topography on lateral fold growth and fold linkage mode*, *Tectonics* **34**, 1622–1645, 2015.
13. A. E. Püskök, B. J. P. Kaus, *Development of topography in 3-D continental-collision models*, *Geochem. Geophys. Geosyst.* **16**, 1378–1400, 2015.
14. T. S. Baumann, B. J. P. Kaus, A. A. Popov, *Constraining effective rheology through parallel joint geodynamic inversion*, *Tectonophysics* **631**, 197–211, 2014.
15. T. S. Baumann, B. J. P. Kaus, *Geodynamic inversion to constrain the non-linear rheology of the lithosphere*, *Geophys. J. Int.* **202**, 1289–1316, 2015.
16. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, *Journal of large-scale research facilities*, **1**, A1, 2015, <http://dx.doi.org/10.17815/jlsrf-1-18>.

Computational Aspects of High-Resolution Global Gravity Field Determination – Numbering Schemes and Reordering

Jan Martin Brockmann and Wolf-Dieter Schuh

Institute of Geodesy and Geoinformation, Department of Theoretical Geodesy,
University of Bonn, Nussallee 17, D-53115 Bonn, Germany
E-mail: {brockmann, schuh}@geod.uni-bonn.de

Estimating high-degree spherical harmonic gravity field models from complementary observation types is computationally demanding. The computational effort depends on the one hand on the maximal resolution of the spherical harmonic expansion (i.e. the number of parameters to be estimated, tens to hundreds of thousands) and on the other hand on the number of observations (up to hundreds of millions). If approximations (e.g. block-diagonal approximations) should be avoided, concepts of high-performance computing have to be used to compute rigorous least-squares solutions from the typically dense systems of equations.

Within this contribution we focus on a technical detail in the context of estimating combined global gravity field models. Typically, dense systems of normal equations of different resolutions and provided in different numbering schemes have to be combined. Due to their dimension, they are mapped to the main memory as block-cyclically distributed matrices. We introduce symbolic numbering schemes, which are used to describe the resolution and the parameter order. This contribution summarises, how these symbolic numbering schemes are used to determine the permutation between the parameter order of two normal equations of different resolution and how the reordering is performed.

1 Introduction and Motivation

Gravity field models of the Earth are highly relevant in many geo-scientific applications¹. With a growing database (through dedicated satellite gravity missions like GRACE² or GOCE³, altimetry or surface data) combined models with higher spatial resolution become possible. The state-of-the-art mathematical parameterisation to describe the global Earth's gravitational potential is a spherical harmonic expansion up to a certain maximal degree l_{\max} . The potential may be written, for some evaluation point (r, θ, λ) in an Earth fixed and centred coordinate system, as⁴

$$V(r, \theta, \lambda) = \frac{GM}{a} \sum_{l=0}^{l_{\max}} \left(\frac{a}{r}\right)^{l+1} \sum_{m=0}^l (c_{lm} \cos(m\lambda) + s_{lm} \sin(m\lambda)) P_{lm}(\cos\theta), \quad (1)$$

where l and m denote the spherical harmonic degree and order (d/o), c_{lm} and s_{lm} the unknown coefficients of the spherical harmonic expansion, a the equatorial radius of the Earth reference ellipsoid, $P_{lm}(\cdot)$ the fully normalised associated Legendre functions, and GM the geocentric gravitational constant. This model comprises $U = (l_{\max} + 1)^2$ unknown coefficients, which can be used to express related geometrical and physical functionals of the Earth's gravity field.

Within global gravity field determination, the unknown coefficients c_{lm} and s_{lm} are estimated in a least-squares adjustment from various data sources, i.e. observations of the

effect of the Earth's gravity field (gravity changes, satellite orbit disturbances, accelerations, ...) ^{1,4,5}. These observations are either globally collected by satellites or locally in terrestrial measurement campaigns. Depending on the used data sets and their characteristics tens to several hundreds of thousands of parameters have to be estimated from the data. The unknown parameters from multiple data sources indicated by subscripts o and n are typically estimated in a joint weighted least squares adjustment ^{5,6}. The parameters \mathbf{x} are determined via the assembly and solution of the combined normal equations ⁷⁻⁹ (NEQs)

$$\left(\sum_o \omega_o \mathbf{A}_o^T \Sigma_{oo} \mathbf{A}_o + \sum_n \omega_n \mathbf{N}_n \right) \mathbf{x} = \left(\sum_o \omega_o \mathbf{A}_o^T \Sigma_{oo} \boldsymbol{\ell}_o + \sum_n \omega_n \mathbf{n}_n \right). \quad (2)$$

The observation groups o are available as raw measurements with design matrices \mathbf{A}_o , stochastic data vectors $\boldsymbol{\ell}_o$, and covariance matrices Σ_{oo} describing the uncertainty characteristics of the observation vector. The observation groups n are assumed to be already available as NEQs, where \mathbf{N}_n are the preprocessed NEQ matrices and \mathbf{n}_n the right hand side vectors. $\omega_{n,o}$ are (unknown) weight factors for the observation groups. The vector \mathbf{x} contains the unknown parameters, i.e. mainly spherical harmonic coefficients c_{lm} and s_{lm} but also additional group specific parameters like biases, or finite element parameters if the dynamic ocean topography is co-estimated from the altimeter data ^{10,11}.

Although it is generally useful to separate the group n and o as they have to be treated separately within the entire procedure ⁹, they can be merged within this contribution. Defining $\mathbf{N}_o := \mathbf{A}_o^T \Sigma_{oo} \mathbf{A}_o$ and $\mathbf{n}_o := \mathbf{A}_o^T \Sigma_{oo} \boldsymbol{\ell}_o$, which assumes the direct computations of the NEQs from groups o , Eq. 2 can be re-written with a single index $i \in \{o, n\}$ as

$$\left(\sum_i \omega_i \mathbf{N}_i \right) \mathbf{x} = \sum_i \omega_i \mathbf{n}_i, \Leftrightarrow \mathbf{N} \mathbf{x} = \mathbf{n}, \text{ with } \mathbf{N} := \sum_i \omega_i \mathbf{N}_i, \mathbf{n} := \sum_i \omega_i \mathbf{n}_i. \quad (3)$$

Due to the dimension of tens to hundreds of thousands of rows and columns ⁹ of the symmetric and positive definite system of NEQs (Eq. 3), their setup and solution is implemented via an integrated use of block-cyclically distributed matrices ^{12,9,8} and SCALAPACK. Although the task of combining the group specific NEQs \mathbf{N}_i and \mathbf{n}_i seems simple, Eq. 3 is only valid if the observation equations and thus the NEQs of all groups i are all assembled for the same and entire target parameter space (in the same parameter order).

Within gravity field determination, this is typically not the case as: (i) the NEQs from individual data sets are typically set up only up to a certain spherical harmonic degree, i.e. a degree determined by the sensitivity of of the measurement concept. For instance this can be $l_{\max} = 5$ for Satellite Laser Ranging ¹³ but $l_{\max} = 280$ for GOCE ¹⁴. (ii) The NEQs may contain group specific parameters, e.g. biases, calibration parameters or additional target parameters which can be determined from specific observation groups only ¹⁰. Finally, (iii) the ordering of the parameters is varying, as it is more or less arbitrary, although different (standard) so called numbering schemes with different properties exist ¹⁵. If an arbitrary number of NEQs should be combined, online reordering is required to guarantee consistent combined NEQs.

This contribution focuses on the combination of different NEQs assuming the parameter space and ordering of the different groups i differs. A reordering strategy is developed to derive a general and flexible framework. Based on a symbolic description of the parameter ordering (numbering scheme) efficient reordering schemes can be applied. As within

global gravity field determination the NEQs are high-dimensional, the concept is applied to block-cyclically distributed matrices. The developed framework is usable for many applications and in different situations, e.g. prerequisites for different solvers^{9,16}.

2 Reordering of Block-Cyclically Distributed Matrices

2.1 Symbolic Numbering Schemes

Instead of relying on rule-defined numbering schemes^{17,15} it is more flexible and more general to define symbolic numbering schemes. A symbolic numbering scheme can be defined as a linear sequence of symbolically described parameters, for instance objects of a type `Parameter`. The linear sequence can be realised as a `std::vector<Parameter>`. The class `Parameter` and their attributes are used to uniquely describe the parameters, the parameter type (e.g. spherical harmonic coefficient, fine element coefficient, ...) plus additional information like d/o for a spherical harmonic coefficient. Every parameter group can be mapped to this scheme. For the later described determination of the reordering operation between two symbolic numbering schemes, it is essential that parameters are comparable such that they are sortable. As every parameter in a numbering scheme is unique, it is simple to define an operator `<(const Parameter & p)`. A numbering scheme, a sequence of parameters is called \mathbf{p} in the following, whereas p is an individual parameter and $\mathbf{p}[i]$ the i th parameter in the numbering scheme. Each numbering scheme, either generated based on rules or arbitrarily chosen, can be represented by such a symbolic numbering scheme. Thus, it is easy to assume that for every NEQ used in the combination, a symbolic numbering scheme can be created online (rule-based) or is available (from a file).

2.2 Reordering of Matrices, Index and Permutation Vectors

Based on two numbering schemes \mathbf{p}_f and \mathbf{p}_t , the goal is to find the index vector and the permutation operator which reorders a vector $\mathbf{x}_{\mathbf{p}_f}$ given in \mathbf{p}_f to the vector $\mathbf{x}_{\mathbf{p}_t}$ which contains the parameters sorted as described by \mathbf{p}_t . The requirements for \mathbf{p}_f and \mathbf{p}_t are that either all coefficients of \mathbf{p}_f are contained in \mathbf{p}_t or vice versa, that all coefficients of \mathbf{p}_t are contained in \mathbf{p}_f . The first case means that \mathbf{p}_f contains fewer parameters than \mathbf{p}_t , i.e. \mathbf{p}_f is a subset of \mathbf{p}_t . The vector to be reordered shall be extended by zeros, corresponding to entries of “missing” coefficients. In the other case, \mathbf{p}_t contains fewer parameters than \mathbf{p}_f , consequently \mathbf{p}_t is a subset of \mathbf{p}_f . In that situation, the vector should be reordered in such a way that the first coefficients correspond to the ordering of \mathbf{p}_t and that the additional coefficients are ordered to the end of the vector, which can then be truncated. The algorithms provided are operational for both cases. Nevertheless, for the combination of NEQs focused on here, the first case is the crucial one. The change of the parameter order (reordering) is nothing else than an interchange of rows for vectors (like the right hand side \mathbf{n}_i) and the interchange of rows and columns for matrices (like the normal matrices \mathbf{N}_i). Mathematically the operation can be described by an index vector $\mathbf{i}_{\mathbf{p}_f \rightarrow \mathbf{p}_t}$, or a permutation operation $\Psi_{\mathbf{p}_f \rightarrow \mathbf{p}_t}$. Whereas the index vector assumes a simultaneous interchange, the permutation assumes a sequential interchange (taking already performed interchanges into account).

Algorithm 1: Computation of index vector from two symbolic numbering schemes.

```
Data:      vector<Parameter>  $\mathbf{p}_{\text{from}}$  Symbolic numbering scheme source matrix is ordered in
           vector<Parameter>  $\mathbf{p}_{\text{into}}$  Symbolic numbering scheme matrix should be reordered to
1  vector<size_t>  $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}$  ( $\mathbf{p}_{\text{into}} \cdot \text{size}()$ , 0)           // initialisation of index vector
2  // start value for fill in indices for parameters in  $\mathbf{p}_{\text{into}}$  but not in  $\mathbf{p}_{\text{from}}$ , inserted at the end
3  size_t  $e = \mathbf{p}_{\text{from}} \cdot \text{size}()$ 
4  // store current index of parameter in auxiliary variable  $i$  of each individual parameter
5  for  $k = 0$  to  $\mathbf{p}_{\text{from}} \cdot \text{size}()$  do
6  |    $\mathbf{p}_{\text{from}}[k] \cdot i() = k$ 
7  end
8  sort( $\mathbf{p}_{\text{from}} \cdot \text{begin}()$ ,  $\mathbf{p}_{\text{from}} \cdot \text{end}()$ )           // sort numbering scheme (operator<)
9  // loop over all parameters in  $\mathbf{p}_{\text{into}}$ 
10 for  $k = 0$  to  $\mathbf{p}_{\text{into}} \cdot \text{size}()$  do
11 |   // find index of parameter  $\mathbf{p}_{\text{into}}(i)$  in  $\mathbf{p}_{\text{from}}$  in sorted numbering scheme
12 |    $i = \text{lower\_bound}(\mathbf{p}_{\text{from}} \cdot \text{begin}(), \mathbf{p}_{\text{from}} \cdot \text{end}(), \mathbf{p}_{\text{into}}(i))$ 
13 |   // if parameter found, insert index  $i$ , otherwise fill in value outside of  $\mathbf{p}_{\text{from}} \cdot \text{size}()$ 
14 |   if  $i < \mathbf{p}_{\text{from}} \cdot \text{size}()$  then
15 |   |    $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}(k) = \mathbf{p}_{\text{from}} \cdot \mathcal{P}(i) \cdot i()$ 
16 |   else
17 |   |    $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}(k) = e$ 
18 |   |    $e++$ 
19 |   end
20 end
21
22 // special case if  $\mathbf{p}_{\text{into}} \subseteq \mathbf{p}_{\text{from}}$ : extend index vector to size of  $\mathbf{p}_{\text{from}}$ 
23 if  $\mathbf{p}_{\text{into}} \cdot \text{size}() < \mathbf{p}_{\text{from}} \cdot \text{size}()$  then
24 |    $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}} \cdot \text{resize}(\mathbf{p}_{\text{from}} \cdot \text{size}())$ 
25 |   // the remaining parameters are sorted to the end as they are not contained in  $\mathbf{p}_{\text{into}}$ ,
26 |   for  $k = \mathbf{p}_{\text{into}} \cdot \text{size}()$  to  $\mathbf{p}_{\text{from}} \cdot \text{size}()$  do
27 |   |    $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}(k) = k$ 
28 |   end
29 end
30 return  $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}$  // index vector performing reordering from  $\mathbf{p}_{\text{from}}$  to  $\mathbf{p}_{\text{into}}$ 
```

Index Vector for Reordering

Within the index vector $\mathbf{i}_{\mathbf{p}_f \mapsto \mathbf{p}_t}$, the entry at position i contains the index, the coefficient $\mathbf{p}_t(i)$ can be found in \mathbf{p}_f , and thus $\mathbf{p}_t(i) = \mathbf{p}_f(\mathbf{i}_{\mathbf{p}_f \mapsto \mathbf{p}_t}(i))$. Given a matrix \mathbf{A} in \mathbf{p}_f , its rows are reordered to \mathbf{p}_t via $\mathbf{A}_{\mathbf{p}_t} = \mathbf{A}_{\mathbf{p}_f}(\mathbf{i}_{\mathbf{p}_f \mapsto \mathbf{p}_t}, :)$, its columns via $\mathbf{A}_{\mathbf{p}_t} = \mathbf{A}_{\mathbf{p}_f}(:, \mathbf{i}_{\mathbf{p}_f \mapsto \mathbf{p}_t})$, and for quadratic matrices rows and columns via $\mathbf{A}_{\mathbf{p}_t} = \mathbf{A}_{\mathbf{p}_f}(\mathbf{i}_{\mathbf{p}_f \mapsto \mathbf{p}_t}, \mathbf{i}_{\mathbf{p}_f \mapsto \mathbf{p}_t})$, using the well known MatLab/Octave like notation. The index vector $\mathbf{i}_{\mathbf{p}_f \mapsto \mathbf{p}_t}$ can be computed with the efficient Alg. 1, which works for both cases mentioned above.

Within each coefficient of \mathbf{p}_f , the original position in the numbering scheme is stored (cf. Alg. 1, l. 5–7) in a auxiliary attribute i . Afterwards \mathbf{p}_f can be sorted. Now, iterating over the parameters in \mathbf{p}_t , they can be efficiently found, if contained, in \mathbf{p}_f as they have to be searched in a sorted vector (cf. Alg. 1, l. 10–20). The original index is stored from the auxiliary variable into the index vector. If the parameter is not contained in \mathbf{p}_f , the index is set to an entry larger than $\mathbf{p}_f \cdot \text{size}()$, which will correspond to the extended zeros when the reordering is performed to a vector/matrix. Parameters contained in \mathbf{p}_f but not in \mathbf{p}_t are arranged to the end, via setting the entries of the index vector to a value larger then the

Algorithm 2: Conversion of an index vector to a permutation vector.

```
Data: vector<size_t>  $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}$  index vector to be converted to permutation vector
1 vector<size_t>  $\boldsymbol{\psi}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}} = \mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}$  // initialisation of permutation vector
2 size_t  $p = 0$ 
3 // auxiliary vector, entry  $\mathbf{h}(k)$  contains index where value  $k$  is stored in  $\boldsymbol{\psi}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}$ 
4 vector<size_t>  $\mathbf{h}(\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}.size(), 0)$ 
5 for  $k = 0$  to  $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}.size()$  do
6 |    $\mathbf{h}(\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}(k)) = k$ 
7 end
8 // loop over entries of index vector
9 for  $k = 0$  to  $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}.size()$  do
10 |    $p = \mathbf{h}(k)$  // index of number  $k$  follows from  $\mathbf{h}$  instead of find operation
11 |   // check if entry  $k$  is in subsequent part of vector
12 |   if  $p > k$  then
13 |   |    $\boldsymbol{\psi}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}(p) = \boldsymbol{\psi}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}(k)$ 
14 |   |    $\mathbf{h}(\boldsymbol{\psi}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}(k)) = p$  // update vector  $\mathbf{h}$ , value  $\boldsymbol{\psi}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}(k)$  is now at position  $p$ 
15 |   end
16 end
17
18 return  $\boldsymbol{\psi}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}$  // sequential permutation vector corresponding to  $\mathbf{i}_{\mathbf{p}_{\text{from}} \mapsto \mathbf{p}_{\text{into}}}$ 
```

dimension of \mathbf{p}_t (cf. Alg. 1, l. 23–29). The complexity of the algorithm is $\mathcal{O}(n \log(n))$. For a test case of two numbering schemes of 520 000 parameters, the serial runtime is 0.2 s.

Permutation Vector for Reordering

An alternative notation/operation, which is better suited for the block-cyclically distributed matrices, is a so called permutation or pivoting vector. In contrast to an index vector, where the column and/or row interchanges are assumed to be performed simultaneously, a permutation vector contains a sequence of serial row and column permutations, i.e. a sequential swapping of two rows/columns starting at the begin of the vector. In contrast to the index vector, already performed swapping operations are taken into account in the representation. An entry in the permutation vector at position i means that the row i is swapped with row $\boldsymbol{\psi}(i)$. To be more precise, the current content of row/column i is swapped with the current content of row/column $\boldsymbol{\psi}(i)$. Note that the content might change with every swapping operation. Now, the old entry of position i is in row $\boldsymbol{\psi}(i)$, thus the index vector needs to be updated. A remaining entry i in the subsequent elements of $\mathbf{i}_{\mathbf{p}_r \mapsto \mathbf{p}_t}$ has to be replaced by the entry $\boldsymbol{\psi}(i)$. The procedure to convert an index vector to a permutation vector is summarised in Alg. 2. The basic idea is to avoid the search operation via introducing a second vector which stores the position of an entry k in the vector. The complexity is $\mathcal{O}(2n)$, its serial runtime is 0.01 s, for an example index vector with 520 000 entries.

To apply a permutation vector to rows and/or columns, the operator $\Psi_{\mathbf{p}_r \mapsto \mathbf{p}_t}(\cdot)$ is defined. This operator performs the serial permutations of rows ($\Psi_{\mathbf{p}_r \mapsto \mathbf{p}_t}^r$) as given by the vector $\boldsymbol{\psi}_{\mathbf{p}_r \mapsto \mathbf{p}_t}$, the operator $\Psi_{\mathbf{p}_r \mapsto \mathbf{p}_t}^c$ performs the column interchanges and $\Psi_{\mathbf{p}_r \mapsto \mathbf{p}_t}^{r,c}$ performs the interchanges for rows and columns.

3 Combined System of NEQs

Assuming that a target numbering scheme \mathbf{p} (associated with \mathbf{N}) exists which covers the entire parameter space to be estimated, Eq. 3 can be rewritten as

$$\left(\sum_i w_i \begin{bmatrix} \mathbf{N}_i & \mathbf{0}_{U_i \times U - U_i} \\ \mathbf{0}_{U - U_i \times U_i} & \mathbf{0}_{U - U_i \times U - U_i} \end{bmatrix} (\mathbf{i}_{\mathbf{p}_i \mapsto \mathbf{p}}, \mathbf{i}_{\mathbf{p}_i \mapsto \mathbf{p}}) \right) \mathbf{x} = \sum_i w_i \begin{bmatrix} \mathbf{n}_i \\ \mathbf{0}_{U - U_i} \end{bmatrix} (\mathbf{i}_{\mathbf{p}_i \mapsto \mathbf{p}}), \quad (4)$$

for the use with an index vector. Using the introduced permutation operator, we obtain

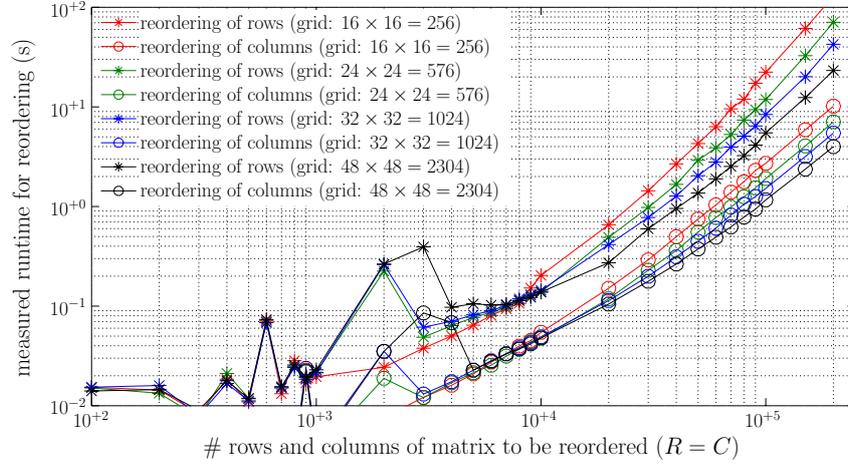
$$\left(\sum_i w_i \Psi_{\mathbf{p}_i \mapsto \mathbf{p}}^{r,c} \left(\begin{bmatrix} \mathbf{N}_i & \mathbf{0}_{U_i \times U - U_i} \\ \mathbf{0}_{U - U_i \times U_i} & \mathbf{0}_{U - U_i \times U - U_i} \end{bmatrix} \right) \right) \mathbf{x} = \sum_i w_i \Psi_{\mathbf{p}_i \mapsto \mathbf{p}}^r \left(\begin{bmatrix} \mathbf{n}_i \\ \mathbf{0}_{U - U_i} \end{bmatrix} \right), \quad (5)$$

assuming \mathbf{N}_i and \mathbf{n}_i to be the original NEQs as they were originally set up. Their numbering scheme is denoted as \mathbf{p}_i . These NEQs for the subset of the parameters are extended with zeros, if required. Afterwards the index vector or the permutation vector is applied to the temporarily extended NEQs. The NEQs can be combined performing a simple addition as the parameter order and parameter space is adjusted to the defined target numbering scheme \mathbf{p} . The solution of the NEQs equations can be performed with SCALAPACK^{8,9}.

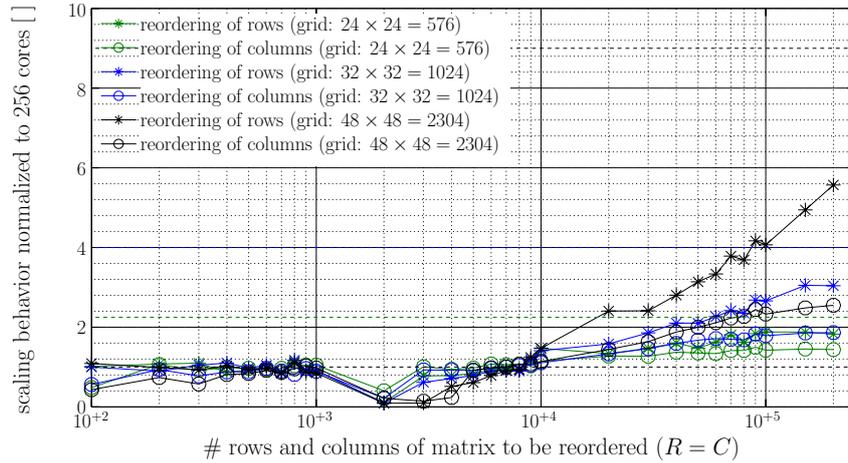
4 Reordering for Block-Cyclically Distributed Matrices

With the known permutation vector between two numbering schemes, the permutation is applied to block-cyclically distributed matrices to compute the combined system of NEQs cf. Eq. 5. With the permutation vector, the operation is simply performed by the SCALAPACK helper routine `pdlapiv`, which is used within SCALAPACK for pivoting during the solution of systems of equations. The subroutine performs the permutation given by $\psi_{\mathbf{p}_i \mapsto \mathbf{p}_i}$ to either rows or columns of a block-cyclically distributed matrix. Applying the function twice, first to permute rows and secondly to permute columns, both are reordered. Beside the standard input of the block-cyclic distribution of the matrix, the function requires the input of the permutation vector as block-cyclically distributed integer vector. Note that before reordering, a symmetric matrix has to be stored in the lower and upper triangle, as during reordering, both are mixed.

Fig. 1 gives an overview of the required runtime for the reordering of rows and columns of distributed matrices of different dimension on different quadratic processor grids (for the distribution parameters default values of $b_r = b_c = 64$ were used) to get an idea of the order of magnitude. The index vector was randomly generated (random shuffle of an index vector). The main conclusions of the test are: i) The reordering of columns is much faster than the reordering of rows (by a factor of three to ten). This could be expected as the column access in memory is much faster using the column major order for matrices for the locally stored matrices. ii) For matrices of dimension lower than $20\,000 \times 20\,000$ the reordering is performed in less than 1 s on all grids. For the reordering of columns, this even holds for matrices smaller than $80\,000 \times 80\,000$. Although there is no real scaling behaviour of the reordering operations with the number of cores (cf. Fig. 1(b)), the most important thing is that the performance increases on larger compute core grids and does not drop to additional organisational requirements (at least for matrices above dimension $10\,000 \times 10\,000$ the scaling is above 1.0 for all cases analysed).



(a) Mean absolute runtime measured for the reordering.



(b) Scaling behaviour of reordering operations (normalised to 256 cores).

Figure 1. Runtime analysis of the row (*) and column (o) reordering operations (index vector randomly generated). The colours represent different dimensions of the processor grid.

5 Summary and Conclusions

The concept of symbolic numbering schemes and the strategy to reorder block-cyclically distributed matrices is used within the framework for global gravity field recovery⁹. The reordering is computed efficiently on the fly and the original NEQs can be kept, and no additional preprocessing and homogenisation of the NEQs into a predefined numbering scheme is necessary. These strategies are also required within global gravity determination using iterative solvers, where the NEQs are required in different numbering schemes for the efficient setup of the observation equations (recursion formulas) and for preconditioning (block diagonal dominance in specific numbering)^{9,16}.

Acknowledgements

Parts of the study and the publication were funded within the DFG project G/O2000+ (SCHU2305/3-1). Most computations within the (entire) study were performed on the JU-ROPA supercomputer at Forschungszentrum Jülich. The computing time was granted by the John von Neumann Institute for Computing (project HBN15). The Open Source HPC computing libraries ATLAS, BLAS, LAPACK, OpenMPI, BLACS, PBLAS and SCALAPACK are gratefully acknowledged.

References

1. R. Rummel, G. Balmino, J. Johannessen, P.N.A.M. Visser, and P. Woodworth, *Dedicated gravity field missions – principles and aims*, J Geodyn, **33**, no. 12, 2002.
2. B. D. Tapley, S. Bettadpur, M. Watkins, and C. Reigber, *The gravity recovery and climate experiment: Mission overview and early results*, Geophys Res Lett, **31**, 2004.
3. R. Floberghagen, M. Fehringer, D. Lamarre, D. Muzi, B. Frommknecht, C. Steiger, J. Pineiro, and A. da Costa, *Mission design, operation and exploitation of the gravity field and steady-state ocean circulation explorer mission*, J Geodesy, **85**, no. 11, 749–758, 2011.
4. B. Hofmann-Wellenhof and H. Moritz, *Physical geodesy*, Springer, Vienna, 2005.
5. N. K. Pavlis, S. A. Holmes, S. Kenyon, and J. K. Factor, *The development and evaluation of the Earth Gravitational Model 2008 (EGM2008)*, J Geophys Res: Solid Earth, **117**, no. B4, 2012.
6. T. Gruber, *High resolution gravity field modeling with full variance-covariance matrices*, J Geodesy, **75**, no. 9-10, 505–514, 2001.
7. K. R. Koch, *Parameter Estimation and Hypothesis Testing in Linear Models*, Springer Berlin Heidelberg, 2nd edition, 1999.
8. J. M. Brockmann, L. Roese-Koerner, and W.-D. Schuh, *Use of high performance computing for the rigorous estimation of very high degree spherical harmonic gravity field models*, in: Proceedings of the International Symposium on Gravity, Geoid and Height Systems (GGHS2012), U. Marti (Ed.), vol. 141 of *International Association of Geodesy Symposia*, pp. 27–33, Springer Berlin Heidelberg. 2015.
9. J. M. Brockmann, *On High Performance Computing in Geodesy – Applications in Global Gravity Field Determination*, PhD thesis, Institute of Geodesy and Geoinformation, University of Bonn, Bonn, Germany, 2014.
10. S. Becker, M. Losch, J. M. Brockmann, G. Freiwald, and W.-D. Schuh, *A Tailored Computation of the Mean Dynamic Topography for a Consistent Integration into Ocean Circulation Models*, Surv Geophys, **35**, no. 6, 1507–1525, 2013.
11. S. Becker, J. M. Brockmann, and W.-D. Schuh, *Mean dynamic topography estimates purely based on GOCE gravity field models and altimetry*, Geophys Res Lett, **41**, no. 6, 2063–2069, 2014.
12. L. S. Blackford, J. Choi, A. Cleary, E. D’Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley, *ScaLAPCK Users Guide*, SIAM, Society for Industrial and Applied Mathematics, Philadelphia, 2nd edition, 1997.

13. A. Maier, S. Krauss, W. Hausleitner, and O. Baur, *Contribution of satellite laser ranging to combined gravity field models*, *Adv Space Res*, **49**, no. 3, 556–565, 2012.
14. J. M. Brockmann, N. Zehentner, E. Höck, R. Pail, I. Loth, T. Mayer-Gürr, and W.-D. Schuh, *EGM_TIM_RL05: An independent Geoid with Centimeter Accuracy purely based on the GOCE Mission*, *Geophys Res Lett*, **41**, no. 22, 8089–8099, 2014.
15. C. Boxhammer and W.-D. Schuh, *GOCE gravity field modeling: computational aspects - free kite numbering scheme*, in: *Observation of the Earth System from Space*, J. Flury et al. (Eds.), pp. 209–224. Springer Berlin Heidelberg, 2006.
16. J. M. Brockmann, L. Riese-Koerner, and W.-D. Schuh, *A concept for the estimation of high-degree gravity field models in a high performance computing environment*, *Stud Geophys Geod*, **58**, 571–594, 2014.
17. W.-D. Schuh, *Tailored Numerical Solution Strategies for the Global Determination of the Earth's Gravity Field*, vol. 81 of *Mitteilungen der geodätischen Institute der Technischen Universität Graz*, TU Graz, Graz, Austria, 1996.

Towards a Better Understanding of Rotating Turbulent Convection in Geo- and Astrophysical Systems

Stephan Stellmach, Jan Verhoeven, Matthias Lischper, and Ulrich Hansen

Institut für Geophysik, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany

E-mail: stellma@uni-muenster.de

Turbulent rotating convection occurs in many geo- and astrophysical bodies, but it is still far from being well understood. Here, we review some recent findings in this field which were obtained using the JUQUEEN system. We present results revealing that tiny viscous boundary layers, so-called Ekman layers, are much more important in rapidly rotating convection than previously thought. We also discuss evidence for upscale kinetic energy transport generating large-scale, coherent structures in rotating convection. Finally, we briefly discuss the geo- and astrophysically relevant case of convective systems in which the fluid parcels in the deeper parts of the convective region get compressed significantly by the weight of the overlying fluid. We show that in the presence of rotation, such compressibility effects can drive alternating jets similar to those observed on Jupiter and other giant planets.

1 Introduction

Buoyancy driven fluid flows, so-called convective flows, are ubiquitous in nature. They stir the outer layer of the sun and play an integral part in transporting the energy generated in the solar interior to the surface, from where it is radiated into space. Convection also occurs in giant planets, and has been proposed to drive the strong zonal winds which organise Jupiter's colourful clouds into the banded structures that dominate the planets' visual appearance. Deep below our feet, the Earth's magnetic field is generated by convective flows in its liquid outer iron core. Convection is also important in daily life and in technical applications, for example when it cools down a cup of tea or when it generates comfortable living room temperatures by redistributing the thermal energy of a radiator.

Geo- and astrophysical convective systems are however special in a number of respects. In particular, *(i)* they typically have a huge spatial extent, *(ii)* they usually spin rapidly, and *(iii)* the fluid in deeper parts of their convective regions can be compressed significantly by the overlying weight. All these features affect the dynamics considerably. The huge system size makes viscous diffusion irrelevant on most flow scales of interest, causing intense turbulence. At the same time, the motions usually feel the effects of rotation, and in many important applications, Coriolis forces dominate the leading order force balance. Finally, fluid particles can get compressed or expanded as they travel between different depth levels, which changes their moment of inertia. As we will discuss below, this can lead to order one changes in the overall dynamics, especially in rapidly rotating systems such as Jupiter, where it is possibly this process driving the zonal winds.

The goal of the research presented here is to gain a better understanding of the basic dynamics of turbulent, rotating convection and of its role in natural systems. In this contribution, we review some recent findings in this field which have been obtained using the JUQUEEN system in Jülich. The presentation style is chosen to appeal to a wider audience, with more details being available in our recent publications¹⁻⁷.

2 Modelling Approach and Applied Methods

Numerical simulations in geo- and astrophysics are often tailored to mimic a specific natural system as closely as possible. All physical processes expected to be relevant are included, and complicated geometries and boundary conditions are adopted. While results from such models can be compared directly with observations, the inherent complexity sometimes obscures the view on the fundamental physics. In this work we therefore focus deliberately on conceptually simple models in plane, Cartesian geometries. A rotating plane layer, heated from below and cooled from above with gravity pointing downwards is considered. This choice of setup allows us to directly relate the results to laboratory experiments which are usually carried out in a similar configuration. Furthermore, the most advanced theoretical models are generally developed for this particular setup. Testing the available theory with numerical simulations allows us to assess our current level of understanding and, as we will show, can also guide future theoretical advances.

Three different levels of approximation are used in our simulations, which in increasing order of complexity are (i) the Boussinesq approximation, which accounts for density variations only in the buoyancy force while otherwise treating the flow field as incompressible, (ii) the anelastic approximation, which includes the effects of an adiabatic increase of density with depth, and (iii) the so-called fully compressible model, which contains no approximations beyond those involved in basic continuum-mechanics. The Boussinesq approximation is usually used in the context of laboratory experiments, and also in the bulk of the existing theoretical work. Unfortunately, it breaks down in large-scale natural convection if the material gets compressed significantly by the weight of the overlying fluid. The anelastic approximation takes this effect into account, while maintaining most computational benefits of the Boussinesq approach. Its general validity, especially in the context of rapidly rotating flows, has however been called into question recently⁸. The fully compressible model provides the most comprehensive description, and remains valid even under super-sonic conditions and for intense thermodynamic fluctuations. An in-depth discussion, including the relevant equations, is beyond the scope of this overview paper and can be found elsewhere^{3,6}.

Different codes^{9,6} are used for the simulations. In the Boussinesq and anelastic case, high order spatial discretisation schemes are employed, based either on spectral Fourier and Chebychev expansions, or on a hybrid spectral / finite-difference formulation. In the fully compressible case, a second order finite difference scheme is used. Time integration is performed by linear multistep methods (AB/BDF2 and AB/BDF3). All codes have been shown to scale well on JUQUEEN up to at least 10^5 cores.

3 Asymptotic Behaviour of Rapidly Rotating Convection

Direct numerical simulations (DNS) of geo- and astrophysical convection are usually unable to cover the full range of temporal and spatial scales occurring in nature. This difficulty is well known to occur in the presence of strong turbulence, where eddy sizes often range from the system scales all the way down to the so-called Kolmogorov micro-scales at which molecular viscosity becomes important. Similar problems are, however, also encountered in rotating fluids, where it is the Coriolis effect that creates dynamical processes on a broad range of temporal and spatial scales. An important control parameter in this

case is the so-called Ekman number $E = \nu/(2\Omega H^2)$, which is defined as the ratio of the rotation time scale $(2\Omega)^{-1}$ to the characteristic time scale of viscous diffusion H^2/ν across the convective layer. Here, ν denotes the kinematic viscosity, Ω is the rotation frequency and H the layer depth. In planetary cores, E is on the order of 10^{-15} , while even the most advanced simulations only reach down to $E = O(10^{-7})$. Viscosity is therefore massively overrepresented in these simulations. The problem is due to an extreme range of spatial and temporal scales that needs to be resolved at small E . While the system scale is H , convective instabilities occur on $O(E^{1/3}H)$ spatial scales and viscous boundary layers have an $O(E^{1/2}H)$ thickness. The time scale of vertical viscous diffusion is H^2/ν , which is $O(E^{-2/3})$ times longer than the horizontal diffusion time scale across convective instabilities, and $O(E^{-1})$ times longer than the period of certain waves appearing in rotating flows. For E as low as 10^{-15} , these scale disparities are too huge to be resolved. Unfortunately, laboratory experiments also have been unable to probe the low Ekman numbers regime beyond the values attainable in simulations, albeit for different reasons^{4,10}.

As both simulations and laboratory experiments cannot reach realistic parameter values, the question arises whether they nevertheless represent the physics of natural systems adequately. Any geophysical interpretation of their results clearly requires an extrapolation over many orders of magnitude in Ekman number. This can only be done if (i) no further bifurcations or regime transitions are encountered within the covered parameter range and (ii) if the relevant scaling relations are known exactly. In essence, the physics in the asymptotic limit $E \rightarrow 0$ needs to be understood rigorously, and both numerical simulations and laboratory experiments have to advance into regions of parameter space where clear asymptotic behaviour becomes apparent.

Over the past decade, there has been considerable progress in mathematical studies of the low Ekman number limit employing asymptotic expansion techniques. Detailed quantitative predictions for rapidly rotating plane layer convection are now available^{11,12}, and these are expected to be valid even in the fully turbulent regime as long as Coriolis forces remain dominant in the force balance. JUQUEEN has allowed us to test these predictions², and thus our current understanding of rapidly rotating convection. As the available theoretical studies cover the Boussinesq case with anti-parallel rotation and gravity, we also employed this configuration in our DNS. Fig. 1 shows the computed convective heat transport as a function of the forcing strength for a range of Ekman numbers. As described in the figure caption, clear convergence against the theoretical predictions is found if the boundary conditions are chosen to explicitly suppress lateral shear stresses on the bounding surfaces. However, for rigid, no-slip boundary conditions, as they arise in laboratory experiments or planetary cores, the available theory appears to break down.

The crucial difference between the two cases is that viscous boundary layers, so-called Ekman layers, form in the no-slip case. These are extremely thin structures - for the simulations at $E = 10^{-7}$ shown in Fig. 1, the Ekman layers cover only the outermost $\sim 0.1\%$ of the fluid layer (but still need to be properly resolved, of course). It is surprising that these tiny structures can have such a strong impact on the overall dynamics - in some cases, they increase the heat transport across the convective region by more than 800%! Even more surprising is that Fig. 1 suggests that the Ekman layers *gain* dynamical importance as E is reduced, despite the fact that both their thickness and the secondary flows they induce are known to *decrease* with decreasing E . Our simulation results are also puzzling in the light of mathematical studies conducted back in the 1960s, which irrevocably showed that at

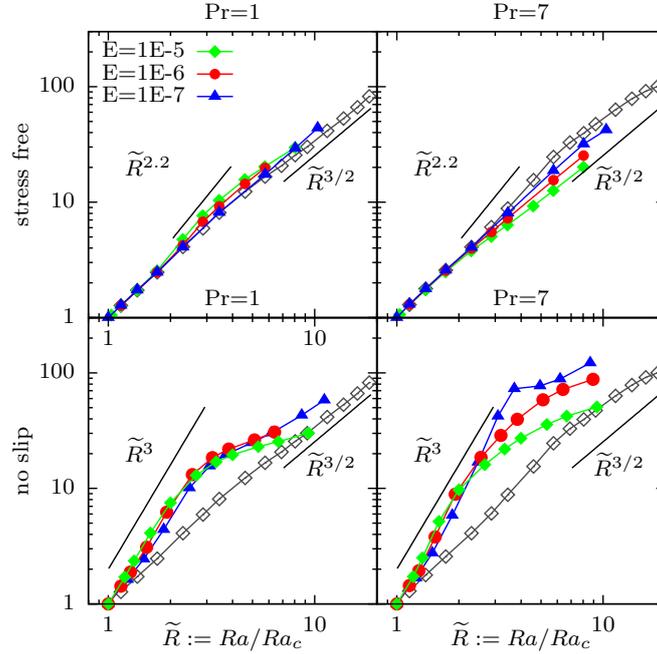


Figure 1. Heat transport, measured in terms of Nusselt number, as a function of super-criticality \tilde{R} . The non-dimensional Nusselt number Nu is defined as the total heat transport normalised by the purely conductive heat transport that would occur in the absence of convection. It is thus a non-dimensional measure of the heat transfer efficiency across the layer. Ra denotes the so-called Rayleigh number, a non-dimensional measure of the driving strength, and Ra_c the value at which convection first sets in. Results are shown for stress-free and no-slip mechanical boundary conditions, and for two different fluids (with Prandtl number $Pr = 1$ and $Pr = 7$), corresponding roughly to air and water. Coloured, full symbols show results obtained in direct numerical simulations carried out on JUQUEEN. Open symbols show theoretical predictions based on asymptotic theory^{11,12} expected to be valid at small E . With decreasing E , even in the highly turbulent regime with $Nu \gg 1$ the results converge to the theoretical prediction for stress-free boundary conditions, but not for no-slip boundary conditions. Interestingly, the maximum deviations from the asymptotical predictions increase for no-slip boundaries as the small Ekman number regime is approached.

the onset of convective motions, Ekman layer effects become asymptotically small in the rapidly rotating regime¹³. This suggests that somewhere above onset, a previously overlooked non-linear effect kicks in that brings the dynamical consequences of the tiny Ekman layers back into the leading order dynamics, with enormous effects on the heat flow.

Using an in-depth analysis of the numerical simulations, a refined version of the asymptotic theory which explicitly includes the Ekman layers has been developed⁷, and detailed computations based on this theory are currently in progress. Computations on JUQUEEN have thus revealed that a key physical process had been missed in previous theories. They also have provided essential information concerning the nature of the missing pieces. At least in the simple geometries considered here, a much more comprehensive understanding of rapidly rotating convection now appears within close reach.

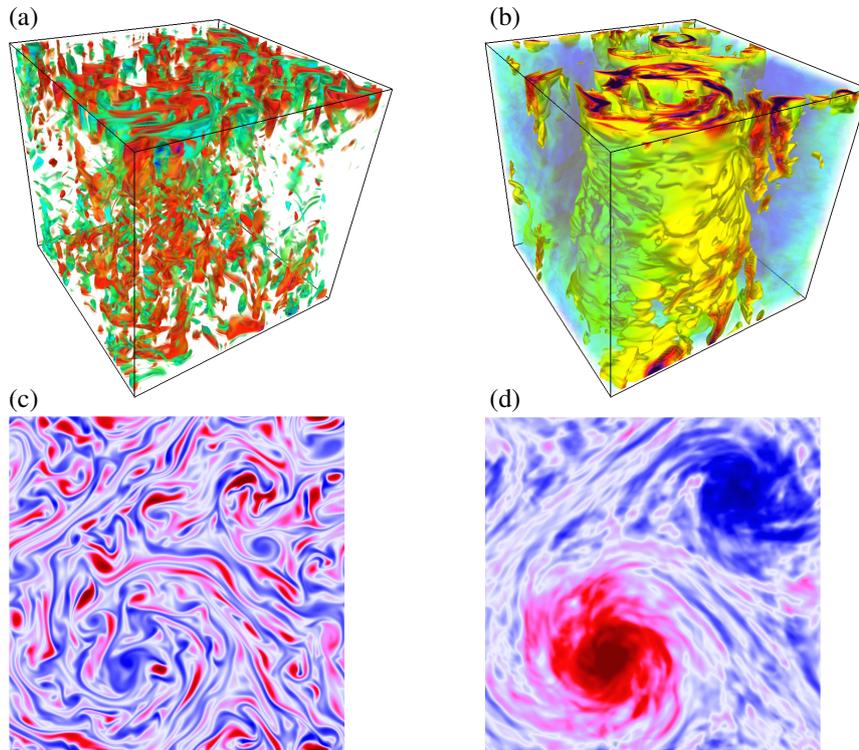


Figure 2. Formation of large-scale, coherent vortices in rotating Rayleigh-Bénard convection ($E = 10^{-7}$, $Pr = 1$ and $\tilde{R} \approx 10$) with stress-free boundaries. Shown is (a) vertical vorticity, (b) kinetic energy of horizontal flow components, (c) temperature anomaly close to the upper boundary and (d) vertically integrated vertical vorticity. As apparent from panels (a) and (c), the convective flow itself exhibits a small-scale, turbulent, three-dimensional structure, but drives system-scale, coherent vortices as shown in panels (b) and (d). Note that the horizontal scales are stretched by a factor of ~ 4.5 in (a) and (b) to aid the visualisation.

4 Upscale Kinetic Energy Transport in Rapidly Rotating Convection

As stated earlier, convective instabilities in rapidly rotating systems occur on small $O(E^{1/3}H)$ length scales, with H denoting the depth of the convective region. For the Earth's core with $E = O(10^{-15})$, this corresponds to a few tens of meters, which is tiny compared to the outer core radius of roughly 3400km. An important question is whether convective turbulence driven at such small scales is able to generate much larger coherent structures on the global scale H . Such structures could then possibly explain large-scale anomalies in the Earth's magnetic field⁵. Similarly, if the zonal winds on Jupiter and other giant planets are indeed convectively driven, a mechanism needs to exist that transports kinetic energy from the local convective injection scales to the global wind scales¹⁴.

Our simulations on JUQUEEN indeed reveal that system scale coherent structures can form in rapidly rotating Boussinesq convection provided the turbulence level is high enough². An example is shown in Fig. 2. Although the convective flow exhibits many small scale features on the $O(E^{1/3}H)$ convective instability scale, it drives an intense pair

of vortices on the largest scale available to the system. The formation of such vortices in rotating convection has recently also been observed in simulations at larger E , where the large-scale energy accumulates predominantly in cyclonic structures^{15,16}. Such symmetry-breaking is theoretically predicted to be absent in the asymptotic case of rapid rotation¹⁷. Indeed, our DNS reveals the generation of both cyclonic and anti-cyclonic vortices, which suggests that symmetry tends to be restored as E is reduced. A further interesting observation is that upscale kinetic energy transport appears to be largely suppressed by no-slip boundaries. An investigation of this effect is currently underway.

5 Modelling Compressibility Effects

So far, we have completely neglected the fluid's compressibility. This is typically justified in laboratory experiments, but not in large-scale geo- and astrophysical systems, where the weight of the overlying fluid often causes a substantial density increase with depth. In Jupiter for example, density varies by about four orders of magnitude from the one bar level down to the region where the atmospheric hydrogen becomes metallic, which is often viewed as a natural lower boundary of an outer convection zone.

There are several ways to model such configurations. One approach is to solve the basic fluid-dynamical equations following from first principles of continuum physics. These equations however also describe sound waves, and the requirement to resolve these can significantly hamper simulations of natural flows which often evolves on much longer time scales. To circumvent this problem, the governing equations can be simplified using the so-called anelastic approximation, which filters out the sound waves. Unfortunately, there are some concerns about the general validity of this approximation, and problems have recently been predicted to occur for low viscosity fluids in rapidly rotating systems⁸. It is therefore necessary (*i*) to check the validity bounds of the approximation, (*ii*) to quantify the arising errors and (*iii*) to quantify the relative computational efficiency for both approaches.

JUQUEEN has recently allowed us to compare both modelling strategies systematically, from the laminar to the turbulent regime, for the first time³. As a first step, an ideal gas has been taken as the working fluid, and only the non-rotating case has been considered so far. Our simulations show that results obtained using the full equations converge to the anelastic case exactly as predicted by theory, even in the turbulent regime. In addition, they provide quantitative error estimates as well as rough guidelines concerning the computational efficiency of each approach. This enables modellers to choose the most accurate and efficient approach for their particular problem at hand. We are currently extending this study to the rapidly rotating case.

6 Are Jupiter's Zonal Winds Driven by Compressibility?

Strong zonal winds organise the colourful clouds on Jupiter's surface into characteristic banded structures that can be observed even with basic amateur-level telescopes. Higher resolution pictures reveal that these winds are embedded in a sea of small scale turbulence. Similar pronounced zonal wind structures have been observed on all gas and ice giants in our solar system, and thus appear to be a very robust feature. Despite the tremendous observational progress that has been made over the last decades, the dynamical origin of the

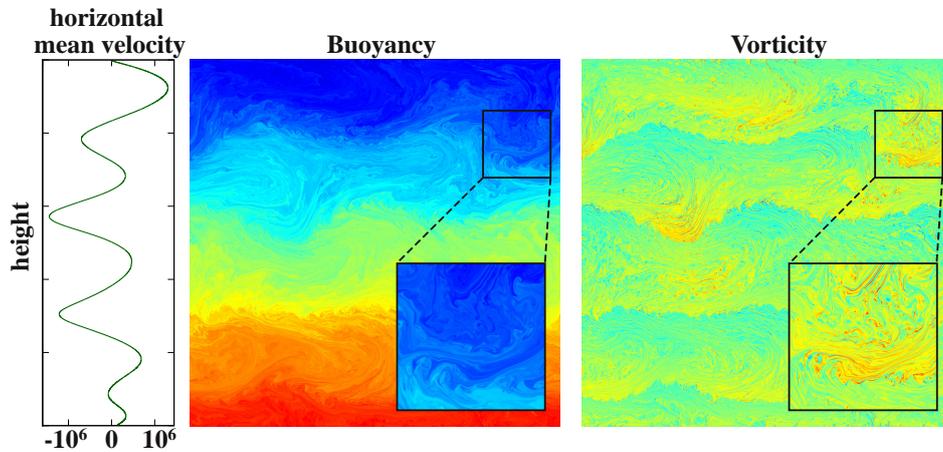


Figure 3. Generation of horizontal jets in a two-dimensional simulation of compressible, rotating convection. A compressible fluid is heated from below and cooled from above, with the rotation axis being perpendicular to the plane of the paper. The relative orientation of gravity and rotation mimics the conditions on the equatorial plane of giant planets. The colour maps range from blue (negatively buoyant, negative vorticity) to red (positive buoyancy and vorticity). Kinetic energy is transported from small-scale convective instabilities to larger spatial scales until the structures become large enough for compressional Rossby waves to be excited, visible here as large-scale, undulating structures. The energy then gets channelled predominantly into strong alternating jets.

winds remains poorly understood. Some models emphasise dynamical effects occurring in the outermost atmospheric layers, invoke approximations valid for shallow atmospheres only and ignore the vast interior of the planet¹⁸. Other models seek the dynamical roots of the differential rotation in convective motions extending deep downwards into the gaseous envelope^{19,14,20}. Today, it is not clear how deeply the jets extend into the planetary interior. The Juno mission that will reach Jupiter in mid-2016 has the potential to better constrain the radial extent of the jets²¹.

Here, we focus on the deep dynamics, and investigate the possibility that the jets are driven by compressibility effects. The idea²² is that in deep planetary convection, an inverse cascade successively transports kinetic energy generated by small-scale convective instabilities to larger scales, in a similar fashion as described in Sec. 4. At some point, the convective eddies become large enough to feel the density increase with depth. When this happens, certain waves, so-called *compressional Rossby waves*, are excited²³. Their basic dynamics is in many respects similar to the classical Rossby waves that create meanders in Earth's jet stream and influence our mid-latitude weather, but in this case it is the compression of the fluid with depth, and not the variation of Coriolis forces with latitude, that causes the wave motions. Theory suggests that the interaction of the turbulent eddies with these waves may channel the kinetic energy into horizontal shear flows, which would correspond to the observed zonal jets. Fig. 3 shows results from a numerical simulation using the anelastic approximation that illustrates the general possibility of this scenario¹.

By running a large number of simulations, we were able to investigate the dynamics in detail¹. The jet thickness could be shown to follow a certain scaling law similar to classical Rhines' theory for beta-plane turbulence. Applied to the giant planets in the solar system,

the observed number of jets is reasonable well predicted, which is remarkable given the simplicity of the model considered here.

7 Conclusions

Rotating convection is a fascinating fluid-dynamical phenomenon with large relevance in geo- and astrophysical systems, but it is still not well understood. Simulations on JUQUEEN have shown that the nature of the boundaries has a much larger impact than previously expected. In rapidly rotating systems, tiny Ekman boundary layers are shown to massively boost the heat transfer, contrary to common theoretical expectations. Also the processes of upscale kinetic energy transport and large-scale coherent structure formation appears to be strongly affected by the mechanical boundary conditions. Both effects need further study. We have also been able to demonstrate very clearly that intense alternating jets can be convectively driven by a complex interplay of compressibility effects and rotation. This should be an important effect in the interiors of gas planets, and must be considered as a possible explanation for the zonal winds observed on all giant planets in our solar system.

Acknowledgements

This work strongly benefited from an intense collaboration with the group of Prof. J. Aurnou (University of California, Los Angeles) and of Prof. K. Julien (University of Colorado, Boulder). Without their state-of-the-art laboratory experiments and advanced theoretical models, the research presented here would have been impossible.

The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN²⁴ at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

We also acknowledge funding by the German Research Foundation (Grant No. STE1976/1-2).

References

1. J. Verhoeven and S. Stellmach, *The compressional beta effect: A source of zonal winds in planets?*, *Icarus*, **237**, 143–158, 2014.
2. S. Stellmach, M. Lischper, K. Julien, G. Vasil, J. S. Cheng, A. Ribeiro, E. M. King, and J. M. Aurnou, *Approaching the asymptotic regime of rapidly rotating convection: Boundary layers versus interior dynamics*, *Phys. Rev. Lett.*, **113**, no. 25, 2014.
3. J. Verhoeven, T. Wiesehöfer, and S. Stellmach, *Anelastic versus fully compressible turbulent Rayleigh-Bénard convection*, *Astrophys. J.*, **805**, no. 1, 2015.

4. J. S. Cheng, S. Stellmach, A. Ribeiro, A. Grannan, E. M. King, and J. M. Aurnou, *Laboratory-numerical models of rapidly rotating convection in planetary cores*, *Geophys. J. Int.*, **201**, no. 1, 1–17, 2015.
5. J. M. Aurnou, M. A. Calkins, J. S. Cheng, K. Julien, E. M. King, D. Nieves, K. M. Soderlund, and S. Stellmach, *Rotating convective turbulence in Earth and planetary cores*, *Phys. Earth Planet. In.*, **246**, 52–71, 2015.
6. J. Verhoeven, *Thermal convection in compressible fluids and its significance for the zonal winds on Jupiter*, PhD thesis, Westfälische Wilhelms-Universität Münster, 2015.
7. K. Julien, J. M. Aurnou, M. Calkins, E. Knobloch, P. Marti, S. Stellmach, and G. Vasil, *A nonlinear model for rotationally constrained convection with Ekman pumping*, submitted to *J. Fluid Mech.*, 2015.
8. M. A. Calkins, K. Julien, and P. Marti, *The breakdown of the anelastic approximation in rotating compressible convection: Implications for astrophysical systems*, *Proc. Phys. Soc. London, Sect. A*, **471**, no. 2175, 2015.
9. S. Stellmach and U. Hansen, *An efficient spectral method for the simulation of dynamos in Cartesian geometry and its implementation on massively parallel computers*, *Geochem. Geophys. Geosyst.*, **9**, no. 5, 2008.
10. R. E. Ecke and J. J. Niemela, *Heat transport in the geostrophic regime of rotating Rayleigh-Bénard convection*, *Phys. Rev. Lett.*, **113**, no. 11, 2014.
11. M. Sprague, K. Julien, E. Knobloch, and J. Werne, *Numerical simulation of an asymptotically reduced system for rotationally constrained convection*, *J. Fluid Mech.*, **551**, 141–174, 2006.
12. K. Julien, A. M. Rubio, I. Grooms, and E. Knobloch, *Statistical and physical balances in low Rossby number Rayleigh-Bnard convection*, *Geophys. Astrophys. Fluid Dyn.*, **106**, no. 4-5, 392–428, 2012.
13. P. P. Niiler and F. E. Bisshopp, *On the influence of Coriolis force on onset of thermal convection*, *J. Fluid Mech.*, **22**, no. 04, 753–761, 1965.
14. M. Heimpel and J. Aurnou, *Turbulent convection in rapidly rotating spherical shells: A model for equatorial and high latitude jets on Jupiter and Saturn*, *Icarus*, **187**, no. 2, 540–557, 2007.
15. C. Guervilly, D. W. Hughes, and C. A. Jones, *Large-scale vortices in rapidly rotating Rayleigh-Bénard convection*, *J. Fluid Mech.*, **758**, 407–435, 2014.
16. B. Favier, L. J. Silvers, and M. R. E. Proctor, *Inverse cascade and symmetry breaking in rapidly rotating Boussinesq convection*, *Phys. Fluids*, **26**, no. 9, 2014.
17. A. M. Rubio, K. Julien, E. Knobloch, and J. B. Weiss, *Upscale energy transfer in three-dimensional rapidly rotating turbulent convection*, *Phys. Rev. Lett.*, **112**, no. 14, 2014.
18. A. R. Vasavada and A. P. Showman, *Jovian atmospheric dynamics: An update after Galileo and Cassini*, *Reports on Progress in Physics*, **68**, no. 8, 1935–1996, 2005.
19. U. R. Christensen, *Zonal flow driven by strongly supercritical convection in rotating spherical shells*, *J. Fluid Mech.*, **470**, 115–133, 2002.
20. T. Gastine, M. Heimpel, and J. Wicht, *Zonal flow scaling in rapidly-rotating compressible convection*, *Phys. Earth Planet. In.*, **232**, 36–50, 2014.
21. Y. Kaspi, W. B. Hubbard, A. P. Showman, and G. R. Flierl, *Gravitational signature of Jupiter’s internal dynamics*, *Geophys. Res. Lett.*, **37**, no. 1, 2010.

22. A. P. Ingersoll and D. Pollard, *Motion in the interiors and atmospheres of Jupiter and Saturn: Scale analysis, anelastic equations, barotropic stability criterion*, *Icarus*, **52**, no. 1, 62–80, 1982.
23. G. A. Glatzmaier, M. Evonuk, and T. M. Rogers, *Differential rotation in giant planets maintained by density-stratified turbulent convection*, *Geophys. Astrophys. Fluid Dyn.*, **103**, no. 1, 31–51, 2009.
24. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, *Journal of large-scale research facilities*, **1**, A1, 2015, <http://dx.doi.org/10.17815/jlsrf-1-18>.

Computer Science and Numerical Mathematics

Computer Science and Numerical Mathematics

Dietmar Kröner

Department of Applied Mathematics, University of Freiburg,
Hermann-Herder-Str. 10, 79104 Freiburg, Germany
E-mail: dietmar.kroener@mathematik.uni-freiburg.de

In this part we have two contributions which concern geophysical flows with the main focus on high scalability, multigrid acceleration and fault tolerance.

The contribution of B. Gmeiner, M. Huber, L. John, U. Råde, C. Waluga, and B. Wohlmuth about “Massively Parallel Large Scale Stokes Flow Simulation” concerns the efficient computation of solutions of the coupled system consisting of a Stokes-type and a convection-dominated transport equation. This coupled system is considered as a mathematical model for geophysical flows in the earth’s mantle, where special convection patterns of solid material evolve on large time and length scales. Laboratory experiments for the investigations of these problems are impossible and therefore a reliable and efficient numerical simulation is necessary. The numerical method is based on finite element discretisations and geometric multigrid methods.

For the solution of a simplified mantle convection problem an iterative coupling of solving the mass-momentum system and the energy equation is used. Different solver strategies like preconditioned Krylov space and multigrid methods for the first system are studied. It turns out that these methods outperform competing methods significantly. Weak scaling results for the JUQUEEN as well as the robustness with respect to the problem size are excellent.

They also study fault tolerance algorithms, i.e. how to recover nodal values which are lost during a fault. They compared the results for global recovery with those obtained without recovery for a large Laplace problem. The project was funded by the DFG.

In the paper “Analysing the Scalability of Climate Codes Using New Features of Scalasca” by M. Harlacher, A. Calotoiu, J. Dennis, and F. Wolf one of the most important challenge in high performance computing is considered: The efficiency, in particular in this paper the scalability, of algorithms on parallel computers with a very large number of threads. In order to detect scalability bugs, a load-balancing simulator is developed. It can be used to compare the advantages of different load balancing strategies, without changing the code: For instance, the performance impact of different domain decomposition strategies, such as block size, aspect ratio and maximal number of blocks that can be designed. The method is applied to a code for climate simulation and it includes in particular sea ice distribution.

Massively Parallel Large Scale Stokes Flow Simulation

Björn Gmeiner¹, Markus Huber², Lorenz John², Ulrich Ruede¹,
Christian Waluga², and Barbara Wohlmuth²

¹ Institute of System Simulation, University Erlangen-Nuremberg, 91058 Erlangen, Germany
E-mail: {bjoern.gmeiner,ulrich.ruede}@fau.de

² Department of Mathematics, Technische Universität München, 85748 Garching, Germany
E-mail: {huber,john,wohlmuth}@ma.tum.de

In many applications, physical models consisting of a Stokes-type equation that is coupled to a convection-dominated transport equation play an important role, e.g., in mantle-convection or ice-sheet dynamics. In the iterative treatment of such problems the computational cost is usually dominated by the solution procedure for the Stokes part. Hence, we focus on massively scalable and fast multigrid solvers for the arising saddle point problem. To gain deeper insight into the performance characteristics, we evaluate the multigrid efficiency systematically and address the methodology of algorithmic resilience. Three methods based on the HHG software framework will be presented and are shown to solve FE systems with half a billion unknowns even on standard workstations. On petascale systems they furthermore exhibit excellent scalability. This together with the optimised performance on each node leads to superior supercomputing efficiency. Indefinite systems with up to ten trillion (10^{13}) unknowns can be solved in less than 13 minutes compute time.

1 Introduction

Thanks to the continuous improvements made in parallel computing technology, current leading-edge supercomputers can provide up to several petaflop/s of performance provided that suitable algorithms and implementations are designed. While this enables the development of increasingly complex computational models with unprecedented resolution, it also requires a novel co-design process that aims at maximal performance at all stages of developing a simulator. This includes the appropriate choice of mathematical models, discretisations, and algorithms, as well as the matter of software implementation, which all – in their interplay – must be carefully analysed, adapted, and possibly revised to avoid unnecessary inefficiencies in data volume, data traffic, and arithmetic cost. To achieve several levels of parallelism with hundreds of thousands of threads all components must be specifically designed for avoiding synchronisation and communication where possible. This often requires the use of complex hybrid programming models.

In this work, we use the hierarchical hybrid grids framework (HHG)^{1,2} that realises a compromise between structured and unstructured grids. It exploits the flexibility of finite elements and capitalises on the algorithmic efficiency of geometric multigrid methods. The HHG package was initially designed with scalar elliptic equations in mind, see Refs. 1, 2. In Refs. 5, 6 an extension to Stokes systems using a pressure-correction scheme was presented and in Ref. 12 we discuss the conservative coupling to transport equations. For the extension to other types of scalable multigrid-based Stokes solvers, see Ref. 10. In the context of multigrid methods, the execution of a parallel smoothing step for the multigrid algorithm consumes a major part of the total computational cost. The performance of such

algorithms can be substantially improved, e.g. by using suitable discretisations that permit a memory-efficient, matrix-free implementation.

Our application problem is motivated by fundamental geophysical questions revolving around the physics of the Earth’s mantle, which extends from some tens of kilometres below the surface down to the core-mantle-boundary at about 2 900 km depth. In this region, convection patterns of solid material evolve on large time- and length-scales; e.g., Fig. 1 shows the rise of a typical plume formation, displaying the characteristic “mushroom” shape of the iso-temperature surfaces.

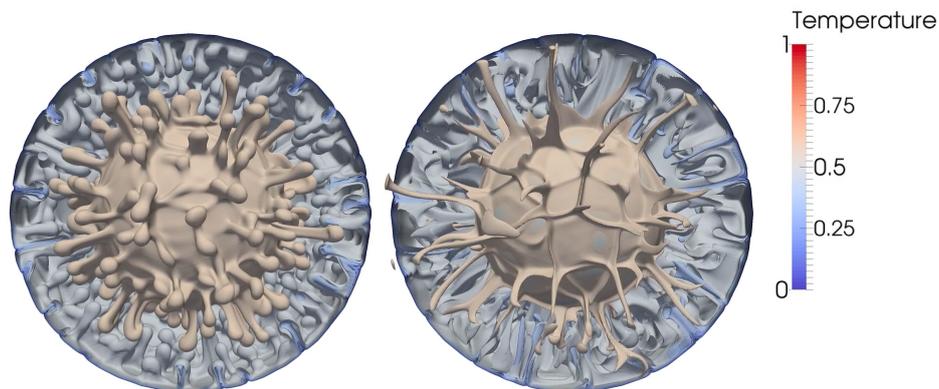


Figure 1. Scaled temperature distribution for a coupled convection simulation. The convective part of the energy balance equation is determined by the velocity solution of a Stokes system.

While the general structure of convection within the mantle is relatively well understood, some important details remain open, including the potential thermo-chemical nature of the convection currents (which are essentially a statement on the buoyancy forces) appropriate rheological parameters and the importance of lateral viscosity variation¹¹. Due to the extreme conditions deep inside the Earth and the large time scales involved, answering these questions is mostly outside the scope of laboratory experiments. Thus, further progress in mantle convection research relies on extracting meaningful answers from the geologic record through a careful assimilation of observations into models by means of fluid dynamics inverse simulations³. There are three aspects making the inversion feasible: the strongly advective nature of the heat transport, the availability of terminal conditions from seismic tomography, providing current temperatures and densities inside the mantle, and the availability of boundary conditions, i.e. surface velocity fields for the past 130 million years, from paleomagnetic reconstructions.

2 Simulation Models

A popular model for mantle convection considers the conservation of mass, momentum and energy in the following form:

$$- \operatorname{div} \boldsymbol{\sigma} = \rho \mathbf{g} \quad (1a)$$

$$\partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0 \quad (1b)$$

$$\partial_t(\rho e) + \operatorname{div}(\rho e \mathbf{u}) = - \operatorname{div} \mathbf{q} + H + \boldsymbol{\sigma} : \dot{\boldsymbol{\epsilon}} \quad (1c)$$

Here, $\boldsymbol{\sigma}$ represents the *stress tensor* while $\dot{\boldsymbol{\epsilon}} = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^\top)$ denotes the *rate of strain tensor* that is defined as the symmetric part of the gradient of the velocity field \mathbf{u} . The vector \mathbf{g} denotes the gravitational acceleration acting in vertical/radial direction. We further denote the internal energy density by e , the heat flux per unit area by \mathbf{q} , and the volumetric radiogenic heat production rate by H . The density $\rho = \rho(p, T)$ is related to the pressure p and the temperature T through an equation of state that is investigated in the field of *mineralogy* and usually represented via lookups or analytical expressions. Finally, the stress tensor $\boldsymbol{\sigma}$ is defined as

$$\boldsymbol{\sigma} = 2\mu(\dot{\boldsymbol{\epsilon}} - \frac{1}{3} \operatorname{tr} \dot{\boldsymbol{\epsilon}} \cdot \mathbf{I}) - p\mathbf{I}, \quad (2)$$

where μ denotes the dynamic viscosity. As mentioned before, the *rheology* of the mantle is not yet well understood⁴, and there are different models for the viscosity. As the focus of this present study is on parallel scalability and efficiency, we only consider numerical examples for a simplified model of mantle convection in the following.

Firstly, we employ the Boussinesq-approximation, i.e., we treat the flow as incompressible everywhere apart from the buoyancy term $\rho \mathbf{g}$. The mass balance (Eq. 1b) then simplifies to the incompressibility constraint $\operatorname{div} \mathbf{u} = 0$. It is well-known that for long term dynamic simulations of coupled problems, local mass conservation is a crucial ingredient. Although we are using stabilised P1 conforming finite element spaces for the velocity and the pressure space, our simulation results do not suffer from the lack of mass conservation. A local post-process^{7,12} guarantees that the velocity flux entering the energy balance equation is locally conservative, and thus no spurious source or sink terms occur. The picture on the right of Fig. 2 shows the vorticity field produced by the modified approach which is in excellent agreement with the reference solution while the one on the left exhibits a physically incorrect structure.

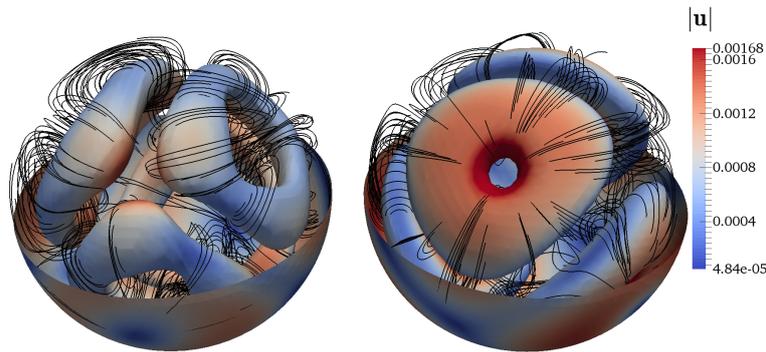


Figure 2. Uncorrected (left) vs. corrected (right) coupling approach: Vorticity contour surfaces for an isoviscous Boussinesq flow with free-slip conditions and a Rayleigh-number of $Ra = 7.7 \cdot 10^4$.

If we additionally assume a constant viscosity, the model can be even further simplified and instead of the symmetric strain operator, we can use the gradient operator in the momentum equation (Eq. 1a) which has the advantage that the different velocity components decouple, and the associated stencils are smaller. By doing so, roughly a factor of two in the time-to-solution can be saved.

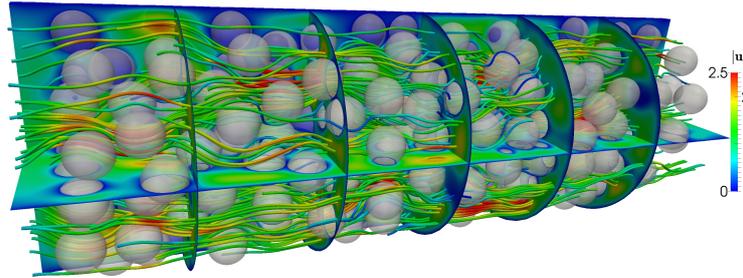


Figure 3. Domain with $n_s = 305$ equally sized, randomly placed spheres and velocity streamlines.

It should be mentioned that our HHG solver is not tailored towards the spherical shell geometry. For instance, in Fig. 3 we present the velocity streamlines of the flow through a cylindrical channel filled with randomly placed spheres on which we impose no-slip boundary conditions. In Fig. 4, we illustrate the effect of different radii. Setups of this type are of relevance for instance when studying infiltration processes.

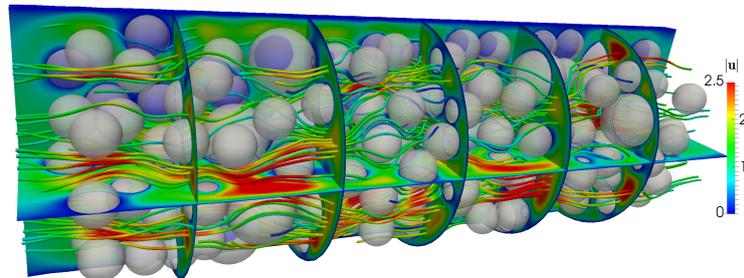


Figure 4. Domain with $n_s = 231$ randomly placed spheres with three different radii and velocity streamlines.

3 Parallel Multigrid Performance

For the solution of the simplified mantle-convection problem on the spherical shell domain we usually consider an iterative coupling, where we solve the mass-momentum system and

the energy equation separately. Since the first part constitutes the most challenging part, we shall in the following employ and compare different solver strategies.

Given a fast solver for a scalar positive definite system, the most natural approach to extend the existing framework to the indefinite Stokes system is to consider the Schur complement for the pressure. After formally performing an elimination of the velocity, the discrete problem for the pressure reads as

$$\mathbf{S}\mathbf{p} = \mathbf{r}. \quad (3)$$

Here \mathbf{S} stands for the pressure Schur complement which is defined by $\mathbf{S} = \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top + \mathbf{C}$, where \mathbf{A} stands for the discrete velocity operator in the momentum balance equation, \mathbf{B} denotes the discrete divergence operator, and \mathbf{C} is the matrix resulting from the stabilisation term that suppresses spurious pressure modes resulting from the equal-order approximation (we use a $P_1 - P_1$ approximation below). The right-hand side is given by \mathbf{r} . In our numerical experiments, we solve Eq. 3 by a preconditioned conjugate gradient method, where we choose a lumped mass matrix preconditioner that is known to be spectrally equivalent to \mathbf{S} ; see e.g. Ref. 8. This simple preconditioning reduces the effects of varying element sizes and shapes and can be extended to account for non-isoviscous flow. Since the direct assembly of the dense matrix \mathbf{S} cannot be performed efficiently, it is applied indirectly by replacing each multiplication of the discrete inverse \mathbf{A}^{-1} by a few cycles of a parallel geometric multigrid algorithm. In addition to the previously described strategy we consider two competing approaches that deal directly with the indefinite nature of the system. The first one is a preconditioned Krylov space method and the second one a multigrid method applied directly to the saddle-point system. This method employs an Uzawa type smoother¹⁰. To obtain a better understanding of the different solvers, we first run the solvers on a conventional, low cost workstation for the serial case with a single Intel Xeon CPU E2-1226 v3, 3.30 GHz and 32 GB shared memory.

		SCG		MINRES		MG Uzawa	
L	DoFs	iter	time	iter	time	iter	time
2	$1.4 \cdot 10^4$	26	0.11	108	0.23	9	0.08
3	$1.2 \cdot 10^5$	28	0.56	83	0.78	8	0.29
4	$1.0 \cdot 10^6$	28	3.33	73	3.99	8	1.79
5	$8.2 \cdot 10^6$	28	24.28	70	26.53	8	12.70
6	$6.6 \cdot 10^7$	31	205.84	67	189.27	8	95.85
7	$5.3 \cdot 10^8$	out of memory		out of memory		8	730.77

Table 1. Iteration numbers and time-to-solution (in sec.) on standard workstation for a constant viscosity.

Roughly speaking, the three approaches based on HHG outperform most competing methods significantly and reach on a workstation the performance when less efficient methods may already require a supercomputer. Beyond this, the multigrid method for the indefinite system with a suitable Uzawa type smoother (MG-Uzawa) outperforms the two alternative approaches with respect to memory and time-to-solution. Tab. 1 shows in detail that with our co-design it is already possible to solve indefinite systems with half a billion unknowns in a few minutes on a standard low cost machine. To further investigate the

characteristics of the different approaches, we identify the required operator applications for all solvers in the same setting. To enable a fair comparison the stopping criteria for all solvers are chosen in exactly the same way. In our results, which we depict in Fig. 5, it can be observed that the Uzawa multigrid method for the indefinite system requires considerably fewer operator evaluations of the operator A and consequently requires a shorter runtime. This effect is expected to become even more significant when considering more complex models, e.g., with varying viscosities.

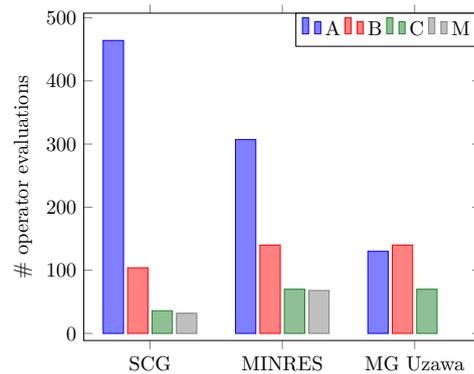


Figure 5. Number of different operator evaluations for the three considered solvers.

Let us next restrict ourselves to the Uzawa multigrid method (MG-Uzawa) and demonstrate some weak scaling results obtained on the JUQUEEN supercomputer (Jülich Supercomputing Centre, Germany), currently listed in the top 10 of the TOP500 list^a. In our numerical results that are presented in Tab. 2, we observe robustness with respect to the problem size and excellent scalability. Additionally to the time-to-solution (time) we present the time without coarse grid (time w.c.g.) and the total amount in % which is needed to solve the coarse grid problems. To approximate the coarse grid problem, we apply a preconditioned Krylov subspace solver. It can be observed that even for the largest example computation, the latter fraction is smaller than one-eighth of the total run-time. The largest problem exceeds 10^{13} unknowns.

Nodes	Threads	DoFs	iter	time	time w.c.g.	time c.g. in %
5	80	$2.7 \cdot 10^9$	10	685.88	678.77	1.04
40	640	$2.1 \cdot 10^{10}$	10	703.69	686.24	2.48
320	5 120	$1.2 \cdot 10^{11}$	10	741.86	709.88	4.31
2 560	40 960	$1.7 \cdot 10^{12}$	9	720.24	671.63	6.75
20 480	327 680	$1.1 \cdot 10^{13}$	9	776.09	681.91	12.14

Table 2. Weak scaling results with and without coarse grid for the spherical shell geometry.

^a<http://top500.org>

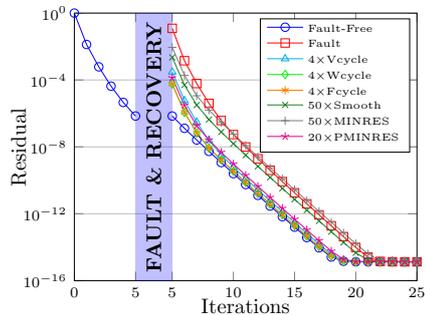
4 Fault Tolerant Algorithms

In the future era of exa-scale computing systems, highly scalable implementations will execute up to billions of parallel threads on millions of compute nodes. In this scenario, fault tolerance will become a necessary property of hardware, software and algorithms. Nevertheless, nowadays commonly used redundancy approaches, e.g., check-pointing, will be too costly, due to the high memory and energy consumption. An alternative and less consuming approach is to incorporate resilient strategies directly into the multigrid solver.

In Ref. 9, we introduce a methodology and data-structure to efficiently recover lost data due to a processor crash (hard fault) when solving elliptic PDEs with multigrid algorithms. We consider a fault model where a processor stores the mesh data of a subdomain including all its refined levels in the multigrid hierarchy. Therefore, in case of a processor failure, we assume that all data is lost in the faulty domain $\Omega_F \subset \Omega$. We further assume that the healthy domain $\Omega_H \subset \Omega$ is unaffected by the fault, and data in this domain remains available. The nodes associated with the interface $\Gamma := \partial\Omega_F \cap \partial\Omega_H$ between the faulty and healthy domain are used to communicate between neighbouring processors by introducing *ghost copies* that redundantly exist on different processors. Therefore, a complete recovery of these nodes is possible without additional storage.

To recover the nodal values $(\mathbf{u}_F, \mathbf{p}_F)$ in Ω_F which are lost during a fault, we propose to solve a local Stokes (subproblem in Ω_F) with Dirichlet boundary conditions on Γ for velocity and pressure, respectively. To guarantee that the local system is uniformly well-posed, we formally include a compatibility condition obtained from the normal components of the velocity. If the local solution in Ω_F is computed while the global process is halted, then this procedure yields a local recovery strategy. In Ref. 9 this local strategy is extended to become a global recovery strategy. To this end, the solution algorithm proceeds asynchronously in the faulty and the healthy domain such that no process remains idle. Temporarily, the two subdomains are decoupled at the interface Γ , and the recovery process in the faulty domain is accelerated by delegating more compute resources to it. This acceleration is termed the *superman strategy*. Once the recovery has proceeded far enough and has caught up with the regular solution process, both subdomains are re-coupled and the regular global iteration is resumed. These approaches result in a time- and energy-efficient recovery. In Fig. 6 (left), we consider a test scenario in $\Omega = (0, 1)^3$ in which we continuously apply multigrid V(3,3)-cycles of the Uzawa multigrid method introduced in Sec. 3. In total 23 iterations are needed to reach the round-off limit of 10^{-15} . During the iteration, a fault is provoked after 5 iterations affecting 2.1% of the pressure and velocity unknowns. As approximate subdomain solvers we compare the fine grid Uzawa-smoother, the minimal residual method, the block-diagonal preconditioned MINRES (PMINRES) method and V, F, W-cycles in the variable smoothing variant from Sec. 3. For the block-preconditioner in case of PMINRES a standard V-cycle and a lumped mass-matrix M are used. Fig. 6 (left) displays also the cases in which no fault appears (fault-free) and when no recovery is performed after the fault.

After the fault, we observe that the residual jumps up and when no recovery is performed, the iteration must start almost from the beginning. A higher pre-asymptotic convergence rate after the fault helps to catch up, so that only four additional iterations are required. This delay can be further reduced by a local recovery computation, but only local multigrid cycles are found to be efficient recovery methods.



Size	DN Strategy $\eta_{\text{super}} = 4$			
	$n_H = 0$	2	3	4
769^3	19.21	0.05	-0.38	4.22
$1\,281^3$	21.27	-0.15	-0.68	3.95
$2\,305^3$	18.50	-0.33	-0.87	3.76
$4\,353^3$	19.74	2.58	4.61	9.24

Figure 6. Laplace problem: left: Convergence of the relative residual for different local recovery strategies. right: Time delay caused by global recovery in terms of a temporary domain decoupling and with a superman factor of $\eta_{\text{super}} = 4$.

The table on the right of Fig. 6 summarises the performance of the global recovery in terms of the time delay (in seconds compute time) as compared to an iteration without faults. The tests are performed for a large Laplace problem discretised with up to almost 10^{11} unknowns. The undisturbed solution is obtained in 50.49 seconds using 14 743 cores on JUQUEEN. Two faults are provoked, one after 5 V-cycles, one after 9. Both faults are treated with both the global recovery strategy and a local *superman* process that is $\eta_{\text{super}} = 4$ times as fast as a regular processor. Faulty and healthy domains remain decoupled for n_H cycles with the Dirichlet-Neumann strategy, i.e., solving a Dirichlet problem on Ω_F and a Neumann problem on Ω_H , then the regular iteration is resumed. The case $n_H = 0$ corresponds to performing no recovery at all and leads to a delay of about 20 seconds compute time. By the superman recovery and the global re-coupling after $n_H = 2$ cycles, the delay can be reduced to just a few seconds. In some cases the fault-affected computation is even faster than the regular one, as indicated by negative time delays in the table.

Acknowledgements

This work was supported (in part) by the German Research Foundation (DFG) through the Priority Programme 1648 “Software for Exascale Computing” (SPPEXA). We are grateful to the Jülich Supercomputing Centre and Leibniz Rechenzentrum for providing computational resources.

References

1. B. Bergen, T. Gradl, U. Rüde, and F. Hülsemann, *A massively parallel multigrid method for finite elements*, Comput. Sci. Engng., 8(6):56–62, 2006.
2. B. Bergen and F. Hülsemann, *Hierarchical hybrid grids: data structures and core algorithms for multigrid*, Numer. Linear Algebra Appl., 11:279–291, 2004.
3. H. P. Bunge, C. R. Hagelberg, and B. J. Travis, *Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography*, Geophys. J. Int., 152(2):280–301, 2003.

4. P. Cordier, J. Amodeo, and P. Carrez, *Modelling the rheology of MgO under Earth's mantle pressure, temperature and strain-rates*, *Nature*, 481:177–180, 2012.
5. B. Gmeiner, U. Rüde, H. Stengel, C. Waluga, and B. Wohlmuth, *Performance and Scalability of Hierarchical Hybrid Multigrid Solvers for Stokes Systems*, *SIAM J. Sci. Comput.*, 37(2):C143–C168, 2015.
6. B. Gmeiner, U. Rüde, H. Stengel, C. Waluga, and B. Wohlmuth, *Towards textbook efficiency for parallel multigrid*, *Numer. Math. Theory Methods Appl.*, 8, 2015.
7. B. Gmeiner, C. Waluga, and B. Wohlmuth, *Local mass-corrections for continuous pressure approximations of incompressible flow*, *SIAM J. Numer. Anal.*, 52(6):2931–2956, 2014.
8. P. P. Grinevich and M. A. Olshanskii, *An iterative method for the Stokes-type problem with variable viscosity*, *SIAM J. Sci. Comput.*, 31(5):3959–3978, 2009.
9. M. Huber, B. Gmeiner, U. Rüde, and B. Wohlmuth *Resilience for multigrid software at the extreme scale*, 2015, submitted.
10. M. Huber, L. John, B. Gmeiner, U. Rüde, and B. Wohlmuth, *Massively parallel hybrid multigrid solvers for the Stokes system*, 2015, in preparation.
11. P. J. Tackley, *Effects of strongly variable viscosity on three-dimensional compressible convection in planetary mantles*, *J. Geophys. Res.*, 101:3311–3332, 1996.
12. C. Waluga, B. Wohlmuth, and U. Rüde, *Mass-corrections for the conservative coupling of flow and transport on collocated meshes*, 2015, submitted.

Analysing the Scalability of Climate Codes Using New Features of Scalasca

Monika Harlacher¹, Alexandru Calotoiu², John Dennis³, and Felix Wolf²

¹ Universität Siegen, 57068 Siegen, Germany
E-mail: monika.luecke@uni-siegen.de

² Technische Universität Darmstadt, 64293 Darmstadt, Germany
E-mail: {calotoiu, wolf}@cs.tu-darmstadt.de

³ National Center for Atmospheric Research, Boulder, CO 80307, USA
E-mail: dennis@ucar.edu

This paper shows how recently developed features of the performance analysis tool Scalasca helped gain important insights into the performance behaviour of state-of-the-art climate codes in the CESM (Community Earth System Model) ensemble. Particular emphasis is given to the load balance of the sea-ice model and the scaling behaviour of the atmospheric model. The presented work is a result of the project *Enabling Climate Simulation at Extreme Scale*, which has been funded through the G8 Research Councils Initiative on Multilateral Research Funding.

1 Introduction

Policy decisions for mitigating climate change or adapting to it are subjects of great discussion throughout the world. Uninformed decisions will impose a heavy cost on future generations, both financial and human. Therefore, it is essential to reduce the current uncertainties about future climate changes and their impact by running climate simulations at 1,000 times larger scales than today. Exascale supercomputers are expected to appear around 2020, featuring a hierarchical design and gathering 100 millions of computing cores. The numerical models of the physics, chemistry, and biology affecting the climate system need to be improved to run efficiently on these extreme systems. Without improvement, these codes will not produce simulation results required to respond to the societal and economical challenges of climate change.

The objective of the G8 ECS (Enabling Climate Simulation at Extreme Scale, 2011-2014) project was to investigate how to run efficiently climate simulations on future exascale systems and get correct results. The project gathered researchers in climate and computer science from Canada, Germany, Japan, Spain, and USA to focus on three main topics: (i) how to complete simulations with correct results despite frequent system failures, (ii) how to exploit hierarchical computers with hardware accelerators close to their peak performance and (iii) how to run efficient simulations with very high numbers of threads.

This article concentrates on the third aspect – the scalability. Subject of the study is the CESM (Community Earth System Model)¹, a fully-coupled ensemble of climate codes maintained at the National Center for Atmospheric Research. It provides state-of-the-art computer simulations of the Earth's past, present, and future climate states. In this study, we analyse the performance of selected CESM codes using new technologies developed in the framework of the Scalasca project², a performance analysis toolset designed for

highly scalable applications. All performance experiments were carried out on the IBM BlueGene/Q system JUQUEEN located at the Jülich Supercomputing Centre.

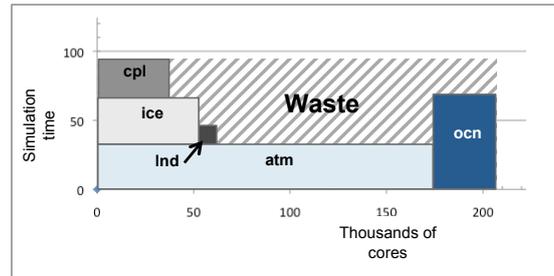


Figure 1. A common configuration of running CESM.

The scalability of CESM is predominantly influenced by the spatial resolution and the performance of the most time-consuming modules, the atmospheric model CAM and the ocean model POP. CAM features very good scalability, while POP scales merely up to a few thousands of cores and is therefore considered as the most severe scalability bottleneck. Since in a common configuration (Fig. 1), the sea-ice model CICE is run alongside POP, one way of compensating for POP's limited scalability is to expand CICE's. This is why major efforts concentrated on this module of CESM. In this article, we focus on CICE and CAM, two studies in which we demonstrate new performance-analysis methods. Whereas we looked at the full model for CICE, we studied CAM using its dynamical core, which is called HOMME³.

Initially, we used only the vanilla version of Scalasca². With the help of Scalasca, we identified critical performance factors and established causal relations between them. During the project, we also used recently added features, such as delay analysis⁴, and added new features on our own whenever necessary. Since load imbalance emerged as a major theme from this initial study, we finally came up with the new idea of a load-balancing simulator that can be used to compare the benefits of different load-balancing strategies without changing the code. We created a design of the simulator, implemented a prototype, and present preliminary results for CICE, which suggest significant optimisation potential. Finally, over the course of the project we developed a novel tool for the automatic detection of scalability bugs. The tool, which is based on automatically generated performance models, was developed in collaboration with the DFG-funded project Catwalk. We applied it to assess HOMME's potential for running at very large scales. More details about the three analysis methods and their results are given below.

2 CICE – The Sea Ice Model

Our study of the sea ice model revealed significant load imbalance. In an early version of the code, all processes receive roughly the same amount of grid cells, which implies huge computational imbalances across processes because the distribution of ice is not uniform. This load imbalance entails wait states in communication sections.

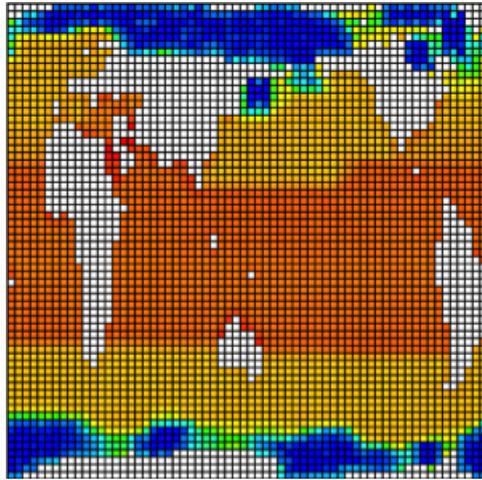


Figure 2. Indirect waiting time for one communication call path in the CESM sea-ice model. Indirect waiting time is waiting time caused by previous wait states as opposed to direct delay. The higher the colour temperature the larger the amount of indirect waiting time.

With our delay analysis extension, we were able to detect a phenomenon caused by the nearest-neighbour exchange pattern used in the application: the propagation of wait states. Near the poles, at the boundary between computationally intensive and less intensive regions, processes with less work wait for messages from processes with more work. This lets them appear overloaded to their neighbours closer to the equator, leading to a chain reaction along which wait states propagate all the way from the poles to the equatorial regions. Fig. 2 depicts the indirect waiting time, which is the waiting time caused by previous wait states and not direct delay through computation overload. It increases with increasing distance from computationally intense regions (poles).

To visualise the performance characteristics with a more advanced domain decomposition, as shown in Fig. 4, we recorded the required decomposition information at runtime. We extended our study to the at that time latest CESM release, 1.0.4, with a decomposition based on space-filling curves as a method for load balancing. This configuration is used by NCAR for high-resolution climate simulations. Already by comparison of the application profiles gained with Scalasca, we were able to measure the performance impact of the different domain decomposition strategies. We experienced a very strong dependence of the application’s execution time on the choice of domain distribution parameters, such as block size and aspect ratio, maximum number of blocks that can be assigned to a process and the kind of the space-filling curve.

Another functionality recently added to Scalasca is the reconstruction of the application’s critical path⁵, the longest execution sequence without wait states in a parallel program. The critical path identifies the activities that determine the overall program runtime, and are therefore preferable candidates for performance optimisation. The higher-level function “evp_haloupdate3dr8” in the dynamics, including communication setup and nearest neighbour data exchange, contributes 27%. Jointly with the computational routines “evp” in the dynamics (with 17%) and “compute_dedd” in the radiation step (16%), these

functions have the largest share of the critical path. The latter two also show high critical imbalance, the time difference between a call path’s contribution to the critical path and the average time spent in the call path across all processes. In a perfectly balanced program, their impact should only amount to 12.5% and 2.9% of the total runtime, respectively. This indicates that the highest potential for performance improvement via better load balancing rests in the function “compute_dedd”.

2.1 Load-Balancing Simulation

In our studies on the scalability and efficiency of CICE, we were able to trace poor scalability of some communication routines back to load imbalance in associated computation phases. Although load balance is crucial for efficient resource utilisation, it is often very difficult for application developers to find a suitable load-balancing strategy that fits their specific problem. The effectiveness of a strategy depends not only on the specific combination of balancing criteria, e.g., computational weight and communication load, but also on the application behaviour, e.g., the communication pattern and sequence of computation and communication phases. Moreover, implementation and test of different load balancing strategies are usually not possible without major code surgery. Therefore, we designed and implemented a load-balancing simulator as a software engineering tool that enables developers to easily test and experiment with different load-balancing strategies. The load-balancing simulator facilitates experiments with different load-balancing strategies and communication patterns without cumbersome analytical comparison or time consuming modifications of the real code and subsequent tests. This information aids developers in choosing a specific method and gives them a guideline whether or not the per-

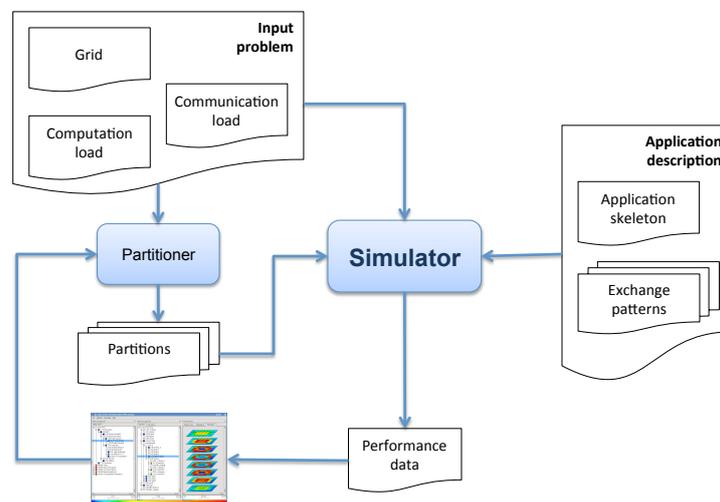


Figure 3. Architecture of the load-balancing simulator. Based on an abstract problem and application description, the simulator re-enacts the application’s communication behaviour under the assumption of different partitions, facilitating an easy comparison in experiments.

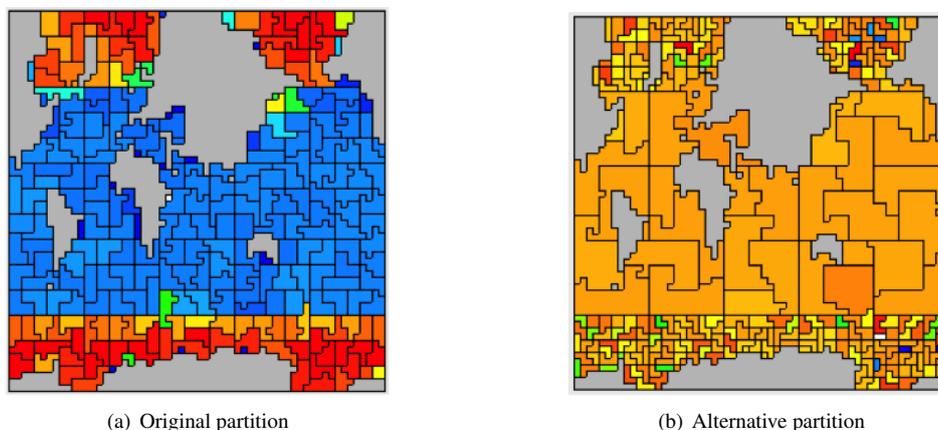


Figure 4. The CESM sea ice model shows severe computational load imbalance between regions of sea ice and open ocean, resulting in large MPI wait time. Using the load-balancing simulator, we replayed this behaviour (left image) and compared it to an alternative partition (right image). The improved load balance reduces MPI wait time and suggests a speedup of up to 2.7 on 256 cores.

formance improvement justifies the effort of software re-engineering. As an example, we applied the load-balancing simulator to CICE and examined its behaviour under different load-balancing strategies.

The architecture of the simulator is depicted in Fig. 3. The simulator operates on a graph structure, where vertices store the computational weight of a geometry item, and edges describe the communication links and volumes between them. This enables representations of arbitrary static grid types, e.g., structured or unstructured meshes of any shape, dimension, and element outline. The simulator also maintains an abstract description of the application’s computation and communication phases to re-enact its behaviour. An application usually has multiple computation phases with communication phases in between. A computation phase is a recurring part in the application where a process executes some local operations. The time required for these operations is specified in the form of vertex weights, one constant weight for each computation phase. Computation is simulated by performing dummy computations lasting the specified amount of time. During communication phases, the simulator re-enacts the specified communication operation, e.g., a stencil operation, with dummy messages of realistic size. The simulator replays the predefined sequence of phases to show how the application would perform under a given input partition. The user can directly compare the runtime of simulations with different partitions or investigate the simulated application behaviour in more detail with performance analysis tools such as Scalasca.

The modular framework of the simulator not only provides pre-defined implementations of common communication patterns, but also allows the user to extend the simulator by adding more complex phase types or by exchanging the underlying communication library. The load-balancing simulator enables scientific application developers to study their application’s behaviour and the associated potential performance yield without cumbersome analytical comparison or time consuming modifications of the real code and subse-

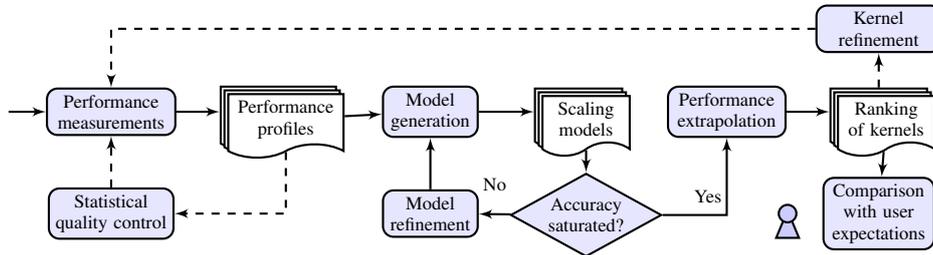


Figure 5. Workflow of scalability-bug detection. Solid boxes represent actions or transformations, and banners their inputs and outputs. Dashed arrows indicate optional paths taken after user decisions.

quent tests. We initially used the simulator to compare the current partition of the sea ice simulation to an alternative partition. In contrast to the current partition, the alternative partition does not restrict block sizes and is based on measured computational weights instead of the probability of sea ice. We estimated a speedup after reconfiguration of 2.7 for 256 cores. Note that the speedup numbers do not yet take into account that to lift the limit on block sizes it would be necessary to allocate the memory dynamically, which might prohibit some compiler optimisations. An investigation of this issue is still in progress.

3 HOMME – The Dynamical Core of the Atmospheric Model

Many parallel applications suffer from latent performance limitations that may prevent them from scaling to larger machine sizes. Often, such scalability bugs manifest themselves only when an attempt to scale the code is actually being made - a point where remediation can be difficult. However, creating analytical performance models that would allow such issues to be pinpointed earlier is so laborious that application developers attempt it at most for a few selected kernels, running the risk of missing harmful bottlenecks. We designed a lightweight performance modelling tool that improves both coverage and speed of this scalability analysis⁶. Generating an empirical performance model automatically for each part of a parallel program, we can easily identify those parts that will reduce performance at larger core counts. Using the HOMME climate simulation as an example, we demonstrated that scalability bugs are not confined to those routines usually chosen as kernels.

Fig. 5 gives an overview of the different steps necessary to find scalability bugs using our method, whose details we explain further below. To ensure a statistically relevant set of performance data, profile measurements may have to be repeated several times - at least on systems subject to jitter. This is done in the optional statistical quality control step. Once this is accomplished, we apply regression to obtain a coarse performance model for every possible program region. These models then undergo an iterative refinement process until the model quality has reached a saturation point. To arrange the program regions in a ranked list, we extrapolate the performance either to a specific target process scale or to infinity, which means we use the asymptotic behaviour as the basis of our comparison. Finally, if the granularity of our program regions is not sufficient to arrive at an actionable recommendation, performance measurements, and thus the kernels under investigation,

can be further refined via more detailed instrumentation.

Going beyond models for just runtime, we now also started to generate empirical models that allow projections for application requirements. Application requirements can be anything such as the required number of floating-point operations, network messages, transmitted bytes, or even memory consumption. System designers can use process-scaling models in tandem with problem-scaling models and the specification of a candidate system to determine the resource usage of an application execution with a certain problem size. Once analytical models are established for an interesting set of requirements, the designer can use them to “play” with configurations such as the amount of memory per node or the network injection speed etc..

To showcase how our tool helps to find hidden scalability bugs in a production code for which no performance model was available, we applied it to HOMME. Being designed with scalability in mind, it employs spectral element and discontinuous Galerkin methods on a cubed sphere tiled with quadrilateral elements. While experiences in the past did not indicate any scalability issues at up to 100,000 processes, HOMME has never been subjected to a systematic scalability study. The result of our analysis was the identification of two scalability issues, one of which was previously unknown⁶. The unknown issue was found in the initialisation and is a reduction needed to funnel data to dedicated I/O processes. It is of the kind usually overlooked when modelling manually. Fig. 6 shows the projected execution times of these two problematic kernels. The initialisation problem is expected to become serious already before hitting one million processes.

The requirements analysis of HOMME revealed some potential future bottlenecks that would need to be mitigated or removed to achieve extreme scaling potential. Specifically, two MPI-collective call paths ending in Allreduce and Bcast show prohibitively growing message sizes ($O(p)$ and $O(p \cdot \log(p))$, respectively).

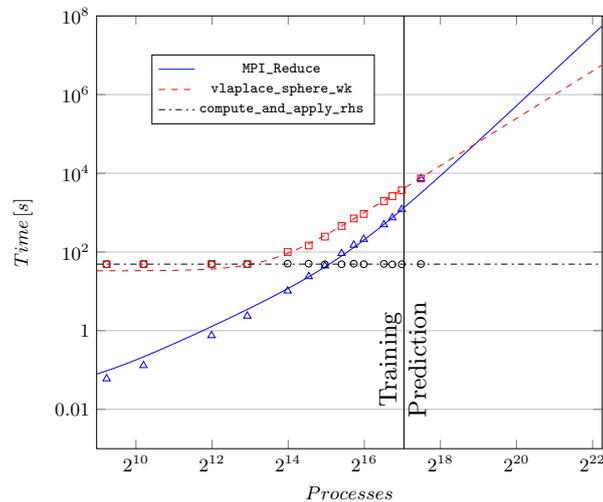


Figure 6. Projected execution times of selected functions in HOMME, two of them representing scalability bottlenecks.

4 Outlook

In the future, we plan to continue our work on load-balance optimisation and requirements analysis for extreme scalability. With respect to the former, we plan to enhance the design of the load-balancing simulator mainly with simplified usage in mind. This includes a more automated way of extracting computation and communication weights from the application and utilities for calibration and validation of the simulation's accuracy. We also plan to enrich the current portfolio of supported load-balancing strategies. Finally, we want to facilitate the simulation of dynamic strategies by modelling them as a sequence of (static) balancing steps.

Acknowledgements

This work received funding through the G8 Research Councils Initiative on Multilateral Research, Interdisciplinary Program on Application Software towards Exascale Computing for Global Scale Issues. Furthermore, the authors gratefully acknowledge the computing time granted by the JARA-HPC Vergabegremium and provided on the JARA-HPC Partition part of the supercomputer JUQUEEN at Forschungszentrum Jülich.

References

1. P. H. Worley, A. A. Mirin, A. P. Craig, M. A. Taylor, J. M. Dennis, and M. Vertenstein, *Performance of the community earth system model*, in: Proc. of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC11), ACM, 2011.
2. M. Geimer, F. Wolf, B. J. N. Wylie, E. Ábrahám, D. Becker, and B. Mohr, *The Scalasca performance toolset architecture*, Concurrency and Computation: Practice and Experience, **22**, no. 6, 702–719, Apr. 2010.
3. J. M. Dennis, J. Edwards, K. J. Evans, O. Guba, P. H. Lauritzen, A. A. Mirin, A. St-Cyr, M. A. Taylor, and P. H. Worley, *CAM-SE: A scalable spectral element dynamical core for the community atmosphere model*, Intl. Journal of High Performance Computing Applications, **26**, no. 1, 74–89, 2012.
4. D. Böhme, M. Geimer, F. Wolf, and L. Arnold, *Identifying the root causes of wait states in large-scale parallel applications*, in: Proc. of the 39th International Conference on Parallel Processing (ICPP), San Diego, CA, USA, pp. 90–100, IEEE Computer Society, Sept. 2010, Best Paper Award.
5. D. Böhme, B. R. de Supinski, M. Geimer, M. Schulz, and F. Wolf, *Scalable Critical-Path Based Performance Analysis*, in: Proc. of the 26th IEEE International Parallel & Distributed Processing Symposium (IPDPS), Shanghai, China, pp. 1330–1340, IEEE Computer Society, May 2012.
6. A. Calotoiu, T. Hoefler, M. Poke, and F. Wolf, *Using Automated Performance Modeling to Find Scalability Bugs in Complex Codes*, in: Proc. of the ACM/IEEE Conference on Supercomputing (SC13), Denver, CO, USA, pp. 1–12, ACM, November 2013.

Fluid Mechanics

Fluid Mechanics

Nikolaus A. Adams

Institute of Aerodynamics and Fluid Mechanics, Technische Universität München,
Boltzmannstraße 15, 85748 Garching, Germany
E-mail: Nikolaus.Adams@tum.de

The variety of subjects reported in this section reflects the significance of fundamental fluid mechanics research for the scientific landscape in Germany. It also reflects the impact high-performance scientific computing has on the further development of research in different areas of fluid mechanics.

The contribution of Valeria Marinova, Iman Kerroumi, Andreas Lintermann, Jens Henrik Göbbert, Charles Moulinec, Sebtiab Rible, Yvan Fournier und Mehdi Behbahani, the University of Applied Science in Jülich, the Jülich Aachen Research Alliance, the Daresbury Laboratory, and Electricite de France perform flow simulations based on the Reynolds-averaged Navier-Stokes equations for a “Numerical Analysis of the FDA Centrifugal Blood Pump”. Operation-condition parameter variations are performed and flow-field data are evaluated in order to identify hemolysis and thrombosis risk due to high local shear forces. The authors conclude that a better resolution of turbulent flow structures is needed for deriving definite conclusions.

The contribution of Sebastian Kreuzahler, Rainer Grauer, Holger Homann, and Yannick Ponty of the Ruhr-Universität Bochum and the Observatoire de la Côte d’Azur employ “Direct Numerical Simulations of Impeller Driven Turbulence and Dynamo Action” for detailed comparisons of well-resolved simulation data with experiments on the MHD dynamo effect. The high resolution properties of the employed pseudo-spectral scheme and its high computational efficiency have allowed to reproduce accurately experimental data and to support experimental findings on the MHD dynamo mechanisms.

The contribution of Fabian Proch, Martin Rieth, Andreas Rittler, and Andreas Kempf of the University of Duisburg-Essen address the “Highly-resolved Numerical Simulation of the Turbulent Combustion Process in Experimental Burners”. Well-resolved LES and DNS of different burner configurations, including those involving spray combustion and particle combustion, are considered, and analysed with respect to general flow structure and mixing efficiencies.

The contribution of Mohammad S. Emran, Paul Götzfried, Anastasiya Kolchinskaya, Janet D. Scheel, and Jörg Schumacher of TU Ilmenau and the Occidental College in Los Angeles perform “Supercomputing Studies in Turbulent Rayleigh-Benard Convection: Challenges and Perspectives”. At typical parameters the simulations reproduce faithfully established flow patterns. In order to address parameter ranges that are more typical for observations in nature or engineering, specifically designed new simulations have been performed at small Prandtl numbers that reveal qualitative differences from the earlier simulations.

The contribution of Siarhei Khirevich and Ulrich Tallarek of the Philipps-Universität Marburg and KAUST “On the Parametrisation of Lattice Boltzmann Method in Pore-scale

Flow Simulations” analyses the effects of different packing types on permeability. Spatial resolution is varied in order to determine practical guidelines for application computations.

Numerical Analysis of the FDA Centrifugal Blood Pump

Valeria Marinova¹, Iman Kerroumi¹, Andreas Lintermann², Jens Henrik Göbbert², Charles Moulinec³, Sebastian Rible¹, Yvan Fournier⁴, and Mehdi Behbahani¹

¹ University of Applied Sciences Aachen, Campus Jülich,
Heinrich-Mußmann-Straße 1, 52428 Jülich, Germany

E-mail: {valeria.vereshchagina, iman.kerroumi}@alumni.fh-aachen.de, behbahani@fh-aachen.de

² Jülich Aachen Research Alliance (JARA-HPC), RWTH Aachen University,
Schinkelstraße 2, 52062 Aachen, Germany

E-mail: {lintermann, goebbert}@jara.rwth-aachen.de

³ Scientific Computing Dept. Science and Technology Facilities Council Daresbury Laboratory,
Sci-Tech Daresbury Warrington, WA4 4AD, United Kingdom

E-mail: charles.moulinec@stfc.ac.uk

⁴ EDF - R&D Fluid Mechanics, Energy and Environment Department,
6 quai Watier - BP 49, 78401 CHATOU Cedex, France

E-mail: yvan.fournier@edf.fr

Ventricular Assist Devices (VADs) are commonly implanted to assist patients suffering from heart diseases. They provide long- and short-term support for the human heart and help patients to recover from heart attacks and from congestive heart failure. It is essential to design blood-sensitive *VADs* to minimise the risk of hemolysis and thrombosis. The blood pump, however, must operate at a wide range of flow rates and pressure heads which makes a low-risk design a challenging task. In this study the flow in a centrifugal blood pump, provided by the *U.S. Food and Drug Administration (FDA)*, is investigated by means of numerical simulations on high performance computers. The simulations are carried out for different operation REYNOLDS numbers. A total of 15 pump revolutions is performed to obtain quasi-steady results. The pressure drop across the pump is considered to study convergence of the solution and to characterise the energy loss of the device. Investigations of the velocity field show that there exist high velocities and strong velocity gradients and shear layers in the outflow region potentially leading to hemolysis. Investigations of the wall-shear stress reveal the existence of thin boundary layers at the blade tips. Finally, the motor torque is investigated to identify the force acting on the blades. All the findings show that there is a strong need to develop more blood-sensitive designs to reduce the risk of hemolysis and thrombosis.

1 Introduction

In 2014, about 26 million people worldwide suffered from heart failure¹, which is the most common fatal disease in developed countries. Heart failure develops gradually as the heart muscle weakens. It is caused by a dysfunction of the heart's ventricle that limits the cardiac performance and can often only be treated by heart transplant. However, only 5,000 donor hearts become available each year, while 50,000 patients remain waiting².

Ventricular Assist Devices (VADs) are the only viable bridging solution allowing patients to survive until a donor is found. Such devices consist of inflow and outflow cannulas and a blood pump to support the ventricle pumping. The major advantage, e.g., over a *Total Artificial Heart (TAH)*, is that the native heart remains in the body while the *VAD* sustains the natural blood flow by means of an external power supply. Two different major kinds

of blood pumps exist, i.e., axial and centrifugal pumps. Axial blood pumps consist of a rotor type impeller mounted in a small housing rotating at 8,000 – 15,000 rpm³. Thereby, the pump accelerates the blood in the streamwise direction based on the principle of the Archimedean screw. High rotational speed may lead to non-physiological flow conditions, i.e., high shear stress and the formation of recirculation zones. Such phenomena may induce hemolysis and thrombosis. High shear rates may damage the red blood cells causing excessive hemolysis which may lead to kidney and hence multiple organ dysfunction. The formation of recirculation zones increases the likelihood of blood clots, which may circulate in the cardiovascular system and obstruct important arteries or veins. In contrast to axial pumps, centrifugal pumps operate at a speed of 1,400 – 4,000 rpm and consist of a non-occlusive pump head with several rotor blades mounted in a pump housing⁴. The fluid enters through the head and is centrifugally accelerated by the perpendicularly mounted pump head. The accelerated flow decelerates in the diverging outflow cannula causing a pressure increase. As fluid is displaced at the discharge side of the pump, more fluid is sucked in to replace it at the suction side leading to flow⁵.

Blood pumps must operate over a wide range of flow rates and pressure heads. Another requirement is minimisation of the damage to blood cells. The intricacy of the problem makes *Computational Fluid Dynamics (CFD)* a powerful tool for the analysis and design of blood pumps. A review of the design and numerical analysis of both centrifugal and axial blood pumps is given in Behbahani *et al.*⁶.

The investigations in this study concentrate on a centrifugal blood pump provided by the *U.S. Food and Drug Administration (FDA)*. The *FDA* aims at characterising the feasibility of *CFD* methods for the design of medical devices. That is, they review the disadvantages and problems of medical devices and call health professionals' attention to their findings. At the end of 2013 the *FDA* invited academical as well as industrial organisations of the *CFD* community to take part in a benchmark study that involves the analysis of the flow in a typical blood pump. The study is composed of six test cases with different pump speeds and REYNOLDS numbers in the range of $Re = 210 \cdot 10^3$ to $Re = 293 \cdot 10^3$. The high REYNOLDS number regime requires large meshes to highly resolve the flow and *High Performance Computing (HPC)* is mandatory to perform the simulations. They have been carried out on the JUQUEEN⁷ system at the *Jülich Supercomputing Centre (JSC)*. Within this work, the velocity field, secondary flow structures, the wall-shear stress, and the motor torque are investigated to develop an understanding of the blood sensitivity of the *FDA* blood pump. For validation of the numerical results the *FDA* has contracted 3 independent laboratories to produce experimental data of the flow through the pump by means of *Particle Imaging Velocimetry (PIV)*.

This paper is organised as follows. The numerical method for the simulation of the flow in the centrifugal *FDA* blood pump and the code scalability are described in Sec. 2. Subsequently, the results of the simulations at several REYNOLDS numbers are presented in Sec. 3. Finally, conclusions are drawn in Sec. 4.

2 Numerical Methods

In this section the numerical methods, i.e., the flow solver, the geometry of the problem and the mesh generation as well as a scalability analysis of the simulation code are presented.

2.1 Flow Solver

To simulate the flow in the *FDA* centrifugal blood pump a *Finite Volume Method (FVM)* is used since it is well suited for *CFD* simulations of high *REYNOLDS* number flows in complex geometries. In addition, *FVM* is known to demonstrate good parallel performance⁸.

Code_Saturne^a implements these advantages and is well suited for three-dimensional calculations of steady or transient single-phase, incompressible or compressible, laminar or turbulent flows⁹. Advanced and accurate turbulence modelling for *Reynolds-Averaged Navier-Stokes (RANS)* and *Large-Eddy Simulations (LES)* represents a key feature of the code. Its flow solver is based on a *FVM*, with a fully co-located arrangement for all variables and a predictor-corrector scheme for the time discretisation of the Navier-Stokes equations. Parallelisation is handled using *MPI*, and some *OpenMP* pragmas are implemented in version 3.2.x and 3.3.x of the code. *Code_Saturne* is developed since 1997 at EDF R&D and has found massive application for a variety of fluid mechanical problems on several high-end machines¹⁰.

Within the framework of this study several *RANS* turbulence models were tested including first order models such as *k- ϵ* and *k- ω* and, a *Reynolds Stress Model (RSM)*, namely the *Rij-SSG (Speziale, Sarkar, Gatzki)*^{11,12}. It was shown that the *RSM* represents an appropriate choice with respect to modeling high-*REYNOLDS* number flow cases as isotropic first order models (*k- ϵ* , *k- ω*) without rotation correction do not correctly model turbulence effects with rotation, which induces anisotropy.

2.2 Geometry and Mesh Generation

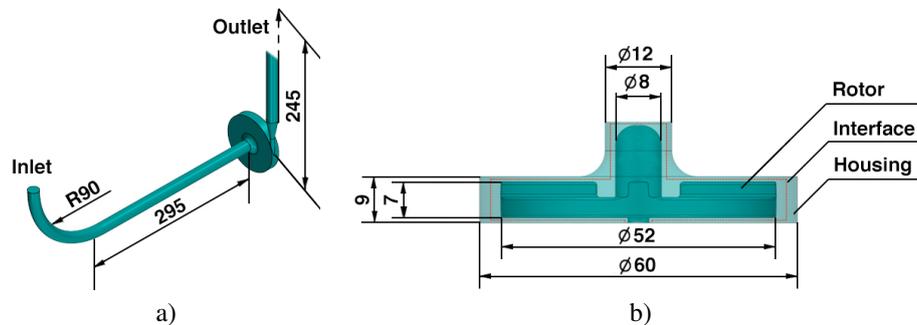


Figure 1. a) Picture of the *FDA* centrifugal blood pump, showing the long curved inlet, the pump and a part of the long outlet; b) cross section of the pump including dimensions and interface (red line) between the rotor and the housing which separates the rotating and static flow regions. All dimensions are given in mm.

The *FDA* centrifugal blood pump consists of two main components: the housing and the rotor positioned inside it. The blood flows through a curved inlet tube into the housing interior, where it meets a hub and then is rotated within the interior of the housing by means of the rotor. Finally, blood exits the pump through a diffuser and continues into the outlet (see Fig. 1).

^a<http://www.code-saturne.org/>

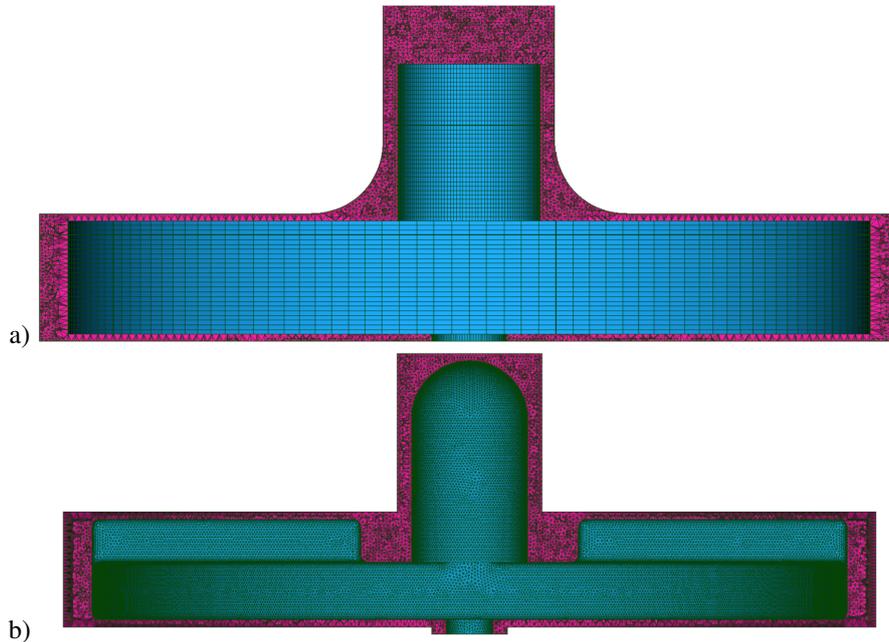


Figure 2. Cross-sectional view (in purple) of a) static domain and b) dynamic domain meshes. At the interface region pyramids can be seen.

For the grid generation SALOME version 6.6.0^b was used. To be able to simulate the motion of the rotor, an interface is introduced and the mesh is split into two sub-meshes, one for the rotor (which is moving) and one for the stator (which is static). The first sub-mesh consists of the mesh of the rotor in a cylinder, and the second sub-mesh of the mesh of the rest of the geometry minus the aforementioned cylinder (see the interface in Fig. 1 b). An efficient and fully parallel mesh joining strategy is used at each time step between both sub-meshes to create a full mesh used to solve the *RANS* equations. Hybrid meshes were generated consisting of pyramids and tetrahedra, which had a defined number of regularly arranged quadrangular surface elements at the interface. The corresponding sub-meshes are shown in Fig. 2. The pump region, i.e., the first sub-mesh, was made of 14 million cells at first and was later refined to 21 million cells for better resolution of critical areas. The final full mesh including also the long inlet and outlet contained a total of 76 million cells. The cell volume ranged between $5.05 \cdot 10^{-15} \text{m}^3$ and $9.14 \cdot 10^{-12} \text{m}^3$. The mesh was designed fine enough to serve for the production runs of the *FDA* simulations, and scalable wall-functions were used to better capture boundary layers. The same 76 million cell mesh was used for all 6 simulations (see Tab. 1), which were conducted on the IBM Blue Gene/Q JUQUEEN at Jülich Supercomputing Centre¹³.

After computing the flows using the *Code_Saturne* version 3.2.x, the flow structures are visualised using ParaView¹⁴ version 3.14.1.

^b<http://www.salome-platform.org/>

2.3 *Code_Saturne* Scalability Study

Code_Saturne has demonstrated extreme scalability up to 1,572,864 *MPI* tasks, using 32 ranks per node on the IBM Blue Gene/Q Mira, Argonne National Laboratory. Fig. 3 shows the speed-up of *Code_Saturne* for three mesh sizes, consisting of 111 million, 889 million, and 13 billion cells, respectively. The first two meshes (blue and green solid lines respectively) have been run for the classical lid-driven cavity test case, using tetrahedral cells. The third test (red solid line) is based on a mesh made of hexahedral cells and was designed for *LES* in tube bundles.

A scalability study was also performed for this case and included test runs for all 6 cases which *FDA* asked for. Effects of partitioning algorithms on the scalability were also investigated, using either serial METIS¹⁵ or PT-Scotch¹⁶. For a number of 1,024 - 65,536 processors METIS turned out to be the most efficient partitioner, requiring 11.7 seconds to compute one iteration.

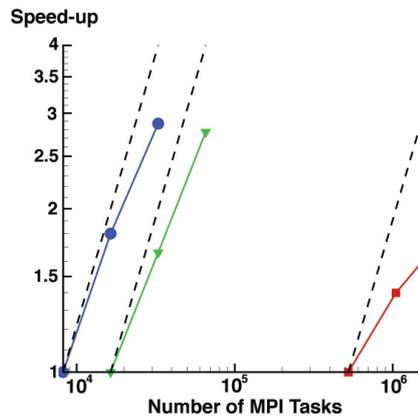


Figure 3. Speed-up of *Code_Saturne* on Blue Gene/Qs, as a function of the number of *MPI* tasks. The blue (111M cells) and green (889M cells) plots with respectively circle and triangle symbols are obtained on Blue Joule (Hartree Centre, STFC, UK) and the red (13B cells) plot (square symbols) on Mira (DOE, Argonne, US).

3 Results of the Centrifugal Pump Analysis

Similar to the REYNOLDS number definition for stirred vessel flow, the *FDA* defined the REYNOLDS number as $Re = \rho v 2\pi d^2 / \eta$, where $\rho = 1035.0 \text{ kg/m}^3$ is the blood density, v is the respective pump's rotational speed, $d = 0.052 \text{ m}$ is the rotor diameter and $\eta = 0.0035 \text{ Pa}$ is the blood viscosity. Blood is known to display non-Newtonian properties¹⁷. However, in experimental studies¹⁸ it has been shown that at shear rates $\dot{\gamma} > 100 \frac{1}{s}$, the viscosity of human blood with physiological hematocrit reaches a constant value, thus justifying the choice of a Newtonian model at moderate and high shear rates. The according flow conditions for all 6 cases are listed in Tab. 1. The data files which had to be

Case	volume flow rate [L/min]	pump speed [rpm]	REYNOLDS number
1	2.5	2500	293,073
2	2.5	3500	209,338
3	4.5	3500	293,073
4	6.0	2500	209,338
5	6.0	3500	293,073
6	7.0	3500	293,073

Table 1. Standard setup parameters of the blood pump benchmark as defined by the *FDA*.

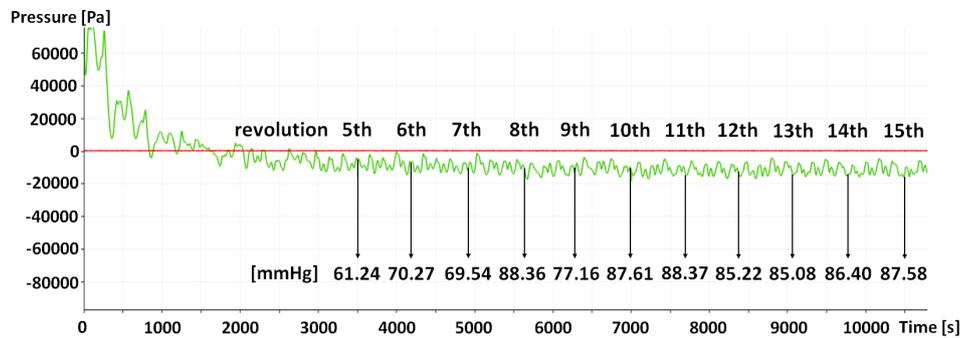


Figure 4. Pressure evolution as a function of time for 15 revolutions for the example of simulation case 6. Pressure values averaged over a full revolution are given in units of mmHg. The red curve represents *zero*-pressure at the outlet and the green curve shows the pressure at the inlet.

submitted to the *FDA* included velocities and pressures in the blade passage plane and in the outlet plane for all 6 simulation cases for prescribed rotor positions.

Depending on the flow case a fixed time step of $2.45 \cdot 10^{-5}$ s or $3.43 \cdot 10^{-5}$ s was applied for typically 10,500 iterations in order to compute 15 revolutions of the rotor to approach a quasi-stationary state which is shown in Fig. 4 for simulation case 6. The computational effort (used CPU time) for computing one revolution of the blood pump while typically using 32,768 MPI tasks for the 76 million cell mesh resulted in 11,279s or 3h 13min. It can be observed that the solution converges towards a quasi-steady state and that the averaged pressures over each revolution approach near-constant values, where a value of 80 – 90mmHg corresponds to physiological pressure ranges.

Velocity and wall shear stress distributions for flow case number 5 are shown in Fig. 5. The highest overall velocities are found in the narrow portion of the diffuser at the pump outlet.

Fig. 6 shows the velocity field inside the blood pump housing. Inside the pump the highest velocities occur behind the rotor blades. From the representation by means of velocity vectors it can be observed that once the fluid has left the low-velocity inflow region, fluid particles follow a circumferential path. Strong radial components are only visible closely behind the rotor blades. Centrifugal forces pushing the particles to the outside are adding an additional radial velocity component in the wake of the blades. The pressure

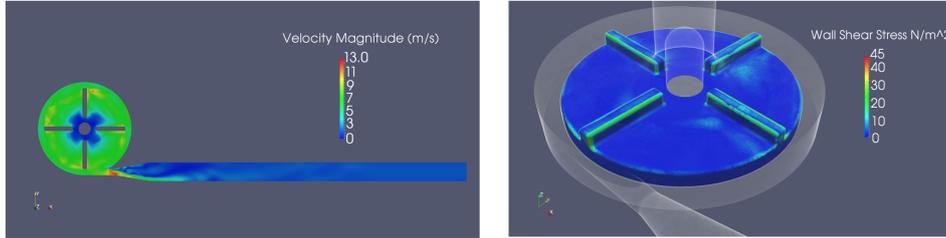


Figure 5. Velocity distribution in the x,y-plane (left) and wall shear stress at the rotor surface (right) for simulation case 5.

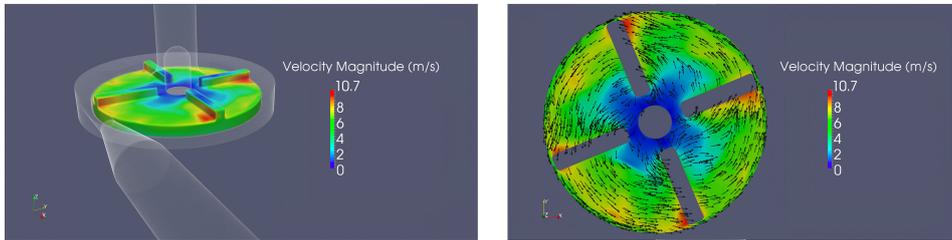


Figure 6. Velocity field inside the *FDA* blood pump housing (left) and velocity vector distribution at close proximity to rotor surface (right) for simulation case 6. Rotor rotates in anti-clockwise direction.

loss caused by the velocity increase further accelerates particles in that region. The pressure distribution (data not shown here) in the pump clearly shows low pressures behind blade tips and high values in front of the tips and thereby confirms this interpretation. Another explanation for the elevated velocity values could be that there exist tip vortices in the wake of the pump blades that rotate at high velocities, which add velocity behind the blade. However, it is possible that *RANS* cannot capture these structures and *LES* computation and analysis is necessary in the future studies.

Tab. 2 shows a summary of characteristic pressure, wall shear stress and torque values. The given pressure head values represent the time-averaged pressure difference between

Case	pressure head [10^3 N/m^2]	wall shear stress [N/m^2]	shaft torque [10^{-3} Nm]
1	77.6	61.2	8.4
2	49.3	27.4	13.3
3	38.6	135.5	18.8
4	-5.1	72.9	15.9
5	24.4	129.0	24.5
6	14.9	119.8	26.7

Table 2. Time-averaged pressure head (between inlet and outlet) over last rotation, time-averaged wall shear stress magnitude over the housing rim and shaft torque for all 6 *FDA* simulation cases.

the outlet and inlet. To that purpose pressure values were averaged over the inlet and outlet areas respectively for a given rotor position. The time-averaged wall shear stress magnitude was computed over the housing rim and was averaged over the last full rotation. Finally, the torque values were computed from the velocity derivatives acting parallel to the rotor surface and pressures perpendicular to the blades. According to the computed values for the pressure head, simulation case 4 is not a feasible operating condition as a very low and negative pressure would not support the required pumping function. All other cases represent operating conditions, where the computed pressure heads lie in the range of typical values reported in literature¹⁹. Values found for commercial blood pumps lie typically in a range of 80 mmHg – 200 mmHg (10,666 – 26,666 Pa) but can go up to values of 700mmHg (93,331 Pa) depending on the operating conditions. The computed wall shear stress values are especially high for the simulation cases 3, 5 and 6. Values between 119.8 – 135.5 Nm may cause hemolysis, especially because in centrifugal blood pump the blood does not necessarily leave the pump after the first rotation. The blood may rotate several times inside the housing, which strongly increases the shearing time for blood cells and can cause strong deformation or rupture of these cells.

Post-processing and visualisation of the simulation data were performed on local computers for distinct time steps and on JUVIS at the JSC if many time steps had to be loaded simultaneously. Besides the various pictures requested by the FDA, it was possible to produce dynamic visualisation of the data including a movie of the blood flow situation inside the FDA pump for the 76 million cell mesh. A snapshot from that movie is shown in Fig. 7. About 1000 frames were generated through ParaView which amounts for over two and a half revolutions. Pathline generation (not shown here) could also be performed.

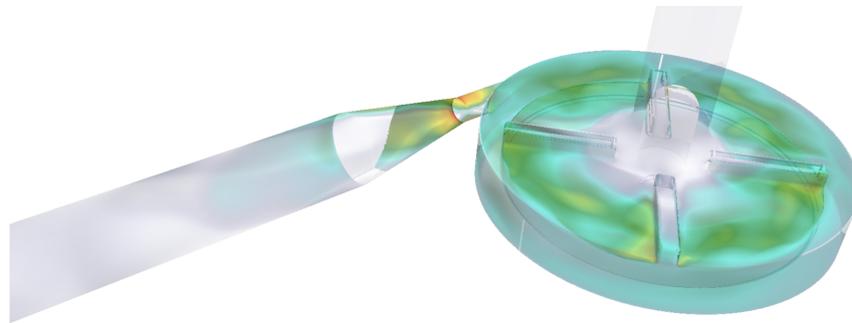


Figure 7. Screenshot from a movie which was made for the 76 million mesh visualising the flow inside the FDA blood pump. Colours represent the velocity magnitude. Rotor rotates in anti-clockwise direction.

4 Conclusion and Outlook

In this study the blood flow could be characterised for different operating conditions of the FDA centrifugal blood pump. High shear stress regions could be identified both inside the housing and in the diffuser region. The high shear values detected for some of the flow cases may lead to hemolysis. The presented CFD study results can therefore serve

to improve existing centrifugal blood pumps to obtain a hemocompatible design. The required shaft torque, an important characteristic value for the operation of a centrifugal blood pump, could equally be determined. All quantitative data and visualisations required for the benchmark could be sent to the *FDA*. The performed flow field characterisation can serve in the future to quantitatively estimate hemolysis using a mathematical model of red blood cell deformation and damage. In the context of this study, *Code_Saturne* could be shown to be an efficient, scalable code which is suitable for the analysis of highly turbulent rotating flows.

One important next step will be to compare the numerical results with real measurements. Until now the *FDA* has not yet published any experimental results from the three laboratories that were asked to perform comparative *PIV* studies but the publication of such results is expected soon. From comparison to literature and from analytical considerations it can already be concluded that the computed results are of the right order.

Another next step will be to perform *Large-Eddy Simulations* (*LES*) for the *FDA* blood pump. *LES* computations will essentially help to understand the unsteady and highly dynamic flow in blood pumps and will support design studies for future *VADs*. The findings will help to reduce the high shear stress in such devices and will hence minimise the risk of hemolysis and thrombosis for the patient. These studies represent a step towards the establishment of reliable methods to accurately predict blood flow and blood damage in fast rotating biomedical devices. The *LES* computations will increase the computational complexity and will at the same time require finer meshes of approximately 200 million elements to capture turbulent effects, thus making the continued use of high performance computing necessary.

Acknowledgements

We would like to thank G. Groten for many fruitful discussions on the intricacies of Fortran 90 and for the coding of a significant part of the interface library.

The authors gratefully acknowledge the *Gauss Centre for Supercomputing* (*GCS*) for providing computing time through the *John von Neumann Institute for Computing* (*NIC*) on the *GCS* share of the supercomputer JUQUEEN⁷ at the *JSC*.

Furthermore, the authors thank the Simulation Laboratory “Highly Scalable Fluids & Solids Engineering” and the Cross-Sectional Group “Immersive Visualisation” of the *Jülich Aachen Research Alliance, High Performance Computing* (*JARA-HPC*) and *JSC*.

The authors would also like to thank the Hartree Centre for using their Blue Gene/Q to generate data for the film and the INCITE PEAC programme of the US DOE for letting them perform extreme scalability studies using MIRA.

References

1. P. Ponikowski, S. D. Anker, K. F. AlHabib, M. R. Cowie, T. L. Force, S. Hu, T. Jaarsma, H. Krum, V. Rastogi, L. E. Rohde, U.C Samal, H. Shimokawa, B. Budi Siswanto, K. Sliwa, and G. Filippatos, *Heart Failure: Preventing Disease and Death Worldwide*, *ESC Heart Failure* **1**, 4-25, 2014.

2. D. O. Taylor, L. B. Edwards, M. M. Boucek, E. P. Trulock, P. Aurora, J. Christie, F. Dobbels, A. O. Rahmel, B. M. Keck, and M. I. Hertz, *Registry of the International Society for Heart and Lung Transplantation: Twenty-fourth Official Adult Heart Transplant Report-2007*, Journal of Heart and Lung Transplantation **8**, 769-781, 2007.
3. D. L. Mann, D. P. Zipes, P. Libby, R. O. Bonow, *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine* Elsevier **10**, 596, 2015.
4. A. D. Kirk, S. J. Knechtle, C. P. Larsen, J. C. Madsen, T. C. Pearson, S. A. Webber, *Textbook of Organ Transplantation Set* Wiley **10**, 559, 2015.
5. C. Allerstorfer, *Centrifugal Pumps* MU Leoben, 2013.
6. M. Behbahani, M. Behr, M. Hormes, U. Steinseifer, D. Arora, O. Coronado, M. Pasquali, *A Review of Computational Fluid Dynamics Analysis of Blood Pumps* European Journal of Applied Mathematics **20**, 363-397, 2009.
7. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities **1**, A1, 2015.
8. Y. Fournier, J. Bonelle, C. Moulinec, Z. Shang, A. G. Sunderland, and J. C. Uribe, *Optimizing Code_Saturne computations on petascale systems*, Computers and Fluids **45**, 103-108, 2011.
9. F. Archambeau, N. Méchitoua, and M. Sakiz, *Code_Saturne: A finite volume code for the computation of turbulent incompressible flows-Industrial applications*, International Journal on Finite Volumes **1**, 1, 2004.
10. C. Moulinec, D. R. Emerson, Y. Fournier and P. Vezolle, *Challenges to be Overcome for Engineering Software to Run Efficiently on Petascale Machines*, in B.H.V. Topping and P. Ivanyi - Saxe-Coburg Publications, 23-40, 2013.
11. C. Speziale, T. Gatski, and N. Mhuiris, *A critical comparison of turbulence models for homogenous shear flows in a rotating frame*, Physics of Fluids **9**, 1678-1684, 1990.
12. C. Speziale, S. Sarkar, and G. Gatski, *Modeling the pressure strain correlation of turbulence - an invariant dynamical systems approach*, Journal of Fluid Mechanics **227**, 245-272, 1991.
13. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities **1**, 1-18, 2015.
14. U. Ayachit, *The ParaView Guide: A Parallel Visualization Application*, Kitware, ISBN 978-1930934306, 2015.
15. G. Karypis and V. Kumar, *A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs*, SIAM Journal on Scientific Computing **20**, 359, 1998.
16. C. Chevalier and F. Pellegrini, *PT-Scotch: A tool for efficient parallel graph ordering*, Parallel Computing **34**, 318-331, 2008.
17. P. Easthope and D. Brooks, *A comparison of rheological and constitutive functions for whole human blood*, Biorheology **17**, 235-247, 1980.
18. S. Chien *Shear dependence of effective cell volume as determinant of blood viscosity*, Science **168**, 977-979, 1970.
19. K. Franco and E. Verrier, *Advanced Therapy in Cardiac Surgery*, in PMPH USA 2nd edition, 2013.

Direct Numerical Simulations of Impeller Driven Turbulence and Dynamo Action

Sebastian Kreuzahler¹, Rainer Grauer¹, Holger Homann², and Yannick Ponty²

¹ Institute for Theoretical Physics I, Ruhr-Universität Bochum,
Universitätsstraße 150, D-44780 Bochum, Germany
E-mail: {grauer, sek}@tp1.ruhr-uni-bochum.de

² Laboratoire Lagrange UMR7293, Observatoire de la Côte d'Azur,
CS34229, 06304 Nice Cedex 4, France

The process, in which a magnetic field is amplified by the flow of an electrically conducting fluid, known as dynamo action, is believed to be the origin of many magnetic fields in the universe including the magnetic field of the earth. A successful laboratory experiment investigating the underlying mechanisms is the Von Kármán Sodium device, consisting of a cylindrical vessel filled with liquid sodium, stirred by two counter-rotating soft-iron impellers. Despite its success, it leaves important questions unsolved and even raises new ones. The aim of this project are detailed high-resolution direct numerical simulations of the VKS experiment. This type of simulations of a three-dimensional turbulent fluid flow in complex geometries supporting a magnetic field are challenging. We designed a massively parallel pseudo-spectral MHD (magneto-hydrodynamics) solver that models the geometry of rotating impellers via a penalisation technique. Benchmarks show a good quantitative agreement with experimental data. The investigation of hydrodynamic properties of the system reveals the generation of conical vortices close to the blades, which may provide a major contribution to dynamo action. We achieve dynamo action in simulations of the full magneto-hydrodynamic system. A variation of the impeller material constants (steel \rightarrow soft iron) towards the experimental value leads to a significant decrease of the dynamo threshold as well as a change of the magnetic field mode as observed in the experiment.

1 Introduction

The self amplification of a magnetic field induced by stretching and twisting magnetic field lines by a fluctuating flow, known as dynamo action, is believed to be the main mechanism for generating magnetic fields in the universe including the magnetic field of the earth. In order to gain a better understanding of the underlying processes different experimental groups have investigated dynamo action in laboratory experiments using liquid sodium^{1,2}.

A very successful experiment is the Von-Kármán-Sodium hosted in Cadarache which was able to reproduce dynamo action in a turbulent flow in 2006. They use impellers to stir liquid sodium in a cylindrical vessel. It has been found that the shape and the material of the impellers play a crucial role for the efficiency of the dynamo action³. However, a detailed understanding of their respective roles is still lacking also because the material properties of sodium make *in situ* diagnostics very difficult. Direct numerical simulations (DNS) provide a unique tool to assess the spatially and temporally resolved flow and magnetic field.

In the frame of this project such simulations are performed. The complex geometry of rotating impellers is modelled via a penalisation or immersed boundary technique (IBM) and implemented in a massive parallel pseudo-spectral Navier-Stokes and magneto-hydrodynamics solver.

One major open problem of the VKS dynamo concerns the observation of an axial dipole magnetic field mode. This mode cannot be explained by a mean flow dynamo model. Two important ingredients have so far been identified: The magnetic boundary conditions of the impellers⁴ and the dynamics of a vortex generated in the vicinity of the impellers⁵. However up to now, only simplifying simulations driven by global or local volume forcing^{6,7} or simulations with an artificial assumption about the dynamo mechanism⁸ have been done. This is the first approach to the problem via DNS of a conducting fluid driven by moving impellers in conjunction with realistic magnetic field boundary conditions. These simulations serve to analyse directly (without further modelling) the resulting flow and magnetic fields. Especially, they allow to investigate the effects of magnetic material properties and the flow structures around the moving disk-blade impeller on the efficiency of the dynamo action.

2 Numerical Methods and Setup

The used code, named LaTu, is of pseudo-spectral type. It solves the three-dimensional incompressible Navier-Stokes and MHD equations and will now be explained in detail. In the MHD approximation considering solid objects with different magnetic properties an incompressible conducting flow is described by the following equations:

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} = \left(\nabla \times \frac{\mathbf{B}}{\mu} \right) \times \mathbf{B} - \nabla p + \mathbf{f}^u + \nu \nabla^2 \mathbf{u}, \quad (1)$$

$$\partial_t \mathbf{B} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \nabla \times \left(\eta \nabla \times \frac{\mathbf{B}}{\mu} \right), \quad (2)$$

$$\nabla \cdot \mathbf{u} = 0, \nabla \cdot \mathbf{B} = 0, \quad (3)$$

with the velocity field $\mathbf{u}(\mathbf{x}, t)$, the magnetic induction field $\mathbf{B}(\mathbf{x}, t)$, the pressure $p(\mathbf{x}, t)$ and the kinematic viscosity ν . The magnetic diffusivity (or electric resistivity) $\eta(\mathbf{x}, t)$ and the magnetic permeability $\mu(\mathbf{x}, t)$ are inhomogeneous in general to take into account solid objects with different material properties. Both the velocity field and the magnetic induction fulfil the incompressibility condition (Eq. 3). \mathbf{f}^u represents additional possible forces such as a penalty force in the immersed boundary technique discussed below. Setting $\mathbf{B} = 0$ gives the Navier-Stokes equations describing hydrodynamic flows. The equation system is solved with a standard Fourier pseudo-spectral method on a regular Cartesian grid. The time step is discretised with a third order Runge Kutta method.

The geometry of the simulations is flexible and is chosen closely to the VKS experimental setup - a cylindrical vessel, filled with liquid sodium, driven by two counter-rotating impellers. For benchmarking purposes we use a configuration resembling that of preceding water experiments. The periodic simulation box has a size of 2π along all three directions. The cylinder has a radius of $R_c = 3.0$ and a height of $H_c = 6.0$. For studies of the turbulent flow we choose an impeller setup close to the ‘‘TM28’’ curved blade impeller⁹. For the numerical investigation of the dynamo, including the magnetic induction field evolution, we switched to the ‘‘TM73’’ impeller¹⁰ setup, used in the final VKS experiment. For all simulations presented here we choose positive turning direction (the direction of convex curvature) and equal angular velocities Ω for both impellers.

The liquid is sticking on the cylinder and impeller surfaces so that the velocity field \mathbf{u} has to fulfil the no-slip boundary condition on boundary surfaces. To impose this condition

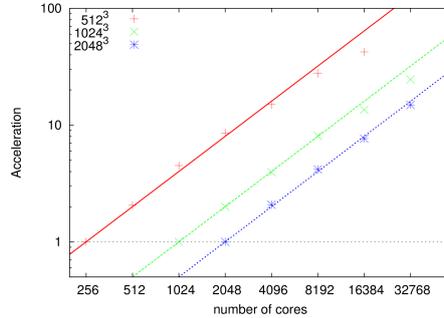


Figure 1. Scaling results of LaTu on the BlueGene/Q system JUQUEEN (NIC, Germany).

we make use of an immersed boundary technique. It consists first in introducing in the right-hand side of the momentum equation Eq. 1 a penalty force $\mathbf{f}^u = \mathbf{f}_b^u(\mathbf{x}, t)$, which acts as a Lagrange multiplier associated to the constraint defined by the boundary condition. In order to compute \mathbf{f}_b^u we make use of a direct forcing method introduced by Ref. 11 where we directly impose the velocity to the grid. This numerical scheme has already been benchmarked and used in the case of stationary and moving finite-size particles in hydrodynamic turbulent flows^{12,13}. For the magnetic induction we use periodic boundary conditions for simplicity, since it has turned out that the outer boundary condition has minor influence compared to material properties⁸. Boundary conditions at the rotating impellers are not imposed explicitly, instead the material is modelled via inhomogeneous material constants, which is known to be sufficiently accurate⁸.

LaTu operates in an entirely parallel manner via MPI communication and has been optimised for the BlueGene/Q system JUQUEEN. The code runs on more than 100k cores. Standard Fast-Fourier-Transform (FFT) libraries such as the FFTW¹⁴ distribute the computation domain over different processes by cutting the domain one-dimensionally into slices. With this strategy one is limited in the maximal number of processes by the largest number of grid-points in one of the three spatial directions. To overcome this restriction for massive parallel architectures such as the BlueGene/Q system, we make use of a parallel pencil-FFT (P3DFFT¹⁵), which divides the domain two-dimensionally into columns and which scales remarkably well. Scaling tests on JUQUEEN showed very good scaling properties up to 32768 cores (see Fig. 1). A further advantage of this P3DFFT-library is that it makes use of the platform-optimised Engineering Scientific Subroutine Library (ESSL)-library.

3 Flow Structures

As the first part of our work we performed a detailed DNS study of impeller driven hydrodynamic flows. There we chose a configuration very close to that of the VKS experiment, in particular the “TM28” impeller configuration⁹. This is the first quantitative comparison of experimental and numerical data on impeller driven turbulence. Here we will present selected results in order to highlight the novelty of our approach.

The typical flow structure is visualised in Fig. 2. We measured the mean flow characteristics and compared it quantitatively with the VKS experimental data for a very similar

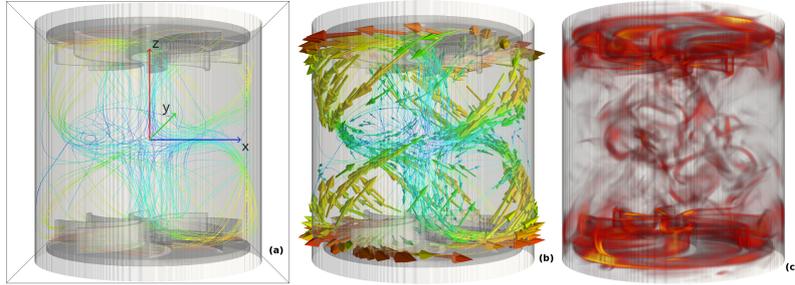


Figure 2. DNS data: (Left) configuration of the cubic computational domain showing the rotating impellers and the flow confining cylinder. (Middle) temporally averaged velocity profile. (Right) snapshot of volume rendering of the turbulent vorticity.

	TM28	256 ³	512 ³	512 ³ hi-Re
$u_{pol,mean}$	0.199	0.174	0.184	0.179
$u_{pol,max}$	0.492	0.425	0.443	0.460
$u_{tor,mean}$	0.281	0.205	0.217	0.217
$u_{tor,max}$	0.691	0.535	0.538	0.509
Γ_{mean}	0.71	0.850	0.847	0.825
Γ_{max}	0.71	0.794	0.824	0.904

Table 1. Quantities from experiment and the simulations: the maximum and the mean of the poloidal and toroidal velocity and the respective ratio $\Gamma_{max} = \frac{u_{pol,max}}{u_{tor,max}}$ and $\Gamma_{mean} = \frac{u_{pol,mean}}{u_{tor,mean}}$. All the velocities are normalised by the maximum velocity of the impellers $V_{max} = \Omega R_d$. A quantification of the poloidal and toroidal components is done by extracting the maximum and mean values in the bulk, in the region $-0.8R_c < z < 0.8R_c$.

geometry. The mean toroidal velocity component of the experimental and numerical data are strikingly similar. The mean quantities are summarised in Tab. 1. Having in mind that the Reynolds number of the experiment is much higher than that of the DNS the values are in very good agreement.

We also performed measurements which are very hard to obtain in experiments. One of them concerns the flow structure very close to the impellers. We found a characteristic outwards spiralling jet in between the blades (Fig. 3). This helicity producing structure might help to increase the dynamo efficiency of the configuration due to a so-called turbulent α effect, which together with the shear flow twists and stretches the magnetic field lines⁵. In this picture the shear flow converts poloidal magnetic energy to toroidal energy (the Ω effect), while the α effect converts toroidal to poloidal energy, closing the dynamo mechanism loop.

4 MHD Dynamo and Underlying Mechanism

For the dynamo study we switched from the TM28 impeller configuration to the so-called TM73 configuration, which is actually used in the successful VKS dynamo experiment¹⁰.

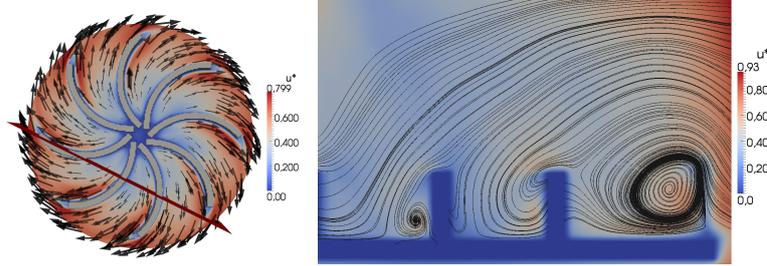


Figure 3. (Left) top view on an impeller in the co-rotating frame of reference. Arrow show the projected velocity onto the plane. The magnitude is given in colours. The slice indicates a second plane on which we give the projected velocity in the right figure. (Right) streamlines of the projected velocity in a plane indicated in the left figure. Colours give the magnitude of the velocity. Note the circular structures.

It is similar to the one previously used except for a smaller radius and a smaller curvature of the mounted blades. We studied in detail the impact of the permeability on global properties of the system. Increasing this material property in the simulation resembles the use of high permeability materials such as soft iron in the experiment. The analysis of local magnetic field structures by means of simulations at high permeability and high resolution lead us to a possible mechanism responsible for the success of the VKS experiment.

Fixing the conductivity of the impellers to $\sigma = 1$, only the kinetic and the magnetic Reynolds numbers Re and Rm and the magnetic permeability μ of the solid impellers remain as basic control parameters. We start all dynamo (MHD) simulations with an impeller-driven velocity field in the statistical stationary regime and a small random seed magnetic field. Under dynamo action, the magnetic energy increases exponentially, meaning $|\mathbf{B}|^2(t) \propto \exp(\gamma t)$ with growth rate γ , up to about 10% of the kinetic energy. Then, quickly the Lorentz force back reaction sets in and stabilises the energy on the saturated level. A first goal is the knowledge of the dynamo onset depending on Re and Rm for different μ . From the dynamo growth rates for different parameters, we found estimates of the dynamo threshold in terms of the magnetic Reynolds number $Rm_c(Re)$ typical for large-scale dynamos such as the Taylor-Green dynamo¹⁶. Starting from low kinetic Reynolds numbers, the likewise low critical magnetic Reynolds number increases, finally reaching a plateau around $Rm = 1000$ for $\mu = 1$ and around $Rm = 600$ for $\mu = 8$. We found that for fixed Reynolds numbers the dynamo growth rate γ also increases with the permeability.

4.1 Magnetic Field Shape

The time-averaged magnetic field for low permeability $\mu = 1$ at $Re = Rm = 506.25$ shows the shape found in previous kinematic simulations using the MND flow⁸ instead of a dynamically forced flow (Fig. 4 (left)). It is the so-called $m = 1$ mode of the field in cylindrical coordinates (r, φ, z) decomposed into

$$\mathbf{B}(r, \varphi, z) = \sum_{m=0}^{\infty} \mathbf{B}_m(r, z) \exp(im\varphi). \quad (4)$$

Increasing the permeability to $\mu = 16$ the shape of the magnetic field changes, showing a rotational symmetric $m = 0$ mode with toroidal structures concentrated around the im-



Figure 4. Isosurfaces of the mean magnetic energy $E = 1/2\mathbf{H} \cdot \mathbf{B}$ at about 10 percent of the peak value and field lines in the saturated regime for $Re = Rm = 506.25$ show a switch from the twisted equatorial dipole with $m = 1$ mode at $\mu = 1$ (left) to an axisymmetric field with $m = 0$ mode concentrated around the impellers at $\mu = 16$ (right).

pellers and a slight contribution from a twisted $m = 1$ mode in the bulk, which is similarly obtained in experiments (Fig. 4 (right)). The same behaviour is found as well for other sets of kinetic and magnetic Reynolds numbers. In previous simulations with stationary flows as well as in our simulations with low permeability most of the magnetic energy is contained in the $m = 1$ -mode, corresponding to an equatorial dipole field⁹. In contrast to this, the field in the successful experiment is an axial dipole field, mainly containing its energy in the rotational symmetric $m = 0$ mode¹⁷. An analysis of the interaction between fluid and solid and the mechanism behind the mode switch to explain the success of the VKS experiment is part of our work.

4.2 Dynamo Enhancement by Winding Mechanism

For a better understanding of the enhancement of the toroidal axisymmetric magnetic energy by the impellers we analysed the magnetic field structure near and inside the impellers. This is a significant advancement compared to experimental measurements, since these only provide information about the (reconstructed) large scale field¹⁷.

We find that the magnetic field lines are refracted at the transition from the fluid to the solid with high permeability (Fig. 5). Field lines approaching from the bulk are pushed downwards to the disk due to the interaction with the flow behind the blades. The refraction effect bends the field lines downwards again. The impeller disk serves as an additional attractor for field lines. Thus the high permeability of the impellers in addition with the specific geometry helps to “trap” field lines in the impeller region, winding them up and producing toroidal magnetic energy via an enhanced Ω effect.

We demonstrate that increasing the permeability reduces the dynamo threshold and additionally changes the first growing dynamo mode. A simulation with $\mu = 1$ ($Re = 1518$, $Rm = 911$) lies below the dynamo threshold for the $m = 1$ mode and hence shows no dynamo action. For $\mu = 8$ we find a dynamo with most magnetic energy contained in the $m = 0$ mode. Like in the experiment it is thus not necessary to start from an already self-amplifying $m = 1$ mode at low permeability to produce the $m = 0$ mode at high permeability.

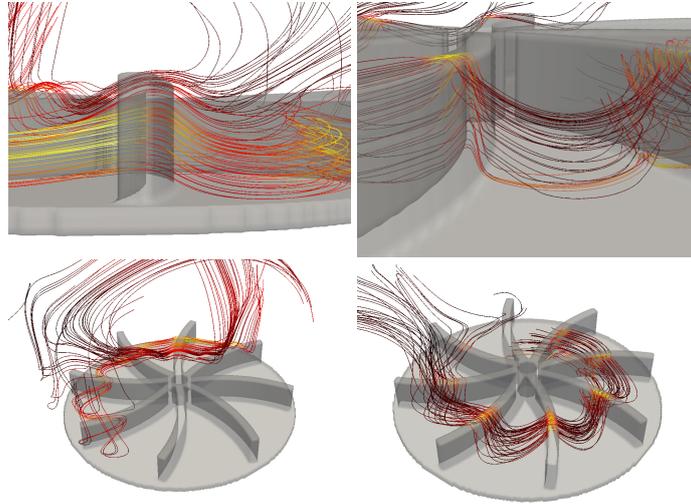


Figure 5. Magnetic field lines from instantaneous fields at $Re = Rm = 506.25$ show different behaviour at permeability $\mu = 1$ (left) and $\mu = 16$ (right). While at low permeability the field lines are only slightly bend by the fluid vortex rolls in between the blades, there is a stronger bending by a refraction effect of the blades at high permeability. This helps to attract field lines approaching from the bulk.

5 Concluding Remarks

Up to now our work has shown that pseudo-spectral solvers in combination with penalisation techniques provide a fast and reliable tool for the simulation of interacting solid objects with different magnetic properties and electrically conducting fluids. The specific implementation in framework LaTu is optimised for large computing clusters such as the BlueGene/G machine JUQUEEN and scales well up to 32k cores, allowing relatively high Reynolds numbers, reaching the turbulent regime or at least the transition regime to turbulence. In particular we performed simulations as close as possible to the Von Kármán Sodium experimental device, a successful liquid metal dynamo in the laboratory. Our investigations revealed possible mechanisms responsible for the success of the experiment, more precisely the vortex generation behind the impeller blades and the effect of field line trapping and winding due to the high permeability of the impeller material. Further detailed investigations are necessary to confirm these findings and to quantify the mechanisms.

Acknowledgements

We acknowledge fruitful discussions with Nicolas Plihon and Arnaud Chiffaudel. Parts of this research were supported by Research Unit FOR 1048, project B2, and the French Agence Nationale de la Recherche under grant ANR-11-BLAN-045, project SiCoMHD. Access to the IBM BlueGene/Q computer JUQUEEN at the Forschungszentrum Jülich was made available through the project HBO40.

References

1. A. Gailitis, O. Lielausis, S. Dementév, E. Platacis, A. Cifersons, G. Gerbeth, T. Gundrum, F. Stefani, M. Christen, H. Hanel, and G. Will, *Detection of a flow induced magnetic field eigenmode in the riga dynamo facility*, Phys. Rev. Lett. **84**:4365, 2000.
2. U. Müller and R. Stieglitz, *Can the earth's magnetic field be simulated in the laboratory?*, Naturwissenschaften **87**:381, 2000.
3. G. Verhille, N. Plihon, M. Bourgoïn, P. Odier, and J.-F. Pinton, *Induction in a von Kármán flow driven by ferromagnetic impellers*, New J. Phys. **12**(3):033006, 2010.
4. C. Gissinger, A. Iskakov, S. Fauve, and E. Dormy, *Effect of magnetic boundary conditions on the dynamo threshold of von Kármán swirling flows*, EPL **82**(2):29001, 2008.
5. F. Petrelis, N. Mordant, and S. Fauve, *On the magnetic fields generated by experimental dynamos*, Geophys. & Astrophys. Flu. Dyn. **101**, 289-323, 2007.
6. K. Reuter, F. Jenko, and C. B. Forest, *Hysteresis cycle in a turbulent, spherically bounded MHD dynamo model*, New J. Phys. **11**, 013027, 2009.
7. J. L. Guermond, J. Léorat, F. Luddens, C. Nore, and A. Ribeiro, *Effects of discontinuous magnetic permeability on magnetodynamic problems*, J. Comp. Phys. **230**(16):6299-6319, 2011.
8. A. Giesecke, F. Stefani, and G. Gerbeth, *Role of soft-iron impellers on the mode selection in the von Kármán-sodium dynamo experiment*, Phys. Rev. L. **104**(4):044503, 2010.
9. L. Marié, J. Burguete, F. Daviaud, and J. Léorat, *Numerical study of homogeneous dynamo based on experimental von Kármán type flows*, Eur. Phys. J. B **33**, 469ff, 2003.
10. F. Ravelet, *Bifurcations globales hydrodynamiques et magnetohydrodynamiques dans un écoulement de von Karman turbulent*, (Phd thesis), Ecole Polytechnique X, 2005, <http://tel.archives-ouvertes.fr/tel-00011016>.
11. E. A. Fadlun, R. Verzicco, P. Orlandi, and J. Mohd-Yusof, *Combined Immersed-Boundary Finite-Difference Methods for Three-Dimensional Complex Flow Simulations*, J. Comp. Phys. **161**, 35ff, 2000.
12. H. Homann and J. Bec, *Finite-size effects in the dynamics of neutrally buoyant particles in turbulent flow*, J. Fluid Mech. **651**:81-91, 2010.
13. H. Homann, J. Bec, and R. Grauer, *Effect of turbulent fluctuations on the drag and lift forces on a towed sphere and its boundary layer*, accepted for J. Fluid Mech., 2013.
14. M. Frigo, and S. G. Johnson, *FFTW*, 2012, <http://www.fftw.org>.
15. D. Pekurovsky, *P3dfft open source library*, San Diego Supercomputer Center, 2012, <http://www.sdsc.edu/us/resources/p3dfft/index.php>.
16. Y. Ponty, P. D. Mininni, D. C. Montgomery, J.-F. Pinton, H. Politano, and A. Pouquet, *Numerical Study of Dynamo Action at Low Magnetic Prandtl Numbers*, Phys. Rev. Lett. **94**, 2005.
17. J. Boisson, and B. Dubrulle, *Three-dimensional magnetic field reconstruction in the VKS experiment through Galerkin transforms*, New J. Phys. **13**:023037, 2011.

Highly-Resolved Numerical Simulation of the Turbulent Combustion Process in Experimental Burners

Fabian Proch, Martin Rieth, Andreas Rittler, and Andreas Kempf

Chair of Fluid Dynamics, Institute for Combustion and Gasdynamics (IVG),
University of Duisburg-Essen, 47057 Duisburg, Germany
E-mail: {fabian.proch, martin.rieth, andreas.rittler, andreas.kempf}@uni-due.de

This paper presents investigations of experimentally well-characterised turbulent flames with highly-resolved **L**arge **E**ddy Simulations (LES) and **D**irect Numerical Simulations (DNS). The combustion process is modelled with a flamelet-based approach, which assumes that the local turbulent flame structure can be described by an ensemble of wrinkled laminar flames. Good agreements between the simulation results and experimental measurement data is achieved. The governing equations are discretised with the **F**inite **V**olume **M**ethod (FVM). The numerical implementation is tailored for massively parallel simulations on a large number of grid cells. The computational efficiency benefits from the applied simple grid structure and the use of non-blocking **M**essage **P**assing **I**nterface (MPI) parallelisation.

1 Introduction

World energy consumption is mainly covered by burning fossil fuels such as natural gas, oil and coal but also sustainable fuels like biomass, biodiesel or biogas. As fossil fuels are limited and their combustion releases pollutants, their efficient usage is essential. To increase efficiency, a detailed understanding of the complex physical and chemical processes involved is required. Besides experiments, numerical analysis increasingly helps to understand combustion, exploiting the rapid increase of computational power in the past decades. As combustion in technical devices mostly occurs under turbulent conditions, the numerical investigation of the interplay between turbulent transport and chemical reactions is of high interest.

The governing equations for fluid flow and combustion are too complex to solve them directly for technical relevant combustors. To enable the simulation of such combustors, the governing equations are averaged in time or locally in space to reduce the computational costs. The averaging in time is termed **R**eynolds **A**veraged **N**avier **S**tokes (RANS) simulation, the filtering in space LES. These averaging operations lead to a loss of information on the smaller scales, which requires additional closure models. These closure models are developed based on theoretical considerations and on DNS of simpler geometries.

This work concentrates on highly resolved LES and DNS of more realistic geometries aiming to investigate and further develop existing combustion models. To avoid the excessive cost of the direct computation of a detailed chemical reaction mechanism, a tabulated chemistry approach is applied.

2 Numerical Approach

The investigations have been performed with the PsiPhi code. The code solves the implicitly filtered Navier Stokes equations for an incompressible flow of constant or variable

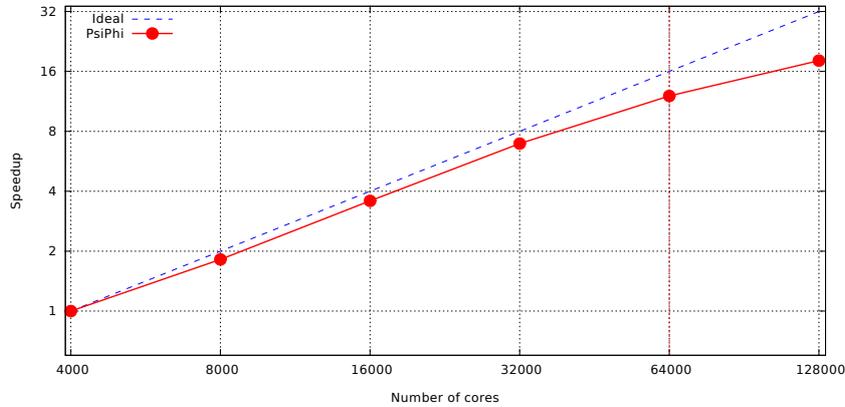


Figure 1. Strong scaling behaviour of PsiPhi on JUQUEEN. This test was performed with 1.6 Billion cells (i.e. 400,000 to 12,500 cells per core). The maximum number of cores used for the presented simulation results was 64,000.

density using a FVM. A Cartesian, equidistant grid is used, providing good numerical accuracy with isotropic filters, parallel efficiency and good vectorisation through avoiding slow, non-sequential memory access. Furthermore, this configuration allows to decompose the domain in blocks for maximum efficiency during communication between CPUs. Specific care has been devoted to efficient MPI parallelisation: PsiPhi is able to handle meshes with billions of cells, enabling highly resolved LES and DNS of laboratory scale configurations. To obtain a high parallel efficiency and performance of the code without the additional complexity of a hybrid approach, non-blocking MPI communication with overlaid computation is used. To assess the parallel efficiency, strong scaling tests have been performed on the JUQUEEN machine. The code performance has been tested on up to 128,000 cores, the respective results are shown in Fig. 1. The code scales well overall, for the maximum used number of 64,000 cores in the productive runs the parallel efficiency is around 70%.

PsiPhi ensures continuity by a pressure-correction scheme using a projection method. The transport equations are discretised with a 2nd-order central difference scheme in space and an explicit 3rd-order Runge-Kutta scheme in time.

The implemented chemical models enable the simulation of complex reacting flows. In the tabulated chemistry approach a small set of transport equations are solved for the conserved quantities, which describe the mixture composition and in addition the combustion progress. All thermochemical quantities are determined *a priori* and tabulated as a function of the control variables.

3 Quasi-DNS of a Bluff-Body Burner

The first presented case is a highly resolved simulation of an experimental bluff-body burner that has been investigated at the University of Cambridge and the Sandia National Labs^{1,2}. The burner features a central bluff body that is surrounded by two co-annular

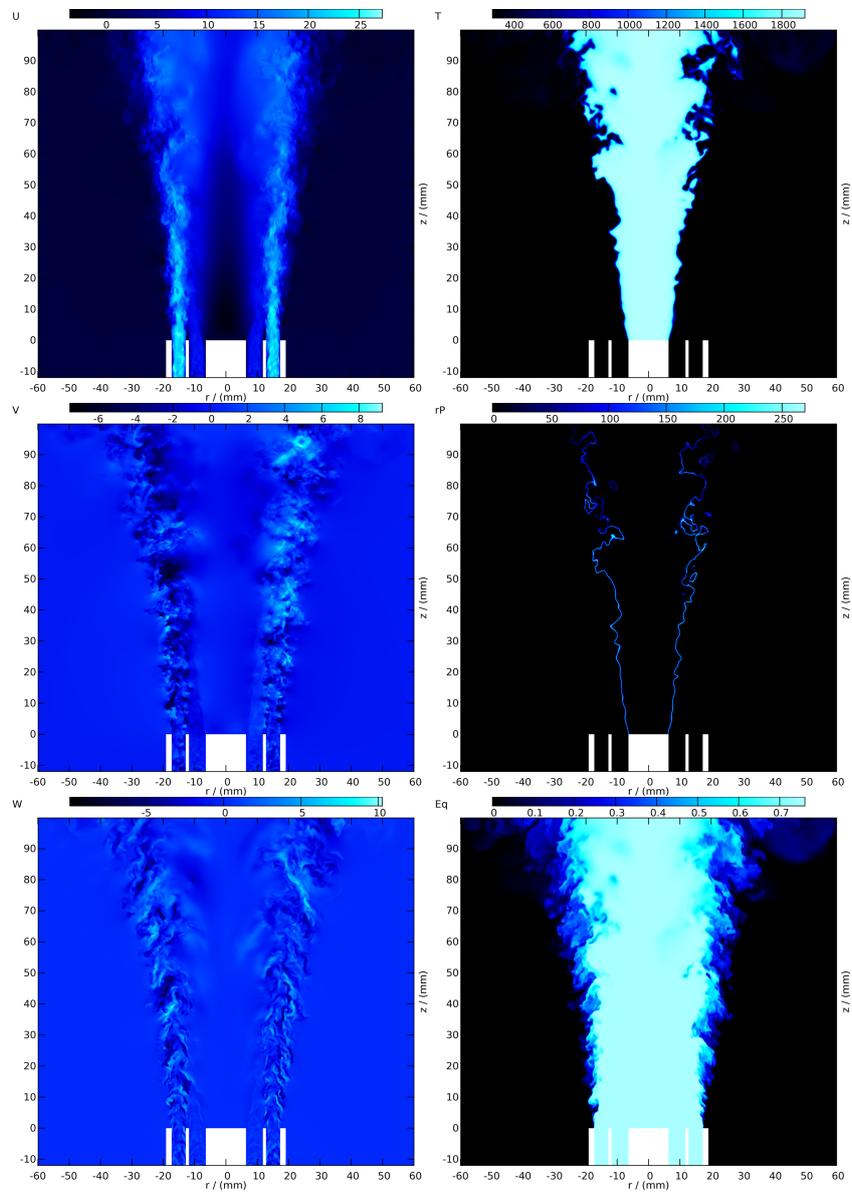


Figure 2. Contours of the velocity vector components (axial direction (U / (m/s)), radial direction (V (m/s)) and circumferential direction (W (m/s))), the temperature (T / (K)), the reaction source term (rP / ($\text{kgm}^{-3}\text{s}^{-1}$)) and the equivalence ratio (Eq).

streams of lean premixed methane/air mixture at an equivalence ratio of 0.75. The whole burner is operated at ambient conditions and embedded in a co-flow of air. The grid resolution has been set to $100\ \mu\text{m}$, which is fine enough to resolve the turbulent flame struc-

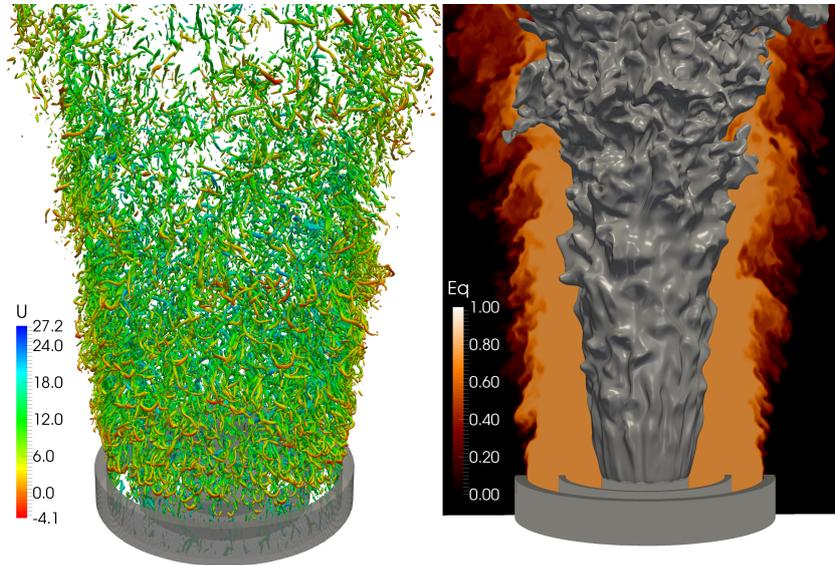


Figure 3. Left: Turbulent structures visualised by the Iso-q-criterion. The colouring corresponds to the axial velocity component (U / (m/s)). Right: Contours of equivalence ratio in the mid-section of the burner. The isosurface for 50% of reaction progress gives the flame location.

ture directly without any need for sub-filter modelling. This led to a total domain size of $1120 \times 1200 \times 1200 = 1.6$ Billion cells. The combustion modelling is done with the **Premixed Flamelet Generated Manifolds (PFGM)** tabulated chemistry approach. The work has been submitted to the *Journal of Fluid Mechanics* and is under review at the moment³. Fig. 2 shows contour plots in the burner mid-section for the velocity vector components (U , V , W), the temperature (T) and the reaction source term (rP). The recirculation zone in front of the bluff-body, which stabilises the flame, becomes clearly visible in the velocity plots. The velocity fluctuation levels vary in the inner and outer stream, and a large amount of the fluctuations is dissipated away in the flame region.

The equivalence ratio field is exposed to strong turbulent mixing due to the shear layer between the outer stream and the co-flow, the mixing layer is getting broader when moving downstream. In contrast, the flame thickness stays relatively constant, due to the balance between chemical reaction and diffusion. The reaction source term plot visualises the inner flame zone, it can be seen that the flame actually burns in two regimes: In the near burner region, the flame front is only mildly wrinkled and relatively laminar. In the region from approximately 35 mm downstream, the flame gets much more wrinkled due to the interaction with the turbulent fluctuations originating from the shear layer between the inner and outer stream. A statistical investigation and comparison to the Peters-Borghgi diagram (not shown here) have confirmed this observation.

Visualisations of the resulting three-dimensional flow- and flame structures are presented in Fig. 3. The left plot shows the iso-q-criterion, which is computed from the second invariant of the velocity gradient tensor and gives a good impression of the rotating structures present in the flow-field. A broad range of vortex structure sizes occurs, which

develop and interact when proceeding downstream. The right plot shows an iso-surface of the inner flame zone and a contour plot of the equivalence ratio. Again, the two flame regimes as discussed for Fig. 2 become visible. The three-dimensional structure of the flame surface develops from surface-wave patterns near the burner exit to distinct three-dimensional pockets far downstream.

The first and second statistical moments (means and variances), obtained from the simulations are compared against the experimental data. Additional correlations, spectra and probability distributions were evaluated to gain a deeper understanding of the process and reference data for the subsequent development of models that can be used in less expensive LES and RANS simulations which are utilised in industry.

4 Large Eddy Simulation of Reactive Multiphase Flows

Liquid fuels, with their many advantages, are utilised in relevant technical applications such as IC-Engines, gas turbines or oil-fired boilers. These devices feature highly turbulent flow structures to increase the mixing of fuel and oxidiser and subsequently increase the overall efficiency. The LES method is a suitable technique for the simulation of premixed and non-premixed flames as outlined by e.g. Janicka and Sadiki⁴, but it also showed the capability to describe reactive multiphase flows.

A well established test-case for the validation of numerical models is the Sydney piloted spray burner, which was investigated by Masri *et al.*⁵ at the University of Sydney. The diluted ethanol spray flame is supported by a concentric premixed pilot flame, which provides the heat for the evaporation of the liquid droplets and subsequently to ignite the partially premixed main flame.

An Eulerian and Lagrangian approach is used to describe the gaseous and liquid phase respectively. The numerical particles, which interact with the gaseous phase, are treated as lists of arrays, where each rank (or core) has a separate list to improve the numerical efficiency. In the present study it was possible to use one numerical particle to describe one liquid droplet. This aforementioned treatment allows to have a better description of the particle distribution in the finite volumes. The PFGM approach combined with the Artificial Thickened Flame method (PFGM/ATF) as described by e.g. Proch and Kempf⁶ was adopted to fulfil the requirements to model the flame in the presence of a diluted spray, details can be found in the studies by Rittler *et al.*⁷. Source terms in the governing equations for the momentum, the mixture fraction and the reaction progress variable are required to account for the evaporating droplets. The unresolved velocity fluctuations are determined with Nicouds Sigma model⁸, that yields the same accuracy as the Smagorinsky model with a dynamic procedure at much lower numerical costs.

As illustrated in Fig. 4 (a), the droplets evaporate due to the impact of the pilot flame and increase the mixture fraction further downstream, starting at $z = 50 - 70$ mm. Due to the higher slip velocity of the particles, that penetrate into the pilot stream from the jet stream, the evaporation is further amplified. The maximum mixture fraction is found at an axial location of $z = 120 - 200$ mm, downstream of this location the figures suggest that evaporation is completed. The spray flame has a lifted character at $z = 50 - 70$ mm, as outlined by the temperature and the OH mass fraction in Fig. 4 (b) and (c). Furthermore the OH mass fraction clearly indicates the flame wrinkling which results from the high turbulence level that increases mixing and the turbulent flame speed and reduces the flame

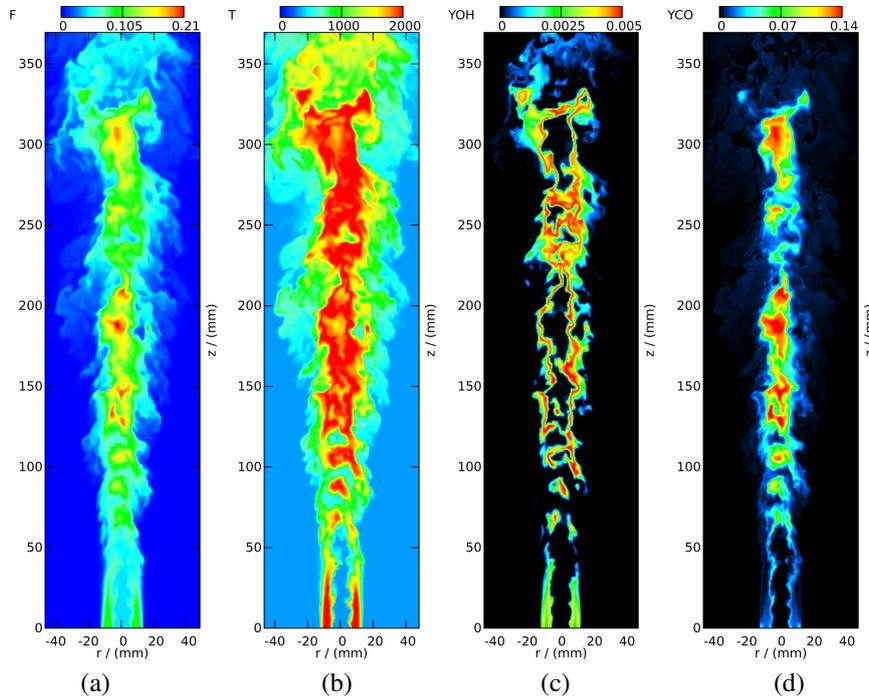


Figure 4. Contour plots of the simulation results for the (a) mixture fraction (F), (b) temperature (T / (K)), and the species mass fractions of (c) OH and (d) CO.

length and width. Besides this, OH indicates the flame front, which was found to be in good agreement with the experiments (not shown here). The applied model is also able to capture some toxic intermediate species like CO, as illustrated in Fig. 4 (d).

The modified PFGM/ATF approach is validated with the available data from the experiments - the first and second statistical moments for the particle properties and gas phase properties are in good agreement. Furthermore, a deeper understanding of the underlying processes in the spray flame could be obtained.

5 Massively Parallel Large Eddy Simulation of Pulverised Coal Combustion

Only since recently, computational power has become sufficient for the LES of realistic Pulverised Coal Combustion (PCC) furnaces. However, as there are still open question regarding the modelling for the LES of PCC, validation against experimental measurements is indispensable. These measurements can only be obtained from laboratory up to semi-industrial experiments. Among these is the experiment by Weber *et al.*⁹ conducted by the International Flame Research Foundation (IFRF). We chose this experiment to validate our models since it offers a good compromise between comprehensive measurements and an environment comparable to large-scale furnaces found in power plants. In the future LES

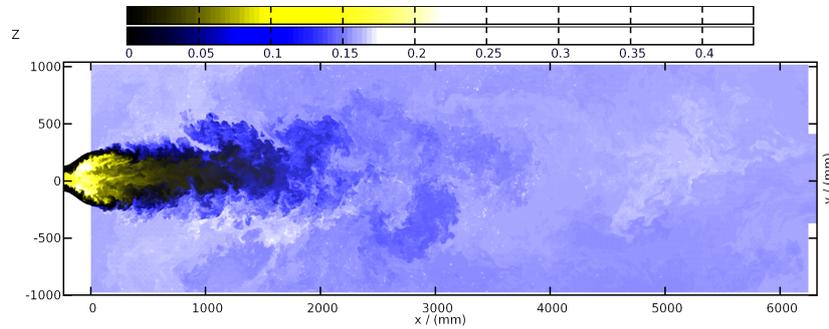


Figure 5. Instantaneous image of volatile gas mixture fraction (yellow scale) and char-off gas mixture fraction (blue scale) from the LES of the IFRF furnace.

of PCC promises to assist engineers and scientists in designing new and efficient furnaces and to obtain additional information that cannot be obtained by experiments.

The presented simulation relies on the non-premixed flamelet model to describe the gas phase. In the case of PCC, two mixture fractions have to be considered which represent: a) volatile gases released as the coal particles are heated up and b) the char-off gas that is released when oxygen diffuses into the porous coal structure during later stages of the coal combustion. Additional to the two mixture fractions, enthalpy and variance are required as parameters to describe the chemical state which leads to a four-dimensional pre-computed table that stores all the chemical information and is accessed during the computation.

An instantaneous image of the mixture fractions is shown in Fig. 5. It illustrates how volatile gases are released close to the inlet in the recirculation zone in the quarl region. Further downstream the burned volatile gases mix with combustion air promoting the slow char oxidation process through providing a hot mixture with oxygen excess.

Fig. 6 shows the CO₂ mass fraction. Intermediate CO₂ mass fractions can be found in the quarl region where mostly volatile gases burn. The CO₂ mass fraction first decreases further downstream as it is mixed with fresh combustion air. The highest CO₂ mass frac-

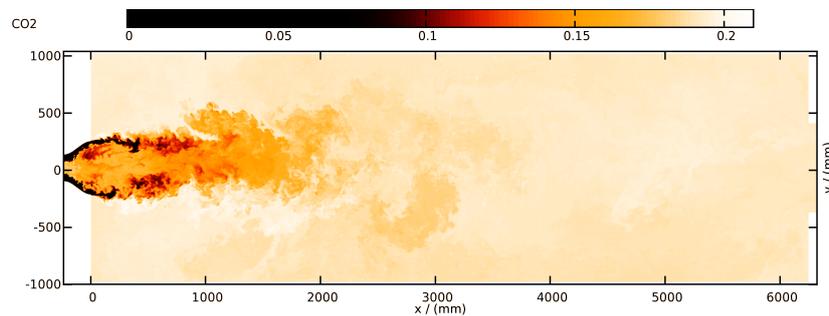


Figure 6. Instantaneous image of CO₂ mass fraction from the LES of the IFRF furnace.

tions are found downstream where char oxidation takes place.

The furnace has been discretised with 1.7 B cells and 40 M particles. The cells have an edge length of 2.5 mm and one numerical particle represents 100 real coal particles. Coarser simulations and the validation of the flamelet model for the LES of realistic PCC are currently under review¹⁰. The data from the simulations is compared to the experiment by means of statistical moments of species concentrations, temperature and velocities. The data is further analysed with a focus on instantaneous particle and gas phase quantities to gain additional insight into the combustion process and to identify possible model improvements.

Acknowledgements

We gratefully acknowledge the funding from the state of North-Rhine-Westphalia and the DFG and the compute time granted on JUQUEEN at Jülich Supercomputing Centre (JSC), through the John von Neumann Institute for Computing (NIC) as well as on SuperMUC at the LRZ and the CCSS CRAY at the University of Duisburg-Essen.

References

1. M. S. Sweeney, S. Hochgreb, M. J. Dunn, and R. S. Barlow, *The structure of turbulent stratified and premixed methane/air flames I: Non-swirling flows*, *Combust. Flame* **159**, 2896–2911, 2012.
2. R. Zhou, S. Balusamy, M. S. Sweeney, R. S. Barlow, and S. Hochgreb, *Flow field measurements of a series of turbulent premixed and stratified methane/air flames*, *Combust. Flame* **160**, 2017–2028, 2013.
3. F. Proch, P. Domingo, L. Vervisch, and A. M. Kempf, *Dynamics of the reaction zones in a premixed bluff-body burner: A quasi Direct Numerical Simulation study of a laboratory turbulent flame*, submitted to *J. Fluid Mech.*
4. J. Janicka, and A. Sadiki, *Large eddy simulation of turbulent combustion systems*, *Proc. Combust. Inst.* **30**, 537, 2005.
5. J. M. Gounder, A. Kourmatzis, and A. R. Masri, *Turbulent piloted dilute spray flames: Flow fields and droplet dynamics*, *Combust. Flame* **159**, 3371, 2012.
6. F. Proch, and A. M. Kempf, *Numerical analysis of the Cambridge stratified flame series using artificial thickened flame {LES} with tabulated premixed flame chemistry*, *Combust. Flame* **156**, 2627, 2014.
7. A. Rittler, F. Proch and A. M. Kempf, *LES of the Sydney piloted spray flame series with the PFGM/ATF approach and different sub-filter models*, *Combust. Flame* **162**, 1575, 2015.
8. F. Nicoud, B. H. Toda, O. Cabrit, S. Bose and J. Lee, *Using singular values to build a subgrid-scale model for large eddy simulations*, *Phys. Fluids* **23**, 085, 2011.
9. R. Weber, J. Dugué, A. Sayre, A. A. F. Peters, B. M. Visser, *Measurements and computations of swirl zone flow field and chemistry in a swirling pulverized coal flame*, *Int. Flame Res. Found. (IFRF)*, F36/y/20, 1992.
10. M. Rieth, F. Proch, M. Rabaçal, B. Franchetti, F. Cavallo Marincola, A. M. Kempf, *Flamelet LES of a semi-industrial pulverized coal furnace*, submitted to *Combust. Flame*.

Supercomputing Studies in Turbulent Rayleigh-Bénard Convection: Challenges and Perspectives

Mohammad S. Emran¹, Paul Götzfried¹, Anastasiya Kolchinskaya¹,
Janet D. Scheel², and Jörg Schumacher¹

¹ Institute of Thermodynamics and Fluid Mechanics, Technische Universität Ilmenau,
98684 Ilmenau, Germany

E-mail: joerg.schumacher@tu-ilmenau.de

² Department of Physics, Occidental College, Los Angeles, California, 90041, USA

Massively parallel supercomputations are an important analysis tool to study the fundamental local and global mechanisms of heat and momentum transfer in turbulent convection. Rayleigh-Bénard convection, which evolves in a fluid layer that is uniformly heated from below and cooled from above, is the simplest setting for a buoyancy-driven turbulent flow and thus a paradigm for many turbulent flows in nature and technology. We discuss two topics of this vital field of fundamental turbulence research – large-scale pattern formation in the turbulent regime and convection at very low Prandtl numbers.

1 Introduction

Turbulent convection is an important area of present research in fluid dynamics with applications to diverse phenomena in nature and technology. They reach from chip cooling devices, heat exchangers in power plants and energy efficient indoor ventilation via convection in the Earth's atmosphere, core and oceans to convection in the Sun and other stars. Often turbulent convection is then combined with other physical processes such as rotation, electromagnetic fields or phase changes. The turbulent Rayleigh-Bénard convection (RBC) model is at the core of all these turbulent flows. It can be studied in a controlled manner, but has enough complexity to contain the key features of turbulence in heated fluids. This flow in cylindrical cells has been investigated intensively over the last few years in several laboratory experiments all over the world^{1,4}. In a RBC, a fluid cell or layer is kept at a constant temperature difference $\Delta\theta = \theta_{bottom} - \theta_{top}$ between top and bottom plates which are separated by a vertical distance H . The dimensionless Rayleigh number, $Ra = CH^3\Delta\theta$, characterises the thermal driving in convective turbulence. Here $C = g\alpha/(\nu\kappa)$ with the acceleration due to gravity, g , the thermal expansion coefficient, α , the kinematic viscosity, ν , and the thermal diffusivity, κ . As the Rayleigh number keeps increasing past a critical value of 1708, a buoyancy-driven instability causes the conducting state to bifurcate via a convection state consisting of straight, parallel rolls and spatiotemporal chaotic states to an eventually fully turbulent flow. The hard turbulence regime in RBC is established for $Ra \gtrsim 10^6$. The Prandtl number $Pr = \nu/\kappa$ and the aspect ratio Γ , which is the ratio of cell diameter or cell length and cell height H , are the two other input parameters.

One of the key questions in RBC is that of the turbulent transport mechanisms of heat and momentum: How does the turbulent fluid carry heat and momentum from the heated plate at the bottom to the cold plate at the top^{1,4}? In the following, we want to discuss two points which are related to this key question: the origin of large-scale flow patterns

in fully turbulent convection in extended layers and the specific turbulence properties of convection at very low Prandtl numbers. We address both points by high-resolution direct numerical simulations (DNS) in which all turbulent scales starting from the system size down to the viscous and diffusive scales which are resolved without any parametrisation or model.

The three-dimensional Boussinesq equations Eqs. 1–3 are used to model Rayleigh-Bénard convection. The first of these is the Navier-Stokes equation. The simplifying assumption is made that the density only varies linearly with the temperature in the buoyancy term. In addition we must include the heat diffusion equation and the incompressibility condition to complete our system of (dimensionless) equations:

$$\nabla \cdot \mathbf{u} = 0, \quad (1)$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \sqrt{\frac{Pr}{Ra}} \nabla^2 \mathbf{u} + \theta \mathbf{e}_z, \quad (2)$$

$$\frac{\partial \theta}{\partial t} + (\mathbf{u} \cdot \nabla) \theta = \frac{1}{\sqrt{RaPr}} \nabla^2 \theta. \quad (3)$$

The variable $\mathbf{u}(\mathbf{r}, t)$ is the velocity field, $p(\mathbf{r}, t)$ is the pressure, and $\theta(\mathbf{r}, t)$ is the temperature. We use no-slip velocity boundary conditions along all walls. The temperature field obeys insulating sidewalls and constant values along the top and bottom plates. Alternatively to the Ra – Pr form of Eqs. 1–3, it is possible to write the Boussinesq equations with the Grashof number $Gr = Ra/Pr$ and Pr . We come back to this point in Sec. 3.

A *spectral element method (SEM)* is applied which is based on the Nek5000 software package¹¹. The order of the Lagrangian interpolation polynomials which is used in each space direction is as high as $N - 1 = 13$. This code scales in our application very well up to $\sim 3 \times 10^5$ cores. Numerical details and comprehensive resolution tests are found in Scheel *et al.*¹². The SEM is preferred if the fine-scale structure, i.e. gradients of the turbulent fields have to be analysed. A *second-order finite difference method (FDM)* for cylindrical^{3,2} or Cartesian¹⁰ nonuniform and staggered grids is additionally applied. This method has been used for the numerical investigations on the large-scale patterns which we will discuss in the subsequent section. In the cylindrical case, we can apply a two-dimensional Poisson solver after a one-dimensional fast Fourier transformation is applied in the azimuthal direction.

2 Large-Scale Circulation Patterns as Turbulent Superstructures of Convection

The classical picture is that turbulent fluid motion is characterised by a cascade of vortices and swirls of different sizes that give rise to a featureless and stochastic fluid motion. Our daily experience shows, however, that turbulent flows in nature and technology are often organised in prominent large-scale and long-living structures, which are termed *turbulent superstructures*. In convection, turbulent superstructures are mean circulation patterns whose coherence does not stop at the natural scale of the height H , but extends over much larger scales. When present, these turbulent superstructures dominate the global transport of heat and momentum. One way to answer the yet open question on their dynamical

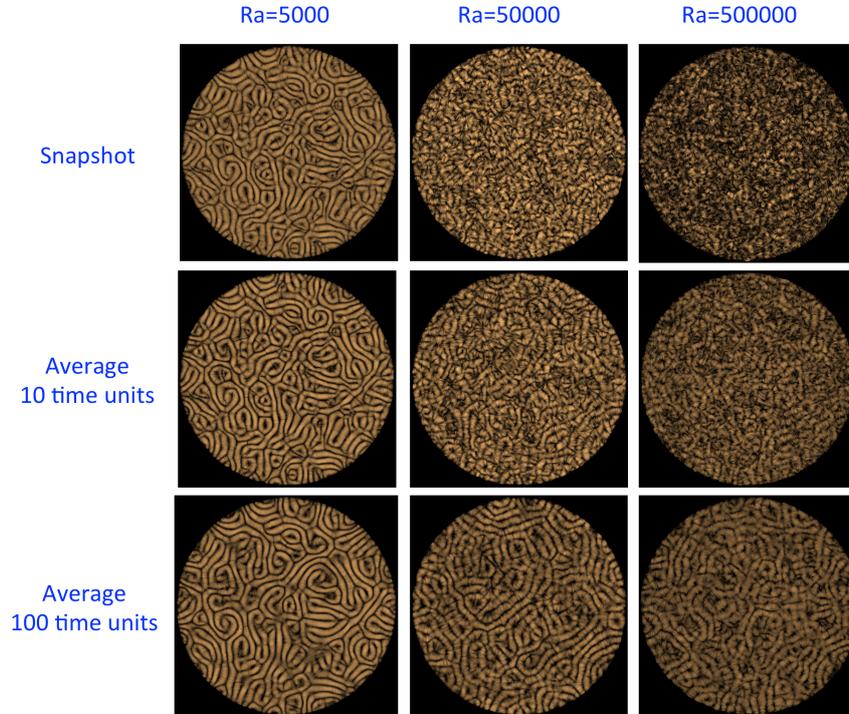


Figure 1. Top view on the velocity field patterns for cylindrical convection cells with an aspect ratio of $\Gamma = 50$. We compare a snapshot and two time averages over 10 and 100 time units T . All runs are at $Pr = 0.7$. Note that the patterns in the upper left and lower right panels of this figure almost agree.

origin is to monitor the structure formation in large aspect ratio cells starting at moderate Rayleigh numbers.

Our simulations showed that large-scale flow patterns, which are well-known from the spiral defect chaos regime of thermal convection at Rayleigh numbers $Ra < 10^4$, continue to exist in our three-dimensional numerical simulations of turbulent Rayleigh-Bénard convection in cylindrical cells with $\Gamma = 50$, $Pr = 0.7$ and $Ra \leq 5 \times 10^5$. In Fig. 1 we display the streamlines of the velocity field viewed from the top for three different Rayleigh numbers and compare instantaneous flow snapshots and time averages which have been obtained for two different intervals. At $Ra = 5000$ (left column) the convection flow is in the weakly nonlinear regime. The streamlines show that rolls are formed between the plates in a pattern of spirals and defects. Differences between a single snapshot and time averages are very small. For Rayleigh numbers $Ra = 50\,000$ (mid column) and $500\,000$ (right column) the magnitude of turbulent fluctuations exceeds the magnitude of the mean flow. Flow snapshots appear rather featureless. When the turbulent fields are averaged in time and turbulent fluctuations are thus removed, regular patterns are revealed, which we term turbulent superstructures of the convection flow⁷. On the one hand, the averaging time should be long enough such that the fluctuations are suppressed. On the other hand, this time should be smaller than the time scale on which such a large-scale

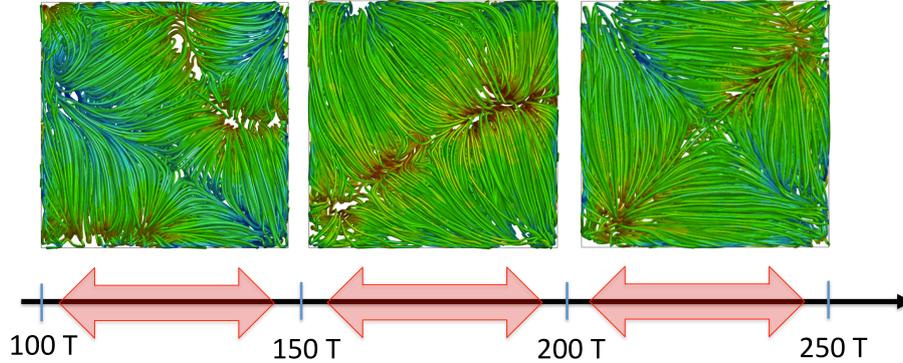


Figure 2. Top view on the mean velocity field patterns for a rectangular convection cell with $\Gamma = 3$. We compare the time averages which have been obtained over three successive intervals of 50 time units T . The run is at $Ra = 10^7$ and $Pr = 0.7$. The stream ribbons are coloured with respect to the vertical velocity (red for upwelling, green for almost zero and blue for downwelling).

pattern evolves which typically grows with $\sim \Gamma^2$.

Although the Rayleigh numbers have been moderate, the simulations are very demanding since the numerical effort grows at least with Γ^2 . Our analysis proceeds in two directions: first, we study how the mean flow structure changes as the Rayleigh number is increased. These simulations have to be conducted in smaller aspect ratio cells. The numerical effort is rapidly increasing as can be seen by the DNS from Ref. 13 for $Ra = 10^8$, for example. There, we compared a run at $\Gamma = 1$ with a single large-scale circulation roll with a run at $\Gamma = 3$. The latter run has just two large-scale circulations roll, i.e., we are still far away from extended patterns as in Fig. 1. The run at $\Gamma = 1$ required $N_e = 256\,000$ spectral elements, a polynomial expansion on each cell of $N = 12$ and thus a total number of mesh cells of $N_e N^3 \approx 4.4 \times 10^8$. The run at $\Gamma = 3$ required $N_e = 2\,304\,000$, $N = 10$ and $N_e N^3 \approx 2.3 \times 10^9$. The latter run is almost 9 times bigger as the $\Gamma = 1$ run and required 32768 MPI tasks on Blue Gene/Q. These studies revealed that the local turbulent heat transport depends sensitively on the shape of the superstructures in RBC.

Secondly, rectangular cells are investigated in order to study how the formation of turbulent superstructures is influenced by the different geometry. While turbulent convection in a cylindrical cell obeys homogeneity of statistical moments in the azimuthal direction, a rectangular cell has a preferred direction which is along the diagonal. First simulation results are displayed in Fig. 2 for a Rayleigh number of 10^7 . Mean patterns which have been obtained over disjoint time intervals switch between the two diagonals of the cell as visible by comparison of left and mid panels.

3 Turbulent Convection at Very Low Prandtl Numbers

Compared to the vast number of investigations at $Pr \gtrsim 1$, the very-low- Pr regime appears almost as a “terra incognita” despite many applications. Turbulent convection in the sun is present at Prandtl numbers $Pr < 10^{-3}$, in the liquid metal core of the earth one finds $Pr \sim 10^{-2}$. Convection flows as present in material processing, nuclear engineering, or

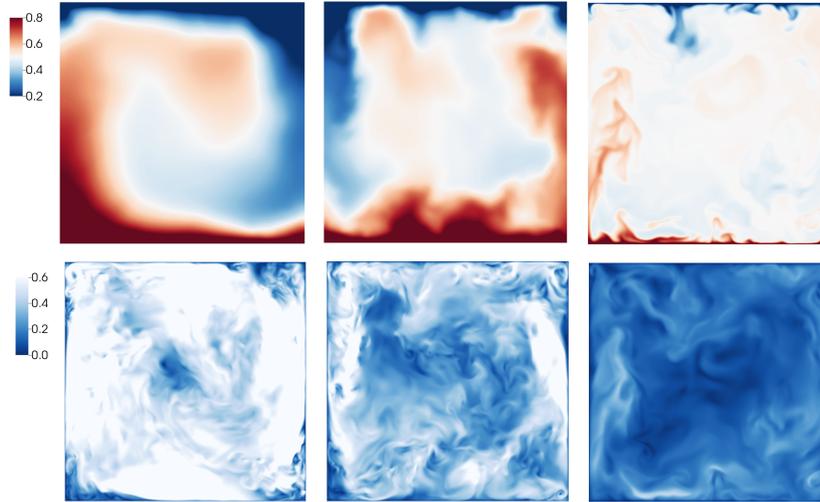


Figure 3. Vertical instantaneous cuts for the temperature field T (top row) and the magnitude of the velocity field $|\mathbf{u}|$ (bottom row). Left: $Ra = 2.38 \times 10^6$ and $Pr = 0.005$ for liquid sodium. Mid: $Ra = 10^7$ and $Pr = 0.021$ for liquid mercury. Right: $Ra = 3.33 \times 10^8$ and $Pr = 0.7$ for convection in air. All data are obtained at the same Grashof number $Gr = 4.76 \times 10^8$. The corresponding colour scales are the same.

liquid metal batteries have Prandtl numbers between 5×10^{-3} and 2×10^{-2} . One reason for significantly fewer studies of convection at low Prandtl numbers is that laboratory measurements have to be conducted in opaque liquid metals, partly operated at temperatures significantly larger than room temperature. The only way to access the statistics in these turbulent flows is via ultrasound or X-ray probes. Examples are mercury⁵ or gallium at $Pr \approx 0.02 - 0.025$ and liquid sodium⁸ at $Pr \approx 0.005 - 0.009$. Liquid sodium provides the lowest Prandtl number which is accessible in controlled conditions of a laboratory experiment.

Direct numerical simulations (DNS) are currently the only way to gain access to the full three-dimensional convective turbulent fields in low- Pr convection and to assure the accuracy of the boundary conditions. These simulations turn out, however, to become very demanding when the small-scale structure of turbulence has to be studied, even for moderate Rayleigh numbers Ra . To give an example: a simulation for convection in liquid sodium at $Ra = 2.38 \times 10^6$ required the same SEM mesh of 4.1×10^9 cells as a simulation for convection in air at $Ra = 10^{10}$, i.e. at a Rayleigh number that is almost four orders of magnitude larger than in sodium. Both DNS runs were conducted on 65 536 MPI tasks on BG/Q.

While the heat transport is reduced in low- Pr convection, the production of vorticity and shear are enhanced significantly. We found that this in turn amplifies the small-scale intermittency in these flows and makes them better comparable to classical Kolmogorov turbulence¹⁴. This result has important consequences in view to parametrisations for the small-scale turbulence in low-Prandtl-number fluids which are necessary in many applications such as in nuclear engineering or astrophysics. In Ref. 14 it was suggested that

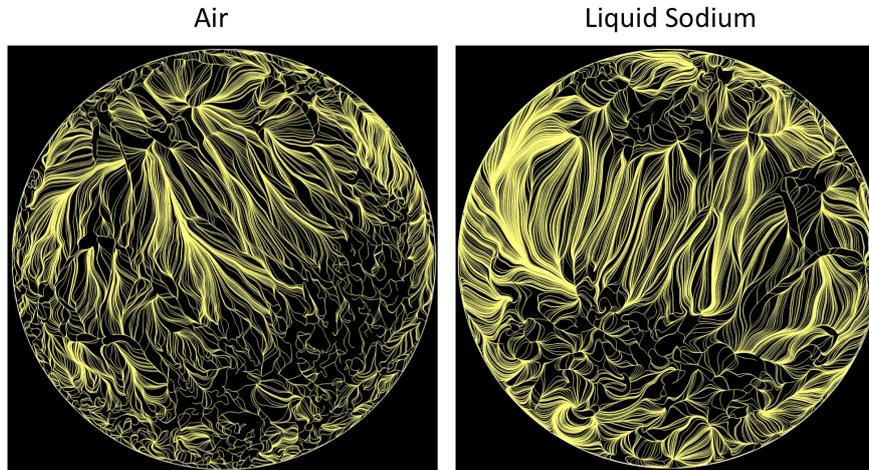


Figure 4. Field line snapshot of the skin friction field $\mathbf{s}(x, y, z = 0, t_0)$ taken at the bottom plate of the convection cell. Left: Air at $Ra = 10^{10}$ and $Pr = 0.7$. Right: Liquid sodium at $Ra = 2.38 \times 10^6$ and $Pr = 0.005$. In both figures we observe a large area fraction of the skin friction field in which the field lines are aligned.

a comparison of the fluid turbulence at different Prandtl numbers should be conducted at the same Grashof number $Gr = Ra/Pr$ rather than at a same Rayleigh number. This can be motivated by an inspection of the dimensionless momentum equation Eq. 2. If Gr is constant for different runs then Eq. 2 remains unchanged. In Fig. 3, we display vertical snapshot cuts of the temperature and the velocity magnitude for three simulations at the same Grashof number. Note that a constant Grashof number implies both, a change of Pr and Ra . With decreasing Prandtl number, the temperature field becomes increasingly washed out due to the enhanced thermal diffusivity while the velocity field is more vigorous thus indicating the enhanced level of fluid turbulence. We have shown that the cascade of the fluid turbulence is extended at the large-scale and small-scale end with decreasing Prandtl number thus enhancing the flow Reynolds number and consequently the turbulent momentum transport in the convection flow. The more vigorous fluid turbulence enhances in the enstrophy production in the fluid.

Since the Rayleigh-Bénard fluid is confined between rigid impermeable walls, tiny boundary layers of the temperature and velocity fields form in the vicinity of the walls. All the heat that is supplied by the boundary conditions has to pass these boundary layers. They form a bottleneck that constrains the global heat transport. How can this bottleneck be widened? One way is to enhance thermal diffusivity in relation to the viscosity, as done in low- Pr fluids. This certainly increases the width of the thermal boundary layer, but reduces the magnitude of the temperature gradients and thus the overall turbulent heat transfer at the same time. An alternative way is achieved when the boundary layers of velocity and temperature become fully turbulent. This regime of convective turbulence is called the *ultimate regime* of convection and goes back to a seminal paper from 1962 by Kraichnan⁹. The existence and the Rayleigh number range at which this transition into the ultimate regime proceeds is currently intensively discussed in the community.

It can however be expected that if the transition to fully turbulent boundary layers exists then the transition Rayleigh number range will depend on the Prandtl number. Our recent high-resolution DNS in the very low- Pr regime provide first hints although the accessible Rayleigh numbers are still too small. In Fig. 4 we display the velocity gradient at the heated plate at $z = 0$. The original nine-component velocity gradient tensor $A_{ij} = \partial u_i / \partial x_j$ is reduced to a two-component vector field $\mathbf{s} = (A_{13}, A_{23})$ at the wall due to the no-slip boundary conditions. Stream line snapshots of this *skin friction field* are shown for the runs at $Ra = 2.38 \times 10^6$ and 10^{10} that were mentioned above. The complexity of both instantaneous velocity boundary layers is by visual inspection very similar: in a larger area fraction the field lines are nearly parallel very similar to the skin friction field in turbulent channel flows⁶. Both velocity boundary layers display area fractions that are turbulent which we have determined by quantitative measures, such as local friction coefficients or local shear Reynolds numbers. For the very low Prandtl number run this turbulence transition should appear at a much smaller Rayleigh number. Thus the level of fluid turbulence is not only enhanced in the bulk¹⁴ but also in the boundary layer when Pr decreases.

4 Perspectives

We have discussed two particular topics in turbulent Rayleigh-Bénard convection which are being addressed by supercomputations:

Origin of turbulent superstructures in convection: Our DNS showed that the mean flow patterns from the weakly nonlinear regime are found to continue to exist when fluctuations in the fully turbulent regime are removed. The patterns thus remain dynamically relevant. As we indicated, one open question is how the turbulent large-scale patterns will vary when the Rayleigh number is increased or the Prandtl number is changed. First studies in this direction in smaller cells have been reported in Ref. 2 and imply further modifications as Ra grows. Such transitions to new mean flow patterns could be analysed by bifurcation analysis, and modelled as phase transitions in the turbulent flow. This suggests one could apply a reduced amplitude equation formalism to describe the evolution of the large-scale patterns in the horizontal directions which are obtained by a decomposition into coarse and fine spatial scales as well as into slow and fast time scales. Computationally, a turbulent convection cell with $\Gamma = 15$ at $Pr = 0.7$ and $Ra = 10^9$ would require at least 57.6 million spectral elements (which are already feasible today with Nek5000) and thus about 10^{11} mesh cells. This requires at least 60 BG/Q racks for production jobs.

Bulk and boundary layer dynamics in liquid metal convection: We showed that the highly diffusive temperature field with coarse thermal plume structures is an efficient driver of fluid turbulence, both, in the bulk and in the boundary layers. Our DNS suggest therefore that the transition to turbulence in the boundary layers of liquid metal convection is triggered at lower Rayleigh numbers than for convection in air. Another open question is if this transition follows the same rules as in wall-bounded turbulent shear flows or if the temperature field alters the transition mechanisms. Simulations are the only approach to answer these questions since they can access the gradients of the fields in the vicinity of the walls. It is also clear that the Rayleigh number has to be increased further. For liquid gallium ($Pr = 0.021$), based on an extrapolation from our currently largest DNS at $Ra = 10^8$, we would need a mesh with 10^{10} cells for a high-resolution run at $Ra = 5 \times 10^8$.

Both topics demonstrate clearly that supercomputing has become an essential tool to reveal the secrets of turbulence, and especially those of turbulent convection.

Acknowledgements

The work is supported by the Deutsche Forschungsgemeinschaft within the Research Unit 1182, the Research Training Group 1567 and project SCHU 1410/18-1. We want to thank the John von Neumann Institute for Computing for the steady support with supercomputing resources that make this work possible, in particular by projects HIL07, HIL08 and HIL09. Further support was received by the PRACE Initiative and the DOE Incite program.

References

1. G. Ahlers, S. Grossmann, and D. Lohse, *Heat transfer and large scale dynamics in turbulent Rayleigh-Bénard convection*, Rev. Mod. Phys. **81**, 503–537, 2009.
2. J. Bailon-Cuba, M. S. Emran, and J. Schumacher, *Aspect ratio dependence of heat transfer and large-scale flow in turbulent convection*, J. Fluid Mech. **655**, 152–173, 2010.
3. R. Camussi and R. Verzicco, *Numerical experiments on strongly turbulent thermal convection in a slender cylindrical cell*, J. Fluid Mech. **477**, 19–49, 2003.
4. F. Chillà and J. Schumacher, *New perspectives in turbulent Rayleigh-Bénard convection*, Eur. J. Phys. E **35**, no. 58, 2012.
5. S. Cioni, S. Ciliberto, and J. Sommeria, *Strongly turbulent Rayleigh-Bénard convection in mercury: comparison with results at moderate Prandtl number*, J. Fluid Mech. **335**, 111–140, 1997.
6. M. S. Chong, J. P. Monty, C. Chin, and I. Marusic, *The topology of skin friction and surface vorticity fields in wall-bounded flows*, J. Turb. **13**, no. 6, 2012.
7. M. S. Emran and J. Schumacher, *Large-scale patterns in extended turbulent convection*, J. Fluid Mech. **776**, 96–108, 2015.
8. P. Frick, R. Khalilov, I. Kolesnichenko, A. Mamykin, V. Pakholkov, A. Pavlinov und S. Rogozhkin, *Turbulent convective heat transfer in a long cylinder with liquid sodium*, Europhys. Lett. **109**, 14002, 2015.
9. R. H. Kraichnan, *Turbulent thermal convection at arbitrary Prandtl number*, Phys. Fluids **5**, 1374–1389, 1962.
10. D. Krasnov, O. Zikanov, and T. Boeck, *Comparative study of finite difference approaches in simulation of magnetohydrodynamic turbulence at low magnetic Reynolds number*, Comput. Fluids **50**, 46–59, 2011.
11. <https://www.nek5000.mcs.anl.gov/>.
12. J. D. Scheel, M. S. Emran, and J. Schumacher, *Resolving the fine-scale structure in turbulent Rayleigh-Bénard convection*, New J. Phys. **15**, 113063, 2013.
13. J. D. Scheel and J. Schumacher, *Local boundary layer scales in turbulent Rayleigh-Bénard convection*, J. Fluid Mech. **758**, 344–373, 2014.
14. J. Schumacher, P. Götzfried, and J. D. Scheel, *Enhanced enstrophy generation for turbulent convection in low-Prandtl number fluids*, Proc. Natl. Acad. Sci. USA **112**, 9530–9535, 2015.

On the Parametrisation of Lattice Boltzmann Method in Pore-Scale Flow Simulations

Siarhei Khirevich^{1,2} and Ulrich Tallarek¹

¹ Department of Chemistry, Philipps-Universität Marburg, 35032 Marburg, Germany
E-mail: {khirevic, tallarek}@staff.uni-marburg.de

² Upstream Petroleum Engineering Research Center, KAUST,
Thuwal 23955-6900, Kingdom of Saudi Arabia
E-mail: siarhei.khirevich@kaust.edu.sa

We analyse impact of parametrisation in lattice Boltzmann simulations of flow in complex geometries. For the input geometries we use four sets of regularly and irregularly packed spheres (“packings”) with known accurate solution for the permeability or drag. All four geometries have porosity equal to 0.366 but different microstructure resulting in their different permeability values. We vary spatial resolution in the range between 5 and 750 lattice nodes per sphere diameter, observe different behaviour of the numerical error for several resolution sub-ranges and address them in detail providing practical guidelines for increasing accuracy in low-resolution simulations, which are typical for practical problems.

1 Introduction

Transport processes (such as flow or diffusion) in porous media occur in many diverse fields of science and engineering, and their accurate prediction and optimisation requires understanding, both qualitative and quantitative, of the underlying physical phenomena. Slow (Stokes) flow of a viscous fluid is one of such key processes, and computer simulations nowadays are promising tools for its study. Among others, the lattice Boltzmann method (LBM) became a powerful computer simulation approach for, in particular, simulations of flows in complex geometries³.

LBM is based on evolution of a “lattice gas”, where classical gas model representation is replaced by the one with discrete space and time: individual molecules are combined into fluxes which do exist at discrete lattice nodes and move along prescribed discrete links connecting each node with its neighbours. During discrete time steps fluxes do propagate between lattice nodes and collide at them. Output of the collision step as well as the simulation result depend on a particular selection of LBM adjustable parameters (relaxation rates).

Despite more than two decades of the method history, there is still a controversy about selection of its parameters even for such a “simple” underlying physical phenomenon as Stokes flow of incompressible isothermal fluids, when we consider complex geometries like porous media. In this study we approach this problem using for the input geometry packing of impermeable spheres densely packed in regular and irregular fashions. We simulate Stokes flow in the voids between spheres, and assess method accuracy against its parametrisation.

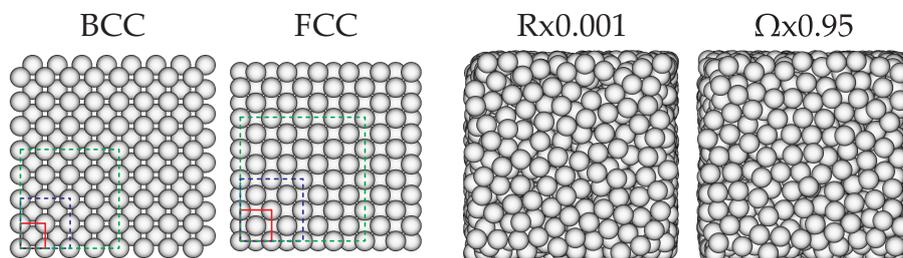


Figure 1. Simulation geometries: sets of spheres (packings) of equal size where spheres are arranged in regular (BCC, FCC) and irregular (Rx0.001, Ωx0.95) fashions; all four packings have equal void space fraction of $\varepsilon = 0.366$. Irregular packings contain about 7000 spheres each.

2 Numerical Simulation Approach

2.1 Input Geometry

We create assemblies of fixed in space, impermeable spheres arranged in regular and irregular (random) fashions, considering two packings for each one. All four packings are spatially periodic and have equal porosity of $\varepsilon = 0.366$, the value close to the ones occurring in many real systems. We intentionally fix porosity to address the impact of heterogeneity on the simulation accuracy while it is assumed that irregular geometries may have lower numerical errors due to “error cancellation” effects. Coordinates of spheres in the regular packings are calculated analytically: first we create standard face-centred and body-centred packings of touching spheres with porosities of ≈ 0.26 and ≈ 0.32 , respectively, and then shrink the sphere radii to obtain the target porosity of 0.366. Irregular packings were generated using Jodrey–Tory² and Monte-Carlo algorithms which are geometrical approaches to distribute spheres in space in irregular fashion while achieving lower porosities ($\varepsilon < 0.40$). These four packings allowed us to create geometries – comprised of identical objects – with equal average porosity but different microstructure and different values of permeability/drag.

2.2 Discretisation Procedure

One of the key differences of this study from previous works is the discretisation procedure. Standard LBM approaches operate with uniform cubic meshes, and to simulate flow in a given packing its geometry is mapped onto the corresponding domain with cubic mesh marking each mesh node as “solid” or “fluid”. In the case of a packing with small amount of spheres (say, less than 10) and low discretisation resolution d_{sp} (≈ 10 or fewer lattice nodes per sphere diameter) the change in the simulated permeability/drag value during step-by-step increase of the domain dimensions can be quite large¹. In this study we replicate considered geometries using their periodicity property and then discretise it. After proper selection of the domain dimensions such approach allows us to create meshes with non-integral dimensions of the initial unit cell and results in dramatic reduction of a scatter in simulated transport coefficients (i.e. drag or permeability)¹.

2.3 Flow Simulation Approach: The Lattice Boltzmann Method

For the simulations of Stokes flow we use the lattice Boltzmann method (LBM), a suitable approach for simulations of flows in complex geometries. LBM is based on the evolution of a “lattice gas” consisting of distribution functions of gas molecule fluxes along a prescribed discrete set of lattice links (19 in this study) connecting each lattice node with its neighbours. Each LBM iteration consists of two steps: i) propagation of fluxes between lattice nodes along lattice links (streaming step), and ii) collision of fluxes at each lattice node (collision step). In LBM simulations with the course of time its distribution function (of fluxes) iteratively approaches an equilibrium, and simulation is done after variation of the distribution function becomes sufficiently small. The streaming step is a universal one for various LBM models, while implementation of the collision may differ. A simple and commonly used version of the collision operator is so-called “BGK”⁴, which contains only one relaxation time τ defining the decay rate of the function towards its equilibrium as well as the viscosity of the simulated fluid. In the case of using so-called “bounce-back” boundary condition (a *de facto* approach for simulations based on micro-CT images of complex medium), a well known drawback of LBM BGK is its dependency of the simulated permeability on τ ⁶.

Recent developments of LBM models resulted in collision operators with collision occurring in the space of hydrodynamic and kinetic moments (like density, momentum, energy, energy flux), with corresponding conversion of the distribution function to/from the momentum space. Such a formulation of the collision operator⁷ allows introduction of multiple (up to the amount of links per lattice node, or 19 in this study, $\tau_{0\dots 18}$) relaxation times – so-called MRT collision operator – for each individual mode, which also poses a question on the particular selection of those parameters, remaining unanswered till now. However, it is suggested that for Stokes simulations individual adjustment of all MRT relaxation rates is not necessary, and in the case of properly grouping the relaxation rates only two prescribed values τ_ν , τ_f are sufficient to provide viscosity independent simulations of permeability. Namely, in this case the multiple-relaxation-time collision operator is reduced to two-relaxation-time (TRT) where, from the point of simulations, τ_ν controls fluid viscosity while τ_f remains apparently free. However, it was demonstrated that when the following combination of these relaxation parameters $\Lambda = (\tau_\nu - 1/2)(\tau_f - 1/2)$ stays fixed, LBM results in viscosity-independent permeability up to machine accuracy in any geometry^{5,6}. Mathematical analysis of LBM behaviour for the simple case of flow between two parallel plates (Poiseuille flow) revealed that Λ controls the location of the zero-velocity boundary between adjacent solid and fluid voxels, and there are some specific values of Λ for this geometry: $\Lambda = 3/16$ results in correct boundary location for horizontally-oriented open channel while $\Lambda = 3/8$ – for diagonal orientation; $\Lambda = 1/8$ gives analytical (up to machine accuracy) flow velocity for horizontal orientation. Analysis of more complex geometries still remains an open topic, and therefore we consider four “basic” values of $\Lambda = 1/8, 3/16, 3/8$, and $1/4$ – the latter providing results of widely used BGK collision with $\tau = 1$ – as well as wider variation of Λ between $1/512$ and 2 .

Here our value of interest is the average drag (F_d) exerted on a sphere, which can be seen as the quantity opposite to permeability (k): for a given geometry higher drag means lower permeability, and, in other words, higher resistance of a given geometry to the flow. Drag and permeability are related as $F_d = \frac{d_{sp}^2}{18(1-\varepsilon)k}$. For the reference values of drag F_d^* we use the accurately determined ones from our previous study¹.

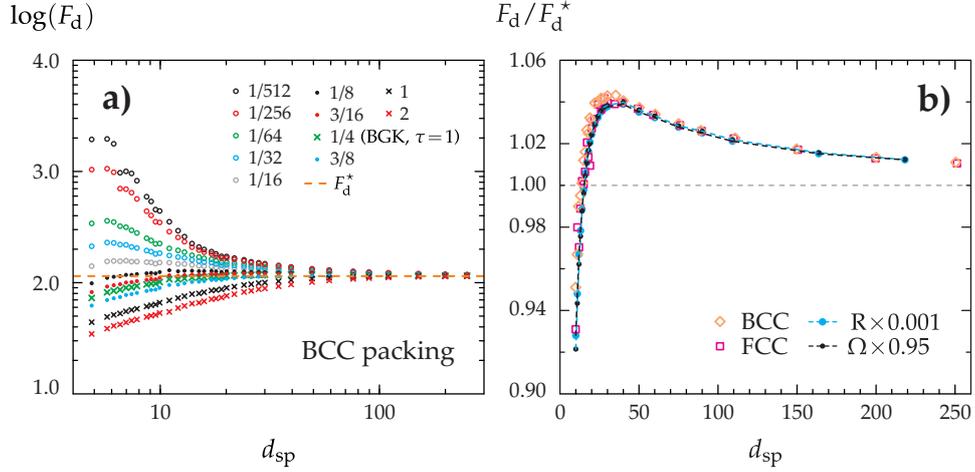


Figure 2. a) Dependence of F_d on discretisation resolution (d_{sp}) in BCC packing; all other packing types (FCC, $R \times 0.001$, $\Omega \times 0.95$) demonstrate a very similar picture (not shown). b) Convergence of the relative drag F_d/F_d^* calculated as average from 4 “basic” Λ values (1/8, 3/16, 1/4, 3/8); all packing types are shown.

2.4 General Impact of Mag on Drag

We start our analysis by addressing the impact of Λ using one of the regular packings (BCC). The results (Fig. 2a) reveal crucial influence of Λ on drag at lower spatial resolutions: variation of Λ may lead to ten-fold over- or underestimation of the accurate value. A similar picture is observed for other three packings¹ (not shown). Convergence towards F_d^* vs. spatial resolution can be non-monotonous, and its behaviour strongly depends on the particular Λ value. On the next step we took four “basic” values (1/8, 3/16, 1/4, 3/8) and calculated the average drag for them. The results shown in Fig. 2b confirm non-monotonous convergence: at low resolutions LBM underestimates F_d^* , then the numerical solution crosses the accurate value and hereafter starts very slow convergence from above. Such a behaviour was observed in various previous works but interpreted differently (suggesting no convergence at all at higher resolutions⁸ or attributing a low-resolution branch of the curve to the discrete porosity error⁹, which, in turn, is significantly lower in this study¹ compared to the work of Maier *et al.*⁹). Fig. 2b motivates to split the following analysis into higher ($d_{sp} > 30$) and lower ($d_{sp} < 30$) resolution regions.

2.5 Analysis of High-Resolution Region

A closer examination of the high-resolution region in Fig. 2a is shown in Fig. 3a,b, where very similar behaviour in relative drag F_d/F_d^* is shown for one regular (BCC) and one irregular ($R \times 0.001$) packings. As both panels a) and b) of Fig. 3 reveal, drag values calculated with all considered Λ values collapse to each other above F_d^* , and similar behaviour is observed for the other two packings (not shown). For the whole range $\Lambda \in [1/512, 2]$ variation of Λ beyond “basic” values of $\Lambda \in [1/8, 3/8]$ does not change the picture qualitatively: for all Λ the solution crosses its accurate value F_d^* , meaning that with an increase of resolution a numerical solution obtained for arbitrary large value of Λ will cross F_d^* .

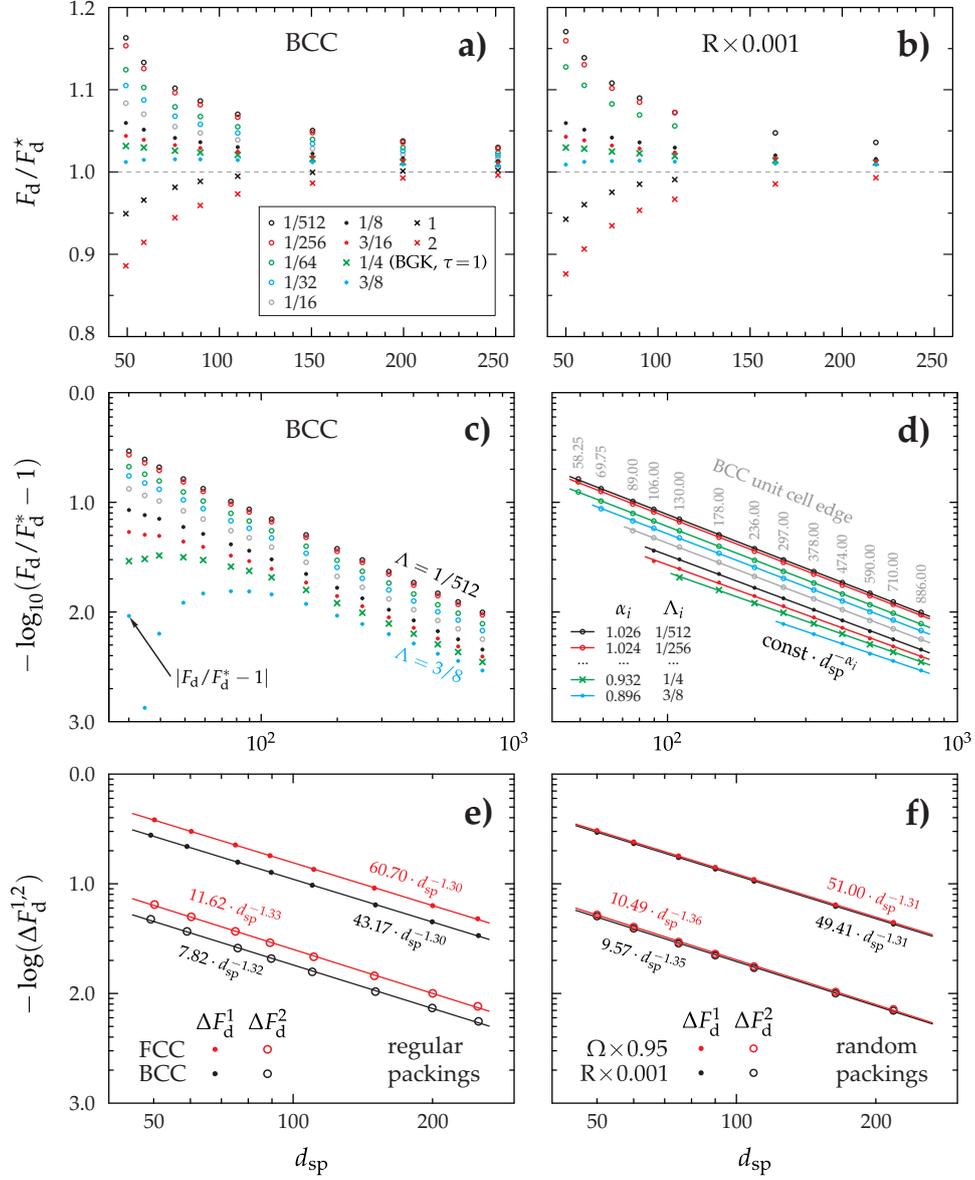


Figure 3. a, b) Dependence of the normalised drag F_d/F_d^* on discretisation resolution at high resolution values for one regular (BCC) and one irregular (R x 0.001) packing. c) The same as b) but with further resolution increase and log-log axis scale. d) Particular points extracted from c) with the corresponding power law fits; inset shows scaling exponents. e) and f) Differences between drag values ΔF_d^i calculated with Λ pairs of (1/512, 2) and (1/8, 3/8) for all packing types: $\Delta F_d^1 = (F_d^{\Lambda=1/512} - F_d^{\Lambda=2})/F_d^*$ and $\Delta F_d^2 = (F_d^{\Lambda=1/8} - F_d^{\Lambda=3/8})/F_d^*$.

Panel c) in Fig. 3 shows the drag error on a double-logarithmic scale. Note that we keep the error sign (except one indicated point) and did not plot its negative values. Log-

log scale clearly demonstrates transient convergence rate of larger Λ values, and for larger Λ a constant convergence rate is observed for higher resolution values. Partially observed constant convergence rates motivate to perform fitting only for the corresponding curve regions, which is shown in Fig. 3d. The obtained convergence rates are -1.0 or slower indicating that selection of larger Λ on the one hand may result in lower absolute errors at high resolutions, but on the other hand approximately constant convergence rates are observed at much higher resolution as well as the rates of convergence becomes slower.

It is straightforward to address convergence of the relative difference in the drag simulated for various Λ values. For this purpose we took two pairs of limiting values: $[1/512, 2]$ for the whole considered Λ range and $[1/8, 3/8]$ for the “basic” Λ values only such that $\Delta F_d^1 = (F_d^{\Lambda=1/512} - F_d^{\Lambda=2})/F_d^*$ and $\Delta F_d^2 = (F_d^{\Lambda=1/8} - F_d^{\Lambda=3/8})/F_d^*$. Our results reveals almost identical convergence rate of ~ 1.3 for all four packings types as well as for two considered difference intervals. This allows to conclude that with the increase of resolution the drag values obtained with virtually any pair of Λ will first collapse to some value above F_d^* and then continue its slow convergence to F_d^* from above.

2.6 Analysis of Low-Resolution Region

Low discretisation resolutions are common in practical LBM simulations^{10,11}. As Fig. 4 shows, Λ has a strong and dispersive impact on the simulated drag values. On the other hand, even for the lower resolution case its impact on the relative drag F_d/F_d^* is almost identical for all four packings. We note that thanks to our improved discretisation approach one can see systematic trends in drag behaviour for $d_{sp} < 20$ in regular packings. Drag obtained with the basic Λ values $[1/8, 3/8]$ systematically under- and then overestimates the accurate value F_d^* . At the same time F_d obtained for Λ values out of the basic region are strictly lower or higher than F_d^* , suggesting to use this property and to calculate a new value of F_d taking the average of drag values obtained for Λ out of the basic range. Among initially chosen Λ values the lowest total error demonstrated the pair of $\Lambda = 1$ and $\Lambda = 1/16$ with the resulting average drag indicated by solid black line in Fig. 4: except for very low spatial resolutions $d_{sp} < 10$ such averaging of two simulations with different Λ improves the final simulation accuracy.

3 Conclusion

We performed simulations of Stokes flow using two-relaxation-time (τ_ν and τ_f) LBM with bounce-back boundary condition, taking as input geometries two regular and two irregular packings of equal spheres fixing their porosity to $\varepsilon = 0.366$. Such a choice of geometries allowed us to study systems with relatively complex pore space and at the same time accurately determined reference permeability/drag values. We studied the impact of the specific combination of LBM relaxation times $\Lambda = (\tau_\nu - 1/2)(\tau_f - 1/2)$ which controls spatial location of zero-velocity boundary in flow simulations. Λ was varied within the wide range of $[1/512, 2]$, giving additional attention to the “basic” values of $1/4$ (widely used in BGK LBM simulations) as well as $1/8, 3/16, 3/8$ which provide exact velocity profiles or average flow rate in slit channel geometry accessible for direct analytical analysis. The “basic” Λ values did not provide any additional gain in accuracy, and fall into the general pattern of solutions with other values from $\Lambda \in [1/512, 2]$. Due to the dispersive impact of Λ

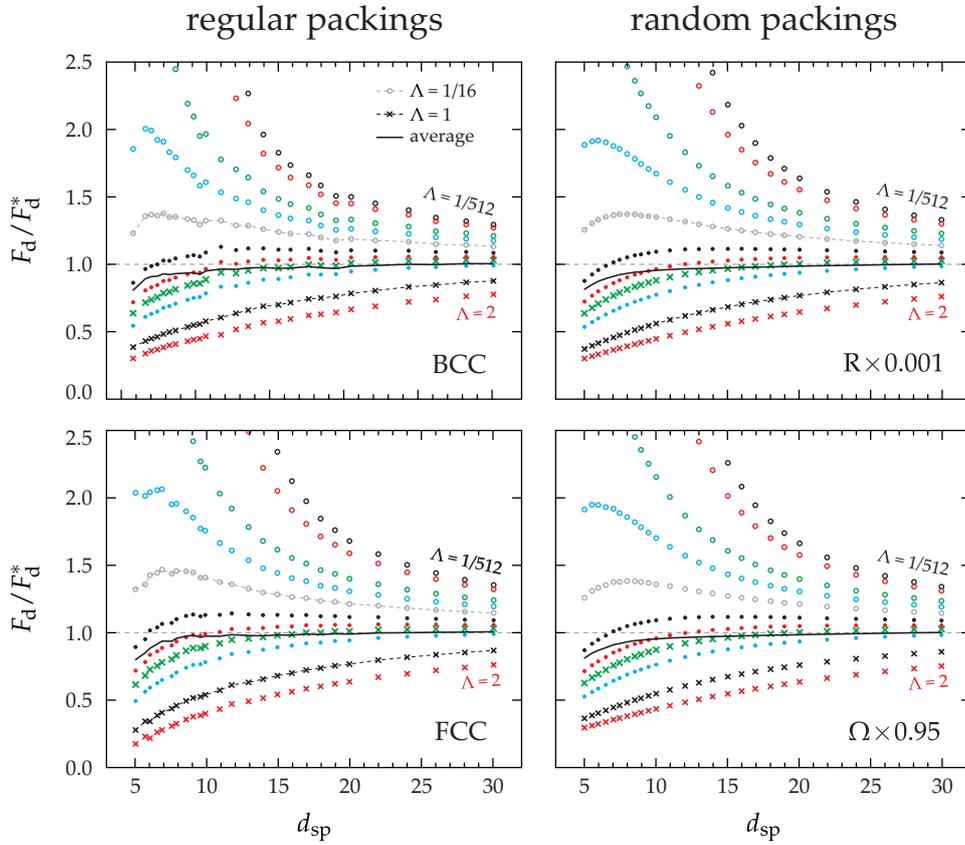


Figure 4. Dependence of the relative drag F_d/F_d^* on discretisation resolution for its lower considered values. The results are shown for $\Lambda = 1/512, 1/256, 1/64, 1/32, 1/16, 1/8, 3/16, 1/4, 3/8, 1, 2$. Some drag values obtained with $\Lambda \leq 1/64$ are not shown due to their larger magnitudes. Black solid line is the drag averaged from two relative drag values obtained with $\Lambda = 1/16$ and $\Lambda = 1$.

on F_d we split our analysis addressing separately high and low resolution regions. High resolution analysis revealed that F_d with virtually any Λ will first achieve some F_d value above the accurate solution and then, with further increase of the resolution, will continue converging to F_d^* at the rate of -1.0 or slower. Addressing lower resolution allowed us to suggest that instead of using any particular Λ value one can achieve higher accuracy after performing two simulations for a pair of (Λ_1, Λ_2) values and then averaging the result; our brief study suggests $\Lambda_1 = 1/16$ and $\Lambda_2 = 1$ to be a good candidate.

Our analysis in this paper is based on variation of spatial resolution within wide range. Although at a first glance higher resolutions are of purely academic interest, they are necessary to obtain complete picture of the solution behaviour (convergence) which later may open the possibility to predict higher resolution results by performing actual simulations in low-to-moderate resolution region only. Variation of the resolution is a very computationally demanding task because the computational effort grows as $O(d_{sp}^5)$ where $O(d_{sp}^3)$

comes from the increase of amount of mesh nodes to be processed on each iteration and $O(d_{sp}^2)$ is due to the higher amount of iterations needed to achieve a given accuracy, which originates from the increase of the distance (in lattice units) between “solid” mesh nodes. In fact, the situation is even worse because at high resolutions the absolute error magnitude becomes small and the simulation accuracy must be further increased to accurately resolve systematic trends in the behaviour of errors with smaller magnitude. To be more specific about computational efforts, a single point in Fig. 3b for $d_{sp} > 200$ required about 5 hours of calculation using 32768 processor cores.

References

1. S. Khirevich, I. Ginzburg, and U. Tallarek, *Coarse- and fine-grid numerical behavior of MRT/TRT lattice-Boltzmann schemes in regular and random sphere packings*, J. Comp. Phys. **281**, 708–742, 2015, DOI: 10.1016/j.jcp.2014.10.038.
2. W. S. Jodrey and E. M. Tory, *Computer simulation of isotropic, homogeneous, dense random packing of equal spheres*, Powder Technol. **30**, 111–118, 1981, DOI: 10.1016/0032-5910(81)80003-4.
3. S. Succi, *The lattice Boltzmann equation for fluid dynamics and beyond*, Oxford University Press, New York, 2001.
4. S. Chen and G. D. Doolen, *Lattice Boltzmann method with regularized non-equilibrium distribution functions*, Ann. Rev. Fluid Mech. **30**, 329–364, 1998, DOI: 10.1146/annurev.fluid.30.1.329.
5. D. d’Humières and I. Ginzburg, *Viscosity independent numerical errors for Lattice Boltzmann models: from recurrence equations to “magic” collision numbers*, Comp. Math. Appl. **58**, 823–840, 2009, DOI: 10.1016/j.camwa.2009.02.008.
6. I. Ginzbourg and P. M. Adler, *Boundary condition analysis for the three-dimensional lattice Boltzmann model*, J. Phys. II France **4**, 191–214, 1994, DOI: 10.1051/jp2:1994123.
7. D. d’Humières, I. Ginzburg, M. Krafczyk, P. Lallemand, and L.-S. Luo, *Multiple-relaxation-time lattice Boltzmann models in three dimensions*, Philos. Trans. R. Soc. A **360**, 437–451, 2002, DOI: 10.1098/rsta.2001.0955.
8. C. Pan, L.-S. Luo, and C. T. Miller, *An evaluation of lattice Boltzmann schemes for porous media simulation*, Comput. Fluids **35**, 898–909, 2006, DOI: 10.1016/j.compfluid.2005.03.008.
9. R. S. Maier, D. M. Kroll, Y. E. Kutsovsky, H. T. Davis, and R. S. Bernard, *Simulation of flow through bead packs using the lattice Boltzmann method*, Phys. Fluids **10**, 60–74, 1998, DOI: 10.1063/1.869550.
10. B. Bijeljic, P. Mostaghimi, and M. J. Blunt, *Insights into non-Fickian solute transport in carbonates*, Water Resour. Res. **49**, 2714–2728, 2013, DOI: 10.1002/wrcr.20238.
11. L. Talon, D. Bauer, N. Gland, S. Youssef, H. Auradou, and I. Ginzburg, *Assessment of the two relaxation time Lattice-Boltzmann scheme to simulate Stokes flow in porous media*, Water Resour. Res. **48**, W04526, 2012, DOI: 10.1029/2011WR011385.

Computational Plasma Physics

Computational Plasma Physics: Two Examples from Active HPC Areas in Plasma Physics

Karl-Heinz Spatschek

Institut für Theoretische Physik, Heinrich-Heine-Universität Düsseldorf,
40225 Düsseldorf, Germany
E-mail: khs@tp1.uni-duesseldorf.de

The two projects being presented here provide new insights into plasma-based acceleration processes and plasma astrophysics, respectively. The first paper¹ deals with a future plasma-based free-electron laser (FEL) which - when being in operation - will have an enormous impact on many scientific disciplines. The path towards a first successful realisation crucially requires advanced simulation codes to model and study the new physics involved. The second paper² is on collisionless shocks. The latter are important in various fields of physics. Perhaps they are most prominent in the context of plasma astrophysics in order to explain the open problem of acceleration of cosmic rays to very high energies. Interestingly, the theory on shock physics can be simulated under laboratory conditions for the generation of energetic ions with applications in inertial fusion or medical physics. Thus, the common bracket of both papers is the plasma-based acceleration physics of particles.

Plasmas are many body systems. Since most of them are not in thermodynamic equilibrium, either a sophisticated non-equilibrium theory is required or high-performance numerical simulations become necessary. The type of modelling depends on the nature of the plasma under consideration. We distinguish between

- plasmas for magnetic fusion,
- plasmas as a result of laser-matter-interaction (including inertial fusion)
- charged beam interaction with matter, radiation, or plasma,
- complex plasmas (including quantum plasmas),
- plasma astrophysics,
- plasma technology (including low-temperature plasmas),

just to name the most prominent examples. The demand on computer time and storage is huge in all these fields.

According to the rich variety of plasmas and their different time and space scales, the theoretical models are multifaceted, e.g. classical, quantummechanical, non-relativistic, or relativistic, respectively. They apply to high-temperature or low-temperature, weakly or strongly coupled plasmas. Nonlinearities play the central role. Numerical approaches encompass single particle dynamics, particle-in cell codes, kinetic, gyro-kinetic, hydrodynamic (plasmadynamic), and magnetohydrodynamic (MHD) treatments, just to name the most popular ones.

The first paper¹ discusses a new approach for a next-generation of free-electron lasers (FELs). The idea is based on the concept of laser-plasma accelerators, promising extremely

compact FEL setups, combined with unique features of the generated X-ray pulses. The experimental setting is as follows. After the laser-plasma interaction, the generated electron beam then emits X-ray pulses while propagating through the subsequent FEL undulator. The process of electron beam generation and acceleration within the plasma has to be controlled. Beam physics itself is multifaceted, containing interaction with radiation, matter and plasmas. Beam self-fields (space charge), focusing, and bunching have to be considered. The processes occurring in the proposed plasma-driven free-electron laser happen on μm length and femtosecond (fs) time scales. These scales require completely new codes³ as an essential tool to study the novel plasma-based FEL concept.

Being presented by a leading group in concept design for new FELs, the paper¹ clearly describes the challenges and possible solutions for a new plasma-driven FEL. The authors report on the remarkable progress based on the interplay between experiment, theory, and high-performance computing (HPC).

The second paper² is written by astrophysicists. It treats the “classical” problem of collisionless shocks, making use of the fully relativistic particle-in-cell code OSIRIS⁴. Collisionless shocks form due to the interaction of particles with electromagnetic fields. Micro-turbulence is seeded due to plasma instabilities, which is then transformed to a large scale structure. The authors² have studied the shock formation process and the mediating instabilities in astrophysical environments, i.e. in electron-positron pair plasmas as well as electron-ion plasmas.

The mediating plasma micro-instabilities play a key role. The authors² present an overview over the shock formation processes in plasmas and treat in detail the evolution of electromagnetic instabilities on different scales, including the related shear flows. They studied the parameter conditions for shock acceleration, which is very efficient in terms of generation of energetic ions.

Besides the interesting applications in astrophysics (e.g. cosmic ray acceleration) the paper also shows that acceleration mechanisms may be similar in large-scale astrophysics and table-top plasma based terrestrial accelerators. Meanwhile the OSIRIS code⁴ has proven to become the “workhorse” of high-intensity plasma dynamics when PIC resolutions are sufficient.

References

1. Irene Dornmair *et al.*, *Towards Plasma-driven free-electron lasers*, this volume.
2. Anne Stockem Novo *et al.*, *Collisionless shocks and mediating instabilities*, this volume.
3. L. T. Campbell and B. W. J. McNeil, *Puffin: A three dimensional, unaveraged free electron laser simulation code*, Phys. Plasmas **19**, 093119, 2012.
4. R. A. Fonseca *et al.*, *OSIRIS: A three-dimensional, fully relativistic particle in cell code for modeling plasma based accelerators*, Lecture Notes in Comput. Sci. **2331**, 342, 2002.

Towards Plasma-Driven Free-Electron Lasers

Irene Dornmair¹, Lawrence T. Campbell^{3,4}, James T. Henderson³, Sören J alas¹,
Oliver Karger², Manuel Kirchen¹, Alexander Knetsch², Grace G. Manhan³,
Georg Wittig², Bernhard Hidding^{2,3}, Brian W. J. McNeil³, and Andreas R. Maier¹

¹ Center for Free-Electron Laser Science & Department of Physics, University of Hamburg,
Luruper Chaussee 149, 22761 Hamburg, Germany
E-mail: andreas.maier@desy.de

² Department of Physics, University of Hamburg,
Luruper Chaussee 149, 22761 Hamburg, Germany

³ Department of Physics, University of Strathclyde, Glasgow, G4 0NG, United Kingdom

⁴ ASTeC, STFC Daresbury Laboratory and Cockcroft Institute, Warrington, United Kingdom

Plasma-based accelerators are promising candidates to drive next-generation compact free-electron lasers (FEL) with unique X-ray properties. The correct modelling of the physics in the laser-plasma accelerator and of the FEL amplification from plasma-generated electron bunches thereby forms the basis for any application driven by these beams. Both problems are, however, extremely challenging. They heavily rely on computer simulations and, in fact, act as science drivers for the development of the latest generation of high-performance, highly parallel and efficient simulation codes. Here, we present our work towards a plasma-driven FEL, based on the particle-in-cell (PIC) codes VSIM and WARP, and the non-averaging 3D FEL code PUFFIN.

1 Introduction

Free-Electron Lasers (FEL) are essential tools for a broad scientific community to study the dynamic of matter on its natural length and time scale. The availability of today's large-scale facilities to users, however, is very limited. A new approach for a next-generation FEL is based on the concept of laser-plasma accelerators (LPA), promising extremely compact FEL setups, combined with unique features of the generated X-ray pulses.

The path towards a plasma-based FEL is extremely challenging. Foremost, the process of electron beam generation and acceleration within the plasma has to be controlled. As the processes within the plasma happen on μm length and femtosecond (fs) time scales, the whole field relies heavily on computer-simulations to (i) predict and study new acceleration techniques, (ii) design new experiments, and (iii) analyse experimental data. Correctly modelling the physics within a plasma accelerator, under the restriction of limited computational resources, is still an area of very active development within the community. In fact, it is one of the scientific drivers for the development of highly parallel, high-performance simulation codes. Following the laser-plasma interaction, the generated electron beam then emits X-ray pulses while propagating through the subsequent FEL undulator. Here, existing codes, that have been successfully used to model the physics of today's large-scale FEL facilities, are no longer sufficient to describe the physics involved. New codes are required, and in active development, as an essential tool to study the novel plasma-based FEL concepts.

Our project, *Towards Plasma-Driven Free-Electron Lasers*, covers crucial aspects of the development of a future plasma-based FEL. The electron beam quality, especially in

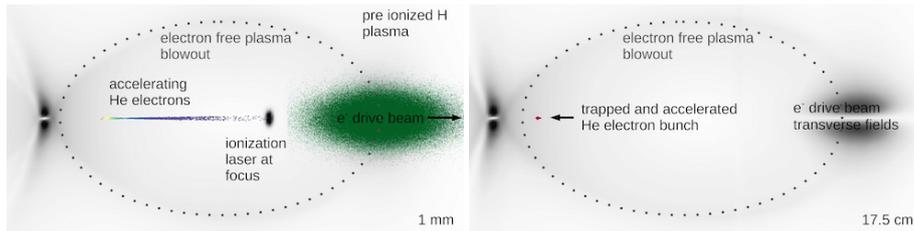


Figure 1. *Left:* Trojan Horse witness bunch generation. An electron beam (green particles) drives a plasma wakefield, represented by its field sum with up to 16 GV/m acceleration fields (gray). The high fields at the rear of the blowout are due to overshooting electrons, which were re-attracted to the drive beam axis. A strongly focused, short pulse laser (black) ionises neutral helium gas to generate free electrons in the wakefield (colour coded particles with 0-1 MeV energy). *Right:* The trapped witness has been accelerated to 1.1 GeV over a distance of 17.5 cm. Further parameters are 2.3 mm.mrad emittance, 10 pC charge, 2.1 kA peak current, and 1.2 % energy spread.

terms of energy spread and emittance, ultimately determines whether a beam is suitable to drive an FEL. It is therefore of utmost importance first, to generate high quality beams, and second, to maintain this quality throughout the acceleration section in the plasma and the subsequent transport to the FEL undulators.

In general, one can distinguish two different approaches in LPA: internal and external injection. In the former, an electron bunch is created inside the plasma from electrons that are pulled from the plasma background and directly accelerated within the same plasma target. In the latter, an external electron beam from a high quality source, such as a conventional accelerator, is injected into another plasma stage where it is further accelerated. In our project, we study both approaches. Electron beams studied in the two sub-projects are then imported into the new FEL simulation code, PUFFIN¹, that we develop and use to describe novel FEL schemes, based on plasma-generated electron beams.

2 Hybrid Systems and Underdense Plasma Photocathode

One path towards the generation of high-quality electron bunches is the underdense plasma photocathode (or “Trojan Horse” scheme). The central idea is to use an electron beam to drive a plasma wakefield, and employ a synchronised, strongly focused, short laser pulse to locally ionise additional electrons from a neutral high-ionisation-threshold background gas within the centre of the wakefield, see Fig. 1. We use the commercial PIC code VSIM to model this scheme, advance it conceptually, and to prepare and assist our granted beamtime at FACET¹⁹ (experiment E-210). There, we will for the first time demonstrate the Trojan Horse concept experimentally.

In the Trojan Horse scheme, a cavity is generated by a high energy, high charge electron driver beam which expels electrons off axis, leaving a positive ion background and providing GV/m level accelerating fields – which is 2 to 3 orders of magnitude larger than in conventional RF cavities. The expelled electrons form a negatively charged electron boundary and become re-attracted towards the axis after one plasma wavelength, forming an electron-free cavity (also known as “blowout”). Electrons created from the synchronised laser pulse can be accelerated to high energies within the plasma wave, if the wakefield is

strong enough to accelerate the electrons to relativistic energies within a length that is substantially smaller than one plasma wavelength. These electrons are then trapped at the rear of the plasma blowout – which is also moving with nearly the vacuum speed of light. The strong localisation of the electron release in combination with the low initial momentum of the electrons create high quality bunches of low emittance, that are then accelerated by the plasma wave. The scheme furthermore gives unprecedented control over the injection, and allows for tuning of charge and other important bunch parameters on a level of control hitherto only known from classical photocathodes.

Since the original publication² we could significantly advance this concept. A major requirement of electron beam driven accelerators is the need for a sufficiently strong electron beam driver, that, so far, requires a conventional large scale accelerator. We are therefore actively investigating laser-plasma accelerators as an alternative and more available source of suitable drive beams³, and study methods to relax the requirements on the electron drive beam. This includes, for example, a plasma density downramp assisted trapping of electrons within the Trojan Horse scheme⁵. Here, the witness bunch electrons are ionised from the beginning of a plasma density downramp, which facilitates trapping, and thus allows for relaxed demands on the drive beam.

We could show that the demands on the driver electron beam quality in our Trojan Horse scheme are substantially lowered, compared for example to the requirements on bunch quality for FEL operation, and readily available by today's laser plasma accelerators⁴. We found, that the most crucial properties for FEL driver beams, energy spread and emittance, are quite unimportant for driving a plasma wakefield. Within the parameter range of today's laser-plasma accelerated bunches, an accelerated witness bunch is hardly affected by changing these properties of the drive beam. Other parameters like beam energy and current are more sensitive for a stable high gradient wakefield, and must not be beyond certain thresholds to guarantee stable acceleration, which is, however, a far easier task than generating small energy spread beams from a laser-plasma accelerator.

As another important achievement, we could propose a novel concept for witness bunch generation in beam driven plasma wakefield accelerators¹⁷, which is easy to set up experimentally and can even help to implement the underdense photocathode mechanism. In this context, we studied the sources of dark current, which can significantly lower the quality of the witness bunch¹⁸.

3 External Injection of Well-Defined Electron Bunches

A second concept to generate high-quality electron bunches, suitable for FEL operation, is based on well-defined electron beams from a conventional accelerator, that are injected into a subsequent plasma accelerator stage. By separating the tasks of electron beam generation, in the external electron source, from the actual acceleration, one could combine the advantages of both technologies: the extremely good control over the electron phase space in a conventional injector, and the huge accelerating field gradients in the plasma.

However, modelling these regimes requires mastering a new set of simulation technologies. External injection schemes typically operate with a significantly lower plasma density, compared to internal injection, which consequently lengthens the plasma target from only a few mm to a cm-scale. The significantly longer target can dramatically increase the computational costs of a PIC simulation. When modelling the problem in a

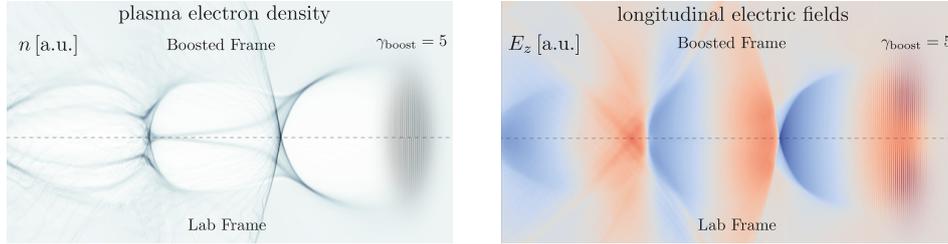


Figure 2. Comparison of the results of a 2D WARP simulation in a boosted frame ($\gamma_{\text{boost}} = 5$) with a reference simulation in the lab frame. The left panel shows the plasma electron density n , the right panel shows the longitudinal electric field E_z , both in arbitrary units. The upper half of the images corresponds to the back transformed results of the boosted frame simulation, whereas the results of the reference simulation are shown in the lower half.

Lorentz-boosted frame the plasma target gets contracted and it also propagates towards the laser. The number of timesteps and consequently the simulation runtime thereby can be reduced by up to a factor $2\gamma_{\text{boost}}^2$, where γ_{boost} is the Lorentz factor of the boosted frame. We are therefore actively contributing to the development of the open-source PIC code WARP, and its capabilities to operate in the Lorentz-boosted frame, which is essential for our work.

For example, to interpret the simulation data, it needs to be transformed from the boosted frame back to the laboratory frame. However, the space and time variables of the boosted and the lab frame are coupled. To create a dataset that covers the simulation space at one moment in lab time, snippets of data from many times and positions in the boosted frame need to be combined. We implemented efficient and performance optimised routines which sort and save the lab frame output on-the-fly in a boosted frame simulation. The agreement between the back-transformed data and a simulation in the lab frame is very good, as can be seen from Fig. 2.

With these techniques we study several crucial physics aspects of a plasma-based FEL: we proposed the adiabatic matching technique for plasma accelerators, as well as a plasma-based phase space diagnostics concept.

One severe challenge for external injection schemes arises from the focusing forces inside the plasma wakefield. In the plasma, the large accelerating fields are accompanied by strong transverse fields that focus the electron beam, and the externally injected beam needs to have a transverse beam size that is matched to these focusing fields. Otherwise, the beam size will oscillate around the matched beam size, like in a harmonic oscillator, which causes strong emittance growth⁶. However, this matched beam size is typically extremely small. In consequence, a very strong focusing optics before the plasma target is required to condition the beam to the plasma entrance, and also after the plasma stage, as the beam will also exit the plasma with a small beam size and consequently with a large divergence. This is not only challenging from a technological point of view, but also leads to emittance growth in the drift after the target⁷.

This crucial issue can be tackled by properly designing the plasma target. Before and after the actual acceleration section a dedicated injection and extraction section is added to the target. There, the focusing forces are slowly increased or decreased and allow the

bunch to adapt from vacuum to plasma and out into vacuum again⁸. The focusing forces are determined by the density profile and by the driver laser envelope, which need to be optimised in order to keep the sections short. For such optimisations extensive parameter scans in PIC simulations are inevitable.

Another challenge arises from the initially ultrashort few-fs bunches that are produced by laser-plasma accelerators. The bunch length is also a crucial parameter for external injection experiments. Since the accelerating field has a slope, different longitudinal slices of the bunch will experience a different accelerating field, which will increase energy spread. In order to minimise this effect, the bunches must be much shorter than one plasma wavelength. However, the measurement of bunch lengths and even current profiles with femtosecond resolution is extremely challenging.

In conventional accelerators, transverse deflecting structures are employed that are expensive, several meters long, and feature a resolution down to a few fs. For external injection experiments, we recently proposed to use a plasma target for the diagnostics of ultrashort electron beams⁹, which promises a resolution of below 100 as, depending on the beam emittance and laser parameters. It does not require any additional instruments, as it employs the already installed driver laser. However, the method is sensitive to beam loading, which distorts the wakefield by the self-fields of the accelerated beam. These effects are only accurately represented in 3D PIC simulations, which are computationally demanding owing to high requirements on both memory and CPU performance.

4 FEL Simulations with PUFFIN

A key component in the design of FEL facilities is the use of numerical codes. Commonly used FEL codes like GENESIS¹⁰ perform various approximations of the mathematical models describing the FEL interaction, such as the slowly varying envelope approximation and averaging of the electron motion over a cycle, which limit their applicability. As such, they are not well suited to simulate some advanced FEL concepts. Neither do these approximations allow proper simulation of electron beams typically produced by laser plasma accelerators, which have quickly varying currents and a large range of energies.

Considering this, the so-called unaveraged 3D FEL simulation code PUFFIN¹ has been developed, which is free of limiting assumptions. The advantage is an enhanced resolution, revealing new physics and enabling proper simulation of some advanced FEL concepts; the downside is the increased computational demand required to model this higher resolution. As such, it is required to run PUFFIN on HPC machines, utilising MPI, when in full 3D mode, and the availability of sufficient HPC resources is now a crucial prerequisite for the design of novel FEL schemes.

Within our project *Towards Laser-Plasma Driven Free-Electron Lasers*, we significantly enhanced the features of PUFFIN to handle some peculiarities, which are intrinsic to novel, plasma-based FEL schemes. PUFFIN previously solved the radiation diffraction in Fourier space, implying periodic boundary conditions, which were causing problems when, under certain conditions, the radiation was propagating significantly outside the boundaries of the model¹¹, for example in proposed demonstration schemes for a first plasma-FEL¹². A first priority was therefore the implementation of absorbing boundary conditions, which have now been utilised in the transverse plane in a method similar to Ref. 13, see Fig. 3.

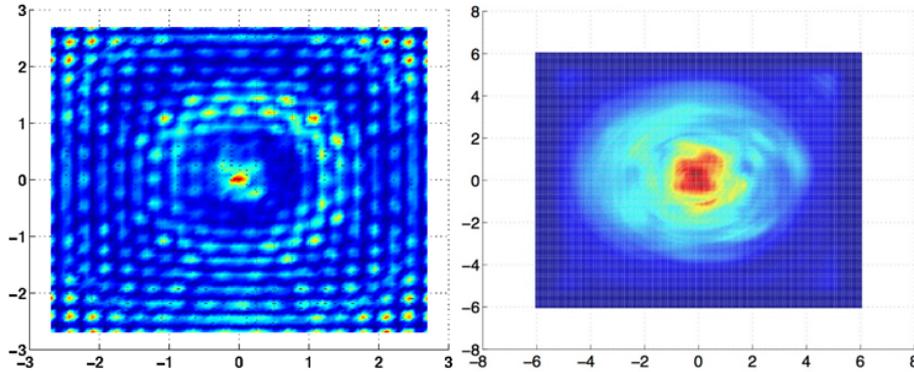


Figure 3. Typical output of the radiation field transverse plane (colour-coded intensity) in extreme diffraction regimes both before (left) and after (right) the implementation of absorbing boundary conditions. One can see a interference pattern before the boundaries were implemented, caused by the interference between the outgoing (towards the edges) emitted radiation and the incoming (from the edges) reflected radiation. This unphysical problem is mitigated by absorbing the radiation (right panel).

Another improvement over the originally presented code¹ is the inclusion of a 3D magnetic undulator field²⁰, that enables us to model realistic experiments. Previously, an artificial focusing channel was super-imposed on the 1D undulator field to approximate a so-called “natural focusing”; now this 3D focusing motion emerges naturally from the 3D magnetic undulator fields. In addition, various algorithmic improvements, such as a slightly modified field driving algorithm, which removes the need for a stiffness matrix (as was previously required and described in Ref. 1), and by paying careful attention to array storage, have reduced the run time by around 75 %.

These new features are crucial to study new FEL regimes, among which is a two-colour FEL scheme with wide frequency separation, as proposed in Ref. 14. In an FEL, the electron beam energy is usually resonant with only one radiated frequency. By switching the resonance condition between two widely separated, strongly non-harmonic frequencies in PUFFIN, the electron beam is shown to bunch at not only these frequencies, but the sum and difference frequencies, and the higher harmonics¹⁴, see Fig. 4. A bunching at a given frequency means the beam will radiate and amplify at that frequency quickly if it is also resonant. This could be exploited in future FEL designs to quickly generate multiple additional colours in the FEL output. This result, involving such widely separated frequencies, cannot be obtained using averaged FEL codes, which are limited to simulating amplification within a narrow bandwidth around a defined central frequency (and the higher harmonics of this frequency). PUFFIN, being an unaveraged code, simulates the full radiation spectrum self-consistently.

Due to this higher resolution, PUFFIN is also ideally suited for simulating the radiation arising from fine structures in the electron pulse. Based on such structures in the bunch¹⁵, we proposed in Ref. 16 a novel scheme that may improve the beam’s condition for lasing, or spontaneously produce high power, coherent radiation.

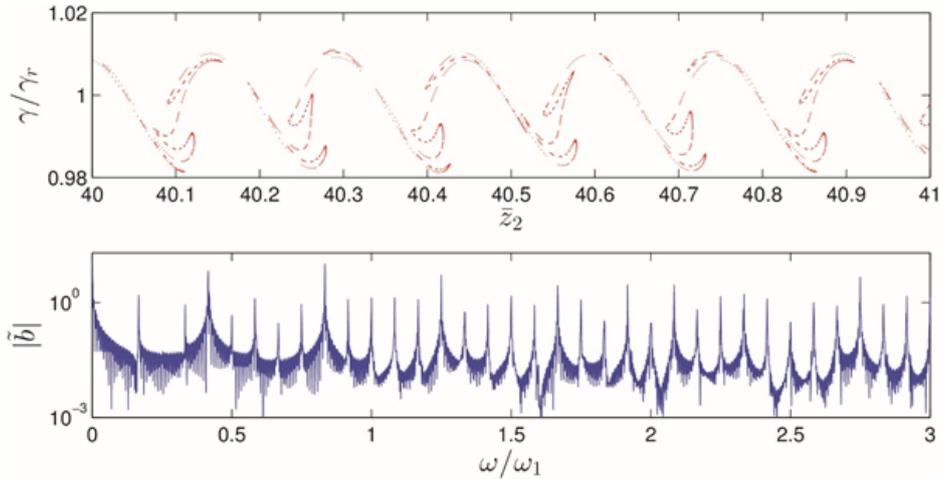


Figure 4. Electron beam phase space (top) and the electron beam bunching parameter as a function of frequency (bottom) at the end of our proposed 2 colour FEL scheme¹⁴. The beam phase space exhibits an oscillation at the shorter wavelength, which is “folded” into the oscillation at the longer wavelength. This results in the beam being bunched at not only these frequencies (the spikes in the bunching spectrum at frequencies 1 and 0.4), but also the sum and difference, and higher harmonics of these frequencies.

5 Concluding Remarks

Future plasma-based FELs would have an enormous impact on many scientific disciplines. The path towards a first successful realisation crucially requires advanced simulation codes, to model and study the new physics involved. To this end, we use a variety of codes, and we are specifically developing some of them to address physical problems, inherent to a plasma-based FEL. Over the past two years, we could show significant progress in addressing crucial aspects towards a future plasma-FEL, including enhanced beam quality from plasma accelerators, beam transport, and correct modelling of novel FEL schemes.

Acknowledgements

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA at Jülich Supercomputing Centre (JSC), project HHH20, and acknowledge EPSRC Grants EP/K000586/1 and EP/MO11607/1. We would like to thank J.-L. Vay and R. Lehe for support with the PIC code WARP, and J. Smith at Tech-X for support and development of VSIM.

References

1. L. T. Campbell, and B. W. J. McNeil, *Puffin: A three dimensional, unaveraged free electron laser simulation code*, Phys. Plasm. **19**, 093119, 2012.

2. B. Hidding, G. Pretzler, J. B. Rosenzweig, et al., *Ultracold Electron Bunch Generation via Plasma Photocathode Emission and Acceleration in a Beam-Driven Plasma Blowout*, Phys. Rev. Lett. **108**, 035001, 2012.
3. B. Hidding, T. Königstein, J. Osterholz, et al., *Monoenergetic Energy Doubling in a Hybrid Laser-Plasma Wakefield Accelerator*, Phys. Rev. Lett. **104**, 195002, 2010.
4. B. Hidding, G. G. Manahan, O. Karger, et al., *Ultrahigh brightness bunches from hybrid plasma accelerators as drivers of 5th generation light sources*, J. Phys. B **47**, 234010, 2014.
5. A. Knetsch, O. Karger, G. Wittig, et al., *Downramp-assisted underdense photocathode electron bunch generation in plasma wakefield accelerators*, arxiv:1412.4844, 2014.
6. T. Mehrling, J. Grebenyuk, F. S. Tsung, K. Floettmann, and J. Osterhoff, *Transverse emittance growth in staged laser-wakefield acceleration*, Phys. Rev. ST-AB **15**, 111303, 2012.
7. K. Floettmann, *Some basic features of the beam emittance*, Phys. Rev. ST-AB **6**, 034202, 2003.
8. I. Dornmair, K. Floettmann, and A. R. Maier, *Emittance conservation by tailored focusing profiles in a plasma accelerator*, Phys. Rev. ST-AB **18**, 041302, 2015.
9. I. Dornmair, C. B. Schroeder, K. Floettmann, B. Marchetti, and A. R. Maier, *Plasma-driven ultrashort bunch diagnostic*, submitted 2015.
10. S. Reiche, *GENESIS 1.3: a fully 3D time-dependent FEL simulation code*, Nucl. Instr. and Meth. A **429**, 243, 1999.
11. L. T. Campbell, A. R. Maier, F. J. Grüner, and B. W. J. McNeil, *Unaveraged Modelling of a LWFA Driven FEL*, Proceedings of FEL **2013**, MOPSO08, 2013.
12. A. R. Maier, A. Meseck, S. Reiche, et al., *Demonstration scheme for a laser-plasma-driven free-electron laser*, Phys. Rev. X **2**, 031019, 2012.
13. A. A. Gonoskov, and I. A. Gonoskov, *Suppression of reflection from the grid boundary in solving the time-dependent Schrodinger equation by split-step technique with fast Fourier transform*, arxiv **physics**, 0607120, 2006.
14. L. T. Campbell, B. W. J. McNeil, and S. Reiche, *Two-colour free electron laser with wide frequency separation using a single monoenergetic electron beam*, New J. Phys. **16**, 103019, 2014.
15. E. Hemsing, G. Stupakov, and D. Xiang, *Beam by design: Laser manipulation of electrons in modern accelerators*, Rev. Mod. Phys. **86**, 897, 2014.
16. J. R. Henderson, L. T. Campbell, and B. W. J. McNeil, *Free Electron Lasers Using "Beam by Design"*, New J. Phys., accepted for publication, 2015.
17. G. Wittig, O. Karger, A. Knetsch, et al., *Optical plasma torch electron bunch generation in plasma wakefield accelerators*, Phys. Rev. ST-AB, accepted for publication, 2015.
18. G. Manahan et al., *Sources and elimination of dark current in advanced plasma wakefield accelerators*, New J. Phys., submitted 2015.
19. M. J. Hogan et al., *Plasma wakefield acceleration experiments at FACET*, New J. Phys. **12**, 055030, 2010.
20. J. Henderson, L. T. Campbell, A. R. Maier, and B. W. J. McNeil, *The Implementation of 3D Undulator Fields in the Unaveraged FEL Simulation Code Puffin*, Proceedings of FEL **2014**, TUP022, 2014.

Collisionless Shocks and Mediating Instabilities

Anne Stockem Novo¹, Paulo Alves², Kevin Schoeffler²,
Luis O. Silva², and Reinhard Schlickeiser¹

¹ Institut für Theoretische Physik IV, Lehrstuhl für Weltraum- und Astrophysik,
Ruhr-Universität Bochum, 44780 Bochum, Germany
E-mail: anne@tp4.rub.de

² Grupo de Lasers e Plasmas – Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico,
Lisboa, Portugal

A shock is characterised by a jump in several physical quantities, such as the particle density, the fluid velocity or the temperature. In contrast to hydrodynamic shocks, collisionless shocks form due to the interaction of particles with electromagnetic fields. Micro turbulence is seeded due to plasma instabilities, which is then transformed to a large scale structure. In different sub-projects, we have studied the shock formation process and the mediating instabilities in astrophysical environments. Furthermore, the findings are applied to the field of laser-plasma physics in order to benefit from an interdisciplinary approach.

1 Introduction

Collisionless shocks are important in various fields of physics. They have been most prominent in the context of space and astrophysics in order to explain the open problem of acceleration of cosmic rays to very high energies. Although having been studied for many decades, many open questions still remain regarding the formation process, the dissipation of energy and the particle acceleration process. In this context, mediating plasma instabilities play a key role. We outline below 5 sub-projects that take advantage of parallel simulations with high performance supercomputing resources in order to investigate either the formation and saturation of shocks, related shear flows and field amplification, or laboratory applications.

Electromagnetic shocks are triggered by the filamentation¹ or Weibel² instability. In anisotropic plasma flows electromagnetic fluctuations develop which can create a strong magnetic field. This leads to a deflection of charged particles and a mass accumulation that triggers the collisionless shock. The process of shock formation is investigated in a sub-project presented in Sec. 2. The large scale formation of the field is studied in a different project which is described in Sec. 3. While the above mentioned projects deal with an idealised system of symmetric flows, it is important to take into account less ideal conditions for understanding the realistic, non-ideal scenario. In a further sub-project, electromagnetic modes in shear flows are investigated and presented in Sec. 4.

In two further sub-projects we connect astrophysics with laser-plasma physics. In Sec. 5, a method is presented which allows us to mimic astrophysical scenarios in the laboratory, helping to understand the processes at stake in the universe. In the second project, the theory on shock physics is simulated under laboratory conditions for the generation of energetic ions with applications in inertial fusion or medical physics, see Sec. 6.

All simulations were performed with the fully relativistic particle-in-cell code OSIRIS^{3,4}.

2 Formation of Electromagnetic Shocks

A collisionless electromagnetic shock forms in an ideal setup of two symmetric, cold counterstreams with relativistic fluid velocities. We studied the formation of such shocks in electron-positron pair plasmas and electron-ion plasmas. It was found that the time to form a steady-state shock in pair plasmas takes two times the saturation time of the mediating plasma instability^{5,6}. The formation process for electron-ion shocks was expected to scale similarly due to the rapid relativistic mass increase of the electrons, which happens on a few ω_{pe}^{-1} . A parameter scan of electron-ion shocks with different beam Lorentz factors and electron to ion mass ratios showed a formation time approximately 3 times longer than for pair shocks. After saturation of the filamentation instability on the ion time scale, an additional merging time of the magnetic field filaments is necessary, prolonging the shock formation process⁷. At the time of magnetic field saturation, the filament size is not big enough in order to significantly deflect the ions. Fig. 1 shows the paths of electrons and ions in the 2D simulation plane.

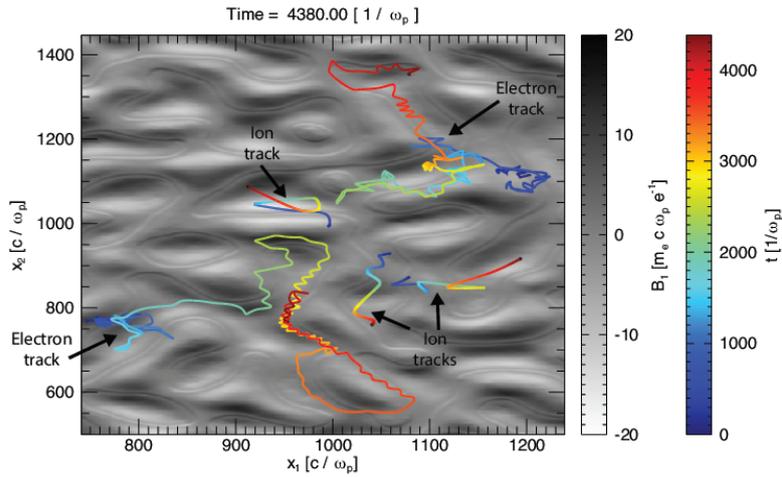


Figure 1. Magnetic field at $t = 4380 \omega_{pe}^{-1}$ (gray) and particle tracks (colour).

3 Formation of Large-Scale Magnetic Fields

The origin of magnetic fields starting from unmagnetised plasmas is an important topic of modern research. Although much of the universe is magnetised (typically of order $10^{-6} G$) such that the magnetic field plays an important role in the dynamics, in the early universe this was not so. During the period of the cosmic microwave background, before recombination, it is widely accepted that there was no magnetic field⁸. Magnetic field growth is generally attributed to the turbulent dynamo^{9,10}, which amplifies a required initial seed field. The source of this initial seed has been speculated to come from the Biermann battery mechanism¹¹, or microinstabilities such as the Weibel² or filamentation instabilities¹.

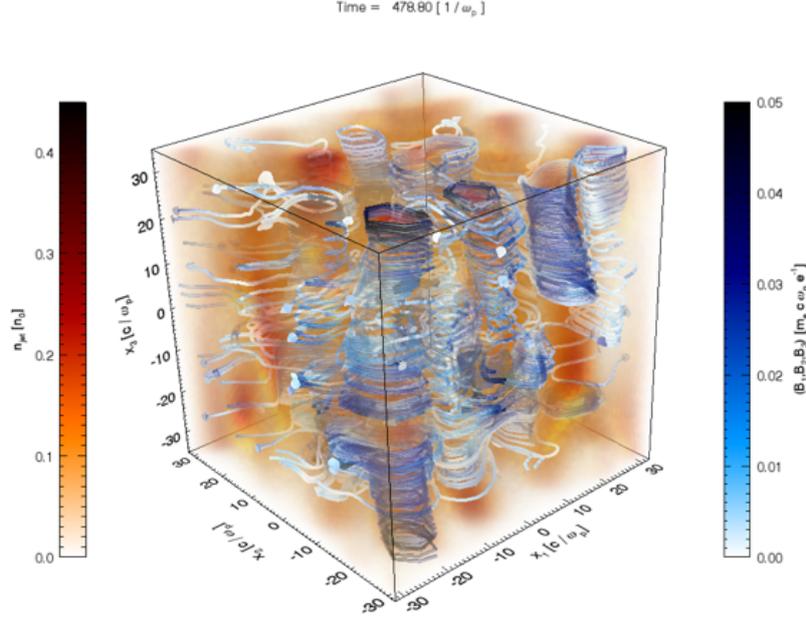


Figure 2. A still frame (at $t\omega_{pe} = 235.2$) from a movie showing the density of the fast moving jet of electrons, and a selection of magnetic field lines which surround the current filaments which have been formed at the electron inertial scale (kd_e) and begun merging to larger scales.

Fields generated by the Weibel instability have the advantage that the saturated fields reach magnetic pressures of the order of the plasma pressure ($\beta \equiv 8\pi P/B^2 \sim 1$) consistent with the $10^{-6}G$ fields seen today, and thus not much amplification is required. A big open question, however, is how these magnetic fields which form on very small spatial scales could change to the much longer observed scales. Besides the question of the origin of astrophysical magnetic fields, which vary over astrophysical length scales, this problem of scales has direct importance in the context of laser-solid experiments. In particular, experiments performed at the Tata Institute of Fundamental Research, Mumbai¹² show turbulent spectra of magnetic fields that reach scales far surpassing the microscopic scales of the Weibel instability.

Using OSIRIS, we have performed 2D and 3D particle-in-cell simulations of the generation of magnetic field, due to the Weibel instability. We modelled a uniform plasma where a percentage of the electrons consist of a stream of electrons flowing near the speed of light, and the remainder flow backward to cancel the current. This setup approximates the hot electrons flowing into the target generated by the laser in the experiment. The stream of fast electrons breaks up into small current filaments (due to the Weibel/filamentation instability) with associated electron inertial scale magnetic fields (wavenumber, $kd_e \sim 1$), which merge in order to form larger scales (see Fig. 2). We can produce a distribution of k which can be compared with the experimental findings, for which we also find exciting

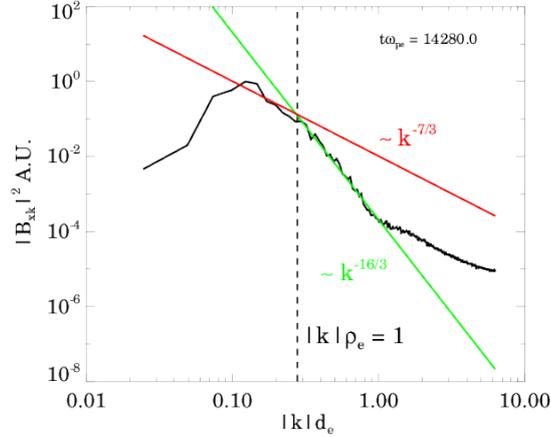


Figure 3. A plot of the magnetic field energy spectra at $t\omega_{pe} = 14280.0$, exhibiting two power law curves, $-7/3$ and $-16/3$, before and after the electron gyroradius scale ($k\rho_e$) respectively.

agreement with gyrokinetic power law predictions of the magnetic spectra¹³ (see Fig. 3).

Although we find excellent agreement in the spectra at the scales simulated, we have measured the merging rate and concluded that it is too slow to explain the experimental results. Therefore effects such as coupling with larger scale features, and flow generated turbulence (which may be required for dynamo) must play an important role in explaining the large scale magnetic fields observed in the laboratory and in nature. These effects, which can only be captured in 3D, provide exciting future work that can be performed using more computationally intensive 3D particle-in-cell simulations.

4 Microphysics of Relativistic Collisionless Shear Flows

Relativistic collisionless shear flows are pervasive in some of the most extreme astrophysical scenarios such as gamma-ray bursts, active galactic nuclei and blazars. While the stability of such shear configurations have been studied at a macroscopic (magnetohydrodynamic) level, the role of microphysical (electron-scale) effects have been overlooked. In addition to providing dissipation on the micro-scale, the operation of microphysical processes and instabilities in shear scenarios can strongly modify the particle distributions (in space and momentum) of the plasma, impacting the long-term macroscopic evolution of such systems.

Using the particle-in-cell code OSIRIS³ we have performed large-scale ($\sim 10^5$ core hours) multidimensional simulations of relativistic collisionless shearing plasma flows, and have identified a fast-growing electron-scale instability that leads to the development of mushroom-like structures in the electron-density, transverse to the flow direction; we have labelled this process the Mushroom instability (MI)¹⁴. Our numerical simulations have shown that the MI efficiently dissipates the immense kinetic energy available in relativistic shear flow scenarios, transforming it into electric and magnetic field energy, and

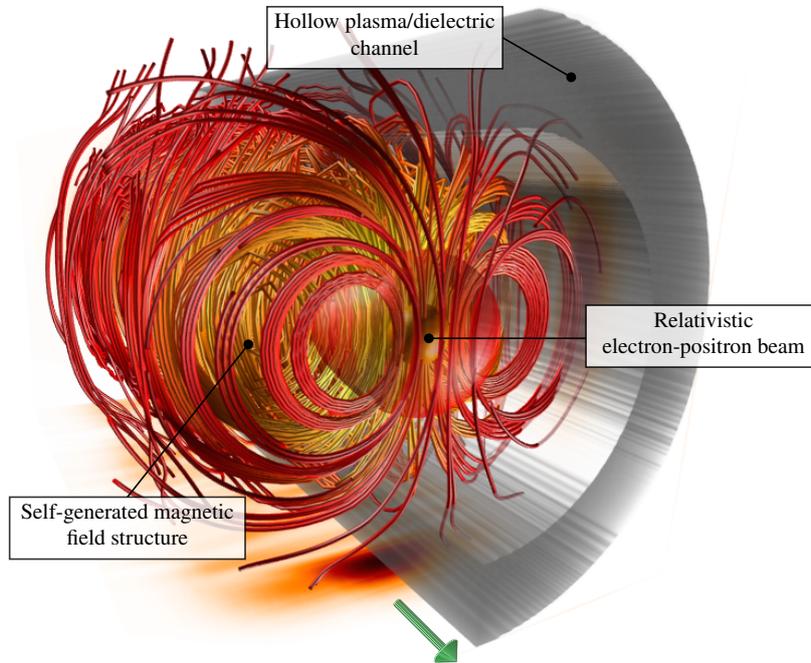


Figure 4. 3D PIC simulation of the interaction between a relativistic, globally neutral e^-e^+ beam (red/orange isosurfaces) and a hollow plasma channel (grey isosurfaces) whose inner diameter is $4\times$ wider than the transverse waist of the beam, to guarantee weak overlap (no contact) between the two; half of the cylindrical structure has been omitted to more clearly illustrate the dynamics of the beam inside. The e^-e^+ beam propagates in the positive x direction. The red lines represent the structure of the MI-generated magnetic field.

nonthermal acceleration of particles and radiation.

In an attempt to identify a laboratory configuration where the development of the MI can be verified, we have explored the dynamics of a relativistic, globally neutral electron-positron (e^-e^+) beam propagating through a hollow plasma channel. We have performed large-scale 3D PIC simulations ($\sim 10^5$ core hours) to investigate the propagation of relativistic e^-e^+ beams (with characteristics that are currently available at SLAC National Accelerator Laboratory) in hollow plasma channels with different diameters, and we have identified the development of the MI even in the absence of overlap (no contact) between the beam and the walls of the hollow channel (Fig. 4)¹⁵. We believe that this configuration reproduces the relativistic collisionless shear flow conditions that are relevant to extreme astrophysical shear scenarios, such as the shear interaction between relativistic astrophysical jets and the interstellar medium.

5 Brillouin Amplification: Towards Next Generation Laser Energy Densities

High-intensity lasers beams allow for the creation of high-energy-density (HED) conditions, that mimic extreme astrophysical scenarios, in controlled laboratory settings. These

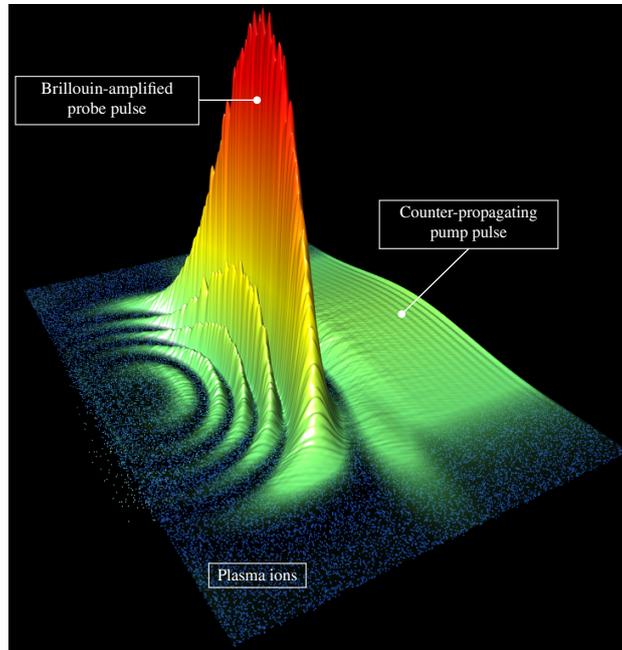


Figure 5. PIC simulation of successful Brillouin amplification. Contour surfaces represent the intensity envelopes of the counter propagating laser pulses, and the spheres represent the plasma ions that oscillate in the ion-acoustic wave. Initial pump and probe intensities are both 10^{16} W/cm² and the plasma density is 30% critical. It is shown that the probe pulse efficiently picks up the pump laser energy, reaching a peak intensity of 1.5×10^{17} W/cm² (15x the pump intensity) while preserving a smooth envelope. Transverse modulations in the probe envelope are observed due to the ponderomotive filamentation instability, but these still remain at a tolerable level at this stage.

exciting experiments give us a closer look at complex astrophysical processes that would otherwise have to be inferred indirectly by their radiation collected by our telescopes. Advances in the generation of higher laser powers and intensities is highly desirable, since these will give access to novel astrophysical conditions and exotic physical regimes of HED science, such as “boiling the vacuum”.

Further increasing current state-of-the-art laser energy densities is highly expensive using solid state optics. Plasma-based laser amplifiers, leveraging on parametric processes like stimulated Brillouin backscattering, overcome the optical damage limitations of solid state optics by several orders of magnitude, providing a promising alternative to the production of next generation laser energy densities.

We have performed large-scale ($\sim 10^6$ core hours), multidimensional PIC simulations in order to understand the main limiting parasitic instabilities (ponderomotive filamentation and Raman scattering instabilities) for Brillouin amplification over a wide range of parameters, including different plasma densities and laser intensities. With the aid of the large-scale PIC simulations and analytic theory, we have identified the optimal parameter regime where parasitic effects are minimised, allowing for efficient Brillouin amplification (Fig. 5). Our work is critical for the design of future Brillouin amplification experiments.

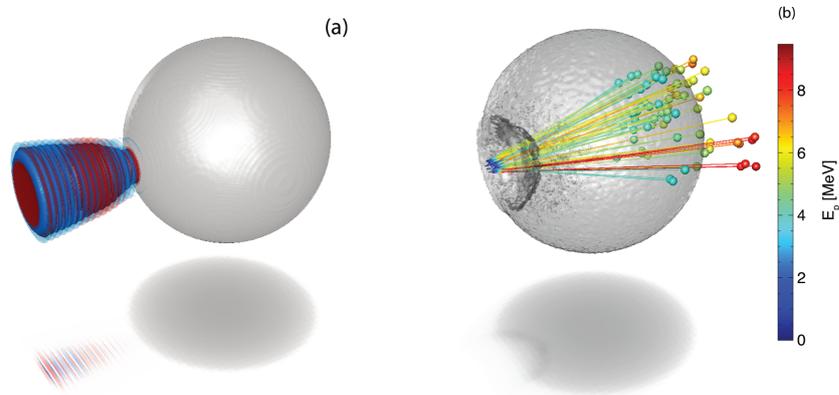


Figure 6. (a) Initial setup of the experiment: a frozen hydrogen pellet is irradiated by a laser. (b) Ions are accelerated through the target.

6 Generation of Energetic Ions with Collisionless Shocks

There are many applications which require energetic ions, e. g. for plasma diagnostics, fusion or medical physics. This acceleration can be efficiently done by irradiating a plasma with a laser. Depending on the actual target setup, different acceleration mechanisms can dominate. With 3D particle-in-cell simulations, we studied the parameter conditions for shock acceleration, which is very efficient in terms of final energy output¹⁶. Fig. 6a shows the setup of the experimental configuration: a frozen hydrogen pellet is irradiated by two laser pulses. The first pulse reduces the target density due to target expansion and the second pulse initiates the acceleration process. On the one hand, the laser has to be well focused in order to guarantee a high laser potential at the target surface, while on the other hand, a wide laser beam would be beneficial for producing a plane shock for a focused ion beam. The transport of the ions from the target surface through the target is shown in Fig. 6b, where only the paths of the most energetic ions have been selected. The density accumulation due to the laser and with the shock ahead (further inside the target), can be observed as well.

7 Concluding Remarks

The global topic of our research project is the investigation of collisionless shock formation. In this context, the mediating plasma micro-instabilities play a key role. We presented five sub-projects in which we have investigated the shock formation process, the evolution of electromagnetic instabilities on different scales, related shear flows and their applications in the laboratory. High performance supercomputing resources are absolutely necessary in order to understand the full physical picture. It is possible to analyse very small spatial scales with a high temporal resolution and to follow the evolution of non-

linear processes on long time scales. Furthermore, idealised setups allow to simplify the physics and to identify separate and basic processes.

Acknowledgements

The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN¹⁷ at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

References

1. B. D. Fried, *Mechanism for Instability of Transverse Plasma Waves*, Phys. Fluids, **2**, 337, 1959.
2. E. S. Weibel, *Stable Orbits of Charged Particles in an Oscillating Electromagnetic Field*, Phys. Rev., **114**, 18, 1959.
3. R. A. Fonseca, L. O. Silva, F. S. Tsung, V. K. Decyk, W. Lu, C. Ren, W. B. Mori, S. Deng, S. Lee, T. Katsouleas, and J. C. Adam, *OSIRIS: A Three-Dimensional, Fully Relativistic Particle in Cell Code for Modeling Plasma Based Accelerators*, Lecture Notes in Comput. Sci., **2331**, 342, 2002.
4. R. A. Fonseca, S. F. Martins, L. O. Silva, J. W. Tonge, F. S. Tsung, and W. B. Mori, *One-to-one direct modeling of experiments and astrophysical scenarios: pushing the envelope on kinetic plasma simulations*, Plasma Phys. Controlled Fusion, **50**, 124034, 2008.
5. A. Bret, A. Stockem, F. Fiuza, C. Ruyer, L. Gremillet, R. Narayan, and L. O. Silva, *Collisionless shock formation, spontaneous electromagnetic fluctuations, and streaming instabilities*, Phys. Plasmas, **20**, 042102, 2013.
6. A. Bret, A. Stockem, R. Narayan, and L. O. Silva, *Collisionless Weibel shocks: Full formation mechanism and timing*, Phys. Plasmas, **21**, 072301, 2014.
7. A. Stockem Novo, A. Bret, R. A. Fonseca, and L. O. Silva, *Shock Formation in Electron-Ion Plasmas: Mechanism and Timing*, Astrophys. J., **803**, L29, 2015.
8. R. M. Kulsrud and E. G. Zweibel, *On the origin of cosmic magnetic fields*, Rep. Prog. Phys., **71**, 046901, 2008.
9. R. M. Kulsrud and S. W. Anderson, *The spectrum of random magnetic fields in the mean field dynamo theory of the Galactic magnetic field*, Astrophys. J., **396**, 606, 1992.
10. A. Brandenburg, D. Sokoloff, and K. Subramanian, *Current Status of Turbulent Dynamo Theory. From Large-Scale to Small-Scale Dynamos*, Space Sci. Rev., **169**, 123, 2012.
11. L. Biermann, *Saturation Mechanisms for the Generated Magnetic Field in Nonuniform Laser-Matter Irradiation*, Z. Naturforsch., **5a**, 65, 1950.

12. Sudipta Mondal, V. Narayanan, Wen Jun Ding, Amit D. Lad, Biao Hao, Saima Ahmad, Wei Min Wang, Zheng Ming Sheng, Sudip Sengupta, Predhiman Kaw, Amita Das, and G. Ravindra Kumar, *Direct observation of turbulent magnetic fields in hot, dense laser produced plasmas*, Proceedings of the National Academy of Sciences, **109**, 8011, 2012.
13. A. A. Schekochihin, S. C. Cowley, W. Dorland, G. W. Hammett, G. G. Howes, E. Quataert, and T. Tatsuno, *Astrophysical Gyrokinetics: Kinetic and Fluid Turbulent Cascades in Magnetized Weakly Collisional Plasmas*, Astrophys. J. Suppl. Ser., **182**, 310, 2009.
14. R. A. Fonseca, E. P. Alves, T. Grismayer and L. O. Silva, Physical Review E, **92**, 021101(R), 2015.
15. E. P. Alves, T. Grismayer, M. G. Silveirinha, R. A. Fonseca, and L. O. Silva, submitted for publication in Plasma Phys. Control. Fusion, 2015.
16. A. Stockem Novo, C. A. Kaluza, R. A. Fonseca, and L. O. Silva, in prep.
17. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities, **1**, A1, 2015, <http://dx.doi.org/10.17815/jlsrf-1-18>.

NIC Series Volume 32

NIC Symposium 2006 - Proceedings

1 - 2 March 2006, Jülich, Germany

edited by G. Münster, D. Wolf, M. Kremer (2006), iv, 365 pages

ISBN: 3-00-017351-X

NIC Series Volume 33

Parallel Computing: Current & Future Issues of High-End Computing

Proceedings of the International Conference ParCo 2005

edited by G. R. Joubert, W. E. Nagel, F. J. Peters, O. Plata, P. Tirado, E. Zapata (2006), xiv, 930 pages

ISBN: 3-00-017352-8

NIC Series Volume 34

**From Computational Biophysics to Systems Biology 2006
Proceedings**

edited by U. H. E. Hansmann, J. Meinke, S. Mohanty, O. Zimmermann (2006),
iv, 214 pages

ISBN-10: 3-9810843-0-6, ISBN-13: 978-3-9810843-0-6

NIC Series Volume 35

**Dreistufig parallele Software zur Parameteroptimierung von
Support-Vektor-Maschinen mit kostensensitiven Gütemaßen**

by T. Eitrich (2007), xiv, 231 pages

ISBN: 978-3-9810843-1-3

NIC Series Volume 36

**From Computational Biophysics to Systems Biology (CBSB07)
Proceedings**

edited by U. H. E. Hansmann, J. Meinke, S. Mohanty, O. Zimmermann
(2007), vi, 312 pages

ISBN: 978-3-9810843-2-0

NIC Series Volume 37

**Parallel Computing: Architectures, Algorithms and Applications -
Book of Abstracts of the International Conference ParCo 2007**

ParCo 2007 Conference, 4 - 7 September 2007

edited by G. R. Joubert, C. Bischof, F. J. Peters, T. Lippert, M. Bucker,
P. Gibbon, B. Mohr (2007), xiv, 195 pages

ISBN: 978-3-9810843-3-7

NIC Series Volume 38

**Parallel Computing: Architectures, Algorithms and Applications -
Proceedings of the International Conference ParCo 2007**

edited by C. Bischof, M. Bucker, P. Gibbon, G. R. Joubert, T. Lippert, B. Mohr,
F. Peters (2007), xx, 810 pages

ISBN: 978-3-9810843-4-4

NIC Series Volume 39

NIC Symposium 2008 - Proceedings

20 - 21 February 2008, Jülich, Germany

edited by G. Münster, D. Wolf, M. Kremer (2008), iv, 358 pages

ISBN: 978-3-9810843-5-1

NIC Series Volume 40

From Computational Biophysics to Systems Biology (CBSB08)

Proceedings

edited by U. H. E. Hansmann, J. H. Meinke, S. Mohanty, W. Nadler,

O. Zimmermann (2008), viii, 430 pages

ISBN: 978-3-9810843-6-8

NIC Series Volume 41

**Multigrid methods for structured grids and their application
in particle simulation**

by M. Bolten (2008), viii, 132 pages

ISBN: 978-3-9810843-7-5

NIC Series Volume 42

Multiscale Simulation Methods in Molecular Sciences - Lecture Notes

Winter School, 2 - 6 March 2009, Forschungszentrum Jülich

edited by J. Grotendorst, N. Attig, S. Blügel, D. Marx (2009), vi, 576 pages

ISBN: 978-3-9810843-8-2

NIC Series Volume 43

Towards the Confirmation of QCD on the Lattice

Improved Actions and Algorithms

by S. F. Krieg (2008), vi, 89 pages

ISBN: 978-3-9810843-9-9

NIC Series Volume 44

NIC Symposium 2010 – Proceedings

24 - 25 February 2010 | Jülich, Germany

edited by G. Münster, D. Wolf, M. Kremer (2012), v, 395 pages

ISBN: 978-3-89336-757-3

NIC Series Volume 45

NIC Symposium 2012 – Proceedings

25 Years HLRZ / NIC

7 - 8 February 2012 | Jülich, Germany

edited by K. Binder, G. Münster, M. Kremer (2012), v, 400 pages

ISBN: 978-3-89336-758-0

NIC Series Volume 46

**Hybrid Particle-Continuum Methods in Computational Materials Physics
Proceedings**

4 - 7 March 2013 | Jülich, Germany

edited by M. H. Müser, G. Sutmann, R. G. Winkler (2013), ii, 232 pages

ISBN: 978-3-89336-849-5

NIC Series Volume 47

NIC Symposium 2014 - Proceedings

12 - 13 February 2014 | Jülich, Germany

edited by K. Binder, G. Münster, M. Kremer (2014), vi, 434 pages

ISBN: 978-3-89336-933-1

NIC Series Volume 48

NIC Symposium 2016 - Proceedings

11 - 12 February 2016 | Jülich, Germany

edited by K. Binder, M. Müller, M. Kremer, A. Schnurpfeil (2015), vi, 417 pages

ISBN: 978-3-95806-109-5

The John von Neumann Institute for Computing (NIC) was established in 1998 by Forschungszentrum Jülich and Deutsches Elektronen-Synchrotron DESY to support the supercomputer-oriented simulation sciences. In 2006, GSI Helmholtzzentrum für Schwerionenforschung joined NIC as a contract partner.

The core task of NIC is the peer-reviewed allocation of supercomputing resources to computational science projects in Germany and Europe. The NIC partners also support supercomputer-aided research in science and engineering through a three-way strategy:

- Provision of supercomputing resources for projects in science, research, and industry.
- Supercomputer-oriented research and development by research groups in selected fields of physics and natural sciences.
- Education and training in all areas of supercomputing by symposia, workshops, summer schools, seminars, courses, and guest programmes for scientists and students.

The NIC Symposium is held biennially to give an overview on activities and results obtained by the NIC projects in the last two years. The contributions for this 8th NIC Symposium are from projects that have been supported by the supercomputers JUROPA, JURECA, and JUQUEEN in Jülich. They cover selected topics in the fields of Astrophysics, Biology and Biophysics, Chemistry, Elementary Particle Physics, Materials Science, Condensed Matter, Soft Matter Science, Earth and Environment, Computer Science and Numerical Mathematics, Fluid Mechanics, and Plasma Physics.



Deutsches
Elektronen-Synchrotron

