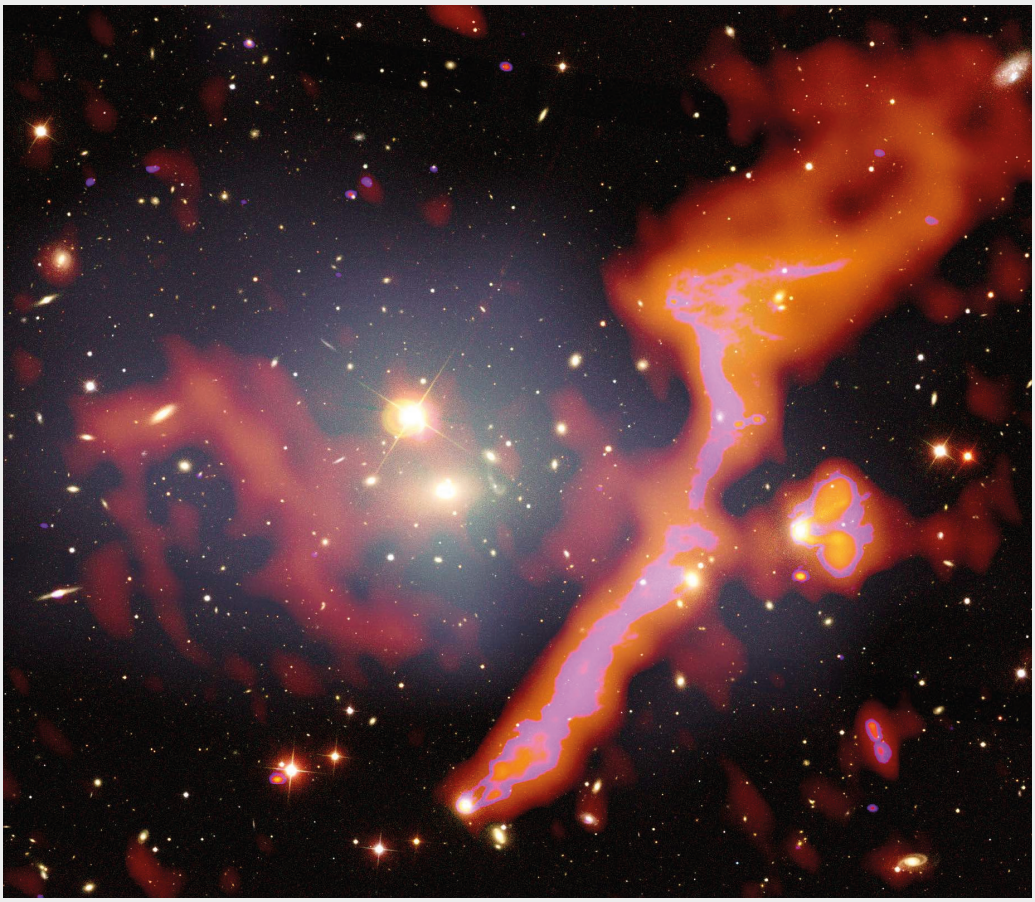
John von Neumann Institute for Computing

NIC Symposium 2020

27 – 28 February 2020 I Jülich, Germany

M. Müller, K. Binder, A. Trautmann (Editors)

Proceedings



NIC Series

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 50

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

NIC Symposium 2020

27 – 28 February 2020 I Jülich, Germany

M. Müller, K. Binder, A. Trautmann (Editors)

Proceedings

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 50

ISBN 978-3-95806-443-0

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte Bibliografische Daten
sind im Internet über http://dnb.d-nb.de abrufbar.

Herausgeber und Vertrieb:	Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag 52425 Jülich Tel.: +49 2461 61-5368 Fax: +49 2461 61-6103 zb-publikation@fz-juelich.de www.fz-juelich.de/zb
Umschlaggestaltung:	Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH
Druck:	Grafische Medien, Forschungszentrum Jülich GmbH
Copyright:	Forschungszentrum Jülich 2020

We thank A. Wilber/LOFAR Surveys Team for the image used in the cover design.

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 50

ISBN 978-3-95806-443-0

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER)
unter www.fz-juelich.de/zb/openaccess.

Preface

Marcus Müller
Institute for Theoretical Physics, Georg-August University, 37077 Göttingen, Germany
mmueller@theorie.physik.uni-goettingen.de

Kurt Binder
Institute for Physics, Johannes-Gutenberg University, 55099 Mainz, Germany
kurt.binder@uni-mainz.de

Alexander Trautmann
John von Neumann Institute for Computing, Jülich Supercomputing Centre,
Forschungszentrum Jülich, 52425 Jülich, Germany
a.trautmann@fz-juelich.de

On February 27 and 28, 2020 computational scientists will present their exciting research results at the 10^{th} NIC symposium in Jülich. About 20 years after the first NIC symposium on December 5 and 6, 2001, this series of biannual meetings has established a valued tradition, highlighting a diverse range of some of the best, modern, computational science at the von Neumann Institute for Computing (NIC). The compilation of recent activities in this accompanying proceedings showcases the extraordinarily broad scope of research on supercomputers, ranging from fundamental aspects of physics such as elementary particle physics, nuclear physics, astrophysics, and statistical physics of hard and soft condensed matter, as well as computational chemistry and life sciences to applied disciplines such as materials physics, fluid-dynamics engineering, and climate research. The symposium and proceedings address both, computational scientists and practitioners as well as the general public that is interested in the advancement of computational science and its applications in diverse, contemporary research fields.

In 1998 the NIC has been founded by the Forschungszentrum Jülich (FZJ) and the Deutsches Elektronen-Synchrotron (DESY), and in 2006 the Helmholtzzentrum für Schwerionenforschung (GSI) joined. Ever since the NIC has served the community by providing state-of-the-art supercomputer resources, training and technical support. Within the framework of the Gauss Centre for Supercomputing (GCS), the federal government, and the states North Rhine-Westphalia, Baden-Württemberg, and Bavaria committed 450 Million Euro until 2025 for the development of supercomputing in Germany with equal shares to the three Tier-1 centres, the Jülich Supercomputing Centre (JSC), the High Performance Computing Center Stuttgart (HLRS), and the Leibniz Supercomputing Centre (LRZ) in Munich. The rapid development of computational science is impressively visible in the hardware development: When the NIC was founded, it provided the community with computing resources of 1 tera-flop peak performance through a parallel Cray T3E system and a Cray SV1ex vector computer. In June 2018, the Jülich Wizard for European Leadership Science (JUWELS) – a versatile cluster module of 2500 dual Intel Xeon Platinum nodes with 2×24 cores each, as well as 56 nodes, equipped with 4 Nvidia

V100 Graphics Processing Units (GPUs) – has been installed at JSC, providing a peak performance of 12 peta-flop – 12 000 times faster than in 1998. In November 2019 this module is placed at rank 31 on the TOP500 list of supercomputers in the world. Thus, it remains challenging to keep up with the fast pace of development in the U.S.A., China, and Japan, as well as other European countries. Continuous support will be crucial to position the NIC (and the entire GCS) in the European High-Performance Computing Joint Undertaking (EuroHPC) that aims at developing a pan-European exascale supercomputing infrastructure.

In November 2019, the JSC announced the significant extension of JUWELS by a booster module (to be available to NIC users in fall 2020), where most of the nodes will be accelerated by Nvidia GPUs. The peak performance of this modular supercomputer, comprised of cluster and booster, is expected to surpass 70 peta-flop. This new concept provides the scientific community simultaneously with an all-round, easy-to-use component and a component for particularly compute-intensive program parts or for analysing big scientific data. This universal tool will push the frontiers of established simulation techniques, *e. g.* in elementary particle physics or extremely large, particle-based simulations in materials physics and engineering and also enable new applications employing, *e. g.* strategies of machine learning.

With these new opportunities come new challenges such as the efficient parallel use of multiple GPUs on a large scale and the development of strategies for efficiently extracting information of large data sets. The training and much appreciated user support that the JSC offers to NIC users play a key role in enabling the community to seize these exciting opportunities. Also the NIC symposium will not only provide a fascinating account of supercomputing-based research at its best, but it will additionally foster cross-fertilisation with respect to application of algorithms and computational strategies in different fields of natural science and engineering. These efforts will also prepare the NIC user community to contribute to and benefit from the Nationale Forschungsdateninfrastruktur (NFDI) that emphasises the efficient (re)use of information contained in the large amount of scientific data, produced by experimental and computational science.

Only a fraction of all projects at the NIC are compiled in this proceedings volume.[a] The contributions in the proceedings and in the NIC symposium have been selected by the reviewers and represent the diversity of outstanding research enabled by the NIC supercomputing resources. The computing time is granted by a stringent peer-review system, considering only the scientific quality of the research. The NIC peer-review board, headed by Prof. D. Wolf, and the international pool of expert reviewers play a crucial role in sustaining the quality of the projects at this very high level and in fairly and effectively allocating the valuable resources. Therefore it is a great pleasure to sincerely thank the distinguished reviewers for their engagement and time that they invest in fulfilling this vital task.

Semiannually, the NIC peer-review board selects one “NIC excellence project” [b] out of

[a] `http://www.john-von-neumann-institut.de/nic/ProjectDistribution`
[b] `http://www.john-von-neumann-institut.de/nic/ExcellenceProjects`

all proposals. These particularly noteworthy projects are compiled in a special section, featuring the following projects:

- M. Moseler, Fraunhofer Institute for Mechanics of Materials, Freiburg – Quantum Chemical Modelling of Water Splitting: From Photoelectrochemistry to Superlubricity
- S. Trebst, University of Cologne – Machine Learning Transport Properties in Quantum Many-Fermion Simulations
- F. Lechermann, University of Hamburg – Cooperation of Many-Body Physics and Defect Chemistry in Transition-Metal Oxides
- L. Pastewka, University of Freiburg – SDS Adsorptions Films at the $H_2O-Au(111)$ Interface: Molecular Dynamics Study of AFM Tip–Surface Contact

The NIC also supports 5 research groups at the member institutions.[c] They cover a spectrum of research disciplines, provide key contributions to the research focus of the respective Helmholtz centre, and develop and apply sophisticated simulation strategies for these important research topics. Their recent results are highlighted in a special chapter comprised of the contributions of Prof. M. Bleicher, GSI Darmstadt, on nuclear matter, Prof. O. Philipsen, Goethe University Frankfurt, on lattice QCD, Prof. R. Sommer, DESY Zeuthen, on elementary particles physics as well as Prof. H. Gohlke, JSC Jülich, on computational biophysical chemistry and Prof. A. Schug, JSC Jülich, on computational structural biology.

We hope that you will enjoy reading about the latest research highlights at the NIC that were made possible by the excellent computing environment and user services that the JSC provides to the NIC community. The contributions are arranged by scientific topics, and it is a pleasure to thank all the authors who have contributed to this book for their careful work. We also thank the eminent experts who wrote the introduction to the different chapters for providing a perspective on the wide spectrum of applications of modern supercomputing to diverse scientific disciplines.

Moreover, we are very grateful to Ms. Martina Kamps who compiled all the texts and produced this high quality book. Finally, we want to thank Dr. Florian Janetzko, Ms. Britta Hoßfeld, Ms. Sarah Hansen, and Ms. Gina Bongartz for their valuable help in organising the 10^{th} NIC Symposium in Jülich.

Jülich, February 2020

M. Müller K. Binder A. Trautmann

[c] http://www.john-von-neumann-institut.de/nic/ResearchGroups

Contents

Developing Exascale Computing at JSC
E. Suarez, W. Frings, S. Achilles, N. Attig, J. de Amicis, E. Di Napoli, Th. Eickermann, E. B. Gregory, B. Hagemeier, A. Herten, J. Jitsev, D. Krause, J. H. Meinke, K. Michielsen, B. Mohr, D. Pleiter, A. Strube, Th. Lippert **1**

The NIC Research Groups

A Study of the Kinetic Freeze-Out in Relativistic Heavy Ion Collisions with the UrQMD/Coarse Graining Approach
G. Inghirami, P. Hillmann, B. Tomášik, M. Bleicher **23**

Cut-Off Effects on the QCD Thermal Transition as a Function of Quark Masses and Chemical Potential
F. Cuteri, O. Philipsen, A. Schön, A. Sciarra **33**

Heavy Quarks, Multi-Level Algorithms and Tensor Networks: Developing Methods for Non-Perturbative Aspects of Quantum Field Theories
K. Jansen, S. Schaefer, H. Simma, R. Sommer **43**

Simulating Thioflavin T and Congo Red Binding to the Fibril Structure of Amyloid-β(1-42)
B. Frieg, H. Gohlke **53**

Massively Parallel Large-Scale Multi-Model Simulation of Tumour Development Including Treatments
M. Berghoff, J. Rosenbauer, A. Schug **63**

The NIC Excellence Projects

Quantum Chemical Modelling of Water Splitting: From Photoelectrochemistry to Superlubricity
L. Mayrhofer, G. Moras, T. Kuwahara, A. Held, M. Moseler **75**

Machine Learning Transport Properties in Quantum Many-Fermion Simulations
C. Bauer, S. Trebst **85**

Cooperation of Many-Body Physics and Defect Chemistry in Transition-Metal Oxides
F. Lechermann **93**

SDS Adsorptions Films at the $H_2O-Au(111)$ Interface: Molecular Dynamics Study of AFM Tip–Surface Contact
J. Hörmann, L. Pastewka **101**

Astrophysics

Introduction
P. L. Biermann **111**

Realising the LOFAR Two-Metre Sky Survey
A. Drabent, M. Hoeft, A. P. Mechev, J. B. R. Oonk, T. W. Shimwell, F. Sweijen, A. Danezi, C. Schrijvers, C. Manzano, O. Tsigenov, R.-J. Dettmar, M. Brüggen, D. J. Schwarz **113**

A Song of Shocks and Dynamo: Numerical Studies of a Galaxy Cluster Merger in the HIMAG Project
D. Wittor, P. Domínguez-Fernández, F. Vazza, M. Brüggen **123**

Computational Biology and Biophysics

Introduction
Y. Goldscheider, W. Wenzel **135**

Conformational Analysis of Dual-Scale Simulations of Ubiquitin Chains
A. Berg, C. Peter **137**

Solvent Effects on the Binding of Fatty Acids to Human Serum Albumin
P. Pant, Y. B. Ruiz-Blanco, E. Sanchez-Garcia **147**

Elementary Particle Physics

Introduction
S. Schaefer **159**

Nuclear Lattice Effective Field Theory: Status and Perspectives
U.-G. Meißner for the NLEFT Collaboration **161**

Flavour Structure of the Baryon Octet
G. S. Bali, S. Collins for the RQCD Collaboration **175**

Scale Setting for QCD with Four Dynamical Quarks
R. Höllwieser, F. Knechtli, T. Korzec **185**

QCD Phase Transition in the Chiral Limit
O. Kaczmarek, F. Karsch, A. Lahiri, L. Mazur, C. Schmidt **193**

Higher Order Fluctuations and Correlations of Conserved Charges from Lattice QCD
S. Borsanyi, Z. Fodor, J. N. Guenther **207**

Materials Science

Introduction
R. O. Jones **219**

A First-Principles Study on the Role of Defects and Impurities in β-In_2S_3
E. Ghorbani, K. Albe **221**

Numerical Simulations of Wave Propagation: Time Reverse Imaging and Defect Mapping in Pipes
E. H. Saenger, C. Werner, L. T. Nguyen, G. K. Kocur, B. Ahrens **231**

Theoretical Condensed Matter

Introduction
K. Binder, F. B. Anders **245**

Optical Pumping of Electron Spins in Quantum Dot Ensembles
I. Kleinjohann, A. Fischer, N. Jäschke, F. B. Anders **247**

Benchmarking Supercomputers with the Jülich Universal Quantum Computer Simulator
D. Willsch, H. Lagemann, M. Willsch, F. Jin, H. De Raedt, K. Michielsen **255**

Numerical Simulations of Strongly Correlated Electron Systems
F. F. Assaad **265**

Computational Soft Matter Science

Introduction
K. Kremer **277**

Do We Understand the Collective Short-Time Dynamics in Multicomponent Polymer Melts?
M. Müller, G. Wang, Y. Ren **279**

Molecular Dynamics Simulations of Curcumin in the Interface Region of Triblock Copolymer Micelles
K. P. K. Luft, S. Gekle **289**

Earth and Environment

Introduction
T. Jung **299**

Data Assimilation with the Integrated Terrestrial System Model TSMP-PDAF
H.-J. Hendricks Franssen, W. Kurtz, B. Naz, S. Gebler **301**

Uncertainty in Terrestrial Water Cycle Simulations
P. Shrestha, M. Sulis, S. Poll, T. Hoar, H.-J. Hendricks Franssen, C. Simmer, S. Kollet **311**

Applications of Seismic Full-Waveform Inversion on Shallow-Seismic and Ultrasonic Data
T. Bohlen, Y. Pan, J. Müller **321**

Computer Science and Numerical Mathematics

Introduction
D. Kröner **333**

Massively Parallel Multigrid with Direct Coarse Grid Solvers
M. Huber, N. Kohl, P. Leleux, U. Rüde, D. Thönnes, B. Wohlmuth **335**

BEAM-ME: Accelerating Linear Energy Systems Models by a Massively Parallel Interior Point Method
T. Breuer, M. Bussieck, K.-K. Cao, F. Fiand, H. C. Gils, A. Gleixner, D. Khabi, N. Kempke, T. Koch, D. Rehfeldt, M. Wetzel **345**

Fluid Mechanics and Engineering

Introduction
C. Stemmer **355**

Machine Learning Applications in Convective Turbulence
R. Kräuter, D. Krasnov, A. Pandey, C. Schneide, K. Padberg-Gehle, D. Giannakis, K. R. Sreenivasan, J. Schumacher **357**

High-Performance Flow Simulation and Scale-Adaptive Turbulence Modelling of Centrifugal Pumps
M. Hundshagen, N. Casimir, A. Pesch, R. Skoda **367**

Sub-Grid Scale Modelling at Scale with Deep Learning and up to 60 Billion Degrees of Freedom
M. Bode, D. Denker, J. Jitsev, H. Pitsch **379**

Size, Shape, Density and Spectral Analysis of Nanoparticles by Analytical Ultracentrifugation Using UltraScan
M. J. Uttinger, S. E. Wawra, J. Walter, W. Peukert **389**

Computational Plasma Physics

Introduction
K.-H. Spatschek **403**

Collective Effect of Radiation Friction in Laser-Driven Hole Boring of Dense Plasma Targets
T. Liseykina, D. Bauer, S. Popruzhenko, A. Macchi **405**

Particle-in-Cell Simulation Studies for Hybrid Laser-Plasma Accelerators and Plasma Eyepieces
M. Zeng, A. Martinez de la Ossa, J. Osterhoff **415**

Developing Exascale Computing at JSC

Estela Suarez[1], Wolfgang Frings[1], Sebastian Achilles[2], Norbert Attig[1], Jacopo de Amicis[1], Edoardo Di Napoli[1], Thomas Eickermann[1], Eric B. Gregory[1], Björn Hagemeier[1], Andreas Herten[1], Jenia Jitsev[1], Dorian Krause[1], Jan H. Meinke[1], Kristel Michielsen[1], Bernd Mohr[1], Dirk Pleiter[1], Alexandre Strube[1], and Thomas Lippert[1]

[1] Jülich Supercomputing Centre (JSC), Institute for Advanced Simulation (IAS), Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {e.suarez, w.frings, th.lippert}@fz-juelich.de

[2] RWTH Aachen University, 52062 Aachen, Germany

The first Exaflop-capable systems will be installed in the USA and China beginning in 2020. Europe intends to have its own machines starting in 2023. It is therefore very timely for computer centres, software providers, and application developers to prepare for the challenge of operating and efficiently using such Exascale systems. This paper summarises the activities that have been going on for years in the Jülich Supercomputing Centre (JSC) to prepare the scientists and users for the arrival of Exascale computing. The Jülich activities revolve around the concept of modular supercomputing. They include both computational and data management aspects, ranging from the deployment and operation of large-scale computing platforms (*e. g.* the JUWELS Booster at JSC) and the federation of storage infrastructures (as for example the European data and compute platform Fenix), up to the education, training and support of application developers to exploit these future technologies.

1 The Exascale Challenge

The numerical simulation of scientific and industrial challenges has become a cornerstone of science and engineering. Sophisticated algorithms are required to deliver solutions that are accurate, fast, safe, and affordable. This traditional modelling is nowadays increasingly accompanied by and fused with data-driven modelling and simulation and complemented by artificial intelligence techniques. As a consequence, the current scientific evolution relies increasingly on advanced computing technologies to process, analyse, and translate information into knowledge and technological innovations.

This observation was the starting point for the latest Scientific Case[1] of PRACE, the Partnership for Advanced Computing in Europe, which conducted a detailed study on a number of areas of major scientific, technical and societal relevance for Europe, including fundamental sciences, climate, weather and Earth sciences, life sciences, energy, infrastructure and manufacturing, future materials and many more. In all these areas, Exascale computing will for sure be required to reach major breakthroughs, such as climate predictions with sub-kilometre resolution, the design and manufacturing of accurate digital twins for new materials, the understanding of molecular biology on biologically relevant time scales, the analysis and simulation of the human brain at the micrometre scale along with the translation of results to advanced medicine and brain inspired AI, up to fully simulating the formation of galaxies and black-holes to advance the understanding of our origin and future.

Exascale Missions

As a consequence of such analyses and conclusions being performed worldwide, High Performance Computing (HPC) systems able to execute 10^{18} floating operations per second will soon be deployed around the world. The first system is already planned for 2020 in China, with the US and Japan following in 2021. In Europe the EuroHPC Joint Undertaking (EuroHPC JU[a]) – an initiative supported already by 29 countries – aims to deploy two Exascale systems starting in 2023.

Germany's Gauss Centre for Supercomputing (GCS[b]) is a leading candidate in Europe to participate in EuroHPC's initiative for Exascale systems. The three GCS partners, together with their funding ministries, the German Federal Ministry of Education and Research (BMBF) and the corresponding ministries in the GCS partner hosting states Baden-Württemberg, Bavaria and North Rhine-Westphalia, are currently working on a participation model competitive under EuroHPC's regulations.

A large part of the mission of the JSC at Forschungszentrum Jülich (FZJ) is to collaborate with hardware and software vendors, including application experts, to conduct research in the field of supercomputer technologies by means of co-design. JSC's specific goal is to advance the novel Modular Supercomputer Architecture (MSA) for the benefit of future computing. The MSA relies on the principle of composability. On the one hand, it allows to integrate future computing modules, like quantum computers, on the other hand, its major advantage is its ability to operate disaggregated resources, which allows to include boosters of highest scalability that are more energy efficient, compute effective and affordable than previous technologies.

To achieve this, JSC has launched the Dynamical Exascale Entry Platform (DEEP[c]) series of EU-funded development projects already in 2011 and is leading them ever since. In addition, JSC is involved in the European Processor Initiative (EPI) for the creation of a European processor. As far as future computing technologies are concerned, JSC is very active in the European Quantum flagship and will host its OpenSuperQ prototype system. JSC's modular architecture and the entire data centre are designed to integrate high-end computing and data analytics with AI and quantum technologies in order to achieve highest flexibility and benefit from a EuroHPC Exascale supercomputer.

To give some examples as to Exascale applications, JSC has become the scalability spearhead of German Earth system modelling for both data and simulation applications, is a founding member of the European Human Brain Project (HBP), where it leads the HPC simulation and data activities, and is deeply liaising with materials sciences. For these fields, JSC creates joint research labs with the respective communities, anchored in its institutional funding program.

Operational Challenges

Let's take a look at the challenges an Exascale system poses to facility operations by inspecting and extrapolating the Top500 list of the fastest supercomputers worldwide. Summit – the number one at the time of this writing located at the Oak Ridge National Laboratory – achieves a peak performance of about 200 PFlop/s with a power consumption

[a] `https://eurohpc-ju.europa.eu`

[b] `https://www.gauss-centre.eu`

[c] `http://www.deep-projects.eu`

of 13 MW on about 520 m^2 usable floor space. A peak-Exaflop-system with the same technology would require more than 60 MW of electrical power and an area of $\sim$10 tennis courts (without GPUs the power and space requirements of the Summit technology would more than triple). Clearly, technological advancement improves power efficiency and integration density. So is the Japanese Fugaku system at RIKEN's R-CCS going into operation with about 500 PFlop/s in 2020, based on new far advanced Fujitsu ARM CPUs, estimated to consume between 30 and 40 MW, *i. e.* between 60 and 80 MW when extrapolated to ExaFlop/s.

Obviously, energy efficiency and integration density will be most crucial aspects of any Exascale system design. As a consequence, extensive infrastructure research and preparations will be needed. Moreover, measures to lower operational costs - *e. g.* through the use of warm water cooling technologies with free-outdoor cooling - will be mandatory. In these areas, JSC has collected broad experience through the operation of prototype systems (*e. g.* DEEP) and Petascale community servers (*e. g.* QPACE in the field of particle physics).

Heterogeneity Challenging Programmability

The emergence of cluster computing, starting as a disruptive technology around the year 2000 with components-off-the-shelf based on standard general purpose processors, has uplifted supercomputing from the Teraflop/s to the Petaflop/s era. However, due to the relatively high power consumption of general purpose CPUs relative to their peak performance, Exaflop-capable clusters built in this manner have extremely high energy requirements. The 3 to 5 times higher Flop/watt ratio of accelerator processors, however – of which GPUs are the most commonly used – made them very attractive and since 2010 has led to cluster systems that are heterogeneous on the node level, combining CPUs with accelerators, but are built as a homogeneous or monolithic system architecture. In fact, all published Exascale roadmaps foresee such heterogeneous nodes with monolithic system architectures – with the very new exception of the purely CPU-based Fugaku system (Japan) as to its node architecture. JSC's strategy for Exascale computing is unique: it is based on the MSA, which is able to combine CPUs and accelerated resources at system level (see Sec. 2).

Data *vs.* Flops

Despite all above arguments as to Exaflop-performance, it would be wrong to drive the Exascale race purely based on the number of floating point operations per second. Efficient data management is critical, especially now that deep learning (DL) is increasingly being used to generate knowledge directly from data rather than from theory or by modelling the data in order to simulate the phenomena based on such models.[1]

It is evident that data-driven approaches lead to applications requiring drastically increased data storage and I/O bandwidth capabilities. These codes and their users need new data-management services as well enabling them to exploit long-term storage and to share their data across sites and communities (Sec. 3 shows some developments driven at JSC to tackle these requirements). Data-intensive applications also benefit from hardware optimised to train neural networks in order to classify data and predict properties.

In particular, systems providing high performance for reduced precision (16 bit or 8 bit integer) operations in the form of tensor products on GPUs are very advantageous for such codes. These requirements, together with the need for high computing performance within a limited power budget call for modular system designs.

Application Challenges

It is becoming increasingly urgent that HPC applications take into account the complex architectures of (future) Exascale systems. A decisive factor will be the correct use of multi-level heterogeneity in parallel programming. The approaches implemented so far in a number of applications, which complement the MPI-based parallelisation with a further parallelisation level on the node, mostly via OpenMP, must be further pursued and extended. Another central goal is the inclusion of programming models for accelerators such as GPUs, for which the traditional node-level strategy still offers only limited support.

The expansion of JSC's JUWELS system with a GPU Booster mid of 2020 represents a first cornerstone for modularity (see Sec. 2.2 for a description of JUWELS, Sec. 4.2 for JSC's activities in GPU porting, and Sec. 4.4 for some examples of applications being adapted to heterogeneous architectures). JSC has designed the JUWELS MSA in such a way that the Booster can be operated as an autonomous, accelerated, homogeneous cluster system with nodes of CPU-driven GPUs. At the same time, the JUWELS Cluster with its CPU nodes can be operated autonomously. In addition, the two systems can be operated as a single machine, with the ability to dynamically reconfigure partitions that can interact with each other in workflows or even within one code on both modules. In this way, the JSC is approaching disaggregation and composability without compromising the advantages of traditional approaches.

Deep Learning is Transforming Scientific Computing

GPU-based acceleration is crucial for the strongly uprising field of DL.[14] Training the models on very large amounts of data requires running in parallel across hundreds or thousands of GPUs to be able to carry out training in reasonable amount of time. Different challenges arising from the demand to massively scale up data-driven DL model training are highlighted in Sec. 4.3. At JSC, distributed training of different DL models over hundreds of GPUs has been performed since 2017, reaching scaling close to linear without loss of task performance. Recently, Exascale-size distributed DL model training was demonstrated, using only GPUs.[2] Further development will be to transfer such successes of distributed training techniques to training of very large models that require heterogeneous accelerators, for instance, when coupling complex physics-based simulations or virtual environments and DL neural networks within closed-loop, end-to-end training pipelines. Modular Exascale supercomputers will provide the hardware and software infrastructure necessary for undertaking this effort.

The research focus and long-term agenda driven by JSC in the AI field aims on enabling large-scale self-supervised, multi-task and active continual learning. This line of research pursues to create methods capable of growing generic models from incoming streams of diverse data, extracting knowledge and skills quickly transferable across different tasks and domains – a grand and open scientific question. Executing and maintaining a continual

learning system over many weeks or months will be the task of future supercomputing centres and require performances and robust stability such as targeted by modular Exascale architectures.

The topics presented are intended to give an impression of the complexity and challenges of Exascale computing for data centres and their users. Aware of its responsibility as a candidate for the Exascale site, the JSC, together with user communities and partners from industry and academia, is making quite a number of efforts to prepare the operation of the first Exascale system.

2 The Modular Supercomputing Architecture

The JSC's strategy for Exascale is based on the MSA,[3,4] which has been developed in the DEEP series of EU-funded projects.[5] The MSA comprises a number of distinct compute modules (see Fig. 1). These modules can be potentially large parallel clusters of CPUs, CPU-accelerator nodes or AI-adapted nodes, whose hardware configuration meets the requirements of a particular type or part of an application, a large storage element, or a future computing system such as a quantum computer. Connecting all of these modules via a high-speed network and a unified software environment allows application codes to be distributed across different modules. This approach gives users full flexibility by allowing them to choose the right mix of computing resources. Of course such a system is well suited for workflows of different, increasingly complex, and emerging applications for a particular research issue.

To maximise the performance in this pool of computer resources, code adjustments are of course required. To support developers, JSC introduced specific support mechanisms to port code to acceleration technologies (especially GPUs) and promote the adoption of the MSA paradigm. In this context, a series of seminars on MSA began in 2019. Among the topics covered were the historical development of the MSA in the DEEP project series, the description of hardware systems used or to be used according to this paradigm, first application results on the modular JURECA system, specific use cases for machine and deep learning, *etc.* The aim of the seminar series was to raise awareness and pass on knowledge about the MSA to all members of the JSC. The high participation in the weekly presentations was a clear sign of the increased interest and triggered additional application porting activities around JSC codes (see some examples in Sec. 4.4), which serve as precursors for future actions for external code developers.

2.1 JURECA: The First Ever (Tier-1) Modular Supercomputer

The JURECA Cluster module put into operation in 2015 was completed with the installation of the Booster module in early 2018.[6] While the cluster uses multicore processors (Intel Xeon *Haswell*) and 100 Gb/s Mellanox (EDR) InfiniBand, the Booster uses multicore processors (Intel Xeon Phi *KNL*) and 100 Gb/s Intel Omni-Path. Bridging the two different high-speed connection technologies is possible in JURECA through a customised development in the ParTec ParaStation Software Suite[d] and is continuously researched and optimised.[7] While the use of different processor technologies for clusters and boosters is a

[d] `http://www.par-tec.com/products/parastationv5.html`

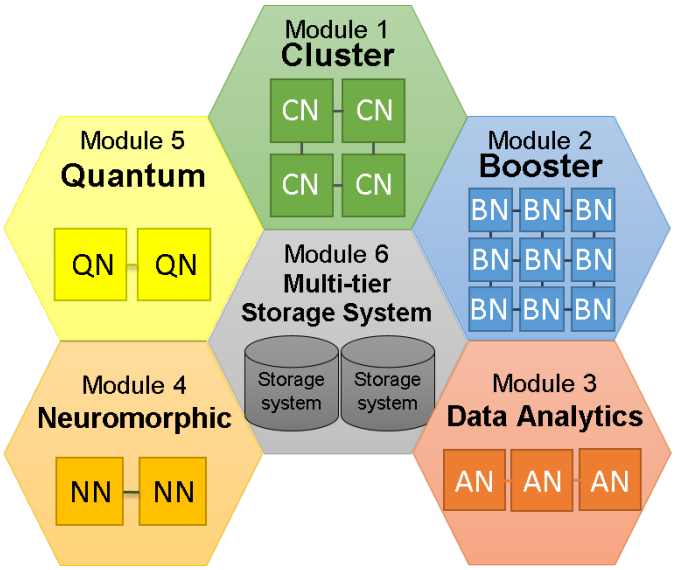


Figure 1. Sketch of the Modular Supercomputing Architecture. Note that this diagram does **not** represent a specific computer, but the general concept. Example modules: Cluster (*CN*: Cluster node), Booster (*BN*: Booster node), Data Analytics (*AN*: Data Analytics node), Neuromorphic (*NN*: Neuromorphic node), and Quantum (*QN*: Quantum node). The amount of modules in the future JSC Exascale platform and their configuration is subject to formal decisions.

key principle of MSA to optimise the highest scalability, it is also possible to benefit from MSA when considering the interaction of two or more partitions on the same machine.

2.2 JUWELS: The First Tier-0 Modular Supercomputer in the World

With the JUWELS system, JSC has further developed the MSA for the design of efficient, scalable and versatile high-performance platforms for national and European users.[8] It builds on the results of the DEEP projects and JURECA developments, with an emphasis on wider application. The Cluster module replaced the aging BG/Q system JUQUEEN in summer 2018, and with the Booster module planned for mid 2020, the modular system continues JSC's tradition of leadership class computing, particularly associated with the three past deployments of the BlueGene/L, /P and /Q generations. Provided by JSC as a GCS member, JUWELS serves as a national Tier 1 and European Tier 0 resource.

The configuration of both modules is summarised in Tab. 1. The JUWELS Cluster was designed to provide computing resources to a very wide range of application and to enable a swift and tight integration with the future Booster module, which targets workloads exposing a very high degree of parallelism.

The high-speed interconnects of Cluster and Booster are based on compatible Mellanox Infiniband interconnect technologies and will be directly integrated to enable easy and efficient inter-module MPI jobs without additional software routing. The system software stack based on the ParTec ParaStation Modulo is used to enable users to fully leverage the capabilities of the JUWELS modular supercomputer.

JUWELS	Cluster	Booster
Time of deployment	2018	2020
Project partner(s)	Atos, ParTec	Atos, ParTec, Nvidia, Mellanox
Node count	2559	>900
Peak performance	10.4 PFlop/s (plus 1.6 PFlop/s from GPU)	>60 PFlop/s
Accumulated mem. capacity	271.6 TB	>500 TB
Avg. power consum.	1.4 MW	2 MW (estimated)
CPU type	Intel Xeon "Skylake"	AMD EPYC "Rome"
Cores per processor	24	24
Accelerator	56 nodes with 4× Nvidia V100	4× Next-gen Nvidia GPUs per node
Network technology	Mellanox 100 Gb/s EDR-IB	Mellanox 200 Gb/s HDR-IB
Network topology	Fat-tree	DragonFly+
Cluster-Booster interface	Single global high-speed network domain	
Bandwidth ...		
... to capacity storage tier	250 GB/s	>400 GB/s
... to performance storage tier	>1 TB/s	>1 TB/s
Integration technology	Atos BullSequana X1000	Atos BullSequana XH2000
Cooling technology	Direct-liquid cooling with warm water (free cooling possible)	
System Software	ParaStation Modulo	
MPI	ParaStation MPI, Open MPI	
Job Scheduler	SLURM	
Resource Manager	SLURM + ParaStation psmgmt	
Others	Various compiler and optimisation libraries	

Table 1. Key features and characteristics of the modular Tier-0/1 supercomputer JUWELS.

It is again emphasised that the JUWELS modules can be operated autonomously without performance losses, *i. e.* as a pure CPU system for the Cluster and as a CPU-driven multi-GPU system for the Booster. The new dimension of modularity leads to the advantage of the combined use of both modules in overlapping workflows and split codes. This can be achieved with very little additional hardware effort.

2.3 From JUWELS to Exascale

The JUWELS system also serves as a prototype of JSC's vision for the first European Exascale system. Despite its much smaller size, JUWELS has the key components required to cross the Exascale limit in 2023 and shows an evolutionary path to an affordable and effective MSA-based Exascale device.

The exact configuration of the modular Exascale platform has not yet been determined, but as with JUWELS, it is expected that a cluster component with standard CPUs will be combined with a booster module characterised by an accelerator-oriented node design and a balanced network architecture. Further modules for data analysis applications could be integrated, while the possible use of disruptive technologies such as quantum computers and neuromorphic devices intended for inclusion in the modular system over a larger period of time is being investigated. A unified common software stack supports the execution of heterogeneous jobs across modules and facilitates applications to map their execution to hardware according to requirements. In addition, tight component integration and an efficient cooling approach will be critical to the cost-effective operation of the future Exascale platform. Following JUWELS' example, the Exaflop-computer will operate with high-density system integration technology and direct hot water cooling.

3 Data-Centric Activities

The Exascale era is by far not just about achieving a certain peak performance. In recent years, there has been a paradigm shift from computational to data-centric HPC applications, and scientists rely heavily on data services in their daily work. This applies both to large amounts of experimental data generated in large-scale facilities and to the mere exchange of everyday documents or intermediate data. The boundaries of these needs disappear over the data life-cycle and require tighter integration between data and HPC infrastructures. Important elements of this combined computing data environment are the secure delivery of the software environment required by data-driven applications, the homogenisation of access to data management services, the optimal organisation of storage diversity and storage technologies in advanced storage infrastructures, and the federation of such storage infrastructures across data centres to facilitate access to and sharing of data sources. The following subsections describe some of the activities carried out at JSC in these areas.

3.1 Containerisation

Containers – a lightweight virtualisation technology – are a convenient way to simplify the deployment of complex software stacks (*e. g.* for deep learning frameworks), increase portability for users using multiple systems, and place specific system software requirements over the bare-metal software stack of supercomputing centres. Today, JSC uses the runtime system *Singularity*[e] to support the execution of containers on its supercomputers. Singularity is a container runtime system designed with the security and functional requirements of HPC environments in mind. In early 2020, a comprehensive container infrastructure will be deployed and research into the most appropriate container support model for the future is underway. Options range from cloud-like *bring-your-own-container* to community-managed images. Pilot applications for containerised workload execution are currently being explored with selected JURECA and JUWELS projects and communities, and accessibility will be extended to the wider JSC user community by 2020.

3.2 OpenStack

JSC has complemented its Tier-0/1 HPC infrastructure with a versatile environment based on the OpenStack cloud platform, enabling users to deploy a variety of community-specific services, which can interact with data or even running applications in the existing HPC infrastructure. Most prominently, three types of infrastructure resources are available: compute, storage, and network.

To interact with the memory in the HPC domain, individual virtual machines of the OpenStack cluster can be connected to the HPC memory layer (see Sec. 3.3), which is also available as a file system on the HPC systems. It is now possible to use data generated in the high-security HPC domain – without compromising security – within the cloud environment and *e. g.* via a service. Conversely, it is possible to offer service interfaces for data interventions that are more sophisticated than a simple copy command. This can

[e] `https://singularity.lbl.gov`

help set up quality assurance processes before the data is finally transferred to a repository. In addition, the approach allows the combination of service-based, community specific environments with the computational HPC environment, a scenario that has been difficult or impossible until now.

3.3 JUST: Storage Hierarchy

Since 2015 the Jülich Storage (JUST) Cluster is the central storage provider for supercomputers at JSC.[9] Over the years, JUST has evolved rapidly in response to technological developments and increasing user demand for capacity, performance and access requirements. The use of the multitude of storage technologies – each with different functions and price categories in terms of capacity and bandwidth per euro – requires customised architectures. JSC has organised these storage options into layers or tiers that are characterised by the performance, capacity, and retention time of the target data.

At the top of the performance hierarchy sits the high-performance storage tier, an I/O acceleration system installed in 2019 to provide very high bandwidth using non-volatile memory technology. A unique feature of this tier will be the tight integration with the client systems, still maintaining a coherent name-space across computer platforms to facilitate high-performance workflows spanning multiple systems in the modular supercomputing facility. The mid-range tiers consist of the existing high-performance storage components – updated in 2018 to achieve $2\times$ the bandwidth and over $3\times$ the capacity – and a new capacity-focused storage cluster – providing a campaign storage file system `DATA` that enables online access to data for longer time frames. The storage hierarchy is rounded off by the archival service, which provides reliable long-term storage.

Portions of the `DATA` file systems can be readily exported to a cloud infrastructure to support the deployment of community data services. Utilising the same underlying storage system, JSC will provide an object-storage data management service in 2020. This service will be accessible to the users of the Fenix federation (see Sec. 3.4) as well as the users of the local facility.

JUST will continue to evolve to meet the increasing performance of the supercomputers and further expand its service portfolio to provide the required ingress and egress points, enable data management and sharing as well as support community services and platforms providing data access and analysis capabilities.

3.4 Data Federation

An important step towards facilitating sharing data between computer sites and communities is the federation of storage infrastructures. JSC is involved in various projects that target this goal at national and European level.

Helmholtz Data Federation (HDF)

In the HDF project, six Helmholtz research centres Alfred-Wegener-Institut (AWI), Deutsches Elektronen-Synchrotron (DESY), Deutsches Krebsforschungszentrum (DKFZ), FZJ, Helmholtzzentrum für Schwerionenforschung (GSI), and Karlsruher Institut für Technologie (KIT) have enhanced their computing and data-management facilities with considerable resource capacities and capabilities to satisfy the additional resource needs of a wide

spectrum of scientific domains. Two major components to satisfy these needs have been installed at JSC: an OpenStack-based cloud environment (see Sec. 3.2) and an extended capacity storage tier (see Sec. 3.3). These new elements complement the already existing HPC resources to build an infrastructure supporting the full life-cycle of data-centric science.

Helmholtz Infrastructure for Federated ICT Services (HIFIS)

Whereas HDF is targeted at data from large-scale scientific challenges, HIFIS[f] has been created to support a more general set of use cases. Eleven Helmholtz centres have partnered in this activity to offer cloud, backbone, and software services for a broader scope of activities. HIFIS aims to integrate services that already exist at a subset of the partner sites and make them available for a wider audience, sharing the load of offering individual services among the partners. At JSC, the existing HDF resources are made available in the HIFIS infrastructure.

Federated European Network for Information eXchange (Fenix)

The Fenix research infrastructure is a federated data and computing infrastructure with components geographically distributed across Europe.

Fenix, built since 2018 as part of the EU-funded ICEI (Interactive Computing e-Infrastructure) project, brings together five major European supercomputer centres: BSC (Spain), CEA (France), JSC (Germany), CSCS (Switzerland) and CINECA (Italy). Fenix combines interactive computing, cloud-based virtual machine hosting and a distributed data infrastructure with or close to supercomputing resources. This unique combination of services enables Fenix to support the construction of higher-level, community-driven platform services, such as an analysis or simulation workload that accesses large data pools controlled by a web front-end. The first user community of this infrastructure is the HBP, where resources in Fenix are available to European neuroscientists. However, Fenix is generally beneficial to European scientists through PRACE, as it offers a new program-driven access model with peer-review-based resource allocation. At the same time, the project will continue the development and deployment of key infrastructure components.

In order to enable a truly federated infrastructure, without introducing a strong operational dependency between centres, a decentralised approach with web and cloud technologies was chosen. A federated authentication and authorisation infrastructure is being devised that shall enable the use of multiple site and community identity providers.

Fenix data stores are separated into two different classes: Local *active data repositories* and *archival data repositories*. While the former are strictly site-local and typically optimised for high-performance, the latter are federated and exposed at all sites via the same type of object-storage interface. This enables distributed data access with a uniform user interface and consistent control of access to data objects.

While resources have been provided since 2019, several components of the Fenix infrastructure are in development and will be incrementally deployed in the next two years.

[f] https://www.helmholtz.de/en/research/information-data-science/helmholtz-federated-it-services-hifis/

This encompasses a centralised web-portal for the management of users and resources of the infrastructure as well as data movement and transfer technologies.

A key focus of the Fenix infrastructure is sustainability, both in the technical design and in the governance structure. Fenix is expected to be an important part of the European supercomputing and cloud strategy well beyond the end of the ICEI project in 2023. Fenix will be able to provide the adhesive between the geographically distributed EuroHPC systems in Europe, including an Exascale system at JSC.

4 Application-Related Activities

The increasing parallelism and complexity of HPC systems has motivated large user groups to modernise their entire code base and apply more flexible and portable programming strategies. At the same time, it has become clear to HPC architects that the requirements and limitations of end users must be carefully considered when building future HPC systems. This close interaction between system, software and application developers – coined with the term *co-design* – has therefore become a decisive element in the HPC landscape. Co-design, application modernisation and porting must be intensified in view of the expected high heterogeneity in Exascale systems on all levels, the desire for compatibility and the still unclear landscape of programming paradigms for support. Some of the activities carried out at JSC in these areas are described below.

4.1 Co-Design

The JSC team is actively driving technology development for Exascale in Europe, in collaboration with key industrial and academic partners. Projects such as the DEEP series, the European EPI[g], Mont-Blanc 2020, Maestro, or Sage are just some examples in which JSC plays a major role in the co-design strategy.

The DEEP projects – coordinated by JSC – have delivered a blue-print for a stringent co-design approach for system-level architectures. This approach included twelve application domains and over twenty individual codes. Their requirements, collected through numerous interviews and intensive co-design discussions, determined the components used in the hardware prototypes and the functionalities to be provided by the software stack, strongly influencing the evolution of the MSA over time.

JSC is also the driving force behind the co-design effort in the EPI project, which aims at developing processor technologies with IP-control in Europe and constitutes a key element of the EuroHPC JU Exascale strategy. A large selection of applications, mini-apps, and benchmarks have been chosen to be part of the EPI benchmark suite. In combination with processor-architecture models, simulators, and emulators the EPI benchmark suite will serve to evaluate design trade-offs and find the best chip configuration for the future HPC users.

4.2 Programming GPUs

The JUWELS Booster (see Sec. 2.2) will be the first large-scale GPU system at JSC, following the GPU partitions of the JUWELS and JURECA Clusters.

[g] https://www.european-processor-initiative.eu

Several programming models are available to enable the use of GPUs. The simplest way is to replace libraries with GPU-accelerated versions, *e. g.* using Nvblas during runtime as a GPU-accelerated drop-in replacement for a dynamically linked BLAS library, or using cuBLAS during compilation for the same effect. If libraries are not available, compiler directives can be used to mark the parts of the code that should run on the GPU, as it is done in CPU code when parallelising loops with OpenMP. In fact, OpenMP provides directives that enable offloading of code segments to GPUs. OpenACC offers an alternative set of directives that enable a higher level of abstraction specifically for GPUs. In both cases it is important to pay attention to memory management and data structures.

Nvidia's own programming environment CUDA is available for multiple languages including C++, Fortran, and Python. It provides tools and documentation, and it is the platform that exposes all new hard- and software features and makes them available first. Unfortunately, CUDA is limited to Nvidia GPUs. A more portable alternative is HIP, developed as an Open Source project by AMD. HIP is very similar to CUDA and compiles for Nvidia and AMD GPUs. Last, but not least, a number of programming models based on C++ programming paradigms (like functors and lambdas, C++ data containers, and basic algorithms) are currently developed to enable programmers to write code that will run efficiently and is portable on GPUs from different vendors and other parallel platforms including CPUs. Examples of these high level abstractions for C++ are Kokkos[h] and SYCL[i].

None of the models discussed above deal with more than a single node, but, luckily, they play nicely with MPI. The MPI versions to be installed on the JUWELS Booster module are GPU-aware. This means that they can transfer data directly from a GPU belonging to rank A to a GPU belonging to rank B. Depending on the locations of ranks A and B, the data is transferred in the most efficient way. This is transparent to the user, making MPI a useful model for programming multi-GPU systems both for intra-node as well as inter-node communication.

While the abundance of programming alternatives is indicative of the thriving GPU ecosystem, it may be hard for the programmer to decide which one to choose for their project. JSC offers courses for learning CUDA and OpenACC. Investigation of the other available programming models are also ongoing to be able to advise our users on the best choice for their applications and to extend the range of training that we can offer. Many of JSC's GPU activities are concentrated through a GPU-related Exascale lab, but consultation on a per-application basis is also provided. In addition, application-oriented high-intensity training for GPU-enabling is offered in the form of week-long GPU Hackathons hosted at JSC. At the multi-node level, a large amount of available parallelism needs to be expressed. As part of the *High-Q club*[j], applications have managed to scale up to about 2 million tasks and threads. This corresponds to the parallelism needed to feed about 100 of the GPU nodes. JSC plans to enable more applications to achieve this kind of parallelism and hopes to see some that will scale far beyond this.

JSC is not alone in its effort to target GPUs for highly scalable systems. CSCS in Switzerland, and the current and upcoming systems in the US are using GPUs to accelerate

[h] `https://github.com/kokkos`

[i] `https://www.khronos.org/sycl/`

[j] `https://www.fz-juelich.de/ias/jsc/EN/Expertise/High-Q-Club/_node.html`

their scientific applications. Many of the improvements made at other sites will benefit JSC's users as well. This is already obvious for users of GROMACS, AMBER, and of a variety of DL methods, who were able to take advantage of the new GPU nodes right away.

4.3 Deep Learning

Especially in DL[14] GPUs became the dominant accelerators used by all major software libraries like TensorFlow, PyTorch, MxNet and Chainer. A widespread class of learning models called convolutional neural networks (CNNs)[15] relies heavily on the efficient execution of very large sets of higher order matrix or tensor operations during learning, which are efficiently supported by GPU libraries such as cuDNN. State-of-the-art DL methods require extensive training on the available data. Even with the latest generation of GPUs, this takes a long time to create a suitable model and achieve a satisfactory level of performance for a particular task, such as visual object recognition or speech understanding. To train such a SOTA DL model, even a workstation equipped with the latest generation NVIDIA graphics processors (V100) will require a runtime of many hours to many days or even weeks, depending on the model type.

Acceleration can be achieved by increasing training to a large number of GPU nodes[10] without compromising trained model accuracy. This so-called distributed training requires efficient communication between the different nodes during learning. In this case, efficient inter-node communication is handled via CUDA-aware MPI routines implemented in low-level libraries like NCCL[k], which in turn can be used by high-level libraries like for instance Horovod[11] to implement efficient `allgather` and `allreduce` operations necessary for the updates of the learning model during distributed training.

At JSC, the Cross-Sectional Team Deep Learning (CST-DL) has employed Horovod since its appearance in 2017, enabling successful distributed training for different DL methods across hundreds of GPUs, reaching almost linear scaling without loss of model accuracy. Recent work suggests that scaling can be extended also to Exascale machines using the same methods.[2] Therefore, JSC is well prepared to perform distributed training of DL models over the thousands of GPUs available with the arrival of the JUWELS Booster. The next challenge will be to employ distributed training on MSAs that will contain heterogeneous accelerators.

With new computational demands put forward by effort to couple physics-based simulations and DL, parallelisation and scaling issues will face new challenges, as the type of computations to be performed efficiently may differ from component to component of such a closed-loop end-to-end simulation-AI pipeline. These challenges have to be met by the design of the MSA that has to provide means to orchestrate the parallel execution of such heterogeneous components in an efficient manner.

4.4 Porting Applications to Modular Supercomputers

A number of applications are being adapted at JSC to test them on the JURECA modular supercomputer (see Sec. 2.1), to demonstrate its capabilities, and to show, on the one hand, the possibility, and on the other hand, the advantages of mapping different parts of an

[k] `https://developer.nvidia.com/nccl`

application onto different compute modules of an MSA system. A selection of applications and results obtained in this area is described below.

4.4.1 1D-NEGF

The 1D-NEGF code, based on the Non-Equilibrium Green's Functions framework, is used to simulate inelastic quantum charge transport phenomena in semiconductors such as novel nano-transistors and quantum photovoltaic devices. The need of open boundary conditions to account for scattering effects, non local dependencies of the Green's function variables, and two nested self-consistent loops results in a large numerical problem, which demands a highly parallel and efficient code, scalable up to hundreds of thousands of cores.

The structure of the NEGF code has a number of distinct tasks exhibiting different computational properties. Tasks are divided into two classes based on their scalability, and the corresponding kernels are partitioned accordingly: Less scalable kernels are mapped to the Cluster and higher scalable kernels to the Booster.

The rationale for this Cluster-Booster mapping comes from the observation that cluster CPUs have fewer cores, higher frequencies, and are better suited to computing tasks that can degrade performance and scalability. On the other hand, booster GPUs have a much higher core count at lower frequencies and are better suited for highly parallel tasks. In a conventional design, the serial tasks set an upper limit for acceleration by Amdahl's law. In modularity, the division of tasks between cluster and booster nodes, which we call "modular execution", minimises the influence of serial tasks on scalability. This leads to an increased overall scalability of the application, better energy efficiency and an improved FLOPs/W performance ratio.

The modular design of the application requires a few further steps compared to the normal cluster-only workflow. Based on the same source code, an executable file with different architecture-dependent compiler options is created for each module. To run a modular job on an MSA platform such as JURECA, two separate jobs must be started simultaneously – one for each executable file – via a co-schedule mechanism.

Since communication between the modules is crucial, the modular implementation of the 1D-NEGF code performs a modular logic that splits the common `MPI_COMM_WORLD` into a sub-communicator for the Cluster, one other for the Booster, and an inter-communicator, which is used for inter-module communication and separating tasks.

The modular execution of 1D-NEGF so far has been tested on 1 Cluster node (Haswell) and by increasing the number of Booster nodes (KNL) to up to 64. The Haswell node has a theoretical peak performance of about 1 TFlop/s while a KNL node provides about 3 TFlop/s. The modular execution used only one additional Haswell node – meaning merely $960/(64 \times 3046.4) \sim 0.5\,\%$ additional resources in terms of peak performance – and remarkably turns out to be $6\,\%$ faster than the conventional execution running on the corresponding number of KNL nodes only.

4.4.2 xPic

xPic is a particle-in-cell code developed by KU Leuven as a new implementation of iPic3D[12] to perform kinetic simulations of non-collisional plasma. The main application field is space weather forecasting, which simulates the interaction of solar plasma with the Earth's magnetosphere in order to predict conditions on Earth and in orbit.

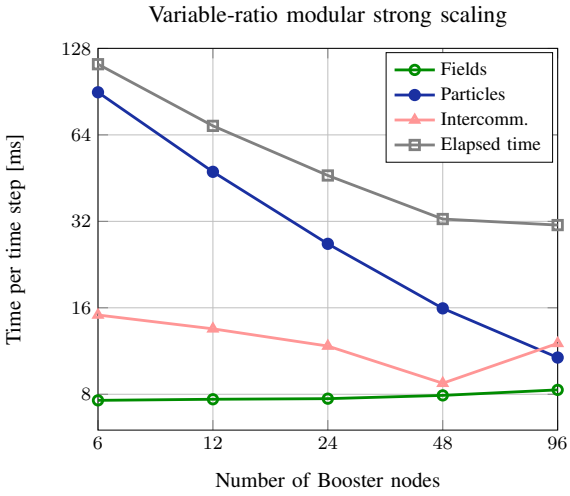


Figure 2. xPiC: strong scaling experiments running the field-solver on four nodes of the JURECA Cluster and the particle-solver on a varying amount of JURECA Booster nodes. In the last point of the graph a configuration is reached, in which a very similar amount of compute time is employed on Cluster and Booster, opening up the opportunity to largely overlap executions. This new option could not yet be exploited in traditional supercomputing.

The xPic code is essentially divided into two components: the field solver and the particle solver. Given the absence of interaction between the computational particles, each particle can be treated independently, making the particle mover in the second component embarrassingly parallel. This makes it particularly suitable for accelerators and co-processors. On the JURECA Booster, a Hybrid MPI/OpenMP approach is used to distribute work over the cores of the KNL processors, while using SIMD directives to exploit the large vector registers available. The field solver consists of a sparse linear solver, and it is expected to be less scalable than the particle solver. Therefore, on JURECA the field solver runs on the Cluster module. The Cluster-Booster implementation of xPic uses MPI (non-blocking point-to-point intercommunication) to exchange data between the field and particle solvers. Depending on the size of the problem and on the numerical algorithm, this intercommunication can take a substantial portion of the runtime.

In previous works it was shown how the particle solver could benefit from the KNL features.[13] More recent tests on JURECA with larger node counts highlight the different scalabilities of field and particle solver, and show how the flexibility of the Cluster-Booster approach could be used to control the scaling of the two sides of the applications independently (see Fig. 2).

4.4.3 Lattice QCD

The SimLab Nuclear and Particle Physics (SLNPP) group prepares LQCD codes for new and future architectures, focusing on GPU porting and MSA.

A large part of the effort to optimise LQCD codes for GPU-usage is based on QUDA[l], a CUDA-based library of LQCD kernels and sparse matrix inverters. Working together with QUDA developers (including Nvidia experts), the SLNPP team aims to identify further code hot spots which could be offloaded to GPUs.

In parallel to the GPU-focused activities, the SLNPP group is preparing community LQCD simulations codes to run on future modular supercomputers. As an initial step, the group has developed the *QMOD* library containing functions that can pass global data structures (*e. g.* a lattice gauge fields, or propagators) from one hardware partition to another. This library interfaces with the widely-used USQCD software stack[m] and its various architecture-specific back-ends, such as QUDA. QMOD should make it possible for anyone using a LQCD code linked against the USQCD software stack to develop and test concurrency schemes in a modular supercomputing environment.

5 User Support

Supporting users is a key element to reach the efficient use of the current HPC system and upcoming Exascale systems. In addition to offering training courses and workshops – described at the end of this section – direct contact with users and cooperation with them in solving problems and optimising application codes are essential.

The three GCS partners Höchstleistungsrechenzentrum Stuttgart, JSC and Leibniz-Rechenzentrum have developed a HPC application support structure as an evolution of well-established user support services, tailored to suit the need of HPC-users on Exascale systems. As shown in Fig. 3 the services are offered at four different levels.

In *support Level 1* the service desk is the most important component. Its staff and user administration can rely on a ticket system as the main tool to process service requests. More complex issues are handled at *support Level 2*. The close collaboration within JSC between HPC experts (cross-sectional teams and system administrators) and experts from various scientific fields (SimLabs) has proven to be an efficient strategy for solving problems or improving user applications by means of code analysis and optimisation. JSC provides tools and web-based services (LLview job-reporting[n]) which enable users to monitor performance and other HPC-relevant metrics of their jobs. In this way staff members as well as users themselves can detect if codes are consistently running inefficiently and initiate support actions. JSC provides in *support Level 3* detailed support in restructuring and optimising user codes, again using in-house tools like Score-P or Scalasca. Within the scope of joint research projects, the priority of *support level 4* is to establish long-term collaborations among the GCS centres, academia, and researchers. Every large-scale project is assigned a *mentor* who acts as a permanent point of contact and has a detailed understanding of the users project, its history, its resource consumption, as well as the projects associated challenges. The mentor coordinates communication between users and support staff and serves as a complementary service to the support offered at levels 2 and 3. The user profits from a more efficient and personalised support structure for his or her project while the centre benefits from more efficient use of valuable HPC resources.

[l] `http://lattice.github.io/quda`
[m] `https://www.usqcd.org/software.html`
[n] `http://www.fz-juelich.de/jsc/llview`

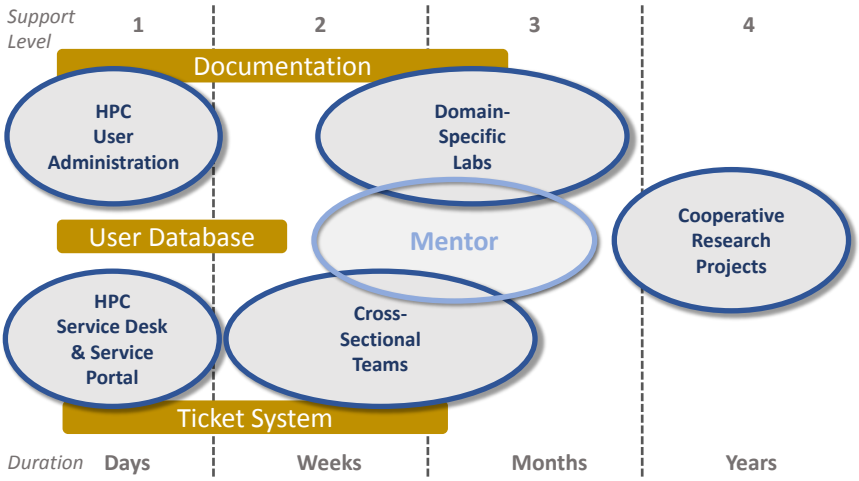


Figure 3. The new HPC application support structure developed by GCS. Level 1: support for basic functionalities and services; Level 2: more complex issues, *e. g.* application code analysis and optimisation; Level 3: detailed support in restructuring and optimising user codes; Level 4: long-term collaboration with users.

To be prepared for Exascale architectures and in particular for data-centric computing, JSC introduced a new user model for the HPC systems in 2018. In contrast to the former model, data management has been moved from a project-centric to a user-centric view that allows for more flexibility, and self-manageable and easier mapping of users to compute and data projects.

In addition to all these support vehicles, JSC provides a wide portfolio of training courses and workshops addressing all relevant topics of HPC from a user perspective. In particular, to foster the collaboration with users on performance optimisations, the VI-HPS workshop series on performance tuning and Porting and Optimisation workshops have been organised. In addition, extreme scaling workshops, Big Blue Gene Weeks and the High-Q Club – all initiated during the Blue Gene era at JSC – focus on extreme scaling of HPC applications.

Finally, JSC is member of the Performance Optimisation and Productivity Centre of Excellence (POP CoE[o]), which offers free performance assessment services for parallel scientific applications, performance measurement and analysis training (also in the form of webinars), and develops new and simplified methods to portably describe and assess the performance of parallel applications.

6 Outlook

JSC is preparing to establish itself as one of the first sites in Europe to operate an Exascale system. Although formal decisions are still pending, the JSC team takes up the responsibility to already prepare for the technical challenges that such a task poses for the data

[o] `https://pop-coe.eu`

centre and its users. In this context, a variety of activities are underway in the areas of hardware, software and application development. This paper describes a subset of important activities ranging from general code modernisation strategies for adaptation to modular systems including GPU programming to approaches aimed at coping with the flood of data emanating from existing applications. Areas and upcoming data-oriented approaches will contribute to the Exascale Computer and Computing Centre.

The overarching concept on which JSC's entire Exascale strategy is based is the MSA. Through the novel and composable arrangement of traditionally heterogeneous computing resources, it offers solutions adapted to a broader range of applications, from traditional, closely coupled simulations to emerging data analysis, machine learning, and artificial intelligence codes. The further development of the MSA, its implementation in the JUWELS Tier-0 system and the preparation of applications are core elements of the JSC Exascale strategy.

Disruptive technologies such as neuromorphic and quantum computer devices are unlikely to provide the computational and data management functions required to build an Exascale system alone and in a short time. However, they evolve rapidly and can already be used very efficiently to solve specific application tasks such as pattern recognition or problem optimisation. JSC's focus is therefore on enabling their integration into an HPC environment, facilitating their introduction, and investigating new use cases for these technologies. Here, too, the MSA concept plays a decisive role, as it allows *disruptive* modules without disturbing the operation of the *traditional* compute modules.

The next immediate step in JSC's Exascale strategy is to fully deploy the JUWELS Booster and help users leverage its capabilities in conjunction with the existing JUWELS Cluster. In parallel to the operational and support tasks, JSC will further develop the MSA and its software and programming environment, understand the specific requirements of the users and bring this knowledge into the design of the modular Exascale platform.

Acknowledgements

The authors gratefully acknowledge the funding provided by the Helmholtz Programme *Supercomputing & Big Data* to realise the JURECA Cluster and Booster, as well as the project funding provided by the Ministry of Education and Research (BMBF) and the State of North Rhine-Westphalia for the procurement of the JUWELS Cluster and Booster (SiVeGCS). Part of the research presented in this paper has received funding from the Deutsche Forschungsgemeinschaft (DFG) under Grant GSC 111. The European Community funded the research performed in the series of DEEP projects through the Seventh Framework Programme (FP7/2007-2013) and the Horizon 2020 (H2020-FETHPC) Programme, under Grant Agreements n° 287530 (DEEP), 610476 (DEEP-ER), and n° 754304 (DEEP-EST). This publication reflects only the authors' views. The European Commission is not liable for any use that might be made of the information contained therein.

References

1. *The Scientific Case for Computing in Europe 2018 2026*, PRACE, E. Lindahl (Editor), published by Insight Publishers, Bristol, UK, 2018, ISBN: 9789082169492, `http://www.prace-ri.eu/third-scientific-case/`

2. N. Laanait *et al.*, *Exascale deep learning for scientific inverse problems*, 2019, arXiv:1909.11150 [cs.LG].
3. E. Suarez, N. Eicker, and Th. Lippert, *Supercomputer Evolution at JSC*, Proceedings of the NIC Symposium 2018, K. Binder, M. Müller, A. Trautmann (Editors), 1–12, 2018, ISBN:978-3-95806-285-6.
4. E. Suarez, N. Eicker, Th. Lippert, *Modular Supercomputing Architecture: from idea to production*, Contemporary High Performance Computing: From Petascale toward Exascale **3**, J. S. Vetterm (Editor), CRC Press, 223–251, 2019, ISBN:978-1-1384-8707-9.
5. N. Eicker, Th. Lippert, Th. Moschny, and E. Suarez, *The DEEP Project - An alternative approach to heterogeneous cluster-computing in the many-core era*, Concurrency and computation: Practice and Experience **28**, 2394-2411, 2016, doi:10.1002/cpe.3562.
6. D. Krause and P. Thörnig, *JURECA: Modular supercomputer at Jülich Supercomputing Centre*, Journal of large-scale research facilities **4**, A132, 2018, doi:10.17815/jlsrf-4-121-1.
7. C. Clauss, Th. Moschny, and N. Eicker, *Allocation-Internal Co-Scheduling - Interaction and Orchestration of Multiple Concurrent MPI Sessions*, Advances in Parallel Computing **28**, 46–68, 2017.
8. D. Krause, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, Journal of large-scale research facilities **5**, A135, 2019, doi:10.17815/jlsrf-5-171.
9. D. Krause, S. Graf, and O. Mextorf, *JUST: Large-Scale Multi-Tier Storage Infrastructure at the Jülich Supercomputing Centre*, Journal of large-scale research facilities **5**, A136, 2019, doi:10.17815/jlsrf-5-172.
10. R. Mayer and H.-A. Jacobsen, *Scalable deep learning on distributed infrastructures: Challenges, techniques and tools*, 2019, arXiv:1903.11314 [cs.DC].
11. A. Sergeev and M. Del Balso, *Horovod: fast and easy distributed deep learning in tensorflow*, 2018, arXiv:1802.05799 [cs.LG].
12. G. Lapenta, S. Markidis, S. Poedts, and D. Vucinic, *Space Weather Prediction and Exascale Computing*, Computing in Science Engineering **15**, 68–76, 2013, doi:10.1109/MCSE.2012.86.
13. A. Kreuzer, N. Eicker, J. Amaya, and E. Suarez, *Application Performance on a Cluster-Booster System*, 2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 69–78, 2018, doi:10.1109/IPDPSW.2018.00019.
14. Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature **521**, 436–444, 2015, doi:10.1038/nature14539.
15. A. Krizhevsky, I. Sutskever, and G. E. Hinton, *Imagenet classification with deep convolutional neural networks*, Proceedings of the 25th International Conference on Neural Information Processing Systems **1**, 1097–1105, 2012.
16. M. Stephan, and J. Docter, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities **1**, A1, 2015, doi:10.17815/jlsrf-1-18.

The NIC Research Groups

A Study of the Kinetic Freeze-Out in Relativistic Heavy Ion Collisions with the UrQMD/Coarse Graining Approach

Gabriele Inghirami[1,2], **Paula Hillmann**[3,4,5,6], **Boris Tomášik**[7,8], **and Marcus Bleicher**[3,4,5,6]

[1] University of Jyväskylä, Department of Physics,
P.O. Box 35, 40014 University of Jyväskylä, Finland

[2] Helsinki Institute of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland

[3] Frankfurt Institute for Advanced Studies (FIAS),
Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany
E-mail: bleicher@fias.uni-frankfurt.de

[4] Institut für Theoretische Physik, Johann Wolfgang Goethe-Universität,
Max-von-Laue-Str. 1, 60438 Frankfurt am Main, Germany

[5] GSI Helmholtzzentrum für Schwerionenforschung GmbH,
Planckstraße 1, 64291 Darmstadt, Germany

[6] John von Neumann Institute for Computing, GSI, Planckstraße 1, 64291 Darmstadt, Germany

[7] Fakulta prírodných vied, Univerzita Mateja Bela, 97401 Banská Bystrica, Slovakia

[8] Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze,
11519 Prague, Czechia

Heavy ion collisions are an excellent tool to test Quantum Chromo-Dynamics, the theory describing the fundamental interaction that allows the formation of the atomic nuclei. Ions of heavy elements like gold or lead are smashed together at relativistic speeds and the particles constituting the debris of the collision are captured by huge detectors to analyse their basic properties. The kinetic freeze-out is the final stage of heavy ion collisions, when the emitted particles cease to interact with each other and their momenta do not change anymore, until the detection by the experimental devices. Therefore, a full understanding of this process is crucial to properly reconstruct the dynamics of the system. In this study we exploit the UrQMD/coarse graining approach to estimate the thermodynamic properties of the system during the kinetic freeze-out. We focus on central Au+Au collisions at five significant reaction energies, from $\sqrt{s_{NN}} = 2.4\,\mathrm{GeV}$ (GSI-SIS) to $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$ (RHIC). We determine the average temperature $\langle T \rangle$ and the average baryon chemical potential $\langle \mu_\mathrm{B} \rangle$ at the moment of the last hadronic interaction, based on a Hadron Gas EoS, and we evaluate what is the distribution of the kinetic freeze-out process with respect to these quantities. The results highlight the nature of the kinetic freeze-out as a continuous process that may take place under very different conditions, an aspect often shadowed in its common representation as a single point in the phase diagram.

1 Introduction

As the name itself suggests, in relativistic heavy ion collisions atomic nuclei of heavy elements, like gold or lead, are accelerated to nearly the speed of light and smashed together. During the initial stages of the collision, the matter becomes very compressed

and extremely hot, with temperatures around 2000 billion Kelvin. Under these conditions, matter is believed to form a new state, called the *Quark and Gluon Plasma* (QGP). The existence of QGP, hypothesised decades ago,[1, 2] is now confirmed by solid experimental evidences.[3, 4] What makes the QGP so interesting is that, under normal conditions, quark and gluons, the elementary particles which also form neutrons and protons, are bound together into composite particles called *hadrons*. The fundamental interaction to which they are sensitive, called *strong* and described by the theory of Quantum Chromo-Dynamics (QCD),[5] has a peculiar property, called *confinement*, that prevents the existence of isolated free quarks or gluons, which, in fact, have never been observed.[6] However, in the QGP quarks and gluons are not confined into hadrons anymore and they exhibit a collective behaviour similar to a fluid with low viscosity,[7] which can be effectively modelled using relativistic hydrodynamics.[8, 9] The QGP is extremely ephemeral: due to the huge gradients of pressure, the QGP medium expands at supersonic speed, quickly cools down and converts again into hadrons. Essentially, in heavy ion experiments we reproduce a state similar to that existing in the whole Universe until a few microseconds after the Big Bang and then we observe how matter clumps together. It's no surprise that sometimes the explosion of the fireball created in heavy ion collision is also referred as the *Little Bang*.[10, 11] Certainly, mastering a full knowledge of how matter forms would represent not only an admirable achievement of the human intellect, but there would be also a plethora of unprecedented technological applications. Indeed, heavy ion collisions are a very powerful tool to accomplish this long term goal, but, unfortunately, their study presents also several difficult challenges. The main problem is that, given the small dimensions of the QGP system under investigation ($\approx 10^{-14}$ m) and its extremely fast evolution ($\approx 10^{-22}$ s), direct observations are not possible. Even with the most advanced experimental apparatus, we are limited to the detection of hadrons at distances many orders of magnitudes larger than the typical size of the colliding ions. Therefore, physicists have to resort on dynamical modelling to connect the theory with the experimental data. Quite often this dynamical modelling is based on numerical simulations that, given the many possible variants in the initial conditions and in the parameter space, require significant computational resources, which only a few centres, like the John von Neumann Institute for Computing, are able to provide.

2 Scope

In this study we focus on the final stage of the collision event, the so called *kinetic* or *thermal freeze-out*,[12, 13] when the emitted hadrons stop interacting and their momenta do not change anymore until detection. This phase follows the *chemical freeze-out*,[14, 15] which, instead, refers to the ceasing of the inelastic scatterings and the stabilisation of the abundances of the hadronic species. In principle, the kinetic and the chemical freeze-out are two different concepts,[16, 17] albeit some models estimate their occurrence at similar temperatures[18] or consider them a single phenomenon.[19, 20] In our model, we identify the kinetic freeze-out with the last interaction of a hadron (inelastic, elastic or strong decay). A common approach to determine the properties of the medium at kinetic freeze-out is based on the so called Blast-Wave model,[21] in which temperature, baryon chemical potential and radial velocity are obtained by fits of experimental data to an assumed phase-space density distribution of hadrons.[22, 23] In the present study[24] we adopt a somewhat opposite

approach.[25, 26] We employ the numerical transport code UrQMD[27, 28] to numerically compute the evolution of the system in terms of microscopic quantities (*i. e.* the scatterings of the hadrons composing the system), we retrieve the corresponding macroscopic quantities by using a coarse-graining method[29, 30] and, finally, we associate them to the kinetic freeze-out event points. Our current approach has a few shortcomings,[24] nevertheless they do not hinder the main goal of this study: highlighting the nature of the kinetic freeze-out as a continuous, dynamical process, by evaluating the distribution of a few basic thermodynamic quantities. We examine collectively the most abundant hadron species and we concentrate on central Au+Au reactions at five relevant collision energies, from $\sqrt{s_{NN}} = 2.4\,\mathrm{GeV}$, slightly below the QGP threshold production, to $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$, already a bit above the limit of applicability of a pure hadronic model without an intermediate hydro phase.[31]

3 Description of the Model

The model adopted for this study is based on the Ultra-relativistic Quantum Molecular Dynamics (UrQMD) transport model, which is able to simulate the dynamics of a heavy ion collision based on the covariant propagation of hadrons, taking into account both elastic and inelastic interactions. For details, the reader is referred to Ref. 27, 28. UrQMD is used, in the first place, to simulate a large number of heavy ion collision events, from which we extract the average thermodynamic properties of the system during its evolution by exploiting a coarse-graining method. Then, UrQMD is used again to determine the time and the position of the kinetic freeze-out of the hadrons. The UrQMD coarse-graining method has been successfully employed to study dilepton and photon production from GSI-SIS to RHIC energies[30, 32] as well as to compute the evolution of the background medium for heavy quark Langevin simulations.[33] In the coarse-graining method one approximates the hadronic phase space distribution function by performing averages on a low resolution spatial grid over the total ensemble of hadrons produced in a large set of heavy ion collision events. For each cell of the coarse grained grid, these averages are made at constant times with respect to the momenta of the hadrons contained inside that cell. To be more precise, we evaluate the net-baryon four current j^μ_B and the energy momentum tensor $T^{\mu\nu}$ as

$$j^\mu_\mathrm{B}(\boldsymbol{x},t) = \frac{1}{\Delta V}\left\langle \sum_{i=1}^{N_h\in\Delta V} B_i \frac{p_i^\mu}{p_i^0} \right\rangle, \qquad T^{\mu\nu}(\boldsymbol{x},t) = \frac{1}{\Delta V}\left\langle \sum_{i=1}^{N_h\in\Delta V} \frac{p_i^\mu p_i^\nu}{p_i^0} \right\rangle \tag{1}$$

in which ΔV stands for the cell volume, B_i and p_i^μ for the baryon number and the μ component of the four momentum of the hadron i, respectively, and the sums are done over all hadrons N_h. We choose the frame defined by the net baryon current j^μ as the reference fluid frame (Eckart's definition[34]). The baryon density ρ_B and the energy density ε, with respect to the fluid rest frame, are then computed from the net-baryon current and the energy momentum tensor by a Lorentz transformation. Once these quantities are known, it is possible to obtain the temperature $T(\varepsilon,\rho_\mathrm{B})$ and the baryon chemical potential $\mu_\mathrm{B}(\varepsilon,\rho_\mathrm{B})$ by interpolation from a tabulated Hadron Resonance Gas EoS.[35] The final step of the procedure consists in associating these quantities with the microscopic freeze-out distribution in space and time computed with UrQMD.

4 Results

We perform numerical simulations of Au+Au collisions with impact parameter $b = 0 - 3.4\,\mathrm{fm}$, approximately corresponding to the 0-5 % centrality class, from $\sqrt{s_{NN}} = 2.4\,\mathrm{GeV}$ to $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$, exploring a range of energies relevant for the HADES[36] at GSI, NA49[37] at CERN and RHIC/BES[38] at BNL experiments. The setup parameters of the simulations are summarised in Tab. 1. For each collision energy we also run 10^4 UrQMD events to calculate the collections of freeze-out points. In this study we limit our interest to the most abundant and important hadron species, *i. e.* pions, kaons, protons, neutrons, lambdas and their antiparticles, including the feeddown of resonance decays.

$\sqrt{s_{NN}}$ (GeV)	$N_{ev.}$	Δt (fm/c)	Δx (fm)	$N_{x,y}$	N_z	t_{max} (fm/c)
2.4	$1.8 \cdot 10^6$	0.5	0.8	70	200	80
4.5	$7.4 \cdot 10^5$	0.5	0.8	80	250	90
7.7	$6.6 \cdot 10^5$	0.5	0.8	80	250	90
19.6	$3.6 \cdot 10^5$	0.5	0.8	86	276	100
200	$6.4 \cdot 10^4$	0.5	1.0	200	402	200

Table 1. List of the main parameters used in the UrQMD/coarse-graining numerical simulations. We report the values of the collision centre-of-mass energy $\sqrt{s_{NN}}\,(\mathrm{GeV})$, the number of events $N_{ev.}$, the time resolution $\Delta t\,(\mathrm{fm/c})$, the spatial resolution $\Delta x\,(\mathrm{fm})$, the number of cells along in the transverse plane ($N_{x,y}$) and in longitudinal direction (N_z), and the time $t_{max}\,(\mathrm{fm})$ after the collision at which we stop the simulations.

4.1 Freeze-Out Time Distributions, Temperature and Baryo-Chemical Potential Variations on the Decoupling Hyper-Surface

As a first step, we explore the distribution of the kinetic freeze-out with respect to the time t, defined in the centre-of-mass frame starting from the beginning of the collision, focusing on central rapidities ($|y| < 0.2$). We recall that the rapidity y of a particle is defined as $1/2 \log[(E+p_z)/(E-p_z)]$, where E and p_z are its energy and its longitudinal momentum, respectively. So a particle with $y \approx 0$ has a small longitudinal momentum (compared to its energy). Fig. 1 shows the normalised time distributions of the hadronic last interaction in central Au+Au collisions from $2.4\,\mathrm{GeV}$ to $200\,\mathrm{GeV}$, corresponding to the decoupling probability of a hadron from the rest of the system. The figure clearly indicates that the kinetic freeze-out is reached within a quite large interval of time, with a typical duration of the process of roughly $15-20\,\mathrm{fm/c}$ and a probability peak between 10 and $30\,\mathrm{fm/c}$. There are two main effects that determine the position of the decoupling peak: I) the transition time of the initial nuclei and II) the dynamics of the rapidly expanding new hot matter created in the collision. At low energies, the dominant effect is due to the transition time, *i. e.* the time needed by two initial nuclei to pass through each other. At higher energies, the higher longitudinal velocity of the nuclei, and the consequent relativistic length contraction that makes the nuclei thinner, strongly reduce the transition time. In the high energy regime, it is rather the dynamics of the meson[a] dominated matter that mostly influences

[a]Loosely speaking, mesons are hadrons formed by two quarks, while baryons are composed by three quarks. The most common mesons are pions and kaons, the most common baryons are protons and neutrons.

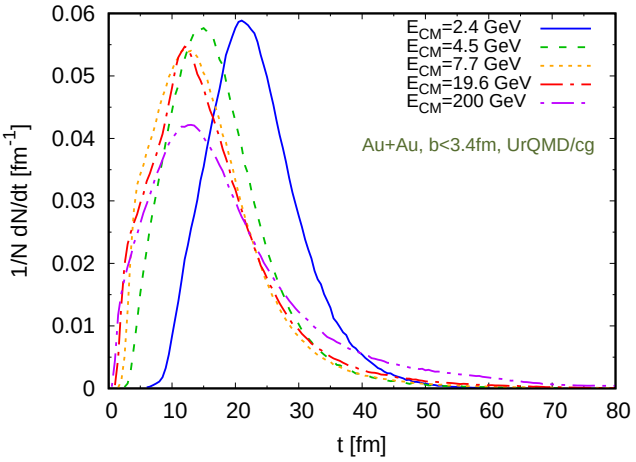


Figure 1. Freeze-out time distribution of hadrons at midrapidity ($|y| < 0.2$) for central Au+Au reaction at centre-of-mass energies of $\sqrt{s_{NN}} = 2.4, 4.5, 7.7, 19.6, 200$ GeV (full line, short dashed line, dashed line, long dashed-dotted line, dotted dashed line). The distributions are normalised to unity.

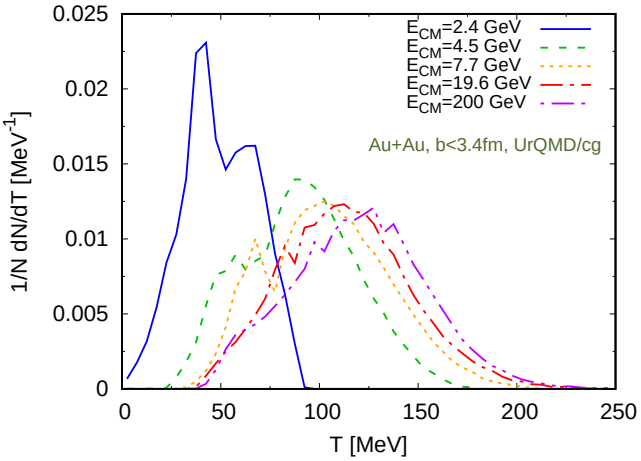


Figure 2. Emission probabilities as a function of temperature at midrapidity ($|y| < 0.2$) for central Au+Au reaction at centre-of-mass energies of $\sqrt{s_{NN}} = 2.4, 4.5, 7.7, 19.6, 200$ GeV (full line, short dashed line, dashed line, long dashed-dotted line, dotted dashed line). The distributions are normalised to unity.

the decoupling time, making it shorter and, at the same time, extending the tail of the distribution with increasing energy.

The UrQMD/coarse graining procedure makes possible to express the decoupling probability also in terms of distributions with respect to the temperature and the baryon chemical potential.

Fig. 2 depicts the normalised distributions of the temperatures at kinetic freeze-out

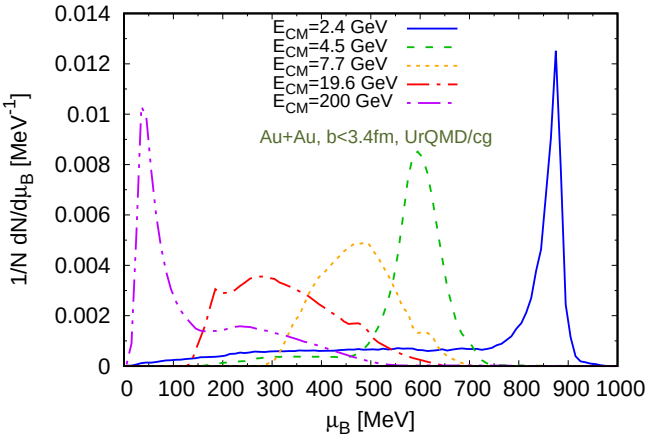


Figure 3. Emission probabilities as a function of baryo-chemical potential at midrapidity ($|y| < 0.2$) for central Au+Au reaction at centre-of-mass energies of $\sqrt{s_{NN}} = 2.4, 4.5, 7.7, 19.6, 200$ GeV (full line, short dashed line, dashed line, long dashed-dotted line, dotted dashed line). The distributions are normalised to unity.

in central Au+Au collisions from $2.4\,\mathrm{GeV}$ to $200\,\mathrm{GeV}$ (from left to right). The plots exhibit well defined maxima, associated with rather broad distributions. We notice that the maximum of the temperature distribution increases with increasing collision energy, in a way similar to the chemical freeze-out curve,[39] yet never exceeding a certain threshold of roughly $150\,\mathrm{MeV}$, even at the highest energies, with a tail that can exceed $170-180\,\mathrm{MeV}$. We recall that we are using a simple Hadron Gas EoS, an appropriate choice for a purely hadronic state, but that might have some issues in describing accurately a system at the limit of the QGP phase or not completely in chemical equilibrium.

Fig. 3 shows the normalised distribution of the kinetic freeze-out points with respect to the baryon chemical potential μ_B in central Au+Au collisions from $2.4\,\mathrm{GeV}$ to $200\,\mathrm{GeV}$ (from right to left). The curves exhibit a clear peak structure, tied with the amount of baryon density in the late evolution of the system. As expected, the distribution shows a sharp peak at high μ_B for low collision energies and at low μ_B for high collision energies. In the intermediate collision energy region, the distributions tend to be somewhat broad due to the transition from a baryon dominated system to a meson dominated system.[b] This kind of behaviour is consistent with the picture of the kinetic freeze-out as the outcome of the non trivial interplay between the complex dynamics of the system and the scattering cross sections of the hadrons, resulting in a process that is not only continuous in time, but also non-isothermal in the temperature and density.

4.2 Temperature and Baryon Chemical Potential Diagrams

Now we summarise the present results for the average kinetic decoupling temperature and chemical potentials in the T-μ_B-diagram, so to make it easy to compare them with chemical

[b] In the collision process, a relevant part of the initial kinetic energy is converted into particles. For a particle at rest, the conversion obeys the famous $E = mc^2$ Einstein's equation. At high collision energy many new particles are produced and most of them are mesons, because in general they have a smaller rest mass than baryons and therefore their production is statistically favoured.

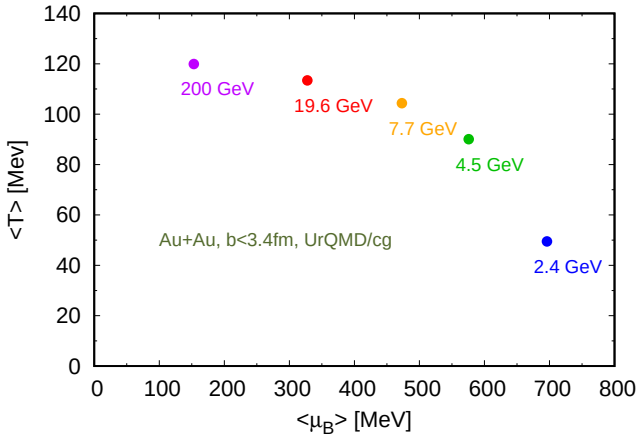


Figure 4. Kinetic freeze-out temperature with respect to the baryon chemical potential in Au+Au reactions at different centre-of-mass energies in the rapidity range $|y| < 0.2$.

freeze-out curves or Blast-Wave fits. Fig. 4 depicts the average temperature and baryon chemical potentials for the five reaction energies analysed in this study. Here we can devise a clear pattern traced by the $(\langle T \rangle, \langle \mu_{\rm B} \rangle)$ points, with increasing $\langle T \rangle$ and decreasing $\langle \mu_{\rm B} \rangle$ as the collision energy increases. Yet, we would like to remark that the common representation of the kinetic freeze-out as a single point in the phase diagram, despite being a convenient way to summarise its key properties, might induce to overlook its complex structure. To this aim, in Fig. 5 we show the density of the kinetic freeze-out events in the $(T, \mu_{\rm B})$ plane for central Au+Au reactions at $\sqrt{s_{NN}} = 19.6\,{\rm GeV}$. It is evident that the process is neither homogeneous nor narrowly localised, but different parts of the system decouple at different $(T, \mu_{\rm B})$ points, albeit a certain degree of correlation between $\langle \mu_{\rm B} \rangle$ and $\langle T \rangle$ can be also identified.

5 Summary and Conclusions

In this work we studied the kinetic freeze-out process using the UrQMD transport model[27, 28] combined with a coarse-graining approach,[30] focusing on Au+Au central collisions at five significant reaction energies in the range $\sqrt{s_{NN}} = 2.4 - 200\,{\rm GeV}$. We evaluated the kinetic decoupling distributions with respect to time, temperature and baryon chemical potential. In addition to showing these distributions, we presented the sequence of the average temperature and baryon chemical potential points at kinetic freeze-out with varying collision energy and a map of the density of the kinetic freeze-out points in the $(T, \mu_{\rm B})$ plane. The results are in reasonable agreements with the expectations based on physical considerations and with the concept of kinetic freeze-out as continuous process happening under different conditions due to the complex intertwining of the hadronic cross sections and the dynamics of the system. More details can be found in Ref. 24.

Finally, we would like to stress that the work that we just presented required a significant effort in terms of CPU time and data storage. High level computational facilities, like

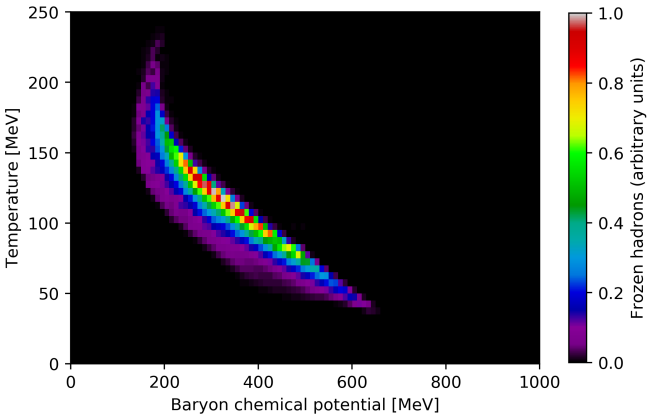


Figure 5. Profile of the kinetic freeze-out temperature and baryon chemical potential at $|y| < 0.2$ for Au+Au collisions at $\sqrt{s_{NN}} = 19.6\,\mathrm{GeV}$.

those available at the John von Neumann Institute for Computing, are not only necessary for the mere feasibility of this kind of studies, but, once the model has been developed and the code optimised in order to avoid any waste of resources, the quality and the accuracy of the results are directly correlated with the available computing power: the more, the better.

Acknowledgements

We gratefully acknowledge Stephan Endres for providing the coarse-graining numerical code which served as a basis for the present work. G. Inghirami is supported by the Academy of Finland, Project no. 297058. P. Hillmann acknowledges support by the GSI in cooperation with the John von Neumann Institute for Computing; she also acknowledges support from the HGS-HIRe and FIGSS graduate schools. B. Tomášik acknowledges support by the grant No. 17-04505S from the Czech Science Foundation. The computational resources were provided by the Center for Scientific Computing (CSC) of the Goethe University Frankfurt and by the Frankfurt Institute for Advanced Studies (FIAS). This work was supported by the COST Action CA15213 (THOR).

References

1. R. Hagedorn, *Statistical thermodynamics of strong interactions at high-energies*, Nuovo Cim. Suppl. **3**, 147–186, 1965.
2. T. Matsui and H. Satz, *J/ψ Suppression by Quark-Gluon Plasma Formation*, Phys. Lett. B **178**, 416–422, 1986.
3. J. Adams *et al.*, *Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR Collaboration's critical assessment of the evidence from RHIC collisions*, Nucl. Phys. A **757**, 102–183, 2005.
4. U. W. Heinz, *The Quark gluon plasma at RHIC*, Nucl. Phys. A **721**, 30–39, 2003.

5. W. Greiner, S. Schramm, and E. Stein, *Quantum chromodynamics*, 3rd Edition, 2007.
6. M. Tanabashi *et al.*, *Review of Particle Physics*, Phys. Rev. D **98**, 030001, 2018.
7. U. W. Heinz, *The Strongly coupled quark-gluon plasma created at RHIC*, J. Phys. A **42**, 214003, 2009.
8. P. Romatschke, *New Developments in Relativistic Viscous Hydrodynamics*, Int. J. Mod. Phys. E **19**, 1–53, 2010.
9. C. Gale, S. Jeon, and B. Schenke, *Hydrodynamic Modeling of Heavy-Ion Collisions*, Int. J. Mod. Phys. A **28**, 1340011, 2013.
10. E. Shuryak, *The sounds of the Little and Big Bangs*, Universe **3**, 75, 2017.
11. U. W. Heinz, *Towards the Little Bang Standard Model*, J. Phys. Conf. Ser. **455**, 012044, 2013.
12. V. K. Magas, C. Anderlik, L. P. Csernai, F. Grassi, W. Greiner, Y. Hama, T. Kodama, Z. I. Lazar, and H. Stoecker, *Kinetic freezeout models*, Acta Phys. Hung. A **9**, 193–216, 1999.
13. E. Molnar, L. P. Csernai, V. K. Magas, A. Nyiri, and K. Tamosiunas, *Covariant description of kinetic freeze out through a finite space-like layer*, Phys. Rev. C **74**, 024907, 2006.
14. J. Cleymans and K. Redlich, *Chemical and thermal freezeout parameters from 1-A/GeV to 200-A/GeV*, Phys. Rev. C **60**, 054908, 1999.
15. P. Braun-Munzinger, J. Stachel, and C. Wetterich, *Chemical freezeout and the QCD phase transition temperature*, Phys. Lett. B **596**, 61–69, 2004.
16. S. Chatterjee, S. Das, L. Kumar, D. Mishra, B. Mohanty, R. Sahoo, and N. Sharma, *Freeze-Out Parameters in Heavy-Ion Collisions at AGS, SPS, RHIC, and LHC Energies*, Adv. High Energy Phys. **2015**, 349013, 2015.
17. M. Bleicher, J. Steinheimer, and R. Stock, *The QCD Phase Diagram from Statistical Model Analysis*, in: Walter Greiner Memorial Volume, World Scientific Publishing Co. Pte. Ltd., 41–64, 2018.
18. D. Prorok, *Thermal freeze-out versus chemical freeze-out revised*, Acta Phys. Polon. B **40**, 2825–2842, 2009.
19. W. Broniowski and W. Florkowski, *Strange particle production at RHIC in a single freezeout model*, Phys. Rev. C **65**, 064905, 2002.
20. A. Baran, W. Broniowski, and W. Florkowski, *Description of the particle ratios and transverse momentum spectra for various centralities at RHIC in a single freezeout model*, Acta Phys. Polon. B **35**, 779–798, 2004.
21. E. Schnedermann, J. Sollfrank, and U. W. Heinz, *Thermal phenomenology of hadrons from 200-A/GeV S+S collisions*, Phys. Rev. C **48**, 2462–2475, 1993.
22. L. Kumar, *Systematics of Kinetic Freeze-out Properties in High Energy Collisions from STAR*, Nucl. Phys. A **931**, 1114–1119, 2014.
23. I. Melo and B. Tomášik, *Reconstructing the final state of Pb+Pb collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, J. Phys. G **43**, 015102, 2016.
24. G. Inghirami, P. Hillmann, B. Tomášik, and M. Bleicher, *Temperatures and chemical potentials at kinetic freeze-out in relativistic heavy ion collisions from coarse grained transport simulations*, arXiv:1909.00643 [hep-ph], 2019.
25. L. V. Bravina *et al.*, *Local thermodynamical equilibration in central Au+Au collisions at AGS*, Phys. Lett. B **434**, 379–387, 1998.
26. L. V. Bravina *et al.*, *Local thermal and chemical equilibration and the equation of state in relativistic heavy ion collisions*, J. Phys. G **25**, 351–361, 1999.

27. S. A. Bass *et al.*, *Microscopic models for ultrarelativistic heavy ion collisions*, Prog. Part. Nucl. Phys. **41**, 255–369, 1998.
28. M. Bleicher *et al.*, *Relativistic hadron hadron collisions in the ultrarelativistic quantum molecular dynamics model*, J. Phys. G **25**, 1859–1896, 1999.
29. P. Huovinen, M. Belkacem, P. J. Ellis, and J. I. Kapusta, *Dileptons and photons from coarse grained microscopic dynamics and hydrodynamics compared to experimental data*, Phys. Rev. C **66**, 014903, 2002.
30. S. Endres, H. van Hees, J. Weil, and M. Bleicher, *Coarse-graining approach for dilepton production at energies available at the CERN Super Proton Synchrotron*, Phys. Rev. C **91**, 054911, 2015.
31. H. Petersen, J. Steinheimer, G. Burau, M. Bleicher, and H. Stocker, *A Fully Integrated Transport Approach to Heavy Ion Reactions with an Intermediate Hydrodynamic Stage*, Phys. Rev. C **78**, 044901, 2008.
32. S. Endres, H. van Hees, J. Weil, and M. Bleicher, *Dilepton production and reaction dynamics in heavy-ion collisions at SIS energies from coarse-grained transport simulations*, Phys. Rev. C **92**, 014911, 2015.
33. G. Inghirami, H. van Hees, S. Endres, J. M. Torres-Rincon, and M. Bleicher, *Elliptic flow and R_{AA} of D mesons at FAIR comparing the UrQMD hybrid model and the coarse-graining approach*, Eur. Phys. J. C **79**, 52, 2019.
34. C. Eckart, *The Thermodynamics of irreversible processes. III. Relativistic theory of the simple fluid*, Phys. Rev. **58**, 919–924, 1940.
35. D. Zschiesche, S. Schramm, J. Schaffner-Bielich, H. Stoecker, and W. Greiner, *Particle ratios at RHIC: Effective hadron masses and chemical freezeout*, Phys. Lett. B **547**, 7–14, 2002.
36. G. Agakishiev *et al.*, *The High-Acceptance Dielectron Spectrometer HADES*, Eur. Phys. J. A **41**, 243–277, 2009.
37. C. Blume, *Review of results from the NA49 collaboration*, J. Phys. G **31**, S685–S692, 2005.
38. G. Odyniec, *RHIC Beam Energy Scan Program: Phase I and II*, PoS **CPOD2013**, 043, 2013.
39. A. Andronic, P. Braun-Munzinger, and J. Stachel, *Hadron production in central nucleus-nucleus collisions at chemical freeze-out*, Nucl. Phys. A **772**, 167–199, 2006.

Cut-Off Effects on the QCD Thermal Transition as a Function of Quark Masses and Chemical Potential

Francesca Cuteri[1], **Owe Philipsen**[1,2], **Alena Schön**[1], **and Alessandro Sciarra**[1]

[1] Institut für Theoretische Physik, Goethe Universität Frankfurt,
Max-von-Laue-Str. 1, 60438 Frankfurt, Germany
E-mail: {*cuteri, philipsen, schoen, sciarra*}*@th.physik.uni-frankfurt.de*

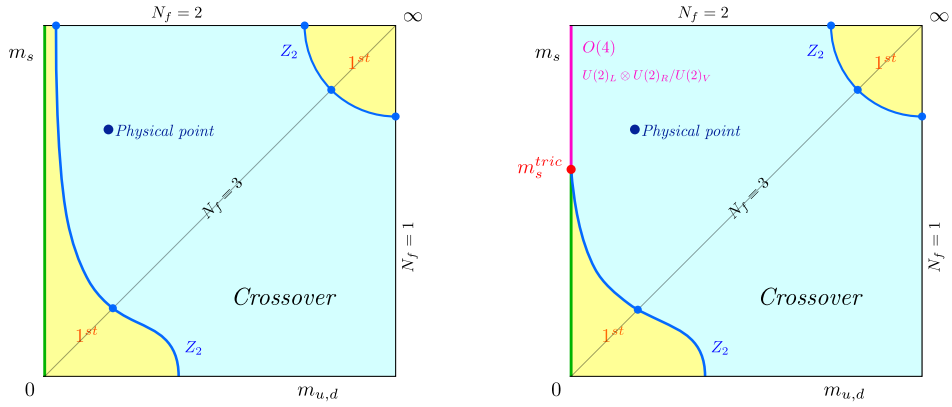
[2] John von Neumann Institute for Computing (NIC), GSI, Planckstr. 1, 64291 Darmstadt, Germany

We report on the status of a long term project to determine the nature of the thermal transition in QCD (the fundamental theory of strongly interacting matter composed of quarks and gluons) as a function of the number of quark flavours, their masses, imaginary chemical potential for baryon number and the lattice spacing. Our knowledge on the order of the thermal QCD transition depending on these parameters is summarised in what is known as *Columbia plot*. Besides showing the structure of the theory, it is important to constrain the QCD phase diagram realised by nature, which cannot be simulated directly due to a severe sign problem at real baryon chemical potentials. Having determined the qualitative structure of the Colombia plot in earlier studies, current efforts focus on reducing the lattice spacing and understanding discretisation effects, which need to be removed to arrive at continuum results.

1 Introduction

Quantum Chromodynamics is the fundamental theory of the strong interactions governing the forces between nuclear and subnuclear particles. Its fundamental degrees of freedom are light u- and d-quarks, a heavier s-quark and gluons, which are the force carriers in this quantum field theory. The coupling strength of the interactions depends 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 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 constituents form a so-called quark gluon plasma. The QCD phase diagram determines the form of matter under different conditions as a function of temperature, T, and matter density parametrised by a chemical potential μ for quark number, which is one third that of the conserved baryon number. Whether and where the hadronic phase and the quark gluon plasma are separated 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. The most reliable approach is by Monte Carlo simulation of a reformulation of the theory on a space-time grid, lattice QCD.

Unfortunately, the so-called sign problem prohibits straightforward simulations at finite baryon density. For this reason, knowledge of the thermal phase transition at zero density as a function of the theory's parameters, also for unphysical values, is of great importance to constrain and anchor research in the finite density direction. It is well-known that the physical point of QCD (with quark mass values realised in nature) displays only an analytic, smooth crossover between the hadronic and the plasma regions, without a non-analytic phase transition.[1] However, the order of the finite temperature phase transition at



(a) First order scenario in the $m_s - m_{u,d}$ plane

(b) Second order scenario in the $m_s - m_{u,d}$ plane

Figure 1. Two possible scenarios for the order of the QCD thermal phase transition as a function of the quarks masses. Indicated in Fig. 1(b) are also plausible universality classes for the second order line at $m_{u,d} = 0$.

zero density depends on the quark masses as is schematically shown in Fig. 1(a), 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 some global symmetry can be defined, and for three degenerate quarks one numerically finds, on rather coarse lattices, 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 model.[2–4] The critical lines delimit 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 $N_\tau = 4$ lattices[5] and puts the physical quark mass configuration in the crossover region. The chiral critical line recedes with decreasing lattice spacing,[6, 7] which is parametrised by increasing temporal lattice extent N_τ (see below): for $N_f = 3$, on the critical point, the quark masses correspond to pion mass values $m_\pi(N_\tau = 4)/m_\pi(N_\tau = 6) \sim 1.8$. 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. 1(b). Clarifying this question is important because of the proximity of the critical line to the physical point.

2 The General Strategy

All numerical simulations have been performed using the publicly available[8] OpenCL-based code CL^{2}QCD,[9] which is optimised to run efficiently on AMD GPUs and provides,

	Crossover	1^{st} triple	Tricritical	$3D$ Ising
B_4	3	1.5	2	1.604
ν	–	1/3	1/2	0.6301(4)

Table 1. Critical values of ν and $B_4 \equiv B_4(\beta_c, X_c, \infty)$ for some universality classes.[11]

among other features, an implementation of the (R)HMC algorithm for unimproved (rooted staggered) Wilson fermions.

In our studies, the chemical potential has been kept fixed at either $\mu = 0$ or at the purely imaginary value $\mu = i\mu_i^{RW}, \mu_i^{RW} = \pi/3$. The latter choice is motivated by the fact that there is no sign problem at imaginary chemical potential, and that μ_i^{RW} constitutes a Roberge-Weiss phase boundary between different centre-sectors of the QCD symmetry group, which are periodically repeated as imaginary chemical potential is increased.[10]

Temperature T is related to the lattice gauge coupling β and the temporal lattice extent according to

$$T = 1/(a(\beta)N_\tau) \tag{1}$$

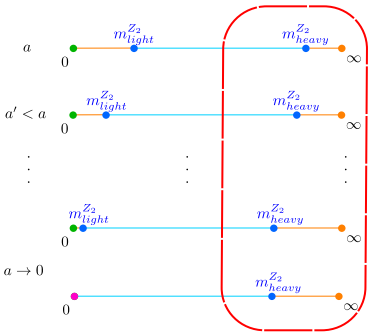
Approaching the continuum limit at fixed temperature is realised by taking $\beta \to \infty$, $a(\beta) \to 0, N_\tau \to \infty$. Our studies in Sec. 5 and in Sec. 4 are conducted at a fixed temporal extent N_τ of the lattices (no continuum limit is attempted in these cases), while for the study described in Sec. 3 three different values of N_τ are considered. The ranges in mass m or hopping parameter κ (parametrising mass in the Wilson discretisation) and gauge coupling constant β are always dictated by our purpose of locating the chiral/deconfinement phase transition.

In order to locate the chiral/deconfinement phase transition and to identify its order, a common strategy is adopted, which consists of a finite size scaling analysis (FSS) of the third and/or fourth standardised moments of the distribution of an order parameter for the transition. The n$^{\text{th}}$ standardised moment, given the distribution of a generic observable $\mathcal{O}$, is expressed as

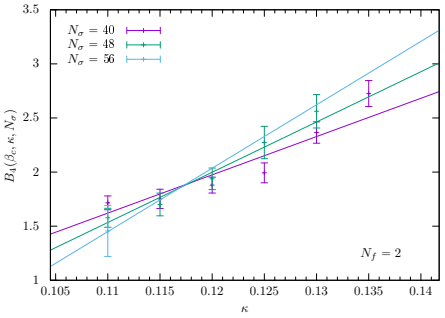
$$B_n(\mathcal{O}) = \frac{\langle (\mathcal{O} - \langle \mathcal{O} \rangle)^n \rangle}{\left\langle (\mathcal{O} - \langle \mathcal{O} \rangle)^2 \right\rangle^{n/2}} \tag{2}$$

and we analyse its dependence on some parameter $X \in \{m, \kappa, \beta\}$ and on the volume. We will introduce an exact or approximate order parameter $\mathcal{O}$ for each investigation in the corresponding sections. However, in all cases, in order to extract the order of the transition as a function of the bare quark mass and/or number of flavours, we considered the kurtosis $B_4(\mathcal{O})$ of the sampled distribution of $\mathcal{O}$.

In the thermodynamic limit $N_\sigma \to \infty$, the universal values taken by the kurtosis B_4 and by the critical exponent ν for our cases of interest are well known results listed in Tab. 1. However, the discontinuous step function characterising the thermodynamic limit is smeared out to a smooth function as soon as a finite volume is considered and a FSS is needed. In all cases we varied the spatial extent of the lattice N_σ such that the aspect ratios, governing the size of the box in physical units at finite temperature, was in the range



(a) Shift of the critical m_{Z_2} masses at $N_f = 2$ towards the continuum limit



(b) FSS of B_4 and fit

Figure 2. Features and preliminary results on the Z_2 boundary in the high mass corner of the Columbia plot.

$N_\sigma/N_\tau \in [2-5]$. In the vicinity of a critical point, the kurtosis can be expanded in powers of the scaling variable $x = (X - X_c)N_\sigma^{1/\nu}$ and, for large enough volumes, the expansion can be truncated after the linear term,

$$B_4(\beta_c, X, N_\sigma) \simeq B_4(\beta_c, X_c, \infty) + c(X - X_c)N_\sigma^{1/\nu} \tag{3}$$

In our case, for the studies described in Sec. 3 and in Sec. 4, the critical value for $X_c = \kappa, m$ corresponds to a second order phase transition in the 3D Ising universality class, so that one can fix $B_4 \approx 1.604$ and $\nu \approx 0.63$ and perform the fit to Eq. 3 with the sole aim of extracting X_c (and c).

3 Updates on the Columbia Plot: Z_2 Boundary in the High Mass Corner

The cut-off effects that quantitatively affect our picture of the QCD phase structure have been investigated in previous studies in the upper right corner of the Columbia plot, and the Z_2 transitions were already observed to shift to smaller masses for $N_f = 2, 2+1, 3$ at $\mu = 0$.[6, 12–14] A sketch of this behaviour for $N_f = 2$ is given in Fig. 2(a). No continuum extrapolation is available yet.

In this case the norm of the Polyakov loop, $||L||$, was used as approximate order parameter for the deconfinement phase transition. The case of $N_f = 2$ degenerate quarks was addressed and, with a scan in κ, the location of the critical $\kappa^{Z_2}_{\text{heavy}}$ endpoint on $N_\tau = 6, 8, 10$

N_τ	$\kappa^{Z_2}_{\text{heavy}}$	a [fm]	am_π	m_π [MeV]	V_{min} [fm^3]	L_{min} [fm]
6	0.0834(64)	[0.1175(6):0.1232(5)]	[3.4709(3):2.2411(3)]	[5910(30):3638(15)]	76	4.23
8	0.1145(32)	[0.0882(4):0.0960(8)]	[2.1310(6):1.2325(4)]	[4830(20):2570(20)]	44	3.54
10	0.1255(43)	[0.0690(10):0.0762(11)]	[1.8734(3):0.9284(5)]	[5430(80):2440(40)]	42	3.49

Table 2. Results on V_{min} from fits of the kurtosis.

was investigated with the aim to monitor and possibly model cut-off effects. A previous report on this project is to be found in Ref. 15, while here we focus on the impact of finite size effects in this investigation. It is an additional difficulty of the heavy mass region that pions cannot be resolved on our lattices up to $N_\tau \approx 10$. And, while $a \to 0$ with growing N_τ, the necessity of keeping the relation $1 \ll N_\tau \ll N_\sigma$ satisfied forces us to use larger N_σ values to keep the size of the box fixed in physical units. This has to be satisfied already for the smallest of the volumes in our FSS analysis.

In view of the increasing cost of simulations, some work was invested in devising an alternative and possibly cheaper strategy to locate $\kappa_{\text{heavy}}^{Z_2}$. One could, indeed, first identify at any fixed value of the bare mass some *minimal physical volume* V_{min}, characterised by that it allows a reliable extraction of $\kappa_{\text{heavy}}^{Z_2}$ from a linear fit of the kurtosis. At a different κ or N_τ, it should then be enough to *e. g.* reweight the effective potential V_{eff} at just one fixed $V \gtrsim V_{\text{min}}$ as in Ref. 13 to locate the phase transition and understand its nature. In practice we start by using a modified fit ansatz for the kurtosis in the vicinity of the critical point[16]

$$B_4(\kappa, N_\sigma) = \left[B_4(\kappa_{\text{heavy}}^{Z_2}, \infty) + c\,(\kappa - \kappa_{\text{heavy}}^{Z_2}) N_\sigma^{(1/\nu)} \right] (1 + B N_\sigma^{y_t - y_h})$$

which incorporates the finite volume effect for generic observables which are a mixture of energy-like and magnetisation-like operator and where the value of the exponent $y_t - y_h$ is fixed by universality. Then we estimate V_{min} by excluding one-by-one the smallest physical volumes in the fit, until the value of the coefficient of the correction term is compatible with zero. Preliminary results are collected in Tab. 2 and one example of the performed fits is provided in Fig. 2(b).

4 Updates on an Alternative $m_{u,d}$–N_{f} Columbia Plot: Z_2 Boundary in the Chiral Limit

In this study we treat N_{f} as a continuous real parameter of some statistical system behaving, at any integer N_{f} value, as QCD at zero density, with N_{f} mass-degenerate fermion species[17]

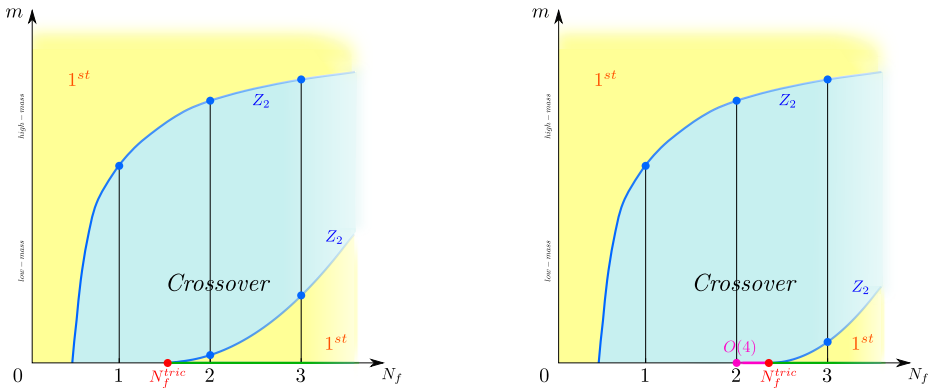
$$Z_{N_{\text{f}}}(m) = \int \mathcal{D}U \left[\det M(U, m)\right]^{N_{\text{f}}} e^{-\mathcal{S}_{\text{G}}} \tag{4}$$

Within this framework, the two considered scenarios for the Columbia plot can be put in one-to-one correspondence with the two sketches for the order of the thermal phase transition in the (m, N_{f})-plane displayed in Fig. 3.

We employ unimproved staggered fermions and use the RHMC algorithm[18] to simulate any number N_{f} of degenerate flavours. Our original strategy was to find out for which tricritical value $N_{\text{f}}^{\text{tric}}$ the phase transition displayed by this system changes from first-order to second-order, by mapping out the Z_2 phase boundary. The extrapolation to the chiral limit with known tricritical exponents can then decide between the two scenarios, depending on whether $N_{\text{f}}^{\text{tric}}$ is larger or smaller than 2.

While the tricritical scaling region was found to be very narrow already on coarse $N_\tau = 4$ lattices, results at larger m and N_{f} were found to feature, over a much wider region, a remarkable linear behaviour, which was not expected on universality grounds, see Fig. 4 .

What our findings suggest is that, if it is reasonable to expect both linearity within some range in N_{f} and tricritical scaling more in the chiral limit, then one would be able to make



(a) First order scenario in the $m_{u,d} - N_{\rm f}$ plane

(b) Second order scenario in the $m_{u,d} - N_{\rm f}$ plane

Figure 3. The two considered possible scenarios for the order of the QCD thermal phase transition as a function of the light-quarks mass and the number of fermion flavours.

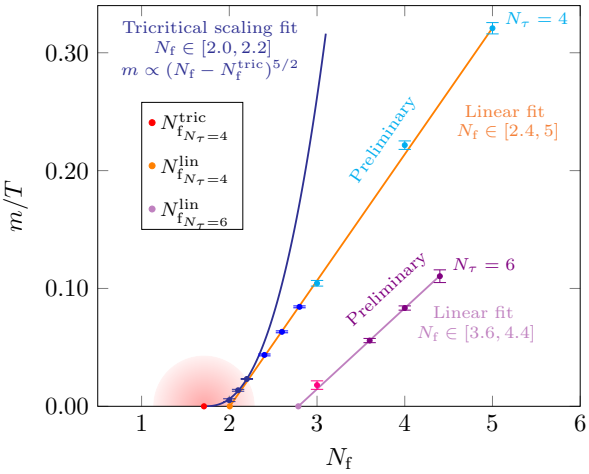


Figure 4. The Z_2 boundary in the $m/T - N_{\rm f}$ plane for $N_\tau = 4, 6$. The dark blue line represents the tricritical extrapolation to the chiral limit as in Ref. 17. The orange line represents a linear extrapolation based on m_{Z_2} in the $N_{\rm f}$ range $2.4 - 5$ using also newly simulated points. The violet line represents a linear extrapolation on the basis of m_{Z_2} in the $N_{\rm f}$ range $3.6 - 4.4$. The magenta point at $N_\tau = 6$ and $N_{\rm f} = 3$ is borrowed from Ref. 6.

use of a linear extrapolation to $m = 0$, to at least get $N_{\rm f}^{\rm lin}$ as an upper bound for $N_{\rm f}^{\rm tric}$, out of much more affordable simulations and possibly without even simulating at noninteger numbers of flavours.

For as long as the upper bound from the linear extrapolation keeps lying at $N_{\rm f} < 2$, while one simulates at larger and larger N_τ values towards the continuum limit, one can

infer that the transition in the $N_{\rm f} = 2$ chiral limit is of first order. However, should our linear extrapolation give $N_{\rm f}^{\rm lin} \gtrsim 2$, then knowledge of the size of the scaling region is necessary to draw conclusions.

Our results are reported in Fig. 4. The first important thing to observe is that, while tricritical extrapolation for $N_\tau = 4$ resulted in $N_{\rm f}^{\rm tric} < 2$, providing an independent confirmation for the first order scenario being realised on coarse lattices, a linear extrapolation to the chiral limit using $N_{\rm f} \in [2.4, 5.0]$, results in $N_{\rm f}^{\rm lin} = 2$ within errors. Strictly speaking, by just considering $N_\tau = 4$ results, one would conclude that the linear extrapolation alone cannot give conclusive answers on the order of the $N_{\rm f} = 2$ transition in the chiral limit. However, results on finer lattices were produced as well. On $N_\tau = 6$, what we observe is that data within the range $N_{\rm f} \in [3.6, 4.4]$ certainly do not fall in the tricritical scaling region, but they do exhibit linear scaling. Moreover, if we consider the result for $N_{\rm f} = 3$ for the same discretisation from the literature,[6] we can see it is fully consistent with our linear extrapolation. Finally, the most important aspect of this result is that, linearly extrapolating at $N_\tau = 6$, we get $N_{\rm f}^{\rm lin} \lesssim 3$, namely quite far to the right of $N_{\rm f} = 2$.

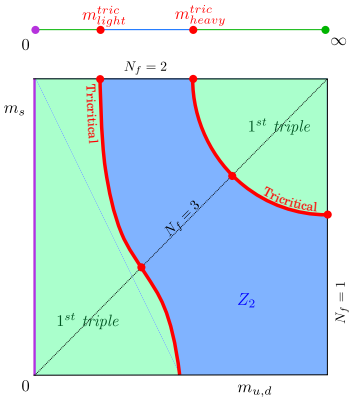
5 Updates on the Extended Columbia Plot: Roberge-Weiss Endpoint

The Columbia plot at $\mu_i = \mu_i^{RW}$ displayed in Fig. 5(a) looks similar to the one in Fig. 1(a), but with the Z_2 lines replaced by tricritical lines, first order triple regions that are wider than at $\mu = 0$ and a second order Z_2 region at intermediate values for the quark masses. In this case the imaginary part of the Polyakov loop $L_{\rm Im}$ was measured as order parameter for the Roberge-Weiss phase transition. Once again, we focused on the case of $N_{\rm f} = 2$ degenerate unimproved staggered quarks, and tried to locate, with a scan in mass, the tricritical points $m_{\rm heavy}^{\rm tricr.}$ and $m_{\rm light}^{\rm tricr.}$ on $N_\tau = 6$ lattices as already done for other discretisations and N_τ values.[19–21] A previous report on this project is to be found in Ref. 23.

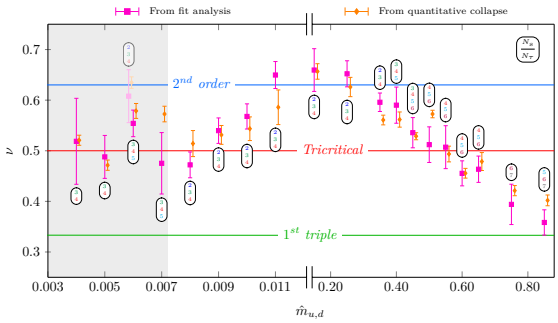
For each value of $m_{u,d}$, simulations were performed at a fixed temporal lattice extent $N_\tau = 6$ and at a fixed value of the chemical potential $a\mu_i^{RW} = \pi/6$. The extraction of the critical exponent ν was accomplished both with the kurtosis fit procedure and with a quantitative data collapse described in Ref. 24. Results for the critical exponent ν are reported in Fig. 5(b). Since results from either kind of analysis happen to agree within a 1σ discrepancy in all (but one) case, they are combined to obtain the final answer on ν. To comment more on our results, it is important to stress that for the FSS at least three volumes are necessary for safe conclusions and we used $N_\sigma^{\rm min} = 12, 18, 24$ and $N_\sigma^{\rm max} = 30, 36, 42$, depending on the mass. For each lattice size, 3 to 8 values of β around the critical temperature were simulated, each with 4 Markov chains.

In order to decide when to stop accumulating statistics, for large (small) masses the kurtosis of the imaginary part of Polyakov loop was required to be compatible on all the chains within 2 (3) standard deviations. Since this condition can be fulfilled also at a very poor statistics, due to large errors, a further empirical requirement is that values of the kurtosis from different chains must span, errors included, an interval not wider than 0.5.

As indicated in Fig. 5(b) it is still not possible to give the two tricritical masses with their statistical error, because in the light mass range no simulation point falls on the first order triple line. This is because larger and larger volumes are needed in ranges where the transition goes from tricritical to weakly first-order. We can however quote a result with asymmetric errors to reflect this uncertainty, $m_{\pi{\rm light}}^{\rm tricr.} = 328^{+44}_{-88}$ MeV.[24] One can now



(a) The Columbia plot of the Roberge-Weiss plane, *i. e.* the order of the thermal phase transition in the $m_s - m_{u,d}$ plane at $\mu_i = \mu_i^{RW}$. The $N_\mathrm{f} = 2$ case is highlighted.



(b) Critical exponent ν as a function of the bare quark mass m. For the sake of readability, the mass axis has been broken and two different scales have been used. Shaded points correspond to preliminary results for which higher statistics is needed.

Figure 5. Features and preliminary results on the tricritical endpoints in the Roberge-Weiss Columbia plot.

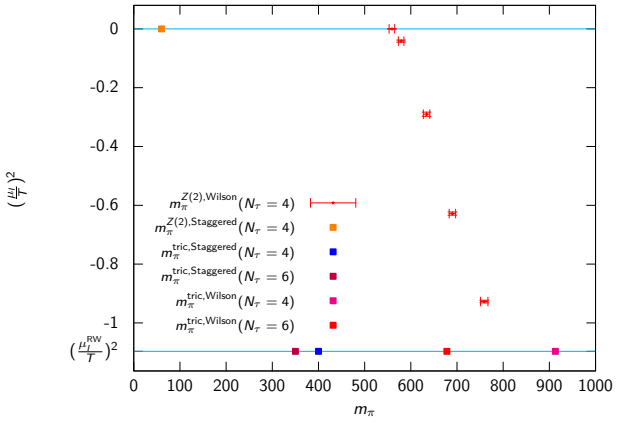


Figure 6. Collection of results on the Z_2 critical line/points in the $m_\pi - \mu^2$ plane. Results from various discretisations and different N_τ can be compared.[19, 20, 22, 25, 26]

compare with results from other discretisations and/or at other N_τ values, as we do in Fig. 6. We now clearly see the shift of $m_{\pi\,\mathrm{light}}^{\mathrm{tricr.}}$ towards smaller masses when going from $N_\tau = 4$ to $N_\tau = 6$. One can also compare the Wilson *versus* staggered discretisations: the shift amounts to 35 % (14 %) of the value for Wilson (staggered). At large masses the value found for $m_{\pi\,\mathrm{heavy}}^{\mathrm{tricr.}}$ is still affected by large cut-off effects, as indicated by $am_\pi > 1$. This means finer lattices are necessary to resolve the pion.

6 Conclusions

Our systematic study of the order of the QCD thermal transition and its dependence on the parameters of the theory as well as on the simulation volume and lattice spacing has revealed two main insights: the first of these is rather unfortunate technically, as it indicates very strong cut-off effects on the second-order boundary lines in the Columbia plot, both at zero and non-zero baryon density. Combined with the large volumes required to decide on the order of the transition, this renders future investigations extremely compute-expensive, and thus slow, before one can extrapolate to the desired continuum results. On the positive side, we have at least arrived at a full understanding of the qualitative features of cut-off effects on the Columbia plot. In particular, the chiral first-order region shrinks strongly with decreasing lattice spacing, for all N_f and at zero as well as non-zero baryon density. This is valuable input also to constrain the physical QCD phase diagram. Finally, at least in the heavy mass region one might hope to find a continuum limit for the second-order boundary within the next couple of years, which would serve as valuable benchmark for, *e. g.* functional renormalisation group methods in the continuum.

Acknowledgements

This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 315477589 -TRR 211. We also acknowledge support by the Helmholtz International Center for FAIR within the LOEWE program of the State of Hesse.

References

1. Y. Aoki, G. Endrődi, 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, 2006.
2. F. Karsch, E. Laermann, and C. Schmidt, *The Chiral critical point in three-flavor QCD*, Phys. Lett. B **520**, 41, 2001.
3. P. de Forcrand and O. Philipsen, *The QCD phase diagram for three degenerate flavors and small baryon density*, Nucl. Phys. B **673**, 170, 2003.
4. S. Kim, P. de Forcrand, S. Kratochvila, and T. Takaishi, *The 3-state Potts model as a heavy quark finite density laboratory*, PoS LAT **2005**, 166, 2006.
5. P. de Forcrand and O. Philipsen, *The Chiral critical line of N(f) = 2+1 QCD at zero and non-zero baryon density*, JHEP **0701**, 077, 2007.
6. P. de Forcrand, S. Kim, and O. Philipsen, *A QCD chiral critical point at small chemical potential: Is it there or not?*, PoS LAT **2007**, 178, 2007.
7. G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabo, *The Nature of the finite temperature QCD transition as a function of the quark masses*, PoS LAT **2007**, 182, 2007.
8. M. Bach, F. Cuteri, C. Czaban, C. Pinke, A. Sciarra *et al.*, *CL^{2}QCD: lattice QCD using OpenCL* (since 2011), `https://github.com/AG-Philipsen/cl2qcd`.
9. O. Philipsen, C. Pinke, A. Sciarra, and M. Bach, *CL^{2}QCD - Lattice QCD based on OpenCL*, PoS LAT **2014**, 038, 2014.

10. A. Roberge and N. Weiss, *Gauge Theories With Imaginary Chemical Potential and the Phases of QCD*, Nucl. Phys. B **275**,734 , 1986, doi:10.1016/0550-3213(86)90582-1.
11. A. Pelissetto and E. Vicari, *Critical phenomena and renormalization group theory*, Phys. Rept. **368**, 549, 2002.
12. P. de Forcrand and O. Philipsen, *The curvature of the critical surface* $(m_{ud}, m_s)^{crit}(\mu)$*: a progress report*, PoS LAT **2008**, 208, 2008, arXiv:0811.3858 [hep-lat].
13. H. Saito, S. Ejiri, S. Aoki, T. Hatsuda, K. Kanaya, Y. Maezawa, H. Ohno, and T. Umeda (WHOT-QCD Collaboration), *Phase structure of finite temperature QCD in the heavy quark region*, Phys. Rev. D **84**, 054502, 2011, [Erratum: Phys. Rev. D **85**, 079902, 2012], arXiv:1106.0974 [hep-lat].
14. M. Fromm, J. Langelage, S. Lottini, and O. Philipsen, *The QCD deconfinement transition for heavy quarks and all baryon chemical potentials*, JHEP **01**, 042, 2012, arXiv:1111.4953 [hep-lat].
15. C. Czaban and O. Philipsen, *The QCD deconfinement critical point for* $N_\tau = 8$ *with* $N_f = 2$ *flavours of unimproved Wilson fermions*, PoS LAT **2016**, 056, 2016, arXiv:1609.05745 [hep-lat].
16. X. Y. Jin, Y. Kuramashi, Y. Nakamura, S. Takeda, and A. Ukawa, *Critical point phase transition for finite temperature 3-flavor QCD with nonperturbatively O(a) improved Wilson fermions at* $N_t = 10$, Phys. Rev. D **96**, 034523, 2017, arXiv:1706.01178 [hep-lat].
17. F. Cuteri, O. Philipsen, and A. Sciarra, *The QCD chiral phase transition from non-integer numbers of flavors*, Phys. Rev. D **97**, 114511, 2018, arXiv:1711.05658 [hep-lat].
18. A. D. Kennedy, I. Horvath, and S. Sint, Nucl. Phys. Proc. Suppl. **73**, 834, 1999, arXiv:hep-lat/9809092.
19. O. Philipsen and C. Pinke, *The nature of the Roberge-Weiss transition in* $N_f = 2$ *QCD with Wilson fermions*, Phys. Rev. D **89**, 094504, 2014, arXiv:1402.0838 [hep-lat].
20. C. Czaban, F. Cuteri, O. Philipsen, C. Pinke, and A. Sciarra, *Roberge-Weiss transition in* $N_f = 2$ *QCD with Wilson fermions and* $N_\tau = 6$, Phys. Rev. D **93**, 054507, 2016, arXiv:1512.07180 [hep-lat].
21. C. Bonati, G. Cossu, M. D'Elia, and F. Sanfilippo, *Roberge-Weiss endpoint in* $N_f = 2$ *QCD*, Phys. Rev. D **83**, 054505, 2011.
22. C. Bonati, P. de Forcrand, M. D'Elia, O. Philipsen, and F. Sanfilippo, *Chiral phase transition in two-flavor QCD from an imaginary chemical potential*, Phys. Rev. D **90**, 074030, 2014.
23. O. Philipsen and A. Sciarra, *Roberge-Weiss transition in* $N_f = 2$ *QCD with staggered fermions and* $N_\tau = 6$, PoS LAT **2016**, 055, 2016, arXiv:1610.09979 [hep-lat].
24. O. Philipsen and A. Sciarra, *Finite size and cut-off effects on the Roberge-Weiss transition in* $N_f = 2$ *QCD with Staggered fermions*, 2019, arXiv:1909.12253 [hep-lat].
25. C. Bonati, G. Cossu, M. D'Elia, and F. Sanfilippo, *The Roberge-Weiss endpoint in* $N_f = 2$ *QCD*, Phys. Rev. D **83**, 054505, 2011, arXiv:1011.4515 [hep-lat].
26. O. Philipsen and C. Pinke, *The* $N_f = 2$ *QCD chiral phase transition with Wilson fermions at zero and imaginary chemical potential*, Phys. Rev. D **93**, 114507, 2016, arXiv:1602.06129 [hep-lat].

Heavy Quarks, Multi-Level Algorithms and Tensor Networks: Developing Methods for Non-Perturbative Aspects of Quantum Field Theories

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

John von Neumann Institute for Computing, DESY, Platanenallee 6, 15738 Zeuthen, Germany
E-mail: rainer.sommer@desy.de

The particle physics group at the John von Neumann Institute for Computing (NIC) is concerned with quantum field theories, in particular with their non-perturbative aspects. Here we describe recent efforts of the group to develop and consolidate methods and concepts to reach the goal to understand non-perturbative phenomena in quantum field theories and to calculate observables with a high precision.

1 Introduction

The group covers a large range of topics, from the exploration of new strategies to treat quantum field theories non-perturbatively (Tensor networks, quantum computations) to the development of ever more efficient algorithms for Markov chain Monte Carlo to high precision applications concerning the determination of fundamental parameters of QCD, the muon anomalous magnetic moment, or weak decays of hadrons. In this contribution we give a short summary of recent work on three subjects: Tensor network methods, multilevel Monte Carlo and heavy sea quarks (quantum effects of heavy quarks).

2 Tensor Network Methods

Nowadays, the path integral is clearly the method of choice to evaluate lattice quantum field theories (LQFT). Still, in the beginning of LQFT it has been the Hamiltonian formalism which was used frequently. In practice, this approach was abandoned, due to the untractably large size of the Hilbert space. However, it has been realised in the last years that it is only a small set of all possible states which is relevant to obtain ground state properties. This small corner of the Hilbert space is build by states which obey the so called area law. These states can, in general, be constructed through tensor network states which become matrix product states (MPS) in one dimension. In this approach, the very high dimensional coefficient tensor needed in the Hamilton formalism is replaced by a product of complex matrices. By systematically increasing the size of these matrices and computing their entries through a variational method, minimising the energy as a cost function, ground state properties can be computed very precisely. In fact, mathematical theorems state that this procedure converges exponentially fast to the ground state of the considered Hamiltonian. In practice, the size of the used matrices are of order 100 which make such computations completely feasible and tensor networks have been used for condensed matter systems very successfully, see *e. g.* Ref. 1 for an introduction into tensor network techniques.

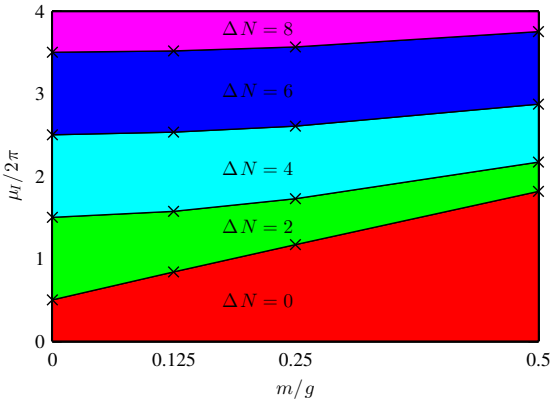


Figure 1. Phase diagram in the m/g - $\mu_I/2\pi$ plane with m denoting the fermion mass, g the coupling and μ_I the chemical potential. The black X's mark the computed data points, the different colours indicate the different phases.

As one of the first, the NIC group has adapted this approach of the Hamilton formalism using MPS for models in high energy physics, see Ref. 2 for a recent review. Since the Hamilton formalism is free of the sign problem, it offers the exciting possibility to study, in principle, questions where conventional Monte Carlo (MC) methods fail and which include non-zero baryon density as relevant for understanding the early universe; topological terms for the matter anti-matter asymmetry and real time evolutions of physical systems.

By computing the low-lying particle spectrum of the Schwinger model,[3] which has one space dimension, as a benchmark, a proof of principle could be provided that MPS can be used also for gauge theories. Followed by a calculation in this model at non-zero temperature,[4] the Schwinger model for two flavours of fermions was studied with a non-zero (isospin) chemical potential.[5] Addressing this question by conventional Monte Carlo methods is impossible due to a severe sign problem. It has therefore been very important to test, whether MPS can overcome this difficulty, or, whether the sign problem re-appears in a different way. Indeed, it was shown that the technique of MPS performs very well also in the situation with a non-zero chemical potential. Working first with a zero fermion mass, MPS results could be confronted with an analytically known expression and a complete agreement was found, demonstrating that MPS solves the problem of a non-zero chemical potential. Switching on a fermion mass, the phase diagram in the plane of chemical potential and fermion mass could be established,[5] see Fig. 1. Here no analytical result is available and only the usage of MPS made it possible to obtain this phase diagram making MPS or tensor networks the most promising tool to address important and so far intractable problems in high energy physics. Tensor network techniques were also used for non-abelian theories[6] and for studying the Schwinger model in presence of a topological term.[7] Still, a warning is in order since the computational cost of calculations for higher than one (space) dimension is presently too large to study systems of realistic size. However, there is a substantial amount of research ongoing to find better techniques for high dimensions as discussed in a recent workshop[a] co-organised by the NIC group. There, sev-

[a] `https://indico.desy.de/indico/event/21941/overview`

eral new ideas were presented which have the potential to make tensor networks practical also in higher dimensions.

3 Towards Multilevel Monte Carlo Methods for QCD

In this section we directly consider three space dimensions and remain within the framework of Monte-Carlo importance sampling methods for lattice field theory computations. In this framework, a major obstacle to progress is the deterioration of the signal of n-point correlation functions with increasing distance of these points. Large distances are needed to effectively study the ground state properties of the theory. Since lattice computations employ Monte Carlo sampling, uncertainties decrease with the inverse square root of the number of measurements – an expensive technique in the face of a noisy signal. By using the locality of the underlying theory, and designing algorithms where this fundamental property is manifest even in the presence of fermions, sampling strategies can be devised which have an improved convergence: depending on the number of regions, a convergence with the inverse number of measurements, or even a higher power becomes possible.

Such methods have been available since some time for pure gauge theory, where the formulation is manifestly local and also observables are typically easily decomposed into products of local contributions.[8, 9] Each of these local components can then be averaged over independently, leading to an exponential speed-up in the size of the observable.

Fermions are fundamentally different from bosons. They are integrated out analytically before formulating the Monte Carlo and thus lead to non-local contributions. Using domain decomposition techniques, we managed to propose a strategy that these multilevel methods become amenable to theories with fermions.[10, 11] In order to achieve this, the hadronic two-point functions as well as the contributions from the fermion sea to the path integral had to be factorised, such that the independent averages of each of the factors can be taken.

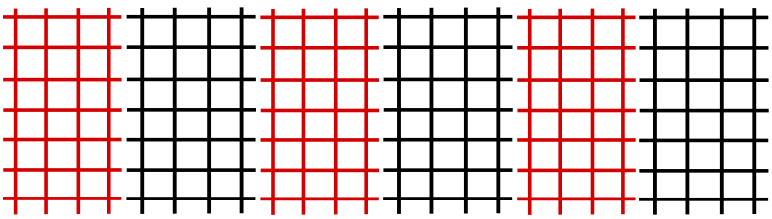


Figure 2. One-dimensional decomposition of the lattice into regions. If the thickness of the black regions is chosen sufficiently large, the red regions can be updated independently.

Contrary to the pure gauge case, the regions in which the independent updates are performed need to be separated by a certain distance, given by the correlation length (the inverse of the pion's mass). An illustration of the geometry is found in Fig. 2, where one envisions independent updates in the red regions, with the field variables in the black regions kept fixed. If one thinks of fermionic contributions in terms of paths on the lattice, it is obvious that as long as the quark does not travel between the red regions forth and back, such a path is trivially factorised in contributions which can be associated to a single red

region, together with the neighbouring black ones. Since longer paths typically contribute less than shorter ones, good approximations can be found which fulfil this property.

It becomes more difficult once we have to consider quark loops, as is necessary for the determinant. Here we managed to rewrite the corresponding terms using integrals over scalar fields, an idea which has already been used in the multiboson algorithm[12] and could demonstrate that the idea works in dynamical simulations.

Ordinary Monte Carlo simulations lead to uncertainties which are reduced with the inverse square root of the number of measurements $1/\sqrt{N}$. In the ideal case, for updates in n independent regions, such a method leads to a scaling with $1/N^{n/2}$. However, as we have seen, for fermions such regions will have to be sufficiently thick in order to profit from this technique. It needs to be seen if for a given observable a decomposition into sufficiently many regions can be found to profit from this algorithm.

4 Effects of Heavy Fermions on Low-Energy Physics and High-Energy Strong Coupling

The discretisation of space-time on a regular lattice leads to a Brilloin zone just like in crystals. Thus the high momentum, high energy behaviour is distorted. Also particles with large masses can't be simulated properly. For this reason lattice simulations include up, down and strange quarks, while the bottom and top quarks with masses above 4 GeV are excluded since they would contribute more distortion effects than physical effects. The charm quark with a mass of around our $M_{\rm charm}$=1.6 GeV is in between the typical scale of hadronic physics $E_{\rm had} \approx 0.5$ GeV and the achievable cutoffs (the edge of the Brilloin zone) of $E_{\rm cut} \approx 4$ GeV. It is thus a good question whether it is better to include the charm quark (often called $2+1+1$ simulations) or not $(2+1)$ and what are the uncertainties introduced by leaving it out.

In general, the answer to this question will depend on many details, from how one discretised QCD to which process one wants to predict from the simulations. Fortunately, there are also some rather universal statements which can be made. These have been the subject of our recent investigations and we give a brief account of them here.

4.1 The Effective Theory: decQCD

At low energies and momenta, say at $E_{\rm had}$ and below, there is a systematic expansion in terms of $y = E/M_{\rm charm}$, given in terms of an effective field theory, which excludes the charm quark, but has a few additional terms in its Lagrangian with coefficients proportional to $1/M^2_{\rm charm}$.

The leading order low energy effective theory is QCD_{n_ℓ}, where n_ℓ is the number of quarks in the Lagrangian. So far we had talked about $n_\ell = 3$. Next-to-leading order terms in the local effective Lagrangian are gauge-, Euclidean- and chiral-invariant local fields. These invariances allow only for fields, $\Phi_i(x)$, of at least dimension six. The Lagrangian may then be written as

$$\mathcal{L}_{\rm dec} = \mathcal{L}_{\mathrm{QCD}_{n_\ell}} + \frac{1}{M^2}\sum_i \omega_i \Phi_i + \mathrm{O(M^{-4})} \tag{1}$$

with dimensionless couplings ω_i which depend logarithmically on the mass M.

At the lowest order in $1/M$, a single coupling[b], $\bar{g}_{\rm dec}$, is adjusted such that the low energy physics of QCD_{n_ℓ} and $\mathrm{QCD}_{n_{\rm f}}$ match for energies $E \ll M$. It then suffices to require one physical low-energy observable to match, *e. g.* a physical coupling. Discussing the issue in perturbation theory,[13] Bernreuther and Wetzel chose the MOM-coupling as a physical coupling and worked out the matching of the $\overline{\rm MS}$ coupling. Meanwhile, the latter is known to high perturbative order,[14–20] which we use.

4.2 Non-Perturbative Investigation for $n_{\rm f} = 2 \;\to\; n_\ell = 0$

The main question is now, whether the effective theory is accurate at values of the quark mass around the charm mass, *i. e.* around 1.6 GeV. A direct test would require to simulate the $2+1+1$ theory which would be very expensive because of the light quarks.

We therefore investigated a very closely related model, namely QCD with $n_{\rm f} = 2$ heavy, mass-degenerate quarks.[21, 22] The decoupling is then $2 \to 0$ and the Lagrangian of the effective theory, $\mathcal{L}_{\rm dec}$, is the Yang-Mills one up to $1/M^2$ corrections. in $n_{\rm f} = 2$ we use quark mass values up to 1.8 GeV, slightly above the charm.

In principle any low-energy hadronic scale $\mathcal{S}(M)$ can be used to test decoupling, but in practice some choices are far superior to others. We want them to have good precision in the MC and have controllable lattice artifacts. In our purely gluonic effective theory, very good scales are defined in terms of the gradient flow.[23, 24] Here we report on two scales explicitly, which probe the theory in the low energy region at $E \approx 1/\sqrt{8t_0} \approx 0.5\,\mathrm{GeV}$ and $E \approx 1/\sqrt{8t_c} \approx 0.7\,\mathrm{GeV}$.

4.2.1 Simulations

In order to avoid the freezing of the topological charge for simulations with lattice spacings below $a = 0.05$ fm,[25] we adopt open boundary conditions in time and use the openQCD[c] package.[26]

Even after solving the topological charge problem, simulations remain difficult. An impression is given in Fig. 3. It shows the integrated autocorrelation times $\tau^O_{\rm int}$ of two observables, O. Their meaning is that (on average) after $2\tau^O_{\rm int}$ MC iterations a statistically independent value of the particular observable is obtained. $\tau_{\rm int}$ itself is difficult to obtain, but the trend in the figure is clear and confirms theoretical expectations, $\tau_{\rm int} \sim a^{-2}$. We then need many thousand MC iterations for reliable and precise results at the smallest a. Our error analysis adds a tail to the autocorrelation function as an estimate of the slow mode [d] contribution.[25] It is thus robust with respect to long autocorrelations.

Note also that due to the expensive nature of the simulations it is very important that projects with different physics goals coordinate and share resources, namely gauge configurations. In our case the group of F. Knechtli extended the simulations to larger quark masses in a separate NIC project whose gauge fields we were able to use.

[b] Again we refer to the theoretical situation where the first n_ℓ flavours are mass-less. In general, also the light quark masses have to be matched.

[c] `http://luscher.web.cern.ch/luscher/openQCD/`

[d] `http://www-zeuthen.desy.de/alpha/public_software/UWerrTexp.html`

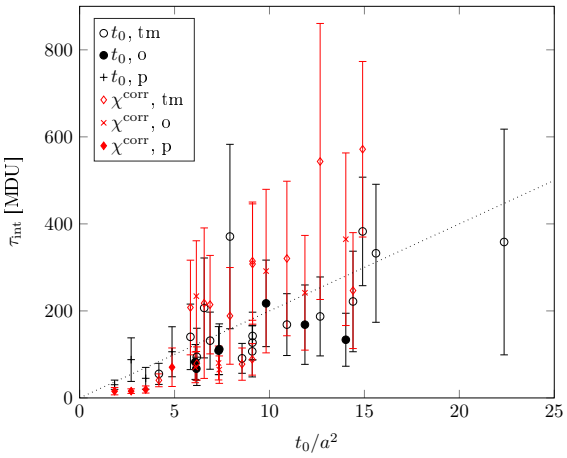


Figure 3. Autocorrelation times derived from observables which are expected to have large overlap with the slowest modes in the simulation are plotted as a function of $t_0(M)/a^2$.

4.3 Results

We turn to the results. The left part of Fig. 4 shows the ratio of two low energy scales in the $n_{\mathrm{f}} = 2$ theory as a function of the squared lattice spacing and for four different quark masses. The continuum extrapolated values are indicated at $a = 0$. In dimensionless ratios, such as the one shown, the value of the gauge coupling in the effective theory is irrelevant and as $M \to \infty$ they approach the $n_\ell = 0$ value. The corresponding behaviour is shown in the right part. Where the data are, a linear behaviour in $1/M$ looks more plausible than the prediction of the EFT, which is $\sim 1/M^2$ for large M. After our pioneering work, Knechtli

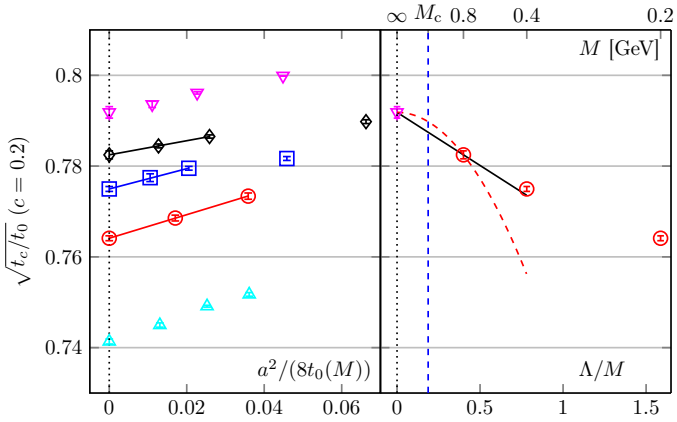


Figure 4. Left: The continuum extrapolation of the ratio $\sqrt{t_c/t_0}$ $(c = 0.2)$ at mass values (from top to bottom) $\Lambda/M = 0, 0.4, 0.78, 1.59, \infty$. Right: Its mass-dependence including a linear and quadratic fit in Λ/M between the largest mass and $N_{\mathrm{f}} = 0$ $(\Lambda/M = 0)$.

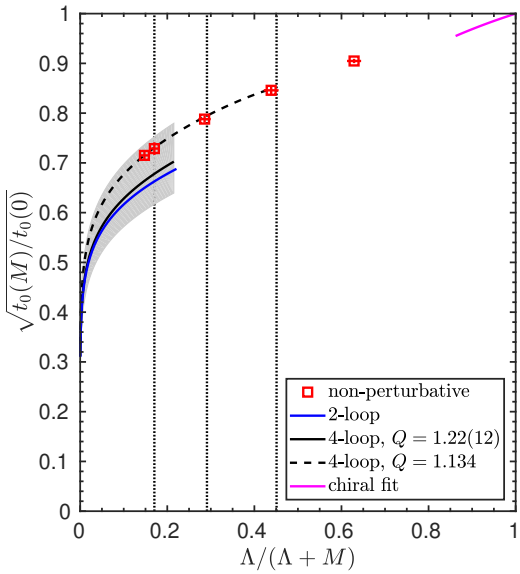


Figure 5. The mass-dependence of the ratio $\sqrt{t_0(M)/t_0(0)}$ in the theory with two mass-degenerate quarks. Monte Carlo data after continuum extrapolation are compared with the perturbative predictions for $1/(QP)$ at large M. The dashed line is the 4-loop curve adjusting the value of Q to go through the point at $M/\Lambda = 5.7781$. The vertical dotted lines mark the values of the quark mass $M_{\rm c}$, $M_{\rm c}/2$ and $M_{\rm c}/4$.

et al.[27] extended the computation to larger mass and found agreement with $1/M^2$. But the most relevant result is that the corrections due to finite mass are very small, few per-mille level, when one rescales to the decoupling of *a single* quark. This holds for the scale ratios which were investigated, which are a few.[21, 27] It is justified to conclude that for low energy physics, also in QCD as realised in nature, one may safely leave out the charm quark and work with the $2+1$ theory.

Also the M-dependence of dimensionfull scales themselves are predictable by the effective theory. Now the matching of the coupling of the fundamental, $n_{\rm f}$, theory and the n_ℓ theory is relevant. It turns out that for large M, the mass scaling function $\eta^{\rm M} = \frac{{\rm d}\log(\mathcal{S})}{{\rm d}\log(M)}$ is computable in perturbation theory and the perturbative series looks very well behaved. A comparison of the shape obtained in the fundamental theory to the one predicted by perturbative matching is given by the comparison of squares and dashed line in Fig. 5. The shape is very well reproduced by PT.

From this non-perturbative test of the quality of perturbative decoupling at the charm quark, we can deduce[22] two important things:

1. The effects of charm, bottom and top-quarks in the running coupling can indeed be added perturbatively as it has been done in Ref. 28 and a number of other works.

2. The heavy quark contribution to the coupling of "scalar dark matter" to hadrons is accurately given by perturbation theory and therefore known more accurately than previously thought.

References

1. R. Orus, *A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States*, Annals Phys. **349**, 117–158, 2014.
2. M. C. Bañuls, K. Cichy, J. I. Cirac, K. Jansen, and S. Kühn, *Tensor Networks and their use for Lattice Gauge Theories*, PoS LAT **2018**, 022, 2018.
3. M. C. Bañuls, K. Cichy, K. Jansen, and J. I. Cirac, *The mass spectrum of the Schwinger model with Matrix Product States*, JHEP **11**, 158, 2013.
4. M. C. Bañuls, K. Cichy, J. I. Cirac, K. Jansen, and H. Saito, *Thermal evolution of the Schwinger model with Matrix Product Operators*, Phys. Rev. D **92**, 034519, 2015.
5. M. C. Bañuls, K. Cichy, J. I. Cirac, K. Jansen, and S. Kühn, *Density Induced Phase Transitions in the Schwinger Model: A Study with Matrix Product States*, Phys. Rev. Lett. **118**, 071601, 2017.
6. M. C. Bañuls, K. Cichy, J. I. Cirac, K. Jansen, and S. Kühn, *Efficient basis formulation for 1+1 dimensional SU(2) lattice gauge theory: Spectral calculations with matrix product states*, Phys. Rev. X **7**, 041046, 2017.
7. L. Funcke, K. Jansen, and S. Kühn, *Topological vacuum structure of the Schwinger model with matrix product states*, 2019, arXiv:1908.00551 [hep-lat].
8. G. Parisi, R. Petronzio, and F. Rapuano, *A Measurement of the String Tension Near the Continuum Limit*, Phys. Lett. **128B**, 418–420, 1983.
9. M. Lüscher and P. Weisz, *Locality and exponential error reduction in numerical lattice gauge theory*, JHEP **09**, 010, 2001.
10. M. Cè, L. Giusti, and S. Schaefer, *A local factorization of the fermion determinant in lattice QCD*, Phys. Rev. D **95**, 034503, 2017.
11. M. Cè, L. Giusti, and S. Schaefer, *Domain decomposition, multi-level integration and exponential noise reduction in lattice QCD*, Phys. Rev. D **93**, 094507, 2016.
12. M. Lüscher, *A New approach to the problem of dynamical quarks in numerical simulations of lattice QCD*, Nucl. Phys. B **418**, 637–648, 1994.
13. W. Bernreuther and W. Wetzel, *Decoupling of heavy quarks in the minimal subtraction scheme*, Nucl. Phys. B **197**, 228–236, 1982, [Erratum: Nucl. Phys. B **513**, 758 (1998)].
14. K. G. Chetyrkin, J. H. Kühn, and C. Sturm, *QCD decoupling at four loops*, Nucl. Phys. B **744**, 121–135, 2006.
15. Y. Schröder and M. Steinhauser, *Four-loop decoupling relations for the strong coupling*, JHEP **01**, 051, 2006.
16. M. Czakon, *The four-loop QCD β-function and anomalous dimensions*, Nucl. Phys. B **710**, 485–498, 2005.
17. P. A. Baikov, K. G. Chetyrkin, and J. H. Kühn, *Five-loop running of the QCD coupling constant*, Phys. Rev. Lett. **118**, 082002, 2017.
18. F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *The five-loop beta function of Yang-Mills theory with fermions*, JHEP **02**, 090, 2017.

19. B. A. Kniehl, A. V. Kotikov, A. I. Onishchenko, and O. L. Veretin, *Strong-coupling constant with flavor thresholds at five loops in the anti-MS scheme*, Phys. Rev. Lett. **97**, 042001, 2006.
20. A. G. Grozin, M. Hoeschele, J. Hoff, and M. Steinhauser, *Simultaneous decoupling of bottom and charm quarks*, JHEP **09**, 066, 2011.
21. M. Bruno, J. Finkenrath, F. Knechtli, B. Leder, and R. Sommer, *Effects of heavy sea quarks at low energies*, Phys. Rev. Lett. **114**, 102001, 2015.
22. A. Athenodorou, J. Finkenrath, F. Knechtli, T. Korzec, B. Leder, M. K. Marinković, and R. Sommer, *How perturbative are heavy sea quarks?*, Nucl. Phys. B **943**, 114612, 2019.
23. M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, JHEP **1008**, 071, 2010.
24. R. Narayanan and H. Neuberger, *Infinite N phase transitions in continuum Wilson loop operators*, JHEP **0603**, 064, 2006.
25. S. Schaefer, R. Sommer, and F. Virotta, *Critical slowing down and error analysis in lattice QCD simulations*, Nucl. Phys. B **845**, 93–119, 2011.
26. M. Lüscher and S. Schaefer, *Lattice QCD with open boundary conditions and twisted-mass reweighting*, Comput. Phys. Commun. **184**, 519–528, 2013.
27. F. Knechtli, T. Korzec, B. Leder, and G. Moir, *Power corrections from decoupling of the charm quark*, Phys. Lett. B **774**, 649–655, 2017.
28. M. Bruno, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Schaefer, H. Simma, S. Sint, and R. Sommer, *QCD Coupling from a Nonperturbative Determination of the Three-Flavor Λ Parameter*, Phys. Rev. Lett. **119**, 102001, 2017.

Simulating Thioflavin T and Congo Red Binding to the Fibril Structure of Amyloid-β(1-42)

Benedikt Frieg[1] **and Holger Gohlke**[1,2]

[1] John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC) and Institute for Complex Systems - Structural Biochemistry (ICS-6), Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {b.frieg, h.gohlke}@fz-juelich.de

[2] Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
E-mail: gohlke@hhu.de

Binding modes for two amyloid-β(1-42) fibril tracers, namely Thioflavin T and Congo red, were identified using unbiased all-atom molecular dynamics simulations and binding free-energy computations. Both dyes bind to primarily hydrophobic grooves on the amyloid fibril surface, perpendicular to its β-strands. Binding affinities computed by the MM-GBSA method are in excellent agreement with experimental values and corroborate the proposed binding modes. The binding modes can guide the rational design of novel biomarkers for amyloid fibrils.

1 Introduction

Alzheimer's disease (AD) is a progressive, unremitting, neurodegenerative disorder, and the leading cause of dementia.[1–4] In 2018 it was estimated that more than 50 Mio. people in the world are living with dementia, with two-thirds associated with AD.[5, 6] There is a new dementia patient around the world every three seconds, and it is estimated that by 2030 more than 80 Mio. will be diagnosed with dementia.[6] The considerable number of dementia patients is directly associated with the healthcare system facing high costs, and it is estimated that the costs will rise to worldwide $2 trillion by 2030.[6] Thus, dementia and AD, in particular, are significant challenges for the modern healthcare system.

Since 1998, more than 100 drug candidates have been tested, but only four have been approved for therapeutic applications in AD.[5, 6] The authorised drugs, however, only help to manage some of the symptoms but do neither stop nor slow the progression.[5, 6] Thus, currently, there is no cure for AD, and considering the underlying pathologic causes of AD, there is not going to be a cure in near future.[5, 6]

The AD type of dementia has been related to an imbalance between the production and elimination of the protein fragment amyloid-beta (Aβ).[7] Accumulation of Aβ outside of the nerve cells (neurons) is associated with cell death in AD.[5, 6] The consequence is that neurons essential for cognitive function are damaged or destroyed and, because AD is a progressive disease, more and more neurons will be destroyed over time.[5, 6]

Currently, there is no test available that conclusively diagnoses AD. Diagnostic strategies focus on multiple tools and aspects, such as detecting key biomarkers for AD.[5, 6] For AD, the accumulation of Aβ has been recognised as a biomarker.[5, 6] As changes in the brain begin 20 years or more before any AD-related symptoms are expected to occur,[8] a conclusive test for AD-related biomarkers will be essential. This is all the more so as a

therapeutic intervention in the very early stages is currently considered necessary to slow or stop the progression of AD.[5, 6]

Under high concentration, Aβ peptides can aggregate to more senior oligomeric complexes or arrange to symmetric and periodic fibrils, which denote pathological hallmarks in AD.[4, 7, 9, 10] Thus, molecular probes that detect amyloid fibrils *in vitro* or *in vivo*, such as Thioflavin T (THT, Fig. 1A) and Congo red (CGR, Fig. 1B),[11, 12] are essential for an accurate and conclusive diagnosis of amyloid fibrils-related diseases. THT and CGR are potent fluorescent dyes that form fluorescent complexes with amyloid and amyloid-like fibrils.[13–15] For both probes, previous studies suggested binding modes,[11, 12, 16] but the exact nature of how both probes bind to the AD-related Aβ(1-42) fibril remained elusive. Such knowledge is, however, essential for the rational search for novel molecular probes.[11, 12]

Recently, the first high-resolution structure of the Aβ(1-42) fibril was obtained by cryo-electron microscopy and nuclear magnetic resonance spectroscopy,[17] opening up the possibility to predict the binding modes of THT and CGR at the atomistic level. Here, we used molecular dynamics (MD) simulations of free diffusion of THT and CGR in the presence of the Aβ(1-42) fibril to derive such binding mode models. The probes were not biased throughout the binding process by any prior knowledge of the binding epitopes. That way, the probes spontaneously recognise energetically preferred binding epitopes, yielding binding mode models that are in agreement with previous observables.

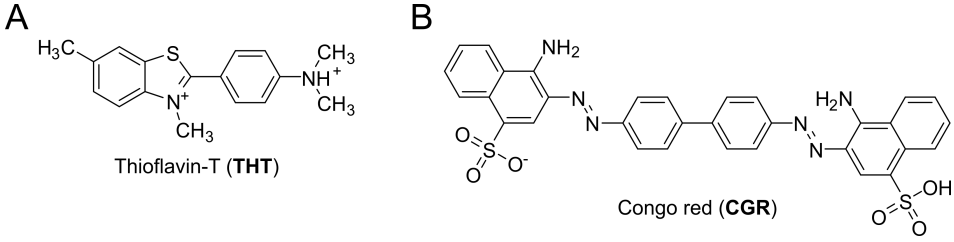


Figure 1. Structures of amyloid biomarkers. Structure of Thioflavin-T (THT, A) and Congo red (CGR, B) with protonation states according to pH 2.

2 Methods

System preparation and molecular dynamics simulations

To investigate THT and CGR binding to the Aβ(1-42) fibril (PDB ID 5OQV[17]), we performed unbiased MD simulations. The protonation states of all residues of the fibril at pH 2 were resolved by NMR spectroscopy and incorporated in our setup. That way, the protonation states in the simulations match those in *in vitro* experiments. THT and CGR were also prepared for pH 2 using Epik[18, 19] (Fig. 1). We randomly placed one dye molecule and the Aβ(1-42) fibril structure into an octahedral box using PACKMOL,[20] and neutralised and solvated the systems with chloride atoms and TIP3P[21] water, respectively. For both dyes, we prepared 45 different initial configurations, in which the distance between the dye molecule and the Aβ(1-42) fibril ranged from 13 Å to 62 Å. The systems were

subjected to production simulations of 1 μs length each, to study THT and CGR binding to the Aβ(1-42) fibril. The minimisation, thermalisation, and equilibration protocol is reported in Ref. 23, which was already applied previously to study ligand binding processes.[24, 25] All minimisation, equilibration, and production simulations were performed with the *pmemd.cuda* module[26, 27] of Amber16[22] on JUWELS.[28]

During visual inspection of the MD trajectories, we observed multiple binding and unbinding events of the dye molecules to and from the Aβ(1-42) fibril (the workflow is exemplarily shown for THT in Fig. 2). We determined all stably bound dye poses by calculating the Root Mean Square Deviation (RMSD) of the dyes after superimposing the Aβ(1-42) fibril structure. That way the RMSD becomes a measure for the spatial displacement of a dye molecule between two snapshots. THT poses with RMSD < 1.2 Å and CGR poses with RMSD < 1.5 Å, respectively, were defined as stably bound. The stably bound poses were subjected to hierarchical clustering, using the minimum distance between the clusters as cluster criterion. Starting from $\epsilon = 2.0$ Å, we gradually increased ϵ in 0.5 Å intervals until the population of the largest cluster remained unchanged ($\epsilon_{\mathrm{THT}} = 5.0$ Å and $\epsilon_{\mathrm{CGR}} = 4.5$ Å).

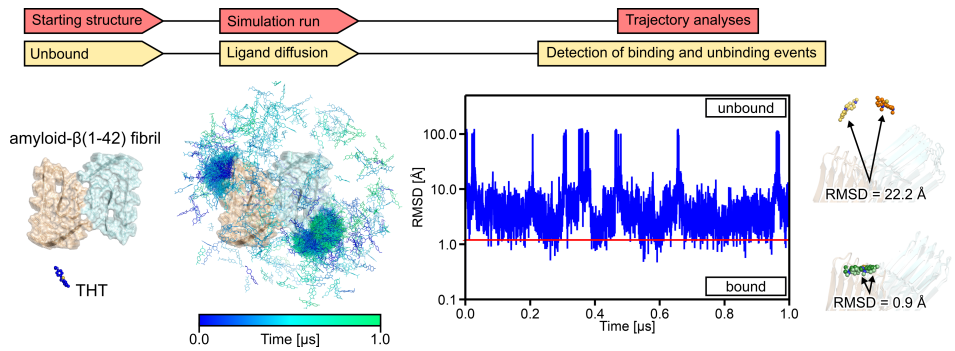


Figure 2. Schematic workflow to identify a binding mode from an MD simulation of ligand binding. From left to right: In the starting configuration, the Aβ(1-42) fibril structure (cartoon-surface representation) and Thioflavin-T (THT, sphere model) are spatially separated. During MD simulations of ligand binding, the Aβ(1-42) fibril and THT molecules diffuse freely and without guiding force (only the diffusion of the THT is shown for clarity purposes). The THT diffusion (shown as stick models) is coloured according to the simulation time (see the colour range). The resulting trajectory is analysed with respect to THT binding to or unbinding from the Aβ(1-42) fibril. In this study, the RMSD between two consecutive THT conformations, after superimposing the fibril, was used to characterise the (un-)binding events. The larger the RMSD, the larger is the spatial displacement between two consecutive THT conformations, as indicated on the right. In this example, THT is considered stably bound, if the RMSD < 1.2 Å (indicated by the red line).

Molecular mechanics generalised Boltzmann surface area (MM-GBSA) calculations

The most-populated clusters were further investigated by binding free energy calculations, performed with MMPBSA.py,[22, 29] yielding ΔG_{bind}.[30, 31] We obtained the standard free energy of binding ΔG^0_{bind} for a standard state of 1 M,[32] which is directly related to the dissociation constant K_{D} according to Eq. (1)

$$\Delta G^0_{bind} = RT\, ln(K_D) \tag{1}$$

where R is the universal gas constant ($R = 0.001987$ kcal K^{-1} mol^{-1}), and $T = 300$ K.

3 Results

Unbiased MD simulations of THT and CGR binding to the Aβ(1-42) fibril and subsequent binding free energy calculations were applied to determine the binding epitope and a binding mode in full atomic resolution for both probes.

During MD simulations of THT and CGR binding, we observed multiple binding and unbinding events of the dyes to and from the Aβ(1-42) fibril structure. The unbinding events are more likely in the case of THT, suggesting that THT binds weaker to the Aβ(1-42) fibril structure than CGR. To identify the binding epitopes of THT and CGR, we focused our analyses on the bound probe conformations. As to THT, stable bound structures were predominant around amino acids V18, F20, and E22. Due to the symmetric organisation of the Aβ(1-42) fibril, high THT concentrations are observed at both protofibrils (Fig. 3A). As to CGR, by contrast, the stably bound poses are distributed across the complete Aβ(1-42) fibril surface (Fig. 3B), and CGR conformations parallel and perpendicular to the fibril axes are observed. In contrast to THT, the area around V18, F20, and E22 is less populated by CGR molecules, suggesting that both probes most likely bind to two distinct epitopes.

To extract the predominant binding pose from all stable bound conformations of THT and CGR, all structures shown in Fig. 3A + B were subjected to hierarchical clustering. The ten most populated clusters were further subjected to binding free energy calculations, yielding a dissociation constant K_D^{comp} (Eq. (1)) for each cluster. Based on both the cluster populations and K_D^{comp}, we extracted a binding mode model for either probe, which revealed several interesting facts.

As to THT, the largest cluster contains 11% of all considered configurations and also shows the most favourable binding affinities. The THT conformations in this cluster bind to V18, F20, and E22, such that the protonated amino function of THT is stabilised by a hydrogen bond interaction with E22, while the aromatic moieties of THT interact with F20 and V18 (Fig. 3C). Interestingly, the side chains of F20 adapt a V-shaped orientation, such that the phenyl rings in F20 align almost parallel to the aromatic rings in THT, forming π-π-stacking interactions. THT binds across four layers of Aβ(1-42) peptides with its axis oriented almost perpendicular to the orientation of the stacked β-strands of the Aβ(1-42) fibril. The observation that multiple segments are essential for THT binding also provides an explanation why THT recognises fibrillary structures, but no single Aβ(1-42) peptides. For the reported binding mode, ΔG^0_{bind} = -9.06 kcal mol^{-1}, yielding K_D^{comp} = 251.21 nM, which is in agreement with experimentally derived binding affinities of THT to Aβ(1-40) fibrils ranging from 790 nM - 1740 nM.[33, 34] As to CGR, clustering resulted in many but rather weakly populated clusters, which is not surprising considering the broad distribution of stably bound CGR conformations. We thus primarily focused on the interpretation of the calculated binding affinities to derive a reasonable binding mode. Experimentally derived binding affinities of CGR to Aβ(1-40) range from 48 nM - 1500 nM[35, 36] and we identified a CGR conformation (Fig. 3D) with ΔG^0_{bind} = -9.78 kcal mol^{-1}, yielding K_D^{comp} = 74.77 nM. In this conformation, CGR binds to the groove between Y10 and V12. In this pose, the amino groups of CGR form hydrogen bond interactions with the backbone carbonyl oxygen of E11, while the charged and polar phosphate groups are exposed to the solvent (Fig. 3D). The CGR biaryl core forms hydrophobic interactions with Y10 and V12, such that the biaryl core and the Y10 side-chain adopt an edge-to-face configuration. CGR

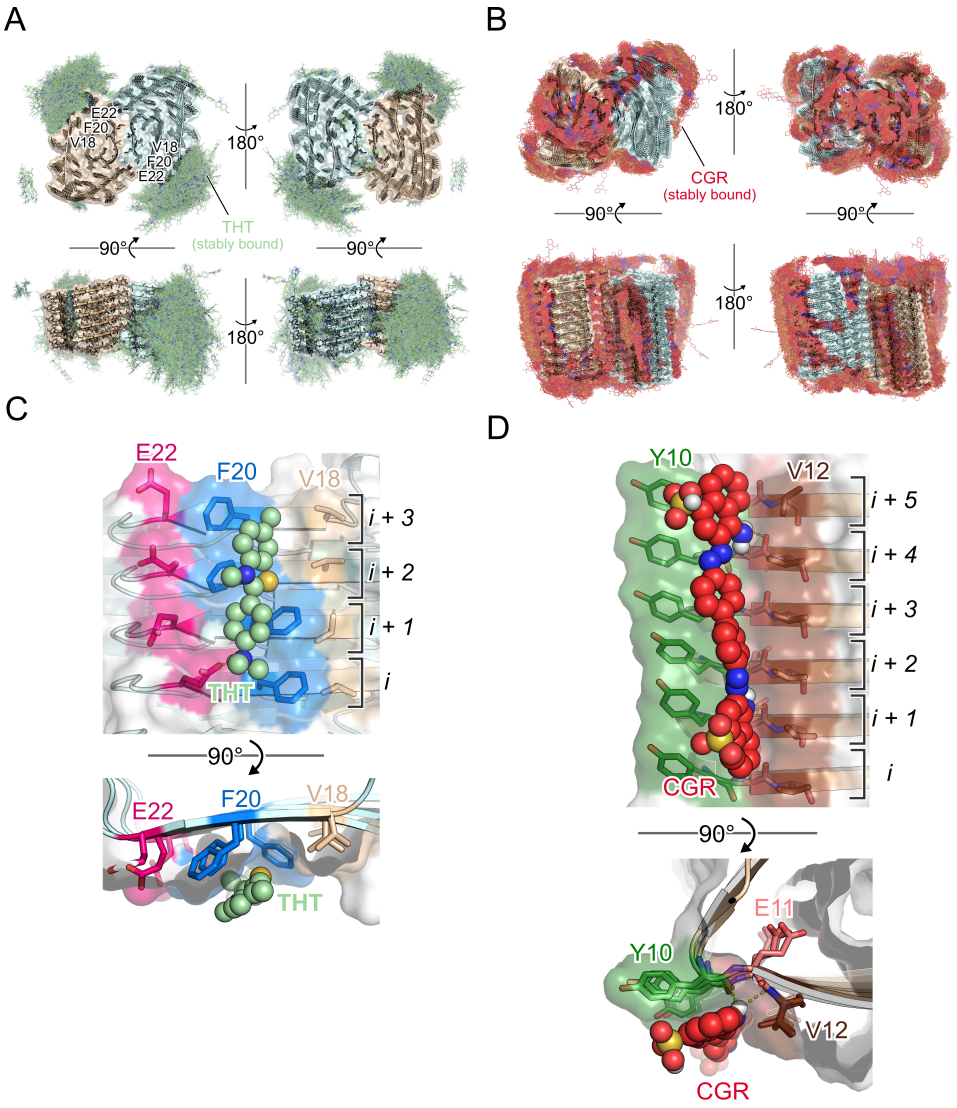


Figure 3. Binding modes of fluorescent dyes at the Aβ(1-42) fibril. A, B: Distribution of Thioflavin-T (THT, A) and Congo red (CGR, B) molecules bound to the Aβ(1-42) fibril. THT is shown as green stick-model and CGR as red stick-model, respectively. The Aβ(1-42) fibril is shown as cartoon-surface representation with each protofibril coloured differently. C, D: Dominant THT (C) and CGR (D) binding poses from side- and top-view shown as sphere models. Amino acids involved in THT or CGR binding are depicted as stick-model.

binds across six layers of Aβ(1-42) peptides with its axis oriented almost perpendicular to the orientation of the stacked β-strands of the Aβ fibril, providing an explanation of why CGR recognises fibrillary structures.

4 Discussion

Binding epitopes and atomistic binding mode models for the fluorescent dyes THT and CGR at the Aβ(1-42) fibril have been identified through unbiased MD simulations and binding free energy calculations. This procedure has been successfully applied to study ligand binding processes before,[24, 37–40] also in the field of amyloid research.[16, 41]

Our binding mode model suggests that THT likely binds on the surface of the Aβ(1-42) fibril to amino acids V18, F20, and E22, which is in agreement with previous observations on multiple accounts. For example, relative fast binding kinetics of THT binding to fibrils suggests that the dye can easily access its binding site,[42] which can be explained by THT binding to a solvent-exposed surface. Former studies on THT binding to fibrils also suggest a minimal binding site on the fibril surface that covers four consecutive β-strands,[43, 44] a feature we also find for our binding mode of THT (Fig. 3C). THT binds fast and specifically to Aβ(1-40) fibrils, but does not bind to monomers or other oligomeric states,[43] supporting the view that multiple β-strands may be involved in THT binding. THT is reported to bind with its long axis almost parallel to the long axis of the fibrils.[45] This orientation was also observed in an X-ray structure of THT bound to a PSAM ladder,[43] during MD simulations of THT binding to protofibrils forming two-layered β-sheets,[16] and in the present study (Fig. 3C).

As to CGR, our results suggest that it binds on the surface of the Aβ(1-42) fibril to amino acids Y10, E11, and V12, which is in excellent agreement with previous findings. Two structural features are essential for CGR binding to Aβ aggregates; first, two negative charges that are separated by a fixed distance of 19 Å,[11] and, second, a biaryl core framework.[11] Modifying the gap between the negative charges reduces the binding affinity,[11] while modifying the substituents of the biaryl core does not influence the binding affinity dramatically.[11, 35] The spacing of 19 Å corresponds to the spacing between five pleated and stacked β-strands,[12, 17] suggesting that CGR binds across multiple layers, thereby forming ionic interactions. The observation was corroborated by Schütz *et al.* who showed that CGR binds to the surface of stacked β-strands, thus creating ionic interactions between the sulphonic groups of CGR and lysine residues of the HET fibril.[46] As the CGR molecule considered in our simulations, however, does not carry two negatively charged groups (Fig. 1B), it is also not surprising that we did not observe the ionic interaction model. Nevertheless, similar to the results from Schütz *et al.*,[46] we find that CGR binds across six layers of Aβ(1-42) peptides (Fig. 3D). In our case, however, the interactions between CGR and the Aβ(1-42) fibril are mainly hydrophobic, which may be explained by a suggested second binding site for CGR-type ligands.[11] For the second binding site, prior studies underline the importance of hydrophobic interactions for CGR binding,[12, 16, 47] as some experimental observables cannot be explained by the ionic model.[11] Interestingly, for CGR derivatives where no ionic state is expected at physiological pH, the binding affinity is even increased relative to CGR,[11] supporting the view that hydrophobic interactions are crucial for CGR binding. Finally, most studies suggest that CGR binds with its long axis perpendicular to the direction of the β-strands,[12, 48] as it was observed in the present study (Fig. 3D).

We validated the binding mode models by comparing computationally derived binding affinities with experimentally derived binding affinities. In general, the binding affinities are in good agreement, although, to the best of our knowledge, there is currently no study available that determined binding affinities for the probes in the presence of the recently

resolved Aβ(1-42) fibril structure at pH 2.[17] This may be of relevance in the case of THT, as the affinity decreases at acidic pH conditions, suggesting a weak pH dependency for THT binding as observed for THT binding to insulin and HET fibrils.[12, 49, 50] As to CGR, there is no evidence suggesting a pH dependency, although most of the affinity studies are done at neutral pH as CGR tends to be insoluble at acidic pH.[12] In our preparation process of CGR, the pK_a values of the sulphonic acids were predicted[18, 19] with pK_a = 2.13 and pK_a = 1.53, such that CGR tends to be in a deprotonated and charged state at pH 2 (Fig. 1B).

In conclusion, we suggest binding sites and binding modes for the fluorescent dyes THT and CGR to the Aβ(1-42) fibril. The binding sites were identified by unbiased MD simulations and subsequently corroborated by calculations of binding affinities. The binding sites and modes agree with previous experimental observations. The binding mode models may provide a starting point for the systematic search and design of novel and improved molecules that bind to Aβ(1-42) fibrils, which is essential for conclusively diagnosing amyloid fibrils-related diseases.

Acknowledgements

The authors gratefully acknowledge fruitful discussion with Lothar Gremer, Henrike Heise, Gunnar Schröder, and Dieter Willbold and the computing time granted by the John von Neumann Institute for Computing (NIC) provided on the supercomputer JUWELS[28] at Jülich Supercomputing Centre (JSC) (project ID: HKF7).

References

1. J. Wang *et al.*, *A systemic view of Alzheimer disease – insights from amyloid-beta metabolism beyond the brain*, Nat. Rev. Neurol. **13(10)**, 612–623, 2017.
2. P. Scheltens *et al.*, *Alzheimer's disease*, Lancet **388(10043)**, 505–517, 2016.
3. C. L. Masters *et al.*, *Alzheimer's disease*, Nat. Rev. Dis. Primers **1**, 1–18, 2015.
4. H. V. Vinters, *Emerging concepts in Alzheimer's disease*, Annu. Rev. Pathol.: Mech. Dis. **10**, 291–319, 2015.
5. Alzheimer's Association, *2019 Alzheimer's disease facts and figures*, 2019.
6. Alzheimer's Disease International, *World Alzheimer Report 2018 – The state of the art of dementia research: New frontiers*, 2018.
7. D. J. Selkoe and J. Hardy, *The amyloid hypothesis of Alzheimer's disease at 25 years*, EMBO Mol. Med. **8(6)**, 595–608, 2016.
8. B. A. Gordon *et al.*, *Spatial patterns of neuroimaging biomarker change in individuals from families with autosomal dominant Alzheimer's disease: a longitudinal study*, Lancet Neurol. **17(3)**, 241–250, 2018.
9. M. Goedert *et al.*, *The propagation of prion-like protein inclusions in neurodegenerative diseases*, Trends. Neurosci. **33(7)**, 317–325, 2010.
10. M. Jucker and L. C. Walker, *Self-propagation of pathogenic protein aggregates in neurodegenerative diseases*, Nature **501(7465)**, 45–51, 2013.
11. L. S. Cai *et al.*, *Radioligand development for PET imaging of beta-amyloid (Abeta) – current status*, Curr. Med. Chem. **14(1)**, 19–52, 2007.

12. M. Groenning, *Binding mode of Thioflavin-T and other molecular probes in the context of amyloid fibrils – current status*, J. Chem. Biol. **3(1)**, 1–18, 2010.
13. A. I. Sulatskaya *et al.*, *Interaction of Thioflavin-T with amyloid fibrils: stoichiometry and affinity of dye binding, absorption spectra of bound dye*, J. Phys. Chem. B **115(39)**, 11519–11524, 2011.
14. C. Xue *et al.*, *Thioflavin T as an amyloid dye: fibril quantification, optimal concentration and effect on aggregation*, Roy. Soc. Open Sci. **4(1)**, 160696, 2017.
15. E. I. Yakupova *et al.*, *Congo red and amyloids: history and relationship*, Biosci. Rep. **39**, 1–22, 2019.
16. C. Wu *et al.*, *Dual binding modes of Congo red to amyloid protofibril surface observed in molecular dynamics simulations*, J. Am. Chem. Soc. **129(5)**, 1225–1232, 2007.
17. C. Wu *et al.*, *Fibril structure of amyloid-beta(1-42) by cryo-electron microscopy*, Science **358(6359)**, 116–119, 2017.
18. Schroedinger, LLC, *Schrödinger Release 2017-3: Schrödinger Suite 2017-3.*, 2017.
19. J. C. Shelley *et al.*, *Epik: a software program for pK a prediction and protonation state generation for drug-like molecules*, J. Comput. Aided Mol. Des. **21(12)**, 681–691, 2007.
20. J. C. Shelley *et al.*, *PACKMOL: a package for building initial configurations for molecular dynamics simulations*, J. of Comp. Chem. **30(13)**, 2157–2164, 2009.
21. W. L. Jorgensen *et al.*, *Comparison of simple potential functions for simulating liquid water*, J. Chem. Phys. **79(2)**, 926–935, 1983.
22. D. A. Case *et al.*, *AMBER 16*, University of California, San Francisco, 2016.
23. B. Frieg *et al.*, *Molecular mechanisms of glutamine synthetase mutations that lead to clinically relevant pathologies*, PLoS Comput. Biol. **12(2)**, e1004693, 2016.
24. S. Bhatia *et al.*, *Targeting Hsp90 dimerization via the C-terminus is effective in imatinib resistant CML and lacks heat shock response induction*, Blood **132(3)**, 307–320, 2018.
25. B. Frieg *et al.*, *Towards restoring catalytic activity of glutamine synthetase with a clinically relevant mutation*, in Proceedings of the NIC Symposium 2016, K. Binder *et al.* (Editors), Jülich, 97–104, 2016.
26. T. Darden *et al.*, *Particle Mesh Ewald: an N-log (N) method for Ewald sums in large systems*, J. Chem. Phys. **98(12)**, 10089–10092, 1993.
27. R. Salomon-Ferrer *et al.*, *Routine microsecond molecular dynamics simulations with Amber on GPUs. 2. Explicit solvent particle mesh Ewald*, J. Chem. Theory. Comput. **9(9)**, 3878–3888, 2013.
28. D. Krause, *JUWELS: Modular Tier-0/1 supercomputer at the Jülich Supercomputing Centre*, JLSRF **5**, A135, 2019.
29. B. R. Miller *et al.*, *MMPBSA.py: an efficient program for end-state free energy calculations*, J. Chem. Theory. Comput. **8(9)**, 3314–3321, 2012.
30. S. Genheden and U. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities*, Expert Opin. Drug. Dis. **10(5)**, 449–461, 2015.
31. N. Homeyer and H. Gohlke, *Free energy calculations by the molecular mechanics Poisson-Boltzmann surface area method*, Mol. Inf. **31(2)**, 114–122, 2012.
32. H. Gohlke *et al.*, *Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RaIGDS complexes*, J. Mol. Biol. **330(4)**, 891–913, 2003.

33. C. Rodriguez-Rodriguez *et al.*, *Thioflavin-based molecular probes for application in Alzheimer's disease: from in silico to in vitro models*, Metallomics **7(1)**, 83–92, 2015.
34. R. L. Yona *et al.*, *Thioflavin derivatives as markers for amyloid-beta fibrils: Insights into structural features important for high-affinity binding*, Chemmedchem **3(1)**, 63–66, 2008.
35. W. Zhen *et al.*, *Synthesis and amyloid binding properties of rhenium complexes: Preliminary progress toward a reagent for SPECT imaging of Alzheimer's disease brain*, J. Med. Chem. **42(15)**, 2805–2815, 1999.
36. Y. Wang *et al.*, *Synthesis and 11C-labelling of (E, E)-1-(3', 4'-dihydroxystyryl)-4-(3'-methoxy-4'-hydroxystyryl) benzene for PET imaging of amyloid deposits*, J. Labelled Compd Radiopharm. **45(8)**, 647–664, 2002.
37. G. Martinez-Rosell *et al.*, *Molecular-simulation-driven fragment screening for the discovery of new CXCL12 inhibitors*, J. Chem. Inf. Model. **58(3)**, 683–691, 2018.
38. N. Plattner and F. Noé, *Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models*, Nat. Commun. **6(7653)**, 1–10, 2017.
39. Y. B. Shan *et al.*, *How does a drug molecule find its target binding site?*, Nature Chem. **136(8)**, 3320–3320, 2014.
40. H. Gohlke *et al.*, *Binding region of alanopine dehydrogenase predicted by unbiased molecular dynamics simulations of ligand diffusion*, J. Chem. Inf. Model. **53(10)**, 2493–2498, 2013.
41. C. König *et al.*, *Binding sites for luminescent amyloid biomarkers from non-biased molecular dynamics simulations*, Chem. Commun. **136(8)**, 3320–3320, 2018.
42. H. LeVine, *Stopped-flow kinetics reveal multiple phases of thioflavin T binding to Alzheimer beta(1-40) amyloid fibrils*, Arch. Biochem. Biophys. **342(2)**, 306–316, 1997.
43. M. Biancalana *et al.*, *Molecular mechanism of Thioflavin-T binding to the surface of beta-rich peptide self-assemblies*, J. Mol. Biol. **385(4)**, 1052–1063, 2009.
44. C. Wu *et al.*, *Binding modes of Thioflavin-T to the single-layer beta-sheet of the peptide self-assembly mimics*, J. Mol. Biol. **394(4)**, 627–633, 2009.
45. M. R. H. Krebs *et al.*, *The binding of thioflavin-T to amyloid fibrils: localisation and implications*, J. Struct. Biol. **149(1)**, 30–37, 2005.
46. A. K. Schütz *et al.*, *The amyloid-congo red interface at atomic resolution*, Angew. Chem. Int. Edit. **50(26)**, 5956–5960, 2011.
47. J. H. Cooper, *Selective amyloid staining as a function of amyloid composition and structure*, Lab. Invest. **31**, 232–238, 1974.
48. L. W. Jin *et al.*, *Imaging linear birefringence and dichroism in cerebral amyloid pathologies*, Proc. Natl. Acad. Sci. USA **100(26)**, 15294–15298, 2003.
49. M. Groenning *et al.*, *Binding mode of Thioflavin T in insulin amyloid fibrils*, J. Struct. Biol. **159(3)**, 483–497, 2007.
50. R. Sabaté *et al.*, *On the binding of Thioflavin-T to HET-s amyloid fibrils assembled at pH 2*, J. Struct. Biol. **162(3)**, 387–396, 2008.

Massively Parallel Large-Scale Multi-Model Simulation of Tumour Development Including Treatments

Marco Berghoff[1], **Jakob Rosenbauer**[2], **and Alexander Schug**[2]

[1] Karlsruhe Institute of Technology, Karlsruhe, Germany
E-mail: marco.berghoff@kit.edu

[2] John von Neumann Institute for Computing and Jülich Supercomputing Centre, Institute for Advanced Simulation, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: al.schug@fz-juelich.de

The temporal and spatial resolution in the microscopy of tissues has increased significantly within the last years, yielding new insights into the dynamics of tissue development and the role of single cells within it. Still, the theoretical description of the connection of single cell processes to macroscopic tissue reorganisations is lacking. Especially in tumour development, single cells play a crucial role in the advance of tumour properties. We developed a simulation framework that can model tissue development up to the centimetre scale with micrometre resolution of single cells. Through parallelisation, it enables the efficient use of high-performance computing systems, therefore enabling detailed simulations on 10.000s of cores. Our generalised tumour model respects adhesion driven cell migration, cell-to-cell signalling, and mutation-driven tumour heterogeneity. We scan the response of the tumour development depending on division inhibiting substances such as cytostatic agents. Furthermore, we are investigating the interaction with radiotherapy to find a suitable therapy plan. Currently, the emergence of ever-more-powerful experimental techniques such as light sheet microscopy already offers unprecedented subcellular insight into tissue dynamics. Combined with powerful machine learning techniques, such large data sets (TB's +) can be effectively evaluated promising realistic parameters for our simulations for topics ranging from cancer development to embryogenesis or morphogenesis with considerable impact both for basic science and applied biomedical fields.

1 Model

Especially in tumour development, the behaviour of single cells plays a significant role in the tumour characteristics and progression. This microscale can be essential for the emergence of macroscopic tumour growth and the emergine tumour properties. Simultaneously, tumours that are clinically relevant consist of millions of cells thereby setting the upper border for the simulated spatial resolution.

Over the last decade the theoretical and computational description of cell and tissue dynamics and processes has developed to be a valuable tool to complement experiments.[1–4] Models are applied in embryogenesis and tissue patterning[5–7] as well as tumour development[8–11] and treatment.[12, 13] At the same time the experimental accessibility of entire tissues and its properties has greatly increased through modern microscopic techniques such as light-sheet microscopy[14] and confocal microscopy[15] as well as CRISPR techniques.[16] The combination of contemporary experimental techniques with the recent advances in machine learning enables the analysis of large datasets in various spacial regimes[17–19] and translating the properties into model parameters.[20, 21]

However, the computational models that are published so far either concentrate on the detailed descriptions on a small number of cells or a coarse grained description of cells for

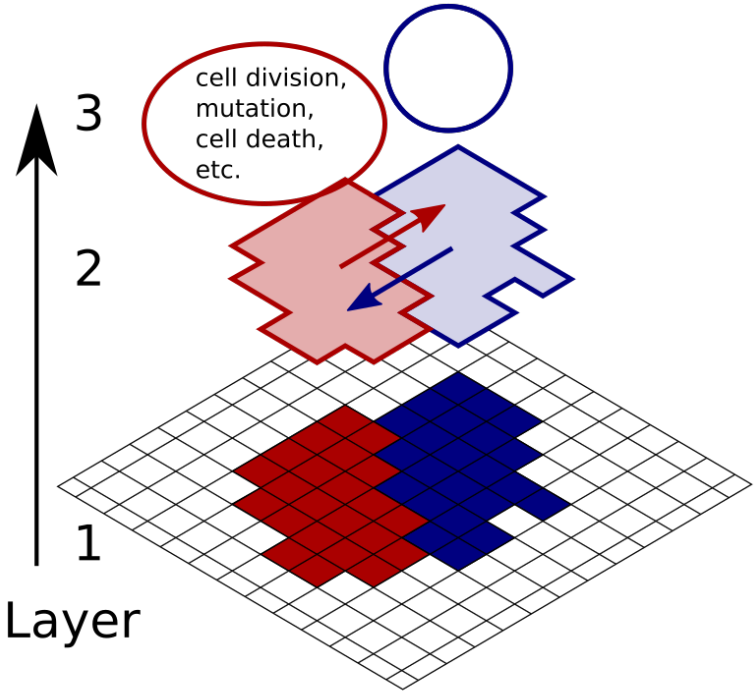


Figure 1. A two-dimensional schematic diagram shows the three different model levels. The lowest layer, the microscale, contains a grid based cellular Potts model to model tissue dynamics. On the highest layer, the macroscale, an agent-based model is acting. In between, on the mesoscale, the transport of signals is processed.

larger tissues. To be able to simulate both, the individual cellular geometry as well as the macroscopic tumour a high degree of parallelism is needed, which we introduce with our model. We use the massively parallel NAStJA framework[22, 23] to implement a multi-model solver with a cell-geometric resolution for the simulation of tissue growth in general and cancer in particular.

The model consists of three layers to accommodate for this large spatial range. On the microscale layer a cellular Potts model simulates adhesion driven cell movements, cell dynamics and cell-to-cell interactions. On the most coarse layer, an agent-based model is used, where each cell corresponds to one agent. On this macroscale layer, the signals are processed, and cell death and cell division, including mutations, are simulated. Another model layer lies between these two scales. It represents the mesoscale and ensures signal diffusion through the surfaces of the cells. Fig. 1 shows an overview of the different model levels. Following the three model layers are described in detail. Each model layer acts on its specific scale.

Microscale The model is based on a regular rectangular grid. Each voxel has an integer number that represents the cell ID. All voxels with the same cell ID are assigned to a biological cell. The cellular Potts model is a Monte Carlo method:[24] The system evolves

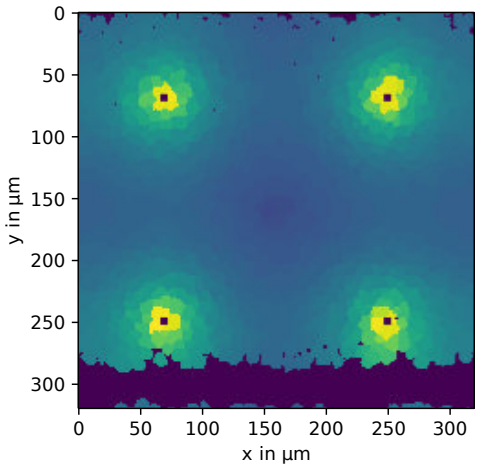


Figure 2. Diffusion of the nutrient signal, the four black squares are blood vessels, that supply the nutrient. Yellow cells have high nutrient concentrations and blue ones low.

under random changes of nearest-neighbour voxels at cell surfaces. These changes are accepted according to a Metropolis acceptance criterion. A Hamiltonian defines the energy of the system. It is the sum of several energy functions:

$$H_{\text{CPM}} = \sum_{i\in\Omega}\sum_{j\in N(i)} J_{\tau(\varsigma_i),\tau(\varsigma_j')}(1-\delta(\varsigma_i,\varsigma_j')) + \lambda_{\text{v}}\sum_{i\in\Omega}\left(v(\varsigma_i)-V(\tau(\varsigma_i))\right)^2 + \lambda_{\text{s}}\sum_{i\in\Omega}\left(s(\varsigma_i)-S(\tau(\varsigma_i))\right)^2 \tag{1}$$

Here, the first sum corresponds to the cell adhesion, the second to the volume, and the last corresponds to the surface of individual cells. While the system evolves, each cell has a target volume V and a target surface S. Possible Monte-Carlo steps are nearest neighbour interactions of adjacent grid points, leading to small changes on the cells surface geometries. Therefore the Monte-Carlo propagation of the system corresponds to a time step, that can be mapped to a physical time. This is modelled by the energy minimisation of the whole non-steady-state system. The implementation of the Hamiltonian is versatile so that the single terms of the energy can be replaced by other terms, *e. g.* different kinds of volume estimations. Besides, the energy sum can be extended by several other terms, *e. g.* adaption to a local or global potential (introduced by the Glazier-Graner-Hogeweg model[25]).

Mesoscale On this scale, the diffusion of signals is simulated. Signals can be among others oxygen, nutrients or therapeutic drugs. Based on the common surface with neighbouring cells, the signals diffuse from one cell to the other, see Fig. 2.

Macroscale At the most coarse level, each cell is assigned to an agent. The individual agents are then responsible for signal processing, *e. g.* usage of nutrient, producing of signals. These agents also control the cell cycle of the individual cell. Events such as cell death and cell division can be induced depending on internal and external parameters of the cell, such as the signal concentration, cell age and therapeutic agents. A set of cell types is defined, each cell type has a specific set of parameters, determining the behaviour of the individual cell belonging to that cell type. Parameters of a single cell type include micro, meso, and macroscale parameters such as the target volume and surface V, S, as well as division rate p_{div} and diffusion constants D, *etc.* After cell division, cells may mutate, depending on external parameters as well as predefined rates. A mutation alters the cell properties, which is implemented by a change of the cell type with defined transition rates between each cell type. The agents take the domain distribution into account and can thus perform cell division correctly, even if the geometric resolution of the cell is distributed over several processes.

2 Scaling

The code is implemented within the NAStJA framework so that it can benefit directly from the excellent scalability. The domain is divided into subdomains and distributed to the MPI ranks. Each MPI rank then holds a block containing the field on the microscale with the cell IDs and additional cell data, which are held for the higher scales, per cell. After each calculation sweep, a halo exchange to the first 26 neighbour blocks is used for the field of cell IDs. Initially, the additional data of the cells must be exchanged globally, since it is not known where the cells are located. That would result in a poorly-scaling collective communication. Domain knowledge about the size of the cells gives the possibility to chose the size of the blocks at least three-fold larger compared to the cells. With this, it can be ensured that a cell is never distributed over more than eight blocks, *i. e.* two per dimension, at the same time, compare Fig. 3. With this knowledge, it can be ensured that

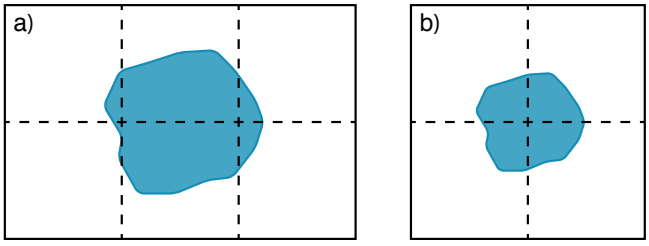


Figure 3. A cell (blue) is distributed to several blocks. (a) shows an inadequate distribution and (b) shows an allowed distribution.

the bookkeeping of the additional data can be handled by local communications to the next 26 neighbours. The efficiency of this exchange is excellent, as shown in Fig. 4(a). Fig. 4(b) right shows an efficiency of $> 80\,\%$ for blocks with an edge length of 100 and even an efficiency of $> 90\,\%$ for an edge length of 200. This result was measured on JURECA with up to 256 nodes corresponding to 6144 processes.

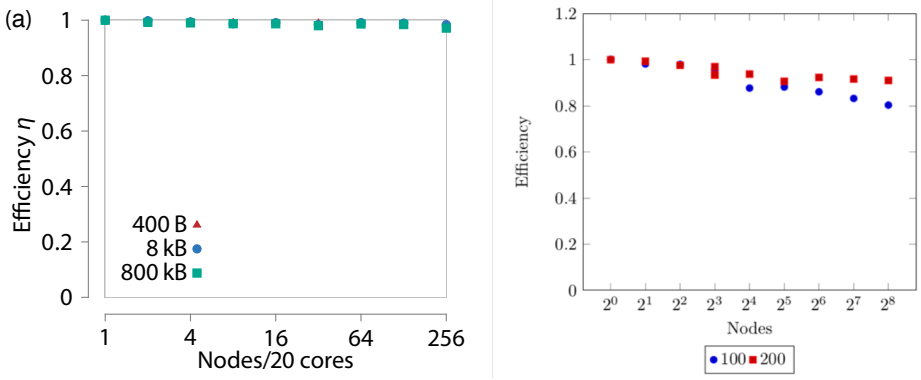


Figure 4. (a) The efficiency of the data exchange on ForHLR II and (b) the efficiency for the whole simulation on JURECA, using up to 256 nodes with 24 cores each. It is shown for a subdomain distribution with a block edge length of 100 and 200.

3 Application

To validate our model implementation we model adhesion driven cell sorting, as it was described in two dimensions in the initial paper of Graner and Glazier.[24] We use a three-dimensional spherical arrangement of randomly mixed cells. Two cell types with differential adhesion are introduced and randomly assigned to the cells. The cell-to-cell adhesion of cells belonging to the same cell type is larger than the adhesion between different cell types. This leads to a demixing of the cell types and formation of clusters as can be seen in Fig. 5. Reproducing the expected behaviour and showing accordance of our three-dimensional implementation with the initial two-dimensional model.

We apply the model on the growth of heterogeneous tumours. A domain is filled with non-dividing tissue cells, that make up the surrounding in which the tumour develops. Each cell has a volume of about 1000 voxels. A network of blood vessels spanning the entire domain is introduced and serves as the source of nutrients and medication. A small nucleus of cancer cells is placed in the centre of the simulated tissue. The cancerous cells proliferate and develop a tumour through cell division. Cells divide when abundant nutrition is present, leading to an increased growth adjacent to blood vessels and hypoxic regions where is a sparse vascularisation. In hypoxic regions cell-division is down-regulated while cell death is up-regulated leading to a negative growth.

Tumour cells can mutate after cell division into one of the 30 predefined cell types. Each cell type provides different properties and therefore enhances or hinders the propagation of the tumour depending on the localisation within the tissue. A heterogeneous tumour forms with its composition of sub-populations of different cell types depending on external conditions. Tumour heterogeneity is a pivotal property of tumours since it enables the specialisation and adaptation of the tumour and leads to complications in tumour treatment. In order to recreate those properties observed *in vivo*, two different treatment types are incorporated into our model:

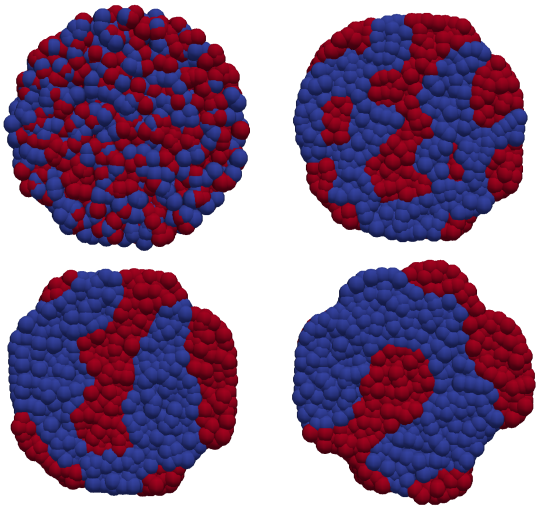


Figure 5. Adhesion driven cell sorting. Time evolution of a two dimensional slice of the three dimensional simulation. The colours represent different cell types. Times are initial condition and after $20 \cdot 10^6$, $45 \cdot 10^6$ and $110 \cdot 10^6$ Monte Carlo sweeps.

Chemotherapy is realised by the introduction of a cytostatic drug that down-regulates cell-division. The drug is distributed via the blood vessels. Since only cell-division is affected by the drug, the tumour shrinks due to cell-death still being present.

Radiation therapy reduces the division rate globally and immediately initiates the cell-death of a fraction of the tumour cells.

With these mechanisms established, treatments can then be optimised for different tumours. This is done by scanning different treatment protocols such as single pulse, multiple pulses and combinations of chemo- and radiotherapy. The effect on the total tumour volume as well as the tumour composition is analysed and can give valuable information about the generation of heterogeneity development of a tumour. This is illustrated in Fig. 6, where chemo- or radiotherapy individually do not lead to success, but a combination does. Multiple pulses of chemo- or radiotherapy lead to the development of a nearly homogeneous tumour, with one dominating cell type.

4 Conclusion

We were able to show that with a highly parallel framework and a multiscale model, the development of tumours can be simulated. We are able to simulate heterogeneous tumour growth and development in a domain measuring $1000 \times 1000 \times 1000$ voxels corresponding to a tissue of $1\,\mathrm{mm}^3$. The cell-geometric resolution was chosen to reproduce the tumour characteristics, which depend significantly on the adhesion between the individual cells. Besides, we have shown that different treatments can be performed. Thus, many different

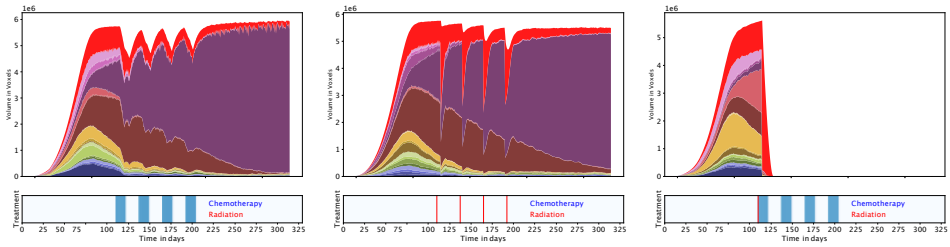


Figure 6. Change of volume of the tumour under different treatment plans of chemotherapy and radiation, single or in combination. The colours represent different cell types.

treatment plans can be simulated before the treatment on the patient, and then the best one can be selected for the patient.

Due to the flexibility of the framework, the model can be easily extended, and further effects can be considered, in future versions.

Acknowledgements

The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA[26] at Jülich Supercomputing Centre (JSC). This work was partially supported by the Helmholtz Analytics Framework of the Helmholtz Association under project number ZT-I-0003.

References

1. L. M. Loew and J. C. Schaff, *The Virtual Cell: a software environment for computational cell biology*, TRENDS in Biotechnology **19**, 401–406, 2001.
2. M. H. Swat, G. L. Thomas, J. M. Belmonte, A. Shirinifard, D. Hmeljak, and J. A. Glazier, *Multi-scale modeling of tissues using compucell3d*, in Methods in cell biology **110**, A. R. Asthagiri and A. P. Arkin (Editors), Elsevier, 325–366, 2012.
3. A. Ghaffarizadeh, R. Heiland, S. H. Friedman, S. M. Mumenthaler, and P. Macklin, *PhysiCell: An open source physics-based cell simulator for 3-D multicellular systems*, PLoS computational biology **14**, e1005991, 2018.
4. P. M. Altrock, L. L. Liu, and F. Michor, *The mathematics of cancer: integrating quantitative models*, Nature Reviews Cancer **15**, 730, 2015.
5. E. Stanganello, A. I. H. Hagemann, B. Mattes, C. Sinner, D. Meyen, S. Weber, A. Schug, E. Raz, and S. Scholpp, *Filopodia-based Wnt transport during vertebrate tissue patterning*, Nature communications **6**, 5846, 2015.
6. B. Mattes, Y. Dang, G. Greicius, L. T. Kaufmann, B. Prunsche, J. Rosenbauer, J. Stegmaier, R. Mikut, S. Özbek, G. U. Nienhaus *et al.*, *Wnt/PCP controls spreading of Wnt/β-catenin signals by cytonemes in vertebrates*, Elife **7**, e36953, 2018.
7. T. Hirashima, E. G. Rens, and R. M. H. Merks, *Cellular Potts modeling of complex multicellular behaviors in tissue morphogenesis*, Development, growth & differentiation **59**, 329–339, 2017.

8. Z. Wang, J. D. Butner, R. Kerketta, V. Cristini, and T. S. Deisboeck, *Simulating cancer growth with multiscale agent-based modeling*, in Seminars in cancer biology **30**, Elsevier, 70–78, 2015.
9. András Szabó and R. M. H. Merks, *Cellular potts modeling of tumor growth, tumor invasion, and tumor evolution*, Frontiers in oncology **3**, 87, 2013.
10. A. Shirinifard, J. S. Gens, B. L. Zaitlen, N. J. Popławski, M. Swat, and J. A. Glazier, *3D multi-cell simulation of tumor growth and angiogenesis*, PloS one **4**, e7190, 2009.
11. B. M. Rubenstein and L. J. Kaufman, *The role of extracellular matrix in glioma invasion: a cellular Potts model approach*, Biophysical journal **95**, 5661–5680, 2008.
12. B. Waclaw, I. Bozic, M. E. Pittman, R. H. Hruban, B. Vogelstein, and M. A. Nowak, *A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity*, Nature **525**, 261, 2015.
13. A. Karolak, D. A. Markov, L. J. McCawley, and K. A. Rejniak, *Towards personalized computational oncology: from spatial models of tumour spheroids, to organoids, to tissues*, Journal of The Royal Society Interface **15**, 20170703, 2018.
14. J. M. Girkin and M. Torres Carvalho, *The light-sheet microscopy revolution*, Journal of Optics **20**, 053002, 2018.
15. J. Jonkman and C. M. Brown, *Any way you slice it – a comparison of confocal microscopy techniques*, Journal of biomolecular techniques: JBT **26**, 54, 2015.
16. M. Adli, *The CRISPR tool kit for genome editing and beyond*, Nature communications **9**, 1911, 2018.
17. C. Ounkomol, S. Seshamani, M. M. Maleckar, F. Collman, and G. R. Johnson, *Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy*, Nature methods **15**, 917, 2018.
18. T.-L. Liu, S. Upadhyayula, D. E. Milkie, V. Singh, K. Wang, I. A. Swinburne, K. R. Mosaliganti, Z. M. Collins, T. W. Hiscock, J. Shea *et al.*, *Observing the cell in its native state: Imaging subcellular dynamics in multicellular organisms*, Science **360**, eaaq1392, 2018.
19. V. Voleti, K. B. Patel, W. Li, C. Perez Campos, S. Bharadwaj, H. Yu, C. Ford, M. J. Casper, R. W. Yan, W. Liang *et al.*, *Real-time volumetric microscopy of in vivo dynamics and large-scale samples with SCAPE 2.0*, Nature methods **16**, 1054–1062, 2019.
20. J. Snoek, O. Rippel, K. Swersky, R. Kiros, N. Satish, N. Sundaram, M. Patwary, Mr. Prabhat, and R. Adams, *Scalable bayesian optimization using deep neural networks*, in International conference on machine learning, 2171–2180, 2015.
21. M. K. Driscoll, E. S. Welf, A. R. Jamieson, K. M. Dean, T. Isogai, R. Fiolka, and G. Danuser, *Robust and automated detection of subcellular morphological motifs in 3D microscopy images*, Nature methods **16**, 1037–1044, 2019.
22. M. Berghoff, I. Kondov, and J. Hötzer, *Massively parallel Stencil Code Solver with Autonomous Adaptive Block Distribution*, IEEE Transactions on Parallel and Distributed Systems **29**, 2282–2296, 2018.
23. M. Berghoff and I. Kondov, *Scalable Global Network Based on Local Non-Collective Communications*, SC Workshop: Latest Advances in Scalable Algorithms for Large-Scale Systems (ScalA'18), 2018.
24. F. Graner and J. A. Glazier, *Simulation of biological cell sorting using a two-dimensional extended Potts model*, Physical Review Letters **69**, 2013, 1992.

25. A. Balter, R. M. H. Merks, N. J. Popławski, M. Swat, and J. A. Glazier, *The glazier-graner-hogeweg model: extensions, future directions, and opportunities for further study*, in Single-Cell-Based Models in Biology and Medicine, A. Anderson and K. Rejniak (Editors), Springer, 151–167, 2007.
26. D. Krause and P. Thörnig, *JURECA: General-purpose supercomputer at Jülich Supercomputing Centre*, Journal of large-scale research facilities **2**, A62, 2016.

The NIC Excellence Projects

Quantum Chemical Modelling of Water Splitting: From Photoelectrochemistry to Superlubricity

Leonhard Mayrhofer[1], Gianpietro Moras[1], Takuya Kuwahara[1], Alexander Held[1,2], and Michael Moseler[1,2,3]

[1] Fraunhofer IWM, MicroTribology Center μTC, Wöhlerstr. 11, 79108 Freiburg, Germany
E-mail: {leonhard.mayrhofer, gianpietro.moras, takuya.kuwahara, michael.moseler}@iwm.fraunhofer.de

[2] Freiburg Materials Research Center FMF, University of Freiburg, Stefan-Meier-Straße 21, 79104 Freiburg, Germany

[3] Physics Department, University of Freiburg, Hermann-Herder-Str. 3, 79104 Freiburg, Germany

Within the project the electronic structure and chemical reactions at interfaces are investigated by using the quantum chemical methods density functional theory (DFT) and density-functional tight binding (DFTB). The research covers a broad spectrum of topics ranging from sustainable energy materials for solar hydrogen production to the wear and friction in tribological contacts including mechanically induced chemical reactions of lubricant and additive molecules at surfaces. A common feature of the different studies are surface reactions of molecules induced by different driving forces such as excited charge carriers or external mechanical load. Here, selected examples of our studies are presented where the splitting of water molecules at surfaces and interfaces plays a key role. We start with the *ab initio* DFT investigation of unconventional U_3O_8//Fe_2O_3 heterostructures which are used to split water molecules with the purpose of solar hydrogen production and conclude with DFTB simulations of water lubricated carbon coatings where water splitting can lead to extremely low friction, also known as superlubricity.

1 Introduction

Materials play a crucial role for the transition of an economy based on fossil fuels to an economy based on renewable energies. This is also true for a more effective usage of resources by reducing energy consumption and wear of mechanical machinery. Atomistic simulations can help to better understand and predict the microscopic processes determining the functional properties and have thus become a main pillar in materials science. Here, we demonstrate with the aid of two selected examples conducted within this project the abilities of DFT and DFTB simulations to describe systems with very different compositions and functionalities. We first give a short report on our research activities in the field of semiconductor heterostructures for solar water splitting applications and then present a recent study where again water molecule degradation at interfaces is the main concern. However, instead of excited charge carriers, in the second example mechanical load drives the splitting of water or lubricant molecules in tribological contacts of tetrahedral amorphous carbon (ta-C). This induces surface modifications which lead to extremely low friction.

2 Unconventional U_3O_8//Fe_2O_3 Metal Oxide Heterostructures for Water Splitting Applications

Solar energy conversion is a safe, eco-friendly and world's most abundant renewable source of energy. Direct synthesis of chemical fuels like hydrogen by photoconversion processes can overcome the strongly unequal temporal and geographic availability of conventional solar technologies such as photovoltaics. Hydrogen is considered to have a large potential to replace fossil fuels as main primary energy source, which will have a major impact on the global energy systems and the environment.[1] There are various ways of producing solar hydrogen such as, thermo-chemical water splitting, use of photobiological systems, and photoelectrochemical (PEC) water splitting. All pathways can contribute to renewable hydrogen production for future green economies. However, hydrogen generation by PEC reactions is regarded as the "Holy Grail" of the hydrogen economy. PEC water splitting by semiconductor materials is based on a simple principle (Fig. 1): upon photo excitation of an electron-hole pair, the excited electron in the conduction band drives the reduction of water to H_2, whereas the hole in the valence band oxidises water to O_2, such that the net effect is the dissociation or splitting of water into gasses of its constituent elements. The minimum theoretical electrical potential required to oxidise water under standard conditions is 1.23 V (*vs.* a standard hydrogen electrode). However, owing to several factors such as cell resistance, polarisation losses, *etc.* in general an over-potential is needed. As a consequence, the actual voltage required to achieve water electrolysis is typically in the range of 1.8 - 2.0 V. Hence, semiconductor systems with sufficiently large band gaps are required for photoelectrochemical (PEC) water splitting. On the other hand, band gaps far beyond 2 eV only exploit a small part of the solar spectrum for photoelectrolysis and therefore result in low solar energy to hydrogen efficiencies.

In addition to band gap requirements, in total a unique combination of properties has to be satisfied by the employed materials to be well suited for large-scale PEC hydrogen production. The most important are: i) high absorption of solar energy, ii) proper alignment of conduction and valence band edges w.r.t. water redox potentials, iii) long life-time without material degradation and finally iv) earth abundance and low production costs.

So far, no single semiconductor system is known which satisfies all of the mentioned preconditions at the same time. Metal oxides such as Fe_2O_3 or TiO_2 are abundant and stable under PEC conditions but have limitations regarding charge transport properties (Fe_2O_3) or optical absorption in the visible range (TiO_2). Material modifications such as doping or hydrogen treatment processes are one way to enhance the PEC performance of metal oxides and several such approaches were studied theoretically in our group.[3–6] Another approach is the combination of different metal oxides in heterostructure devices, where synergetic effects can drastically increase the PEC performance. Within this project the mechanisms leading to an enhanced water splitting performance of two different metal oxide heterostructures, $BiVO_4$//TiO_2 and U_3O_8//Fe_2O_3 were investigated by DFT calculations in collaborative experimental/theoretical studies.[7, 8] The latter study is presented in the following.

The use of uranium oxides for renewable energy conversion might look peculiar on a first glance. However, there is a huge amount of depleted and hence low-radiation uranium available as waste from the nuclear economy.[9] Due to the problems with the long-term storage of the uranium waste in the form of highly toxic UF_6, it is highly desirable to find

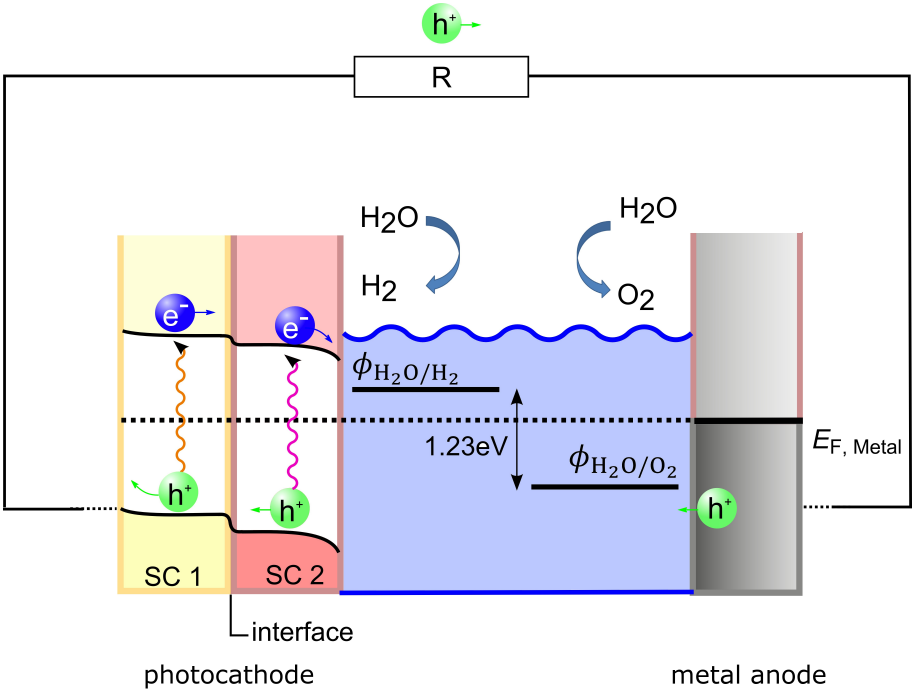


Figure 1. Water splitting is based on the separation of electrons and holes after the photo excitation of electron/hole pairs in semiconductors immersed in an aqueous electrolyte. After photoexcitation and electron-hole separation water molecules can be dissociated if the redox potentials of the water splitting reactions are properly aligned with the semiconductor band edges. The use of semiconductor heterostructures based on two semiconductors with the proper band alignment leads to enhanced electron/hole separation and helps to increase the solar energy to hydrogen conversion efficiencies.

applications of the uranium in a closed environment such that the deconversion of UF_6 to uranium oxide becomes economically viable. Hence, the application of uranium oxides as semiconductors or catalysts has been an active research field over the last twenty years.[10–13] Together with our experimental partners from the research group of Prof. Sanjay Mathur at the University of Cologne, the properties of U_3O_8//Fe_2O_3 heterostructures as photoanodes for solar water splitting were investigated.[8] Here, the Fe_2O_3 side of the heterojunction was exposed towards the aqueous electrolyte and supposed to catalyse the water oxidation. Experimentally, a strong enhancement of the PEC performance of U_3O_8//Fe_2O_3 based photoanode was found compared to the conventional PEC material Fe_2O_3. Pure U_3O_8 even showed negligible photocurrents. Transient absorption spectroscopy (TAS) measurements revealed largely improved charge carrier separation properties of the U_3O_8//Fe_2O_3 heterojunction. Using the DFT+U approach,[2] extensive simulations of U_3O_8//Fe_2O_3 interfaces were conducted. The calculated interfacial electronic structure revealed a so called type II alignment between the U_3O_8 and Fe_2O_3 band edges which favours the fast separation of electrons and holes, see Fig. 2. The holes favour transport towards the Fe_2O_3 surface to promote the oxidation of water. The electrons are transferred to the counter electrode

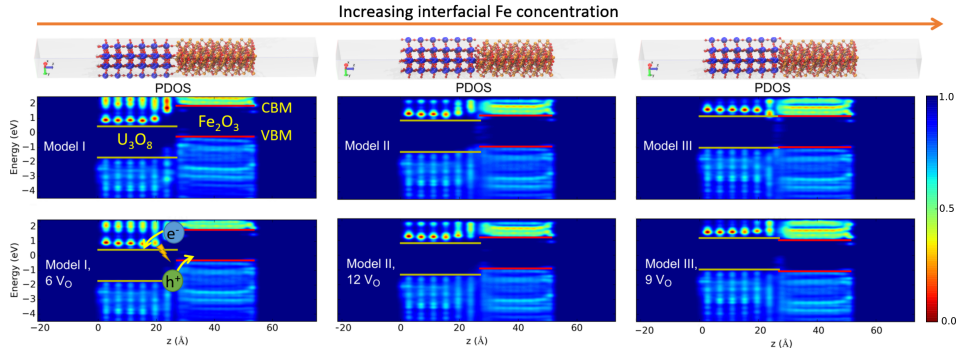


Figure 2. Interfacial electronic structure of different U_3O_8//Fe_2O_3 interface models with varying interfacial Fe and O ion concentration. The projected density of states (PDOS) over the interface shows that the conduction band minimum (CBM) and the valence band maximum (VBM) in Fe_2O_3 are higher than in U_3O_8 except at high interfacial Fe ion concentration and low oxygen concentration (Model III, 9 V_O, where V_O indicates the number of oxygen vacancies in comparison to the original structure). This so called type II alignment favours efficient charge carrier separation of photoexcited electron/hole pairs and leads to enhanced water splitting performance. The atomistic structures of the various interface models are shown on the top. U ions are indicated by blue, Fe ions by brownish and O ions by red spheres. Hydrogen atoms (white spheres) are used to saturate the free U_3O_8 surface. Each interface model contained approximately 700 atoms.

where the actual hydrogen evolution reaction takes place. Interestingly, our simulations revealed that the relative band alignment between Fe_2O_3 and U_3O_8 can be controlled by the concentration of Fe ions at the interface. This behaviour can be traced back to the multivalency of U ions within U_3O_8. At low enough interfacial concentrations of iron, all Fe ions are in the 3+ state, whereas U_3O_8 intrinsically contains U^{5+} and U^{6+} ions and the 5+ valence state readily switches to the 6+ state due to oxidation by interfacial oxygen species. Hence an interfacial dipole is building up due to charge transfer from O to U ions that tends to lift the Fe_2O_3 electronic states with respect to U_3O_8. The reverse effect only appears at rather high Fe ion concentrations when also Fe^{2+} ions are present which can be oxidised. Our DFT calculations predicted that high interfacial Fe concentrations are favoured under more reducing conditions and the interesting question arises whether the band alignment can thus be controlled by changing the synthesis conditions. The dependence of the band alignment on the Fe ion concentration could also be quantitatively captured by a simple classical electrostatic model taking into account the valency of the interfacial Fe and U ions and allowed a simple interpretation of the DFT results.

3 Tribochemical Water Splitting for Superlow Friction

The minimisation of frictional losses in mechanical components is of paramount importance to a more efficient use of energy.[14] Very significant reductions of fuel consumption and CO_2 emissions could be achieved by reducing friction in passenger cars to ultralow (friction coefficient μ in the $0.1 - 0.01$ range) or even to superlow levels ($\mu < 0.01$).[15] Thanks to the increased scientific effort devoted to this subject in recent years, such extremely low friction levels can be nowadays achieved in tribological systems that are technologically relevant. For instance, in order to optimise the tribological performance of

components that operate under severe environmental or loading conditions, there is an ever increasing use of hard surface coatings made of carbon, both with crystalline (nanocrystalline diamond, NCD) and amorphous (tetrahedral amorphous carbon, ta-C) structure. Not only do these coatings offer very high resistance to wear and corrosion but they can also yield superlow friction coefficients when used in combination with water[16] and other OH-containing molecules, such as hydrogen peroxide or glycerol.[17]

Understanding the atomic-scale mechanisms leading to superlow friction is crucial to be able to control the friction behaviour of the system (*e. g.* by introducing the optimal quantity of OH-containing molecules in the system) and to understand whether superlow friction conditions can be transferred to other materials interfaces. However, tribological interfaces are hardly accessible by spectroscopy and microscopy techniques and these have to be used *ex situ*, on surfaces after they have experienced the tribological load. This is why atomistic simulations that can directly "look into" tribologically loaded interfaces are an ideal tool to complement experimental characterisation. Following this concerted approach, and combining experimental and atomistic simulations results,[16–19] first steps forward were taken towards an explanation of superlubricity of water-lubricated carbon coatings. In vacuum, covalent bonds form between unpassivated diamond or ta-C surfaces. This results in plastic deformation of the interface region through a sp^3-to-sp^2 rehybridisation[20–22] and in very high shear stresses, needed to make one surface slide against the other. In presence of water at the sliding interface, tribochemical dissociative chemisortion reactions lead to H/OH-passivation of the carbon surfaces. This prevents the formation of covalent bonds across the sliding interface and results in superlow friction.

In this project, we combined efficient density-functional tight-binding (DFTB) molecular dynamics (MD) and density functional theory (DFT) electronic structure simulations to gain further insights into the mechanisms leading to superlow friction of water-lubricated diamond films.[23, 24] In particular, our goal was to extend the time scale spanned by previous DFT simulations[19] and to systematically investigate how the quantity of water molecules at the sliding interface can influence such mechanisms. First, we considered the unreconstructed (111) diamond surface, which is the most stable fracture surface and the least prone to amorphisation. We performed sliding MD simulation under a normal load of 5 GPa and a sliding speed of 100 ms^{-1} for water contents that vary from dry sliding conditions to a quantity of water molecules that is larger than the quantity needed to chemically terminate the two surfaces with H atoms and OH groups.

The results of the simulations are summarised in Fig. 3. We observed five different friction regimes that depend on the water quantity. Three of these friction regimes (I, IV and V) are compatible with the conclusions of previous experimental and simulation studies. For dry sliding or very low amounts of water, we observe cold welding and amorphisation of the sliding region with resulting very high friction (friction regime I). If the quantity of water molecules at the sliding interface is sufficient to passivate the two diamond surfaces by dissociative chemisorption, friction can reach ultralow levels, in particular in the case of full passivation of the surface dangling bonds (friction regime IV). The presence of further water leads to the formation of a thin water film between the passivated surfaces with an associated friction coefficient that is higher than in the case of full passivation without water film (friction regime V).

Interestingly, our simulation unveils two friction regimes that were not mentioned in the previous literature (II and III). Despite appearing at the same, low number of water

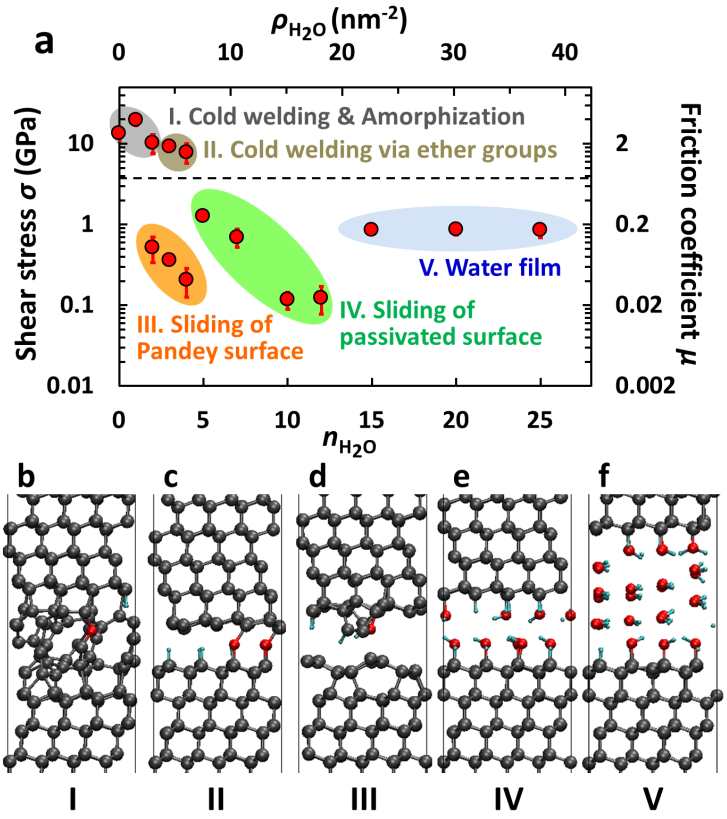


Figure 3. (a) Shear stress σ and friction coefficient μ as a function of the number of water molecules n_{H_2O} and of the water surface density ρ_{H_2O} for two initially unreconstructed diamond (111) surfaces. Five friction regimes (I-V) are observed. Panels (b) to (f) show representative snapshots for friction regimes I-V. Gray, red and cyan spheres represent carbon, oxygen and hydrogen atoms, respectively. Reused with permission from Ref. 23.

molecules (well below the quantity needed to partially passivate the surfaces), the two friction regimes are characterised by completely different friction coefficients. Friction regime II is obtained when, after tribochemical splitting of the water molecules, ether groups form across the sliding interface which remains crystalline. This is a cold-welding regime with very high friction coefficient. Conversely, friction regime III can potentially lead to superlow friction. After splitting of the water molecules, mechanochemical reactions involving oxygen atoms cause local rearrangements of the surface structure and trigger the reconstruction of one of the diamond surfaces. The surface reconstruction we observe is the so called Pandey reconstruction, which is more stable than the unreconstructed (111) surface when only a few surface dangling bonds are terminated. The aromatic character of the Pandey reconstruction makes it chemically very stable, thus preventing the formation of covalent bonds across the sliding interface.

To understand whether these friction regimes can be found on other crystallographic

orientations of the diamond surface, we adopted the same approach to study the (110) and (100) diamond surfaces.[24] We found that the cold-welding regimes I and II, the H/OH-passivation regime IV and the water film regime V are general to all of these low index surfaces. Furthermore, new friction regimes were found on these two crystallographic orientations. Due to the energetic stability of ether and keto groups on the (110) and (100) surfaces, ether- and keto-passivation of these surfaces is possible and can lead to friction coefficients that are very similar to those obtained in regime IV. Moreover, cross-linking ether groups are much less probable at these two interfaces and only give rise to mild cold-welding situations. Finally, superlow friction by aromatic passivation (regime III) was not observed because the (110) and (100) surfaces do not show aromatic reconstructions. However, since an amorphous layer forms on most diamond surfaces after deposition[18] or during tribological load,[20] we performed further sliding simulations on ta-C sliding interfaces. The simulations showed that passivating aromatic structures can form on ta-C surfaces in sliding contact with glycerol,[25] thus proving the general importance of friction regime III for hard-carbon coatings whenever OH-containing molecules at the sliding interface cannot provide full chemical termination of the surface dangling bonds.

4 Concluding Remarks

We hope that we could convince the reader that water splitting at surfaces and interfaces is a fascinating research topic with a broad range of applications and possibly a large impact on renewable energy conversion and in reducing energy consumption. Quantum chemical methods proved as powerful tools to investigate the mechanisms which drive the water splitting reactions under diverse conditions and for a plethora of different material systems.

Acknowledgements

We gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputers JURECA and JUWELS at Jülich Supercomputing Centre (JSC). We feel honoured that our project was selected as John von Neumann Excellence Project 2017. T. K. acknowledges financial support by JSPS Overseas Research Fellowships. A. H. and M. M. are grateful for funding by the Deutsche Forschungsgemeinschaft within Grant No. Mo 879/17. A. H., L. M., and M. M. kindly acknowledge financial support from the Insol Project (BMBF, Grant 01DQ14011) and the SOLAROGENIX Project (EC-FP7-Grant Agreement No. 310333).

References

1. T. Bak, J. Nowotny, M. Rekas, and C. C. Sorrell, *Photo-electrochemical hydrogen generation from water using solar energy. Materials-related aspects*, Int. J. of Hydrogen Energy **27**, 991–1022, 2002.
2. S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, *Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study*, Phys. Rev. B **57**, 1505–1509, 1998.

3. M. Mehta, N. Kodan, S. Kumar, A. Kaushal, L. Mayrhofer, M. Walter, M. Moseler, A. Dey, S. Krishnamurthy, S. Basu, and A P. Singh, *Hydrogen treated anatase TiO_2: a new experimental approach and further insights from theory*, J. Mater. Chem. A **4**, 2670–2681, 2016.
4. A. Mettenbörger, Y. Gönüllü, T.Fischer, T.Heisig, A.Sasinska, C.Maccato, G.Carraro, C.Sada, D.Barreca, L.Mayrhofer, M.Moseler, A.Held, and S.Mathur, *Interfacial insight in multi-junction metal oxide photoanodes for water-splitting applications*, J. Mater. Chem. A **4**, 2670–2681, 2016.
5. D. Primc, G. Zeng, R. Leute, M. Walter, L. Mayrhofer, and M. Niederberger, *Chemical Substitution – Alignment of the Surface Potentials for Efficient Charge Transport in Nanocrystalline TiO_2 Photocatalysts*, Chem. Mater. **28**, 4223–4230, 2016.
6. X. Wang, L. Mayrhofer, M. Höfer, S. Estrade, L. LopezConesa, H. Zhou, Y. Lin, F. Peiró, Z. Fan, H. Shen, L. Schäfer, M. Moseler, G. Bräuer, and A. Waag, *Facile and Efficient Atomic Hydrogenation Enabled Black TiO2 with Enhanced PhotoElectrochemical Activity via a Favorably Low Energy Barrier Pathway*, Adv. Energy Mater. **9**, 1900725, 2019.
7. A P. Singh, N. Kodan, B. R. Mehta, A. Held, L. Mayrhofer, and M. Moseler, *Band Edge Engineering in $BiVO_4/TiO_2$ Heterostructure: Enhanced Photoelectrochemical Performance through Improved Charge Transfer*, ACS Catal. **6**, 5311–5318, 2016.
8. J. Leduc, Y. Gönüllü, T.-P. Ruoko, Th. Fischer, L. Mayrhofer, N. Tkachenko, C. L. Dong, A. Held, M. Moseler, and S. Mathur, *Electronically Coupled Uranium and Iron Oxide Heterojunctions as Efficient Water Oxidation Catalysts*, Adv. Funct. Mater. **29**, forthcoming, 2019.
9. A. K. Burrell, T. M. McCleskey, P. Shukla, H. Wang, T. Durakiewicz, D. P. Moore, C. G. Olson, J. J. Joyce, and Q. Jia, *Controlling Oxidation States in Uranium Oxides through Epitaxial Stabilization*, Adv. Mater. **19**, 3559–3563, 2007.
10. A. R. Fox, S. C. Bart, K. Meyer, and C. C. Cummins, *Towards uranium catalysts*, Nature **455**, 341–349, 2008.
11. G. J. Hutchings, C. S. Heneghan, I. D. Hudson, and S. H. Taylor, *Uranium-oxide-based catalysts for the destruction of volatile chloro-organic compounds*, Nature **384**, 341–343, 1996.
12. T. T. Meek and B. von Roedern, *Semiconductor devices fabricated from actinide oxides*, Vacuum **83**, 226–228, 2008.
13. Z. Sofer, O. Jankovsky, P. Simek, K. Klimova, A. Mackova, and M. Pumera *Uranium- and Thorium-Doped Graphene for Efficient Oxygen and Hydrogen Peroxide Reduction*, ACS Nano **8**, 7106–7114, 2014.
14. S. W. Zhang, *Green tribology: fundamentals and future development*, Friction **1**, 186–194, 2013.
15. K. Holmberg, P. Andersson, and A. Erdemir, *Global energy consumption due to friction in passenger cars*, Tribol. Int. **47**, 221–234, 2012.
16. A. R. Konicek, D. S. Grierson, A. V. Sumant, T. A. Friedmann, J. P. Sullivan, P. U. P. A. Gilbert, W. G. Sawyer, and R. W. Carpick, *Influence of surface passivation on the friction and wear behavior of ultrananocrystalline diamond and tetrahedral amorphous carbon thin films*, Phys. Rev. B **85**, 155448, 2012
17. J.-M. Martin, M.-I. De Barros Bouchet, C. Matta, Q. Zhang, W. A. Goddard III, S. Okuda, and T. Sagawa, *Gas-phase lubrication of ta-C by glycerol and hydrogen*

peroxide. Experimental and computer modeling, J. Phys. Chem. C **114**, 5003–5011, 2010.

18. M.-I. De Barros Bouchet, G. Zilibotti, C. Matta, M. C. Righi, L. Vandenbulcke, B. Vacher, and J.-M. Martin, *Gas-phase lubrication of ta-C by glycerol and hydrogen peroxide. Experimental and computer modelingFriction of diamond in the presence of water vapor and hydrogen. Coupling gas-phase lubrication and first-principles studies*, J. Phys. Chem. C **116**, 6966–6972, 2012.
19. G. Zilibotti, S. Corni, and M. C. Righi, *Load-induced confinement activates diamond lubrication by water*, Phys. Rev. Lett. **111**, 146101, 2013.
20. L. Pastewka, S. Moser, P. Gumbsch, and M. Moseler, *Anisotropic mechanical amorphization drives wear in diamond*, Nat. Mater. **10**, 34–38, 2011.
21. T. Kunze, M. Posselt, S. Gemming, G. Seifert, A. R. Konicek, R. W. Carpick, L. Pastewka, and M. Moseler, *Wear, plasticity and rehybridization in tetrahedral amorphous carbon*, Tribol. Lett. **53**, 119–126, 2014.
22. G. Moras, A. Klemenz, T. Reichenbach, A. Gola, H. Uetsuka, M. Moseler, and L. Pastewka, *Shear melting of silicon and diamond and the disappearance of the polyamorphic transition under shear*, Phys. Rev. Mater. **2**, 083601, 2018.
23. T. Kuwahara, G. Moras, and M. Moseler, *Friction regimes of water-lubricated diamond (111): role of interfacial ether gorups and tribo-induced aromatic surface reconstructions*, Phys. Rev. Lett. **119**, 096101, 2017.
24. T. Kuwahara, G. Moras, and M. Moseler, *Role of oxygen functional groups in the friction of water-lubricated low-index diamond surfaces*, Phys. Rev. Mater. **2**, 073606, 2018.
25. T. Kuwahara, P. A. Romero, S. Makowski, V. Weihnacht, G. Moras, and M. Moseler, *Mechano-chemical decomposition of organic friction modifiers with multiple reactive centres induces superlubricity og ta-C*, Nat. Commun. **10**, 151, 2019.

Machine Learning Transport Properties in Quantum Many-Fermion Simulations

Carsten Bauer and Simon Trebst

Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany
E-mail: {bauer, trebst}@thp.uni-koeln.de

In computational condensed matter physics, the influx of algorithms from machine learning and their combination with traditional numerical many-body approaches is one of the most enticing recent developments. At this confluence novel techniques have been developed that allow to characterise many-body wave functions and discriminate quantum phase of matter by adapting concepts from computer science and statistics, which have proved tremendously practical in completely different contexts. However, in order to actually turn into a productive and widely accepted tool for obtaining a deeper understanding of microscopic physics these novel approaches must allow for meaningful, comprehensible inference and go beyond the applicability of their traditional counterparts. In this contribution, we report on significant progress made in this direction by discussing a novel algorithmic scheme using machine learning techniques to numerically infer the transport properties of quantum many-fermion systems. This approach is based on a quantum loop topography (QLT), and capable of distinguishing conventional metallic and superconducting transport in quantum Monte Carlo simulations by learning current-current correlations from equal-time Green's functions. We showcase this approach by studying the emergence of s- and d-wave superconducting fluctuations in the negative-U Hubbard model and a spin-fermion model for a metallic quantum critical point. The presented results, combined with the numerical efficiency of the QLT approach, point a way to identify hitherto elusive transport phenomena such as non-Fermi liquids using machine learning algorithms.

1 Introduction

State-of-the art machine learning techniques have not only become ubiquitous in our daily life (sorting emails, suggesting movies to watch, or identifying users by the touch of a button or the scan of a face), they also promise to become a powerful tool in quantum statistical mechanics. Their core functions – dimensional reduction and feature extraction – are a perfect match to the goal of identifying essential characteristics of a quantum many-body system, which are often hidden in the exponential complexity of its many-body wavefunction or the abundance of potentially revealing correlation functions. The basic idea of interpreting collective states of matter, such as superfluids, superconductors, or insulating quantum liquids, as a source of complex data with unknown intrinsic structure immediately opens up all of condensed matter physics as a playground for established machine learning algorithms implementing supervised, unsupervised, or reinforcement learning schemes.

Groundbreaking initial steps in this direction[1] have demonstrated that convolutional neural networks can indeed be trained to learn sufficiently many features from the correlation functions of classical and quantum many-body systems[2] such that distinct phases of matter can be discriminated and the parametric location of the phase transition between them identified. In parallel, it has been demonstrated that machine learning of wave functions is possible,[3,4] which can lead to a variational representation of quantum states based on artificial neural networks like restricted Boltzmann machines that, for some cases, out-

performs entanglement-based variational representations.[3] In a parallel development, efforts to integrate machine learning perspectives into established numerical schemes have led to many methodological improvements such as new quantum Monte Carlo flavours with dramatically reduced autocorrelation times.[5, 6]

This contribution reports on recent progress made in the use of machine learning to identify transport characteristics of itinerant electron systems. We will showcase that an innovative technique dubbed "quantum loop topography" (QLT), initially introduced by Eun-Ah Kim's group at Cornell to detect topological order in integer and fractional Chern insulators,[7] is capable of reliably distinguishing conventional metallic and superconducting transport by machine learning the essential features of longitudinal imaginary time current-current correlations. It should be noted that electronic transport properties are notoriously difficult to calculate – in the context of Monte Carlo approaches this typically asks for an analytic continuation, which is numerically ill posed, yielding no controlled framework to probe transport properties. As such, the QLT approach offers an intriguing alternative that is also vastly more efficient; the scaling of the computational cost differs by multiple orders of the considered system size. When compared to other machine learning protocols for quantum state recognition, the QLT stands out as a preprocessing step that, by using a correlation loop topography as a filter, selects and organises the input data with the physical response characteristic of the target phase in mind. It thereby distinguishes itself from, *e. g.* the application of convolutional neural networks (CNNs) whose motivation has been primarily rooted in image recognition techniques. Crucially, while making equally good transport predictions as CNNs, QLT requires a considerably smaller fraction of the input data, which is given by equal-time Green's functions produced in quantum Monte Carlo simulations. It has thereby become the method-of-choice for future applications in quantum statistical physics.

Although QLT is generally applicable, we will restrict our demonstration in this article to the particularly interesting case of transport in a quantum critical metal.[8] In the vicinity of an antiferromagnetic quantum critical point, which marks the onset of magnetic order at zero temperature, quantum critical fluctuations can interact with gapless excitations on a finite Fermi surface. In previous work, our group has established in numerically exact studies[9–12] of sign-free microscopic models that this interplay can give rise to novel non-Fermi liquid regimes and can lead to the emergence of unconventional d-wave superconductivity, making it a candidate mechanism responsible for some of the physics of many actively research materials such as the electron-cuprates, iron-pnictides, and heavy fermion materials - the high-temperature superconductors. We have reported on these activities in a previous contribution to this series.[13] As we will show in this article, QLT offers a new route for the exploration of transport properties across the phase diagram of such quantum critical metals[8] – an insight made possible through our large-scale simulations on the Jülich supercomputing facilities.

2 From Convolutional Neural Networks to Quantum Loop Topography

Arguably, one of the most striking applications of machine learning is pattern recognition in images, the prime examples being handwritten digit classification - for instance

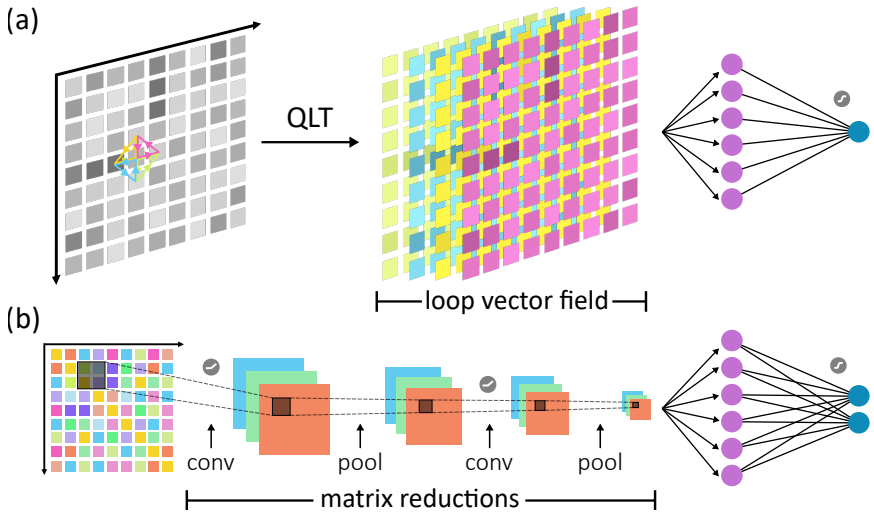


Figure 1. **Neural network architectures**. (a) QLT used as an input to a feed-forward fully-connected shallow neural network with one hidden layer of sigmoid neurons. Only triangular loops $L^{\triangle}_{jkl}$ are illustrated. (b) Deep convolutional neural network that convolves and pooles the unprocessed Green's functions $\widetilde{P}_{jk|\alpha}$ before threading them through a fully-connected layer of hidden neurons. Taken from Ref. 8.

of the famous MNIST data - and facial recognition. Conceptually, the employed strategy is straightforward: given a set of pre-classified images, a statistical model is trained until it has learned enough essential features to be able to distinguish between images of different categories. Hereby, the training amounts to minimising a cost function, which quantifies the deviation from the expected prediction across the training dataset, by a variant of gradient descent. However, despite its simplicity the performance of this approach will generally largely depend on the specific machine learning model and in particular its capability to extract and memorise correlation characteristics. As a consequence, many different architectures like multi-layer perceptrons, support vector machines, and feed-forward neural networks (FFNN) have been proposed over time. A key step in the success story of automated image recognition has been to combine such artificial neural networks with convolutional layers that systematically act as filters or preprocessors on the original data set. This way the identification of patterns is facilitated by exploiting the locality of correlations. In Fig. 1b we illustrate the typical architecture of such a so-called deep convolutional neural network, which consists of multiple successive convolutional layers eventually feeding into a regular FFNN.

Transferring this convolutional prefiltering approach to physics applications has allowed our group to take the first steps in identifying and discriminating different quantum states of many-fermion systems.[2] By interpreting the equal-time Green's functions of a spinful Hubbard-model, produced in quantum Monte Carlo simulations, as *image* snapshots of the physical system – real and imaginary parts are taken as separate colour channels, similar to the well-known RGB colour scheme – we were able to machine learn the distinction between an antiferromagnetic insulator and a semi-metal.[2] This original ap-

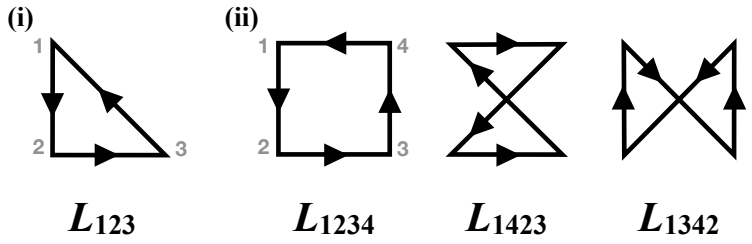


Figure 2. Illustration of the (i) triangular and (ii) quadralateral **loop operators** employed in the loop topography of the longitudinal current-current correlation function. Taken from Ref. 8.

proach, which employed a supervised learning strategy can be further generalised into an unsupervised setting, where the number of distinguishable quantum phases is not a priori known and learned in the process[14] – allowing for a semi-automatic determination of entire phase diagrams.

Some of the more intriguing many-fermion states like superconducting states and topologically ordered states, however, ask for more sophisticated, tailor-made "physics filters" in lieu of the simple convolutional layers in order to allow for a deeper understanding. In a recent work,[8] we have expanded one such approach dubbed "quantum loop topography" (QLT) by its developers,[7] which substitutes the generic convolutional layers by a loop topography scheme based on the specific physical response function of interest. More concretely, targeting longitudinal transport properties we consider the zero-frequency current-current correlation function

$$\Lambda_{xx}(\mathbf{r}_1,\mathbf{r}_2;\omega_n=0)\equiv\int d\tau\left\langle \hat{j}_x\left(\mathbf{r}_1,\tau\right)\hat{j}_x\left(\mathbf{r}_2,0\right)\right\rangle \tag{1}$$

where $\hat{j}_x(\mathbf{r}_1,\tau) = e^{H\tau}\hat{j}_x(\mathbf{r}_1)e^{-H\tau}$ with the current density operator $\hat{j}_x(\mathbf{r_1}) = -i[H(\mathbf{r}_1),\hat{x}]$. The Fourier transform of Λ_{xx} is famously related to the superfluid density ρ_s which, when surpassing a certain critical value, indicates superconducting transport.[15] For a gapped system it can readily be shown[8] that the current-current correlation at zero temperature breaks into a weighted combination of four-vertex loops and triangular loops of Green's functions $P_{\mathbf{r}'\mathbf{r}}=\langle c^\dagger_{\mathbf{r}'}c_{\mathbf{r}}\rangle$,

$$\begin{aligned}\Lambda_{xx}(\mathbf{r}_1,\mathbf{r}_2;\omega_n=0)=&\sum_{\mathbf{r}_3\mathbf{r}_4}P_{\mathbf{r}_2\mathbf{r}_4}P_{\mathbf{r}_4\mathbf{r}_1}P_{\mathbf{r}_1\mathbf{r}_3}P_{\mathbf{r}_3\mathbf{r}_2}\left(x_1-x_4\right)\left(x_2-x_3\right)\\&-\sum_{\mathbf{r}_4}P_{\mathbf{r}_2\mathbf{r}_4}P_{\mathbf{r}_4\mathbf{r}_1}P_{\mathbf{r}_1\mathbf{r}_2}\left(x_1-x_4\right)\left(x_2-x_1\right)\end{aligned} \tag{2}$$

The key idea of QLT is to approximate these current-current correlations by small loop operators (Fig. 2),

$$L^{\triangle}_{jkl}\equiv\widetilde{P}_{jk}|_\alpha\widetilde{P}_{kl}|_\beta\widetilde{P}_{lj}|_\gamma \tag{3}$$

and

$$L^{\square}_{jklm}\equiv\widetilde{P}_{jk}|_{\alpha'}\widetilde{P}_{kl}|_{\beta'}\widetilde{P}_{lm}|_{\gamma'}\widetilde{P}_{mj}|_{\delta'} \tag{4}$$

in which the Green's functions $\widetilde{P}_{jk}|_\alpha$ are typically obtained from individual samples α of quantum Monte Carlo simulations. Pictorially, as shown in Fig. 1a, the QLT amounts to constructing a loop vector field from the expectation values of $L^{\triangle}_{jkl}$ and $L^{\square}_{jklm}$ for different lattice sites in a preprocessing step before feeding the loop topography information into a regular feed-forward neural network. By processing the loop operators during the sampling process and avoiding an *a posteriori* Monte Carlo averaging, we quickly pass these fluctuation-laden data, which encodes partial information of the current-current correlation function, to the machine learning step.

Compared to regular convolutional layers, the QLT as a preprocessing unit has two important advantages. First, since being based on a particular physical correlation function, the information that is passed to the FFNN is transparent and physical. Hence one can draw inferences about the importance of certain physical correlations and can systematically improve the predictive powers of QLT by including higher order loop operators. Second, the QLT approach is vastly more efficient as it operates only a small fraction of the full Green's function data: in two spatial dimensions it scales as $L^2 \times D(d_c)$, where L is the linear system size and $D(d_c)$ denotes the loop vector dimension for a given maximal loop length d_c, compared to the full $L^2 \times L^2$ information. Remarkably, as we demonstrate below, this increased efficiency does not come with a loss in prediction accuracy and thus gives QLT a significant competitive edge over CNNs in transport characterisation applications.

3 Quantum Criticality in a Nearly Antiferromagnetic Metal

To showcase the power of the QLT approach for extracting transport characteristics of itinerant electron systems, we will apply it to the case of a two-dimensional metal near the onset of commensurate antiferromagnetic spin-density wave (SDW) order. Quantum phase transitions in metals pose a substantial theoretical and computational challenge since, in contrast to insulating systems, the order parameter modes can interact with gapless fermion excitations. The quantum critical physics of the system can be captured[11] by an effective Euclidean action $S = S_\psi + S_\phi + S_\lambda$ in terms of fermionic operators $\psi^\dagger, \psi$ (spin and other indices are left implicit) with dispersion $\epsilon_{\mathbf{k}}$ and chemical potential μ coupled to a bosonic order parameter ϕ,

$$\begin{aligned} S_\psi &= \int_{\tau,\mathbf{k}} \psi^\dagger_{\mathbf{k}} \left(\partial_\tau + \epsilon_{\mathbf{k}} - \mu\right) \psi_{\mathbf{k}} \\ S_\phi &= \int_{\tau,\mathbf{x}} \left[\frac{r}{2}\phi^2 + \frac{1}{2}\left(\nabla\phi\right)^2 + \frac{1}{2c^2}\left(\partial_\tau\phi\right)^2 + \frac{u}{4}(\phi^2)^2\right] \\ S_\lambda &= \lambda \int_{\tau,\mathbf{x}} e^{i\mathbf{Q}\cdot\mathbf{x}} \vec{\phi} \cdot \psi^\dagger(\mathbf{x}) \vec{\sigma} \psi(\mathbf{x}) + \text{h.c.} \end{aligned} \tag{5}$$

Here, $\vec{\sigma}$ are spin Pauli matrices, $\mathbf{Q} = (\pi, \pi)$ is the antiferromagnetic ordering wavevector and $\vec{\phi}$ is a 1-, 2-, or 3-component vector (depending on whether the SDW order parameter has easy-axis, easy-plane, or isotropic character, respectively). S_ϕ is a usual Landau-Ginzburg-Wilson ϕ^4-theory, written as an expansion in powers of the order parameter ϕ and its derivatives, where the tuning parameter r drives the system through a quantum critical point (QCP). The interaction term, S_λ, is of Yukawa type, linear in ϕ and quadratic in the fermion operators.

As usual, the central quantity of interest is the partition function $\mathcal{Z} = \int D\left(\psi, \psi^\dagger, \varphi\right) e^{-S}$ which after tracing out the fermion degrees of freedom can be written in determinant quantum Monte Carlo (DQMC) form as[16]

$$\mathcal{Z} = \int D\varphi \, e^{-S_\varphi} \det G_\phi^{-1} + \mathcal{O}\left(\Delta\tau^2\right) \tag{6}$$

Here, $G_\phi = \widetilde{P}_{jk}|_\phi = \langle \psi_j \psi_k^\dagger \rangle_\phi$ is the equal-time Green's function and $\Delta\tau^2$ is the usual controlled Trotter decomposition error.

Generally, quantum Monte Carlo simulations of the partition function for a model of interacting fermions will be hindered by a severe fermion-sign-problem - the integral kernel cannot be safely interpreted as a probability weight, as G_ϕ and its determinant are complex valued. However, as identified by Erez Berg and collaborators in a seminal work,[17] this issue is lifted for model (5) by an intrinsic anti-unitary symmetry (much like time reversal) which renders it amenable to a numerically exact analysis.

As our group was able to show in the first unbiased, numerically exact studies[9, 12] of model (5), which have been conducted on the JURECA and JUWELS supercomputers over multiple funding periods (see also our contribution[13] to this series in 2018), the phase diagram of the antiferromagnetic quantum critical metal clearly reveals a dome-shaped d-wave superconducting phase with a comparably high superconducting transition temperature $T_c \approx E_F/30$ compared to BCS theory (see the inset of Fig. 3). This transition from metallic to superconducting transport, which has been established by (costly) explicit calculations of the superfluid density, serves as a real-world benchmark for the QLT approach.

Applying a supervised learning scheme, we consider a finite-temperature scan cutting into the superconducting dome close to the maximal T_c and train a basic feed-forward neural network on QLT preprocessed equal-time loop correlations (Fig. 1a) at the extremal temperatures $T \approx 0.2$ and $T \approx 0.03$, where the system shows regular metallic and superconducting transport character, respectively. In Fig. 3 we show the neural output of the QLT architecture, which quantifies the confidence that the metal is in a superconducting phase, as a function of the inverse temperature. Clearly, the QLT model is able to reliably distinguish both transport regimes in the high and low temperature limits. Furthermore, it correctly identifies superconducting transport properties over an extended temperature range, as indicated by a high-confidence plateau. Remarkably, the entire trend of the neural output is in great agreement with quantum Monte Carlo results:[9] the notable increase of the confidence coincides with the onset of diamagnetic superconducting fluctuations (grey shaded region) before it indicates an extended superconducting phase at low temperatures, $\beta > \beta_c \approx 12.5$ (dashed line). In comparison to a conventional CNN (Fig. 3), which has been trained on the entire equal-time Green's function data, the QLT analysis provides similar accuracy while only relying on a subset of the data, namely physical small-loop correlations (Fig. 2). Overall, these findings suggest that the QLT model has efficiently identified and captured the superconducting transport characteristics across the phase diagram.

4 Concluding Remarks

In this article we have reported on a recent advance in the application of machine learning in condensed matter research: the probing of transport properties in itinerant quantum

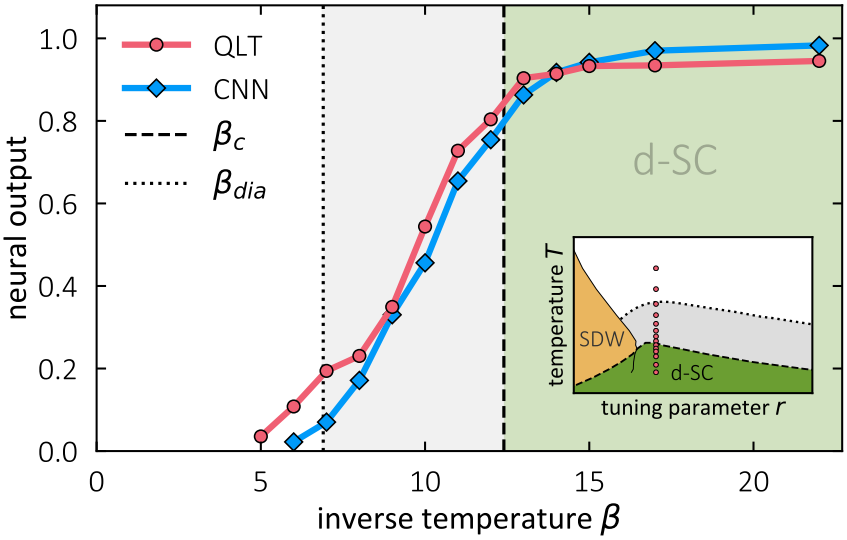


Figure 3. Neural network output for the superconducting transition near a **metallic easy-plane antiferromagnetic quantum critical point**. The training points were at $\beta = 30$ for superfluid transport and $\beta = 5$ for metallic transport. The vertical lines indicate the superconducting transition temperature $\beta_c \sim 12.5$ (dashed) derived from the superfluid density measurements and the onset of diamagnetic fluctuations $\beta_{\rm dia} \sim 6.9$ (dotted) where the orbital magnetic susceptibility changes sign.[9] The inset (modified from Ref. 9) illustrates the chosen finite-temperature scan cutting into the superconducting dome. Taken from Ref. 8.

many-fermion systems by quantum loop topography.[7,8] In contrast to directly applying feature extraction techniques (often originally developed for pattern recognition in images) to physics problems, the QLT approach stands out as a preprocessing scheme that filters the relevant correlations of a physical response function of interest. As we have showcased for a quantum critical metal the QLT approach is capable of reliably identifying a change in transport to reasonable accuracy while being vastly more efficient than traditional calculations, based on the superfluid density, and convolutional neural network architectures. Building upon these results, a natural question is whether, for a sufficiently suppressed superconducting dome, QLT can be used for the exploration of a potential fan-shaped quantum critical regime where transport is governed by the underlying QCP at $T = 0$. In this regard, we believe that the transparent physical basis of QLT is a valuable step in the implementation of machine learning methods that, when combined with analytical insight, will allow us to draw inferences about the physical mechanisms at play.

Acknowledgements

We are indebted to our collaborators Eun-Ah Kim, Yi Zhang, Erez Berg, and Yoni Schattner for many insightful discussions. We gratefully acknowledge the computing time on the JUWELS/JURECA clusters at the Forschungszentrum Jülich, which have made these numerical studies possible in the first place. Our research has been supported, in part, by

the collaborative research grants SFB 1238 (projects C02 and C03) and Transregio CRC 183 (project B01) by the Deutsche Forschungsgemeinschaft (DFG).

References

1. J. Carrasquilla and R. G. Melko, *Machine learning phases of matter*, Nature Physics **13**, 431–434, 2017.
2. P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, *Machine learning quantum phases of matter beyond the fermion sign problem*, Scientific Reports **7**, 8823, 2017.
3. G. Carleo and M. Troyer, *Solving the quantum many-body problem with artificial neural networks*, Science **6325**, 602–606, 2017.
4. G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, *Neural-network quantum state tomography*, Nature Physics **5**, 447–450, 2018.
5. J. Liu, H. Shen, Y. Qi, Z. Y. Meng, and L. Fu, *Self-learning Monte Carlo method and cumulative update in fermion systems*, Phys. Rev. B **24**, 241104, 2017.
6. L. Huang and L. Wang, *Accelerated Monte Carlo simulations with restricted Boltzmann machines*, Phys. Rev. B **95**, 035105, 2017.
7. Y. Zhang and E.-A. Kim, *Quantum Loop Topography for Machine Learning*, Phys. Rev. Lett. **118**, 216401, 2017.
8. Y. Zhang, C. Bauer, P. Broecker, S. Trebst, and E.-A. Kim, *Probing transport in quantum many-fermion simulations via quantum loop topography*, Phys. Rev. B. **99**, 161120, 2019.
9. Y. Schattner, M. H. Gerlach, S. Trebst, and E. Berg, *Competing Orders in a Nearly Antiferromagnetic Metal*, Phys. Rev. Lett. **117**, 097002, 2016.
10. Max H. Gerlach, Yoni Schattner, Erez Berg, and Simon Trebst, *Quantum critical properties of a metallic spin density wave transition*, Phys. Rev. B. **95**, 035124, 2017.
11. E. Berg, S. Lederer, Y. Schattner, and S. Trebst, *Monte Carlo Studies of Quantum Critical Metals*, Annual Review of Condensed Matter Physics **10**, 63–84, 2019.
12. C. Bauer, Y. Schattner, S. Trebst, and E. Berg, *Hierarchy of energy scales in an antiferromagnetic quantum critical metal: a Monte Carlo study*, in preparation.
13. S. Trebst, *Exact Results for the Many-Electron Problem: Competing Orders in a Nearly Antiferromagnetic Metal*, in NIC Symposium 2018 - Proceedings, NIC Series Vol. 49, 289–296, 2018.
14. P. Broecker, F. F. Assaad, and S. Trebst, *Quantum phase recognition via unsupervised machine learning*, 2017, arXiv:1707.00663 [cond-mat.str-el].
15. D. J. Scalapino, S. R. White, and S. Zhang, *Insulator, metal, or superconductor: The criteria*, Phys. Rev. B **47**, 7995–8007, 1993.
16. R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, *Monte Carlo calculations of coupled boson-fermion systems. I*, Phys. Rev. D **24**, 2278, 1981.
17. E. Berg, M. Metlitski, and S. Sachdev, *Sign-Problem-Free Quantum Monte Carlo of the Onset of Antiferromagnetism in Metals*, Science **338**, 1606–1609, 2012.

Cooperation of Many-Body Physics and Defect Chemistry in Transition-Metal Oxides

Frank Lechermann

I. Institut für Theoretische Physik, Universität Hamburg,
20355 Hamburg, Germany
E-mail: Frank.Lechermann@physnet.uni-hamburg.de

The problem of defects in correlated materials is at the heart of the fascinating phenomenology of many of these compounds. A vast number of prominent features of strongly correlated systems, such as *e. g.* high-temperature superconductivity in cuprates and iron pnictides, or heavy-fermion physics in Ce-based compounds is often directly associated with a defect-crystal state. Already the very concept of a doped Mott-insulator builds up on the understanding of impurities implanted in an otherwise perfect crystal lattice. However, a deeper understanding of the realistic physics is then connected to a faithful description of the defect chemistry underlying the material under consideration. We here show that the combination of density functional theory (DFT) with dynamical mean-field theory (DMFT) provides a proper tool to elucidate this realistic interplay between many-body physics and defect chemistry. Focus is on transition-metal oxides which are well known to harbour diverse manifestations of electronic correlations. Two prominent concrete examples, the paramagnetic metal-to-insulator transition in V_2O_3 driven by chromium doping, and the long-standing issue of lithium-doped NiO will be addressed in some detail.

1 Introduction

Defects in crystal lattices drive very important phenomenology in condensed matter physics. Prominent examples are for instance the impurity doping of semiconductors such as silicon to facilitate conductivity or the movement of dislocations in metals as the origin for plasticity. Traditionally, one characterises such defects by their dimensionality: zero-dimensional point defects (*e. g.* vacancies or antisite atoms), one-dimensional line defects (*e. g.* dislocations) and two-dimensional planar defects (*e. g.* grain boundaries). In this brief overview, we will focus on the aspect of point defects in prominent correlated transition-metal oxides.

The effect of realistic point defects are believed to be at the root of many challenging correlation phenomena. For instance, the high-temperature superconductivity of cuprates[1] such as La_2CuO_4 or $YBa_2Cu_3O_7$ originates from hole doping the CuO_2 layers in these systems. This is realised either by directly achieving oxygen deficiency in those layers, or by introducing different-valence elements in the region in between, *e. g.* partially replacing La^{3+} by Sr^{2+} in the case of La_2CuO_4. As another example, the Ce defect compound $CeCu_{6-x}Au_x$ is a paramount case of a heavy-fermion system, possibly harbouring a quantum-critical point. Thereby, the concentration balance x between Cu and Au is a key control parameter. Furthermore, the introduction of point defects is a proper and versatile tool to modify, tune and engineer correlated materials properties in view of possible technological applications. For instance, in the areas of photovoltaic materials and battery materials, defect engineering of charge-gap sizes or redox potentials is an important optimisation procedure. Thus, a deeper realistic understanding of the interplay of electronic

correlations and defect properties is relevant from both perspectives, basic research as well as technological applications.

Of course, the theoretical modelling of point defects in materials has a long history. Yet especially for correlated materials, the tools for an investigation on a realistic level have been limited until recently. There are several limiting perspectives that are often utilised to account for the impact of defects in quantum materials. These include *e. g.* volume or lattice-parameter modifications, shifting of the chemical potential, effective-medium or virtual-crystal considerations, or the sole formation of impurity bands. While those approaches to realistic doping may provide reasonable results in certain cases, they are surely restricted, miss relevant aspects and even can be misleading. It is however also noteworthy to state that a perfectly satisfying approach does yet not exit. Randomly-distributed point defects break the translational invariance of the crystal lattice, which render usual $\mathbf{k}$-space approaches only approximate. Still, a sound first-principles modelling should be able to tackle both, local as well as long-ranged effects of the defect introduction to the lattice. Therefore, supercell approaches are generally believed to mark the best-tailored approximations to the problem. Though they necessarily introduce a periodical reproduction and hence an artificial ordering pattern of defects, this failure can in principle be converged to negligible contribution by further and further enlargement of the supercell. In practise, since a quantum-physical relevant defect impact usually arise only for concentrations $> 1\%$, supercell sizes of up to 100-200 atomic sites turn out sufficient for most materials under consideration to obtain reasonable results. Note that we here do not aim at *e. g.* metallurgical problems such a dislocation dynamics or crack propagation on an atomistic level, which often ask for more than 10^6 atoms in a (semi)classical molecular-dynamics simulation. First-principles supercell approaches enable the description of local effects, such as *e. g.* local distortions or local symmetry breakings, and global effects, such as *e. g.* lattice-parameter changes or modifications of the lattice electronic structure.

2 Theoretical approach: multi-site DFT+DMFT

The proper first principles approach for defects in correlated materials asks for the following ingredients. A faithful description of the chemical bonding and lattice effects, a general and accurate account of many-body physics, both with reasonable numerical performance in a supercell architecture. The charge self-consistent combination of density functional theory (DFT) with dynamical mean-field theory (DMFT) in a multi-site or so-called "real-space" setting fits to these demands. We here provide only a brief review of the DFT+DMFT framework, more detailed descriptions can be found elsewhere (*e. g.* Ref. 2).

Key to DFT+DMFT is the interfacing of both methods, usually realised by subsequent down- and upfolding between a general Bloch space and a chosen correlated subspace $\mathcal{C}$. The Kohn-Sham DFT treatment of the problem covers the whole Bloch space, while DMFT is explicitly active only in $\mathcal{C}$. Charge self-consistency via an update of the Kohn-Sham potential by including DMFT self-energy effects for the charge density, ensures a thorough coupling within this hybrid scheme. The correlated subspace is obtained from a Wannier(-like) construction for the correlation-relevant quantum space. The d shell of transition metals or the f shell of rare-earth atoms usually form the latter, if they actively take part in the charge fluctuations within the system. Locally, a multi-orbital Hubbard Hamiltonian is applied $\mathcal{C}$, usually parametrised in a spherical form by a Hubbard interaction

U and a Hund's exchange J_{H}.

In the supercell approach, the correlated subspace spans over all correlated sites, whereby each of these sites defines an impurity problem in the DMFT sense.[3] The coupling is realised via the DFT+DMFT self-consistency condition invoking the computation of the complete lattice Green's function. This multi-site or real-space framework generalises DMFT to the problem of various coupled single-site impurity problems on a given lattice. Note that it is only necessary to explicitly compute the impurity self-energy on symmetry-inequivalent sites and then translate the self-energy to the equivalent sites.

The results discussed here were obtained with a mixed-basis pseudopotential code[4] for the DFT part and by hybridisation-expansion continous-time quantum Monte Carlo (CT-QMC)[5] as implemented in the TRIQS package[6] for the DMFT impurity solution. For more technical details on the DFT+DMFT implementation the reader is referred to Ref. 2. The computational effort scales with the number of inequivalent sites to be treated in DMFT, the CT-QMC part marking the main bottleneck of the calculation. However, the DFT part of the problem may also become demanding in the supercell approach. Note that computationally, the calculations are challenging from both aspects, numerical performance as well as memory issues.

3 Cr and Ti doping of V_2O_3

The phase diagram of V_2O_3[7–9] (see Fig. 1) poses a sophisticated condensed-matter problem for about 50 years. The correlated oxide at stoichiometry is a paramagnetic metal (PM) at ambient conditions, but transforms into an antiferromagnetic insulator (AFI) well below 200 K. Application of pressure as well as doping with Ti stabilises the PM phase. Doping with Cr leads to a paramagnetic insulator (PI) at a certain concentration threshold. Measurements show that the introduction of Cr expands furthermore the lattice. Therefore, to a first approximation, doped V_2O_3 has long been thought as a realistic case of a canonical Mott system: applied pressure leads to strong hybridisation and hence stronger metallicity, while negative pressure strengthens the localisation tendencies, hence stabilises the Mott

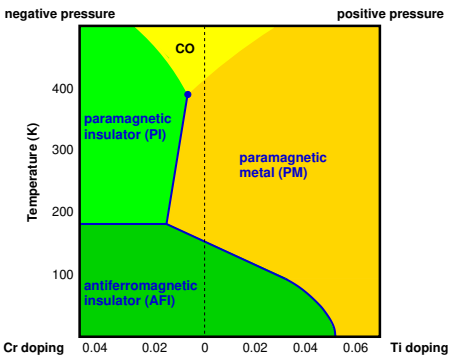


Figure 1. Experimental phase diagram of V_2O_3 as given by Refs. 7–9. Paramagnetic metal (PM) at ambient conditions and with Ti doping, paramagnetic insulator (PI) with Cr doping and at negative pressure, antiferromagnetic insulator (AFI) at low temperature.

insulator. The localised spin degree furthermore orders antiferromagnetically at low T due to the well-known kinetic-exchange mechanism.

However based on further experimental facts[10] and recent DFT+DMFT calculations,[2, 11–14] it has become clear that this appealing simplistic picture falls too short. For instance, it has long been shown that the Ti doping also expands the lattice,[15] though not that strongly as Cr doping. This is an important lesson: though applied pressure and Ti doping have the same principal effect on the conductivity, that does not immediately mean that the physical reason behind is identical, *i. e.* that both processes contract the lattice. Furthermore, the recent DFT+DMFT[14] study revealed that for reasonable local Coulomb interactions the lattice-expansion effect in the Cr-doped case alone cannot account for the appearance of the PI phase. Thus there has to be something more to both doping scenarios, in order to explain the long-standing phase-diagram findings. We here briefly review the key content of Ref. 14, and refer to the original publication for further details.

The DFT+DMFT investigation was performed with using the V$(3d)$-t_{2g} threefold orbital sector, *i. e.* two e_g^π and one a_{1g} orbital, as the local correlated subspace. A Slater-Kanamori Hamiltonian parametrised by $U = 5\,\mathrm{eV}$ and $J_\mathrm{H} = 0.7\,\mathrm{eV}$ governs that space. The doping scenarios are realised by 80-atom supercells, where one V site is replaced by Cr or Ti, respectively (see Fig. 2a). This amounts to a defect concentration of 3.1 %, matching experimental reality. Structural relaxation of the defect cells is performed within DFT+U. Fig. 2b/c summarise the key findings of the theoretical study at $T = 194\,\mathrm{K}$. The stoichiometric reference phase is verified metallic with sizeable correlations, visible by the clear formation of lower and upper Hubbard bands in line with a renormalised quasiparticle

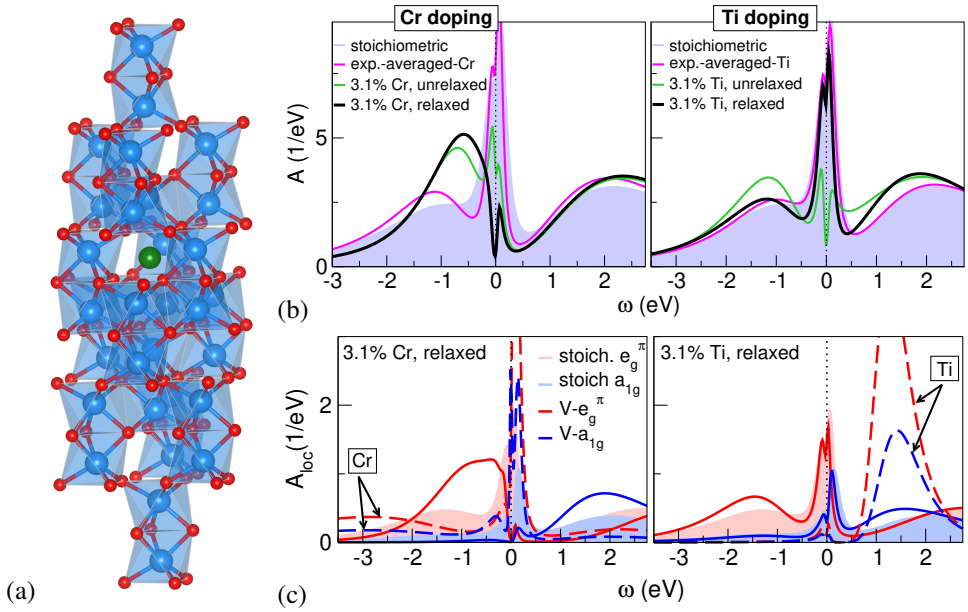


Figure 2. Theoretical modelling of 3.1 % Cr or Ti doping of V_2O_3. (a) 80-atom supercell with V (blue), O (red) and impurity (green). (b,c) DFT+DMFT result for the total (b) and local (c) spectral function for various settings at $T = 194\,\mathrm{K}$.

(QP) peak at the Fermi level. On the local level, there is polarisation towards e_g^π, slightly enhanced compared to the orbital polarisation within DFT.

Let's then focus first on the Cr doping. The simplest treatment is still performed in the original small primitive cell, but using the experimental averaged lattice data that comes with the Cr-induced expansion. Correlations are further enhanced within this picture, yet an insulating regime may not be realised. Next, a supercell calculation with still unrelaxed atomic positions already leads to a further substantial spectral-weight transfer and a weakened QP peak. Thus the sole substitution of V by non-isovalent Cr already leads to charge fluctuations/transfers and symmetry reduction in the system, and increases electronic correlations. Additional structural relaxations finally lead to the opening of charge gap of size $\sim 0.1\,\mathrm{eV}$ in good agreement with experimental results.[16, 17] The local spectral functions show finally a very strong orbital polarisation towards e_g^π, which qualitatively also matches with an experimental comparison between the stoichiometric and the Cr-doped phase.[18] The chromium site nonetheless essentially remains in the expected Cr^{3+} state, with its t_{2g} states located around $\sim -2.5\,\mathrm{eV}$. However, fluctuations with the host atoms are still very strong, as evidenced by a strong Cr resonance just above the charge gap. The key reason for the Mott-insulating instability with Cr doping are local symmetry breakings from trigonal to monoclinic that occur upon structural relaxation. Thereby, the degeneracy of both e_g^π states is lifted and the strong correlations are then very effective in driving the system towards the insulating regime.

Turning to the Ti-doped case, the physics is quite different. The introduction of Ti leads in the end to a correlated metal, in agreement with experiment. The Ti-t_{2g} states are well above the Fermi level, hence one is facing a Ti^{4+} oxidation state. In other words, the titanium impurity is in a $3d^0$ state and effectively dopes its surrounding with one additional electron. From a low-temperature perspective, this charge-doping effect of Ti renders Mott-insulating V_2O_3 eventually metallic and prohibits insulating tendencies also at higher temperatures. Note that upon structural relaxation, the local monoclinic distortions that take place for Cr doping are absent in the Ti-doped case.

This essential differences for Cr and Ti doping explain the key features of the experimental V_2O_3 phase diagram. It becomes very clear that albeit the trends with doping seem to roughly match with a simplistic lattice expansion *vs.* contraction picture, the underlying physics is much more subtle.

4 Li doping of NiO

In the second application to be discussed in this brief review we want to focus on the late transition-metal oxide NiO. Contrary to early transition-metal oxides like V_2O_3, the charge-transfer energy Δ is much smaller in late transition-metal oxides, therefore a charge gap with correlations is often set by Δ, and charge-transfer insulators instead of Mott-Hubbard insulators are realised.[19] For NiO indeed $U > \Delta$ holds, and the O$(2p)$ states reside in between the lower and upper Hubbard band of the Ni$(3d)$ states.[20] Yet Mott-Hubbard physics is still also a vital player in determining the correlated electronic structure for the insulator with a charge gap of $\sim 4\,\mathrm{eV}$. Replacing part of Ni by Li in nickel oxide has long been known to be effective in hole doping the compound.[21, 22] Recently, doped NiO gained renewed interest in view of photovoltaic applications,[23, 24] since pronounced in-gap states (IGS) appear at $\sim 1.2\,\mathrm{eV}$.[25, 26]

Here, we briefly review the recent first-principles many-body study of Li-doped NiO, for further details we refer to Ref. 27. In order to account for the stoichiometric and doped electronic spectrum of NiO with state-of-the-art DFT+DMFT, the charge-transfer tendencies have to be reasonably described. This means not only the electronic correlations emerging from Ni($3d$), but also from O($2p$) have to be described beyond DFT. Yet a full many-body account treatment of those ligand states together with the ones from the transition-metal sites proves cumbersome for various reasons. A simpler static handling of the O($2p$) correlations is sufficient to a good approximation. In our DFT+sicDMFT approach this is achieved by applying a self-interaction correction (SIC) to the O($2s$) and O($2p$) states on a pseudopotential level within a complete charge self-consistent DFT+DMFT scheme. This is very efficient since it does not ask for a major change of the existing self-consistency procedure, and boils down to only one additional α parameter for the SIC of O($2p$). The latter is chosen $\alpha = 0.8$ and should hold for a wide range of transition-metal oxides. The DMFT correlated subspace consists still of the relevant transition-metal orbital, *i. e.* here the complete Ni($3d$) shell, with a nominal $3d^8$ filling for present Ni^{2+}. A full Slater-Condon 5-orbital Hamiltonian for $U = 10\,\text{eV}$ and $J_\text{H} = 1.0\,\text{eV}$ acts in the correlated subspace. Since Li can be doped at rather large amounts into NiO, we choose a doping level of $x = 0.125$ for $\text{Li}_x\text{Ni}_{1-x}\text{O}$ (see Fig. 3a).

Fig. 3b shows that we indeed find the IGS essentially right at the experimentally known position of 1.2 eV. Also the further features of the stoichiometric and doped spectrums agree to a very large extent with the experimental data .[20, 25, 26] Moreover, it can be revealed that the formation of the IGS is at the expense of part of the competing correlated charge-transfer states, *i. e.* d^8 and ligand hole as well as Zhang-Rice doublet, in NiO.

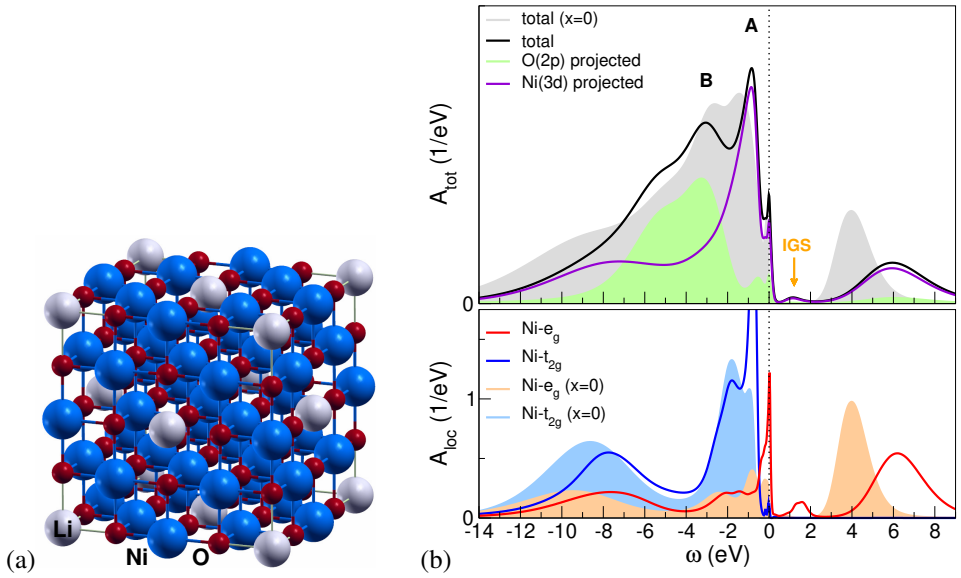


Figure 3. Theoretical modelling of 12.5 % Li doping of NiO. (a) 16-atom supercell with Ni (blue), O (red) and Li (grey). (b) DFT+sicDMFT result for the stoichiometric and doped case: total (top) and local (bottom) spectral function at $T = 582\,\text{K}$.

5 Summary

It was shown that state-of-the-art DFT+DMFT as well as its extension to DFT+sicDMFT, represent powerful tools to tackle prominent point-defect problems in correlated materials. The results underline furthermore the often underrated very important effects of dopants, sometimes referred to as "dirt". Working with this "dirt" is an indispensable way to tailor, tune and engineer correlated materials, and a deeper understanding of the associated physics inevitable for designing new states of matter in the future.

Acknowledgements

Financial support from the DFG project "Design of strongly correlated materials" LE 2446/4-1 is acknowledged. Computations were performed at the University of Hamburg and the JURECA/JUWELS Clusters of the Jülich Supercomputing Centre (JSC) under project number hhh08.

References

1. J. G. Bednorz and K. A. Müller, *Possible highTc superconductivity in the Ba-La-Cu-O system*, Z. Physik B – Condensed Matter **64**, 189, 1986.
2. D. Grieger, C. Piefke, O. E. Peil, and F. Lechermann, *Approaching finite-temperature phase diagrams of strongly correlated materials: A case study for V_2O_3*, Phys. Rev. B **86**, 155121, 2012.
3. M. Potthoff and W. Nolting, *Surface metal-insulator transition in the Hubbard model*, Phys. Rev. B **59**, 2549, 1999.
4. B. Meyer, C. Elsässer, F. Lechermann, and M. Fähnle, "Fortran 90 program for mixed-basis-pseudopotential calculations for crystals".
5. P. Werner, A. Comanac, L. de' Medici, M. Troyer, and A. J. Millis, *Continuous-time solver for quantum impurity models*, Phys. Rev. Lett. **97**, 076405, 2006.
6. P. Seth, I. Krivenko, M. Ferrero, and O. Parcollet, *TRIQS/CTHYB: A continuous-time quantum Monte Carlo hybridisation expansion solver for quantum impurity problems*, Comput. Phys. Commun. **200**, 274, 2016.
7. D. B. McWhan, T. M. Rice, and J. B. Remeika, *Mott Transition in Cr-Doped V_2O_3*, Phys. Rev. Lett. **23**, 1384, 1969.
8. D. B. McWhan, J. B. Remeika, T. M. Rice, W. F. Brinkman, J. P. Maita, and A. Menth, *Electronic Specific Heat of Metallic Ti-Doped V_2O_3*, Phys. Rev. Lett. **27**, 941, 1971.
9. D. B. McWhan, A. Menth, J. B. Remeika, T. M. Rice, and W. F. Brinkman, *Metal-Insulator Transitions in Pure and Doped V_2O_3*, Phys. Rev. B **7**, 1920, 1973.
10. I. Lo Vecchio, J. D. Denlinger, O. Krupin, B. J. Kim, P. A. Metcalf, S. Lupi, J. W. Allen, and A. Lanzara, *Fermi Surface of Metallic V_2O_3 from Angle-Resolved Photoemission: Mid-level Filling of e_g^π Bands*, Phys. Rev. Lett. **117**, 166401, 2016.
11. D. Grieger and F. Lechermann, *Effect of chromium doping on the correlated electronic structure of V_2O_3*, Phys. Rev. B **90**, 115115, 2014.
12. X. Deng, A. Sternbach, K. Haule, D. Basov, and G. Kotliar, *Shining light on transition-metal oxides: unveiling the hidden fermi liquid*, Phys. Rev. Lett. **113**, 246404, 2014.

13. I. Leonov, V. I. Anisimov, and D. Vollhardt, *Metal-insulator transition and lattice instability of paramagnetic V_2O_3*, Phys. Rev. B **91**, 195115, 2015.
14. F. Lechermann, N. Bernstein, I. I. Mazin, and R. Valentí, *Uncovering the Mechanism of the Impurity-Selective Mott Transition in Paramagnetic V_2O_3*, Phys. Rev. Lett. **121**, 106401, 2018.
15. S. Chen, J. E. Hahn, C. E. Rice, and W. R. Robinson, *The effects of titanium or chromium doping on the crystal structure of V_2O_3*, J. of Solid State Chem. **44**, 192, 1982.
16. A. S. Barker Jr. and J. P. Remeika, *Optical properties of V_2O_3 doped with chromium*, Solid State Comm. **8**, 1521, 1970.
17. S.-K. Mo, H.-D. Kim, J. D. Denlinger, J. W. Allen, J.-H. Park, A. Sekiyama, A. Yamasaki, S. Suga, Y. Saitoh, T. Muro, and P. Metcalf, *Photoemission study of $(V_{1-x}M_x)O_3$ (M=Cr, Ti)*, Phys. Rev. B **74**, 165101, 2006.
18. J.-H. Park, L. H. Tjeng, A. Tanaka, J. W. Allen, C. T. Chen, P. Metcalf, J. M. Honig, F. M. F. de Groot, and G. A. Sawatzky, *Spin and orbital occupation and phase transitions in V_2O_3*, Phys. Rev. B **61**, 11506, 2000.
19. J. Zaanen, G. A. Sawatzky, and J. W. Allen, *Band gaps and electronic structure of transition-metal compounds*, Phys. Rev. Lett. **55**, 418, 1985.
20. G. A. Sawatzky and J. W. Allen, *Magnitude and origin of the band gap in NiO*, Phys. Rev. Lett. **53**, 2339, 1984.
21. R. R. Heikes and W. D. Johnston, *Mechanism of Conduction in Li-Substituted Transition Metal Oxides*, J. Chem. Phys. **26**, 582, 1957.
22. J. B. Goodenough, D. G. Wickham, and W. J. Croft, *Some magnetic and crystallographic properties of the system $Li_x^+Ni_{1-2x}^{++}Ni_x^{+++}O$*, J. Phys. Chem. Solids **5**, 107, 1958.
23. J. Y. Jeng, K. C Chen, T. Y. Chiang, P. Y Lin, T. D Tsai, Y. C Chang, T. F Guo, P. Chen, T. C. Wen, and Y. J. Hsu, *Nickel Oxide Electrode Interlayer in $CH_3NH_3PbI_3$ Perovskite/PCBM Plana-Heterojunction Hybrid Solar Cells*, Adv. Mater. **26**, 4107, 2014.
24. W. Chen, Y. Wu, Y. Yue, J. Liu, W. Zhang, X. Yang, H. Chen, E. Bi, I. Ashraful, M. Grätzel, and L. Han, *Efficient and stable large-area perovskite solar cells with inorganic charge extraction layers*, Science **350**, 944, 2015.
25. J. van Elp, H. Eskes, P. Kuiper, and G. A. Sawatzky, *Electronic structure of Li-doped NiO*, Phys. Rev. B **45**, 1612, 1992.
26. J. Y. Zhang, W. W. Li, R. L. Z. Hoye, J. L. MacManus-Driscoll, M. Budde, O. Bierwagen, L. Wang, Y. Du, M. J. Wahila, L. F. J. Piper, T.-L. Lee, H. J. Edwards, V. R. Dhanak, and K. H. L. Zhang, *Electronic and transport properties of Li-doped NiO epitaxial thin films*, J. Mater. Chem. C **6**, 2275, 2018.
27. F. Lechermann, W. Körner, D. F. Urban, and C. Elsässer, *Interplay of charge-transfer and Mott-Hubbard physics approached by an efficient combination of self-interaction correction and dynamical mean-field theory*, Phys. Rev. B **100**, 115125, 2019.

SDS Adsorptions Films at the H_2O – Au(111) Interface: Molecular Dynamics Study of AFM Tip–Surface Contact

Johannes Hörmann and Lars Pastewka

Department of Microsystems Engineering, University of Freiburg,
Georges-Köhler-Allee 103, 79110 Freiburg, Germany
E-mail: lars.pastewka@imtek.uni-freiburg.de

On the example of the anionic model surfactant sodium dodecyl sulphate at the idealised water-gold interface, our present molecular dynamics simulations illuminate the interdependence of adsorption film morphology and its mechanical properties at the nanoscale. Distinct film phases exhibit unique force response features in modelled atomic force microscopy experiments.

1 Introduction

In aqueous solution, amphiphilic surfactants adsorb on immersed surfaces. Dependent on the specific molecule and concentration, they form surface structures of varying morphology: monolayers of flat-lying molecules, hemicylindrical stripes or full cylinders have all been observed. These films exhibit morphology-dependent mechanical properties, such as nanoscopic response under the atomic force microscope (AFM)[1] as well as macroscopic lubrication performance.[2] Our work focuses on the model surfactant sodium dodecyl sulphate (SDS) at the H_2O – Au(111) interface, known to form flat-lying monolayers at low concentrations and stripe-like hemicylindrical aggregates with increasing adsorption mass.[3] This phase transition alters the boundary film's mechanical properties, and it was hypothesised that this facilitates stick-slip response[4] under frictional loading. We use an existing SDS parametrisation to simulate AFM experiments by means of classical all-atom molecular dynamics (MD), as illustrated in Fig. 1. Systematic exploration of the parametric space spanned by concentration, film morphology, as well as AFM probe approach velocity reveals distinctive behaviour along each of these dimensions.

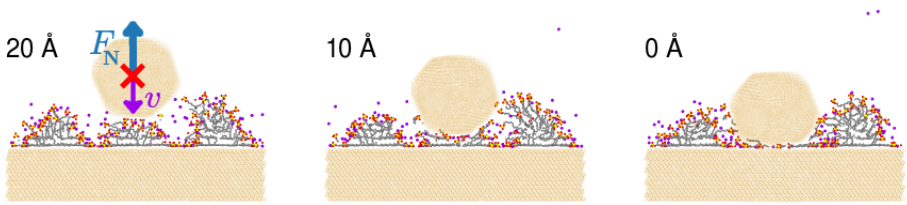


Figure 1. $v = 1\ \mathrm{m\,s^{-1}}$ approach of model AFM probe towards hemicylindrical SDS aggregates on Au(111). A purple velocity vector v implies the probe's linear approach, while a blue force vector F_N illustrates the film's resistance felt by the probe. Snapshot labels indicate surface–surface distance between probe and substrate.

2 Methods

We carry out classical MD simulations employing a valence force-field that allows modelling of physisorption processes but does not include breaking of covalent bonds. Parameters are taken from CHARMM[5] 36 and the INTERFACE force-field.[6] CHARMM has been reported to reproduce micellar properties of SDS well.[7] The rigid water model TIP3P (CHARMM standard) constitutes our explicit solvent. An embedded atom method potential by Grochola[8] describes the Au-Au interaction. A cubic Au block with one of its $\{111\}$ surfaces exposed to the solution containing the SDS serves as substrate. It is sized as to accommodate three hemicylindrical surfactant aggregates of ~ 2.5 nm experimentally determined radius[9] on its surface. Sampling the full adsorption process from solution lies out of reach within the timescales accessible by brute-force MD. Hence, we preassemble a set of monolayers and hemicylinders at appropriate surface concentrations covering the adsorption isotherm's region of film phase transition. After minimisation, NVT-, and NPT-equilibration, all systems evolve for 10 ns, with only the substrate Langevin-tempered while pressurising the solution along the surface normal.

AFM probe models are prepared by melting gold spheres of 2.5 nm initial radius and subsequent slow quenching from 1800 K down to 298 K over a time span of 100 ns at 5 fs time step. Insertion at surface–surface distance $d = 3$ nm – sufficiently far as not to disturb the adsorption film structure – follows the parallel alignment of a probe's $\{111\}$ facet with the substrate's (111) surface. Note that all surface–surface distances are stated with respect to perfect overlap of substrate's and probe's outermost atomic layer as zero reference. Succeeding repeated minimisation, NVT- and NPT-equilibration, substrate and probe are tempered by a dissipative particle dynamics (DPD) thermostat[10, 11] while holding the system volume fixed. After 0.5 ns, the probe model's frozen core of 2.4 nm diameter is instantaneously set into motion at prescribed velocity, approaching the substrate with its bottommost 1.4 nm layer immobile.

All calculations at the interface apply a 2 fs time step, 3D-periodic boundary conditions, and particle-particle particle-mesh Ewald summation treatment of long interactions. In the above, standard conditions of temperature $T = 298$ K and pressure $P = 1013$ hPa hold wherever applicable.

3 Results & Discussion

Estimating the drag force of our model AFM probe in water at standard conditions by means of Stokes' law for a bead in viscous fluid, $F_d = 6\pi\eta r v$, with probe radius

Stokes drag of the AFM probe			
velocity v [m s^{-1}]	10	1	0.1
analytical estimate F_d [nN]	0.15	0.02	~ 0
MD average F_d [nN]	1.27	0.15	0.02

Table 1. Viscous drag in TIP3P water as estimated for a bead of $r = 2.5$ nm by Stokes' law and as recorded via MD for our AFM tip model in bulk solution.

$r = 2.5$ nm and dynamic viscosity of the TIP3P rigid water model,[12] $\eta = 0.321$ mPa s, yields non-negligible drag for $v = 10$ m s^{-1} compared to the expected magnitude of adsorption film response. Tab. 1 shows estimated and measured viscous drag at slower speeds to diminish in comparison to the encountered film resistance. Viscous drag was measured as the average force acting on the probe for distances $d > 2.3$ nm above the substrate. Deviations from Stokes estimate are likely due to hydrodynamic interactions between tip and surface at the distances probed here.

Characteristic force response features are unidentifiable in the fast approach case of $10\,\mathrm{m\,s^{-1}}$, as apparent in Fig. 2. Even the two slower approaches still exhibit non-negligible quantitative discrepancies. Yet, the two slowest approaches produce comparable qualitative features. From inset A of Fig. 2 (distance $d \sim 17$ Å) to Fig. 2B ($d \sim 13$ Å), the slightly compressed monolayer begins to reorder laterally under applied load. This is accompanied by a repulsive force on the indenter. For the slowest approach rate (0.1 m s^{-1}), there is a slight increase in force between snapshots A and B that we attribute to the initial triggering of the lateral tilt of the molecules. From Fig. 2C ($d \sim 11$ Å) to D ($d \sim 9$ Å), a forcefully

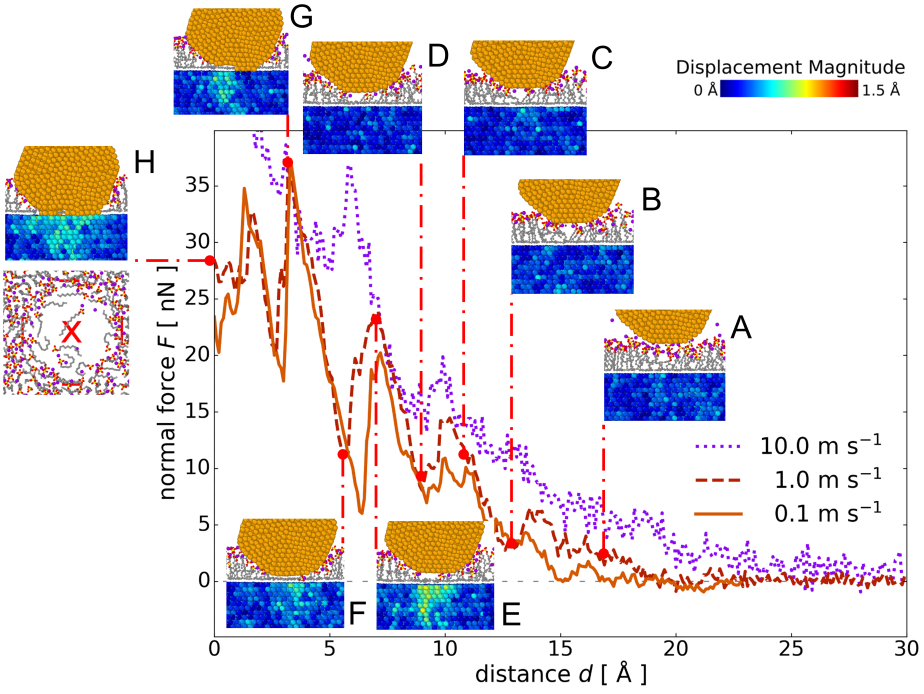


Figure 2. Velocity dependency of AFM force response. Figure shows normal force F on the AFM tip as a function of tip–surface distance d. Values at large d correspond to Stokes drag on the AFM tip reported in Tab. 1. The colour-coding of substrate atoms in the insets shows the displacement of each atom from its initial position for this particular snapshot. Varying substrate colouring for insets at $d \geq 10\,$Å (A-D) reflects thermal fluctuations. Slight substrate deformation is evident for $d < 10\,$Å (E-H). Panel H shows a top-view of only the SDS molecules in the contact area.

restructured film pushes the surfactant's hydrocarbon tails towards the surface, where they partially adhere to the substrate. This is accompanied by a slight drop in the repulsive force on the tip. At Fig. 2E ($d \sim 7$ Å), the head groups of molecules remaining within the gap protrude above a monomolecular hydrocarbon layer. Strong repulsion arises from the resistance against becoming a fully flat-lying monolayer, which needs to be accompanied by the movement of molecules out of the contact.

The colour-coded magnitude of substrate atom displacement (measured with respect initial position) reveals an indentation of the substrate clearly distinguishable from thermal fluctuations at this distance. While the colours in panels A-D are purely due to thermal fluctuations, panels E-H (forces $\gtrapprox 10$nN) reveal and extended region of finite displacement that is due to the elastic deformation of the substrate. This means the surface does not feel a significant force from the indenter during initial approach A-D, as the whole deformation is localised within the thin adsorption film.

At Fig. 2F ($d \sim 5.5$ Å), only a monomolecular layer remains between tip and substrate. The stark drop in repulsive force is due to the onset of Au-Au interaction. Note that the employed EAM potential[8] has a cutoff radius of 5.5 Å, corresponding roughly to the position of the minimum in repulsive force marked by panel F. Events at closer distances, Fig. 2G to H, correspond to the probe's deformation around a few surfactant molecules trapped within the contact area, visualised in the contact area's top view of Fig. 2H at $d = 0$. It is remarkable that most surfactant molecules have been "squeezed out" of the contact. We believe that at even slower approach rates the remaining molecules should also be able to leave the contact by surface diffusion. Compromising between computational costs and accuracy, we use $1\,\mathrm{ms}^{-1}$ as reference velocity for the simulations described in the following.

Fig. 3 shows the variation of force–distance curves with surface concentration during approach of the AFM tip for homogeneous films. With increasing surface number

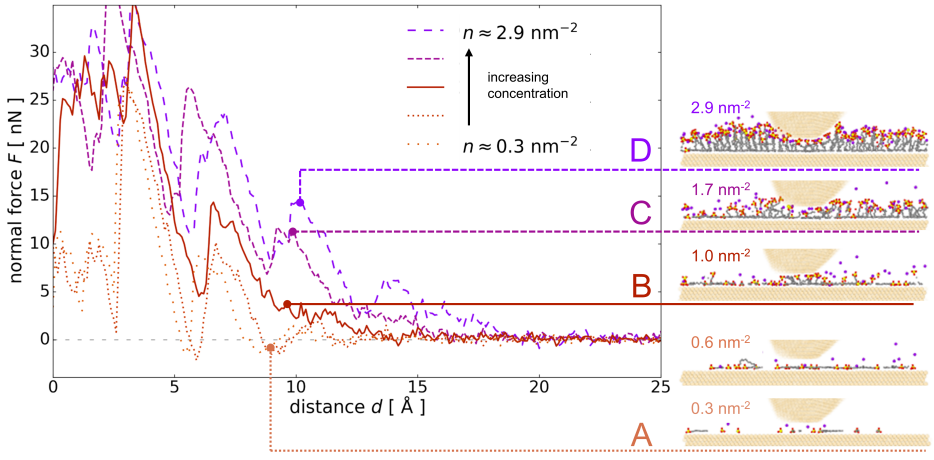


Figure 3. Distinguished AFM response features of SDS monolayers for increasing concentrations at 1 m s^{-1} approach velocity. Insets show snapshots around surface–surface distance $d \sim 10$ Å. At surface number concentrations as high as 2.9 nm^{-2}, the morphology probed here does not constitute a realistic system anymore: The monolayer will have undergone a phase transition to hemicylindrical aggregates, as discussed below in Fig. 4.

concentration n, the simulated AFM force–distance curves exhibit transitions in characteristic features. In Fig. 3, the onset of a repulsive force felt by the approaching tip moves from below 15 Å at $n < 1\ \mathrm{nm}^{-2}$ towards 20 Å with increasing concentration up to $n \sim 2.9\,\mathrm{nm}^{-2}$. Highlighting only the most apparent distinction between the different concentrations probed here, we attribute the attractive contribution for low coverage $n < 1\ \mathrm{nm}^{-2}$ (Fig. 3A) between surface–surface distances of 8 to 10 Å to the hydrophobic interaction between bare hydrocarbon tails lying flat at the substrate surface and the approaching probe model. Note the absence of any interaction between the tip's and the surface's gold atoms at this distance. At $n \sim 1\ \mathrm{nm}^{-2}$ (Fig. 3B), the surface is homogeneously covered with surfactants and some head groups protrude into solution. The tip has just made contact with the adsorption film and the attractive contribution at this position has vanished. For even higher concentrations $n > 1\ \mathrm{nm}^{-2}$ (Fig. 3C and D), the probe compresses a densely packed film and forces surfactant molecules to tilt, thereby requiring a rearrangement of their environment. Because of this, the feature that was attractive at low concentrations is repulsive at high concentrations.

In experiments, homogeneous monolayers such as shown in Fig. 3D are not observed at surface concentrations as high as $2.9\ \mathrm{nm}^{-2}$. Instead, the film will have completed its phase transition to hemicylindrical aggregates. Within the timescales probed in these calculations, both phases are (meta)stable when evolving from artificially preassembled initial configurations. Fig. 4 contrasts probing of these different morphologies under otherwise equivalent conditions, in particular under equal surface concentration. The curves exhibit distinct behaviour within the $d \in [10, 15]$ Å region. Cross-sectional views (Fig. 4 A1, B1)

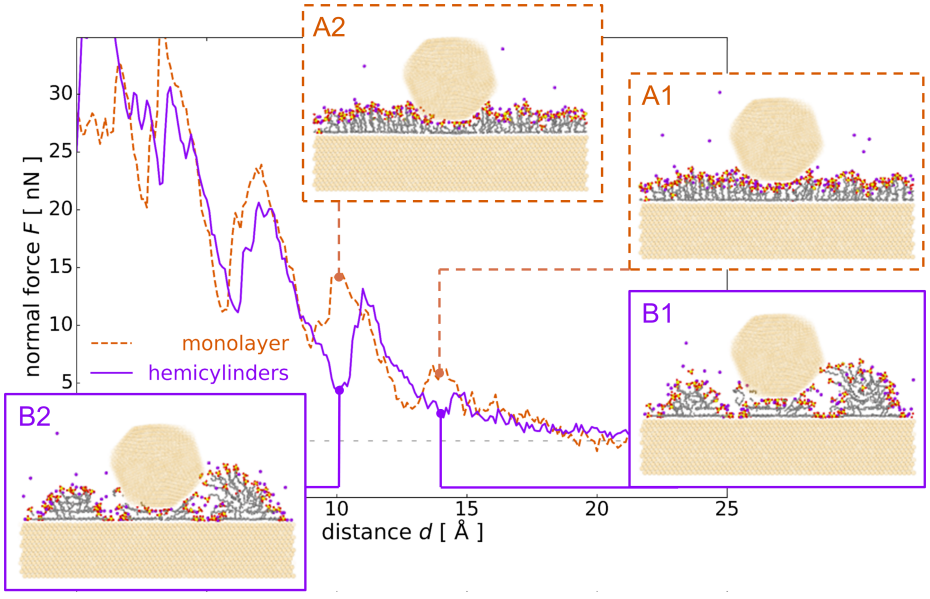


Figure 4. $1\ \mathrm{m\,s}^{-1}$ AFM probe approach above hypothetical monolayer and experimentally observed hemicylindrical aggregates of SDS at $2.9\,\mathrm{nm}^{-2}$. Insets show snapshot cross sections at gap widths $d \sim 14$ Å and $d \sim 10$ Å.

at $d \sim 14$ Å and (Fig. 4 A2, B2) at $d \sim 10$ Å reveal different ordering of the surfactant molecules that remain in the narrow gap between tip and surface: While surfactant molecules strive to stand upright on the surface within the artificially dense monolayer (Fig. 4 A1, A2), hemicylindrical configurations accommodate molecules that are aligned parallel to the surface (Fig. 4 B1, B2). Only head groups are in contact with the tip in the dense monolayer case. Contrastingly, the indented hemicylindrical aggregate allows tangentially aligned hydrocarbon tails to wet the probe surface at $d \sim 10$ Å (Fig. 4 B2). Constituents of neighbouring aggregates stretch to adhere to the unpassivated probe with their head groups in the case of striped aggregates (Fig. 4 B1, B2). The resulting force–distance curves have qualitatively different features, in particular a peak in the repulsive force at $d \sim 14$ Å for the monolayer (Fig. 4 A1) that is absent for the striped aggregates (Fig. 4 B1).

4 Summary & Conclusions

For the example of the anionic surfactant SDS, we have shown that MD-simulated AFM force–distance curves exhibit morphology-dependent characteristics. We systematically explored the impact of probe velocity and surfactant concentration on measured normal forces, showing that with changing surfactant concentration we expect qualitative changes in the force–distance curves. Subsequently, comparison between the experimentally observed hemicylindrical aggregates and purely hypothetical monolayers at high surface concentrations revealed distinct force response features. Thus, we have highlighted the impact of three different parametric dimensions, namely probe velocity, surfactant concentration, and film morphology on the manifestation of force–distance curves in MD simulations of AFM experiments. These calculations demonstrate MD to be a promising tool in understanding the nanoscopic mechanisms behind mechanical properties of surfactant adsorption films.

Acknowledgements

All simulations were carried out with LAMMPS.[13] For visualisation, we used OVITO.[14] VMD[15] and PACKMOL[16] were used for pre- and post-processing. FireWorks[17] assisted in work flow administration and data provenance tracking. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC). The authors acknowledge support by the state of Baden-Württemberg through the federative initiative *bwCloud SCOPE*.

References

1. I. Burgess, C. A. Jeffrey, X. Cai, G. Szymanski, Z. Galus, and J. Lipkowski, *Direct Visualization of the Potential-Controlled Transformation of Hemimicellar Aggregates of Dodecyl Sulfate into a Condensed Monolayer at the Au(111) Electrode Surface*, Langmuir **15**, 2607–2616, 1999.

2. J. Zhang and Y. Meng, *Boundary Lubrication by Adsorption Film*, Friction **3**, 115–147, 2015.
3. M. Chen, I. Burgess, and J. Lipkowski, *Potential Controlled Surface Aggregation of Surfactants at Electrode Surfaces - A Molecular View*, Surf. Sci. **603**, 1878–1891, 2009.
4. J. Zhang and Y. Meng, *Stick–Slip Friction of Stainless Steel in Sodium Dodecyl Sulfate Aqueous Solution in the Boundary Lubrication Regime*, Tribol. Lett. **56**, 543–552, 2014.
5. R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, *Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain X_1 and X_2 Dihedral Angles*, J. Chem. Theory Comput. **8**, 3257–3273, 2012.
6. H. Heinz, T. J. Lin, R. Kishore Mishra, and F. S. Emami, *Thermodynamically Consistent Force Fields for the Assembly of Inorganic, Organic, and Biological Nanostructures: The INTERFACE Force Field*, Langmuir **29**, 1754–1765, 2013.
7. X. Tang, P. H. Koenig, and R. G. Larson, *Molecular Dynamics Simulations of Sodium Dodecyl Sulfate Micelles in Water – The Effect of the Force Field*, J. Phys. Chem. B **118**, 3864–3880, 2014.
8. G. Grochola, S. P. Russo, and I. K. Snook, *On Fitting a Gold Embedded Atom Method Potential Using the Force Matching Method*, J. Chem. Phys. **123**, 204719, 2005.
9. M. Jaschke, H. Butt, H. E. Gaub, and S. Manne, *Surfactant Aggregates at a Metal Surface*, Langmuir **13**, 1381–1384, 1997.
10. R. D. Groot and P. B. Warren, *Dissipative Particle Dynamics: Bridging the Gap between Atomistic and Mesoscopic Simulation*, J. Chem. Phys. **107**, 4423–4435, 1997.
11. T. Soddemann, B. Dünweg, and K. Kremer, *Dissipative Particle Dynamics: A Useful Thermostat for Equilibrium and Nonequilibrium Molecular Dynamics Simulations*, Phys. Rev. E **68**, 46702, 2003.
12. M. A. González and J. L. F. Abascal, *The Shear Viscosity of Rigid Water Models*, J. Chem. Phys. **132**, 096101, 2010.
13. S. Plimpton, *Fast Parallel Algorithms for Short-Range Molecular Dynamics*, J. Comput. Phys. **117**, 1–19, 1995.
14. A. Stukowski, *Visualization and Analysis of Atomistic Simulation Data with OVITO–the Open Visualization Tool*, Model. Simul. Mater. Sc. **18**, 015012, 2009.
15. W. Humphrey, A. Dalke, and K. Schulten, *VMD: Visual Molecular Dynamics*, J. Mol. Graphics **14**, 33–38, 1996.
16. L. Martinez, R. Andrade, E. G. Birgin, and J. M. Martínez, *PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations*, J. Comput. Chem. **30**, 2157–2164, 2009.
17. A. Jain, S. P. Ong, W. Chen, B. Medasani, X. Qu, M. Kocher, M. Brafman, G. Petretto, G. M. Rignanese, G. Hautier, D. Gunter, and K. A. Persson, *FireWorks: a dynamic workflow system designed for highthroughput applications*, Concurrency Computat.: Pract. Exper. **27**, 5037–5059, 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), Karlsruhe, Germany

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

Many new exciting discoveries, such as gravitational waves showing mergers of black holes with each other, or the detection of the shadow of the black hole in the galaxy M87, or a new observational window at very low radio frequency, the European-wide LOFAR[3] telescope array, share the requirement of the use of supercomputers such as at Jülich Supercomuting Centre (JSC). This is nowadays in parallel to the use of supercomputers to simulate the cosmos, such as a merger of galaxy clusters,[4] including all non-thermal phenomena, such as the production of magnetic fields and ultra-high energy particles at energies far beyond the LHC at CERN.

1 Introduction

The main excitement in Astrophysics today is from observations, that would not even be possible without the use of supercomputers, such as the detection of the merger of black holes through their gravitational waves,[1] and the detection of the shadow of the monstrous black hole in the galaxy M87 through the data analysis of the interference fringes of mm-telescopes the world over.[2] A main example here is the low radio frequency telescope array LOFAR, with telescope stations across all of Europe, that uses the supercomputers at JSC. With LOFAR the sensitivity to weak non-thermal features is so strong, that we can actually observe the merger of black holes through their spatial emission patterns due to the spin-flip of the merging black holes. In contrast to that are the classical simulation calculations of exploding stars, of binary orbits in very dense clusters of stars, and here the merger calculation of galaxy clusters which may allow us to discern the effect of dark matter so clearly, that perhaps we can finally discern what dark matter is; *a fortiori* such calculations could perhaps begin to constrain the properties of dark energy on the larger scales, and point to the origin of ultra high energy particles.

2 Observations and Supercomputers

The European-wide low frequency radio array LOFAR needs supercomputers, such as at JSC to analyse their data, showing for instance radio galaxies during and after super-massive black hole mergers, as well as the mergers of clusters of galaxies and the formation of bridges between them. LOFAR detects active starbursts as well as normal galaxies to high redshift, and can so trace their evolution from small to large galaxies via merging. When black holes merge, both stellar mass and super-massive black holes, the final merged black hole can get a substantial kick in some cases, also visible in the observations. LOFAR

needs supercomputers in a first stage just to develop the key software, in a second stage to run a pipeline for the huge amounts of data, and in a third to manage the data archive. Observations by LOFAR may then lead to a better understanding of the origin of magnetic fields and ultra high energy particles. Another aspect is to constrain where in the universe life such as ours had a long peaceful time to develop.

3 Cosmological Simulations

The mergers of clusters of galaxies can be simulated and then be compared to observations, such as by LOFAR, as regards both the galaxies and the intergalactic medium with all its magnetic fields and energetic particle populations. This allows to constrain the properties and statistics of the mergers of galaxies with both super-massive black holes, as well as the mergers of stellar mass black holes, including their spin-flip generating huge sweeps through the sky by the relativistic jets. As soon as rotation and fluid flow combine we get a dynamo to generate magnetic fields; oblique shocks are just one example, accretion to the centre of a galaxy another. Energetic particle populations are produced, lose their energy in adiabatic expansion as well as via synchrotron and inverse Compton losses, and get revitalised through intergalactic shocks. Going to larger scales the constraints due to the dark matter distribution can be modelled and use to limit models of what dark matter could be.

4 Conclusions

Today we need all three applications of supercomputers, (i) the analysis of the very complex data, such as in interferometry (*e. g.* LOFAR), (ii) the simulation library of what the data might show (*e. g.* in mergers of black holes of various masses, spins, and relative spin orientations), and (iii) the theoretical simulations of scenarios such as merging clusters of galaxies including the generation of magnetic fields and ultra high energy particles.

References

1. B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs*, Phys. Rev. X **9**, 031040, 2019.
2. The Event Horizon Telescope Collaboration, *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, The Astrophysical Journal Letters, **875:L1**, 17pp, 2019.
3. A. Drabent, M. Hoeft, A. P. Mechev, J. B. R. Oonk, T. W. Shimwell, F. Sweijen, A. Danezi, C. Schrijvers, C. Manzano, O. Tsigenov, R.-J. Dettmar, M. Bruggen, and D. J. Schwarz, *Realising the LOFAR Two-Metre Sky Survey*, this volume.
4. D. Wittor, P. Domínguez-Fernández, F. Vazza, and M. Brüggen, *A Song of Shocks and Dynamo: Numerical Studies of a Galaxy Cluster Merger in the HIMAG Project*, this volume.

Realising the LOFAR Two-Metre Sky Survey

Alexander Drabent[1], Matthias Hoeft[1], Alex P. Mechev[2], J. B. Raymond Oonk[2,3,4], Timothy W. Shimwell[2,3], Frits Sweijen[2], Anatoli Danezi[4], Coen Schrijvers[4], Cristina Manzano[5], Oleg Tsigenov[5], Ralf-Jürgen Dettmar[6], Marcus Brüggen[7], and Dominik J. Schwarz[8]

[1] Thüringer Landessternwarte, Sternwarte 5, 07778 Tautenburg, Germany
E-mail: {alex, hoeft}@tls-tautenburg.de

[2] Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands

[3] Netherlands Institute for Radio Astronomy (ASTRON), Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

[4] SURFsara, P.O. Box 94613, 1090 GP Amsterdam, The Netherlands

[5] Jülich Supercomputing Centre, Institute for Advanced Simulation (IAS), Forschungszentrum Jülich, 52425 Jülich, Germany

[6] Ruhr-Universität Bochum, Astronomisches Institut, Universitätsstraße 150, 44780 Bochum, Germany

[7] Hamburger Sternwarte, Gojenbergweg 112, 21029 Hamburg, Germany

[8] Fakultät für Physik, Universität Bielefeld, Universitätsstraße 25, 33615 Bielefeld, Germany

The new generation of high-resolution broad-band radio telescopes, like the Low Frequency Array (LOFAR), produces, depending on the level of compression, between 1 to 10 TB of data per hour after correlation. Such a large amount of scientific data demand powerful computing resources and efficient data handling strategies to be mastered. The LOFAR Two-metre Sky Survey (LoTSS) is a Key Science Project (KSP) of the LOFAR telescope. It aims to map the entire northern hemisphere at unprecedented sensitivity and resolution. The survey consist of 3 168 pointings, requiring about 30 PBytes of storage space. As a member of the German Long Wavelength Consortium (GLOW) the Forschungszentrum Jülich (FZJ) stores in the Long Term Archive (LTA) about 50 % of all LoTSS observations conducted to date. In collaboration with SURFsara in Amsterdam we developed service tools that enables the KSP to process LOFAR data stored in the LTA at the Jülich Supercomputing Centre (JSC) in an automated and robust fashion. Through our system more than 500 out of 800 existing LoTSS observations have already been processed with the `prefactor` pipeline. This pipeline calibrates the direction-independent instrumental and ionospheric effects and furthermore reduces the data size significantly. For continuum imaging, this processing pipeline is the standard pipeline that is executed before more advanced processing and image reconstruction methods are applied.

1 Introduction

The **Low Frequency Array**[1, 2] (LOFAR) is a novel radio telescope. It has been developed, built, and is operated by the Netherlands Institute for Radio Astronomy (ASTRON). In addition, there is a strong contribution by institutes in other European countries. LOFAR consists of 38 stations in the Netherlands, and 13 stations outside the Netherlands, six of them in Germany, three in Poland, and one in each of France, UK, Sweden, and Ireland. The array is still expanding with a new station in Latvia and Italy.[3] All LOFAR partners together form the International LOFAR Telescope (ILT).

LOFAR is an interferometer observing in the poorly explored frequency range between 10 and 240 MHz with unprecedented sensitivity and spatial resolution in comparison to preceding telescopes operating at the low-frequency regime. It therefore opens up a new window to the Universe.

2 The LOFAR Two-Metre Sky Survey (LoTSS)

For LOFAR several Keys Science Projects (KSPs) have been defined. These large projects demand a large fraction of the available observing time to be granted. The LOFAR Two-metre Sky Survey (LoTSS)[4, 5] is a project of the LOFAR Surveys KSP (SKSP) and is observing the entire northern sky. At optimal declinations LoTSS produces 6″ images with sensitivities below 100 µJy beam^{-1} at a central frequency of 144 MHz, using a bandwidth of 48 MHz, see Fig. 1. To cover the entire northern sky 3 168 observations will be carried out, each of which requires 8 hours of observing time. This will eventually lead to an extraordinarily detailed map with 120 billion pixels and an expected amount of about 30 PByte of data to be stored in the archives.

The first data release was published in early 2019[6] and covers the HETDEX field, which has a size of 424 deg^2, *i. e.* 2 % of the entire survey, see Fig. 2. The source catalogue derived from the first data release contains 325 694 sources with 231 716 optical and/or infrared identifications.[7] A special issue of the scientific journal Astronomy & Astrophysics

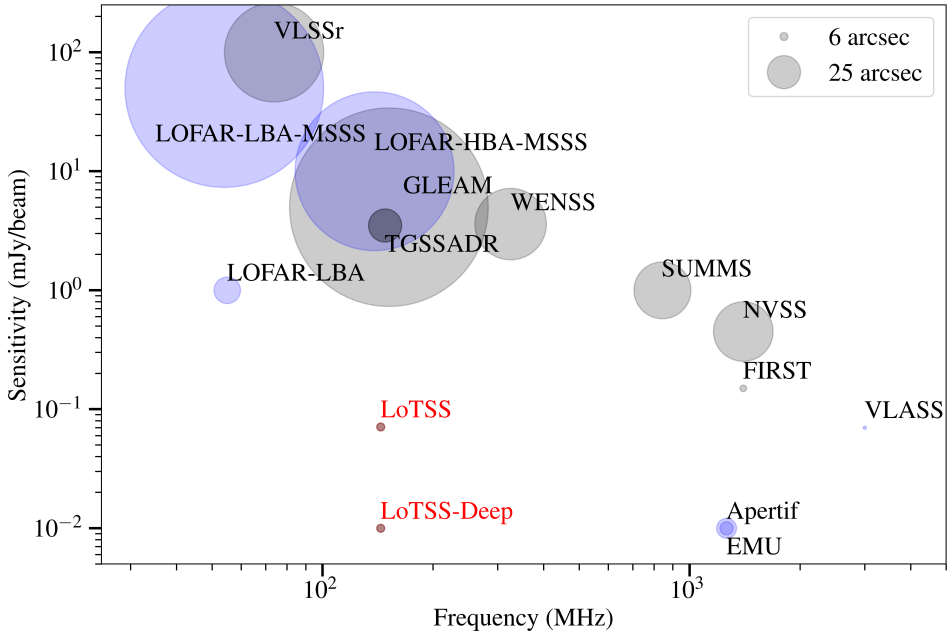


Figure 1. Sensitivity versus frequency of selected recent (grey), forthcoming (blue), and LOFAR (red) radio sky surveys. The size of the dot resembles the square root of the resolution of the survey. Image from T. W. Shimwell.[6]

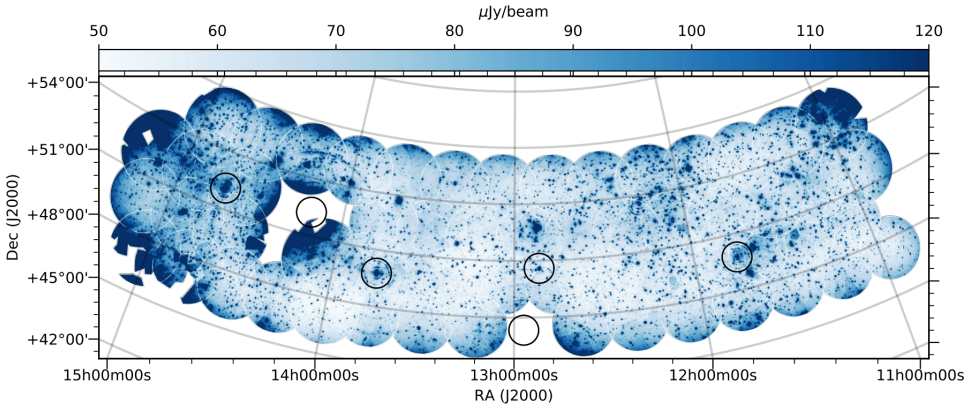


Figure 2. Noise image of the HETDEX field published in the first LoTSS data release. The median noise level is 71 μJy beam^{-1}. Black circles mark the locations of bright radio sources from the 3Cr catalogue and indicate problematic areas for the current calibration techniques. Image from Shimwell *et al.* (2019).[6]

(A&A) has been dedicated to the first twenty-six research papers describing the survey and its first results.[8] Discoveries made in the framework of the LoTSS cover a multitude of research topics, *e. g.* detailed studies of black holes in Active Galactic Nuclei (AGN), diffuse radio emission in clusters of galaxies, and cosmic-rays in nearby galaxies, to mention a few (see also Fig. 3). The second data release of LoTSS will offer the largest source catalogue at radio frequencies to date. The current status of the processing is shown in Fig. 4.

The high sensitivity, resolution and survey speed of LOFAR comes with a price: the

Figure 3. Total intensity radio maps from the LoTSS (in red colour scale) overlaid on top of an optical image. Credit: LOFAR Surveys Team. **Left panel:** The radio galaxy 3C 31. The observation reveals that the size of the galaxy is more than 3 million light years (Heesen *et al.*, 2018).[12] **Right panel:** The galaxy cluster Abell 1314. LOFAR reveals the presence of large-scale diffuse radio emission that traces highly-relativistic electrons in the intracluster medium embedded in a $\sim$μG strong cluster-wide magnetic field. It is believed that such emission was caused by a merger with another cluster. Thermal bremsstrahlung emission detected with the Chandra X-ray observatory is shown in grey (Wilber *et al.*, 2019).[13]

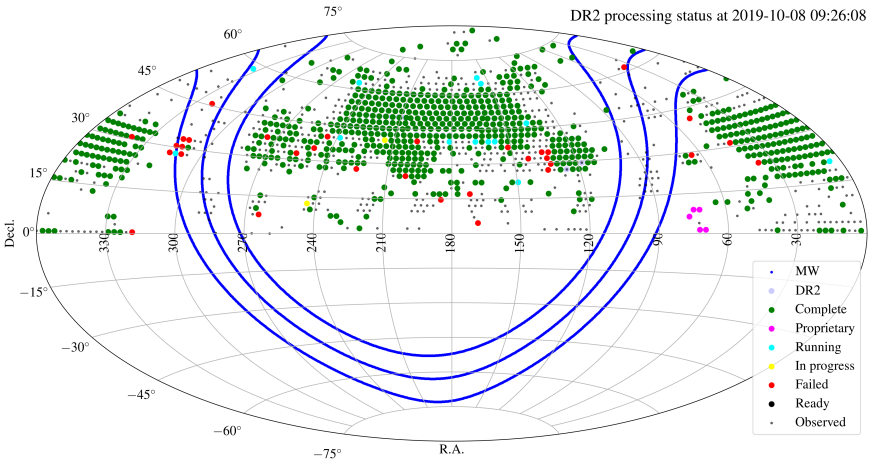


Figure 4. Status of the LoTSS as of 8th October 2019. Green dots mark pointings that are already observed and processed.[14]

large bandwidths, the large distance of stations and the large field of view require entirely novel calibration techniques. LOFAR operates at low frequencies and thus suffers severely from ionospheric disturbances, which can cause decorrelation of the signal and thus limits the sensitivity of the instrument. These challenges are met with non-standard and computational intense methods like advanced directional-dependent calibration techniques.[9–11] Moreover, averaging in time and frequency can only be performed moderately to permit recovering the sky brightness distribution over the entire field of view. This ensemble of large amounts of data and complex algorithms does not only demand big and robust data archives, but also an efficient and automated data handling in order to achieve high-fidelity radio maps within a reasonable amount of time and computational effort.

3 The LTA at JSC

The German institutes participating in LOFAR have formed the *German Long Wavelength Consortium*[15] (GLOW), which coordinates the German LOFAR activities. Besides the archival sides in Amsterdam and Poznań, the JSC operates a large part of the LOFAR LTA. About 50 % of all LoTSS observations are archived in the tape archive at JSC, *i. e.* 800 observations occupying about 12 PByte of storage space on tape. Such an amount of data needs to be processed in an automated, well reproducible and organised fashion within a reasonable amount of time. An efficient processing therefore demands computing facilities with a fast connection to the data storage in order to avoid unreasonably long transfer times of the data to local computing resources via the internet. Thus, the storing of a large amount of data and the processing of the data at JSC is a significant German contribution to the ILT. The data in the LOFAR LTAs are among the largest collections of astronomical data. Operating these LTAs allows facilities to explore strategies how to handle the vast amounts of scientific data generated by radio interferometers. In particular, the staging

and the retrieval of the data from tape to the disk pool is a time critical step for the whole processing.

4 The Automated Processing Scheme for JUWELS

In the framework of LoTSS the LOFAR Surveys Key Science Project (SKSP) developed an automatised job management and execution system for the grid-computing facilities at SURFsara in Amsterdam. These tools are described in Mechev *et al.* (2017)[16] and Mechev *et al.* (2018ab).[17, 18] Since the biggest fraction of LOFAR data observed in the framework of the LoTSS is stored in the LTA at JSC, an efficient and fast processing of the data is crucial. A regular 8 hour pointed observation using the full bandwidth of the telescope at a time resolution of 4 seconds and a frequency resolution of 48.82 kHz requires about ∼8-16 TB of disk space, depending on the level of compression. Transferring such an amount of data from the LTA at JSC to the computing facilities in Amsterdam via public network or internet is slow and thus not feasible.

The fast disk cache of the LTA is directly connected to the Data Access Server at JSC (JUDAC) and thus provides an order of magnitude higher throughput of data than data transfer via internet. Therefore, we extended the already established automated and coherent processing system at the SURFsara grid-computing facilities in Amsterdam to JUWELS. To achieve this goal we developed cluster-specific tools that allow us to retrieve data from the LTA at JSC and subsequently process them directly on JUWELS. These tools comprise of:

- **Software installation:** We share a pre-compiled LOFAR installation hosted in a `cvmfs`-directory at SURFsara. This allows us to perform regular updates via a simple synchronisation of the software. The software (parrot-connector) needed to be compiled in user space. Furthermore, cluster-specific environment parameters are adjusted in the retrieved software copy.

- **Job management and execution:** Service monitoring scripts, running continuously on JUDAC and the head node of JUWELS, act as an interface between the LTA, an external database, and the supercomputer JUWELS, respectively.

The external database uses Apache's `couchdb` and is operated by SURFsara. It serves as an "order book" (a so-called "pilot job framework") in order to define and manage all the planned jobs, see Fig. 5. All the jobs are organised in so-called "tokens" that are comprised of a job description (usually a pre-defined list of tasks or pipelines) and their corresponding parameters (*e. g.* location of input/output data, observation IDs, software version to be used, current processing status of the job). Thanks to such a central database, from the user's perspective it is not relevant where a LOFAR observation to be processed is actually located and on which or what kind of machine it will be processed. The user does not even need to have any knowledge about the structure of the archive or the supercomputing system to run a job. Moreover, the database allows the user to track the status of the current processing and even inspect all diagnostics and logfiles provided by already finished tasks. Thus, our system is a realisation of the "near-data processing" principle.

In Fig. 6 we present the structure of the processing pipeline. Thanks to the ability of JUDAC to access non-standard ports via HTTP and the availability of grid-tools like `voms`,

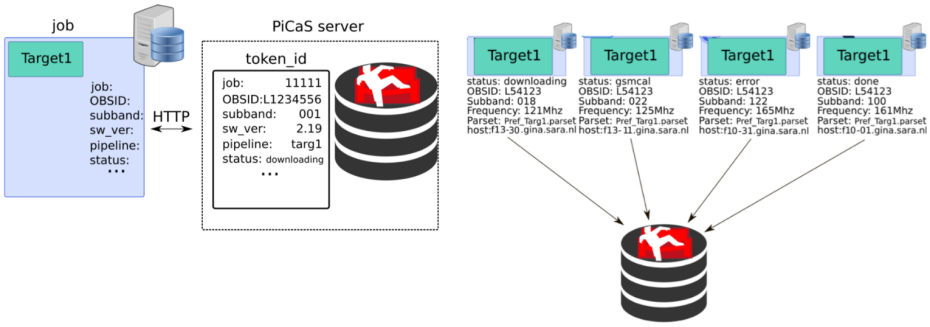


Figure 5. Sketch of the internal structure of the job management system developed by the SURFsara group in Amsterdam. Jobs are organised in tokens, comprising of the job description and their corresponding parameters to be used during the processing. All tokens are stored in a database, using Apache couchdb, which can be accessed via HTTP. Image by courtesy of A. Danezi.

`srmls`, or `globus-copy-url`, the monitoring service script running on JUDAC takes over the entire communication with the external database and the data retrieval with the LTA at JSC. It essentially incorporates the following services:

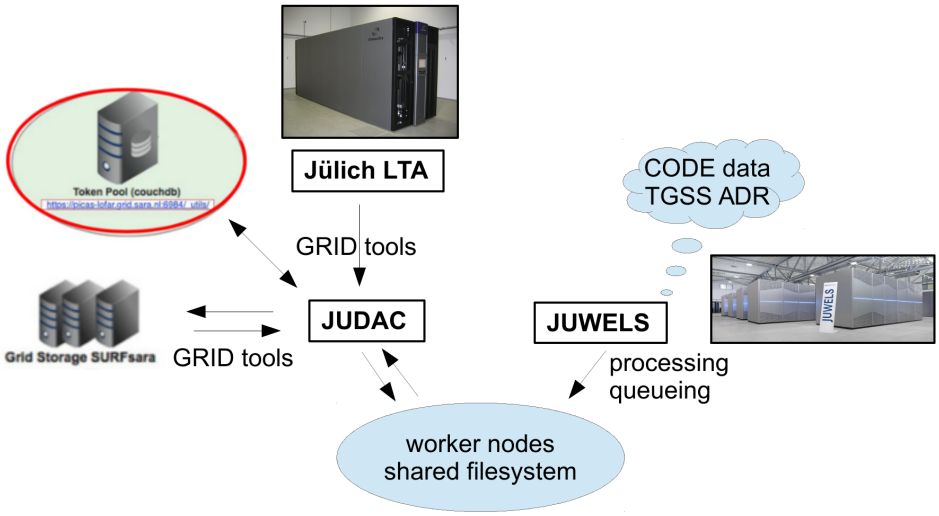


Figure 6. Sketch of the automated processing scheme implemented on JUDAC and JUWELS. Interactions that require access to the grid and the external database (here called as Token Pool) are realised on JUDAC. The head node of JUWELS prepares the tasks provided by JUDAC and submits it to the job queue of JUWELS. Before it needs to retrieve online data necessary for the calibration (here denoted as CODE data, a GPS satellite, and TGSS ADR, a skymodel database). Then it awaits execution. The communication between JUDAC and the JUWELS head node is done via the shared filesystem.

(1) interaction with the external database, *i. e.* looking for new tokens

(2) checking whether data of a new or active job are staged and transferring staged data from the LTA at JSC to the shared file system, which is accessible for both JUDAC and JUWELS

(3) preparing environment for the job and writing a so-called "submission script" to a dedicated location of the shared filesystem

(4) monitoring of submitted jobs via the logfile of an active job, and if applicable reporting back success or failure to the external database

(5) uploading diagnostic plots created during the calibration of the data and the logfile of the active job to the external database

(6) transferring the output data to a grid-storage at SURFsara in Amsterdam if job was run successfully

A second monitoring script is in place on the head node of JUWELS. It continuously checks for a "submission script" put in a dedicated location of the shared filesystem and executes it. Since the worker nodes of JUWELS do not have any internet access, data necessary for the calibration that is only available from online services have to be downloaded before the actual submission of the calibration task to the job queue. So the "submission script" contains:

(1) downloading of all necessary data needed for the particular calibration tasks, *e. g.* sky-models from an online repository

(2) submitting the job to the JUWELS queue

After submission the "submission script" is deleted and the job will be traced only by the monitoring script on JUDAC via the job ID returned from the JUWELS queueing system, in particular if in execution via the written logfile.

After processing the reduced data and all derived solution tables are transferred to a central grid-network storage hosted by SURFsara where they are accessible for all members of the consortium for further advanced processing. The code and its documentation are available on github.[22]

5 LOFAR Data Reduction at JSC: prefactor

The `prefactor`[20, 21] pipeline is a core component of the LOFAR software and performs the first direction-independent calibration, which is of fundamental importance for the success of all upcoming direction-dependent calibration and image reconstruction steps. An average LoTSS observation with a duration of 8 hours is bookended with calibrator observations with a duration of ten minutes. The processing of the data on JUWELS with `prefactor` is therefore split into two main parts:

(a) **calibrator pipeline:** extraction of instrumental effects from the calibration of the calibrator observation

(b) **target pipeline:** correcting for the instrumental effects (derived from the calibrator pipeline) in the actual target data, flagging, further averaging of the target data and further calibration off a model of the field

In case of (a) it derives corrections for instrumental effects like the polarisation alignment, the bandpass, and the clock offsets (mainly for remote stations). To achieve this the instrumental effects have to be separated from the ionospheric effects, *i. e.* phase signal delays and Faraday Rotation of the phase signals due to the time-dependent variation of the total electron content (TEC) at the location of the LOFAR station and its pointing direction. `prefactor` calibrates and corrects for all these effects in the order of their interference with the signal along its path from the sky to the telescope and even uses the full bandwidth if necessary to increase the signal-to-noise ratio.

In the target pipeline (b) the derived corrections from the calibrator pipeline are applied to the target data. These corrections allow for moderate averaging and thus reducing the data size without decorrelating the signal. Moreover, in the course of the averaging the international stations are discarded. In a second step an initial direction-independent phase-only self-calibration with a skymodel is performed to reduce still remaining residuals in the phase signal. The reduction in data size achieved is usually about a factor of 64 and thus allows for a much easier data transfer and handling in the following direction-dependent calibration steps. Future full-resolution surveys, making use of the international stations, will be more challenging, since they will not allow for such high levels of averaging.

`prefactor` offers many diagnosis tools to assess the quality of the calibration in a purposeful way. Dynamic spectra plots of the calibration solutions of each individual antenna for every calibration step allows the user to easily spot failures in the processing. Missing data (of the calibrator or target) or data of low quality are identified, *e. g.* through flagging of stations that are severely affected by Radio Frequency Interference (RFI), or large frequency regimes that are affected by broad-band RFI. This mitigates the need for repeating the processing with adapted parameters and will lead to a successful calibration run of a random observation in more than 90% of the cases.

6 Concluding Remarks

Efficient and robust processing of large amounts of scientific data is a key challenge for the future of radio interferometry. Large-scale surveys like LoTSS allow for deep insights into previously unveiled regimes of the radio sky. With about 30 PB of required storage space, LoTSS will be among the biggest scientific data collections in the world. We implemented a set of service scripts for JUWELS that have so far enabled us to retrieve and calibrate the data of 500 LoTSS observations stored in the LTA at JSC in an automated, robust and efficient manner with prefactor. Hereby, we make use of an already established framework hosted and maintained at SURFsara in Amsterdam, which was originally developed for processing LoTSS data at their grid-computing facilities. This allows an external user to manage the jobs for processing the data with a single interface independent on the used hardware or software infrastructure and is thus a realisation of the principle of "near-data processing". The ILT has identified the need to develop all three LOFAR LTAs, including the one operated by the JSC, into integrated LOFAR Science Data Centres, which would combine archiving and computing. The Science Data Centres would offer standard

pipelines to all LOFAR users and coordinate the pipeline development. The work done by the SKSP at SURFsara and JSC represents a first step towards achieving this goal.

Acknowledgements

AD acknowledges support by the BMBF Verbundforschung under the grant 05A17STA. JBRO acknowledges financial support from NWO Top LOFAR-CRRL project, project No. 614.001.351. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC). This work was carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative through grants e-infra 160022 & 160152. The authors would like to thank the staff at SURFsara for all assistance received during this project. LOFAR (van Haarlem *et al.* 2013) is the LOw Frequency ARray designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources) and are collectively operated by the ILT foundation under a joint scientific policy. The ILT resources have benefited from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, Forschungszentrum Jülich, MIWF-NRW, MPG, Germany; Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council (STFC), UK.

References

1. https://lofar.org
2. M. P. van Haarlem, M. W. Wise, A. W. Gunst, G. Heald, J. P. McKean, J. W. T. Hessels, A. G. de Bruyn *et al.*, *LOFAR: The LOw-Frequency ARray*, A&A **556**, A2, 2013.
3. https://www.astron.nl/lofar-crosses-alps-italy-joins
4. https://lofar-surveys.org
5. T. W. Shimwell, H. J. A. Röttgering, P. N. Best, W. L. Williams, T. J. Dijkema, F. de Gasperin, F., M. J. Hardcastle, G. H. Heald *et al.*, *The LOFAR Two-metre Sky Survey. I. Survey description and preliminary data release*, A&A **598**, A104, 2017.
6. T. W. Shimwell, C. Tasse, M. J. Hardcastle, A. P. Mechev, W. L. Williams, P. N. Best, H. J. A. Röttgering, J. R. Callingham, T. J. Dijkema, F. de Gasperin *et al.*, *The LOFAR Two-metre Sky Survey. II. First data release*, A&A **622**, A1, 2019.
7. W. L. Williams, M. J. Hardcastle, P. N. Best, J. Sabater, J. H. Croston, K. J. Duncan, T. W. Shimwell, H. J. A. Röttgering *et al.*, *The LOFAR Two-metre Sky Survey. III. First data release: Optical/infrared identifications and value-added catalogue*, A&A **622**, A2, 2019.
8. A&A special issue – *LOFAR Surveys: a new window on the Universe*, 2019, https://www.aanda.org/component/toc/?task=topic&id=920.
9. C. Tasse, *Applying Wirtinger derivatives to the radio interferometry calibration problem*, 2014, arXiv:1410.8706 [astro-ph.IM].

10. R. J. van Weeren, W. L. Williams, M. J. Hardcastle, T. W. Shimwell, D. A. Rafferty, J. Sabater, G. Heald, S. Sridhar, T. J. Dijkema, G. Brunetti *et al.*, *LOFAR Facet Calibration*, ApJS **233**, 2, 2016.
11. C. Tasse, B. Hugo, M. Mirmont, O. Smirnov, M. Atemkeng, L. Bester, M. J. Hardcastle, R. Lakhoo, S. Perkins, and T. W. Shimwell, *Faceting for direction-dependent spectral deconvolution*, A&A **611**, A87, 2018.
12. V. Heesen, J. H. Croston, R. Morganti, M. J. Hardcastle, A. J. Stewart, P. N. Best, J. W. Broderick, M. Brüggen, G. Brunetti, K. T. Chyży, J. J. Harwood *et al.*, *LOFAR reveals the giant: a low-frequency radio continuum study of the outflow in the nearby FR I radio galaxy 3C 31*, MNRAS **474**, 5049-5067, 2018.
13. A. Wilber, M. Brüggen, A. Bonafede, D. A. Rafferty, T. W. Shimwell, R. J. van Weeren, H. Akamatsu, A. Botteon, F. Savini, H. Intema, L. Heino, V. Cuciti *et al.*, *Evolutionary phases of merging clusters as seen by LOFAR*, A&A **622**, A25, 2019.
14. `https://www.lofar-surveys.org/status.html`
15. `https://www.glowconsortium.de`
16. A. P. Mechev, J. B. R. Oonk, A. Danezi, T. W. Shimwell, C. Schrijvers, H. Intema, A. Plaat, and H. J. A. Röttgering, *An Automated Scalable Framework for Distributing Radio Astronomy Processing Across Clusters and Clouds*, in Proceedings of the International Symposium on Grids and Clouds (ISGC) 2017, Taipei, Taiwan, 2017.
17. A. P. Mechev, J. B. R. Oonk, T. W. Shimwell, A. Plaat, H. T. Intema, and H. J. A. Röttgering, *Fast and Reproducible LOFAR Workflows with AGLOW*, 2018, arXiv:1808.10735 [astro-ph.IM].
18. A. P. Mechev, A. Plaat, J. B. R. Oonk, H. T. Intema, and H. J. A. Röttgering, *Pipeline Collector: Gathering performance data for distributed astronomical pipelines*, Astronomy and Computing **24**, 117–128, 2018.
19. `http://doc.grid.surfsara.nl`
20. `https://github.com/lofar-astron/prefactor`
21. F. de Gasperin, T. J. Dijkema, A. Drabent, M. Mevius, D. A. Rafferty, R. J. van Weeren, M. Brüggen, J. R. Callingham, K. L. Emig, G. Heald, H. T. Intema *et al.*, *Systematic effects in LOFAR data: A unified calibration strategy*, A&A **622**, A5, 2019.
22. `https://github.com/adrabent/JURECA_tools`

A Song of Shocks and Dynamo: Numerical Studies of a Galaxy Cluster Merger in the HIMAG Project

Denis Wittor[1,2,3], Paola Domínguez-Fernández[3], Franco Vazza[1,2,3], and Marcus Brüggen[3]

[1] Dip. di Fisica e Astronomia, Universitá di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy

[2] INAF, Istituto di Radioastronomia di Bologna, Via Gobetti 101, 41029 Bologna, Italy

[3] Hamburger Sternwarte, Gojenbergsweg 112, 21029 Hamburg, Germany
E-mail: dwittor@hs.uni-hamburg.de

With *ENZO* simulations run on the Jülich supercomputers, we have investigated the evolution of magnetic fields in the largest cosmic structures (namely galaxy clusters and filaments connecting them) with unprecedented dynamical range. These simulations revealed the full development of the small-scale dynamo in Eulerian cosmological magneto-hydrodynamical simulations. The turbulent motions developed during the formation of clusters are energetic enough to foster the growth of magnetic fields by several orders of magnitude, starting from weak magnetic fields up to strengths of $\sim \mu$G as observed. Furthermore, shock waves are launched during cluster formation and they are able to accelerate cosmic-ray electrons, that emit in the radio wavelengths. Radio observations of this emission provide information on the local magnetic field strength. We have incorporated, for the first time, the cooling of cosmic-ray electrons when modelling this emission. In this contribution, we present our advances in modelling these physical processes. Here, we mostly focus on the most interesting object in our sample of galaxy clusters, which shows the complexity of magnetic fields and the potential of existing and future multi-wavelengths observations in the study of the weakly collisional plasma on $\sim$ Megaparsecs scales.

1 Introduction

Galaxy clusters are the largest plasma laboratories in the Universe. They contain hundreds of galaxies that sit at the nodes of the cosmic web, which consists primarily of Dark Matter ($\sim$ 83 %), as well as a hot, ionised, weakly collisional plasma called the intracluster medium (ICM, $\sim$ 26 %). Note that stars only make up about a few percent of the total mass of galaxy clusters.

While the star light of the cluster is observable in the optical light, the ICM is not and one has to use X-ray and radio observations to shed light on the state and properties of this plasma. The ICM hosts a variety of thermal and non-thermal phenomena that are driven during the evolution of the cluster by the accretion of matter and via the merging with other clusters. These mergers drive both shock waves and turbulence in the ICM (*e. g.* Refs. 1, 2). Moreover, radio observations reveal the presence of diffuse radio emission on large-scales, giving evidence for the presence of large-scale magnetic fields and relativistic particles. Shock waves are expected to accelerate cosmic-ray electrons via diffusive shock acceleration and to produce radio relics, while turbulence is assumed to produce radio halos via turbulent re-acceleration of cosmic-ray electrons. Radio relics are large and elongated sources that are found on the cluster periphery. Moreover, relics have a high degree of polarisation. On the other hand, radio halos are diffuse sources located at the cluster centre and they do not show any sign of polarised emission (*e. g.* Ref. 3, for a recent review).

Many questions still surround the ICM. While both radio relics and halos are produced by cosmic-ray electrons, there has not been a signal, in form of γ-rays, of cosmic-ray protons that should be accelerated by the same mechanisms (*e. g.* Ref. 4). Moreover, it is still debated if the origin of the large-scale magnetic fields is primordial or astrophysical (*e. g.* Ref. 5). Since the characteristic time scales of the systems are large, $\sim \mathrm{Gyr}$, it is impossible to study their evolution using observations. Hence, realistic numerical simulations are required to connect theory and observations in a quantitative way.

Simulating the ICM poses some technical and computational challenges for which different numerical algorithms have been designed. Beside the needs of properly resolving the scales at which Dark Matter and baryonic matter form clusters, the study of non-thermal phenomena in the ICM and their related radio emission makes it mandatory to resolve in time and space the small-scale dynamo amplification of magnetic fields, triggered by turbulent motions (*e. g.* Ref. 6). Hence, the physical scales that need to be considered span several orders of magnitude, *i. e.* going from the Giga-parsec to the parsec[a] scale. When resolving these small scales, additional physical process related to galaxy formation (such as feedback from active galactic nuclei or radiative processes) must be included. Hence, these computational challenges can only be tackled by using high-performance computing.

In this contribution, we present our recent contributions to the study of galaxy clusters and magnetic fields within them using high-resolution simulations and supercomputing at the Forschungszentrum Jülich. First, will present our numerical methods and benchmark tests performed on JUWELS in Sec. 2. In Sec. 3, we present our cosmological simulation in greater detail and provide an overview over our scientific highlights. We give our concluding remarks in Sec. 4.

2 Numerical Methods & Hardware Considerations

2.1 Cosmological Magneto-Hydrodynamics with ENZO

Several numerical codes have been developed to simulate the cosmic web and its components. In this work, we used the cosmological grid code *ENZO*.[7] *ENZO* is an adaptive mesh refinement (AMR) code that solves the equations of magneto-hydrodynamics (MHD) in the cosmological framework. It is being developed as a collaborative effort of scientists at many universities and national laboratories. *ENZO* uses a particle-mesh N-body method (PM) to follow the dynamics of the collision-less Dark Matter component, and an adaptive mesh method for ideal fluid dynamics. The Dark Matter component is coupled to the baryonic matter (gas) via gravitational forces, calculated from the total mass distribution (Dark Matter + gas) by solving the Poisson equation with an FFT based approach. In its basic version, the gas component is described as a perfect fluid ($\gamma = 5/3$) and its dynamics are calculated by solving conservation equations of mass, energy and momentum over a computational mesh, using a Eulerian solver based on the Piecewise Parabolic Method (PPM) or the lower order Piecewise Linear Method (PLM). These schemes are higher-order extensions of Godunov's shock capturing method.

ENZO is parallelised by domain decomposition into rectangular sub-grids, including the top/root grid (which is the only level in a non-AMR run). Message passing paradigm is

[a]One parsec is 3.1×10^{16} metre.

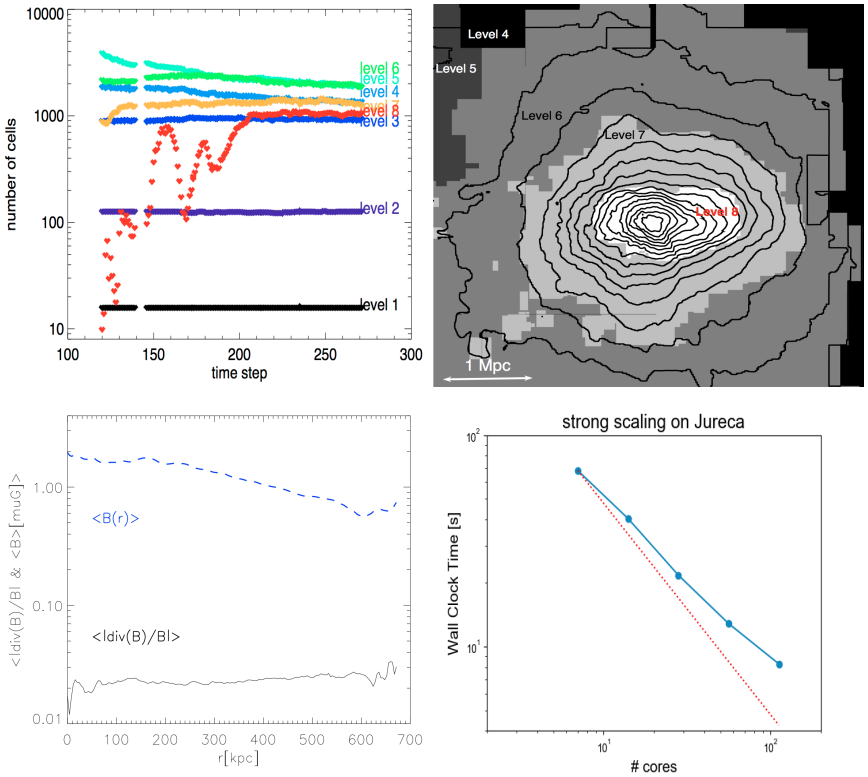


Figure 1. A few computing/numerical tests on our AMR strategy. Top left: progression in the number of cells at different AMR levels produced by a typical 8-levels *ENZO* cluster simulation on JUWELS. Top right: thin slice giving the distribution of AMR levels through the centre of a simulated cluster. Bottom left: radial profile of magnetic field and residual divergence of magnetic field in a simulated cluster at $z = 0$. Bottom right: strong scaling test on JURECA.

adopted and implemented by means of the MPI library, while I/O makes use of the HDF5 data format.

The MHD extension used in our runs is based on the conservative Dedner formulation of MHD equations,[8] which uses hyperbolic divergence cleaning to preserve the divergence of the magnetic field as small as possible. The MHD solver adopts the PLM reconstruction, fluxes at cell interfaces are calculated using the local Lax-Friedrichs Riemann solver (LLF) and time integration is performed using the total variation diminishing (TVD) second-order Runge-Kutta (RK) scheme. The resulting solver is somewhat more diffusive than the PPM approach, but allows a more efficient treatment of the electromagnetic terms.

2.1.1 Adaptive Mesh Refinement Strategy

The mesh refinement approach in *ENZO* adopts an adaptive hierarchy of grids at varying levels of resolution. *ENZO*'s implementation of structured AMR poses no fundamental

restrictions on the number of grids at a given refinement level or on the number of refinement levels. Moreover, it also allows arbitrary integer ratios of parent and child grid resolution. The patches are organised through a Hierarchical Tree structure, responsible for the management of adding or removing patches, specifying the neighbouring boxes at the same refinement level, the parent of each box, *etc.* In *ENZO*, no parallelisation is implemented on the Hierarchical Tree and the hierarchy data structure has to be replicated on each processor. This often yields a large memory usage overhead for our production runs.

In order to cope with the memory bounds on computing nodes on JURECA and JUWELS, we imposed a maximum AMR level of 8, considering a fixed refinement ratio of 2, which gave a $2^8 = 256$ refinement of the initial mesh resolution, *i. e.* from $\Delta x_0 = 1014\,\mathrm{kpc}$ to $\Delta x_8 = 3.9\,\mathrm{kpc}$. It shall be noticed that obtaining the highest possible spatial resolution when dealing with small-scale dynamo of magnetic fields is mandatory, as a coarse spatial resolution limits the effective Reynolds number of the simulated flows, and hence can severely hamper the development of the small-scale dynamo (*e. g.* see discussion in Ref. 6).

In order to maximise the dynamical range for each simulated object, in HIMAG/2 we started from two levels of static uniform grids with 256^3 cells each and using 256^3 particles each to sample the Dark Matter distribution, with a mass resolution per particle of $m_{\mathrm{DM}} = 1.3 \cdot 10^{10} M_{\odot}$ at the highest level. Then, we further refined the innermost $\sim (25\ \mathrm{Mpc})^3$ volume, where each cluster forms, with additional 7 AMR levels (refinement $= 2^7$). The refinement was initiated wherever the gas density was $\geq 1\%$ higher than its surroundings. This ensures that most of the central volume of each simulated clusters (where the small-scale dynamo is expected to develop while gas and dark matter continue to be assembled via mergers and accretions) is resolved up to the maximum possible resolution, in a rather uniform way. Fig. 1 shows the increase in the number of levels as a function of time in a typical cluster simulation in HIMAG/2, as well as the map of AMR levels in a cluster at the end of the simulation: the innermost $\approx 2^3\ \mathrm{Mpc}^3$ is typically resolved down to $\Delta x_8 = 3.9$ kpc, while the entire cluster volume is refined from $\Delta x_7 = 7.8$ kpc to $\Delta x_8 = 15.6$ kpc. With this approach, our AMR strategy was able to limit the spurious generation of the divergence of magnetic fields to a few percent in the entire volume of simulated galaxy clusters, as shown in the bottom left panel of Fig. 1.

2.1.2 Optimisation for Jülich Supercomputers

JURECA and JUWELS are well-suited for this project, given that the AMR implementation in *ENZO* requires a relatively small number of cores but a large memory per node. The bottom right panel of Fig. 1 shows an example of a strong scalability test for a 256^3 root grid, with AMR triggered by overdensity refinement for 3 AMR levels in the innermost 64^3 region, using from 7 to 112 cores. The behaviour of the code after 200 timesteps is very stable, and the average wallclock time per timestep is a factor ~ 2 above the theoretical scaling for the 112 cores case.

In full production, we used up to ~ 680 cores on 64 nodes to cope with the memory bounds posed by our runs, which can require up to ~ 1.8 Tb of memory towards the end of the simulation, due to the large number of cells generated by the AMR scheme and by the data stored in the hierarchy tree. Each cluster of our sample was simulated independently, using $\sim 30\,000$–$50\,000$ core hours on JURECA (HIMAG) and JUWELS (HIMAG2).

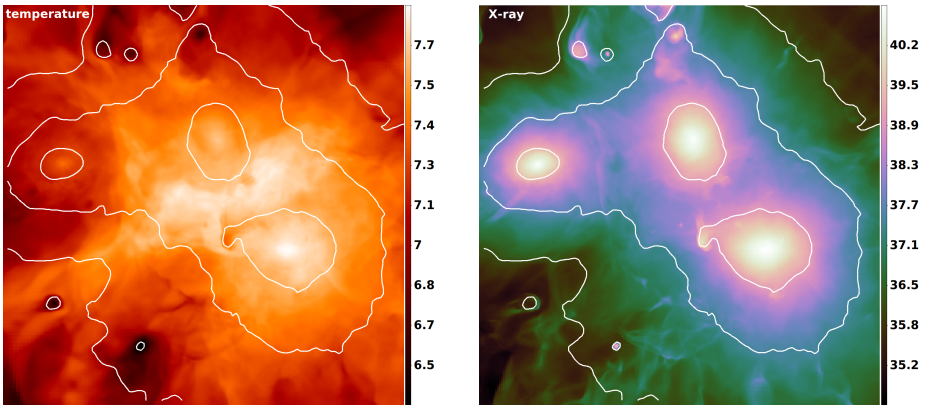


Figure 2. Projected gas temperature (left), in K, and X-ray luminosity (right), in erg/s, of the simulated galaxy cluster E5A. Both maps are given in log-scaling and they are overlayed with the baryonic density contours.

3 Scientific Results

Using the above strategies, we simulated a set of galaxy clusters within the HIMAG/2 projects. The main results were presented in several papers.[9–15] Here, we want to highlight some of our results by focusing on one peculiar cluster in our sample (identified as "E5A") which displays a number of remarkable properties.[b] Fig. 2 shows the projected gas temperature and the projected X-ray luminosity of E5A at the epoch of $z \approx 0.1$.

For the simulation of E5A, as well as for all other simulations, we assumed the same ΛCDM cosmological model, with $h = 0.72$, $\Omega_{\rm M} = 0.258$, $\Omega_{\rm b} = 0.0441$ and $\Omega_\Lambda = 0.742$ (see *e. g.* Ref. 10). A public repository of cluster snapshots from this catalog is available.[c]

3.1 The Evolution of Magnetic Fields in an Interesting Cluster Pair

In Ref. 9, we have presented evidence for *resolved* dynamo growth of intracluster magnetic fields in a Coma-like galaxy clusters from our sample. The unprecedented dynamical range achieved in the innermost cluster regions provided evidence of a small-scale dynamo and local amplification of magnetic fields up to values similar to what is found in observations ($\sim \mu G$), approaching energy equipartition with the kinetic energy flow on ≤ 100 kpc scales. We could constrain an overall efficiency of order ~ 4 % in the transfer between turbulent kinetic energy (in the solenoidal component) and the magnetic energy field.

Furthermore, the topology of the magnetic fields seem to be consistent with observations of the Coma cluster (*e. g.* Ref. 16 and references therein). In particular, the Faraday Rotation of background polarised sources is in good statistic agreement with the observations of real sources in the background of Coma, even if our simulation produce significant non-Gaussian tails in the distribution of magnetic fields components, which stems

[b] See `https://vimeo.com/266941122` for a movie showing the evolution of gas density and magnetic fields for this system.
[c] `https://cosmosimfrazza.myfreesites.net/amr_clusters`

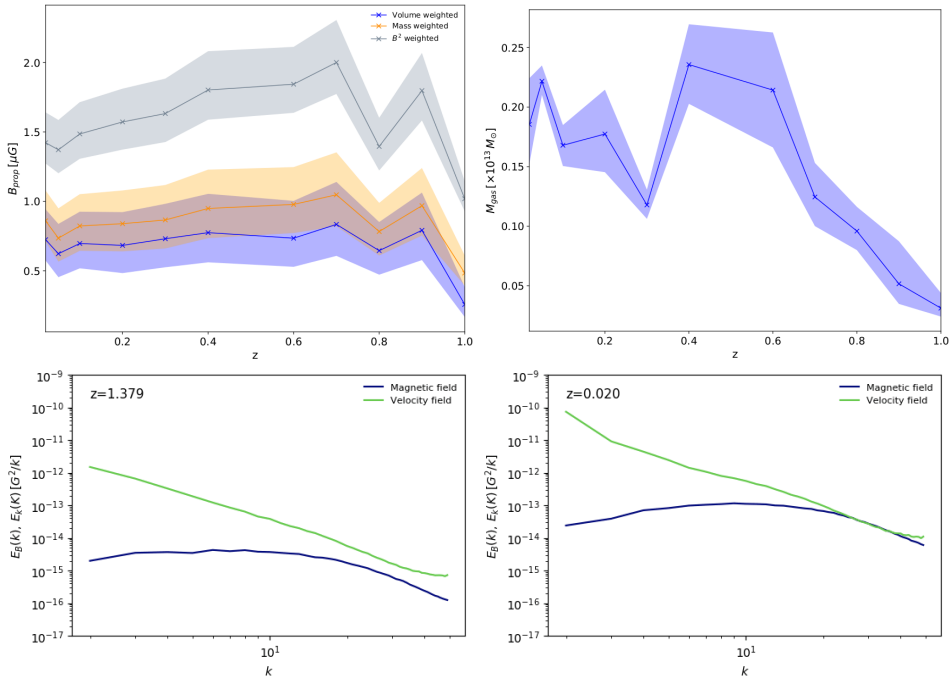


Figure 3. Top panels: evolution of the physical magnetic field strength (left) and gas mass (right) of E5A with redshift. The magnetic field is shown with a volume, mass and B^2 weight. The shadowed areas in both panels correspond to scatter of values within a radius of 1 Mpc at each redshift. Bottom panels: E5A magnetic (purple) and kinetic (green) spectra evolution at different redshifts.

from the superposition of different amplification patches which mix in the ICM. This result has an important implication since a Gaussian distribution of magnetic field components is often assumed in the interpretation of the Faraday Rotation (*e. g.* Ref. 16).

In Ref. 10, we generalised this analysis to a larger set of clusters and snapshots, produced in HIMAG and HIMAG2. There we studied how the spectral properties of magnetic fields are affected by mergers, and we could relate the measured magnetic energy spectra to the dynamical evolution of the ICM. Our results show that the magnetic growth rate is larger for merging systems and that relaxed systems generally have a larger outer scale. Since this is the scale that characterises the magnetic spectrum, larger values of the outer scale suggest that a dynamo has acted for a longer time in such systems. In Ref. 10, we also studied in more detail the evolution of E5A. As an example, we show in the top panels of Fig. 3 how the physical magnetic field and the gas mass of the cluster are changing with redshift (analysis made in a Mpc^3 cube moving with the cluster centre of mass). The magnetic energy of E5A grows by a factor of $\sim 40 - 50$ in a time-span of ~ 9 Gyr and equipartition between kinetic and magnetic energy occurs on a range of scales (< 160 kpc at all epochs), depending on the turbulence state of the system. Several minor-merger and major-merger events affected the magnetic properties of E5A cluster and shaped its magnetic spectrum in a complicated manner (see bottom panels of Fig. 3). During mergers, the

peak in the magnetic spectra shifts to *smaller spatial scales* and thus the magnetic spectrum broadens. A major-merger event affects the magnetic amplification of the cluster by delaying it for $\lesssim 1$ Gyr, in contrast, continuous minor mergers seem to promote the steady growth of the magnetic field.

3.2 Magnetic Fields Illuminated by Cosmic Shocks

In Refs. 15 and 14, we used two clusters of the HIMAG/2 sample to study the properties of radio relics. To compute their radio relic emission in cosmological simulations, one has to follow a simple recipe: First, one has to find shock waves in the simulation. Therefore, we used *Velocity-Jump* method that searches for jumps in the three-dimensional velocity field to detect shock waves.[12] Second, one has to assign the radio emission to the detected shock. Cosmic-ray electrons are expected to have short live times ($\leq 10\,\mathrm{Myr}$) and, hence, they loose most of their energy within $100\,\mathrm{kpc}$ of their injection site. Moreover, their energy budget seems to be dynamically unimportant. Hence, one can assume a quasi-stationary balance between the shock acceleration and radiative losses and, compute the cosmic-ray electron energy budget and associated radio power using the information of a single timestep.[17]

With the increasing resolution of cosmological simulations, the assumption of a quasi-stationary balance becomes invalid as the resolution drops below the electron cooling length. Hence, one has to properly resolve the downstream profile of the radio emission. In Ref. 15, we have developed an algorithm that takes the cooling of electrons into account and attaches a radio emission profile to each shocked cell. In addition to it, we have implemented an algorithm, following Ref. 18, that computes the polarised emission of radio relics. For details on the implementation we point to Ref. 15. In the left panel of Fig. 4, we give an example of the polarisation fraction of E5A. Using the new model, we have produced the most detailed simulation to date of radio polarisation in cluster shocks in Refs. 15 and 14. In Ref. 15, we found that **the observed polarised emission from radio relics should strongly depend on the environment**, and that the direction of polarisation and its variation across the relic depends on the properties of the upstream gas. Laminar gas flows in the upstream result in a parallel alignment of polarisation vectors, while disturbances in the upstream will cause a random orientation. These observations might reflect the local correlation length of the magnetic field. Furthermore in Ref. 14, we combined VLA observation of the western relic in RXCJ1314.4-2515 and cosmological simulations of a similar relic to study the history of the radio relic. We found that the relic power, measured at 3 GHz, can only be explained by the re-acceleration of an fossil population of cosmic-ray electrons.

3.3 Intracluster Bridges

Massive binary mergers as the one in E5A are rare and powerful events, in which a large amount of kinetic energy is concentrated between the two main clusters, and gas matter gets strongly compressed in the region connecting them. This collapse is rather fast ($\leq \mathrm{Gyr}$), leading to transonic turbulence and weak shocks.[12] This process leads to the compression and **"boosting" of the X-ray emission of the Warm Hot Intergalactic Medium (WHIM)** in the interacting regions, which is otherwise invisible.

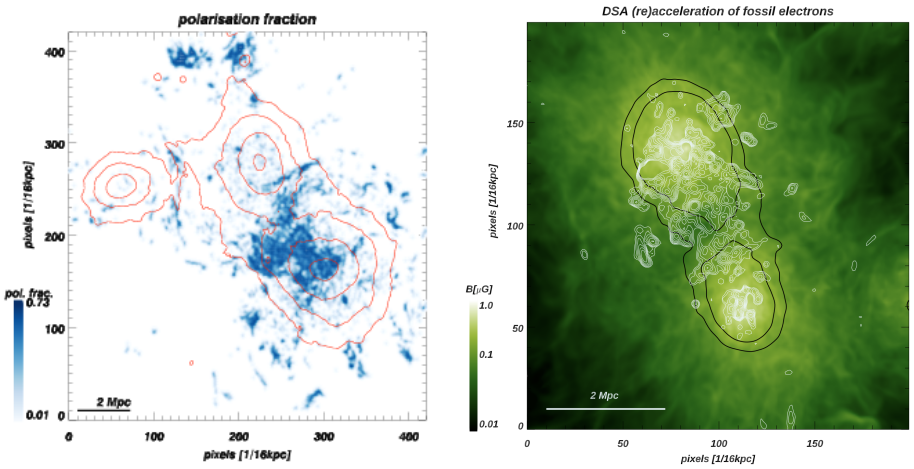


Figure 4. Left: Density contours and polarisation fraction (colours) of E5A. Right: Projected magnetic fields (colog), radio emission at 200 MHz (white contours) and projected Sunyaev-Zeldovich decrement (black contours) for the same object.

We produced synthetic radio and X-ray observations of E5A, tailored to reproduce existing (LOFAR, MWA and XMM) or future instruments (SKA-LOW and SKA-MID, ATHENA and eROSITA). Our simulations show that with sufficiently long ($\sim 10^2$ rm ks) exposures in the soft ($0.3 - 1.2$ keV) X-ray band on candidates similar to E5A it should be possible to **measure important thermodynamic properties of intracluster bridges, providing complementary constraints to radio data.** For example, with X-IFU it should be possible to derive the Mach number of radio shocks entirely from spectroscopically-derived information of the local gas velocity dispersion and of the local sound speed, in a temperature regime which is difficult to find in galaxy clusters.[12]

As a follow up, in Ref. 13 we mock-observed E5A with a LOFAR-HBA configuration, and compared it to real observations of the A399-A401 system of galaxy clusters. Ref. 13 reported the LOFAR-HBA detection of spectacular ridge of radio emission extending several Megaparsecs between A399 and A401. Using E5A as a testbed for particle acceleration models, we studied under which circumstances the intergalactic magnetic fields between the two clusters could become bright enough at radio wavelengths to match the real LOFAR observation (see right panel of Fig. 4). We reported that standard Diffusive Shock Acceleration alone cannot provide enough acceleration of cosmic-ray electrons here to explain the observed radio emission, and thus that some more volume filling (and/or) efficient acceleration mechanisms should operate in the intergalactic medium. This implies that systems like E5A, although rare in the Universe, can be more radio luminous than expected, while it remains challenging to be detected at other wavelengths.

4 Concluding Remarks

Achieving a large dynamical range in the numerical description of evolving magnetic fields in large-scale structures is crucial to allow the quantitative modelling of complex radio

data. Thanks to the allotted HIMAG/2 project on JURECA and JUWELS, our group produced the most resolved MHD simulation to date of magnetic fields in galaxy clusters with *ENZO*,[7] focusing with unprecedented detail on the evolution of intracluser magnetic fields under the action of the small-scale turbulent dynamo.[9, 10]

Among the investigated systems, the simulated merger in E5A has a peculiar evolutionary track, owing to its complex cycle of shock waves, turbulence injection and dynamo amplification episodes. As discussed in this contribution, E5A has thus highlighted a number of very interesting and little explored features of magnetic fields in cosmology: from the sequence of complex amplification cycles following mass accretions[10] to the powering of polarised radio emitting shocks (following Ref. 15, but not analysing E5A), and from the association of detectable radio and X-ray features in intracluser bridges[12] to the hints for more efficient particle acceleration mechanisms in such regions.[13]

Despite the improvements in numerical study of galaxy clusters presented in this work, there are still several physical processes, such as AGN feedback and galaxy formation, that have to be incorporated, making the simulations more complex and computational challenging. Thanks to the availability of high-performance computers, such as JUWELS, it will be possible to overcome these challenges and to make new exciting scientific discoveries in the future.

Acknowledgements

The cosmological simulations described in this work were performed using the *ENZO* code (`http://enzo-project.org`). The authors acknowledge the usage of computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC), under projects hhh42, hhh44 and stressicm. D. W., F. V. and P. D. F. acknowledge financial support from the European Union's Horizon 2020 program under the ERC Starting Grant "MAGCOW", no. 714196. We also acknowledge the usage of online storage tools kindly provided by the Inaf Astronomica Archive (IA2) initiative (`http://www.ia2.inaf.it`).

References

1. D. Ryu, H. Kang, E. Hallman, and T. W. Jones, The Astrophysical Journal **593**, 599–610, 2003.
2. F. Miniati, *The Matryoshka Run: A Eulerian Refinement Strategy to Study the Statistics of Turbulence in Virialized Cosmic Structures*, The Astrophysical Journal **782**, 21, 2014.
3. R. J. van Weeren, F. de Gasperin, H. Akamatsu, M. Brüggen, L. Feretti, H. Kang, A. Stroe, and F. Zandanel, *Diffuse Radio Emission from Galaxy Clusters*, Space & Science Review **215**, 16, 2019.
4. D. Wittor, F. Vazza, and M. Brüggen, *Testing cosmic ray acceleration with radio relics: a high-resolution study using MHD and tracers*, Monthly Notices of the Royal Astronomical Society **464**, 4448–4462, 2017.
5. F. Vazza, M. Brüggen, C. Gheller, S. Hackstein, D. Wittor, and P. M. Hinz, *Simulations of extragalactic magnetic fields and of their observables*, Class. Quantum Grav. **34**, 234001, 2017.

6. J. Donnert, F. Vazza, M. Brüggen, and J. ZuHone, *Magnetic Field Amplification in Galaxy Clusters and its Simulation*, Space & Science Review **214**, 122, 2018.
7. The Enzo Collaboration: G. L. Bryan *et al.*, *Enzo: An Adaptive Mesh Refinement Code for Astrophysics*, The Astrophysical Journal **211**, 19, 2014.
8. A. Dedner, F. Kemm, D. Kröner, C.-D. Munz, T. Schnitzer, and M. Wesenberg, *Hyperbolic Divergence Cleaning for the MHD Equations*, Journal of Computational Physics **175**, 645–673, 2002.
9. F. Vazza, G. Brunetti, M. Brüggen, and A. Bonafede, *Resolved magnetic dynamo action in the simulated intracluster medium*, Monthly Notices of the Royal Astronomical Society **474**, 1672–1687, 2018.
10. P. Domínguez-Fernández, F. Vazza, M. Brüggen, and G. Brunetti, *Dynamical evolution of magnetic fields in the intracluster medium*, Monthly Notices of the Royal Astronomical Society **486**, 623–638, 2019.
11. S. Hackstein, F. Vazza, M. Brüggen, J. G. Sorce, and S. Gottlöber, *Simulations of ultra-high energy cosmic rays in the local Universe and the origin of cosmic magnetic fields*, Monthly Notices of the Royal Astronomical Society **475**, 2519–2529, 2018.
12. F. Vazza, S. Ettori, M. Roncarelli, M. Angelinelli, M. Brüggen, and C. Gheller, *Detecting shocked intergalactic gas with X-ray and radio observations*, Astronomy & Astrophysics **627**, A5, 2019.
13. F. Govoni, E. Orrù, A. Bonafede, M. Iacobelli, R. Paladino, F. Vazza, M. Murgia, V. Vacca, G. Giovannini, and L. Feretti, *A radio ridge connecting two galaxy clusters in a filament of the cosmic web*, Science **364**, 981–984, 2019.
14. C. Stuardi, A. Bonafede, D. Wittor, F. Vazza, A. Botteon, N. Locatelli, D. Dallacasa, N. Golovich, M. Hoeft, R. J. van Weeren, M. Brüggen, and F. de Gasperin, *Particle re-acceleration and Faraday-complex structures in the RXC J1314.4-2515 galaxy cluster*, Monthly Notices of the Royal Astronomical Society **489**, 3905–3926, 2019.
15. D. Wittor, M. Hoeft, F. Vazza, M Brüggen, and P Domínguez-Fernández, *Polarisation of Radio Relics in Galaxy Clusters*, Monthly Notices of the Royal Astronomical Society **490**, 3987–4006, 2019.
16. A. Bonafede, F. Vazza, M. Brüggen, M. Murgia, F. Govoni, L. Feretti, G. Giovannini, and G. Ogrean, *Measurements and simulation of Faraday rotation across the Coma radio relic*, Monthly Notices of the Royal Astronomical Society **433**, 3208–3226, 2013.
17. M. Hoeft and M. Brüggen, *Radio signature of cosmological structure formation shocks*, Monthly Notices of the Royal Astronomical Society **375**, 77–91, 2007.
18. B. J. Burn, *On the depolarization of discrete radio sources by Faraday dispersion*, Monthly Notices of the Royal Astronomical Society **133**, 67, 1966.

Computational Biology and Biophysics

Computational Biology and Biophysics: Meeting the Challenges of Biomolecular Structure Formation and Function with High Performance Computing

Yu Goldscheider, Wolfgang Wenzel

Institute of Nanotechnology, Karlsruhe Institute of Technology,
Kaiserstr. 12, 76021 Karlsruhe, Germany
E-mail: wolfgang.wenzel@kit.edu

Biomolecular structure prediction and biomolecular function are among the main research challenges in biophysics as it remains very important to fully characterise the structure of various biological entities in the context of their functional role. Most biomolecules do not work in isolation but as part of complexes and structure prediction of those constitutes an even more difficult problem. As these functional roles are often difficult to characterise on the experimental side, simulation methods, which have reached an increasing degree of maturity to predict biomolecular structure and function, increasingly complement experiment.

The highly dynamic nature of multi-domain protein complexes necessitates the development of analytical approaches for the interpretation of the available experimental data in terms of representative ensembles. As these highly dynamic assemblies represent a significant fraction of all proteomes, deciphering the physical basis of their interactions represents a considerable challenge.[1] An example of such a study is the work by Berg and Peter in this volume, who present an approach to analyse long time-scale simulations to characterise conformational ensembles of flexibly-linked multidomain proteins at the example of covalently conjugated ubiquitin chains. With this multiscale simulation and analysis approach, it is possible to identify characteristic properties of ubiquitin chains in solution, which have been subsequently correlated with experimentally observed linkage- and chain length-specific behaviour. Simulations of these systems are complicated by the large conformational space that these proteins occupy, 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 research challenge addressed using modern day HPC resources is to elucidate the alignment and interaction of biomolecules with other biological components. Solvent interactions can influence the properties and function of complex biomolecules and solvent mixtures have been extensively used to unravel macromolecular and recognition events such as protein binding hotspots identification, ligand binding, protein folding, and the elucidation of biological mechanisms.[2, 3] The effects of the solvent on stearic acid – Human serum albumin (HAS) binding are explored in the work of Pant *et al.* in this volume. Their investigation shows that altering solvent composition can be a useful tool for modulating biomolecular function, in this case the binding specificity of human serum albumin as a

carrier of fatty acids. Their findings have special relevance for the regulation of biological function since they show that the targeted manipulation of protein – ligand binding events can be achieved by rationally tuned solvent-driven interactions. This work 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.

In order to understand fully the molecular basis of interactions between physiological partners it is necessary to map the modulation of the free energy landscapes of the interacting molecules throughout the interaction trajectory. Simulations were tied closely to the availability of structural data[4] to eventually fill the remaining gaps in understanding, which experimental methods could not yet resolve.[5]

These two studies serve as excellent examples to exploit the availability of HPC resources to address outstanding challenges in the life-sciences. They illustrate that molecular simulations offer a computational microscope to investigate dynamic natures of biomolecules. Given the inherent complexities of these systems in the context of their environment both advance in simulation methods, *e. g.* with regards coarse graining (CITE) and code development are strongly required.

References

1. E. Delaforge, S. Milles, J.-R. Huang *et al.*, *Investigating the Role of Large-Scale Domain Dynamics in Protein-Protein Interactions*, Front. Mol. Biosci. **3**, Article No. 54, 2016, doi:10.3389/fmolb.2016.00054.
2. L. Yang, J. S. Dordick, and S. Garde, *Hydration of Enzyme in Nonaqueous Media Is Consistent with Solvent Dependence of Its Activity*, Biophys. J. **87**, 812–821, 2004.
3. S. E. Graham, N. Leja, and H. A. Carlson, *MixMD Probeview: Robust Binding Site Prediction from Cosolvent Simulations*, J. Chem. Inf. Model. **58**, 1426–1433, 2018.
4. A. Miranker, S. E. Radford, M. Karplus, and C. M. Dobson, *Demonstration by NMR of folding domains in lysozyme*, Nature **349**, 633–636, 1991.
5. L. V. Bock, C. Blau, G. F. Schröder, I. I. Davydov, N. Fischer, H. Stark, M. V. Rodnina, A. C. Vaiana, and H. Grubmüller *Energy barriers and driving forces in tRNA translocation through the ribosome*, Nat. Struct. Mol. Biol. **20**, 1390–1396, 2013.

Conformational Analysis of Dual-Scale Simulations of Ubiquitin Chains

Andrej Berg and Christine Peter

Department of Chemistry, Universität Konstanz, 78464 Konstanz, Germany
E-mail: christine.peter@uni-konstanz.de

The analysis of large-scale simulations of (bio)molecular systems generated on high performance computer (HPC) clusters poses a challenge on its own due to the sheer amount of high-dimensional data. To make sense of these data and extract relevant information, techniques such as dimensionality reduction and clustering are used. They can be applied to characterise the sampling of conformational phase space as well as to bridge between simulations on different levels of resolution in multiscale setups. Here, we present an approach to analyse long-timescale simulations and to characterise conformational ensembles of flexibly-linked multidomain proteins using the example of differently covalently conjugated ubiquitin chains. We have analysed exhaustive coarse grained (CG) and atomistic simulations with the help of collective variables (CVs) that are particularly suitable to describe the mutual orientation of different subunits and the protein-protein interfaces between them. These data have been further processed through different dimensionality reduction techniques (relying on multidimensional-scaling like approaches as well as neural network autoencoders). The resulting low-dimensional maps have been used for the characterisation of conformational states and the quantitative comparison of conformational free energy landscapes (from simulations at different levels of resolution as well as of different chain types). With this multiscale simulation and analysis approach it is possible to identify characteristic properties of ubiquitin chains in solution which can be subsequently correlated with experimentally observed linkage- and chain length-specific behaviour.

1 Introduction

A special form of post translational modification of proteins is ubiquitylation, which describes the covalent attachment of the C-terminus of the Ubiquitin (Ub) protein to a lysine side chain in a substrate protein by an isopeptide bond. Ub itself possesses seven lysine (K) residues and additionally its N-terminal methionine (M) for further ubiquitylation which allows for the formation of different types of polyubiquitin (polyUb) chains. These chains can be selectively recognised by ubiquitin binding domains (UBDs). For some UBDs, relative or absolute selectivity for distinct linkage types and chain lengths was shown.[1, 2] Since all possible homotypic chain types, and even mixed and branched polyUb, were found in cells so far, it is supposed that the Ub signalling system serves as a tool to store and transmit information inside the eukaryotic cellular system.[3, 4]

The specific Ub recognition mechanisms are largely unknown, although some first insights can be obtained from experimentally determined (static) structures. Recent studies, however, are indicating that Ub dimers (diUb), which are the shortest possible Ub chains, can assume linkage specific, dynamic ensembles of various conformations in solution.[5–7] In the following we present our previous and ongoing work on the conformational characterisation of Ub conjugates from molecular dynamics (MD) simulations. We are using a dual-scale framework linking protein simulations on the atomistic and coarse grained (CG) level to obtain free-energy landscapes and to identify low free-energy configurations, which are representatives for the thermodynamically weighted conformational ensemble of

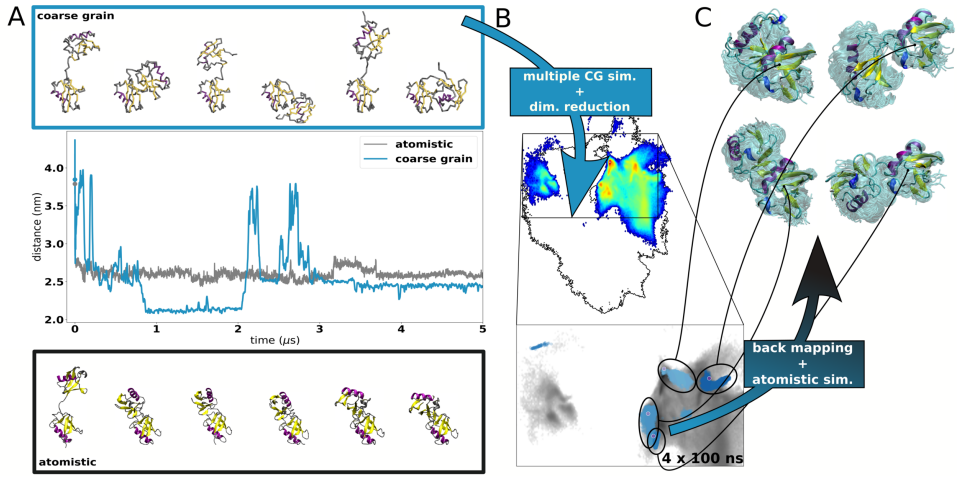


Figure 1. Schematic representation of the multi dimensional procedure. A: Centre of geometry distance between the two Ub moieties inside of diUb as function of time. Atomistic simulation (grey) exhibits a single aggregation event. Corresponding snapshots are shown below in cartoon representation. CG simulation (blue) exhibits repeated transitions between different conformations. Snapshots are shown above as sphere and stick model of back bone beads. B: Two-dimensional representation of a diUb ensemble for a certain linkage type (K27) obtained from multiple independent CG simulations (histogram coloured blue to red for low to high densities of conformations projected to the map). This landscape is shown in zoom inset in grey scale. Blue scatters show atomistic simulations which were started from back-mapped CG low energy simulations. C: Cartoon representation of diUb conformations from atomistic simulations shown in B (blue scatters) aligned on proximal chain.

Ub chains.[8,9] Atomistic simulations have a high level of accuracy but are not able to yield a proper sampling of the conformational equilibrium for diUb, with transitions between different collapsed structures that exhibit contact interfaces between the Ub subunits. For example, Fig. 1A (gray graph) illustrates how one observes a single aggregation event in a 5 μs simulation of K48 linked diUb where a stable contact interface forms that does not open up again. It is therefore not feasible to obtain an equilibrated atomistic ensemble for diUb with a correctly weighted representation of different collapsed structures and protein-protein interfaces. In contrast, on the CG level we can observe repeated transitions between open and aggregated configurations on the μs time scale (Fig. 1A, blue graph).

We used large conformational ensembles from CG MD simulations, a suitable set of collective variables (CVs) and dimensionality reduction to obtain a two-dimensional map of the conformational space (Fig. 1B) that can be interpreted as a free-energy landscape if one assumes that the sampling is converged. This representation can be easily used for structural clustering and intuitive interpretation of complex data sets. From the landscapes we identified low free-energy conformations for back-mapping to the atomistic level for visualisation and as input for MD simulations to obtain high resolution ensembles of these low free-energy states (Fig. 1C). We used the same dimensionality reduction technique to project the conformations from atomistic MD simulations on top of the landscape from the CG ones to validate, that atomistic simulations are staying in the low free-energy area from where they were started on ns time scales.[9] Additionally, such two-dimensional representations can be analysed in many different ways. Certain conformational properties

and features can be assigned to certain landscapes areas.[9] In contrast to high dimensional representations, two-dimensional distributions can be compared easily qualitatively but also quantitatively, for example with the Earth-Mover-Distance[10] algorithm. Clustering using k-means or other density based algorithms can be applied to such low dimensional representations to identify conformational states.

For dimensionality reduction we have successfully used the Sketch-map algorithm in the past, which is a multi-dimensional scaling like method which generates a map by iteratively minimising a nonlinear distance metric for a subset of representative data points (landmarks), focusing on an intermediate range of distances between data points with the help of a sigmoid function. All other points are subsequently projected into the map based on their relative positions to those landmarks.[11, 12] We found that Sketch-map gives good results in combination with a suitable set of input CVs to characterise the interaction of two Ub proteins/subunits to each other.[8, 9] However, this algorithm can be computationally demanding for very large data sets. Therefore, we extended the analysis and applied a new algorithm called EncoderMap, which was developed and improved recently in our group.[13, 14] This method combines the advantage of the pairwise distance based cost function of Sketch-map with a neuronal network autoencoder for dimensionality reduction. Here, we show that with this machine learning technique we were able to reproduce our results with higher efficiency. Thus, the results are robust with respect to the exact dimensionality reduction method used. Furthermore the use of EncoderMap allows to extend the analysis to substantially larger data sets such as simulations of ubiquitin trimers (triUb) and longer chains.

2 Methods

MD simulations

All simulations were performed with the GROMACS simulation package v5.[15] A modified version of the MARTINI force field v2.2 was used for CG simulations.[8, 16] For further details on the simulation setup and force field parameters please see Refs. 8 and 9. For each diUb linkage type (8) 20 simulations (of 10 μs each) were performed (200 μs simulation time for each linker). Two unlinked Ub were simulated in 80 simulations (10 μs each). Ub trimers (triUb) were simulated for 100 μs in 20 simulations which sums up to 2000 μs for each linker type. Structures were used every 100 ps for analysis.

EncoderMap

A detailed description on the functionality of EncoderMap can be found in Ref. 13 which is available as a python package on *github.com*.[a] The network which was used to obtain projections presented in Fig. 2 was trained on diUb conformations only. Based on the high-dimensional distance distribution of CG data, the sigmoid function parameters $\sigma = 5.9$, A = 12, B = 4, a = 2, b = 4 were chosen. This network was build up from 144 input nodes followed by three hidden layers of 150 neurons for each layer and two bottle neck nodes for the encoder part. The decoder part of the network was mirrored at the bottle neck layer.

[a] `https://github.com/AG-Peter/encodermap`

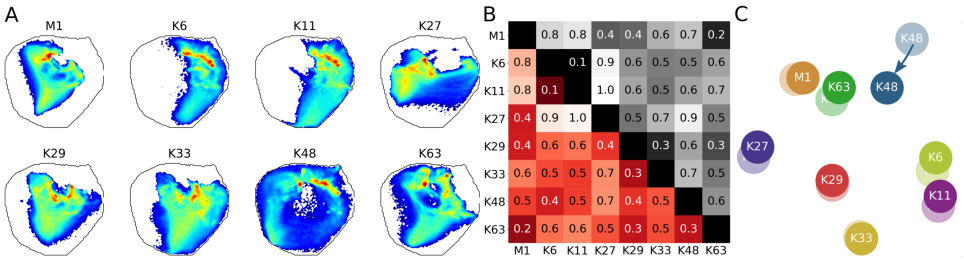


Figure 2. EncoderMap projections of diUb and similarity map. A: Histograms of two-dimensional representations of diUb conformations obtained by EncoderMap dimensionality reduction separated by linkage type (coloured blue to red for low to high densities of conformations projected to the map). Black circular line represents the outer rim of the sum over all linkage types. B: Relative pair-wise EMDs calculated between histograms shown in A (red lower half matrix) and histograms obtained by Sketch-map in Ref. 8 (gray upper half matrix). C: Similarity between two-dimensional conformational landscapes in A is represented by their distance to each other (rich coloured markers). Positioning was obtained by MDS of pair-wise EMD shown in lower half matrix in B. Initial positions for MDS optimisation were taken from Ref. 8 (pale markers).

The network was trained on $1.6 \cdot 10^6$ randomly selected conformations from diUb CG simulations over 10^4 training steps. The network which was used to obtain projections presented in Fig. 3 was trained on randomly selected configurations ($n = 4.6 \cdot 10^6$) from CG simulations of two unlinked Ub proteins. The network topology was the same as for the first neuronal network, but sigmoid function parameters were adjusted to σ = 10, A = 6, B = 10, a = 2, b = 6.

Miscellaneous

To obtain atomistic representations of simulated CG conformations BACKWARD was used for back-mapping.[17] All figures were created using Python v.3.5 and Matplotlib v.2.2.2. Two-dimensional distributions were quantitatively compared using the EMD algorithm as it is implemented in Pyemd v.0.5.1[10] to give relative pair-wise EMDs between two-dimensional distributions of diUb simulations (Fig. 2B). These distances were then used together with metric multi-dimensional scaling (implemented in *sklearn.manifold.MDS*) to image the similarity between two-dimensional representations (Fig. 2C).

3 Results

Dual Scale Simulations of Ub Dimers

We have used CG simulations, which we validated by comparison with atomistic and experimental data, to characterise the conformation of diUb in solution.[8] To this end, we performed extensive MD simulations on the CG level which by now consist of 200 μs simulation time for each of the 8 possible linkage positions for diUb. Residue-wise minimum distances (RMD), calculated for each dimer conformation every 100 ps, have been found to give a good set of CVs for dimensionality reduction. The RMD vector consists of 144 minimum distances between the two subunits, this means that for each of the 72 backbone beads (the last four residues of the highly flexible C-terminus are excluded) in the distal

moiety the minimum distance to any backbone bead in the proximal moiety is calculated and *vice versa*. We used Sketch-map to obtain a low dimensional (2D) representation of this 144 dimensional RMD space. Conformations which have similar RMD values are positioned nearby in 2D while dissimilar conformations are separated. We calculated two dimensional free energy landscapes for each linkage type which we used for qualitative and quantitative interpretation of how the linkage positions influence the conformation of diUb. Quantitative analysis was done by calculation of Earth-Mover-Distances (EMD), which is a metric developed for image recognition but can also be used to compare two-dimensional distributions. This combination of state of the art methods yields a versatile tool for the investigation of protein-protein interaction.

While Sketch-map is rather good for dimensionality reduction of non linear high dimensional data it has also some drawbacks.[18] The most striking for us was, that for training of the model only a limited amount of representative structures, so called landmarks can be used (a few hundreds in the present case). They have to be selected carefully since they should span the whole phase space of the data set. Once the model is trained, each sample from the data set is projected into the map based on optimised positions relative to the landmark points. Although, this method is much faster than conventional MDS with all samples at the same time (which is only feasible with some thousand data points), it can take several days to project a microsecond CG simulation data set. EncoderMap solves these two problems by replacing the MDS algorithm with a neural network autoencoder.[13] A neural network can be trained with a much larger data set (up to several million samples), *i. e.* no landmarks are needed, and such a network, once trained, can be used to project a huge amount of samples with an efficiency, that read and write operations (IO) become the limiting factor. By adding a MDS like pairwise-distance cost contribution to the normal autoencoder cost function upon training, EncoderMap aims (similarly to Sketch-map) at reproducing relative distances between data points in the low-dimensional projection (latent space).

Here, we have applied EncoderMap to the CG data set from Ref. 8 to see if this alters the results of our analysis of diUb conformational landscapes (Fig. 2A). Using very similar parameters for the sigmoid function in the Sketch-map cost contribution of EncoderMap, which determine the outcome of the projection to a large extent, we were able to obtain conformational landscapes which are strikingly similar compared to the previously obtained Sketch-map projections (see Fig. 2 in Ref. 8). A significant difference is observed only for the K48 linked dimer which now spans almost the whole conformational space accessible to diUb. However, this is not due to the dimensionality reduction technique but due to the fact that in the meantime we have also expanded the conformational ensembles by 66 % for each linkage type by addition of multiple new independent runs. It turns out that the K48 ensemble had not been fully converged. As in Ref. 8 the eight two-dimensional distributions are compared to each other by calculating pair-wise EMD which are normalised to a range between zero (for identical distributions) and one, for two very dissimilar maps, K11 and K27 in this case. In Fig. 2B the EMDs for the EncoderMap projections are compared to the data obtained with Sketch-map. Although, some values are different for EncoderMap projections (red bottom half matrix) compared to Sketch-map projections (upper gray half matrix), the overall similarity map for diUb is only slightly affected (Fig. 2C). This similarity map is obtained by again using MDS with the pair-wise EMDs as input for optimisation (the positioning from Ref. 8 has been used

for initialisation). The good agreement between the two different dimensionality reduction methods validates our assumption that RMDs are a suitable set of CVs for a structure based analysis of protein-protein interactions. Furthermore, EncoderMap enables us to analyse much larger data sets, which is required for our ongoing work on the characterisation of simulation data obtained from simulation of Ub trimers and tetramers.

Impact of Linkage on Protein-Protein Interface Formation

In the previous section we demonstrated how a suitable set of CVs together with dimensionality reduction can be used to obtain a (conformational) free energy landscape for differently linked diUb. Since this set of CVs is independent of the linker positions on the Ub surface, this method can be also applied to unlinked Ub, or other proteins in general. Therefore, to characterise the various binding modes Ub can exhibit to a second Ub, we performed extensive CG simulations (80×10 μs) of two unlinked Ub. A Sketch-map model was trained on conformations from these simulations only.[9] We obtained a landscape which was able to depict a larger interaction phase space than if diUb data was used

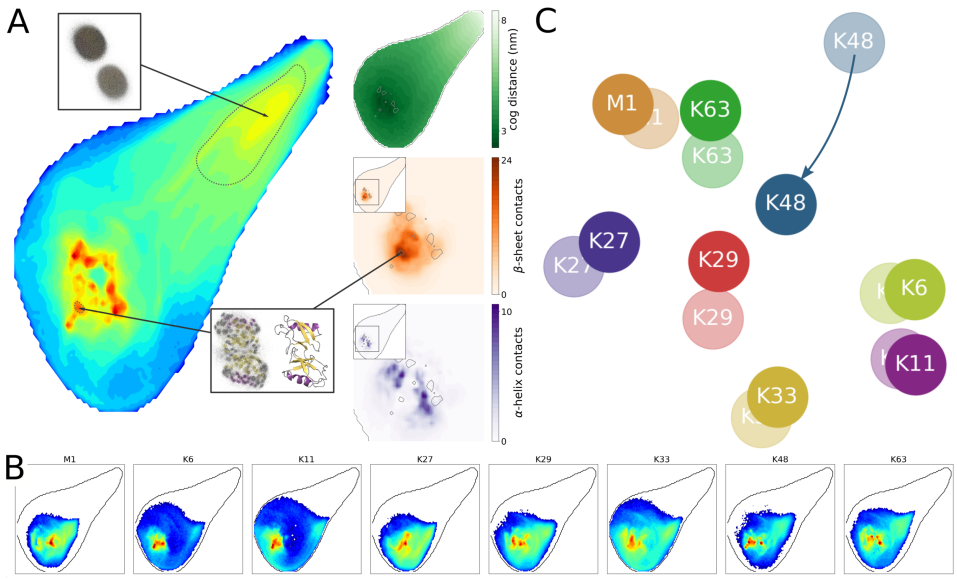


Figure 3. EncoderMap projections of two unlinked Ub and diUb and similarity map. A: Histogram of two-dimensional representations of two unlinked Ub configurations obtained by EncoderMap dimensionality reduction (coloured blue to red for low to high densities of conformations projected to the map). Panels on the right show certain properties of configurations dependent on their position on the landscape: Centre of geometry distance between the two proteins (green); Number of β-sheet contacts (orange) and α-helix contacts between the two proteins. Structure bundles from CG simulations extracted from certain areas on the map demonstrate conformational clustering. Cartoon representation shows back-mapped atomistic structure. B: Histograms, separated by linkage type, of two-dimensional representations of diUb conformations obtained by an EncoderMap network which was trained on unlinked configurations. Black circular line represents the outer rim of landscape in A. C: Similarity between two-dimensional conformational landscapes in B is represented by their distance to each other (rich coloured markers). Positioning was obtained by MDS of pair-wise EMDs. Initial positions for MDS optimisation were taken from Ref. 8 (pale markers).

for training. We found, that the set of CVs which we used as input for dimensionality reduction introduces an arbitrary ordering of chain A and B, which had no evident impact while used for diUb. In a dimer topology chain A is called distal moiety and is linked via its C-terminus to the lysine side chain of moiety B, also called proximal. For two unlinked Ub, which are obviously identical, this is not true any more and the high dimensional data set contains samples which are identical if one would change the chain order from AB to BA. Interestingly, this feature manifests as a two-fold symmetry in the low dimensional projection (as discussed and demonstrated in detail in Ref. 9).

With this model we were able to obtain proper projections for diUb and investigate the immediate (linkage specific) effect of covalent linkage formation on the protein-protein interface between two Ub proteins. Since conformations of diUb are projected into the two-dimensional representation of two unlinked Ub, one can compare these two landscapes directly and observe that some linkage types (*e. g.* K63) select already present conformations while others (*e. g.* K6) obtain conformations which are not stable without any linker. This might be critical for the investigation of the Ub ligase mechanism which happens by a cascade of reactions involving a linkage and substrate specific combination of three enzymes.[3] Additionally, by training the model on a set of configurations which can be formed by two unlinked proteins and should therefore contain all possible configurations at least to some extent, the chance is reduced to obtain unreasonable positions in 2D if new diUb conformations are added to the projection.

In our present work, thanks to the ability of EncoderMap to process a much larger data set for training, we were able to improve the projection of Ub conformations even further – and consequently enhance the characterisation of the protein-protein interface formed by two Ub proteins. Exploiting the inherent symmetry of the projection problem, due to the equivalence of the Ub proteins, we can further improve the map by doubling the high dimensional RMD data set for two unlinked Ub by just changing the chain order. With EncoderMap we are able to use this large data set ($4.6 \cdot 10^6$ samples) for training to obtain an almost perfectly symmetric landscape (Fig. 3A). This projection divides clearly configurations with no contacts between the two chains (high centre of geometry distance, see small green histogram) from aggregated structures (the red highly populated spots in the main map). One can find configurations with a symmetric interaction interface along the symmetry axis and – as it was shown in Ref. 9 already – configurations from two minima which are mirrored at this symmetry axis can be aligned almost perfectly if the chain order is interchanged. Certain interaction interface features, like a high amount of β-sheet contacts (see small orange histogram) can be assigned to certain areas on the map and enhances the interpretation of the configurational space.

The EncoderMap neuronal network, once trained, can be now used to project diUb configurations in a very effective way (shown in Fig. 3B for all linkages). In these landscapes, diUb conformations occupy only the area with lower centre of geometry distance which is obviously due the linkage constraint. We can observe similar overlaps for certain linkage types and we can repeat the quantitative comparison by calculating pair-wise EMD for these distributions, which are then used for MDS optimisation to obtain a similarity map (Fig. 3C). Initial positions were again taken from Ref. 8. Although all markers (saturated colours) are shifted slightly compared to the reference data (pale colours), the overall picture persists, except for K48 which we already identified to be different due to additional sampling compared to the previously published data.

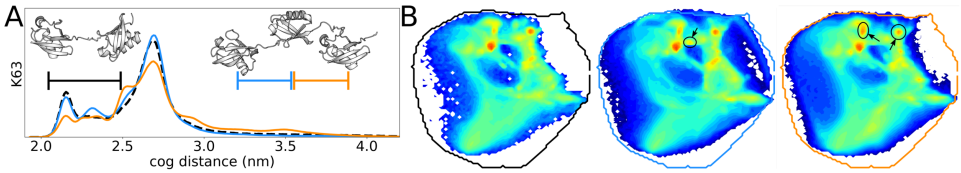


Figure 4. Conformation of K63 linked polyUb. A: Centre of geometry distance distributions between distal and proximal chain in diUb (black) and adjacent chains in triUb (blue, orange) from CG simulations. B: Two-dimensional representations obtained by EncoderMap network which was trained on RMD data from diUb. Conformational landscape of K63 linked diUb (left panel). Interaction landscapes of adjacent Ub moieties in K63 linked triUb (middle and right panel).

Protein-Protein Interface Characterisation in Ub Trimers

To investigate the linkage specific conformational space of longer Ub chains we performed CG simulations of all eight homotypic triUb. In this section we will present preliminary results from these simulations, where we obtained 2000 μs of simulation time for each linkage type (topology). Fig. 4A illustrates how centre of geometry distance distributions between two adjacent chains are altered in triUb compared to a diUb with the same linkage type (K63). Although, this gives already some insight how polyUb chain elongation affects the interaction between Ub moieties, we can use the EncoderMap neuronal network, which was trained to characterise diUb conformations, to obtain interaction maps for triUb.

As discussed above, the two Ub proteins in diUb are not equivalent, which means that the distal moiety ($diUb_d$) is bound by its C-terminus to the proximal moieties ($diUb_p$) lysine side-chain. TriUb has an additional Ub moiety in the middle ($triUb_m$) which is constrained by a covalent bond in both directions, by its lysine side-chain to $triUb_d$ and by its C-terminus to $triUb_p$. Consequently, while diUb has a single interaction interface between $diUb_d$ and $diUb_p$, triUb has two interfaces (for the moment neglecting, that $triUb_d$ and $triUb_p$ can in principle make contact, too). For diUb we had used the RMD vector which describes distances between the distal and proximal moiety to train an EncoderMap neuronal network to obtain a conformational landscape, which can be considered at the same time as an interaction map (Fig. 2A). We applied this trained network to triUb to obtain two interaction maps for each triUb linkage type by providing two different RMD vectors for each triUb conformation, *i. e.* one RMD vector between $triUb_d$ and $triUb_m$ and a second one between $triUb_m$ and $triUb_p$. These maps are shown in Fig. 4B for K63 linked diUb and triUb, where one can see that the overall appearance of the two triUb interaction landscapes is comparable with the conformational landscape from diUb. A closer view reveals that differences in distance distributions are also imaged as altered intensities of certain states in these maps. Since we are able to assign representative structures to each state (red areas in 2D distribution), we can obtain a very detailed picture about the interaction pattern of polyUb chains. By this, we will observe for each linkage type which interaction modes are altered in a triUb compared to diUb and we will conclude on reasons for chain length specificity of certain UBDs.

4 Concluding Remarks

In the present work we present the application of EncoderMap to efficiently obtain low-dimensional representations of complex conformational ensembles of protein conjugates. These representations allow for conformational clustering but also for the visualisation of certain properties and quantitative comparison between histograms of different data sets. With this, intuitive examination of large simulated data sets is possible. Furthermore, selected CVs (RMD) as input for dimensionality reduction are highly transferable and can be used to characterise protein-protein interaction of unlinked proteins as well as certain parts of larger conjugates. In our ongoing work we use the combination of tools presented here to characterise polyUb (up to tetramers) conformations for all homotypic chain types. By this we hope to extend our knowledge about the linkage type and chain length specific conformational properties of polyUb and with this contribute to the understanding of the Ub recognition mechanism.

Acknowledgements

We would like to thank Tobias Lemke for support and helpful suggestions regarding the use of EncoderMap and for critically reading this manuscript. AB is funded by the DFG through SFB969 (Project B09). We are grateful for computational resources of the bwHPC project (DFG grant INST 35/1134-1 FUGG and the state of Baden-Württemberg) for the initial simulations of diUb. We also gratefully acknowledge computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

References

1. F. Ikeda, N. Crosetto, and I. Dikic, *What Determines the Specificity and Outcomes of Ubiquitin Signaling?*, Cell **143**, 677–681, 2010.
2. X. Zhang, A. H. Smits, G. B. A. van Tilburg, P. W. T. C. Jansen, M. M. Makowski, H. Ovaa, and M. Vermeulen, *An Interaction Landscape of Ubiquitin Signaling*, Molecular Cell **65**, 941–955.e8, 2017.
3. D. Komander and M. Rape, *The Ubiquitin Code*, Annual Review of Biochemistry **81**, 203–229, 2012.
4. K. N. Swatek and D. Komander, *Ubiquitin Modifications*, Cell Research **26**, 399–422, 2016.
5. C. Castañeda, A. Chaturvedi, C. M.Camara, J. E.Curtis, S. Krueger, and D. Fushman, *Linkage-specific Conformational Ensembles of Non-canonical Polyubiquitin Chains*, Physical Chemistry Chemical Physics **18**, 5771–5788, 2016.
6. C. A. Castañeda, E. K. Dixon, O. Walker, A. Chaturvedi, M. A. Nakasone, J. E. Curtis, M. R. Reed, S. Krueger, T. A. Cropp, and D. Fushman, *Linkage via K27 bestows ubiquitin chains with unique properties among polyubiquitins*, Structure **24**, 423–436, 2016.
7. A. Kniss, D. Schuetz, S. Kazemi, L. Pluska, P. E. Spindler, V. V. Rogov, K. Husnjak, I. Dikic, P. Güntert, T. Sommer, T. F. Prisner, and V. Dötsch, *Chain Assembly*

and Disassembly Processes Differently Affect the Conformational Space of Ubiquitin Chains, Structure **26**, 249–258.e4, 2018.
8. A. Berg, O. Kukharenko, M. Scheffner, and C. Peter, *Towards a molecular basis of ubiquitin signaling: A dual-scale simulation study of ubiquitin dimers*, PLoS computational biology **14**, e1006589, 2018.
9. A. Berg and C. Peter, *Simulating and analysing configurational landscapes of protein-protein contact formation*, Interface Focus **9**, 20180062, 2019.
10. O. Pele and M. Werman, *A linear time histogram metric for improved sift matching*, in Computer Vision–ECCV 2008, Springer, 495–508, 2008.
11. M. Ceriotti, G. A. Tribello, and M. Parrinello, *Simplifying the Representation of Complex Free-energy Landscapes Using Sketch-map*, Proceedings of the National Academy of Sciences **108**, 13023–13028, 2011.
12. O. Kukharenko, K. Sawade, J. Steuer, and C. Peter, *Using Dimensionality Reduction to Systematically Expand Conformational Sampling of Intrinsically Disordered Peptides*, Journal of Chemical Theory and Computation **12**, 4726–4734, 2016.
13. T. Lemke and C. Peter, *EncoderMap: Dimensionality Reduction and Generation of Molecule Conformations*, Journal of Chemical Theory and Computation **15**, 1209–1215, 2019, PMID: 30632745.
14. T. Lemke and C. Peter, *EncoderMap(II): Visualizing important molecular motions with improved generation of protein conformations*, Journal of Chemical Information and Modeling, 2019, doi:10.1021/acs.jcim.9b00675.
15. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers*, SoftwareX **1**, 19–25, 2015.
16. L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S.-J. Marrink, *The MARTINI Coarse-Grained Force Field: Extension to Proteins*, J. Chem. Theory and Comput. **4**, 819–834, 2008.
17. T. A. Wassenaar, K. Pluhackova, R. A. Böckmann, S. J. Marrink, and D. P. Tieleman, *Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models*, Journal of Chemical Theory and Computation **10**, 676–690, 2014.
18. G. A. Tribello and P. Gasparotto, *Using Dimensionality Reduction to Analyze Protein Trajectories*, Frontiers in Molecular Biosciences **6**, 46, 2019.

Solvent Effects on the Binding of Fatty Acids to Human Serum Albumin

Pradeep Pant, Yasser B. Ruiz-Blanco, and Elsa Sanchez-Garcia

Computational Biochemistry, Center of Medical Biotechnology,
University of Duisburg-Essen, 45117 Essen, Germany
E-mail: elsa.sanchez-garcia@uni-due.de

Solvent interactions can influence the properties and function of complex biomolecules. Among others, altering the solvent composition has consequences for the recognition of binding partners. Human serum albumin (HSA) is one of the most enigmatic biomolecules, known as an efficient carrier of biological materials, such as hormones, fatty acids and drugs.

Here we explored the effects of the solvent on stearic acid-HSA binding. To this end, we performed all-atom molecular dynamics (MD) simulations in explicit solvent ($\sim$ 2.9 µs in total). These MD simulations were carried out in explicit water and in a 20 % ethanol-water mixture. The sampling in both systems was processed with the MM-PBSA binding free energy approach, which allowed us to investigate the effects of the solvent composition on the binding of stearic acid molecules to seven binding sites of HSA. Using this computational approach, we were able to reproduce the experimental preference of fatty acid's binding sites for albumin in water. **Site 5** $>$ **site 4** $>$ **site 2** were calculated as high affinity fatty acid binding sites, in agreement with the experimental reports.[1] Interestingly, we observed that **site 1** becomes the most prominent binding pocket in the 20 % ethanol-water mixture, with overall binding affinity towards stearic acid: **site 1** $>$ **site 5** $>$ **site 2.** Our simulations in explicit solvent also provided a rationale for this effect. Importantly, we achieved weak binding-to strong binding conversion by using a solvent mixture, with repercussions for the specific binding properties and the manipulation of HSA properties as biological carrier.

1 Introduction

Human serum albumin (HSA) is a 585 residues protein found in most fluids of the body.[2–4] HSA acts as depot and carrier for a broad spectrum of compounds like fatty acids and affects the pharmacokinetics of many drugs such as penicillin and sulphonamides.[5, 6] HSA also acts as a toxic waste handler.[7] Furthermore, HSA displays pseudo enzymatic properties and it is a valuable biomarker in many diseases.[8, 9] HSA features several binding sites for medium to long chain fatty acids, wherein lauric acid (C12), myristic acid (C14), palmitic acid (C16), and stearic acid (C18) are commonly found.[10]

Changes in solvent media[11–13]can strongly affect molecular interactions that occur routinely in aqueous solution. Protein-protein, protein-DNA, and protein-RNA recognition are progressed by water-mediated hydrogen bonds.[14–16] Thus, the hydrogen bond donor and acceptor capabilities of water often play a crucial role in biological recognition.[17–19] Solvent mixtures have been extensively used to unravel macromolecular and recognition events such as protein binding hotspots identification, ligand binding, protein folding, and the elucidation of biological mechanisms.[13, 20–22] HSA in particular displays three high-affinity fatty acid binding sites and four weakly binding sites in water, which were confirmed by several experimental and theoretical studies.[23–25] However, to the best of our knowledge, the solvent-driven regulation of stearic acid binding to albumin has not been addressed.

Here, we explored the binding of stearic acid to HSA (Fig. 1) in water and in a 20 % ethanol-water mixture. The simulations in water were used as reference and control system. There, we were able to reproduce the experimental order of binding of stearic acid to the seven binding pockets of HSA, validating our computational protocol. A different binding site (**site 1**) emerged as preferred in the 20 % ethanol-water mixture. This can be explained by an enhanced hydrogen bond network between stearic acid and HSA in **site 1**, involving nearby ethanol molecules. Energy calculations using the MM-PBSA[26] approach also indicated enhanced electrostatics effects in the ethanol-water system.

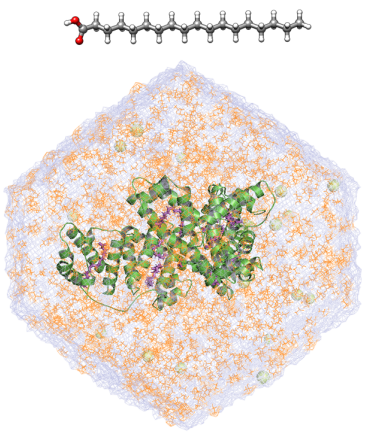


Figure 1. Top: Stearic acid (C18), carbon atoms are shown in grey, oxygen in red, hydrogen in white. Bottom: Representative structure of HSA in 20 % ethanol-water, water is shown as a blue mesh surface, ethanol molecules in orange, stearic acid in violet and ions as green spheres.

2 Methods

A crystal structure of human serum albumin (RCSB PDB:[27] 1E7I, resolution: 2.7 Å) was taken as starting structure for the computational studies. The first two N-terminal residues (aspartic acid, alanine) and the C-terminal leucine were missing in the crystal structure and hence were subsequently modelled by us using I-Tasser.[28] The Root Mean Square Deviation (RMSD) differences between the energy minimised structure and the crystal structure was 0.2 Å. The available crystallographic data of palmitic acid in HSA[29] (PDB ID: 4BKE) was used to set the initial orientations of the stearic acid in HSA since, in the 1E7I complex, some of the fatty acid molecules missed the carboxyl group. The protonation states of the titrable residues of HSA were determined using the H++ server[30] and corroborated by visual inspection. MD simulations were performed using the AMBER 14[31] and NAMD 2.12b software suites.[32] An octahedral water box of 15 Å was generated as both reference and control system along with the requisite counter ions to neutralise the system. The modified ff99SB force field[33] was used for the protein in the simulations with AMBER 14, and the CHARMM36m force field[34] was used in the simulations with

NAMD 2.12b. Stearic acid parameters for AMBER were generated at the HF/6-31G*[35] level of theory using Gaussian09.[36, 37] PME was used to treat long-range electrostatic interactions.[38] The system was first minimised (15 000 CG + 15 000 SD) and then equilibrated for 5 ns (at room temperature) followed by 50 ns of production run to allow the structure to relax in the aqueous environment. The last frame comprising HSA with seven stearic acid molecules in their binding pockets was used to generate seven starting systems (HSA with only one ligand). Each of those systems was again simulated (3 independent replicas of 50 ns each). The MM-PBSA approach was used for computing the binding free energies. The Cpptraj[39] module of AMBER was utilised for the analysis of the MD trajectories (RMSD fluctuations, hydrogen bonding). The 20 % ethanol-water mixture was generated using Packmol,[40] a freely accessible tool that creates initial configurations for MD simulations. For HSA in the 20 % ethanol-water mixture, we followed a protocol analogous to that described for HSA in water. The equilibrated structure of albumin in the 20 % ethanol-water mixture is shown in Fig. 1.

3 Results and Discussion

The molecular dynamics simulation trajectories were visualised using the VMD tool.[41] HSA with seven stearic acid molecules occupying simultaneously all the possible binding sites was first simulated for 50 ns in an octahedral box with counterions added to neutralise the whole system. The backbone RMSD fluctuations throughout the trajectory were below 3.0 Å indicating no overall large dynamical changes in the structure. The final frame of this trajectory was used to generate seven initial systems, with a single stearic acid molecule in each of the seven binding pockets of HSA. These systems were further relaxed by performing 50 ns long MD simulations as per the protocol discussed in the Methods section. We observed stable control plots of temperature, pressure, density and total energy (kinetic and potential) along with low RMSD fluctuations. The last frame of this MD provided the initial coordinates for performing three replicas of 50 ns of each system.

From the sampling in water, we extracted 500 frames, which were used to calculate the binding free energies of stearic acid in the fatty acid-binding sites of HSA using the MM-PBSA approach, which has been extensively applied to estimate binding free energies of small molecules in the binding pocket of macromolecular systems[25] (Tab. 1). We observed that in water, **site 5** has the highest binding affinity for stearic acid (-24.5 $\pm$ 0.59 kcal/mol (Tab. 1)), followed by **site 4** (-18.96 $\pm$ 0.76 kcal/mol) and **site 2** (-18.15 $\pm$ 0.82 kcal/mol). These three regions are also reported as high affinity fatty acid-binding sites in the literature.[1] Other binding spots, classified as low affinity, have binding energies $>$ -15 kcal/mol, also in agreement with experimental observations[24, 25] and validating our protocol of simulations and energy estimations using MM-PBSA. We note that the change in the dielectric constant of the system from water to the 20 % ethanol-water mixture is not expected to affect significantly the evaluation of the Poisson's equation for continuum electrostatics. Also, the contributions of non-polar groups (which is proportional to the variation of the solvent accessible surface area) is expected to be lower than for more polar systems like pure water. Hence the MM-PBSA approach for calculating binding free energy remains a suitable protocol in water-organic solvent mixtures at moderate or low organic solvent fraction.

Binding sites	Water	20 % ethanol-water
Site1	-13.6 (0.93)	-29.9 (0.81)
Site2	-18.2 (0.82)	-23.9 (0.81)
Site3	-6.5 (0.84)	-6.2 (0.57)
Site4	-19.0 (0.76)	-13 (1.71)
Site5	-24.5 (0.59)	-24.0 (0.62)
Site6	-12 (1.02)	-16 (1.01)
Site7	-14.8 (0.68)	-12 (1.09)

Table 1. Energetics of binding of stearic acid molecules (in kcal/mol) in the seven fatty acid binding sites of HSA in water and in the 20 % ethanol-water mixture. SE of mean indicated in parentheses.

Further, we examined the binding proclivity of stearic acid in the binding pockets of HSA in the 20 % ethanol-water system. **Site 1** emerged as the preferred spot with the most favourable binding free energy for stearic acid (-29.88 $\pm$ 0.81 kcal/mol) followed by **site 5** (-24.01 $\pm$ 0.62 kcal/mol) and **site 2** (-23.89 $\pm$ 0.81 kcal/mol). These results evidence how altering solvent composition induces a low-binding to high binding transition and overall enhances the binding affinity of stearic acid towards HSA.

The analysis of the contributions to the free energy of binding of stearic acid in **site 1** (Fig. 2) indicates a remarkable increment in the total electrostatic contribution in the 20 % ethanol-water mixture with respect to water, leading to an enhanced binding of stearic acid to HSA in ethanol-water.

The 2-D interaction diagrams (plotted using the LigPlot+ software[42]) allowed us to rationalise the enhanced electrostatics contributions of stearic acid in **site 1** of HSA in the

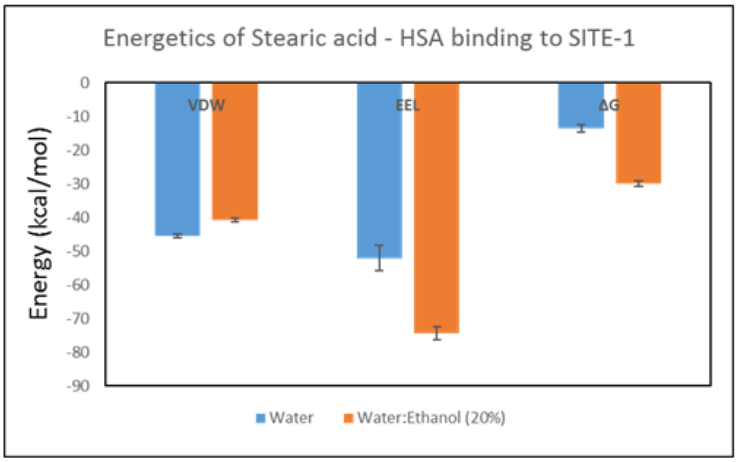


Figure 2. Electrostatic and van der Waals contributions to the total free energy of binding for the stearic acid (**site 1**)-HSA system in water (blue) and in the 20 % ethanol-water mixture.

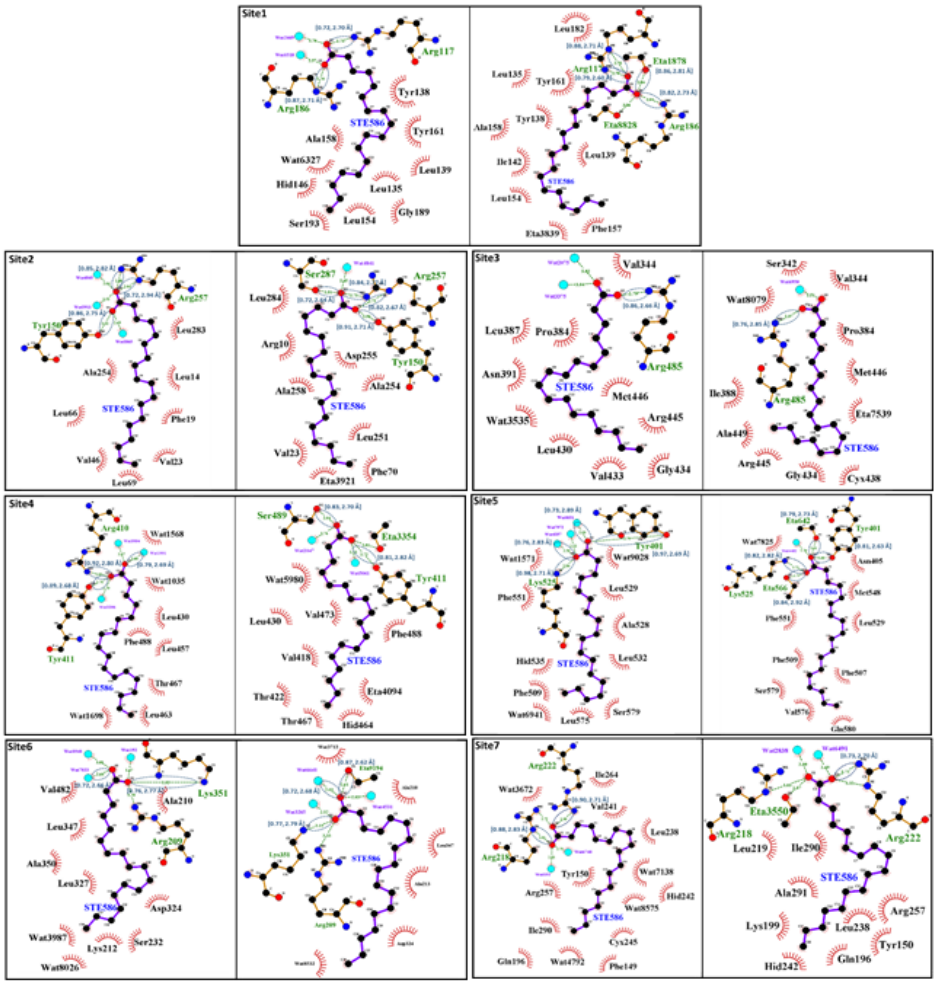


Figure 3. 2-D interaction plots of stearic acid in the seven fatty acid binding pockets of human serum albumin. All hydrogen bonds with lifetime > 0.70 are highlighted. In all cases, the left panel corresponds to the simulations in water and the right panel to the simulations in the 20 % ethanol-water mixture. The lifetime fraction of these hydrogen bonds and the average h-bond distances are shown in parenthesis (format: [lifetime fraction, average distance], see Tab. 2).

ethanol-water mixture (Fig. 3, hydrogen bond interactions with a fraction half-life time larger than 0.70 are highlighted). In water environment, stearic acid establishes in **site 1** two conserved hydrogen bonds with HSA residues (Arg 117 and Arg 186). However, in the 20 % ethanol-water mixtures, the stearic acid molecule in **site 1** is able to establish four conserved hydrogen bonds (Arg 117, Arg 186 and two ethanol molecules) (Tab. 2).

On the other hand, **site 4** shifts from a high affinity site to a low affinity site upon addition of methanol to the water solvent environment. The analysis of the interactions in this site allows us to rationalise this observation (Fig. 3). In water, stearic acid displays

	Water			**20 % ethanol-water**		
Binding site	Residue	Average h-bond distance (Å)	Fraction of h-bond lifetime	Residue	Average h-bond distance (Å)	Fraction of h-bond lifetime
Site1	Arg186	2.71	0.87	Arg186	2.73	0.82
	Arg117	2.70	0.72	Arg117	2.60	0.79
				Arg117	2.71	0.88
				ETA	2.81	0.86
Site2	Arg257	2.94	0.72	Arg257	2.67	0.82
	Arg257	2.82	0.85	Arg257	2.77	0.84
	Tyr150	2.75	0.86	Tyr150	2.67	0.82
				Ser287	2.64	0.72
Site4	Tyr411	2.68	0.89	Tyr411	2.82	0.81
	Arg410	2.80	0.92	Ser489	2.70	0.83
	WAT	2.69	0.79			
Site5	Tyr401	2.69	0.92	Tyr401	2.63	0.81
	Lys525	2.71	0.98	Lys525	2.82	0.82
	WAT	2.89	0.73	ETA	2.92	0.84
	WAT	2.83	0.76	ETA	2.73	0.79

Table 2. Average hydrogen bond distance and fraction of hydrogen bond lifetime for high affinity binding sites.

three hydrogen bonds (Tyr 411, Arg 410 of HSA and one molecule of water) in **site 4**. However, in the solvent mixture, the stearic acid molecule in **site 4** is only able to establish two conserved hydrogen bonds with Tyr 411 and Ser 489 of HSA. The 3-D representation of the three high affinity fatty acid binding sites (**site 1**, **site 2**, **site 5**) in both systems (water and 20 % ethanol-water) also evidence the role of solvent interactions on the stabilisation of specific binding motifs (Fig. 4).

Next, we compare representative snapshots from the MD simulations of the stearic acid-albumin complex (with stearic acid in **site 1**) under different solvent conditions with the reported crystal structure of stearic acid-albumin (PBD ID: 1E7I) (Fig. 5). The comparison indicates that the stearic acid molecule adopts a less compact conformation in the 20 % ethanol-water mixture with respect to the control system (water). A higher conformational flexibility of the stearic acid molecule in the solvent mixture may enable it to establish more hydrogen bond contacts with strong affinity.

Overall, we observed that the solvent has modulating effects on the binding of the fatty acid to a particular site of human serum albumin. **Site 1**, which was amongst the low affinity fatty acid binding sites in water, emerged as a high affinity-binding site in the ethanol-water mixture.

4 Conclusions

Here we show that altering solvent composition can be a useful tool for modulating biomolecular function, in this case the binding specificity of human serum albumin as a carrier of fatty acids. Low affinity to high affinity fatty acid binding site conversion was achieved when using, instead of water as solvent, a 20 % water-ethanol mixture. Conversely, a high affinity fatty acid binding site transitioned to low affinity under the same conditions. Our findings, which we rationalise at the molecular level, have special relevance for the regulation of biological function since we show that the targeted manipula-

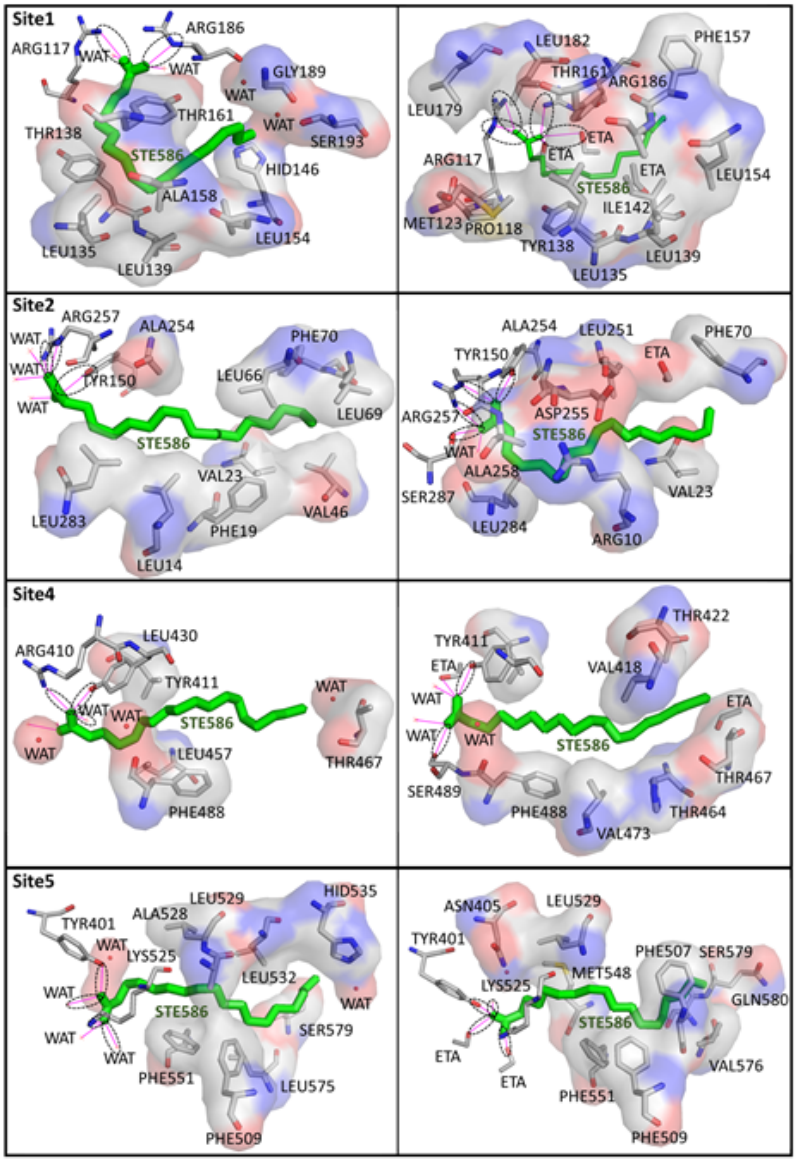


Figure 4. Structure of HSA in water (left panel) and in 20 % ethanol-water for **site 1**, **site 2, site 4** and **site 5**. Residues/molecules involved in van der Waals interactions are shown in surface representation (vdW colour cloud with red for oxygen, blue: nitrogen, gray: carbon). The stearic acid molecule is shown in green sticks. Conserved hydrogen bond contacts are circled.

tion of protein-ligand binding events can be achieved by rationally tuned solvent-driven interactions.

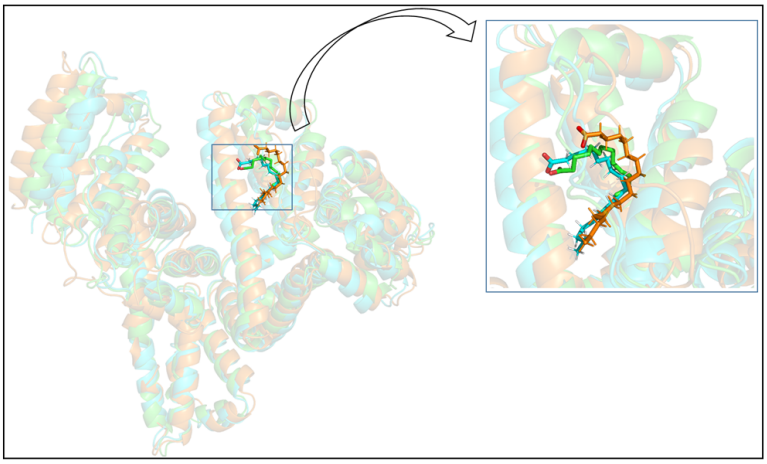


Figure 5. Overlay of the structures of human serum albumin in the crystal structure (green), in water (cyan) and in the 20 % ethanol-water mixture (orange) with stearic acid bound to **site 1** (close-up as inset). Water and counterions were removed for clarity.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2033 – Projektnummer 390677874 and by the FP-RESOMUS program. E. S.-G. also acknowledges the support of the Boehringer Ingelheim Foundation (Plus-3 Program). Y. R.-B and E. S.-G. 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). E. S.-G. acknowledges the working group Albunet for Albumin studies and Dr. Katja Waterstradt for experimental collaboration.

References

1. S.-I. Fujiwara, T. Amisak *et al.*, *Identification of High Affinity Fatty Acid Binding Sites on Human SerumAlbumin by MM-PBSA Method*, Biophys. J. **94**, 95–105, 2008.
2. Q. Qian, Z. You, L. Ye, J. Che, Y. Wang, S. Wang, and B. Zhong, *High-efficiency production of human serum albumin in the posterior silk glands of transgenic silkworms, Bombyx mori L*, PLoS ONE **13**, e0191507, 2018.
3. C. Cai, K. Zhou, Y. Wu, and L. Wu, *Enhanced liver targeting of 5-fluorouracil using galactosylated human serum albumin as a carrier molecule*, J. Drug Targ. **14**, 55–61, 2006.
4. A. Lamousé-Smith, G. M. Subramanian, M. Fiscella, S. Zeuzem, and J. G. McHutchison, *Albinterferon alpha-2b: a genetic fusion protein for the treatment of chronic hepatitis C*, Nat. Biotech. **25**, 1411–1419, 2007.
5. D. Bratlid and T. Bergan, *Displacement of albumin-bound antimicrobial agents by bilirubin*, Pharmacology **14**, 464–472, 1976.

6. A. Dalhoff, *Seventy-Five Years of Research on Protein Binding*, Antimicrob Agents Chemother **62**, e01663-17, 2018.
7. S. Yuge, M. Akiyama, and T. Komatsu, *An Escherichia coli trap in human serum albumin microtubes*, Chem. Commun. **50**, 9640–9643, 2014.
8. U. Kragh-Hansen, *Molecular aspects of ligand binding to serum albumin*, Pharmacol. ReV. **33**, 17–53, 1981.
9. J. Seetharamappa and B. J. Kamat, *Study of the interaction between fluoroquinolones and bovine serum albumin*, J. Pharm. Biomed. **39**, 1046–1050, 2005.
10. S. Curry, P. Brick, and N. P. Franks, *Fatty acid binding to human serum albumin: new insights from crystallographic studies*, Biochim. Biophys. Acta **1441**, 131–140, 1999.
11. P. M. Wiggins, *Role of water in some biological processes*, Microbiol Rev. **54**, 432–449, 1990.
12. D. Roccatano, *Computer simulations study of biomolecules in non-aqueous or cosolvent/water mixture solutions*, Current Protein & Peptide Science **9**, 407–426, 2008.
13. L. Yang, J. S. Dordick, and S. Garde, *ydration of enzyme in nonaqueous media is consistent with solvent dependence of its activity*, Biophys. J. **87**, 812–821, 2004.
14. S. K. Pal and A. H. Zewail, *Dynamics of Water in Biological Recognition*, Chem. Rev. **104**, 2099–2123, 2004.
15. G. A. Papoian, J. Ulander, and P. G. Wolynes, *Role of Water Mediated Interactions in ProteinProtein Recognition Landscapes*, J. Am. Chem. Soc. **125**, 9170–9178, 2003.
16. B. Jayaram and T. Jain, *The role of water in Protein-DNA recognition*, Ann. Rev. Biophy. Biomol. Struc. **33**, 343–361, 2004.
17. S. Wong, R. E. Amaro, and J. A. McCammon, *MM-PBSA Captures Key Role of Intercalating Water Molecules at a ProteinProtein Interface*, J. Chem. Theory Comput. **5**, 422–429, 2009.
18. D. Russo, G. Hura, and T. Head-Gordon, *Hydration Dynamics Near a Model Protein Surface* Biophys. J. **86**, 1852–1862, 2004.
19. P. Ball, *Water as an active constituent in cell biology*, Chem. Rev. **108**, 74–108, 2008.
20. Y. Yu, J. Wang, Q. Shao, J. Shi, and W. Zhu, *The effects of organic solvents on the folding pathway and associated thermodynamics of proteins: a microscopic view*, Sci. Rep. **6**, 19500, 2016.
21. G. Carrea and S. Riva, *Properties and Synthetic Applications of Enzymes in Organic Solvents*, Angew. Chem. Int. Ed. **39**, 2226–2254, 2000.
22. S. E. Graham, N. Leja, and H. A. Carlson, *MixMD Probeview: Robust Binding Site Prediction from Cosolvent Simulations*, J. Chem. Inf. Model. **58**, 1426–1433, 2018.
23. P. Lee and X. Wu, *Review: modifications of human serum albumin and their binding effect*, Curr. Pharm. Des. **21**, 1862–1865, 2015.
24. J. R. Simard, P. A. Zunszain, C.-E. Ha, J. S. Yang, N. V. Bhagavan, I. Petitpas, S. Curry, and J. A. Hamilton, *Locating high-affinity fatty acid-binding sites on albumin by x-ray crystallography and NMR spectroscopy*, Proc. Natl. Acad. Sci. USA **102**, 17958–17963, 2005.
25. D. S. Goodman, and E. Shafrir, *The Interaction of Human Low Density Lipoproteins with Long-chain Fatty AcidAnions*, J. Am. Chem. Soc. **81**, 364–370, 1959.
26. H. Gohlke, and D. A. Case, *onverging free energy estimates: MM-PB(GB)SA studies on the protein protein complex Ras-Raf*, J. Comput. Chem. **25**, 238–250, 2004.

27. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, *The Protein Data Bank*, Nucleic Acids Res. **28**, 235–242, 2000.
28. J. Yang and Y. Zhang, *I-TASSER server: new development for protein structure and function predictions*, Nucleic Acids Res. **43**, W174–W181, 2015.
29. A. Sivertsen, J. Isaksson, H.-K. S. Leiros, J. Svenson, J.-S. Svendsen, and B.-O. Brandsdal, *Synthetic Cationic Antimicrobial Peptides Bind with Their Hydrophobic Parts to Drug Site II of Human Serum Albumin*, Bmc Struc. Biol. **14**, 4–13, 2014.
30. R. Anandakrishnan, B. Aguilar, and A. O. Onufriev, *H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations*, Nucleic Acids Res. **40**, W537–541, 2012.
31. D. A. Case *et al.*, *AMBER 14*, University of California, San Francisco, 2014.
32. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. Hauser, and C. Simmerling, *ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB*, J. Chem. Theory Comput. **11**, 3696–3713, 2015.
33. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *Scalable molecular dynamics with NAMD*, J. Comput. Chem. **26**, 1781–1802, 2005.
34. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.-L. de Groot, H. Grubmüller, and M. A. MacKerell Jr., *CHARMM36m: an improved force field for folded and intrinsically disordered proteins*, Nat. Methods. **14**, 71–73, 2017.
35. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. MacKerell Jr., *CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields*, J. Comput. Chem. **31**, 671–690, 2009.
36. E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak, and F. Y. Dupradeau, *R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments*, Nucleic Acids Res. **39**, W511–W517, 2011.
37. W. D. Cornell *et al.*, *A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids and Organic Molecules*, J. Am. Chem. Soc. **117**, 5179–5197, 1995.
38. T. E. Cheatham III, J. L. Miller, T. Fox, T. A. Darden, and P. A. Kollman, *Molecular Dynamics Simulations on Solvated Bio-molecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA and Proteins*, J. Am. Chem. Soc. **117**, 4193–4194, 1995.
39. D. R. Roe and T. E. Cheatham III, *PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data*, J. Chem. Theory Comput. **9**, 3084–3095, 2013.
40. L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, *PACKMOL: a package for building initial configurations for molecular dynamics simulations*, J. Comput. Chem. **30**, 2157–2164, 2009.
41. W. Humphrey, A. Dalke, and K. Schulten, *VMD: visual molecular dynamics*, J. Molec. Graphics **14**, 33–38, 1996.
42. R. A. Laskowski and M. B. Swindells, *LigPlot+: multiple ligand-protein interaction diagrams for drug discovery*, J. Chem. Inf. Model. **51**, 2778–2786, 2011.

Elementary Particle Physics

Elementary Particle Physics

Stefan Schaefer

John von Neumann Institute for Computing, DESY, Platanenallee 6, 15738 Zeuthen, Germany
E-mail: stefan.schaefer@desy.de

Strong interaction physics is believed to be described by Quantum Chromodynamics (QCD), part of the Standard Model of particle physics. It is the theory of quarks and gluons, which interact forming protons, neutrons and a plethora of other hadrons. While there is little doubt that this theory is correct, predictions of low-energy strong interaction phenomena are difficult.

With qualitative insight available for some time, in recent years precision calculations have started to provide crucial input for the understanding of experiments as the ones conducted at the LHC at CERN in Geneva or the Belle experiment at KEK in Japan. The motivation of these experiments is the verification of the standard model and possibly finding phenomena beyond its reach. The experimental programmes using the collision of heavy ions, as they are conducted at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory (BNL) or the ALICE experiment at CERN, aim at understanding the behaviour of QCD at the high temperatures that are believed to have existed at the beginning of the universe. The interpretation of all these experiments needs specific theory input which are the target of investigations in this chapter.

Since such computations are necessarily done from first principles, and the underlying theory is strongly coupled, significant computer resources as well as human commitment over a long time are required to arrive at accuracies which make the results relevant. The contributions in this volume show the rich spectrum of phenomena described by QCD, with effects spanning energy scales of many orders of magnitude. While the questions addressed in these investigations vary widely, the underlying methodology unites them: definition of the theory on a lattice in Euclidean space-time and evaluation of the resulting path integral using Markov Chain Monte Carlo, including the construction of observables which such an approach can evaluate.

This section starts with a contribution by Meißner, that is different from the ones which follow. The authors do not attempt to solve directly the fundamental theory of QCD, but use a nuclear effective theory to describe phenomena which are far beyond the reach of QCD calculations, *e. g.* the binding energy of light nuclei. The contribution highlights some algorithmic developments which make the computations more efficient and exploit specific physical properties of the system.

Bali and Collins then report on investigations of the baryon spectrum and their structure. The dependence of the masses of the baryon octet on the quark masses are investigated, along with observables which encode their structure. Such computations can have impact, *e. g.* on dark matter searches and are an example of quantities, which are difficult to get from experiment and therefore theory input is required. The physics of the charm quark, which is the fourth quark if ordered by mass, is the subject of the investigations presented by Höllwieser, Knechtli and Korzec. Here the first steps of a programme simulating

QCD with four dynamical quarks are presented, with first results reproducing the masses of charmed mesons.

The final two contributions provide essential input for the interpretation of heavy ion collision experiments, investigating the behaviour of QCD at high temperature and its phase transition. Kaczmarek *et al.* present results for the chiral phase temperature in the situation where two quarks are massless and the strange quark at its physical value. Borsanyi *et al.* in the final contribution to this section investigate the behaviour of QCD away from zero baryon chemical potential, as is the case in heavy ion experiments. As direct computations in this situation are not feasible, the direction of imaginary chemical potential is explored. Together with a Taylor expansion around zero chemical potential this can be used to make contact to experimental data from the STAR collaboration at the RHIC accelerator at BNL.

Nuclear Lattice Effective Field Theory: Status and Perspectives

Ulf-G. Meißner[1,2,3]
for the NLEFT Collaboration

[1] Helmholtz-Institut für Strahlen- und Kernphysik and Bethe Center for Theoretical Physics, Universität Bonn, 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

[3] Tbilisi State University, 0186 Tbilisi, Georgia

I give an outline of the recent developments in 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. This remarkable progress has been made possible by recent dramatic increases in HPC resources, as well as advances in computational methods and algorithmic developments.

1 Introduction

Recent advances in high-performance computing (HPC) have 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, see *e. g.* Ref. 1. 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 ^{12}C and ^{16}O). 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 and beyond. 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 order interaction 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) solutions of the Lippmann-Schwinger or Faddeev-Yakubowsky 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 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, particularly when nucleons (baryons) are involved.

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 leading-order (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. More complicated initial states, such as α-clusters or alike, are also possible. 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

$$Z(t) \equiv \langle \Psi_A | \exp(-tH) | \Psi_A \rangle = \mathrm{Tr}\{M^{L_t}\} \tag{1}$$

in the limit of large Euclidean projection time t. M is the usual normal-ordered transfer -matrix operator and L_t is the number of Euclidean time steps. 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. Similarly, the breaking of Galilean invariance in moving frames is taken care of. A much more detailed description is given in the recent monograph.[2]

4 New Algorithms and 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. We start with the latter.

4.1 The Pinhole Algorithm

Auxiliary-field Monte Carlo (AFMC) simulations are efficient for computing the quantum properties of systems with attractive pairing interactions. By the calculating the exact

quantum amplitude for each configuration of auxiliary fields, we obtain the full set of correlations induced by the interactions. However, the exact quantum amplitude for each auxiliary field configuration involves quantum states which are superpositions of many different centre-of-mass positions. Therefore information about density correlations relative to the centre of mass is lost. The pinhole algorithm[3] is a new computational approach that allows for the calculation of arbitrary density correlations with respect to the centre of mass. As this was not possible in all previous AFMC simulations, adaptations of this technique should have wide applications to hadronic, nuclear, condensed matter, and ultracold atomic simulations.

Let $\rho_{i,j}(\mathbf{n})$ be the density operator for nucleons with spin i and isospin j at lattice site $\mathbf{n}$,

$$\rho_{i,j}(\mathbf{n}) = a^\dagger_{i,j}(\mathbf{n}) a_{i,j}(\mathbf{n}) \tag{2}$$

We construct the normal-ordered A-body density operator

$$\rho_{i_1,j_1,\cdots i_A,j_A}(\mathbf{n}_1,\cdots \mathbf{n}_A) = :\rho_{i_1,j_1}(\mathbf{n}_1)\cdots\rho_{i_A,j_A}(\mathbf{n}_A): \tag{3}$$

In the A-nucleon subspace, we note the completeness identity

$$\sum_{i_1,j_1,\cdots i_A,j_A}\sum_{\mathbf{n}_1,\cdots \mathbf{n}_A} \rho_{i_1,j_1,\cdots i_A,j_A}(\mathbf{n}_1,\cdots \mathbf{n}_A) \;= A! \tag{4}$$

The new feature of the pinhole algorithm is that MC importance sampling is performed according to the absolute value of the expectation value

$$\langle\Psi_f|M_*^{L_t'} M^{L_t/2}\rho_{i_1,j_1,\cdots i_A,j_A}(\mathbf{n}_1,\cdots \mathbf{n}_A) M^{L_t/2} M_*^{L_t'}|\Psi_i\rangle \tag{5}$$

Here, M^t is the transfer matrix at time t of dimension $A\times A$ and M_*^t is referring to the SU(4) invariant transfer matrix used at the beginning and at the end of the time evolution to tame the sign problem, see *e. g.* Ref. 2. Due to the completeness identity Eq. (4), the sum of the amplitude in Eq. (5) over $\mathbf{n}_1,\cdots \mathbf{n}_A$ and $i_1,j_1,\cdots i_A,j_A$ gives $A!$ times the amplitude without any insertion of the A-body density,

$$\langle\Psi_f|M_*^{L_t'} M^{L_t} M_*^{L_t'}|\Psi_i\rangle \tag{6}$$

The pinhole locations $\mathbf{n}_1,\cdots \mathbf{n}_A$ and spin-isospin indices $i_1,j_1,\cdots i_A,j_A$ are sampled by Metropolis updates, while the auxiliary fields are sampled by the hybrid MC algorithm (see also below). In the left panel of Fig. 1 we show a sketch of the pinhole locations and spin-isospin indices for the operator $\rho_{i_1,j_1,\cdots i_A,j_A}(\mathbf{n}_1,\cdots \mathbf{n}_A)$ inserted at time $t = L_t a_t/2$. We obtain the ground state expectation value by extrapolating to the limit of infinite projection time.

For spatial lattice spacing a, the coordinates $\mathbf{r}_i$ of each nucleon on the lattice is an integer vector $\mathbf{n}_i$ times a. We do not consider mass differences between protons and neutrons for the moment. Since the centre of mass is a mass-weighted average of A nucleons with the same mass, the centre-of-mass position $\mathbf{r}_{\rm CM}$ is an integer vector $\mathbf{n}_{\rm CM}$ times a/A. Therefore the density distribution has a resolution scale that is A times smaller than the lattice spacing. In order to determine the centre-of-mass position $\mathbf{r}_{\rm CM}$, we minimise the squared radius

$$\sum_i |\mathbf{r}_{\rm CM} - \mathbf{r}_i|^2 \tag{7}$$

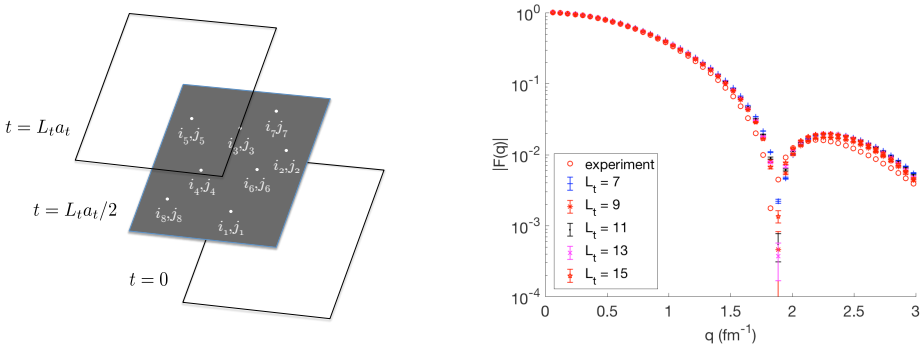


Figure 1. Left panel: A sketch of the pinhole locations and spin-isospin indices $(i_n j_n)$ at time $t = L_t a_t/2$. Right panel: The elastic form factor $F(q)$ for ^{12}C calculated with the pinhole algorithm. Open circles are the experimental values, the other symbols refer to NLEFT results for various Euclidean times L_t.

where each term $|\mathbf{r}_{\rm CM} - \mathbf{r}_i|$ is minimised with respect to all periodic copies of the separation distance on the lattice. As a first application, we show in the right panel of Fig. 1 the calculated elastic form factor of ^{12}C (which is the Fourier transform of the charge density, calculated using the pinhole algorithm[3]). The agreement with the data is quite satisfactory. This paves the way for detailed nuclear structure investigations.

4.2 The Shuttle Algorithm

In most NLEFT simulations, the hybrid MC (HMC) algorithm is used to update the configurations, which consist of nucleon world-lines coupled to external fields. In Ref. 4, we presented the so-called shuttle algorithm. More precisely, the auxiliary fields $s(n_t, \boldsymbol{n})$ that represent the interactions between the nucleons are updated using the shuttle algorithm, where only one time slice is updated at a time. In Fig. 2 a schematic plot sketching the difference between the shuttle algorithm and the HMC algorithm which performs an update of all time slices is shown. The shuttle algorithm works as follows:

1. Choose one time slice n_t, record the corresponding auxiliary field as $s_{\rm old}(n_t, \boldsymbol{n})$.

2. Propose the new auxiliary fields $s_{\rm new}(n_t, \boldsymbol{n})$ at each lattice site $\boldsymbol{n}$ according to the probability distribution $P\left[s_{\rm new}(n_t, \boldsymbol{n}) = \phi_k\right] = \omega_k$ for $k = 1, 2, 3$. We note that $\omega_1 + \omega_2 + \omega_3 = 1$.

3. Calculate the determinant of the $A \times A$ correlation matrix M_{ij} using $s_{\rm old}(n_t, \boldsymbol{n})$ and $s_{\rm new}(n_t, \boldsymbol{n})$, respectively.

4. Generate a random number $r \in [0, 1)$ and perform the following "Metropolis test". If
$$\left| \frac{\det\left[M_{ij}\left(s_{\rm new}(n_t, \boldsymbol{n})\right)\right]}{\det\left[M_{ij}\left(s_{\rm old}(n_t, \boldsymbol{n})\right)\right]} \right| > r \tag{8}$$
accept the new configuration $s_{\rm new}(n_t, \boldsymbol{n})$ and update the wave functions accordingly, otherwise keep $s_{\rm old}(n_t, \boldsymbol{n})$.

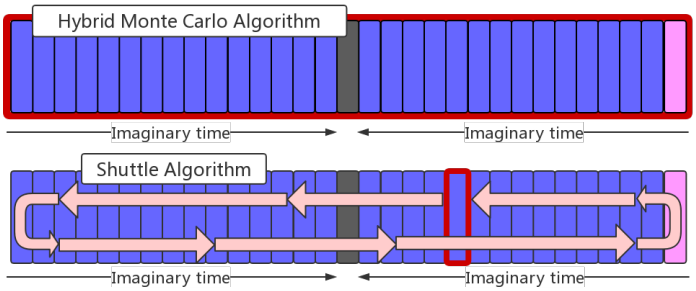


Figure 2. Schematic plot for the HMC algorithm (upper panel) and the shuttle algorithm (lower panel). The red squares denote the time slices to be updated in each run. The pink, blue and black boxes refer to the centre-of-mass projection, the SU(4) invariant transfer matrix and the density operator under consideration, respectively.

5. Proceed to the next time slice, repeat steps 1)-4), and turn around at the end of the time series.

As shown in Fig. 2, the program runs back-and-forth like a shuttle bus and all the auxiliary fields are updated after one cycle is finished. The shuttle algorithm is well suited for small values of the temporal lattice spacing a_t. In this case the number of time slices is large and the impact of a single update is small. In each update the new configuration is close to the old one, resulting in a high acceptance rate. For example, in the recent works discussed below, the temporal lattice spacing is $a_t = 0.001$ MeV^{-1} and the acceptance rate is around 50 % in most cases. We compared the results with the HMC algorithm and found that the new algorithm is more efficient. In most cases the number of independent configurations per hour generated by the shuttle algorithm is three or four times larger than that generated by the HMC algorithm.

4.3 Essential Elements of Nuclear Binding

NLEFT and other *ab initio* methods encounter problems when going to larger mass numbers. More precisely, the three-body forces (and eventually four-body forces) become more important with increasing A, which makes the computations more difficult and harder to control. This led *e. g.* our collaboration to introduce an effective four-nucleon term when considering alpha-type nuclei up to ^{28}Si in 2014.[5] In Ref. 4, we presented a highly improved LO action that overcomes these problems and allows to elucidate the essential mechanisms underlying nuclear binding. For related investigations in the continuum, see *e. g.* Ref. 6 (and references therein).

We start with a simple SU(4)-invariant leading order EFT without explicit pions on a periodic L^3 cube with lattice coordinates $\boldsymbol{n} = (n_x, n_y, n_z)$. The Hamiltonian is

$$H_{\mathrm{SU(4)}} = H_{\mathrm{free}} + \frac{1}{2!} C_2 \sum_{\boldsymbol{n}} \tilde{\rho}(\boldsymbol{n})^2 + \frac{1}{3!} C_3 \sum_{\boldsymbol{n}} \tilde{\rho}(\boldsymbol{n})^3 \tag{9}$$

where H_{free} is the free nucleon Hamiltonian with nucleon mass $m = 938.9$ MeV. The

density operator $\tilde{\rho}(\boldsymbol{n})$ is defined in the same manner as in Ref. 3,

$$\tilde{\rho}(\boldsymbol{n}) = \sum_i \tilde{a}_i^\dagger(\boldsymbol{n})\tilde{a}_i(\boldsymbol{n}) + s_L \sum_{|\boldsymbol{n}'-\boldsymbol{n}|=1} \sum_i \tilde{a}_i^\dagger(\boldsymbol{n}')\tilde{a}_i(\boldsymbol{n}') \tag{10}$$

where i is the joint spin-isospin index and the smeared annihilation and creation operators are defined as

$$\tilde{a}_i(\boldsymbol{n}) = a_i(\boldsymbol{n}) + s_{NL} \sum_{|\boldsymbol{n}'-\boldsymbol{n}|=1} a_i(\boldsymbol{n}') \tag{11}$$

The summation over the spin and isospin implies that the interaction is SU(4) invariant. Note that we perform local (L) as well as non-local smearing (NL). In the latter case, the nucleon creation/annihilation operators are distributed over a lattice point and its six neighbours. The parameter s_L controls the range of the local part of the interaction, while s_{NL} controls the range of the nonlocal (*i. e.* velocity-dependent) part of the interaction. The parameters C_2 and C_3 give the strength of the two-body and three-body interactions, respectively. In what follows, we use a lattice spacing $a = 1.32$ fm, which corresponds to a momentum cutoff $\Lambda = \pi/a \approx 465$ MeV. The dynamics with momentum Q much smaller than Λ can be well described and residual lattice artifacts are suppressed by powers of Q/Λ, see Ref. 7. For systems with more than three nucleons, we use AFMC lattice simulations for a cubic periodic box with length L, see Ref. 2. For nuclei with $A < 30$ nucleons, we take $L \geq 8$, with larger values of L for cases where more accuracy is desired. For nuclei with $A \geq 30$ we take $L = 9$. The temporal lattice spacing is 0.001 MeV^{-1} and the projection time is set to 0.3 MeV^{-1}. We find that these settings are enough to provide accurate results for systems with $A \leq 48$. We also use the recently-developed pinhole algorithm[3] in order to calculate density distributions and charge radii. We use few-body data with $A \leq 3$ to fix the interaction coefficients C_2 and C_3, while the range of the interactions are controlled by the parameters s_{NL} and s_L. In the few-body sector, the two smearing parameters s_{NL} and s_L produce very similar effects and are difficult to distinguish from few-body data alone. Therefore, to pin down s_L and s_{NL} more precisely, we require fits to heavier nuclei as described in detail in Ref. 4. The full set of optimised parameters are $C_2 = -3.41 \times 10^{-7}$ MeV^{-2}, $C_3 = -1.4 \times 10^{-14}$ MeV^{-5}, $s_{NL} = 0.5$, and $s_L = 0.061$.

Using this highly improved action, we have calculated the binding energies for 86 bound even-even nuclei (even number of protons, even number of neutrons) with up to $A = 48$ nucleons, see the left panel of Fig. 3. Due to the SU(4) symmetry, there is no sign problem and all of the MC error bars are smaller than the size of the symbols. The remaining errors due to imaginary time and volume extrapolations are also small, less than 1 % relative error, and thus are also not explicitly shown. The general trends for the binding energies along each isotopic chain are well reproduced. In particular, the isotopic curves on the proton-rich side are close to the experimental results. The discrepancy is somewhat larger on the neutron-rich side and is a sign of missing effects such as spin-dependent interactions. With the same interaction, one can also calculate the charge density profile of a given nucleus. This has been done in Ref. 4 for ^{16}O and ^{40}Ca and the results are quite accurate for such a simple nuclear interaction.

Further, predictions can be made for pure neutron matter (NM), see the right panel of Fig. 3. There, we show the calculated NM energy as a function of the neutron density and the comparison with other calculations using next-to-next-to-next-to-leading-order

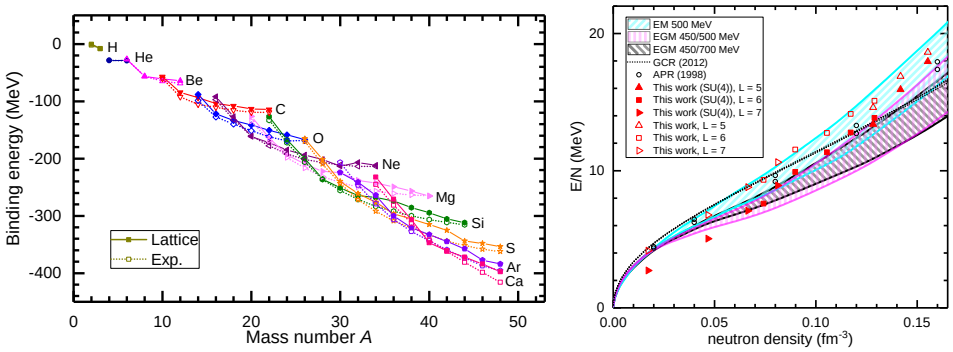


Figure 3. Left panel: The calculated binding energies from ^{3}H to ^{48}Ca. The solid symbols denote the lattice results and the open symbols denote the experimental values. Different symbols and colours denote different elements. The experimental values are taken from Ref. 8. Right panel: The pure neutron matter (NM) energy as a function of neutron density calculated using the NL50[4] interaction with box size L=5 (upright triangles), L=6 (squares), L=7 (rightward-pointing triangles), respectively. The filled red polygons show results for the leading-order SU(4)-symmetric interaction. The open red polygons show an improved calculation with a short-range interaction to reproduce the physical neutron-neutron scattering length as well as a correction to improve invariance under Galilean boosts. For comparison we also show results calculated with full N^{3}LO chiral interactions (EM 500 MeV, EGM 450/500 MeV and EGM 450/700 MeV),[9] the results from variational (APR)[10] and Auxiliary Field Diffusion MC calculations (GCR).[11]

(N^{3}LO) chiral interactions. In the lattice results we vary the number of neutrons from 14 to 66. The data for three different box sizes L=5 (upright triangles), L=6 (squares), L=7 (rightward-pointing triangles) are marked as filled red polygons. We see that our results are in general agreement with the other calculations at densities above 0.05 fm^{-3}, though calculations at higher orders are needed and are planned in future work to estimate uncertainties. At lower densities the discrepancy is larger as a result of the SU(4)-invariant interaction having the incorrect neutron-neutron scattering length. The open red polygons, again L=5 (upright triangles), L=6 (squares), L=7 (rightward-pointing triangles), show an improved calculation with a short-range interaction to reproduce the physical neutron-neutron scattering length as well as a correction to improve invariance under Galilean boosts. The restoration of Galilean invariance on the lattice is described in Ref. 12. Overall, the results are quite good in view of the simplicity of the four-parameter interaction.

The computational effort needed for the auxiliary-field lattice Monte Carlo simulations scales with the number of nucleons, A, as somewhere between A^1 to A^2 for medium mass nuclei. The actual exponent depends on the architecture of the computing platform. The SU(4)-invariant interaction provides an enormous computational advantage by removing sign oscillations from the lattice Monte Carlo simulations for any even-even nucleus. Coulomb interactions and all other corrections can be implemented using perturbation theory or the recently-developed eigenvector continuation method if the corrections are too large for perturbation theory.[13] Given the mild scaling with nucleon number and suppression of sign oscillations, the methods presented here provide a new route to realistic lattice simulations of heavy nuclei in the future with as many as one or two hundred nucleons. By realistic calculations we mean calculations where one can demonstrate order-by-order convergence in the chiral expansion going from LO to NLO, NLO to N^{2}LO, and N^{2}LO to N^{3}LO and so on, while maintaining agreement with empirical data.

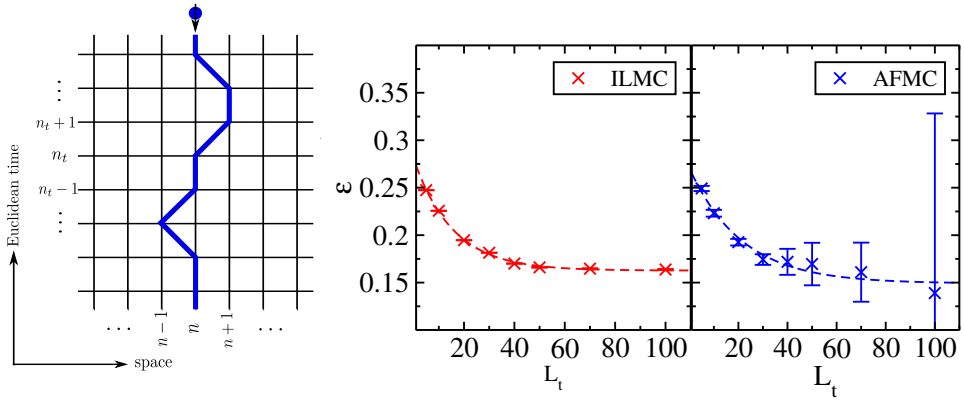


Figure 4. Left panel: A segment of a worldline configuration for a ↓ impurity in a sea of ↑-particles on a 1+1 dimensional Euclidean lattice. Right panel: The ground state energy as a function of Euclidean projection time for ILMC simulations versus AFMC simulations for $L^3 = 10^3$ and $N = 10$.

4.4 Impurity Lattice Monte-Carlo Algorithm and Polaron Physics

Hyper-nuclear (or strangeness nuclear) physics investigates nuclei, in which one (or two) nucleon(s) are substituted by a Λ or Σ hyperon (or two of these). This extends the nuclear chart into a third dimension, because such nuclei are catalogued in terms of proton number Z, neutron number N and strangeness $S = -1, -2, ...$ for one, two, ... hyperons. Such systems are in particular interesting as only very few hyperon-nucleon and hyperon-hyperon scattering data points (~ 50) exist, in contrast to the about $10\,000$ data on nucleon-nucleon scattering below pion production threshold. This is, of course, a consequence of the fact that no beams of hyperons exist and scattering data can only be inferred from final-state interactions. Therefore, bound systems including hyperons give important information on these elusive baryon-baryon interactions. It should also be noted that hyperon-nucleon scattering lengths are not unnaturally large (as far as they are known), see *e. g.* Ref. 14.

So the question naturally arises how to treat these systems in NLEFT? A first approach is to simply enlarge the degrees of freedom from the nucleon doublet to the hyperons with strangeness $S = -1$. A LO computation using just the contact terms and fitting the low-energy constants to certain cross section threshold ratios showed that such an approach is feasible and the volume dependence of the scattering lengths is consistent with the Lüscher formula.[15] However, due to the missing SU(4) Wigner symmetry in this type of systems, this approach was not further developed. So is there another method to treat hyperons in nuclei? The answer leads us to the so-called *impurity lattice MC algorithm* (ILMC), where the hyperon(s) are treated as an impurity(ies) in a sea of nucleons.

The basic idea of the ILMC algorithm was developed in Ref. 16. Consider a worldline configuration for the impurity as shown in the left panel of Fig. 4 and then compute the reduced transfer matrix by integrating out the impurity. To be specific, consider a system of one ↓-particle in a sea of ↑-particles with $m_\uparrow = m_\downarrow = m$. The lattice Hamiltonian

$H = H_0 + V$ is given by:

$$H_0 = H_0^{\uparrow} + H_0^{\downarrow} \,, \quad V = C_0 \sum_{\vec{n}} \rho_{\uparrow}(\vec{n})\, \rho_{\downarrow}(\vec{n})$$

$$H_0^s = \frac{1}{2m} \sum_{l=1}^{3} \sum_{\vec{n}} \Big[2a_s^{\dagger}(\vec{n}) a_s(\vec{n}) - a_s^{\dagger}(\vec{n}) a_s(\vec{n}+\hat{l}) - a_s^{\dagger}(\vec{n}) a_s(\vec{n}-\hat{l}) \Big] \tag{12}$$

with $s = \uparrow, \downarrow$, $\hat{l}$ is a unit vector on the lattice and C_0 is tuned to the binding energy of the shallow dimer consisting of one $\uparrow$ and the $\downarrow$-particle. It is advantageous to work in the occupation number basis,

$$\left| \chi_{n_t}^{\uparrow}, \chi_{n_t}^{\downarrow} \right\rangle = \prod_{\vec{n}} \left\{ \left[a_{\uparrow}^{\dagger}(\vec{n}) \right]^{\chi_{n_t}^{\uparrow}(\vec{n})} \left[a_{\downarrow}^{\dagger}(\vec{n}) \right]^{\chi_{n_t}^{\downarrow}(\vec{n})} \right\} \,, \;\; \chi_{n_t}^{s}(\vec{n}) = 0 \text{ or } 1 \tag{13}$$

This allows to calculate the elements of the reduced transfer matrix,

$$\langle \chi_{n_t+1}^{\uparrow}, \chi_{n_t+1}^{\downarrow} |\, M \,| \chi_{n_t}^{\uparrow}, \chi_{n_t}^{\downarrow} \rangle \tag{14}$$

In computing the transfer matrix, one has to deal with two different cases. First, if the impurity performs one spatial hop,

$$M_{\vec{n}'' \pm \hat{l}, \vec{n}''} = \left(\frac{\alpha_t}{2m} \right) \,:\, \exp\left[-\alpha_t H_0^{\uparrow} \right] \,: \tag{15}$$

with $\alpha_t = a_t/a$. If the impurity worldline remains stationary, the corresponding transfer matrix reads

$$M_{\vec{n}'', \vec{n}''} = \left(1 - \frac{3\alpha_t}{m} \right) \,:\, \exp\left[-\alpha_t H_0^{\uparrow} - \frac{\alpha_t C_0}{1 - 3\alpha_t/m} \rho_{\uparrow}(\vec{n}'') \right] \,: \tag{16}$$

Note that these reduced transfer matrices are just one-body operators on the linear space of the $\uparrow$-particles. With that, one can proceed as usual in NLEFT.

To show the power of the ILMC, we present in the right panel of Fig. 4 results for the ground state energy of a system of $9\,|\uparrow\rangle + 1\,|\downarrow\rangle$ particles in a volume $L^3 = 10^3$ ($N = 10$) with a zero-range interaction, using both conventional AFMC and the ILMC.[17] The ground state energy ϵ is given by:

$$\epsilon = \frac{1}{a_t} \lim_{L_t \to \infty} \ln \frac{Z(L_t - 1)}{Z(L_t)} \tag{17}$$

The ILMC clearly outperforms AFMC, because it is computationally simpler and faster and also has far smaller sign oscillations (see the detailed discussion in Ref. 17).

As a first application of the ILMC, we considered the polaron problem in two and three dimensions.[17] As a benchmark of the method, we calculated the universal polaron energy in three dimensions (3D) in the scale-invariant unitarity limit and found agreement with published results, see the left panel of Fig. 5. Next, consider attractive polarons in two dimensions. There is a very interesting and important question as to whether a polaron-molecule transition occurs in the ground state as a function of the interaction strength. At present cold-atom experiments are not yet conclusive on the question of a transition. The existence and nature of such a transition impacts the overall phase diagram for spin-imbalanced 2D Fermi gas. The theoretical investigations on this issues were so far inconclusive, see the discussion in Ref. 17. We have investigated this problem considering

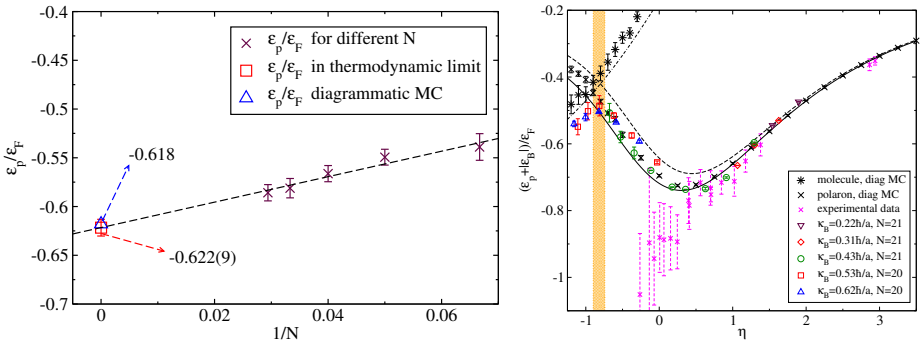


Figure 5. Left panel: The 3D unitarity limit polaron energy ϵ_P in units of the Fermi energy ϵ_F as a function of the inverse particle number $1/N$. The square is the ILMC result in the thermodynamic limit $N \to \infty$, while the triangle gives the diagrammatic Monte Carlo result from Ref. 18. Right panel: Ground state energy as a function of the dimensionless parameter $\eta = (1/2)\ln((2\epsilon_F/|\epsilon_B|))$ in comparison with diagrammatic MC results.[19] The data are from Ref. 20. The vertical band represents the region where the crossover transition from a polaron to a molecule occurs.

again one down-spin impurity and N up-spin particles in the limit of zero-range attractive interactions. The delta function interaction is tuned according to the two-body bound state energy, $|\epsilon_B|$. Using ILMC, we calculated the polaron energy as a fraction of the up-spin Fermi energy. The coupling constant was tuned in order to get the two-body bound states with binding momentum $\kappa_B = \sqrt{m|\epsilon_B|}$ equal to $0.22/a$, $0.31/a$, $0.43/a$, $0.53/a$, and $0.62/a$. We ran simulations for several different lattice areas, L^2, and several different particle numbers, N. The lattice sizes go from $L^2 \times L_t = 20^2 \times 100$ to $L^2 \times L_t = 80^2 \times 700$. For each L^2 and N, the ground state energy was obtained by extrapolating to the limit $L_t \to \infty$ by fitting the Euclidean time projection amplitude to the asymptotic function $\epsilon_0 + \alpha \exp(-\delta \cdot t)$. To magnify the details, we subtracted the dimer energy in vacuum, ϵ_B, from the polaron energies and scaled by ϵ_F, the majority up-spin particle Fermi energy. In Fig. 5 we show the subtracted-scaled polaron energy $(\epsilon_p + |\epsilon_B|)/\epsilon_F$ versus the dimensionless parameter $\eta \equiv \frac{1}{2}\ln\left(\frac{2\epsilon_F}{|\epsilon_B|}\right)$, which characterises the strength of the interaction. The simulations are done with $N = 21$ and $N = 20$ up-spin particles. For comparison we have plotted the diagrammatic Monte Carlo results from Ref. 19 and variational results from Ref. 20. We find good agreement with the data in the weak-coupling region, $\eta > 1$. In the very strong coupling limit $\eta \to -\infty$, one expects $(\epsilon_p + |\epsilon_B|)/\epsilon_F$ to approach -1 from above. This corresponds to a tightly-bound molecule that has pulled one up-spin from the Fermi sea and is only weakly repelled by the remaining up spins. Perhaps most interesting is that the lattice results show a smooth dependence on energy in the intermediate region $-0.90 < \eta < -0.75$. We interpret this as evidence for a smooth crossover from fermionic polaron to bosonic molecule. In order to uncover the underlying nature of the polaron-molecule transition, we have used ILMC to measure the density-density correlation function between the impurity and the majority particles. As discussed in Ref. 17, there is no sign of a sharp phase transition such as a divergence of the correlation function or non-analytic dependence on the interaction strength. This shows that the polaron-molecule transition is a smooth crossover. These interesting results clearly show the power of the ILMC and time is ripe now to apply it to the physics of hyper-nuclei.

5 Outlook: What is next?

At this time, NLEFT has matured into a well-established *ab initio* framework at the forefront of the 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 issues of the mechanism underlying pairing in nuclei, the microscopic origin of the so-important spin-orbit force in nuclei or the precise location of the drip lines of the nuclear chart. A slight modification of the pinhole algorithm, the so-called *trace pinhole algorithm*, allows one to perform first principles calculations in nuclear thermodynamics. Questions of high importance are the precise location of the liquid-gas phase transition in the phase diagram of baryonic matter or the cluster-fragment distribution in the collisions of mid-mass nuclei measured at various facilities world-wide. Scattering processes at stellar energies can also be addressed, in particular the so-called "holy grail" of nuclear astrophysics, *i. e.* the $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$ reaction at stellar energies, is within reach. In short, NLEFT appears headed towards exciting times of progress and discoveries.

Acknowledgements

I thank my NLEFT colleagues for their contributions to the results presented here. This work is supported in part by the DFG (Grant No. TRR110) through the funds provided to the Sino-German CRC 110 "Symmetries and the Emergence of Structure in QCD", by the Chinese Academy of Sciences (CAS) President's International Fellowship Initiative (PIFI) (grant no. 2018DM0034) and by VolkswagenStiftung (grant no. 93562). The computational resources were provided by the Jülich Supercomputing Centre (JSC) at Forschungszentrum Jülich and by RWTH Aachen.

References

1. E. Epelbaum, H. W. Hammer, and U.-G. Meißner, *Modern Theory of Nuclear Forces*, Rev. Mod. Phys. **81**, 1773, 2009.
2. T. A. Lähde and U.-G. Meißner, *Nuclear Lattice Effective Field Theory : An Introduction*, Lect. Notes Phys. **957**, Springer, 2019.
3. S. Elhatisari *et al.*, *Ab initio Calculations of the Isotopic Dependence of Nuclear Clustering*, Phys. Rev. Lett. **119**, 222505, 2017.
4. B. N. Lu, N. Li, S. Elhatisari, D. Lee, E. Epelbaum, and U.-G. Meißner, *Essential elements for nuclear binding*, Phys. Lett. B **797**, 134863, 2019.
5. T. A. Lähde, E. Epelbaum, H. Krebs, D. Lee, U.-G. Meißner, and G. Rupak, *Lattice Effective Field Theory for Medium-Mass Nuclei*, Phys. Lett. B **732**, 110, 2014.
6. S. König, H. W. Grießhammer, H. W. Hammer, and U. van Kolck, *Nuclear Physics Around the Unitarity Limit*, Phys. Rev. Lett. **118**, 202501, 2017.
7. N. Klein, D. Lee and U.-G. Meißner, *Lattice Improvement in Lattice Effective Field Theory*, Eur. Phys. J. A **54**, 233, 2018.
8. M. Wang *et al.*, *The AME2016 atomic mass evaluation (II). Tables, graphs and references*, Chin. Phys. C **41**, 030003, 2017.

9. I. Tews, T. Krüger, K. Hebeler, and A. Schwenk, *Neutron matter at next-to-next-to-next-to-leading order in chiral effective field theory*, Phys. Rev. Lett. **110**, 032504, 2013.
10. A. Akmal, V. R. Pandharipande, and D. G. Ravenhall, *The Equation of state of nucleon matter and neutron star structure*, Phys. Rev. C **58**, 1804, 1998.
11. S. Gandolfi, J. Carlson, and S. Reddy, *The maximum mass and radius of neutron stars and the nuclear symmetry energy*, Phys. Rev. C **85**, 032801, 2012.
12. N. Li, S. Elhatisari, E. Epelbaum, D. Lee, B. Lu, and U.-G. Meißner, *Galilean invariance restoration on the lattice*, Phys. Rev. C **99**, 064001, 2019.
13. D. Frame, R. He, I. Ipsen, D. Lee, D. Lee, and E. Rrapaj, *Eigenvector continuation with subspace learning*, Phys. Rev. Lett. **121**, 032501, 2018.
14. J. Haidenbauer, U.-G. Meißner, and A. Nogga, *Hyperon-nucleon interaction within chiral effective field theory revisited*, accepted for publication in Eur. Phys. J. A, 2019, arXiv:1906.11681 [nucl-th].
15. S. Bour, Master Thesis, University of Bonn, unpublished (2009).
16. S. Elhatisari and D. Lee, *Fermion-dimer scattering using an impurity lattice Monte Carlo approach and the adiabatic projection method*, Phys. Rev. C **90**, 064001, 2014.
17. S. Bour, D. Lee, H.-W. Hammer, and U.-G. Meißner, *Ab initio Lattice Results for Fermi Polarons in Two Dimensions*, Phys. Rev. Lett. **115**, 185301, 2015.
18. N. V. Prokofev and B. V. Svistunov, *Bold diagrammatic Monte Carlo: A generic sign-problem tolerant technique for polaron models and possibly interacting many-body problems*, Phys. Rev. B **77**, 125101, 2008.
19. J. Vlietinck, J. Ryckebusch, and K. Van Houcke, *Diagrammatic Monte Carlo study of the Fermi polaron in two dimensions*, Phys. Rev. B **89**, 085119, 2014.
20. J. Levinsen and M. M. Parish, *Strongly interacting two-dimensional fermi gases*, in Annu. Rev. Cold At. Mol. **3**, World Scientific, Chapter 1, 2015.

Flavour Structure of the Baryon Octet

Gunnar S. Bali and Sara Collins
for the RQCD Collaboration

Institut für Theoretische Physik, Universität Regensburg, 93040 Regensburg, Germany
E-mail: {*gunnar.bali, sara.collins*}*@ur.de*

We investigate aspects of the structure of different baryons via simulations of quantum chromodynamics in lattice regularisation (Lattice QCD). In particular, we study the mass spectrum, (generalised) isovector charges as well as moments of light cone distribution amplitudes. The charges correspond to moments of parton distribution functions (PDFs). Almost all visible matter in the universe consists of nucleons (*i. e.* protons and neutrons), which are also the prime probes for new physics, be it in accelerator experiments or dark matter and neutrino detectors. The results will increase the precision of the relation between experimental cross sections and decay rates and the underlying fundamental theory, which describes interactions on the quarks and gluon level rather than interactions with the nucleons (which are composed of quarks and gluons). By extending the study to so-called hyperons that contain strange quarks, in addition to the up and down quarks of the nucleon, at many different quark mass combinations, the validity range of chiral perturbation theory (ChPT) and quark flavour symmetry relations can be checked and low energy constants predicted.
The simulations are carried out in $N_f = 2 + 1$ QCD, neglecting the mass difference between up and down quarks and the electric charges of the quarks. We employ Coordinated Lattice Simulations (CLS) gauge ensembles. These were generated in a Markov chain using the hybrid Monte Carlo (HMC) algorithm with open boundary conditions in time, on several European supercomputers including JUWELS and JUQUEEN. The Wilson fermion discretisation is used, with non-perturbative removal of lattice spacing effects that are proportional to the lattice constant (order a improvement). The main computational task in the analysis that is carried out on the Xeon-Phi Booster module of JURECA of these gauge ensembles with volumes ranging from $64 \cdot 32^3$ up to $192 \cdot 96^3$ points, encompassing 1000–2000 configurations each, is the multiple solution of sparse large linear systems with a dimension of up to $(2 \cdot 10^9)^2$ complex variables. This is achieved by an adaptive algebraic multigrid algorithm. A novel stochastic method allows us to obtain results for four different baryons and many momentum combinations with little computational overhead, relative to just computing the structure of the nucleon at a few momenta.

1 Introduction

With the discovery of the Higgs boson at the LHC at CERN in 2012 the existence of the last missing particle predicted by the Standard Model was confirmed. Since then (and also before then) the Standard Model, which is the accepted theory of elementary particles and their interactions, successfully passed an impressive number of precision tests, *e. g.* at the LHC, B-factories and low energy β decay experiments. At present only a few observables show tensions with respect to experiment on the level of three standard deviations. Yet we know that the Standard Model in its present form is incomplete. For instance, from cosmological observations one can infer the existence of the so-called Dark Matter that at present has only been detected via its gravitational interaction. This matter is distributed differently in galaxies than one would expect from the particles that are included in the Standard Model and interactions between this matter and Standard Model matter must be very weak. Another issue is the abundance of matter over antimatter in the early universe,

which requires a significant breaking of the so-called CP-symmetry and a first order phase transition, which the Standard Model fails to explain. Finally, there are unresolved questions in the neutrino sector and the puzzle why differences between masses of elementary particles are so large, ranging from the top quark ($m_t \approx 1.7 \cdot 10^{11}$ eV) down to the lightest neutrino ($m_{\nu_1} < 0.9 \cdot 10^{-1}$ eV[1]).

The Standard Model unifies the electromagnetic and weak interactions and contains the theory of strong interactions (QCD) that bind quarks and gluons into hadrons (*e. g.* protons and pions). In order to discover deviations between experiment and Standard Model predictions, a new level of precision is required, and once deviations are seen, theory input will be needed to relate experimental (differential) cross sections and life times to the couplings between fundamental particles. The accuracy of this is often limited by theoretical rather than experimental uncertainties and, in particular, by hadronic uncertainties: almost all visible matter of the universe consists of nucleons, which are also the prime probe for new physics, be it in collider experiments, long baseline neutrino oscillation experiments or direct dark matter detection. These are made up of strongly interacting quarks and gluons, whose interactions are often elusive to analytical approaches and hence many aspects of hadron structure and in particular of nucleon structure are not very well determined. Fits to experimental data, sometimes employing flavour symmetry arguments and/or chiral perturbation theory (ChPT), still dominate our knowledge. As pointed out above, clarifying the structure of nucleons in terms of their quark and gluon components in a rigorous way is of great importance. In principle this can be achieved by *ab initio* Lattice QCD simulations of the underlying theory.

In the last decade Lattice QCD simulations have reached the quark masses, lattice volumes and spacings and statistical precision required to make an impact in this field. Since then discrepancies with phenomenological estimates have shown up. For instance the nucleon σ term, which determines the strength of the coupling of the proton to hypothetical new matter either via Higgs exchange or directly, when extracted from fits to pion nucleon scattering[2, 3] comes out by a factor of almost two larger than when determined directly in lattice simulations close to the physical quark mass point.[4–7] Clearly, the systematics need to be understood better, also on the phenomenological side.

Here, we study several observables related to baryon structure to shed light on these questions. Below we list some of the highlights of our present and ongoing large scale simulation programme.

- We achieve an excellent coverage of the plane spanned by the strange quark mass m_s and the light quark mass $m_\ell = m_u = m_d$, enabling us not only to extrapolate or interpolate to the physical point but also to verify effective field theory predictions and in particular covariant baryon chiral perturbation theory (BChPT).
- We cover a large bracket of lattice spacings reaching down to $a < 0.04$ fm $= 4 \cdot 10^{-17}$ m. Usually, the hybrid Monte Carlo (HMC) and other updating algorithms based on local changes will slow down exponentially with the inverse lattice spacing beyond a value $a \lesssim 0.06$ fm, due to the freezing of the topological charge.[8] Further reducing the lattice constant while retaining ergodicity of the algorithm became possible, employing open boundary conditions in the Euclidean time direction,[9] allowing so-called instantons to flow into and out of the simulated volume.
- We analyse not just the nucleon but all octet baryons, *i. e.* the nucleon (N, quark content $\ell\ell\ell$, where $\ell \in \{u, d\}$) as well as the Σ ($\ell\ell s$), the Λ ($\ell\ell s$) and the Ξ ($\ell s s$). This

is possible by employing a novel stochastic method[10–13] that was efficiently implemented by us on Xeon and Xeon-Phi platforms, employing a very general factorisation approach[14, 15] into a "spectator part" defining properties of the external baryon states and an "insertion part", corresponding to the current with which the baryon is probed.

This article is organised as follow. First we explain our simulation strategy in Sec. 2, before we show results on the quark mass dependence of octet baryon masses in Sec. 3. We then proceed to investigate flavour symmetry breaking effects on isovector charges in Sec. 4 and briefly discuss light cone distribution amplitudes in Sec. 5, before we summarise.

2 Overview of the Simulation

We employ $N_f = 2 + 1$ flavours of non-perturbatively order a improved Wilson fermions and the tree-level Symanzik improved gauge action. For details on the action and the HMC simulation, see Ref. 16. Since that publication more CLS simulation points have been added. The ensembles were generated using OPENQCD[17, 18] within the CLS[19] effort.[a] Some of this was carried out on JUWELS[21] and JUQUEEN.[22] So far six values of the inverse coupling constant $\beta = 6/g^2$ have been realised, along three quark mass plane trajectories. These cover lattice spacings $a = 0.039\,\text{fm}$ ($\beta = 3.85$) up to $a = 0.099\,\text{fm}$ ($\beta = 3.34$) and range from $\{48, 128\} \cdot 24^3$ (A, U ensembles) over $\{64, 96, 128\} \cdot 32^3$ (B, H, S), $\{96, 128\} \cdot 48^4$ (C, N), $\{128, 192\} \cdot 64^3$ (D, J) up to $192 \cdot 96^3$ (E) lattice sites. For an overview, see Fig. 1. The lattice spacing was set using $\sqrt{8t_0^*} = 0.413(6)\,\text{fm}$, with the scale t_0^* defined in Ref. 24. This value was obtained using pion and kaon decay constants. Here, we obtain a consistent but more precise value using baryon masses.

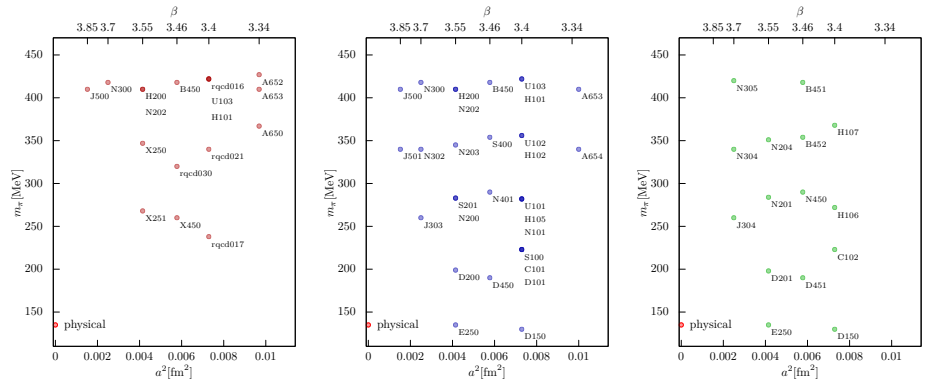


Figure 1. Analysed ensembles: the different quark mass trajectories (left: $m_s = m_\ell$, centre: $m_s + 2m_\ell = \text{const}$, right: $\widehat{m}_s \approx \text{const}$) have been analysed at six (four for $\widehat{m}_s \approx \text{const}$) different lattice spacings.

The naive scaling of the simulation cost is $\propto a^{-4}$ due to the increasing number of points if the size of the four-dimensional physical box is kept fixed. In addition autocorrelation

[a] A few additional $m_\ell = m_s$ ensembles were generated by the RQCD group employing the BQCD-Code.[20]

times will be subject to critical slowing down $\propto a^{-z}$ with $z \approx 2$, further increasing the computational effort towards the continuum limit. However, also reducing the quark mass m_ℓ does not come for free. First of all, finite size effects are suppressed exponentially with the smallest mass gap, which is the pion mass $M_\pi \propto \sqrt{m_\ell}$. If we wish to keep the lattice extent $L \gtrsim 4/M_\pi$ constant in units of the pion mass, then the volume has to be increased in proportion to M_π^{-4}. Moreover, the step size in the HMC algorithm with an Omelyan integrator scales like $L/M_\pi \propto M_\pi^{-3/2}$ while the condition number of the discretised Dirac matrix increases $\propto m_\ell^{-1} \propto M_\pi^{-2}$. The last aspect becomes less severe when employing a multigrid solver[25, 26] (see also Ref. 27) as we do for our analysis of the gauge ensembles on JURECA-Booster.[23] Due to using open boundary conditions in time,[9, 17] indeed we do not observe any severe slowing down, see Fig. 2, except for the finest lattice spacing where at present we are accruing more statistics on JUWELS. In general, autocorrelation times increase towards finer lattices, the exception being our coarsest lattice which is at the limit of what is possible to simulate with our action.

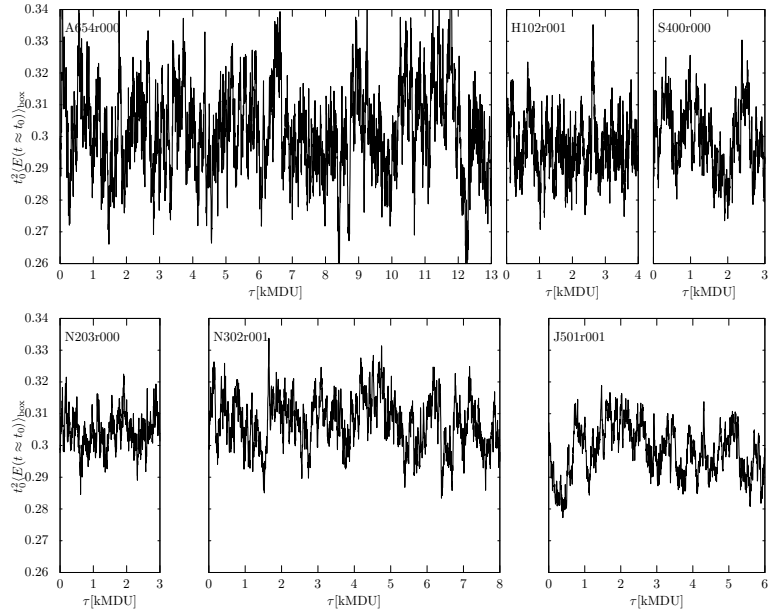


Figure 2. History of the Wilson flow action density,[28] multiplied by t_0^2,[29] after a flow time close to t_0, inside a central sub-volume of approximately $1\,\text{fm} \cdot (aN_s)^3$, along a line of $M_\pi \approx 340\,\text{MeV}$ (for $m_s + 2m_\ell = \text{const}$) from coarse to fine lattice spacings. This quantity has the largest integrated autocorrelation time. The amplitude of the fluctuation varies, *e. g.* due to somewhat different physical volumes. Autocorrelations increase from top left to bottom right, with the exception of A654 at $\beta = 3.34$ where we observe larger autocorrelation times than at $\beta = 3.4$. For the cases where more than one Monte-Carlo chain exists, only one replica is shown.

As mentioned above the ensembles are distributed along three lines in the quark mass plane: along $m_\ell = m_s$, along $2m_\ell + m_s = \text{const}$ and along a line of an approximately constant renormalised strange quark mass.[30] The simulation points for the latter two quark

mass trajectories intersect close to the physical point while the former one approaches the chiral limit $m_s = m_\ell = 0$, which is of particular interest for ChPT which is an expansion in terms of pseudoscalar meson masses about this limit. Some relations in the mesonic sector, in particular the Gell-Mann–Oakes–Renner relation, which implies the approximate proportionality $M_\pi^2 \propto m_\ell$, have been verified to surprising accuracy in lattice simulations, even at quite large pion masses. In the baryonic sector, however, BChPT often does not appear to converge very well, and for some quantities like axial charges or σ terms (*i. e.* scalar charges) it has been suggested that pion masses smaller than 200 MeV may be necessary to make contact with the regime where SU(2) BChPT is of any use.[31] This poses serious questions regarding the validity of SU(3) BChPT since the kaons and η mesons of the real world are much heavier than that. Strong interactions do not distinguish between different quark flavours and flavour symmetry is only broken by quark mass differences and by electroweak interactions, the latter being small corrections. This means that SU(3) BChPT is much more constrained than SU(2) BChPT: while the number of hadronic observables explodes, the number of independent low energy constants (LECs) increases only very moderately since $\delta m = m_s - m_\ell$ is the only flavour symmetry breaking parameter.

3 The Baryon Spectrum

In order to set the lattice scale a, to test BChPT and to determine LECs, we determine the light octet baryon spectrum. The results are shown in Fig. 3. The fit corresponds to the parametrisation

$$m_B(M_\pi, M_K, a) = m_B(M_\pi, M_K, 0)\left[1 + \bar{c}a^2\overline{M}^2 + \delta c_B a^2 \delta M^2\right] \qquad (1)$$

where everything is scaled in units of the parameter $\sqrt{8t_0}$. The squared pseudoscalar mass combinations $\overline{M}^2 = (2M_K^2 + M_\pi^2)/3 \approx 2B_0\overline{m}$ and $\delta M^2 = 2M_K^2 - 2M_\pi^2 \approx 2B_0\delta m$

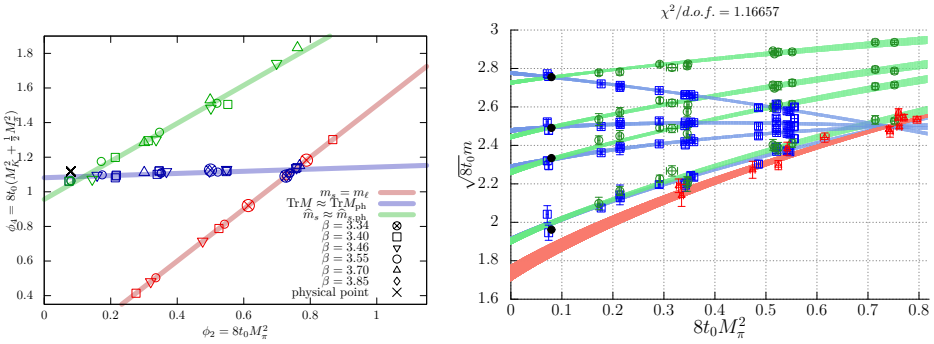


Figure 3. Left: the three lines in the quark mass plane: $m_s = m_\ell$ (red), $m_s + 2m_\ell = \mathrm{const}$ (blue) and $\widehat{m}_s \approx \mathrm{const}$ (green). The horizontal axis is roughly proportional to the light quark mass, the combination on the vertical axis to the sum of quark masses. Right: baryon masses as a function of the squared pion mass: Ξ, Σ, Λ, N from top to bottom. Along the $m_s = m_\ell$ trajectory (red) these are degenerated. The other two sets (blue and green) intersect at the physical pion mass. The scale t_0 was set such that the Ξ mass corresponds to the experimental value (black points).

are proportional to the average quark mass $\overline{m} = (m_s + 2m_\ell)/3$ and the flavour symmetry breaking parameter δm. There is no relation between the four δc_B so that the continuum limit behaviour is encapsulated into six parameters. The quark mass dependence of the continuum limit baryon masses shown in the figure corresponds to NNLO, *i. e.* $\mathcal{O}(p^3)$ BChPT[32] in the EOMS scheme.[33] To this order four parameters are needed to describe this dependence so that in total the combined fit only contains ten parameters. The data points shown are shifted along the fit to $a = 0$ and interpolated to the strange quark masses that exactly correspond to the trajectories. We also carried out a multitude of other fits, with similar results. We find $\sqrt{8t_{0,\mathrm{ph}}} = 0.4128(22)$ MeV, with a systematic error that still needs to be determined. From these fits one can also infer the σ term $\sigma_{\pi N} = 41(2)$ fm, where again the error given is purely statistical for the moment being. Nevertheless, it is clear by now that the total error will be very competitive. The number agrees with earlier lattice QCD determinations[4–7] but disagrees with current phenomenological estimates.[2, 3]

Clearly, the data are in agreement with experiment and well described by the fit. It still needs to be studied to what extent the LECs are universal, *i. e.* if this agreement with the BChPT parametrisation is accidental and higher order terms need to be included. Two LECs appear at order p^2, which corresponds to a linear fit. The curvature is due to the remaining two LECs F and D that appear at $\mathcal{O}(p^3)$ in SU(3) BChPT. These are related to axial charges of octet baryons, for instance the axial charge of the nucleon in the chiral limit reads $\mathring{g}_A = F + D$.

4 Baryon Isovector Charges

Isovector charges are defined as matrix elements of local operators at zero momentum transfer. Here we consider two kinds of (generalised) charges:

$$g_J^B = \langle B|O(\Gamma_J)|B\rangle, \quad J \in \{V, A, T, S\} \tag{2}$$

$$m_B\langle x\rangle_J^B = \langle B|O(\Gamma_J)|B\rangle, \quad J \in \{u^+ - d^+, \Delta u^- - \Delta d^-, \delta u^+ - \delta d^+\} \tag{3}$$

where $q^\pm = q \pm \bar{q}$. The Γ-structures of the latter currents that correspond to the second Mellin moment of parton distribution functions contain one derivative and in both cases we use isovector combinations $O(\Gamma_J) = \bar{u}\Gamma_J u - \bar{d}\Gamma_J d$, for details see Ref. 15.

Assuming SU(3) flavour symmetry, the axial vector charges can be expressed as combinations of just two variables:

$$g_A^N = F + D, \qquad g_A^\Sigma = 2F, \qquad g_A^\Xi = F - D \tag{4}$$

In the chiral limit these correspond to the LECs discussed above. We can directly extract these LECs, extrapolating the combinations $F = (g_A^N + g_A^\Xi)/2$ and $D = (g_A^N - g_A^\Xi)/2$ along the $m_\ell = m_s$ trajectory as illustrated in Fig. 4. This is the only lattice determination so far of these LECs in the SU(3) chiral limit. We ultimately aim at carrying out a combined BChPT fit both to the spectrum and the charges in the continuum limit.

In the case of flavour symmetry, *i. e.* $m_s = m_\ell$, also all the other octet charges g_J can each be parameterised in terms of two parameters F_J and D_J. According to Eq. 4, in this limit $(g_J^N + g_J^\Xi)/g_J^\Sigma = (2F_J)/(2F_J) = 1$ holds. Consequently, We can quantify the SU(3) breaking effect in terms of the parameter

$$\delta_{\mathrm{SU(3)}}^J = \left(g_J^N + g_J^\Xi\right)/g_J^\Sigma - 1 \tag{5}$$

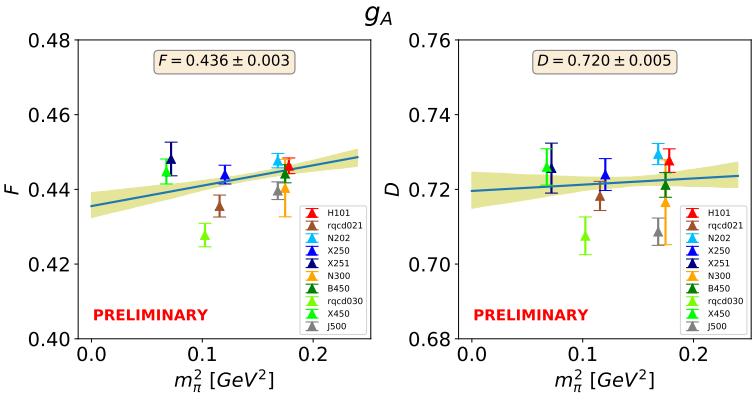


Figure 4. The parameters F and D, extrapolated along the $m_s = m_\ell$ line to the chiral limit. Lattice spacing and finite volume effects are not as yet resolved and the data shown are raw data that have not been shifted to the continuum, infinite volume limit.

It turns out that while these symmetry breaking effects amount to about 10 % at the physical quark mass point for $g_A = \langle \mathbb{1} \rangle_{\Delta u^+ - \Delta d^+}$ and $\langle x \rangle_{\Delta u^- - \Delta d^-}$, these appear to be negligible for g_T and $\langle x \rangle_{u^+ - d^+}$.[15] At present we are carrying out chiral and continuum limit extrapolations similar to those shown for the masses above.

5 Baryon Distribution Amplitudes

Within the parton model description of deep inelastic scattering PDFs can be viewed as probability densities of a particular quark or gluon to carry a fraction x of the longitudinal momentum of the baryon. Radiative QCD corrections and higher twist contributions complicate this simplistic picture. Moments of PDFs, namely the generalised charges introduced above, can be expressed in terms of matrix elements involving local operators which can be computed in Euclidean spacetime using Lattice QCD. PDFs are important in the description of inclusive processes. In contrast, light cone distribution amplitudes (DAs) correspond to probability wavefunctions, where transverse momenta are integrated out. These are needed for the description of exclusive processes and are much harder to constrain experimentally than PDFs. Interestingly, in this case at leading twist the longitudinal momentum fractions of the valence quarks q_1, q_2 and q_3 sum up to 1 ($x_1 + x_2 + x_3 = 1$), *i. e.* effects of the gluons and sea quarks are suppressed by powers of the inverse momentum transfer. Moreover, in this case at large scales an asymptotic shape is approached (in the baryon case $120 x_1 x_2 x_3$). Deviations from this form can be parameterised in terms of the first few moments since higher moments are suppressed due to their larger anomalous dimensions, for details, see Refs. 34, 35. This is quite different from the PDF situation. It turns out that lattice spacing effects are very significant for baryon DAs and, for the first time, we have determined their normalisations and first moments in the continuum limit.[35] The normalisations f^B also play a role in hypothetical proton decay involving operators that convert two quarks into an antiquark. We have reconstructed the three leading twist

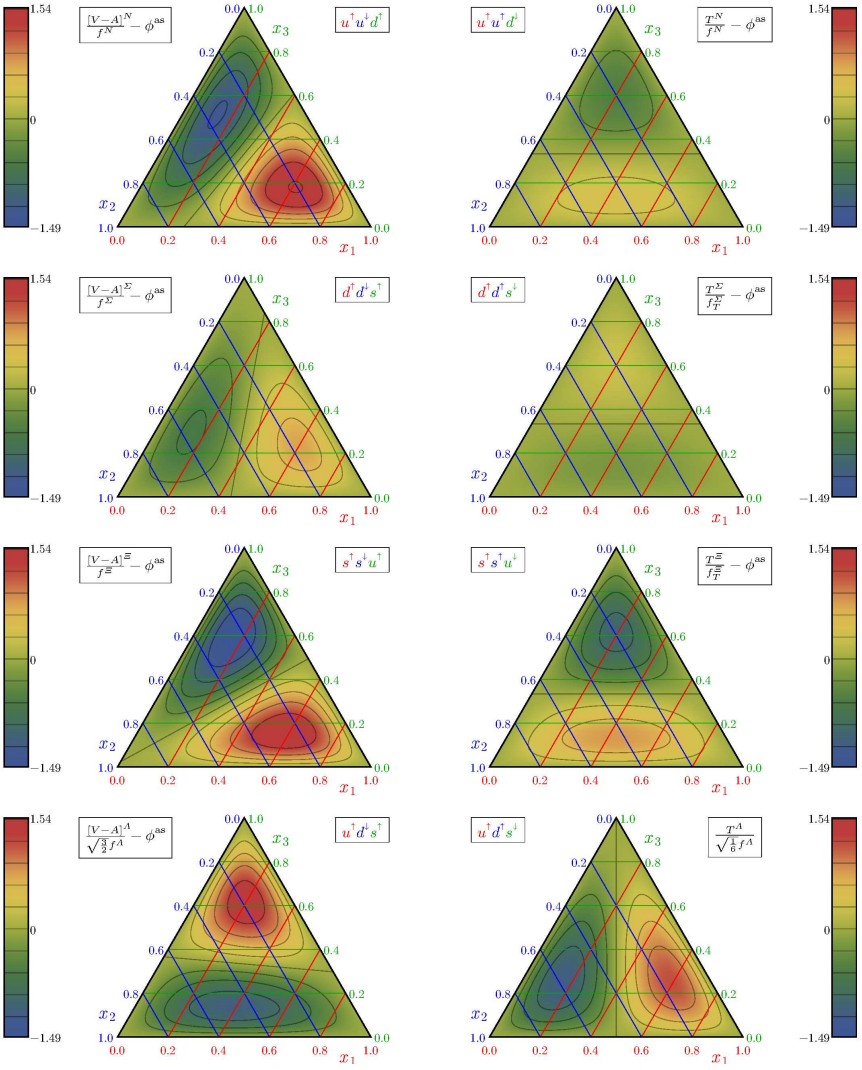


Figure 5. Barycentric figures ($x_1 + x_2 + x_3 = 1$) showing the deviation of the DAs $[V - A]^B$ (left) and T^B (right) from the respective asymptotic limits $120 x_1 x_2 x_3$ and 0, respectively.[35]

DAs V^B, A^B and T^B using the first moments in x_i at a scale $\mu = 2\,\text{GeV}$. Of particular interest are deviations from the (symmetric) asymptotic shape. These are plotted in Fig. 5, where the combinations $V^B - A^B$ and T^B are accompanied by definite helicities of the quarks involved. Within the proton (first row) the "leading" u quark with its helicity aligned to that of the proton carries a larger momentum fraction: $x_1 > x_2 \approx x_3$. This is the only lattice determination of these quantities in the continuum limit.

6 Summary

We were able to obtain for the first time controlled continuum limit results for various baryonic observables. Here we highlighted just a few results of an extensive simulation programme and we did not even comment on the non-perturbative renormalisation and order a improvement that was carried out as well or other technical details.

Acknowledgements

The work of GB is funded by BMBF grant 05P18WRFP1. Support was also provided by the DFG SFB/TRR 55 and the EU ITN EuroPLEx (grant 813942). We thank all our collaborators in RQCD and CLS. We gratefully acknowledge computing time granted by the John von Neumann Institute for Computing (NIC), provided on the Booster partition of the supercomputer JURECA[23] at Jülich Supercomputing Centre. Additional simulations were carried out at the QPACE 3 Xeon Phi cluster of SFB/TRR 55.

References

1. A. Loureiro *et al.*, *On the upper bound of neutrino masses from combined cosmological observations and particle physics experiments*, Phys. Rev. Lett. **123**, 081301, 2019.
2. M. Hoferichter, J. Ruiz de Elvira, B. Kubis, and U. G. Meißner, *High-precision determination of the pion-nucleon σ term from Roy-Steiner equations*, Phys. Rev. Lett. **115**, 092301, 2015.
3. J. Ruiz de Elvira, M. Hoferichter, B. Kubis, and U. G. Meißner, *Extracting the σ-term from low-energy pion-nucleon scattering*, J. Phys. G **45**, 024001, 2018.
4. S. Dürr *et al.*, *Lattice computation of the nucleon scalar quark contents at the physical point*, Phys. Rev. Lett. **116**, 172001, 2016.
5. Y. B. Yang *et al.* [χQCD Collaboration], *πN and strangeness sigma terms at the physical point with chiral fermions*, Phys. Rev. D **94**, 054503, 2016.
6. A. Abdel-Rehim *et al.* [ETM Collaboration], *Direct evaluation of the quark content of nucleons from Lattice QCD at the physical point*, Phys. Rev. Lett. **116**, 252001, 2016.
7. G. S. Bali *et al.* [RQCD Collaboration], *Direct determinations of the nucleon and pion σ terms at nearly physical quark masses*, Phys. Rev. D **93**, 094504, 2016.
8. S. Schaefer *et al.* [ALPHA Collaboration], *Critical slowing down and error analysis in Lattice QCD simulations*, Nucl. Phys. B **845**, 93–119, 2011.
9. M. Lüscher and S. Schaefer, *Lattice QCD without topology barriers*, JHEP **1107**, 036, 2011.
10. R. Evans, G. Bali, and S. Collins, *Improved semileptonic form factor calculations in Lattice QCD*, Phys. Rev. D **82**, 094501, 2010.
11. C. Alexandrou *et al.* [ETM Collaboration], *A stochastic method for computing hadronic matrix elements*, Eur. Phys. J. C **74**, 2692, 2014.
12. G. S. Bali *et al.*, *Nucleon structure from stochastic estimators*, PoS LAT **2013**, 271, 2014.

13. Y. B. Yang, A. Alexandru, T. Draper, M. Gong, and K. F. Liu, *Stochastic method with low mode substitution for nucleon isovector matrix elements*, Phys. Rev. D **93**, 034503, 2016.
14. G. S. Bali *et al.*, *Baryonic and mesonic 3-point functions with open spin indices*, EPJ Web Conf. **175**, 06014, 2018.
15. G. S. Bali *et al.*, *Hyperon couplings from $N_f = 2 + 1$ Lattice QCD*, PoS LAT **2019**, 099, 2019.
16. M. Bruno *et al.*, *Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions*, JHEP **1502**, 043, 2015.
17. M. Lüscher and S. Schaefer, *Lattice QCD with open boundary conditions and twisted-mass reweighting*, Comput. Phys. Commun. **184**, 519, 2013.
18. `http://luscher.web.cern.ch/luscher/openQCD/`
19. `https://wiki-zeuthen.desy.de/CLS/`
20. Y. Nakamura and H. Stüben, *BQCD – Berlin quantum chromodynamics program*, PoS LAT **2010**, 040, 2010.
21. D. Krause, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, JLSRF **5**, A135, 2019.
22. M. Stephan and J. Docter, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, JLSRF **1**, A1, 2015.
23. D. Krause and P. Thörnig, *JURECA: Modular supercomputer at Jülich Supercomputing Centre*, JLSRF **4**, A132, 2018.
24. M. Bruno, T. Korzec, and S. Schaefer, *Setting the scale for the CLS $2 + 1$ flavor ensembles*, Phys. Rev. D **95**, 074504, 2017.
25. S. Heybrock, M. Rottmann, P. Georg, and T. Wettig, *Adaptive algebraic multigrid on SIMD architectures*, PoS LAT **2015**, 036, 2016.
26. P. Georg, D. Richtmann, and T. Wettig, *DD-αAMG on QPACE 3*, EPJ Web Conf. **175**, 02007, 2018.
27. A. Frommer, K. Kahl, S. Krieg, B. Leder, and M. Rottmann, *Adaptive aggregation based domain decomposition multigrid for the lattice Wilson Dirac operator*, SIAM J. Sci. Comput. **36**, A1581–A1608, 2014.
28. M. Lüscher, *Trivializing maps, the Wilson flow and the HMC algorithm*, Commun. Math. Phys. **293**, 899–919, 2010.
29. M. Lüscher, *Properties and uses of the Wilson flow in Lattice QCD*, JHEP **1008**, 071, 2010, [Erratum: JHEP **1403**, 092, 2014].
30. G. S. Bali *et al.* [RQCD Collaboration], *Lattice simulations with $N_f = 2+1$ improved Wilson fermions at a fixed strange quark mass*, Phys. Rev. D **94**, 074501, 2016.
31. G. S. Bali *et al.*, *Nucleon isovector couplings from $N_f = 2$ Lattice QCD*, Phys. Rev. D **91**, 054501, 2015.
32. P. J. Ellis and K. Torikoshi, *Baryon masses in chiral perturbation theory with infrared regularization*, Phys. Rev. C **61**, 015205, 2000.
33. J. Gegelia and G. Japaridze, *Matching heavy particle approach to relativistic theory*, Phys. Rev. D **60**, 114038, 1999.
34. G. S. Bali *et al.*, *Light-cone distribution amplitudes of the baryon octet*, JHEP **1602**, 070, 2016.
35. G. S. Bali *et al.* [RQCD Collaboration], *Light-cone distribution amplitudes of octet baryons from lattice QCD*, Eur. Phys. J. A **55**, 116, 2019.

Scale Setting for QCD with Four Dynamical Quarks

Roman Höllwieser, Francesco Knechtli, and Tomasz Korzec

Department of Physics, Bergische Universität Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany
E-mail: {hoellwieser, korzec}@uni-wuppertal.de, knechtli@physik.uni-wuppertal.de

Monte-Carlo Simulations of QCD with dynamical up, down, strange and charm quarks are carried out on lattices with large volumes and fine lattice spacings using a new massive renormalisation and improvement scheme that is tailored towards situations with a dynamical charm quark. The lattice spacings of the generated ensembles are determined and first physical results, concerning the mass spectrum of mesons built from a charm and anti-charm quark, are presented.

1 Introduction

1.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of strong interactions. The building blocks are quark fields, that come in six different flavours, each with a different mass, and gluon fields that are the force carriers of the strong interactions. Due to the strength of the interactions and the phenomenon of confinement, the approximation that works very well in the electroweak sector, perturbation theory, is only of limited use for QCD, especially at low energies. This is however the energy region where the standard model makes many easily testable predictions. The whole hadron mass spectrum including exotic states, various properties of hadrons, like the magnetic moments of proton and neutron or their charge radii, decay constants, and a plethora of other quantities can in principle be calculated in QCD from first principles. But the only method to carry these calculations out without relying on major uncontrolled approximations is by a numerical evaluation of the QCD path integral by Monte-Carlo methods.

1.2 Lattice QCD

A formulation of QCD that is particularly well suited for numerical methods, is the lattice-regularised Euclidean path integral. A four dimensional piece of space-time of size $T \times L^3$ is discretised. A hypercubic lattice with lattice-spacing a is introduced and quark fields are only defined on the sites of this lattice instead of being continuous functions of the spacetime position. In this formulation predictions are obtained from high dimensional integrals of the form

$$\langle \mathcal{O} \rangle = \int DU \, \frac{1}{Z} e^{-S_g[U]} \det[D_u] \det[D_d] \det[D_s] \det[D_c] \det[D_b] \det[D_t] \, \mathcal{O}[U] \quad (1)$$

The integration is over all gauge field configurations

$$\int DU \equiv \prod_{x_0=0}^{T/a-1} \prod_{x_1,x_2,x_3=0}^{L/a-1} \prod_{\mu=0}^{3} \int dU(x,\mu) \quad (2)$$

where $U(x,\mu)$ is a $SU(3)$ matrix on the link between site $x = (x_0, x_1, x_2, x_3)$ and its neighbour in the positive μ direction, and dU is the Haar measure of $SU(3)$. This amounts to a very high dimensional (*e. g.* around 500×10^6 dimensional for ensemble B of Tab. 1), compact, integral. $S_g[U]$ is the gauge-action. There are many different possibilities to discretise the continuum Yang-Mills action, one particular being the tree-level Symanzik improved action[1]

$$S_g[U] = \frac{1}{g_0^2}\left(\frac{5}{3}\sum_p \mathrm{tr}[1 - U_p] - \frac{1}{12}\sum_r \mathrm{tr}[1 - U_r]\right) \tag{3}$$

where g_0 is the bare gauge coupling constant, p runs over all oriented elementary squares and r over all rectangles of size 1×2 of the lattice. U_p (U_r) are ordered products of the SU(3) gauge field variables around the squares (rectangles). The contributions of the sea-quarks are encoded in the fermionic determinants $\det[D_i]$. The huge ($\approx 190{\cdot}10^6 \times 190{\cdot}10^6$ for ensemble B) matrices D_i are discretised Dirac operators, and again there are many different choices. For clover improved Wilson fermions[2] they are given by

$$D_i[U] = D_\mathrm{w} + m_i + \delta D_\mathrm{v} \tag{4}$$

$$D_\mathrm{w} = \frac{1}{2}\{\gamma_\mu(\nabla_\mu^* + \nabla_\mu) - \nabla_\mu^*\nabla_\mu\} \tag{5}$$

$$\delta D_\mathrm{v}\psi(x) = c_\mathrm{sw}\frac{i}{4}\sigma_{\mu\nu}\hat{F}_{\mu\nu}(x)\psi(x) \tag{6}$$

Here $\gamma_0, \ldots, \gamma_3$ are Dirac matrices, $\sigma_{\mu\nu} = i\gamma_\mu\gamma_\nu$, m_i is the bare mass parameter for quark flavour i and ∇_μ (∇_μ^*) are covariant forward and backward lattice derivative operators with respect to x_μ. $\hat{F}_{\mu\nu}$ is a discretised field strength tensor. See *e. g.* Ref. 3 for an exact definition of all quantities. Z is a normalisation constant, such that $\langle 1 \rangle = 1$, and finally $\mathcal{O}[U]$ is an "observable". For every quantity that one wants to compute, an appropriate observable has to be found, *e. g.* meson masses can be extracted from the t dependence of correlation functions with

$$\mathcal{O}[U] = \sum_{\vec{x},\vec{y}} \mathrm{tr}[\Gamma D_i^{-1}(x_0, \vec{x}; x_0 + t, \vec{y})\Gamma D_j^{-1}(x_0 + t, \vec{y}; x_0, \vec{x})] \tag{7}$$

Here the space-time indices of the Dirac operators are written out explicitly, $x = (x_0, \vec{x})$, and Γ are matrices like *e. g.* $\Gamma = \gamma_5 \equiv \gamma_0\gamma_1\gamma_2\gamma_3$, that select a particular symmetry channel. Note, that the inverse operators are not sparse matrices anymore!

Results of lattice QCD depend on the lattice spacing a and ultimately the continuum limit $a \to 0$ has to be taken (numerically). Due to asymptotic freedom in the continuum limit the bare coupling approaches $g_0 \to 0$. L and T are usually chosen to be large enough, that one is effectively working in the infinite volume limit.

1.3 Algorithms

The factor $p[U] \equiv \frac{e^{-S_g}}{Z}\prod_i \det[D_i]$ is real, positive (for mass-degenerate u and d quarks and heavy enough remaining quarks) and normalised, which makes it a valid probability density function for the distribution of gauge field configurations. This is the perfect starting point for Monte-Carlo methods. If a sequence of gauge field configurations

$U^{(1)}, \ldots, U^{(N)}$ distributed according to p can be generated, the integral in Eq. 1 can be estimated by $\langle \mathcal{O} \rangle = \frac{1}{N} \sum_i \mathcal{O}[U^{(i)}] + O(N^{-1/2})$.

The generation of these configurations proceeds according to a variant of the Hybrid Monte-Carlo (HMC) algorithm.[4] Even-odd and Hasenbusch mass preconditioning[5] are applied to the determinants, before the molecular dynamics equations are integrated using multi-level higher order integrators.[6] Linear systems involving the Dirac operators are delt with using deflated[7] and SAP-preconditioned[8] Krylov space solvers. All simulations are carried out using `openQCD-1.6`.[9]

2 Dynamical Charm Quarks

According to the Appelquist-Carazzone decoupling theorem,[10] heavy quarks have only a minor influence on low energy physics. In practice this means that replacing $\det D_i$ by 1 (a.k.a. quenching the quarks) will only introduce $O(m_i^{-2})$ errors on quantities with energies $E \ll m_i$, as long as the coupling and the remaining quark masses are chosen correctly. This is almost always applied to the heaviest quarks, the bottom and the top quark. The simplified setup spares a lot of tuning efforts, and more importantly makes a multi-scale problem more manageable. The next lighter quark, charm, is a bit of an edge case, and is treated differently by different collaborations. Its impact on low energy observables is believed to be small, at around 2 permille[11] level. At the same time, a dynamical charm quark may introduce large lattice artifacts of order $(am_c)^2$ or even $O(am_c)$ with unimproved or partially improved Wilson fermions. Moreover the costs of the simulations are increased and the tuning of simulation parameters is substantially more difficult. Hence *e. g.* the CLS consortium opted for simulations with a three flavour action.[12] There are however applications, where a dynamical charm quark would be desirable, provided that the increased lattice artifacts can be controlled. One such application is charm physics, where the energies of the quantities of interest are not necessarily smaller than m_c. Although binding energies seem to be reproduced well by the effective theory,[13] it is expected that quantities that depend on so-called disconnected charm quark diagrams are quite sensitive to the presence of a dynamical charm quark. Another is the determination of fundamental parameters of QCD. With a dynamical charm quark one has direct access to the four flavour Λ-parameter, without relying on perturbation theory.

Unimproved Wilson fermions have leading lattice artifacts of $O(a)$. This can be brought down to $O(a^2)$ by implementing the Symanzik improvement programme.[14] The action is augmented by an improvement term[2] that, when its coefficient, $c_{\rm sw}$ in Eq. 6, is tuned correctly, cancels most of the $O(a)$ artifacts. In mass-independent renormalisation schemes, renormalisation and improvement factors do not depend on the quark masses. To achieve full $O(a)$ improvement in such schemes often many additional improvement coefficients need to be determined, that multiply terms proportional to quark masses. Examples are given by the b, $\overline{b}$ and d, $\overline{d}$ coefficients in the expression for the renormalised, improved quark mass[15]

$$\overline{m}_i = Z_m(\tilde{g}_0^2, a\mu)\Bigg[m_{q,i} + (r_m(\tilde{g}_0^2) - 1)\frac{\mathrm{tr}[M_q]}{N_{\rm f}} + a\Bigg\{b_m(g_0^2)m_{q,i}^2 + \overline{b}_m(g_0^2)\mathrm{tr}[M_q]m_{q,i} + (r_m(g_0^2)d_m(g_0^2) - b_m(g_0^2))\frac{\mathrm{tr}[M_q^2]}{N_{\rm f}} + (r_m(g_0^2)\overline{d}_m(g_0^2) - \overline{b}_m(g_0^2))\frac{(\mathrm{tr}[M_q])^2}{N_{\rm f}}\Bigg\}\Bigg] \quad (8)$$

In the past these coefficients were determined approximately in perturbation theory. This could be justified by the smallness of the quark masses. With dynamical charm quarks a non-perturbative determination of these coefficients is called for.[16] Alternatively – and this is what we are exploring in this project – one can depart from massless schemes, and define mass-dependent renormalisation and improvement factors, resulting in a largely simplified $O(a)$ improvement pattern.

$$\overline{m}_i = \tilde{Z}_m(g_0^2, a\mu, aM)\left[m_i - \tilde{m}_{\rm crit}(g_0^2, a{\rm tr}[M_q])\right] \tag{9}$$

The clover coefficient $c_{\rm sw}$ for a physical value of the charm quark mass and degenerate light quark masses corresponding to their average in nature was determined in Ref. 17 to a high accuracy, and here we use this action for the first time in large volume simulations.

3 Scale Setting

The lattice spacing a is not a parameter of the simulation. What can be adjusted, are the dimensionless bare coupling g_0 and the bare quark masses in lattice units am_u, am_d, am_s, am_c, non of which have a physical meaning. The results of a simulation are then (physical) quantities expressed in multiples of the unknown lattice spacing, *e. g.* the masses of pions am_π, kaons am_K, D-mesons am_D and all other hadron masses, or their decay constants af_π *etc.* In principle the mass parameters can be tuned such that a set of ratios of meson masses takes their physical values. One last quantity, *e. g.* f_π can then be used to determine the lattice spacing, *i. e.* to "set the scale", $a = af_\pi/f_\pi^{\rm experimental}$. Simulations at physical masses are extremely expensive and feasible only at rather coarse lattice spacings, if at all. The scale setting therefore is complicated by chiral extrapolations in the light quark masses. For high precision also electro-magnetic effects need to be taken into account. Fortunately some quantities are known to a high precision also at non-physical quark masses, from earlier simulation projects, which can be used to greatly simplify the scale setting procedure. One such quantity is the gradient flow scale t_0[18] at a mass point where $\overline{m}_u = \overline{m}_d = \overline{m}_s$ and $12m_\pi^2 t_0 = 1.11$.

$$\sqrt{8t_0^\star} = 0.413(5)(2)\ {\rm fm} \tag{10}$$

To set the scale with a new action, one only needs to carry out simulations at this unphysical mass point (denoted by the $\star$ symbol), and to determine $t_0^\star/a^2$.

4 Simulations and Results

The goal is to find simulation parameters that correspond to quark masses implicitly given by $m_u = m_d = m_s$ and

$$\phi_4 \equiv 8t_0\left(m_{\rm K}^2 + \tfrac{m_\pi^2}{2}\right) = 1.11 \tag{11}$$

$$\phi_5 \equiv \sqrt{8t_0}\,(m_{\rm D_s} + 2m_{\rm D}) = 11.94 \tag{12}$$

and lattice spacings $a \in \{0.054, 0.041, 0.032\}$ fm. The QCD formulation is a tree-level Symanzik improved gauge action paired with non-perturbatively $O(a)$ improved Wilson fermions, in a massive scheme.[17] Once the correct simulation parameters are found, high

ens.	$\frac{T}{a} \times \frac{L^3}{a^3}$	$6/g_0^2$	$am_{u,d,s}$	am_c	a[fm]	$Lm_\pi^\star$	N_{traj} (MDUs)
A0	96×16^3	3.24	-0.27996	-0.088861	0.054	1.75	1000 (2000)
A1	96×32^3	3.24	-0.27996	-0.088861	0.054	3.5	3908 (7816)
A2	128×48^3	3.24	-0.27996	-0.088861	0.054	5.3	3868 (7736)
B	144×48^3	3.43	-0.32326	-0.17971	0.041	4.4	3030 (6060)
C	192×64^3	-	-	-	0.032	-	0 (0)

Table 1. Summary of finished and planned ensembles.

ens.	N_{ms}	t_0/a^2	$am_{\pi,K}$	am_{D,D_s}	ϕ_4	ϕ_5
A0	500	8.83(23)	0.310(6)	0.614(17)	10.22(90)	15.48(43)
A1	1954	7.43(4)	0.1137(8)	0.5247(7)	1.159(17)	12.168(40)
A2	1934	7.36(3)	0.1107(3)	0.5228(4)	1.087(6)	12.059(20)
B	1515	11.55(6)	0.0910(4)	0.4183(7)	1.148(12)	12.0627(32)

Table 2. Tuning results for the Ensembles available so far.

statistics ensembles are generated and (among other things) $t_0^\star/a^2$ is measured. Tab. 1 summarises the finished and planned ensembles.

Ensembles A0, A1 and A2 have the same lattice spacing and serve as a check of finite volume effects. The result is, that these are negligible for our purposes, as long as $m_\pi L > 4$, see Fig. 1.

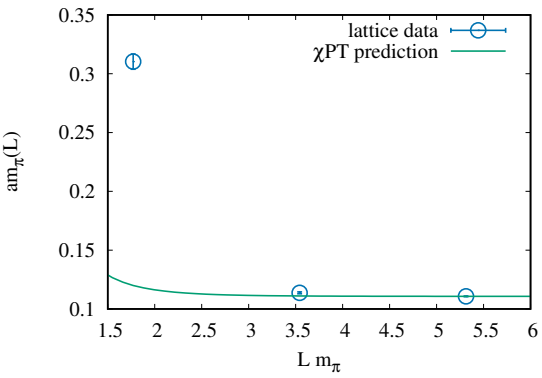


Figure 1. Finite volume effects on the pion mass, from Ensembles A0, A1 and A2.

Although we have not gathered the full statistics of ensemble B yet, we already roughly know the lattice spacing and can judge, how well the novel action copes with lattice artifacts. In Fig. 2 a quantity is plotted, that is supposed to vanish in the continuum limit. The rate at which it does so, is compatible with leading scaling violations of $O(a^2)$, which is very encouraging, given that these are the two coarsest ensembles that we consider.

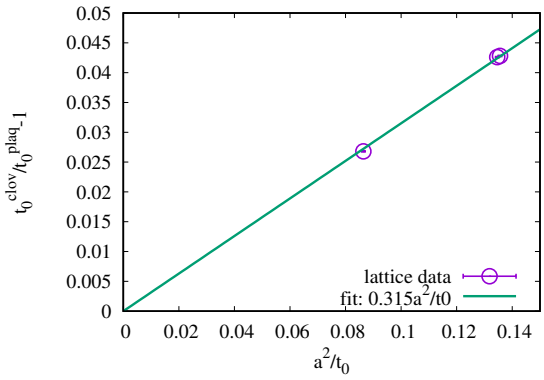


Figure 2. The relative difference of two different definitions of the gradient flow scale t_0, which should vanish in the continuum limit at a rate proportional to a^2.

5 Future Applications

5.1 Charmonium Spectrum

Although the ensembles generated so far are all at a non-physical mass point, the chiral extrapolations are expected to be very flat for quantities that do not contain light (u,d,s) valence quarks. The reason for this is, that the sum of the (renormalised) light quark masses has already the physical value. Therefore the sum of the differences to the physical masses is zero $(\Delta_u + \Delta_d + \Delta_s) = 0$. The derivatives of these quantities with respect to the light quark masses are equal to each other at the mass-symmetric point, so in the expression for the correction between symmetrical and physical mass point, the leading term vanishes, *e. g.* for the mass of the η_c meson

$$m_{\eta_c}^{\text{phys}} = m_{\eta_c} + (\Delta_u + \Delta_d + \Delta_s)\frac{dm_{\eta_c}}{dm_u} + O(\Delta^2) \tag{13}$$

With the fine lattice spacings, our ensembles are very well suited for a study of charmonia and in fact, already at the coarsest lattice, the charmonia masses that we obtain, are very close to their values in nature. One should however note, that we are neglecting disconnected contributions to these masses at the moment. Tab. 3 summarises the findings for ensemble A2.

	η_c	J/ψ	χ_{c0}	χ_{c1}	h_c
am	0.8180(2)	0.8489(2)	0.9398(86)	0.9833(72)	0.9902(81)
m [GeV]	2.9890(7)	3.1019(7)	3.434(31)	3.593(26)	3.618(30)
PDG [GeV]	2.9834(5)	3.096900(6)	3.4148(3)	3.51066(7)	3.52538(11)

Table 3. Masses of charmonium states on ensemble A2 together with their PDG[19] values.

5.2 Strong Coupling

In determinations of the strong coupling constant as done by the ALPHA collaboration, an intermediate result is the Λ-parameter in multiples of a simulated box size L. The L/a values of the finite volume simulations that enter this determination are of course known, but to obtain the physical value of Λ, the lattice spacings need to be known, too. For the planned case of four flavours they are provided by this project.

5.3 Chiral Trajectories

If, at some point, $N_{\mathrm{f}} = 2 + 1 + 1$ simulations of QCD with improved Wilson quarks are desired, this work provides a formidable starting point. The quantities ϕ_4 and ϕ_5 are chosen such, that $\overline{m}_c$ and $\overline{m}_u + \overline{m}_d + \overline{m}_s$ have nearly their physical values. The physical mass point can be approached along chiral trajectories where $\overline{m}_u = \overline{m}_d$ is decreased and $\overline{m}_s$ is increased, while the trace of the quark mass matrix is kept constant. This is a strategy which is particularly well suited for Wilson fermions, and which has been already successfully employed in Ref. 12, 20. No expensive tuning of simulation parameters is required anymore.

6 Conclusions

First Monte-Carlo simulations of QCD with a novel action that is tailored towards simulations including a dynamical charm quark have been carried out on the mass-symmetrical points of relatively fine lattices. The lattice artifacts seem to be relatively mild and more importantly in agreement with $O(a^2)$ scaling, already at the coarsest ensemble. The scale has been set and the ground state masses of a few charmonia have been computed. Already at the coarsest lattice a reasonable agreement with their values in nature is found. All of this encourages us to carry on with the last ensemble with finest lattice spacing.

Acknowledgements

We gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computer time at the supercomputer JUWELS at the Jülich Supercomputing Centre (JSC) under GCS/NIC project ID HWU35. R. H. was supported by the Deutsche Forschungsgemeinschaft in the SFB/TR55.

References

1. M. Lüscher and P. Weisz, *On-Shell Improved Lattice Gauge Theories*, Commun. Math. Phys. **97**, 59, 1985, [Erratum: Commun. Math. Phys. **98**, 433, 1985].
2. B. Sheikholeslami and R. Wohlert, *Improved Continuum Limit Lattice Action for QCD with Wilson Fermions*, Nucl. Phys. B **259**, 572, 1985.
3. M. Lüscher, S. Sint, R. Sommer, and P. Weisz, *Chiral symmetry and O(a) improvement in lattice QCD*, Nucl. Phys. B **478**, 365–400, 1996.

4. S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Hybrid Monte Carlo*, Phys. Lett. B **195**, 216–222, 1987.
5. M. Hasenbusch, *Speeding up the hybrid Monte Carlo algorithm for dynamical fermions*, Phys. Lett. B **519**, 177–182, 2001.
6. 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**, 272–314, 2003.
7. M. Lüscher, *Local coherence and deflation of the low quark modes in lattice QCD*, JHEP **07**, 081, 2007.
8. M. Lüscher, *Solution of the Dirac equation in lattice QCD using a domain decomposition method*, Comput. Phys. Commun. **156**, 209–220, 2004.
9. M. Lüscher and S. Schaefer, *Lattice QCD with open boundary conditions and twisted-mass reweighting*, Comput. Phys. Commun. **184**, 519–528, 2013.
10. T. Appelquist and J. Carazzone, *Infrared Singularities and Massive Fields*, Phys. Rev. D **11**, 2856, 1975.
11. F. Knechtli, T. Korzec, B. Leder, and G. Moir, *Power corrections from decoupling of the charm quark*, Phys. Lett. B **774**, 649–655, 2017.
12. M. Bruno *et al.*, *Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions*, JHEP **02**, 043, 2015.
13. S. Cali, F. Knechtli, and T. Korzec, *How much do charm sea quarks affect the charmonium spectrum?*, Eur. Phys. J. C **79**, 607, 2019.
14. K. Symanzik, *Continuum Limit and Improved Action in Lattice Theories: I. Principles and φ^4 Theory*, Nucl. Phys. B **226**, 187–204, 1983.
15. T. Bhattacharya, R. Gupta, W. Lee, S. R. Sharpe, and J. M. S. Wu, *Improved bilinears in lattice QCD with non-degenerate quarks*, Phys. Rev. D, **73**, 034504, 2006.
16. P. Korcyl and G. S. Bali, *Non-perturbative determination of improvement coefficients using coordinate space correlators in $N_f = 2 + 1$ lattice QCD*, Phys. Rev. D **95**, 014505, 2017.
17. P. Fritzsch, R. Sommer, F. Stollenwerk, and U. Wolff, *Symanzik Improvement with Dynamical Charm: A 3+1 Scheme for Wilson Quarks*, JHEP **06**, 025, 2018.
18. M. Bruno, T. Korzec, and S. Schaefer, *Setting the scale for the CLS $2 + 1$ flavor ensembles*, Phys. Rev. D **95**, 074504, 2017.
19. M. Tanabashi *et al.*, *Review of Particle Physics*, Phys. Rev. D **98**, 030001, 2018.
20. W. Bietenholz *et al.*, *Flavour blindness and patterns of flavour symmetry breaking in lattice simulations of up, down and strange quarks*, Phys. Rev. D **84**, 054509, 2011.

QCD Phase Transition in the Chiral Limit

Olaf Kaczmarek, Frithjof Karsch, Anirban Lahiri, Lukas Mazur, and Christian Schmidt

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

We present a lattice QCD based determination of the chiral phase transition temperature in QCD with two massless (up and down) and one strange quark having its physical mass. We propose and calculate two novel estimators for the chiral transition temperature for several values of the light quark masses, corresponding to Goldstone pion masses in the range of $58\ \mathrm{MeV} \lesssim m_\pi \lesssim 163\ \mathrm{MeV}$. The chiral phase transition temperature is determined by extrapolating to vanishing pion mass using universal scaling analysis. After thermodynamic, continuum and chiral extrapolations we find the chiral phase transition temperature $T_c^0 = 132^{+3}_{-6}$ MeV. We also present some preliminary calculations on interplay of effective $U_A(1)$ restoration and chiral phase transition towards chiral limit.

1 Introduction

For physical values of the two light (up and down) and one heavier (strange) quark masses strongly interacting matter undergoes a crossover from a low temperature hadronic regime to a high temperature region that is best described by quark and gluon degrees of freedom. This smooth crossover between the two asymptotic regimes is not a phase transition.[1] It is characterised by a pseudo-critical temperature, T_{pc}, that has been determined in several numerical studies of Quantum Chromodynamics (QCD).[2–4] A recent determination of T_{pc} using the maximal fluctuations of several chiral observables gave $T_{pc} = (156.5 \pm 1.5)$ MeV.[5] On the contrary to the calculation for physical masses, till date one of the outstanding challenges in QCD thermodynamics is to clarify the nature of the QCD phase transition that exists in the chiral limit, *i. e.* in the limit of vanishing light quark masses, $(m_u, m_d) \to (0, 0)$. While it is widely expected that the chiral phase transition at vanishing values of the two light quark masses is a second-order transition, belonging to the $O(4)$ universality class, a subtle role is played by the $U(1)_A$ axial anomaly.[6] If the $U(1)_A$ symmetry, which is broken in the QCD vacuum, does not get "effectively restored" at high temperature, the transition indeed will be in the universality of 3-d, $O(4)$ spin models. However, if the $U(1)_A$ symmetry breaking effects are small already at the chiral phase transition, at which the chiral condensate vanishes, the phase transition may turn out to be first order,[6] although a second-order transition belonging to larger 3-d universality class[7–10] may become of relevance. If the chiral phase transition is first order then a second order phase transition, belonging to the 3-d $Z(2)$ universality class, would occur for $m_l^c > 0$. When decreasing the light to strange quark mass ratio, $H = m_l/m_s$, towards zero, this would give rise to diverging susceptibilities already for some critical mass ratio $H_c = m_l^c/m_s > 0$.

2 Determination of Chiral Phase Transition Temperature

2.1 Observables

In QCD chiral symmetry is spontaneously broken at low temperatures The corresponding order parameter named chiral condensate is obtained as the derivative of the partition function, $Z(T, V, m_u, m_d, m_s)$, with respect to the quark mass m_f of flavour f,

$$\langle\bar{\psi}\psi\rangle_f = \frac{T}{V}\frac{\partial \ln Z(T, V, m_u, m_d, m_s)}{\partial m_f} \tag{1}$$

In the chiral limit, $m_l \to 0$, the light quark chiral condensate, $\langle\bar{\psi}\psi\rangle_l = (\langle\bar{\psi}\psi\rangle_u + \langle\bar{\psi}\psi\rangle_d)/2$, is an exact order parameter for the chiral phase transition. We take care of additive and multiplicative renormalisation by introducing[3] a combination made out of the light and strange quark chiral condensates,

$$M = 2\left(m_s\langle\bar{\psi}\psi\rangle_l - m_l\langle\bar{\psi}\psi\rangle_s\right)/f_K^4 \tag{2}$$

where $f_K = 156.1(9)/\sqrt{2}$ MeV, is the kaon decay constant, which we use as normalisation constant to define a dimensionless order parameter M. The derivative of M with respect to the light quark masses defines the renormalised chiral susceptibility,

$$\begin{aligned}\chi_M &= m_s(\partial_{m_u} + \partial_{m_d})M|_{m_u=m_d} \\ &= m_s(m_s\chi_l - 2\langle\bar{\psi}\psi\rangle_s - 4m_l\chi_{su})/f_K^4\end{aligned} \tag{3}$$

with $\chi_{fg} = \partial_{m_f}\langle\bar{\psi}\psi\rangle_g$ and $\chi_l = 2(\chi_{uu} + \chi_{ud})$. M and χ_M are renormalisation group invariant quantities when terms proportional to the logarithm of light quark masses can be neglected.

Near a 2^{nd} order phase transition M and χ_M can be described by universal finite-size scaling functions $f_G(z, z_L)$ and $f_\chi(z, z_L)$ where the scaling variables in arguments are defined as $z = t/h^{1/\beta\delta}$ and $z_L = l_0/(Lh^{\nu/\beta\delta})$, with $t = (T - T_c^0)/(t_0T_c^0)$ being the reduced temperature; $h = H/h_0$ with $H = m_l/m_s$ denotes the symmetry breaking field and L being the linear extent of the system, $L \equiv V^{1/3}$. The normalisation constants $t_0,\ h_0$ and l_0 that appear in the definition of the scaling variables are non-universal parameters. Before approaching the chiral limit, one also needs to take a proper thermodynamic limit, $V \to \infty$ of any calculation.

Although close to a critical point at $(z, z_L) = (0, 0)$, M and χ_M can be described by the universal scaling functions, but away from the critical point they also receive contributions from corrections-to-scaling[11, 12] and regular terms. With this we may write

$$\begin{aligned}M &= h^{1/\delta}f_G(z, z_L) + f_{sub}(T, H, L) \\ \chi_M &= h_0^{-1}h^{1/\delta-1}f_\chi(z, z_L) + \tilde{f}_{sub}(T, H, L)\end{aligned} \tag{4}$$

where $f_{sub}(T, H, L)$ and $\tilde{f}_{sub}(T, H, L)$ denotes the above-mentioned sub-leading contributions for M and χ_M respectively.

A commonly defined estimator for pseudo-critical temperature corresponds to the peak in the scaling function $f_\chi(z, z_L)$ for large enough system sizes. Within the scaling regime the peak is located at $z = z_p(z_L)$, which defines T_p,[3]

$$T_p(H, L) = T_c^0\left(1 + \frac{z_p(z_L)}{z_0}H^{1/\beta\delta}\right) + \text{ sub-leading} \tag{5}$$

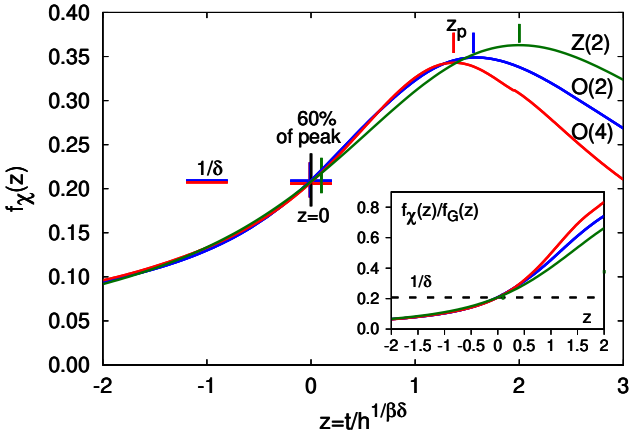


Figure 1. Scaling functions for the 3-d $Z(2)$, $O(2)$ and $O(4)$ universality classes. The position z_p of the peak of the scaling functions (vertical lines) and the position z_{60} where the scaling function attains 60 % of its maximal value (crosses) are shown. Lines close to $z = -1$ show $1/\delta$ for these three universality classes, which agree to better than 1 %. The inset shows the ratio of scaling functions, $f_\chi(z)/f_G(z)$, used in determinations of the chiral phase transition temperature.

with $z_0 = h_0^{1/\beta\delta}/t_0$. The first term accounts for the universal quark mass dependence of T_p. Apart from that contributions from corrections-to-scaling and regular terms will be there, shifting the peak-location of the chiral susceptibilities. These contributions together have been denoted as "sub-leading" in Eq. 5.

Depending on the magnitude of $z_p/z_0 \equiv z_p(0)/z_0$, $T_p(H, L)$ may change significantly with H, when approaching towards the chiral limit.[3] So the contribution from the sub-leading terms will be non-negligible which in turn makes the chiral extrapolation non-trivial. Moreover one has to deal with more non-universal parameters. Thus it will be really advantageous to determine T_c^0 from scaling of such estimators which are defined with $z \simeq 0$. Then by construction $T_p(H, L)$ will have a way milder H dependence which results in rather easier to control $H \to 0$ extrapolation to calculate T_c^0.

Here we consider two such estimators[13, 14] for T_c^0, defined close to or at $z = 0$, in the thermodynamic limit. We define temperatures T_δ and T_{60} through the following conditions:

$$\frac{H\chi_M(T_\delta, H, L)}{M(T_\delta, H, L)} = \frac{1}{\delta} \tag{6}$$

$$\chi_M(T_{60}, H) = 0.6\chi_M^{max} \tag{7}$$

The ratio on the LHS of Eq. 6 has already been introduced in Ref. 15 as a tool to analyse the chiral transition in QCD. T_{60} will correspond to the temperature left to the peak of χ_M; *i. e.* $T_{60} < T_p$. These pseudo-critical temperatures T_X, defined through Eq. 6 and Eq. 7, are already close to T_c^0 for finite H and L^{-1} because they involve $z_X(z_L)$ which either vanishes or stays close to zero in the $L^{-1} \to 0$ limit. To be precise, $z_\delta \equiv z_\delta(0) = 0$ and $z_{60} \equiv z_{60}(0) \simeq 0$. Some values for z_{60}, along with the corresponding scaling functions in the thermodynamic limit have been shown in Fig. 1 for relevant universality classes.

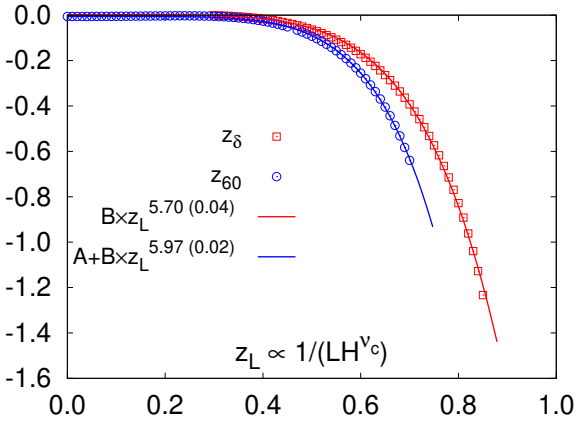


Figure 2. Universal functions $z_\delta(z_L)$ and $z_{60}(z_L)$ calculated from the finite-size scaling functions $f_G(z, z_L)$, $f_\chi(z, z_L)$ determined in Ref. 16.

Ignoring possible contributions from corrections-to-scaling, and keeping in f_{sub} only the leading T independent, infinite volume regular contribution proportional to H, we then find for the pseudo-critical temperatures,[13]

$$T_X(H, L) = T_c^0 \left(1 + \left(\frac{z_X(z_L)}{z_0}\right) H^{1/\beta\delta}\right) + c_X H^{1-1/\delta+1/\beta\delta} \quad , \quad X = \delta,\ 60 \qquad (8)$$

The universal functions, $z_X(z_L)$ may directly be determined from the ratio of the scaling functions $f_\chi(z_\delta, z_L)/f_G(z_\delta, z_L) = 1/\delta$ and $f_\chi(z_{60}, z_L)/f_\chi(z_p, z_L) = 0.6$, respectively. In Fig. 2 we show the calculation of the universal functions $z_X(z_L)$ along with the optimal parameterised form, for the 3-d, $O(4)$ universality class using the finite-size scaling functions $f_G(z, z_L)$, $f_\chi(z, z_L)$ determined in Ref. 16.

We will present here results on T_δ and T_{60} obtained in lattice QCD calculations.[13] We calculated the chiral order parameter M and the chiral susceptibility χ_M (Eqs. 2 and 3) in $(2+1)$-flavour QCD with degenerate up and down quark masses ($m_u = m_d$). Our calculations are performed with the Highly Improved Staggered Quark (HISQ) action[17] in the fermion sector along with the Symanzik improved gluon action. The strange quark mass has been tuned to its physical value[18] and the light quark mass has been varied in a range $m_l \in [m_s/160 : m_s/20]$ corresponding to Goldstone pion masses in the range $58\ \text{MeV} \lesssim m_\pi \lesssim 163\ \text{MeV}$. At each temperature we performed calculations on lattices of size $N_\sigma^3 N_\tau$ for three different values of the lattice cut-off, $aT = 1/N_\tau$, with $N_\tau = 6,\ 8$ and 12. The spatial lattice extent, $N_\sigma = L/a$, has been varied in the range $4 \le N_\sigma/N_\tau \le 8$. For each N_τ we analysed the volume dependence of M and χ_M in order to perform controlled infinite volume extrapolations.

2.2 Results

We show results for χ_M in Fig. 3, on lattices with temporal extent $N_\tau = 8$ for 5 different values of the quark mass ratio, $H = m_l/m_s$, and the largest lattice available for each H.

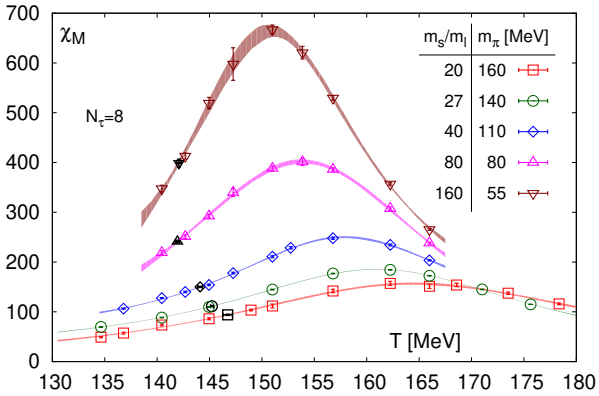


Figure 3. Quark mass dependence of the chiral susceptibility on lattices with temporal extent $N_\tau = 8$ for several values of the light quark masses. The spatial lattices extent N_σ is increased as the light quark mass decreases: $N_\sigma = 32$ $(H^{-1} = 20,\ 27)$, 40 $(H^{-1} = 40)$, 56 $(H^{-1} = 80,\ 160)$. Black symbols mark the points corresponding to 60 % of the peak height. Figure taken from Ref. 13.

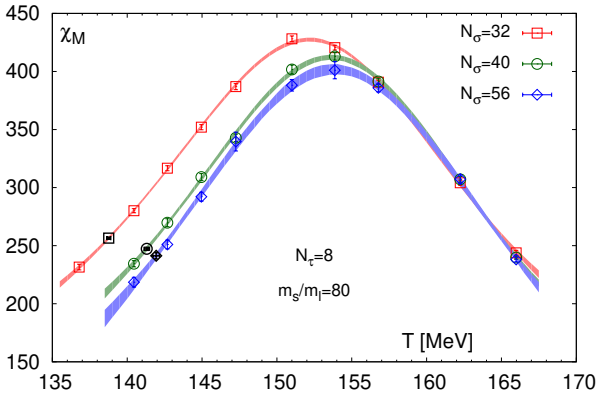


Figure 4. Volume dependence of the chiral susceptibility on lattices with temporal extent $N_\tau = 8$ for three different spatial lattice sizes at $H = 1/80$. Black symbols mark the points corresponding to 60 % of the peak height. Figure taken from Ref. 13.

The increase of the peak height, χ_M^{max}, with decreasing H is consistent with the expected behaviour, $\chi_M^{max} \sim H^{1/\delta - 1} + const.$, with $\delta \simeq 4.8$ within rather large uncertainty which restricts a precise determination of δ.

In Fig. 4 we show the volume dependence of χ_M for $H = 1/80$ on lattices with temporal extent $N_\tau = 8$ and for $N_\sigma/N_\tau = 4,\ 5$ and 7. Similar results have also been obtained for $N_\tau = 6$ and 12. It is important to note that χ_M^{max} decreases slightly with increasing volume, contrary to what one would expect to find at or close to a 1^{st} order phase transition. In fact, this trend seems to be consistent with the behaviour seen for

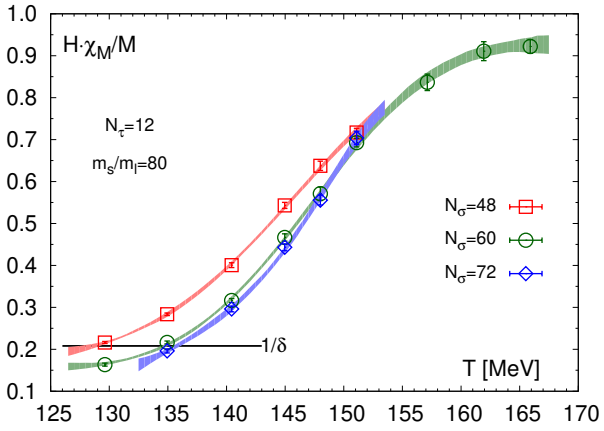


Figure 5. The ratio $H\chi_M/M$ versus temperature for $N_\tau = 12$, $m_l/m_s = 1/80$ and different spatial volumes. Figure taken from Ref. 13.

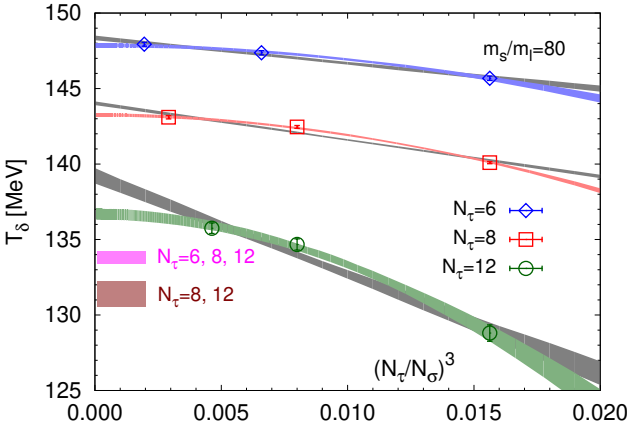


Figure 6. Infinite volume extrapolations based on an $O(4)$ finite-size scaling ansatz (coloured bands) and fits linear in $1/V$ (grey bands). Horizontal bars show the continuum extrapolated results for $H = 1/80$. Figure taken from Ref. 13.

$O(4)$ universality class finite-size scaling functions.[16] Our current results, thus, suggest a continuous phase transition at $H_c = 0$.

Using results for χ_M and M we constructed the ratios $H\chi_M/M$ for different lattice sizes and several values of the quark masses. In Fig. 5 this ratio has been shown on the $N_\tau = 12$ lattices with $H = 1/80$ which is the lowest mass for this N_τ. The intercepts with the horizontal line at $1/\delta$ define $T_\delta(H, L)$. For $H = 1/80$ and each of the N_τ, we have results for three different volumes on which we can extrapolate $T_\delta(H, L)$ to the thermodynamic limit. We performed such extrapolations using (i) the $O(4)$ ansatz given

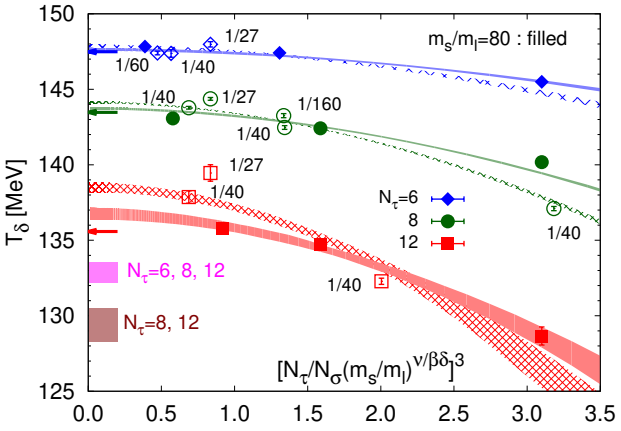


Figure 7. Finite size scaling fits for T_δ based on all data for $H \leq 1/27$ and all available volumes. Arrows show chiral limit results at fixed N_τ and horizontal bars show the continuum extrapolated results for $H = 0$. Figure taken from Ref. 13.

in Eq. 8 which is appropriate when the singular part dominates the partition function and (ii) an extrapolation in $1/V$ which is appropriate if, for large L, the volume dependence predominantly arises from regular terms. In the former case we use the approximation $z_\delta(z_L) \sim z_L^{5.7}$, as shown in Fig. 2. This parameterises well the finite-size dependence of T_δ in the scaling regime. The resulting volume extrapolations are shown in Fig. 6. For fixed H the results tend to approach the infinite volume limit more rapidly than $1/V$, which is in agreement with the behaviour expected from the ratio of finite-size scaling functions. The resulting continuum limit extrapolations in $1/N_\tau^2$ based on data with and without $N_\tau = 6$ are shown as horizontal bars in this figure. A similar analysis is performed for $H = 1/40$. Finally, we extrapolate the continuum results for $T_\delta(H, \infty)$ with $H = 1/40$ and $1/80$ to the chiral limit using Eq. 8 with $z_\delta(0) = 0$. Results obtained from these extrapolation chains, which involve either an $O(4)$ or $1/V$ ansatz for the infinite volume extrapolation, and continuum limit extrapolations performed using all three N_τ or the two largest N_τ, lead to chiral transition temperatures T_c^0 in the range (128-135) MeV. A complete summary of the resulting values for T_c^0 are depicted in Fig. 8.

Since the fits in Fig. 6 suggest that the $O(4)$ scaling ansatz is appropriate to account for the finite volume effects already at finite N_τ, we can attempt a joint infinite volume and chiral extrapolation of all data available for different light quark masses and volumes at fixed N_τ. This utilises the quark mass dependence of finite-size corrections, expressed in terms of z_L. The main difference of this method from the earlier one is that chiral extrapolation is done before taking the continuum limit. Using the scaling ansatz given in Eq. 8, it also allows to account for the contribution of a regular term in a single fit. Fits for fixed N_τ based on this ansatz, using data for all available lattice sizes and $H \leq 1/27$, are shown in Fig. 7. Bands for $H = 1/40$ and $1/80$ are shown in the figure. As can be seen, for $H = 1/80$, these bands compare well with the fits shown in Fig. 6. For each N_τ an arrow shows the corresponding chiral-limit result, $T_\delta(0, \infty)$. We extrapolated these chiral-limit results to the continuum limit and estimated systematic errors again by including or leaving

out data for $N_\tau = 6$. The resulting T_c^0, shown in Fig. 8, are in complete agreement with the corresponding numbers obtained by first taking the continuum limit and then taking the chiral limit. Within the current accuracy these two limits are interchangeable.

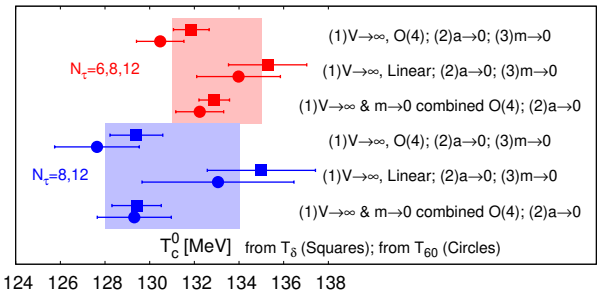


Figure 8. Summary of fit results. The order of different limits taken, described in the main text, is written beside each pair of closest points. For such a pair squares stands for calculations using T_δ and circles stands for calculations using T_{60}. The upper half of the points corresponds to the calculations when continuum extrapolation uses all three N_τ and is presented in red. Similarly the lower half, presented in blue, corresponds to the calculations where only two largest N_τ have been used in the continuum extrapolation. Figure taken from Ref. 13.

Similarly, we analysed results for T_{60} on all data sets following the same strategy as for T_δ. As can be seen in Fig. 8, we find for each extrapolation ansatz that the resulting values for T_c^0 agree to better than 1 % accuracy with the corresponding values from T_δ. This suggests that the chiral susceptibilities used for this analysis reflect basic features of the $O(4)$ scaling functions.

Performing continuum extrapolations by either including or discarding results obtained on the coarsest ($N_\tau = 6$) lattices leads to a systematic shift of about (2-3) MeV in the estimates for T_c^0. This is reflected in the displacement of the two coloured bands in Fig. 8, which show averages for T_c^0 obtained with our different extrapolation ansätze. Averaging separately over results for T_δ and T_{60} obtained with both, continuum extrapolation procedures and including this systematic effect, we find for the chiral phase transition temperature,

$$T_c^0 = 132^{+3}_{-6}\ \mathrm{MeV} \tag{9}$$

3 Towards an Understanding of Anomalous $U_A(1)$ Symmetry Restoration

3.1 Observable

The chiral susceptibility, defined after Eq. 3 receives contributions from a disconnected and connected part which are related to quark-line disconnected and connected Feynman diagrams,[3] *i. e.*

$$\chi_l = \chi_{l,disc} + \chi_{l,conn} \tag{10}$$

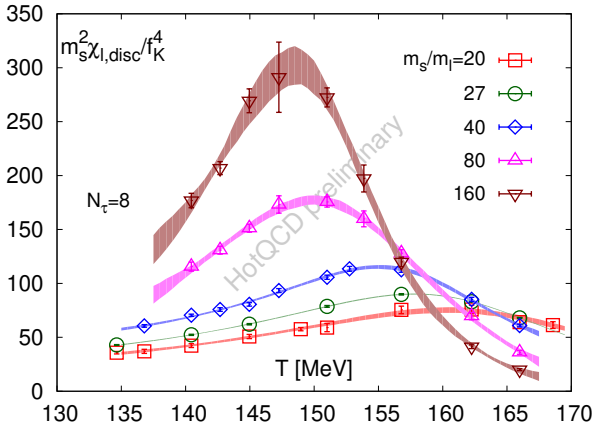


Figure 9. The disconnected chiral susceptibility obtained from calculations in (2+1)-flavour QCD with the HISQ action on lattices of $N_\tau = 8$, at several values of the light quark mass. The susceptibility has been multiplicatively renormalised by multiplying with the square of the strange quark mass and the kaon decay constant f_K has been used to set the scale for the susceptibility as well as the temperature.

with

$$\chi_{l,disc} = \frac{N_f^2}{16 N_\sigma^3 N_\tau} \left\{ \langle \left(\mathrm{Tr} D_l^{-1} \right)^2 \rangle - \langle \mathrm{Tr} D_l^{-1} \rangle^2 \right\} \quad \text{and} \tag{11}$$

$$\chi_{l,conn} = -\frac{N_f}{4 N_\sigma^3 N_\tau} \langle \mathrm{Tr} D_l^{-2} \rangle \tag{12}$$

Here D_l denotes the light quark, staggered fermion Dirac matrix. A crucial role in the analysis of the temperature dependence of the axial anomaly and its overall strength is played by the disconnected chiral susceptibility in the following way.

$U(1)_A$ symmetry relates the susceptibilities of the pion and the scalar iso-triplet delta meson. In a chirally symmetric system one has the degeneracy between pion and iso-scalar meson whose susceptibility is the same as the chiral susceptibility. The disconnected chiral susceptibility, $\chi_{l,disc}$, is directly related to $U(1)_A$ in the following way,

$$\begin{aligned} \chi_\pi - \chi_\delta &= \chi_{l,disc} + (\chi_\pi - \chi_\sigma) \\ &= \chi_{l,disc}, \quad \text{in chirally restored phase} \end{aligned} \tag{13}$$

which says in a chirally symmetric background $\chi_{l,disc}$ vanishes when $U(1)_A$ is effectively restored. This is the strategy we are going to take and here we present some preliminary results in this direction.

3.2 Results

Existing calculations with staggered fermions[19–21] as well as overlap and Möbius domain wall[22–24] and Wilson[25] fermions show evidence for effective restoration of anomalous $U_A(1)$ symmetry above the pseudo-critical temperature T_{pc}, *i. e.* at about $(1.2 - 1.3)T_{pc}$.

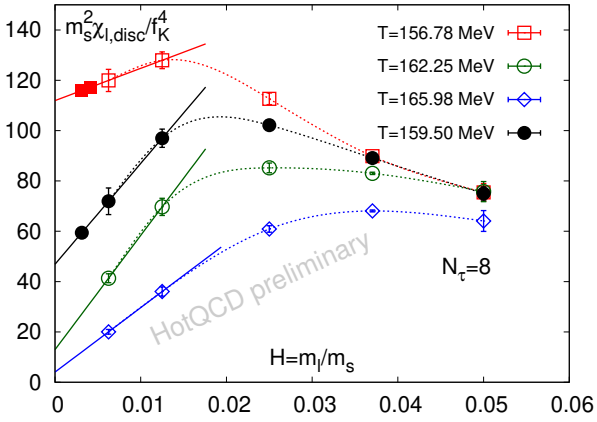


Figure 10. The quark mass dependence of the disconnected chiral susceptibility at three values of the temperature obtained on lattices of size $N_\sigma^3 \times 8$ with $N_\sigma = 32 - 56$. Shown are results for a light (m_l) to strange (m_s) quark mass ratio $1/160 \leq m_l/m_s \leq 1/20$. The open symbols are the data points on which direct calculation is available and solid symbols correspond to those points where interpolation or extrapolation has been used. Dotted lines are drawn to guide the eye and the solid lines are drawn from a linear fit using the points corresponding to $H = 1/80$ and $H = 1/160$.

At non-zero values of the light quark masses the disconnected chiral susceptibility $\chi_{l,disc}(T)$ and the total chiral susceptibility $\chi_m(T)$ share the same qualitative properties, *i. e.* $\chi_{dis}(T)$ has a pronounced peak close to that of the total chiral susceptibility which also rises with decreasing values of the quark mass. When one approaches the chiral limit these two peaks start coming more close to each other and in the chiral limit both total and disconnected chiral susceptibility diverges at T_c^0. In particular, the location of the peak in $\chi_{l,disc}$ has also been used in calculations with chiral fermions (Domain Wall Fermions) to locate the QCD transition temperature.[26] Also in the chiral limit, it generally is expected that $\chi_{l,disc}$ remains non-zero in the high temperature chirally symmetric phase and that it diverges as the chiral phase transition temperature is approached from above. In Fig. 9 we show the quark mass dependence of the disconnected chiral susceptibility obtained from calculations in (2+1)-flavour QCD with the HISQ action on lattices of $N_\tau = 8$, where the above-mentioned properties are apparent. While the figure shows in the vicinity of the transition temperature the expected quark mass dependence, *i. e.* $\chi_{l,disc}$ rises with decreasing quark mass, the figure also shows that at temperatures $T \gtrsim 165$ MeV the quark mass dependence in the chiral susceptibility seems to be inverted. The disconnected chiral susceptibility becomes smaller as the quark mass decreases.

In Fig. 10 we show preliminary results of $\chi_{l,disc}$ as a function of scaled light quark masses for various temperatures which are above chiral critical temperature, obtained on lattices with temporal extent $N_\tau = 8$ and spatial lattice sizes up to $N_\sigma = 56$. Some of these high temperature results, in particular checks on their insensitivity to finite volume effects have been confirmed.[13] One can see at $T \geq 162$ MeV that $\chi_{l,disc}$ almost monotonically decreases and the trend suggests that in the chiral limit within uncertainty it will vanish. Although the situation becomes more involved for $T \leq 160$ MeV when $\chi_{l,disc}$

first increases with decreasing quark mass and attains a maximum and then sharply decreases. Preliminary calculations of Fig. 10 suggest that for a range of temperatures below 160 MeV that $\chi_{l,disc}$ will still go to zero after proper chiral extrapolation but for even lower temperatures chiral extrapolation will give a finite intercept. In other words, from Fig. 10, it seems to be quite unlikely that $\chi_{l,disc}$ will vanish in chiral limit down to $T = T_c^0$. More calculations are needed to confirm this.

4 Conclusions

Based on two novel estimators, we have calculated the chiral phase transition temperature in QCD with two massless light quarks and a physical strange quark. Eq. 9 lists our thermodynamic-, continuum- and chiral- extrapolated result for the chiral phase transition temperature, which is about 25 MeV smaller than the pseudo-critical (crossover) temperature, T_{pc} for physical values of the light and strange quark masses. Preliminary calculations of disconnected chiral susceptibility suggests that in $U_A(1)$ the symmetry remains broken at chiral phase transition.

Acknowledgements

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) through the grant 315477589-TRR 211, the grant 05P18PBCA1 of the German Bundesministerium für Bildung und Forschung, grant 283286 of the European Union, the National Natural Science Foundation of China under grant numbers 11775096 and 11535012, Furthermore, this work was supported through Contract No. DE-SC0012704 with the U.S. Department of Energy, through the Scientific Discovery through Advanced Computing (SciDAC) program funded by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research and Nuclear Physics and the DOE Office of Nuclear Physics funded BEST topical collaboration, and a Early Career Research Award of the Science and Engineering Research Board of the Government of India. Numerical calculations have been made possible through PRACE grants at CSCS, Switzerland, and at the Gauss Centre for Supercomputing and NIC-Jülich, Germany as well as grants at CINECA, Italy. These grants provided access to resources on Piz Daint at CSCS, at JUWELS at NIC as well as on Marconi at CINECA. Additional calculations have been performed on GPU clusters of USQCD, at Bielefeld University, the PC^2 Paderborn University and the Nuclear Science Computing Center at Central China Normal University, Wuhan, China. Some data sets have also partly been produced at the TianHe II Supercomputing Center in Guangzhou.

References

1. For a recent review see: H. T. Ding, F. Karsch, and S. Mukherjee, *Thermodynamics of strong-interaction matter from Lattice QCD*, Int. J. Mod. Phys. E **24**, 1530007, 2015, arXiv:1504.05274 [hep-lat].
2. Y. Aoki, S. Borsanyi, S. Durr, Z. Fodor, S. D. Katz, S. Krieg, and K. K. Szabo, *The QCD transition temperature: results with physical masses in the continuum limit II.*, JHEP **0906**, 088, 2009, arXiv:0903.4155 [hep-lat].

3. A. Bazavov *et al.*, *The chiral and deconfinement aspects of the QCD transition*, Phys. Rev. D **85**, 054503, 2012, arXiv:1111.1710 [hep-lat].
4. C. Bonati, M. D'Elia, M. Mariti, M. Mesiti, F. Negro, and F. Sanfilippo, *Curvature of the chiral pseudocritical line in QCD: Continuum extrapolated results*, Phys. Rev. D **92**, 054503, 2015, arXiv:1507.03571 [hep-lat].
5. A. Bazavov *et al.* [HotQCD Collaboration], *Chiral crossover in QCD at zero and non-zero chemical potentials*, Phys. Lett. B **795** 15, 2019, arXiv:1812.08235 [hep-lat].
6. R. D. Pisarski and F. Wilczek, *Remarks on the Chiral Phase Transition in Chromodynamics*, Phys. Rev. D **29**, 338, 1984.
7. A. Butti, A. Pelissetto, and E. Vicari, JHEP **0308**, 029, 2003.
8. M. Grahl and D. H. Rischke, *Functional renormalization group study of the two-flavor linear sigma model in the presence of the axial anomaly*, Phys. Rev. D **88**, 056014, 2013, arXiv:1307.2184 [hep-th].
9. A. Pelissetto and E. Vicari, *Relevance of the axial anomaly at the finite-temperature chiral transition in QCD*, Phys. Rev. D **88**, 105018, 2013, arXiv:1309.5446 [hep-lat].
10. T. Sato and N. Yamada, *Linking $U(2) \times U(2)$ to $O(4)$ model via decoupling*, Phys. Rev. D **91**, 034025, 2015, doi:10.1103/PhysRevD.91.034025, arXiv:1412.8026 [hep-lat].
11. M. Hasenbusch, *Eliminating leading corrections to scaling in the three-dimensional O(N) symmetric ϕ^4 model: N=3 and 4*, J. Phys. A **34**, 8221, 2001, arXiv:cond-mat/0010463].
12. J. Engels, S. Holtmann, T. Mendes, and T. Schulze, *Equation of state and Goldstone mode effects of the three-dimensional O(2) model*, Phys. Lett. B **492**, 219, 2000, arXiv:hep-lat/0006023.
13. H. T. Ding *et al.* [HotQCD Collaboration], *Chiral Phase Transition Temperature in (2+1)-Flavor QCD*, Phys. Rev. Lett. **123**, 062002, 2019, arXiv:1903.04801 [hep-lat].
14. H.-T. Ding, P. Hegde, F. Karsch, A. Lahiri, S.-T. Li, S. Mukherjee, and P. Petreczky, *Chiral phase transition of (2+1)-flavor QCD*, Nucl. Phys. A **982**, 211, 2019, arXiv:1807.05727 [hep-lat].
15. F. Karsch and E. Laermann, *Susceptibilities, the specific heat and a cumulant in two flavor QCD*, Phys. Rev. D **50**, 6954, 1994, arXiv:hep-lat/9406008.
16. J. Engels and F. Karsch, *Finite size dependence of scaling functions of the three-dimensional O(4) model in an external field*, Phys. Rev. D **90**, 014501, 2014, arXiv:1402.5302 [hep-lat].
17. E. Follana *et al.* [HPQCD and UKQCD Collaborations], *Highly improved staggered quarks on the lattice, with applications to charm physics*, Phys. Rev. D **75**, 054502, 2007, arXiv:hep-lat/0610092.
18. For details on the tuning of quark masses, determination of lines of constant physics and scale setting using the kaon decay constant used in our calculations see: A. Bazavov *et al.* [HotQCD Collaboration], *Equation of state in (2+1)-flavor QCD*, Phys. Rev. D **90**, 094503, 2014, arXiv:1407.6387 [hep-lat].
19. A. Bazavov *et al.*, *Meson Screening Masses in (2+1)-Flavor QCD*, arXiv:1908.09552 [hep-lat].
20. M. Cheng *et al.*, *Meson screening masses from lattice QCD with two light and the strange quark*, Eur. Phys. J. C **71**, 1564, 2011, doi:10.1140/epjc/s10052-011-1564-y, arXiv:1010.1216 [hep-lat].

21. H. Ohno, U. M. Heller, F. Karsch, and S. Mukherjee, $U_A(1)$ *breaking at finite temperature from the Dirac spectrum with the dynamical HISQ action*, PoS LAT **2012** , 095, 2012, doi:10.22323/1.164.0095, arXiv:1211.2591 [hep-lat].
22. M. I. Buchoff *et al.*, *QCD chiral transition,* $U(1)_A$ *symmetry and the dirac spectrum using domain wall fermions*, Phys. Rev. D **89**, 054514, 2014, doi:10.1103/PhysRevD.89.054514, arXiv:1309.4149 [hep-lat].
23. A. Bazavov *et al.* [HotQCD Collaboration], *The chiral transition and* $U(1)_A$ *symmetry restoration from lattice QCD using Domain Wall Fermions*, Phys. Rev. D **86**, 094503, 2012, doi:10.1103/PhysRevD.86.094503, arXiv:1205.3535 [hep-lat].
24. A. Tomiya, G. Cossu, S. Aoki, H. Fukaya, S. Hashimoto, T. Kaneko, and J. Noaki, *Evidence of effective axial U(1) symmetry restoration at high temperature QCD*, Phys. Rev. D **96**, 034509, 2017, doi:10.1103/PhysRevD.96.034509, [Addendum: Phys. Rev. D **96**, 079902, 2017, doi:10.1103/PhysRevD.96.079902], arXiv:1612.01908 [hep-lat].
25. B. B. Brandt, A. Francis, H. B. Meyer, O. Philipsen, D. Robaina, and H. Wittig, *On the strength of the* $U_A(1)$ *anomaly at the chiral phase transition in* $N_f = 2$ *QCD*, JHEP **1612**, 158, 2016, doi:10.1007/JHEP12(2016)158, arXiv:1608.06882 [hep-lat].
26. T. Bhattacharya *et al.*, *QCD Phase Transition with Chiral Quarks and Physical Quark Masses*, Phys. Rev. Lett. **113**, 082001, 2014, doi:10.1103/PhysRevLett.113.082001, arXiv:1402.5175 [hep-lat].

Higher Order Fluctuations and Correlations of Conserved Charges from Lattice QCD

Szabolcs Borsanyi[1], **Zoltan Fodor**[1,2,3], **and Jana N. Guenther**[1,4]

[1] University of Wuppertal, 42119 Wuppertal, Germany
E-mail: fodor@physik.uni-wuppertal.de

[2] Eötvös University, 1117 Budapest, Hungary

[3] Jülich Supercomputing Centre, 52425 Jülich, Germany

[4] University of Regensburg, 93053 Regensburg, Germany

We calculate several diagonal and non-diagonal fluctuations of conserved charges in a system of 2+1+1 quark flavours with physical masses, on a lattice with size $48^3 \times 12$. Higher order fluctuations at $\mu_B = 0$ are obtained as derivatives of the lower order ones, simulated at imaginary chemical potential. From these correlations and fluctuations we construct ratios of net-baryon number cumulants as functions of temperature and chemical potential, which satisfy the experimental conditions of strangeness neutrality and proton/baryon ratio. Our results qualitatively explain the behaviour of the measured cumulant ratios by the STAR collaboration.

1 Introduction

One of the most challenging goals in the study of Quantum Chromodynamics (QCD) is a precise mapping of the phase diagram of strongly interacting matter. First principle, lattice QCD simulations predict that the transition from hadrons to deconfined quarks and gluons is a smooth crossover, taking place in the temperature range $T \simeq 145 - 165$ MeV. Lattice simulations cannot presently be performed at finite density due to the sign problem, thus leading to the fact that the QCD phase diagram is still vastly unexplored when the asymmetry between matter and antimatter becomes large.

With the advent of the second Beam Energy Scan (BES-II) at the Relativistic Heavy Ion Collider (RHIC), scheduled for 2019-2020, there is a renewed interest in the heavy ion community towards the phases of QCD at moderate-to-large densities. A rich theoretical effort is being developed in support of the experimental program; several observables are being calculated, in order to constrain the existence and location of the QCD critical point and to observe it experimentally.

Fluctuations of conserved charges (electric charge Q, baryon number B and strangeness S) are among the most relevant observables for the finite-density program for several reasons. One possible way to extend lattice results to finite density is to perform Taylor expansions of the thermodynamic observables around chemical potential $\mu_B = 0$. Recent results can be found in Ref. 1. Fluctuations of conserved charges are directly related to the Taylor expansion coefficients of such observables, thus, they are needed to extend first principle approaches to the regions of the phase diagram relevant to RHIC. An other popular method to extend observables to finite density is the analytical continuation from imaginary chemical potentials. The agreement between the analytical continuation and Taylor expansion was shown for the transition temperature by Bonati *et al.* in Ref. 2.

Fluctuations can also be measured directly, and a comparison between theoretical and experimental results allows to extract the chemical freeze-out temperature T_f and chemical potential μ_{Bf} as functions of the collision energy. Such fluctuations have been recently calculated and extrapolated using the Taylor method in Ref. 3. Finally, higher order fluctuations of conserved charges are proportional to powers of the correlation length and are expected to diverge at the critical point, thus providing an important signature for its experimental detection.[4]

In this project, we calculate several diagonal and non-diagonal fluctuations of conserved charges up to sixth-order and give estimates for higher orders, in the temperature range 135 MeV $\leq T \leq$ 220 MeV, for a system of 2+1+1 dynamical quarks with physical masses and lattice size $48^3 \times 12$. We simulate the lower-order fluctuations at imaginary chemical potential and extract the higher order fluctuations as derivatives of the lower order ones at $\mu_B = 0$. This method has been successfully used in the past and proved to lead to a more precise determination of the higher order fluctuations, compared to their direct calculation.[5,6] The direct method (see *e. g.* Ref. 7) requires the evaluation of several terms and is affected by a signal-to-noise ratio which is decreasing as a power law of the spatial volume V, with an exponent that grows with the order of the susceptibility.

We also construct combinations of these diagonal and non-diagonal fluctuations in order to study the ratio of the cumulants of the net-baryon number distribution as functions of temperature and chemical potential by means of their Taylor expansion in powers of μ_B/T. We discuss their qualitative comparison with the experimental results from the STAR collaboration, as well as the validity of the truncation of the Taylor series.

2 Fluctuations and Imaginary Chemical Potentials

The chemical potentials are implemented on a flavour-by-flavour basis, their relation to the phenomenological baryon (B), electric charge (Q) and strangeness (S) chemical potentials are given by

$$\mu_u = \frac{1}{3}\mu_B + \frac{2}{3}\mu_Q,\ \mu_d = \frac{1}{3}\mu_B - \frac{1}{3}\mu_Q,\ \mu_s = \frac{1}{3}\mu_B - \frac{1}{3}\mu_Q - \mu_S \tag{1}$$

The observables we are looking at are the derivatives of the free energy with respect to the chemical potentials. Since the free energy is proportional to the pressure, we can write:

$$\chi^{B,Q,S}_{i,j,k} = \frac{\partial^{i+j+k}(p/T^4)}{(\partial\hat{\mu}_B)^i(\partial\hat{\mu}_Q)^j(\partial\hat{\mu}_S)^k} \tag{2}$$

with $\hat{\mu}_i = \frac{\mu_i}{T}$. These are the generalised fluctuations we calculated around $\mu = 0$ in our previous work.[8]

The fermion determinant $\det M(\mu)$ is complex for real chemical potentials, prohibiting the use of traditional simulation algorithms. For imaginary μ, however, the determinant stays real. The chemical potential is introduced through weighted temporal links in the staggered formalism: $U_0(\mu) = e^{\mu}U_0$ and $U_0^\dagger(\mu) = e^{-\mu}U_0^\dagger$. Thus, an imaginary μ translates into a phase factor for the antiperiodic boundary condition in the Dirac operator. Due to the $Z(3)$ symmetry of the gauge sector, there is a non-trivial periodicity in the imaginary quark chemical potential $\mu_q \to \mu_q + i(2\pi/3)T$, which translates to the baryochemical potential as $\mu_B \to \mu_B + i2\pi T$, the Roberge-Weiss symmetry. This is independent of the charge conjugation symmetry $\mu_B \leftrightarrow -\mu_B$.

At $\mu_B = i\pi T$ there is a first order phase transition at all temperatures above the Roberge-Weiss critical end point T_{RW}.[9] When μ_B crosses $i\pi T$ in the imaginary direction, the imaginary baryon density is discontinuous. At low temperature the Hadron Resonance Gas model predicts $\langle B\rangle \sim \sinh(\mu_B/T)$, thus for imaginary values we expect a sine function below T_c: $\mathrm{Im}\langle B\rangle \sim \sin(\mathrm{Im}\mu_B/T)$. At temperatures slightly above T_c, we observe that further Fourier components appear in addition to $\sin(\mathrm{Im}\mu_B/T)$ with alternating coefficients, these are consistent with a repulsive interaction between baryons.[10] At very high temperatures, on the other hand, $\langle B\rangle$ is a polynomial of μ_B since the diagrams contributing to its $\sim \mu_B^5$ and higher order components are suppressed by asymptotic freedom.[11] The Stefan-Boltzmann limit is non-vanishing only for two Taylor coefficients of $\mathrm{Im}\,\langle B\rangle$, giving $\mathrm{Im}\langle B\rangle|_{\mu_B/T=i\pi-\epsilon} = 8\pi/27$. At finite temperatures above T_{RW} this expectation value is smaller but positive, which implies a first order transition at $\mu_B = i\pi T$.

We have only the range $\mu/T \in [0, i\pi)$ to explore the μ-dependence of the observables. Recent simulations in this range include the determination of the transition line, where the slope was determined on the negative side of the $T - \mu_B^2$ phase diagram. Using analyticity arguments, this coefficient gives the curvature of the transition line on the real $T - \mu_B$ phase diagram.[12–14] Apart from the transition temperature, we used imaginary chemical potentials also to extrapolate the equation of state to real μ_B,[5] which serves as an alternative approach to the Taylor extrapolation.[15] In a recent study D'Elia *et al.* have used the low order fluctuations at imaginary chemical potentials to calculate generalised quark number susceptibilities.[6]

3 Analysis Details

3.1 Lattice Setup

In this work we calculate high order fluctuations by studying the imaginary chemical potential dependence of various generalised quark number susceptibilities.

We use a tree-level Symanzik improved gauge action, with four times stout smeared ($\rho = 0.125$) staggered fermions. We simulate $2+1+1$ dynamical quarks, where the light flavours are tuned in a way to reproduce the physical pion and kaon masses and we set $\frac{m_c}{m_s} = 11.85$.[16] For the zero-temperature runs that we used for the determination of the bare masses and the coupling, the volumes satisfy $Lm_\pi > 4$. The scale is determined via f_π. More details on the scale setting and lattice setup can be found in Ref. 8.

Our lattice ensembles are generated at eighteen temperatures in the temperature range 135…220 MeV. We simulate at eight different values of imaginary μ_B given as: $\mu_B^{(j)} = iT\frac{j\pi}{8}$ for $j \in \{0, 1, 2, 3, 4, 5, 6, 7\}$. In this work the analysis is done purely on a $48^3 \times 12$ lattice, we leave the continuum extrapolation for future work.

In terms of quark chemical potentials we generate ensembles with $\mu_u = \mu_d = \mu_s = \mu_B/3$. In each simulation point we calculate all derivatives in Eq. 2 up to fourth order. Thanks to our scan in $\mathrm{Im}\,\hat{\mu}_B$, we can calculate additional μ_B derivatives. Ref. 6 uses various "trajectories" in the $\mu_B - \mu_Q - \mu_S$ space, allowing the numerical determination of higher *e. g.* μ_Q and μ_S derivatives. We find a relatively good signal for the μ_Q and μ_S derivatives by directly evaluating Eq. 2 within one simulation. We recently summarised the details of the direct calculation in Ref. 8.

3.2 Correlated Fit with Priors

We start with the analysis for $\chi_2^B(T)$, $\chi_4^B(T)$ and $\chi_6^B(T)$. Our goal is to calculate these quantities at zero chemical potential, using the imaginary chemical potential data up to $\chi_B^4(T, \hat{\mu}_B)$. In this work we extract these derivatives at a fixed temperature. Results for different temperatures are obtained completely independently, an interpolation in temperature is not necessary at any point. Thus, the error bars in our results plot will be independent. The errors between the quantities $\chi_2^B(T)$, $\chi_4^B(T)$ and $\chi_6^B(T)$ will be highly correlated, though, since these are extracted through the same set of ensembles at the given temperature. This correlation will be taken into account when combined quantities are calculated, or when an extrapolation to real chemical potential is undertaken.

Thus we consider the ensembles at a fixed temperature T. For each value of imaginary $\mu_B \neq 0$ we determine χ_1^B, χ_2^B, χ_3^B and χ_4^B from simulation, while for $\mu_B = 0$ only χ_2^B and χ_4^B can be used, since χ_1^B and χ_3^B are odd functions of μ_B and therefore equal to zero.

We make the following ansatz for the pressure:

$$\chi_0^B(\hat{\mu}_B) = c_0 + c_2\hat{\mu}_B^2 + c_4\hat{\mu}_B^4 + c_6\hat{\mu}_B^6 + c_8\hat{\mu}_B^8 + c_{10}\hat{\mu}_B^{10} \tag{3}$$

where the Taylor expansion coefficients c_n are related to the baryon number fluctuations χ_n^B by: $n!c_n = \chi_n^B$. Our data do not allow for an independent determination of c_8 and c_{10}. Nevertheless, in order to have some control over these terms we make assumptions on the higher order terms. A detailed discussion on our assumptions can be found in Ref. 17 In the fit function we keep the terms up to $\chi_{10}^B = 10!c_{10}$. Without this term the statistical errors on χ_8^B were clearly smaller, but the fit would be less controlled. As the highest order in the function, the resulting χ_{10}^B probably contains severe contamination from even higher order terms. For this reason, and since we fit χ_{10}^B with large statistical errors we do not give results on that quantity. For simplicity, we use the same prior distribution for χ_{10}^B/χ_4^B as for χ_8^B/χ_4^B.

We can then rewrite our ansatz as

$$\chi_0^B(\hat{\mu}_B) = c_0 + c_2\hat{\mu}_B^2 + c_4\hat{\mu}_B^4 + c_6\hat{\mu}_B^6 + \frac{4!}{8!}c_4\epsilon_1\hat{\mu}_B^8 + \frac{4!}{10!}c_4\epsilon_2\hat{\mu}_B^{10} \tag{4}$$

where ϵ_1 and ϵ_2 are drawn randomly from a normal distribution with mean -1.25 and variance 2.75. We use the same distribution for all temperatures. In effect, our c_8 and c_{10} coefficients are stochastic variables. The used distribution for $\epsilon_{1,2}$ actually implements a prior for χ_8^B and χ_{10}^B.

For this ansatz we calculate the following derivatives, which are the actually simulated lattice observables:

$$\chi_1^B(\hat{\mu}_B) = 2c_2\hat{\mu}_B + 4c_4\hat{\mu}_B^3 + 6c_6\hat{\mu}_B^5 + \frac{4!}{7!}c_4\epsilon_1\hat{\mu}_B^7 + \frac{4!}{9!}c_4\epsilon_2\hat{\mu}_B^9 \tag{5}$$

$$\chi_2^B(\hat{\mu}_B) = 2c_2 + 12c_4\hat{\mu}_B^2 + 30c_6\hat{\mu}_B^4 + \frac{4!}{6!}c_4\epsilon_1\hat{\mu}_B^6 + \frac{4!}{8!}c_4\epsilon_2\hat{\mu}_B^8 \tag{6}$$

$$\chi_3^B(\hat{\mu}_B) = 24c_4\hat{\mu}_B + 120c_6\hat{\mu}_B^3 + \frac{4!}{5!}c_4\epsilon_1\hat{\mu}_B^5 + \frac{4!}{7!}c_4\epsilon_2\hat{\mu}_B^7 \tag{7}$$

$$\chi_4^B(\hat{\mu}_B) = 24c_4 + 360c_6\hat{\mu}_B^2 + c_4\epsilon_1\hat{\mu}_B^4 + \frac{4!}{6!}c_4\epsilon_2\hat{\mu}_B^6 \tag{8}$$

We perform a correlated fit for the four measured observables, thus obtaining the values of c_2, c_4 and c_6 for each temperature, and the corresponding χ_2^B, χ_4^B and χ_6^B. We repeat

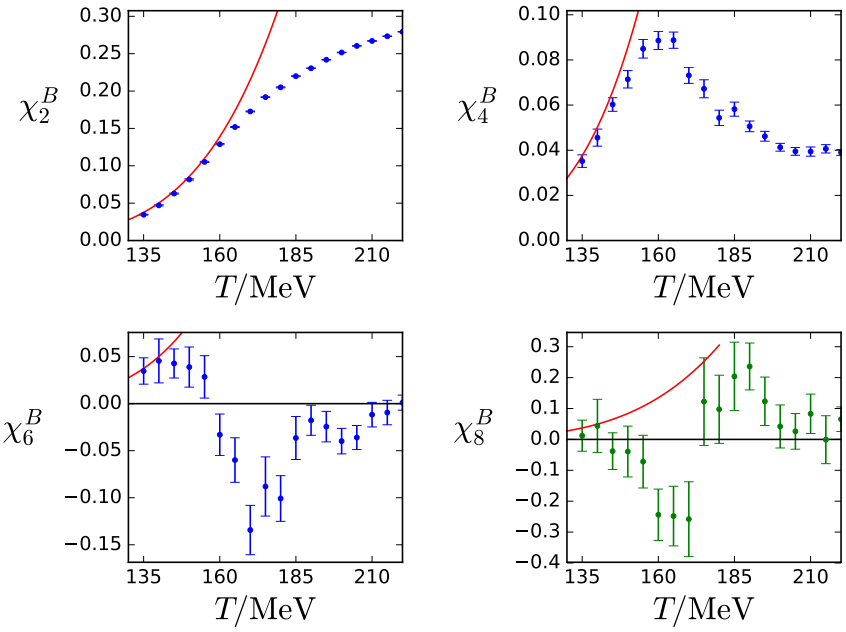


Figure 1. Results for χ_2^B, χ_4^B, χ_6^B and an estimate for χ_8^B as functions of the temperature, obtained from the single-temperature analysis. We plot χ_8^B in green to point out that its determination is guided by a prior, which is linked to the χ_4^B observable by Eq. 4. The red curve in each panel corresponds to the Hadron Resonance Gas (HRG) model result.

the fit for 1000 random draws for ϵ_1 and ϵ_2. The result is weighted using the Akaike Information Criterion.[18] Through these weights we get a posterior distribution from the prior distribution. Our final estimate for χ_8^B represents this posterior distribution. These results are shown in Fig. 1, together with an estimate of χ_8^B, related to χ_4^B by Eq. 4. With a similar fit procedure we determine a number of diagonal and off-diagonal fluctuations.

3.3 Error Analysis

For a reliable comparison between experimental measurements and theoretical calculations, the error estimate is an important ingredient. Our statistical error is estimated through the jackknife method. For our systematic error there are several sources. We determine our systematic error by the histogram method described in Ref. 19, where each analysis is weighted with the Akaike information criteria. We include the influence of the number of points in the μ_B direction, by either including or ignoring the data from our highest value of μ_B. A very important source for our systematic error is the influence of the higher order contributions in μ_B. This effect was estimated by adding the higher order terms with pre-factors ϵ_1 and ϵ_2 as described in Sec. 3.2. We consider 1000 different ϵ pairs and add the different analyses to our histogram. The width of the histogram using Akaike weights corresponding to the fit quality gives the systematic errors for the fit coefficients, and from the same histogram we obtain the posterior distributions for ϵ_1. The physical quantities

that are constrained only by the posterior distribution are plotted with green symbols.

These histograms are built independently for each number (j and k) of μ_S and μ_Q derivatives. When calculating the systematics for the cumulant ratios (Sec. 4) we need to calculate different combinations of diagonal and non-diagonal fluctuations from the available analyses. Though these fits (corresponding to the same temperature) are carried out separately we keep track of the statistical correlation, by maintaining the jackknife ensembles throughout the analysis. The correct propagation of systematic errors is a more elaborate procedure. When χ^{BSQ}_{ijk} coefficients are combined with different j, k pairs, different histograms have to be combined. If we had only two variables to combine, each of the 2000 first fit variants should be combined with each of the 2000 second fit variants and use the product of the respective probability weights. Instead, we combine the fit results by drawing 'good' fits by importance sampling from each histogram independently. In this way, $\mathcal{O}(100)$ random combinations of χ^{BSQ}_{ijk} results already give convergence for each discussed quantity and its error bar. For the results in this paper we used 1000 such random combinations. This procedure assumes that between different j, k pairs the prior distribution is uncorrelated.

4 Phenomenology at Finite Chemical Potential

For a comparison with heavy ion collision experiments, the cumulants of the net-baryon distribution are very useful observables. The first four cumulants are the mean M_B, the variance σ^2_B, the skewness S_B and the kurtosis κ_B. By forming appropriate ratios, we can cancel out explicit volume factors. However, the measured distributions themselves may still depend on the volume, which one should take into account when comparing to experiments.

Heavy ion collisions involving lead or gold atoms at $\mu_B > 0$ correspond to the following situation $\langle n_S \rangle = 0$ and $\langle n_Q \rangle = 0.4 \langle n_B \rangle$. For each T and μ_B pair, we have to first calculate μ_Q and μ_S that satisfy this condition. The resulting $\mu_Q(\mu_B)$ and $\mu_S(\mu_B)$ functions, too, can be Taylor expanded,[20, 21] introducing

$$q_j = \frac{1}{j!} \frac{d^j \hat{\mu}_Q}{(d\hat{\mu}_B)^j}\bigg|_{\mu_B=0}, \; s_j = \frac{1}{j!} \frac{d^j \hat{\mu}_S}{(d\hat{\mu}_B)^j}\bigg|_{\mu_B=0} \tag{9}$$

We investigate three different ratios of cumulants:

$$\frac{M_B}{\sigma^2_B} = \frac{\chi^B_1(T, \hat{\mu}_B)}{\chi^B_2(T, \hat{\mu}_B)} = \hat{\mu}_B r^{B,1}_{12} + \hat{\mu}^3_B r^{B,3}_{12} + \ldots \tag{10}$$

$$\frac{S_B \sigma^3_B}{M_B} = \frac{\chi^B_3(T, \hat{\mu}_B)}{\chi^B_1(T, \hat{\mu}_B)} = r^{B,0}_{31} + \hat{\mu}^2_B r^{B,2}_{31} + \ldots \tag{11}$$

$$\kappa_B \sigma^2_B = \frac{\chi^B_4(T, \hat{\mu}_B)}{\chi^B_2(T, \hat{\mu}_B)} = r^{B,0}_{42} + \hat{\mu}^2_B r^{B,2}_{42} + \hat{\mu}^4_B r^{B,4}_{42} + \ldots \tag{12}$$

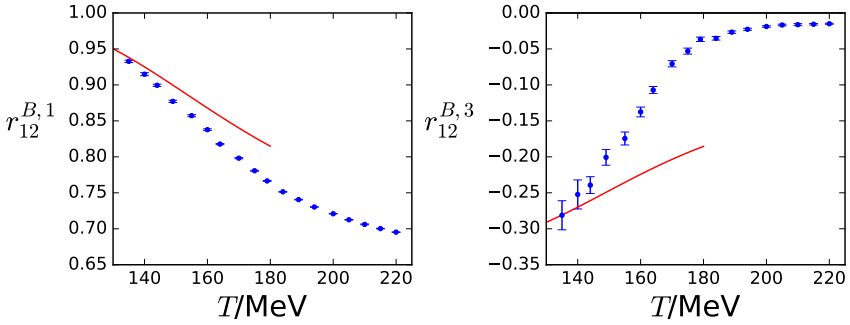


Figure 2. Taylor expansion coefficients for $\frac{M_B}{\sigma_B^2} = \frac{\chi_1^B(T,\hat{\mu}_B)}{\chi_2^B(T,\hat{\mu}_B)}$ as functions of the temperature: $r_{12}^{B,1}$ (left panel) and $r_{12}^{B,3}$ (right panel).

The μ_B-dependence of the $\chi_i^B(T,\hat{\mu}_B)$ can again be written as a Taylor series:

$$\begin{aligned}\chi_{i,j,k}^{BQS}(\hat{\mu}_B) = \chi_{i,j,k}^{BQS}(0) + \hat{\mu}_B \left[\chi_{i+1,j,k}^{BQS}(0) + q_1\chi_{i,j+1,k}^{BQS}(0) + s_1\chi_{i,j,k+1}^{BQS}(0)\right] \\ + \frac{1}{2}\hat{\mu}_B^2 \left[\chi_{i+2,j,k}^{BQS}(0) + q_1^2\chi_{i,j+2,k}^{BQS}(0) + s_1^2\chi_{i,j,k+2}^{BQS}(0)\right. \\ \left. +2q_1s_1\chi_{i,j+1,k+1}^{BQS}(0) + 2q_1\chi_{i+1,j+1,k}^{BQS}(0) + 2s_1\chi_{i+1,j,k+1}^{BQS}(0)\right] + \ldots \quad (13)\end{aligned}$$

The χ coefficients that we determined in Sec. 3 include derivatives up to sixth order, and we have estimates for the eighth order, too. The fit coefficients corresponding to the tenth order are likely to be contaminated by higher orders, that we did not include into the ansatz. These χ_{ijk}^{BQS} coefficients, however, are given for $j+k \leq 4$, which is the highest order that we used in μ_Q and μ_S.

This list of coefficients allows us to calculate the $r_{ij}^{B,k}$ coefficients from Eqs. 10, 11 and 12. The results for the $r_{12}^{B,k}$ coefficients are shown in Fig. 2. We confirm the observation from Ref. 3 that the coefficient $r_{42}^{B,2}$ has a similar temperature dependence as $r_{31}^{B,2}$ but it is ~ 3 times larger in magnitude.

For higher order coefficients, higher order derivatives in μ_S and μ_Q are needed. The direct simulations have a rapidly increasing error with the order of the derivative, and very large statistics would be needed to improve our calculations at this point. Another possibility would be to simulate new ensembles with finite μ_S and μ_Q and do a similar fit as for the μ_B direction. This approach has been used in Ref. 6.

After calculating the Taylor coefficients for $S_B\sigma_B^3/M_B$ and $\kappa_B\sigma_B^2$, we use these results to extrapolate these quantities to finite chemical potential. They are shown in Fig. 3. In the left panel, $S_B\sigma_B^3/M_B$ is shown as a function of the chemical potential for different temperatures. The Taylor expansion for this quantity is truncated at $\mathcal{O}(\hat{\mu}_B^2)$. The black points in the figure are the experimental results from the STAR collaboration from an analysis of cumulant ratios measured at mid-rapidity, $|y| \leq 0.5$, including protons and anti-protons with transverse momenta 0.4 GeV $\leq p_t \leq$ 2.0 GeV.[22, 23] The beam energies were translated to chemical potentials using the fitted formula of Ref. 24. Even if we do

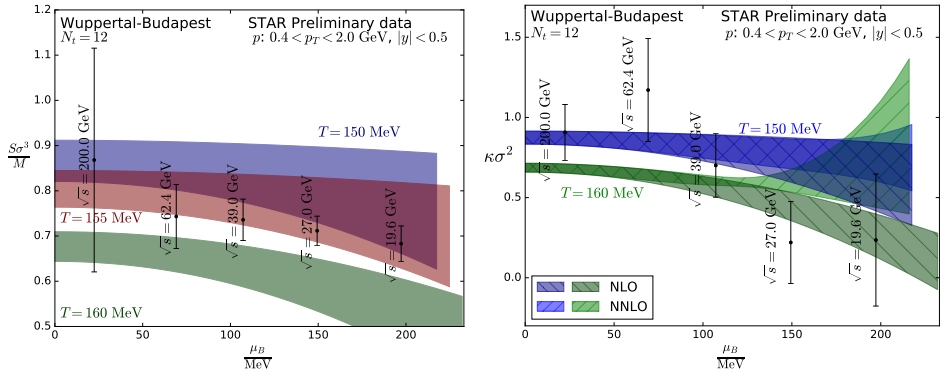


Figure 3. $S_B\sigma_B^3/M_B$ (left panel) and $\kappa_B\sigma_B^2$ (right panel) extrapolated to finite chemical potential. The left panel is extrapolated up to $\mathcal{O}(\hat{\mu}_B^2)$. In the right panel, the darker bands correspond to the extrapolation up to $\mathcal{O}(\hat{\mu}_B^2)$, whereas the lighter bands also include the $\mathcal{O}(\hat{\mu}_B^4)$ term.

not quantitatively compare the lattice bands to the measurements to extract the freeze-out parameters, as experimental higher order fluctuations might be affected by several effects of non-thermal origin and our lattice results are not continuum extrapolated, we notice that the trend of the data with increasing μ_B can be understood in terms of our Taylor expansion.

In the right panel, we show $\kappa_B\sigma_B^2$ as a function of μ_B/T for different temperatures. The darker bands correspond to the extrapolation up to $\mathcal{O}(\hat{\mu}_B^2)$, whereas the lighter bands also include the $\mathcal{O}(\hat{\mu}_B^4)$ term. Also in this case, the black points are the experimental results from the STAR collaboration with transverse momentum cut 0.4 GeV$\leq p_t \leq$ 2.0 GeV.[22, 23] Notice that, due to the fact that the $r_{42}^{B,4}$ is positive in the range 160 MeV$\leq T \leq$195 MeV, we observe a non-monotonic behaviour in $\kappa_B\sigma_B^2$ for $T = 160$ MeV at large chemical potentials. By comparing the two different truncations of the Taylor series we can conclude that, as we increase the temperature, the range of applicability of our Taylor series decreases: while at $T = 150$ MeV the two orders agree in the whole μ_B/T range shown in the figure, at $T = 160$ MeV the central line of the next-to-next-to-leading order bends upwards and is not contained in the next-to-leading order band. To make the NLO prediction precise substantially more computer time would be needed.

5 Conclusions and Outlook

In this manuscript, we have calculated several diagonal and non-diagonal fluctuations of electric charge, baryon number and strangeness up to sixth-order, in a system of 2+1+1 quark flavours with physical quark masses, on a lattice with size $48^3 \times 12$. The analysis has been performed simulating the lower order fluctuations at zero and imaginary chemical potential μ_B, and extracting the higher order fluctuations as derivatives of the lower order ones at $\mu_B = 0$. The chemical potentials for electric charge and strangeness have both been set to zero in the simulations. From these fluctuations, we have constructed ratios of baryon number cumulants as functions of T and μ_B, by means of a Taylor series which

takes into account the experimental constraints $\langle n_S \rangle = 0$ and $\langle n_Q \rangle = 0.4 \langle n_B \rangle$. These ratios qualitatively explain the behaviour observed in the experimental measurements by the STAR collaboration as functions of the collision energy.

We focused on observables (baryon distribution, ratios of cumulants) that are less sensitive to lattice artefacts. An obvious extension of our work will be the use of finer lattices and a continuum extrapolation. The other extension is to use a two- or even three-dimensional mapping of the space of the imaginary chemical potentials using non-vanishing μ_S and μ_Q. That would not only improve the μ_S- and μ_Q-derivatives, but would allow us to study the melting of states with various strangeness and electric charge quantum numbers. Our first study in this direction using strangeness chemical potentials was published in Ref. 25.

Acknowledgements

This project was funded by the DFG grant SFB/TR55. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project by providing computing time on the GCS Supercomputer JUQUEEN[26] at Jülich Supercomputing Centre (JSC) as well as on HAZELHEN at HLRS Stuttgart, Germany.

References

1. A. Bazavov *et al.*, *Chiral crossover in QCD at zero and non-zero chemical potentials*, Phys. Lett. B **795**, 15–21, 2019.
2. C. Bonati, M. D'Elia, F. Negro, F. Sanfilippo, and K. Zambello, *Curvature of the pseudocritical line in QCD: Taylor expansion matches analytic continuation*, Phys. Rev. D **98**, 054510, 2018.
3. A. Bazavov *et al.*, *Skewness and kurtosis of net baryon-number distributions at small values of the baryon chemical potential*, Phys. Rev. D **96**, 074510, 2017.
4. R. V. Gavai and S. Gupta, *QCD at finite chemical potential with six time slices*, Phys. Rev. D **78**, 114503, 2008.
5. J. N. Guenther, R. Bellwied, S. Borsanyi, Z. Fodor, S. D. Katz, A. Pasztor, C. Ratti, and K. K. Szabó, *The QCD equation of state at finite density from analytical continuation*, Nucl. Phys. A **967**, 720–723, 2017.
6. M. D'Elia, G. Gagliardi, and F. Sanfilippo, *Higher order quark number fluctuations via imaginary chemical potentials in $N_f = 2 + 1$ QCD*, Phys. Rev. D **95**, 094503, 2017.
7. C. R. Allton, S. Ejiri, S. J. Hands, O. Kaczmarek, F. Karsch, E. Laermann, C. Schmidt, and L. Scorzato, *The QCD thermal phase transition in the presence of a small chemical potential*, Phys. Rev. D **66**, 074507, 2002.
8. R. Bellwied, S. Borsanyi, Z. Fodor, S. D. Katz, A. Pasztor, C. Ratti, and K. K. Szabo, *Fluctuations and correlations in high temperature QCD*, Phys. Rev. D **92**, 114505, 2015.
9. A. Roberge and N. Weiss, *Gauge Theories With Imaginary Chemical Potential and the Phases of QCD*, Nucl. Phys. B **275**, 734–745, 1986.

10. V. Vovchenko, A. Pasztor, Z. Fodor, S. D. Katz, and H. Stoecker, *Repulsive baryonic interactions and lattice QCD observables at imaginary chemical potential*, Phys. Lett. B **775**, 71–78, 2017.
11. A. Vuorinen, *Quark number susceptibilities of hot QCD up to $g^6 lng$*, Phys. Rev. D **67**, 074032, 2003.
12. C. Bonati, M. D'Elia, M. Mariti, M. Mesiti, F. Negro, and F. Sanfilippo, *Curvature of the chiral pseudocritical line in QCD: Continuum extrapolated results*, Phys. Rev. D **92**, 054503, 2015.
13. R. Bellwied, S. Borsanyi, Z. Fodor, J. Guenther, S. D. Katz, C. Ratti, and K. K. Szabo, *The QCD phase diagram from analytic continuation*, Phys. Lett. B **751**, 559–564, 2015.
14. P. Cea, L. Cosmai, and A. Papa, *Critical line of 2+1 flavor QCD: Toward the continuum limit*, Phys. Rev. D **93**, 014507, 2016.
15. A. Bazavov *et al.*, *The QCD Equation of State to $\mathcal{O}(\mu_B^6)$ from Lattice QCD*, Phys. Rev. D **95**, 054504, 2017.
16. C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage, *High-Precision c and b Masses, and QCD Coupling from Current-Current Correlators in Lattice and Continuum QCD*, Phys. Rev. D **82**, 034512, 2010.
17. S. Borsanyi, Z. Fodor, J. N. Guenther, S. K. Katz, K. K. Szabo, A. Pasztor, I. Portillo, and C. Ratti, *Higher order fluctuations and correlations of conserved charges from lattice QCD*, JHEP **2018**, 205, 2018.
18. H. Akaike, *Information theory and an extension of the maximum likelihood principle*, in Selected Papers of Hirotugu Akaike, E. Parzen, K. Tanabe, G. Kitagawa (Editors), Springer, 119–213, 1998.
19. S. Dürr *et al.*, *Ab Initio Determination of Light Hadron Masses*, Science **322**, 1224–1227, 2008.
20. A. Bazavov *et al.*, *Freeze-out Conditions in Heavy Ion Collisions from QCD Thermodynamics*, Phys. Rev. Lett. **109**, 192302, 2012.
21. S. Borsanyi, Z. Fodor, S. D. Katz, S. Krieg, C. Ratti, and K. K. Szabo, *Freeze-out parameters: lattice meets experiment*, Phys. Rev. Lett. **111**, 062005, 2013.
22. X. Luo, *Energy Dependence of Moments of Net-Proton and Net-Charge Multiplicity Distributions at STAR*, PoS **CPOD2014**, 019, 2015.
23. J. Thäder, *Higher Moments of Net-Particle Multiplicity Distributions*, Nucl. Phys. A **956**, 320–323, 2016.
24. A. Andronic, P. Braun-Munzinger, and J. Stachel, *Hadron production in central nucleus-nucleus collisions at chemical freeze-out*, Nucl. Phys. A **772**, 167–199, 2006.
25. Paolo Alba *et al.*, *Constraining the hadronic spectrum through QCD thermodynamics on the lattice*, Phys. Rev. D, **96**, 034517, 2017.
26. M. Stephan, J. Docter, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, JLSRF **1**, A1, 2015.

Materials Science

Materials Science

Robert O. Jones

Peter Grünberg Institut PGI-1, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: r.jones@fz-juelich.de

It has often been said – by me and by others – that the challenges faced by materials scientists are often of immediate technological relevance. All fields represented in this volume have benefited immensely from the dramatic improvements of computer capability in the past decades, and the paths to new knowledge in most areas include simulations. Materials science simulations can provide detailed information that is not available to experimentalists, such as the coordinates and velocities of individual atoms of a disordered sample, and there are obvious advantages if we study materials that are poisonous, radioactive, metastable, or even unstable. The expansion and availability of computer resources has revolutionised materials simulations, particularly in the field of density functional (DF) calculations, where simulations requiring millions of calculations can now be performed on systems with hundreds of atoms, and in the area of classical molecular dynamics (MD), where long simulations on very large samples have become routine.

The DF formalism[1] is the basis of almost all calculations in condensed matter physics that do not use adjustable parameters, and it now plays a dominant role in computational chemistry. This approach allows us to calculate the total energy of an arrangement of atoms and molecules, so that we can determine structural information and cohesive properties without input from experiment. Such calculations are often termed "*ab initio*", but I reserve this for solutions of the Schrödinger equation, because we still need an approximation for the "exchange-correlation" part of the total energy. There has been much progress in the development of such approximations, but many incorporate input from measurements, and it is a common view (incorrect, in my opinion), particularly among chemists, that DF calculations are essentially "semi-empirical". Nevertheless, it is possible to perform combined MD/DF calculations for millions of time steps (several nanoseconds) for up to 1000 atoms.[2] Modern computer hardware and mathematical algorithms have made possible the study of properties of whole *families* of materials, which is often essential in materials screening. Such length and time scales are completely inadequate in some cases, where the energy and forces are usually described by a classical force field with parameters chosen to reproduce the results of experiment or DF calculations on smaller systems.

The first of the materials science studies presented in this volume is a good example of a density functional simulation of a real materials problem, where answers are obtained for questions concerning specific materials. In this case, the focus is on role of defects and other impurities on the properties of β-In_2S_3, which is a possible replacement for CdS, which is used as a buffer material in solar cells and many other electronic devices. The role of vacancies and Cu, Na, O, and Cl impurity atoms has been studied in detail. The approximation to the exchange-correlation energy used here involves a so-called hybrid functional, where the local exchange energy is augmented by a fraction of the screened Hartree-Fock exchange. The use of hybrid functionals often leads to improved agreement

with measured excitation energies, but at increased computational cost. The second article differs from the mainstream of materials simulations on the Jülich supercomputers, but it falls without question into this category. It involves the numerical solution of wave propagation through three-dimensional models using finite-difference codes and adapting the time reverse imaging (TRI) method used in seismic studies. The applications are to the location and other properties of defects in models of a 3D-pipe.

The work described in these two articles are examples of the materials science projects being carried out on JUWELS and JURECA, and there is little doubt that computer simulations will play an essential role in future computational research in this area. It is somewhat disconcerting to note, however, that materials science and the related fields of condensed matter physics and chemistry play much smaller roles in allocations of resources on Jülich supercomputers than they did 20-30 years ago. It may be that this is a consequence of the difficulties of coding programs in the density functional field for massively parallel computers. While there are clear exceptions to this statement, such as the KKRNANO program for solids,[3] many current codes run on hundreds or thousands of cores, but not efficiently on the hundreds of thousands of cores that are now available and can be used by projects in other areas. The diminished role of materials science and related areas is certainly not due to the lack of relevant problems of technological importance or the ability of computer simulations to help solve them.

References

1. R. O. Jones, *Density functional theory: Its origins, rise to prominence, and future*, Rev. Mod. Phys. **87**, 897-923, 2015.
2. J. Kalikka, J. Akola, and R. O. Jones, *Crystallization processes in the phase change material $Ge_2Sb_2Te_5$: Unbiased density functional/molecular dynamics simulations*, Phys. Rev. B **94**, 134105, 2016.
3. R. Zeller, *Towards a linear-scaling algorithm for electronic structure calculations with the tight-binding Korringa-Kohn-Rostoker Green function method*, J. Phys.: Condens. Matter **20**, 294215, 2008.

A First-Principles Study on the Role of Defects and Impurities in β-In_2S_3

Elaheh Ghorbani and Karsten Albe

Fachgebiet Materialmodellierung, Institut für Materialwissenschaft, TU Darmstadt,
Otto-Berndt-Straße 3, 64287 Darmstadt, Germany
E-mail: {ghorbani, albe}@mm.tu-darmstadt.de

CdS is a well-established buffer layer for Cu(In, Ga)(S, Se)$_2$ (CIGS)-based thin film solar cells. However, because of its toxicity, low quantum efficiency at blue-wavelength region, and the drawbacks of the chemical bath deposition technique used for its growth, looking for an alternative buffer material has been a matter of debate in recent years. In this context, β-In_2S_3 is considered as a promising substitution for CdS. β-In_2S_3 crystallises in an ordered vacancy spinel-like structure, which can accommodate impurities diffusing from the absorber and/or front contact layers. Due to the existence of structural vacancies in its crystalline matrix, the electronic and optical properties of β-In_2S_3 can be effectively tuned through (un)intentional doping with a third element. In this contribution, we will report on the origin of n-type conductivity of β-In_2S_3, the influence of Cu and Na incorporation, the thermodynamic stability and electronic properties of β-In_2S_3, and the influence of O and Cl on electronic properties of β-In_2S_3.

1 Introduction

β-In_2S_3 is an n-type semiconductor which crystallises in a defective spinel-like structure[1, 2] with four structural tetrahedral vacancies per unit cell. The interesting properties of this chalcogenide semiconductor, such as light transmission in the blue wavelength region, high open circuit voltage and fill factor qualify β-In_2S_3 for photovoltaic applications.[3] To date, the highest reported efficiencies[4] for copper indium gallium diselenide (CIGS)-based thin film solar cells have been achieved through the deposition of cadmium sulphide (CdS) as buffer layer. The major drawbacks of CdS buffer layers are: toxicity of cadmium, its conventional chemical bath deposition method, which is difficult to combine with the deposition techniques for CIGS, and its low external quantum efficiency in the short wavelength region.[5] Thus, there is an interest in replacing CdS by a transparent buffer system, which can be deposited by various techniques and has a wider band gap than CdS. This is why, there is a strong focus on β-In_2S_3 which can be deposited by various techniques[6] and has a band gap in the range of 2.0 to 2.9 eV.[1, 3, 7–9]

Due to the existence of structural vacancies, the electronic and optical properties of β-In_2S_3 can be effectively engineered and optimised through (un)intentional doping with selected atom types. Several studies have documented an intermixed In_2S_3/CIGS interface[10–12] mainly containing Na, Cu, O, and Cl impurities. Therefore, a detailed knowledge of the role of point defects and impurities in this material is of key importance. In this contribution, we present a complete assessment of intrinsic and extrinsic point defects in In_2S_3.[13–15]

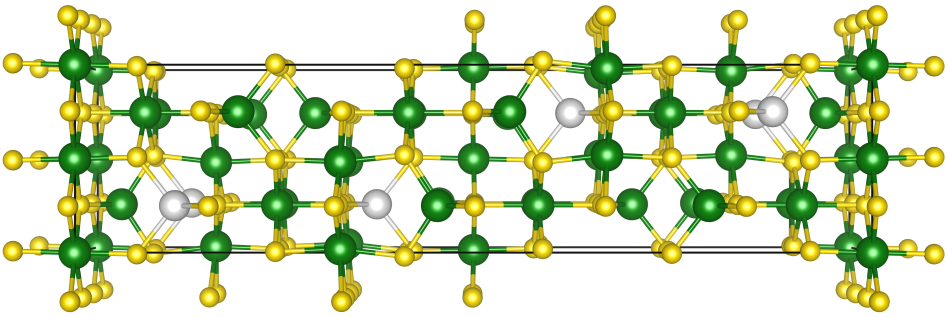


Figure 1. Crystal structure showing the unit cell of β-In_2S_3. Green and yellow spheres represent In and S atoms respectively. The tetrahedral structural vacancies are shown as white spheres (taken from Ref. 13).

2 Methodology

All calculations were carried out within the framework of density functional theory (DFT) using a hybrid functional approach[16] as implemented in the Vienna Ab initio Simulation Package (VASP).[17, 18] For an accurate description of the formation energies, as well as atomic and electronic structures, 25% of non-local Hartree-Fock screened exchange (also known as mixing parameter, α) has been added to 75% of the semi-local exchange energy of PBE functional. The screening range of the electron interaction has been treated by adjusting the Thomas-Fermi screening parameter (w) to 0.13 $Å^{-1}$. The Kohn-Sham equations were solved through employing the projector augmented-wave method,[19, 20] expanding the wave functions up to a cut-off energy of 300 eV. Spin-polarised calculations have been performed for systems with unpaired spins. Defect properties were investigated using a 2×2×1 supercell containing 320 atoms, and the Brillouin-zone sampling was performed using a Γ-centred tetragonal mesh of 2×2×1.

Point defect formation energies have been calculated at theoretical equilibrium volume as

$$E_f[d^q] = \Delta E[d^q] \pm \Sigma_i n_i \mu_i + q[E_{\mathrm{VBM}} + \mu_e] + E_{corr} \quad (1)$$

Since the changes in volume and entropy are tiny in the dilute regime, the effects of formation volume and formation entropy have been skipped in this study. In all calculations the Hellmann-Feynman force convergence was reached only when the largest residual force component on each atom falls below 0.05 eV/Å. $\Delta E[d^q]$ is the energy difference between defect-free and defect-containing structures. n_i corresponds to the number of atoms added or removed from the system. $\mu_i = \mu_i^{ref} + \Delta\mu_i$ is the chemical potential of the constituent i in the reservoir, where μ_i^{ref} denotes the chemical potential of the constituent elements in their stable elemental phases. E_{VBM} is the energy of the valence band maximum (VBM) of the ideal supercell and μ_e is the chemical potential of electrons with respect to E_{VBM}, also known as Fermi level. E_{corr} stands for the artificial electrostatic interactions of point defects with their periodic images and the introduced background charge, which are corrected using FNV scheme provided by Freysoldt, Neugebauer, and Van de Walle.[21, 22]

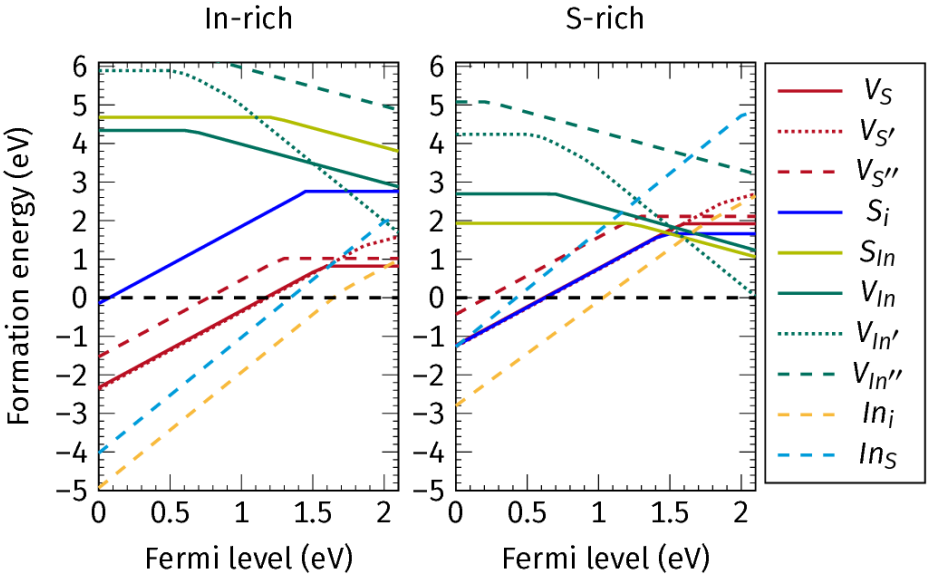


Figure 2. The calculated formation energies of indium and sulphur vacancies as a function of the chemical potential of the electrons, E_f, where E_f=0 eV and E_f=2.1 eV correspond to p-type and n-type conditions, respectively (taken from Ref. 13).

3 The Origin of the n-Type Conductivity of In_2S_3

For obtaining a comprehensive understanding of the role of point defects in indium sulphide, we have calculated formation energies of vacancies (V_S and V_{In}), self-interstitials (S_i and In_i) and anti-sites (S_{In} and In_S) defects as a function of the Fermi level under In-rich and S-rich growth conditions, as shown in Fig. 2. The formation energies shown in Fig. 2 are plotted for the entire range of possible Fermi levels representing various doping conditions. Charge transition levels are positions at points, where the slope of the calculated formation energy changes.

There are three non-equivalent crystallographic environments for indium and sulphur atoms in the β-phase of In_2S_3. The indium substitutional sites in the crystalline matrix of In_2S_3 are tetrahedral $8e$ (denoted In in the following), octahedral $8c$ (denoted In′), and octahedral $16h$ (denoted In″). In a similar fashion, the three different sulphur sites are shown by S (four-fold coordinated to one In, two In′, and one In″), S′ (three-fold coordinated to one In and two In″), and S″ (four-fold coordinated to one In and three In″).

We can see from Fig. 2 that for In-rich samples, there is a strong preference for creation of sulphur vacancies over indium vacancies under both p-type and n-type conditions. However, for S-rich samples, as the Fermi level raises towards the conduction band minimum (CBM), the formation energies and consequently the concentrations of vacant sulphur and vacant indium sites follow a reverse trend. In In-rich samples, the indium interstitial located on the position of a tetrahedral structural vacancy (In_i) has the lowest formation energy at the VBM, lower than all other intrinsic defects. The formation energy of the

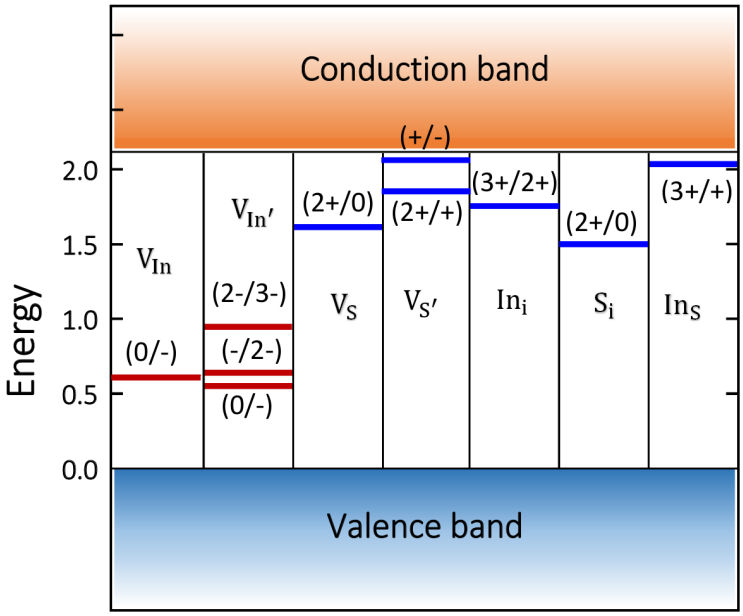


Figure 3. Calculated defect charge transition energies formed in the band gap of β-In_2S_3 using the HSE06[13] functional.

sulphur vacancy (V_S) under the same condition is 0.82 eV. Accordingly, In_i and V_S may play a significant role under thermodynamic equilibrium conditions.

In order to reveal the influence of defects on the electronic structure of the buffer layer, we now discuss the position of charge transition levels as shown in Fig. 3. We see from Fig. 3 that indium vacancies create acceptor levels. However, none of the induced acceptor levels are shallow enough to ionise at room temperature. In S-rich samples, indium atoms need to attain less energy in order to form a vacancy. Consequently, the introduced acceptor levels can be destructive for carrier transport through producing p-conductive acceptor levels. In contrast to indium vacancies, which adopt an acceptor configuration, sulphur vacancies produce donor levels acting as hole killers.

The hybrid functional calculations show that indium in the In_S configuration produces a very shallow (3+/+) donor level at 2.01 eV and a (+/0) level at 0.21 eV above the CBM. Therefore, these intrinsic donor levels can get easily ionised and donate electrons to the conduction band. In_i has also (2+/+) and (+/0) donor levels at 0.03 eV and 0.95 eV above the CBM, which are both resonant within the host bands and can ionise spontaneously, releasing electrons to the system. In consequence, excess indium on anti-sites and tetrahedral vacancies are the main electron-producers, which can effectively dope the material n-type. Besides In_S and In_i, $V_{S'}$ contributes to the n-type conductivity of the system via creating a shallow (+/-) donor-to-acceptor transition level at 0.03 eV below the conduction band. The shallow nature of the (+/-) level allows it to release the two captured electrons at very low temperatures and further boost the free-electron population in the β-In_2S_3.

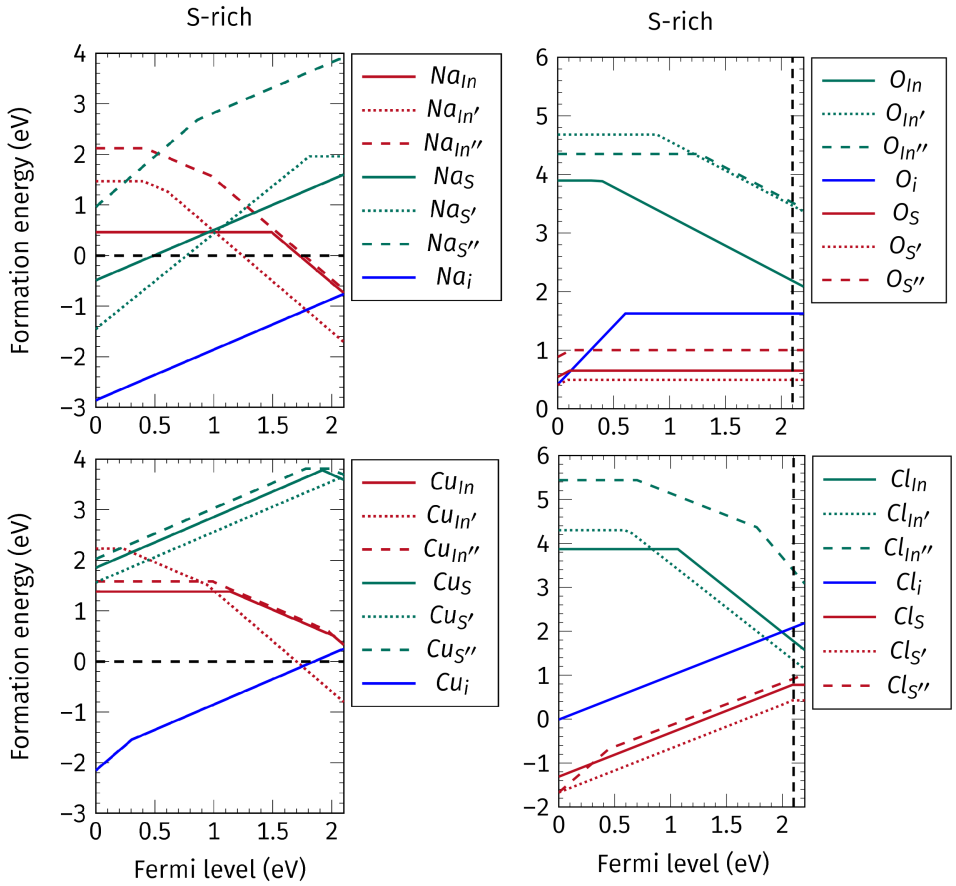


Figure 4. The calculated formation energies of sodium (upper-left), copper (bottom-left), oxygen (upper-right), and chlorine (bottom-right) as a function of the chemical potential of the electrons μ_e, where μ_e=0 eV and μ_e=2.1 eV correspond to p-type and n-type conditions, respectively (taken from Refs. 14, 15).

4 Influence of Impurities on Thermodynamic Stability and Electronic Properties of In_2S_3

When indium sulphide is deposited on CIGS, interdiffusion of Na and Cu from the absorber into the buffer layer leads to the formation of an intermixed absorber-buffer interface involving a high concentration of Na and Cu. Depending on the preparation technique, the presence of O atoms in In_2S_3 may, for instance, stem from side reactions involving decomposition of indium acetylacetonate (used as precursor in the atomic layer deposition (ALD) technique)[23] or high amounts of hydroxide formed during chemical bath deposition (CBD).[23, 24] A significant amount of Cl in In_2S_3 is also detected in the films grown using a Cl-containing precursor (for instance $InCl_3$). In order to put these experimental findings into the context of our theoretical calculations, we have studied the formation enthalpy of Na, Cu, O, and Cl in substitutional and interstitial sites of In_2S_3.[14, 15] Fig. 4 shows the

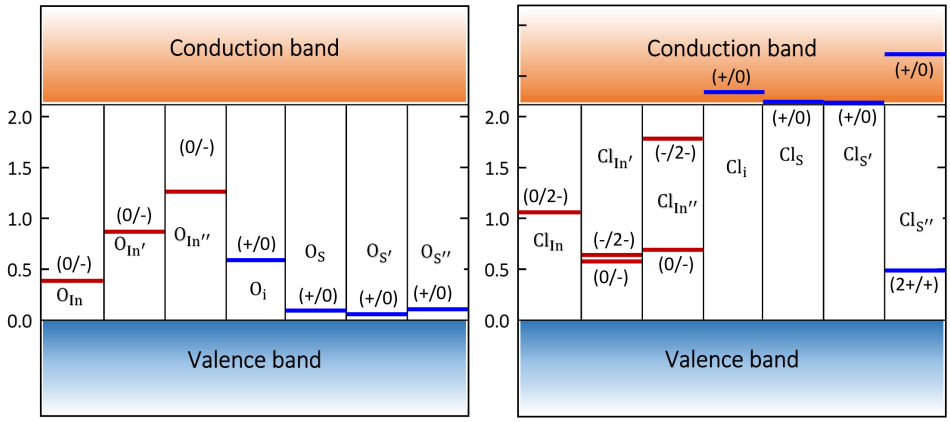


Figure 5. The calculated thermodynamic charge transition levels of O and Cl in different lattice sites of the In_2S_3 (taken from Ref. 15).

calculated formation energies of extrinsic defects in various lattice sites of In_2S_3.

As it can be seen in Fig. 4, the formation enthalpy of Na in the tetrahedral vacancy is negative under both p-type and n-type conditions; this finding unfolds the instability of the system when it comes into contact with a Na reservoir. By Na diffusion into In_2S_3, the system releases energy, and consequently there is a major driving force that triggers the occurrence of side reactions and formation of a Na-containing, interfacial secondary phase on the buffer-side. As a result, the CIGS-In_2S_3 interface is thermodynamically unstable if a Na supply exists in the structure of the multilayer system (*e. g.* Na-containing substrate).

For Cu a slightly different scenario emerges. As long as the Fermi level lies close to the CBM, In_2S_3 can take up a considerable amount of Cu without forming any secondary phases, such as $CuIn_5S_8$; note that the formation energy of Cu_i is positive close to the CBM. However, then the system becomes unstable with respect to Cu_{In} antisite formation, and there is only a small stability window for In-rich conditions, which are experimentally not met. Thus, in the presence of a Cu reservoir, *i. e.* the absorber layer, there is a strong driving force to chemical modifications of the buffer layer. Thus, we can conclude from our data that a thermodynamically stable CIGS-In_2S_3 interface cannot be formed.

Oxygen atoms prefer to occupy S sites rather than In sites. Under n-type conditions, the formation energies of O_S, $O_{S'}$ and $O_{S''}$ are 0.76 eV, 0.49 eV, and 1.0 eV respectively. The very low formation energies of O on different S sites indicates that O substituting anionic sites can occur in In_2S_3 as intrinsic defects. On the Cl side, we see from Fig. 4 that the formation of Cl on sulphur sites is energetically favourable, allowing for a high solubility of Cl in the anionic sub-lattices of In_2S_3. For n-type In_2S_3, the isolated Cl ions on different S sites have formation energies below 1.0 eV, implying that significant concentrations of Cl on S lattice sites can be obtained. $Cl_{S''}$ forms an extremely deep (2+/+) donor state at 1.53 eV below the CBM, and a shallow resonant (+/0) level at 0.65 eV above the CBM, which can readily release electrons to the system.

Fig. 5 shows the calculated charge transition levels for O- and Cl-related point defects. It is seen that O on all S sites induces an extremely deep (+/0) donor level close to the VBM.

The (+/0) level for all three substitutional O on S sites have ionisation energies of around 2.0 eV. In order to effectively dope electrons to the system, donor levels are expected to form near the CBM. However, these deep lying (+/0) levels can neither deliver electrons to the CB nor trap the free holes. In principle, impurity levels close to the valence band can capture free holes. However, since In_2S_3 is an n-type material, the Fermi level lies in the upper part of the band gap. Therefore, the (+/0) levels cannot remove a hole from the valence band. Cl_S and $Cl_{S'}$ induce slight resonances in the conduction band continua, giving rise to a hydrogenic-like (+/0) transition level. The fact that Cl on S sites acts as a shallow donor, suggests that Cl is a suitable n-type dopant in In_2S_3. Therefore, while O on S sites is electrically inactive, Cl on S sites plays a beneficial role in enhancing the concentration of free electrons. This finding allows us to conclude that under low sulphur pressure, intentional doping of In_2S_3 with Cl is an effective way to achieve high levels of n-type conductivity.

5 Concluding Remarks

The formation enthalpy, equilibrium defect transition levels and doping behaviour of a series of native defects and impurities have been studied for bulk β-In_2S_3. Our calculations were done for $2\times2\times1$ supercells containing 320 atoms by means of hybrid functional calculations. Based on our calculations, we identify In_S, In_i, and $V_{S'}$ as dominant sources of n-type conductivity in β-In_2S_3. To understand the stability of the In_2S_3-CIGS interface, we studied the inclusion of extrinsic defects in the In_2S_3 buffer layer. We found that there is a massive driving force for the occurrence of side reactions leading to the formation of Na-containing and Cu-containing interfacial secondary phases. Accordingly, we conclude that the absorber-buffer interface is thermodynamically unstable against the diffusion of Na and Cu into In_2S_3. With regard to the positive formation energies of O- and Cl-related impurities in the crystalline matrix of In_2S_3, we conclude that these anionic impurities do not form secondary phases in the buffer side of the pn-junction. Therefore, the absorber-buffer interface is stable against O and Cl. We found that incorporation of O and Cl on the S sites are energetically favourable. However, while O on S sites creates nonconducting (+/0) impurity levels, Cl on S sites induces hydrogenic-like conductive states, which can actively assist in raising the number of free electrons in the Cl-doped samples.

Acknowledgements

This work has been financially supported by German Federal Ministry for Economic Affairs and Energy (BMWi) through EFFCIS project under contract No. 0324076D. The computing time was provided by Jülich Supercomputing Centre (Project HDA30) and Lichtenberg HPC computer resources at TU Darmstadt.

References

1. W. Rehwald and G. Harbeke, *On the conduction mechanism in single crystal β-indium sulfide In_2S_3*, Journal of Physics and Chemistry of Solids **26(8)**, 1309–1324, 1965.

2. P. Pistor, J. M. Merino Alvarez, M. Leon, M. Di Michiel, S. Schorr, R. Klenk, and S. Lehmann, *Structure reinvestigation of α-, β- and γ-In_2S_3*, Acta Crystallogr. Sect. B: Struct. Sci., Cryst. Eng. Mater. **72**, 410–415, 2016.
3. N. Naghavi, S. Spiering, M. Powalla, B. Cavana, and D. Lincot, *High-efficiency copper indium gallium diselenide (CIGS) solar cells with indium sulfide buffer layers deposited by atomic layer chemical vapor deposition (ALCVD)*, Progress in Photovoltaics: Research and Applications **11(7)**, 437–443, 2003.
4. M. A. Green, K. Emery, Y. Hishikawa, W. Warta, and E. D. Dunlop, *Solar cell efficiency tables (version 48)*, Progress in Photovoltaics: Research and Applications **24(7)**, 905–913, 2016.
5. W. Witte, S. Spiering, and D. Hariskos, *Substitution of the CdS buffer layer in CIGS thin-film solar cells*, Duenne Schichten **26(1)**, 23–27, 2014.
6. N. Naghavi, D. Abou-Ras, N. Allsop, N. Barreau, S. Bücheler, A. Ennaoui, C.-H. Fischer, C. Guillen, D. Hariskos, J. Herrero, R. Klenk, K. Kushiya, D. Lincot, R. Menner, T. Nakada, C. Platzer-Björkman, S. Spiering, A.N. Tiwari, and T. Törndahl, *Buffer layers and transparent conducting oxides for chalcopyrite Cu(In,Ga)(S,Se)$_2$ based thin film photovoltaics: present status and current developments*, Progress in Photovoltaics: Research and Applications **18(6)**, 411–433, 2010.
7. R. Nomura, K. Konishi, and H. Matsuda, *Single-source organometallic chemical vapour deposition process for sulphide thin films: Introduction of a new organometallic precursor BunIn(SPri)2 and preparation of In_2S_3 thin films*, Thin Solid Films **198(1)**, 339–345, 1991.
8. J. F. Trigo, B. Asenjo, J. Herrero, and M. T. Gutierrez, *Optical characterization of In_2S_3 solar cell buffer layers grown by chemical bath and physical vapor deposition*, Solar Energy Materials and Solar Cells **92(9)**, 1145–1148, 2008.
9. J. Sterner, J. Malmström, and L. Stolt, *Study on ald In_2S_3/Cu(In,Ga)Se_2 interface formation*, Progress in Photovoltaics: Research and Applications **13(3)**, 179–193, 2005.
10. P. Pistor, R. Caballero, D. Hariskos, V. Izquierdo-Roca, R. Waechter, S. Schorr, and R. Klenk, *Quality and stability of compound indium sulphide as source material for buffer layers in Cu(In,Ga)Se_2 solar cells*, Solar Energy Materials and Solar Cells **93(1)**, 148–152, 2009.
11. M. Bär, N. Allsop, I. Lauermann, and Ch.-H. Fischer, *Deposition of In_2S_3 on Cu(In,Ga)Se_2 thin film solar cell absorbers by spray ion layer gas reaction: Evidence of strong interfacial diffusion*, Applied Physics Letters **90(13)**, 132118, 2007.
12. M. Bär, N. Barreau, F. Couzinie-Devy, L. Weinhardt, R. G. Wilks, J. Kessler, and C. Heske, *Impact of annealing-induced intermixing on the electronic level alignment at the In_2S_3/Cu(In,Ga)Se_2 thin-film solar cell interface*, ACS Applied Materials & Interfaces **8(3)**, 2120–2124, 2016.
13. E. Ghorbani and K. Albe, *Intrinsic point defects in β-In_2S_3 studied by means of hybrid density-functional theory*, Journal of Applied Physics **123(10)**, 103103, 2018.
14. E. Ghorbani and K. Albe, *Influence of Cu and Na incorporation on the thermodynamic stability and electronic properties of β-In_2S_3*, J. Mater. Chem. C **6**, 7226–7231, 2018.
15. E. Ghorbani and K. Albe, *Role of oxygen and chlorine impurities in In_2S_3: A first-principles study*, Phys. Rev. B **98**, 205201, 2018.
16. J. Heyd, G. E. Scuseria, and M. Ernzerhof, *Hybrid functionals based on a screened Coulomb potential*, J. Chem. Phys. **118(18)**, 8207, 2003.

17. G. Kresse and J. Furthmueller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*, Phys. Rev. B **54**, 11169, 1996.
18. G. Kresse and J. Furthmueller, *Efficiency of ab initio total energy calculations for metals and semiconductors using a plane-wave basis set*, Comput. Mater. Sci. **6(1)**, 15, 1996.
19. G. Kresse and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method*, Phys. Rev. B **59**, 1758, 1999.
20. P. E. Bloechl, *Projector augmented-wave method*, Phys. Rev. B **50**, 17953, 1994.
21. C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Fully ab initio finite-size corrections for charged-defect supercell calculations*, Phys. Rev. Lett. **102**, 016402, 2009.
22. C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Electrostatic interactions between charged defects in supercells*, physica status solidi (b) **248(5)**, 1067–1076, 2011.
23. N. Naghavi, R. Henriquez, V. Laptev, and D. Lincot, *Growth studies and characterisation of In_2S_3 thin films deposited by atomic layer deposition (ALD)*, Applied Surface Science **222(1–4)**, 65–73, 2004.
24. D. Hariskos, M. Ruckh, U. Ruehle, T. Walter, H. W. Schock, J. Hedstroem, and L. Stolt, *A novel cadmium free buffer layer for Cu(In,Ga)Se_2 based solar cells*, Solar Energy Materials and Solar Cells **41**, 345–353, 1996.

Numerical Simulations of Wave Propagation: Time Reverse Imaging and Defect Mapping in Pipes

Erik H. Saenger[1,2], **Claudia Werner**[1,2], **Luan T. Nguyen**[1], **Georg K. Kocur**[3], **and Benedikt Ahrens**[1]

[1] International Geothermal Center Bochum, Lennershofstr. 140, 44801 Bochum, Germany
E-mail: {erik.saenger, claudia.werner, luan.nguyen, benedikt.ahrens}@hs-bochum.de

[2] Ruhr-University Bochum, Universitätsstr. 150, 44801 Bochum, Germany

[3] Institute of General Mechanics, RWTH Aachen University,
Templergraben 64, 52062 Aachen, Germany
E-mail: kocur@iam.rwth-aachen.de

Time reverse imaging (TRI) is evolving into a standard technique for locating and characterising seismic events. In recent years, TRI has been employed for a wide range of applications from the lab scale, to the field scale and up to the global scale. No identification of events or their onset times is necessary when locating events with TRI; therefore, it is especially suited for locating quasi-simultaneous events and events with a low signal-to-noise ratio. However, in contrast to more regularly applied localisation methods, the prerequisites for applying TRI are not sufficiently known. To investigate the significance of station distributions, complex velocity models and signal-to-noise ratios with respect to location accuracy, numerous simulations were performed using a finite difference code to propagate elastic waves through three-dimensional models.

Moreover, we present a reverse-time imaging technique by cross-correlating the forward wavefield with the reverse wavefield for the detection, localisation, and sizing of defects in pipelines. The presented technique allows to capture the wavefield reflectivity at the places of ultrasonic wave scattering and reflections. Thus, the method is suitable for detecting pipe defects of either point-like or finite-size types using data from a pulse-echo setup. By using synthetic data generated by 3D spectral element pipe models, we show that the 3D wavefield cross-correlation imaging is capable in the case of cylindrical guided ultrasonic waves. With a ring setup of transducers, we analyse the imaging results obtained from the synthetic single-transducer and all-transducer firings. The presented pipe flaw imaging method is straightforward to carry out using a suitable wave equation solver.

1 Time Reverse Imaging

The localisation and characterisation of seismic events in the subsurface are crucial for understanding physical processes in the Earth. Well-established methods are able to locate most seismic events in a fast and reliable manner; however, they rely on the identifiable onsets of events. Time reverse imaging (TRI) is a method especially suited for locating and characterising events which are indistinguishable in traces because they occur quasi-simultaneously and/or are superposed by noise. The prerequisites for more regularly applied localisation methods are very well known. However, the station distributions, the degree of complexity in the velocity model and the level of noise that hinder or enhance locating events with TRI are not sufficiently known. Therefore, this study systematically tests different station distributions for their localisation capabilities while considering complex velocity models and low signal-to-noise ratios. TRI uses the whole recorded waveform rendering the identification of events and their onsets obsolete. It can be applied as long as

the wave propagation can be described with a time-invariant wave equation. Seismic traces are reversed in time and back propagated through a medium until they focus on the initial event location. Imaging conditions are used to visualise aspects of the back-propagating wave field and obtain the source location. TRI has been applied in earth sciences as well as medical sciences for some time.[1] In recent years multiple studies have shown that TRI is a reliable, easy-to-use localisation tool: it has been used to retrieve source information on various scales from the lab scale in nondestructive testing,[2–5] over the field scale, for example in volcanic tremor[6] and non-volcanic tremor applications,[7, 8] and above hydrocarbon reservoirs,[9] up to the global scale.[10, 11] However, to apply TRI, a fairly accurate velocity model is needed to precisely back propagate the wave field. With the increasing availability of high resolution large three-dimensional velocity models and growing knowledge about prerequisites, TRI has the potential to locate seismic events, which could previously not be reliably located, over a wide range of applications.

1.1 Numerical Simulations for Time Reverse Imaging

The estimation of location accuracy is one major challenge when applying TRI. Therefore, a common approach is to perform a preliminary synthetic study to test if the given velocity model and station distribution enable reliable locations. If the synthetic study fails, the set-up is adjusted until either the study is abandoned or a sufficiently reliable result is achieved. This adjustment phase can be time-consuming because there are multiple characteristics appearing at once that may hinder the localisation.

A previous study by Werner and Saenger[12] aimed to find station distributions that produce reliable source locations with TRI. Finding these optimal station distributions is crucial to estimate the success rate of TRI with a given set of stations. In addition, the time needed to adjust station distributions may be decreased. Furthermore, the prerequisites of the method should be known when designing an array for TRI. Therefore, the influence of the station distribution, the complexity of the velocity model and the effect of different levels of noise on the location accuracy were investigated. Werner and Saenger[12] performed numerous simulations to systematically analyse different station distributions and their influence on the location accuracy of sources at different depths as well as analysing the influence of different source types. A typical example is shown in Fig. 1. Therefore, the focus was on the distance between receivers, the symmetry of the array in relation to the source position, the azimuthal gap between receivers and the number of stations. Simulations were first performed with a homogeneous velocity model and then with a complex velocity model. Location accuracy was investigated when the velocity model was known and when it was not specifically known. The authors[12] investigated the ability of TRI to cope with very low signal-to-noise ratios. The found guidelines using the methodical tests were applied to an actual real-life example in southern California. The ability to locate events at a target depth was investigated while using the existing array as well as the actual velocity model for a specific region in California.[13]

1.2 Discussion and Conclusion of Time Reverse Imaging

Werner and Saenger[12] showed that the inter-station distance should not be larger than the source depth, and that the aperture of the array should be at least twice the source depth.

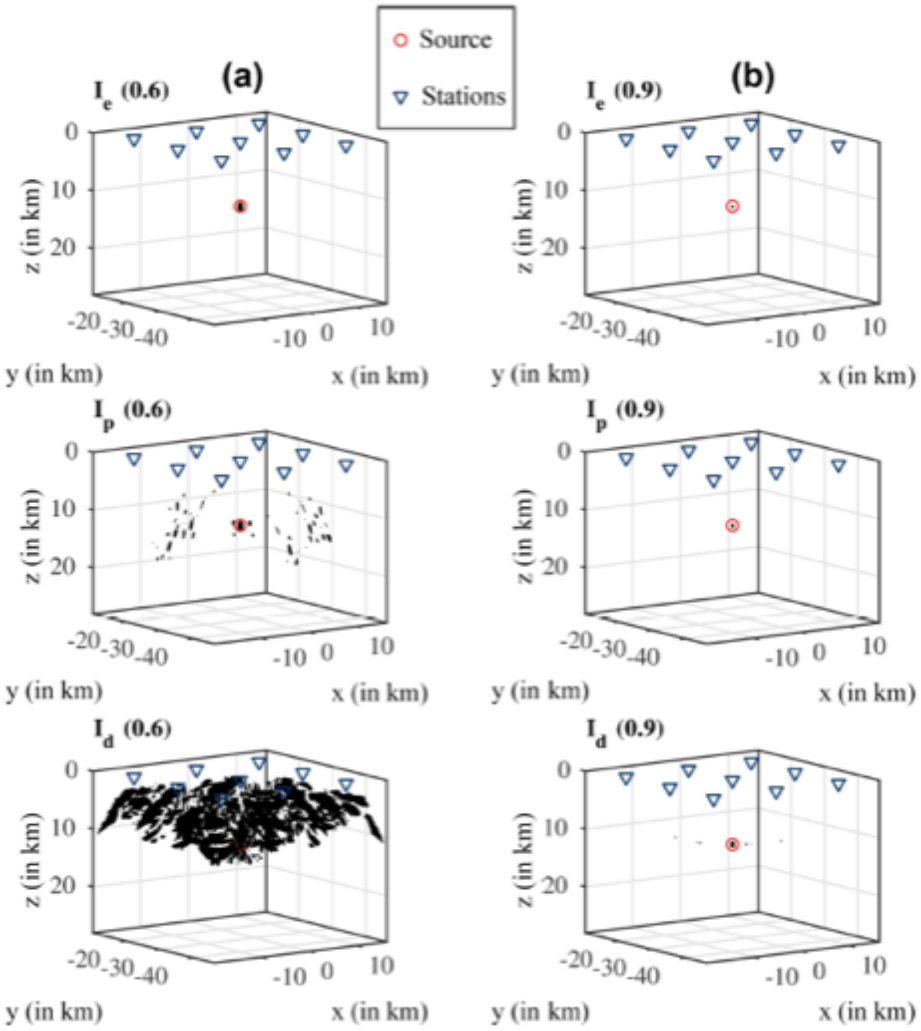


Figure 1. Imaging conditions plotted for one example simulation. Black spots are points with the values of 0.6 (in column a) and 0.9 (in column b) in the respective imaging conditions.[12]

Additionally, sources could be most accurately located when they were underneath the array. When using more stations, locations outside of the array became possible. In general, stations should be spaced regularly, although the deeper the source the more heterogeneously the stations can be and still achieve reliable locations. In addition to this strong sensitivity of the source location accuracy to the station distribution, the study found a strong influence of the velocity model on the localisation. However, complex velocity structures, like low velocity zones, do not hinder the localisation as long as they are adequately incorporated in the model. However, a high level of noise could inhibit the source

location even if the velocity model was homogeneous. An increased number of stations seemed to increase location accuracy in the presence of noise. Furthermore, an increase in location accuracy in some cases with an increase in complexity was observed, such as adding noise, more irregular stations or a more complex velocity model. This suggests that the additional scattering of waves may help the localisation. Therefore, the authors advise performing preliminary synthetic studies not only with homogeneous velocity models and regular stations but also incorporating as many complexities as possible to get an accurate estimation of the achievable source location. It was shown that only a few simulations are necessary when performing a preliminary synthetic study to estimate if the given set-up will work for TRI. Additionally, it was demonstrated that when designing an array to be used with TRI the target area should be considered. The depth of expected events then dictates the inter-station distance and the aperture of the array. If considering a range of depths, the inter-station distance should fit the shallowest events and the aperture of the deepest events.

2 Pipe Inspection

Guided ultrasonic waves are stress waves that propagate in a bounded solid medium. The use of ultra high frequency guided waves in nondestructive testing (NDT) of plate-like structures allows long-distance inspection capability, thus reducing the time and effort in collecting the ultrasound data and later in imaging possible defects. However, those guided waves propagate multi-modally and dispersively,[14] making interpretation of the ultrasound recordings complicated. Most often, ultrasonic excitations of the first few guided wave modes are used for ease of the ultrasound data analysis and interpretation.[15] If higher wave modes cannot be avoided (due to constraints of the inspected structures or transducers' bandwidths, and wave mode conversions), dispersion compensation and separation of propagation modes are sometimes needed.[16] Analytical, experimental, and numerical analyses of guided elastic waves in cylindrical rods and hollow cylinders[17] revealed complicated dispersive and multi-modal characteristics of cylindrical waves. Compared with Lamb waves in thin plate-like structures, cylindrical waves exhibit more complicated guided modes and dispersion characteristics (see Nguyen *et al.*[18] for more details). Flaw detection in pipework using guided waves has received great attention.[19] Most often, the degree of a defect is quantified based on a reflection coefficient expressed as a function of the geometrical extent of the defect.[20] To map defects in pipes, Leonard and Hinders[21] applied a type of travel time "cross-hole" tomography on the helical wave paths for building a tomographic flaw map of pipes. Gaul *et al.*[22] applied a synthetic focusing algorithm to the measured signals of the guided elastic waves in pipes. In another approach, Wang *et al.*[23] used a non-ring sparse sensor network of transducers to image pipe damage in an unwrapped grid based on the correlation of the first arrival waves. Also based on a sensor network of a non-ring configuration, Dehghan- Niri and Salamone[24] proposed to enhance damage-sensitive features, which are input to a reconstruction algorithm based on a waveform difference in baseline and measurement, by using multiple helical paths of the cylindrical waves. Hayashi[25] introduced a fast pipe thinning mapping method using amplitude spectrum peaks obtained from a laser based ultrasonic scanning system. Bagheri *et al.*[26] used the continuous wavelet and Hilbert transforms to extract damage-sensitive features and combined with a metaheuristic optimisation and a probabilistic approach to

construct an image of defects in pipe.

Time reversal in the manner of Fink[27] has also been found useful for NDT of thin-walled structures. By the time reversal principle, the recorded acoustic/elastic waveforms, if being time-reversed and reemitted at the receiving positions, form a wavefield that retraces its path back to the source location. When focused, the reverse wavefield becomes spatially concentrated and temporally compressed to reconstruct the original excitation pulse. Ing and Fink[28] experimentally showed that this time reversal invariance also holds for the dispersive Lamb waves. These features are important for the extrapolated reverse wavefield used in this work to be correctly back-propagated toward scattering and reflecting points. Other reports on structural health monitoring[29] showed that the time reversal principle can be applied to plates and composite structures for achieving baselinefree flaw detection even though dispersion does disturb refocusing of time reversed wave propagation. Xu and Giurgiutiu[29] and Park *et al.*[30] pointed out that time reversal creates sidebands around the true reconstructed impulse if more than one mode of the guided waves is excited. For pipe inspection, Hayashi and Murase[31] used the torsional T(0,1) mode for building a tomographic map of the pipe defects using the time reversal principle. Using the same cylindrical wave mode, Davies and Cawley[32] applied synthetic focusing based on time reversal operation in the frequency-wavenumber domain for pipe defect mapping. To the best of the authors' knowledge, in NDT applications for both plates and pipes, however, the time reversal principle is mostly limited to the location of point-like acoustic sources and scatterers and the reconstruction of the source wavelets. Nguyen *et al.*[18] introduced the use of the wavefield cross-correlation method for flaw detection and sizing in waveguides with a particular application for the inspection of pipework. It was shown that the so-called reverse-time migration, which is based on a zero-lag cross-correlation between the forward wavefield and the reverse wavefield, can be efficiently used for detecting, locating, and sizing flaws of abrupt material changes such as cracks, eroded/corroded voids, and local material damages caused by, for example, impacts in pipes. As this is a method for imaging of reflectors, it can be used to detect and locate both point-like scatterers and finite-size defects.

2.1 Guided Wave Propagation in Pipes

The motions of elastic wave propagation in a pipe, as for other thin-walled structures, are guided within the pipe wall. However, unlike in plates, there exist three distinct families of wave modes in pipes: the axisymmetric longitudinal modes L(0,m), the non-axisymmetric flexural modes F(n,m), and the torsional modes T(0,m) (where m and n are the radial and circumferential mode parameters, respectively[33]). The longitudinal wave modes L(0,1) and L(0,2) correspond to the fundamental anti-symmetric A0 and symmetric S0 Lamb modes, respectively. The torsional mode T(0,1) is equivalent to SH waves in a plate. Detailed solutions of the elastic wave problem in hollow and solid cylinders are studied in Ref. 17 and their modal relation to those in plates are discussed in Ref. 34. Generally, the results show, for the same ratio of thickness-to-diameter, the dispersion characteristics of the axisymmetric longitudinal modes in cylindrical waves are comparable to those of Lamb waves in a plate at the high-frequency range. The similar behaviour is observed if the thickness-to-diameter ratio decreases. The significant difference in the dispersion characteristics between the cylindrical modes and Lamb modes lies in the low-frequency

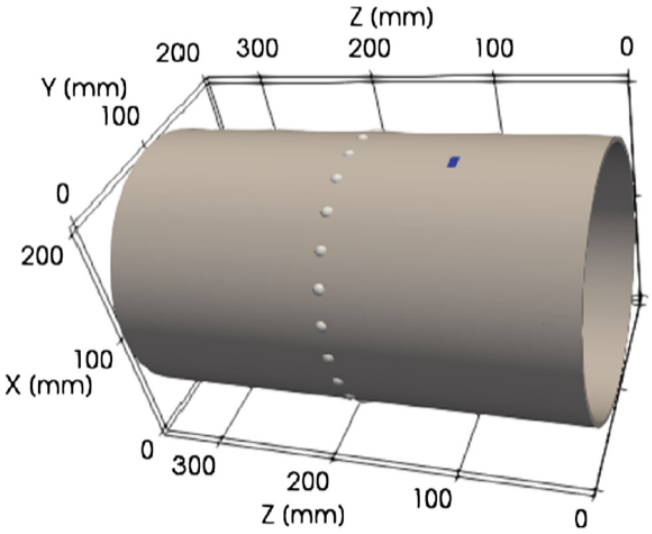


Figure 2. 3D model of a 350 mm long portion of a straight steel pipe with a through-thickness rectangular defect (the blue coloured area).[18] The dots around the pipe circumference represent the positions of the ultrasound transducers arranged at 15 degree intervals.

range. For example, as the frequencies decrease, the phase velocity of the L(0,1) mode of cylindrical waves approaches the S0 Lamb mode instead of approaching zero propagating velocity as for the A0 Lamb mode.

Nguyen *et al.*[18] examined wave propagation in a 350 mm long steel pipe having a small-size through-thickness voided defect as shown in Fig. 2. Because the imaging method requires access to the full ultrasonic wavefields, an efficient 3D elastic wave solver that can handle complex geometries is desirable. 3D guided elastic wavefields were simulated for the pipe model using SPECFEM3D.[35] In addition to an efficient forward solver, SPECFEM3D also provides reversetime simulation and adjoint capabilities instrumental for waveform tomography and imaging. The discretised model consists of 92160 hexahedral elements and is simulated with time stepping of 35 ns. All boundaries including the pipe ends are modelled as traction-free. Based on the dispersion characteristics of guided waves, it is considered advantageous to select a source time function that is narrowband to reduce the number of the excited wave modes and dispersive behaviour during wave propagation. However, a narrowband excitation also means long-duration propagating wave packets in the time domain, which may cause a loss of focus of the reverse wavefield in the imaging. Here, the study examined the effects of narrowband Hanning windowed toneburst and the broadband Ricker wavelet as input signals, both centred at 200 kHz. The choice of the central frequency depends on the imaging resolution one wants to resolve taken into consideration that a higher frequency may excite higher order guided wave modes. Although not formally studied in this work, central frequency around 200 kHz appears to be satisfactory for imaging the simulated defects.

2.2 Wavefield Cross-Correlation Imaging

In reverse-time imaging, the used imaging condition governs the reconstruction capability of an imaging method. To reconstruct a reflectivity image of possibly multiple reflecting and scattering defects in a pipe, the forward wavefield was cross-correlated with the reverse wavefield, which is the working principle behind reverse-time migration (RTM).[36] This imaging condition is based on the princle that wave reflection and scattering happen at places where the source wavefield is time coincient with the receiver re-emited wavefield. Hence, this imaging technique differs greatly from a type of Kirchhoff migration, which is based on the ray assumption of propagating waves, commonly used in ultrasound NDT within implementations of the delay-and-sum synthetic aperture focusing technique (SAFT).[37] RTM, as based on simulating the full wave equation of a possibly heterogeneous velocity model, can be superior to SAFT in delivering higher resolution and more accurate imaging results. RTM also differentiates itself from the popular time reversal multiple signal classification (MUSIC), which calculates scattering intensity at the image points based on the inner products of the Green's function and the noise vectors obtained from singular value decomposition of the time reversal operator.[38]

In ultrasound NDT, RTM of the body waves has recently been used for locating damaged features, hidden interfaces, and boundaries of the civil engineering structures.[39, 40] Concerning NDT of plate-like structures, a small number of published articles reported on the use of RTM for NDT of laminated composite plates within implementations of the excitation-time imaging condition[41] and the zero-lag wavefield cross-correlation imaging condition.[42] The imaging condition by RTM in this work is based on the zero-lag wavefield cross-correlation.[36] So, the final reflectivity image is obtained by performing an entrywise product in space of the forward (source) wavefield and the reverse (receiver) wavefield and cross-correlating them in time following an imaging algorithm (more details are presented in the Appendix A of Nguyen *et al.*[43]). The cross-correlated image adds up constructively at time when the reverse wavefield is spatially refocused and at place where wave scattering and reflection take place.

The homogeneous intact pipe model is used to simulate the forward and reverse wavefields required in the imaging step. Fig. 3 illustrates the stepwise reverse-time imaging process at three selected timesteps: Before focusing, at focusing, and long after focusing of the reverse wavefield takes place. The cross-correlation results in Fig. 3 show that the spatial reconstruction of the defect is captured at the time the forward wavefield and the re-focused reverse wavefield coincide (the second selected timestep in Fig. 3b). After the refocusing of the reverse wavefield, the cross-correlated reconstruction remains relatively unchanged (Fig. 3c). Although the wavefields and the resulting cross-correlated image are three-dimensional, only the fields at the outer surface of the pipe are displayed in Fig. 3. The reconstructed images are normalised to their reflectivity range. This style of plotting applies to the rest of wavefield plots and reflectivity images.

2.3 Discussion and Conclusion of Wavefield Cross-Correlation Imaging

Nguyen *et al.*[18] have presented the wavefield cross-correlation whose imaging principle is sound and its implementation is straightforward.[18] The cross-correlation of the simulated full 3D ultrasonic wavefields honours the kinematics of the wave phenomena. The synthetic verification of this imaging technique for defect mapping in pipes demonstrates

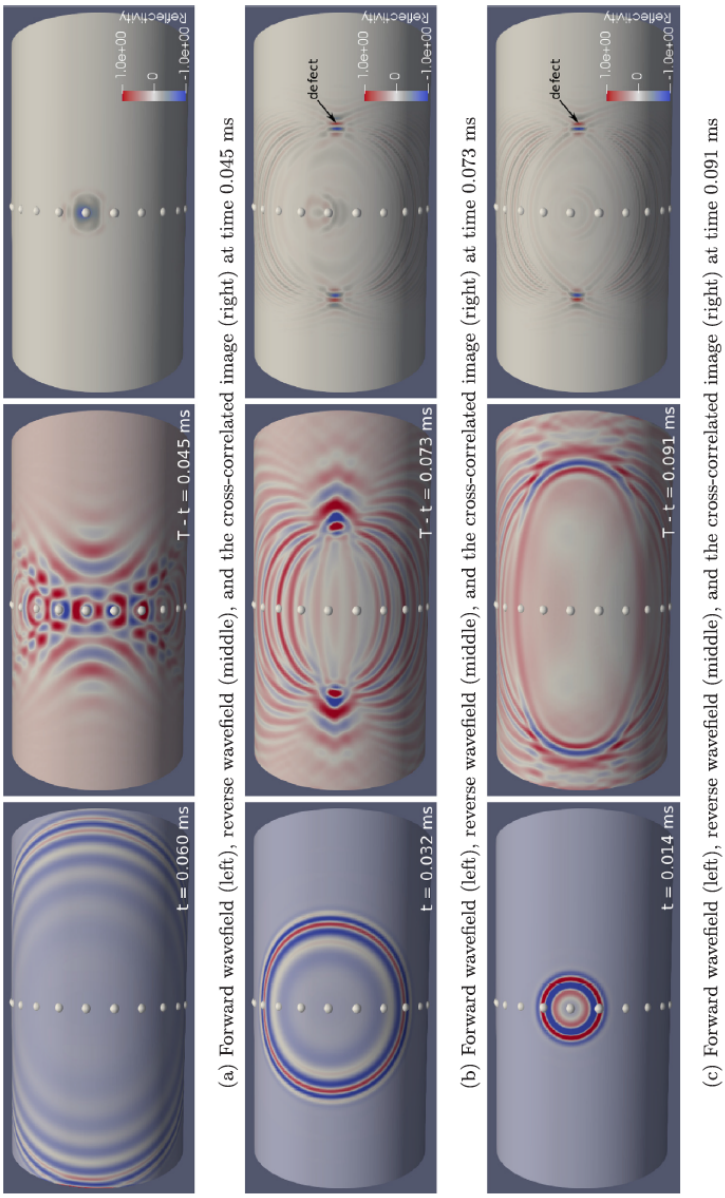


Figure 3. Wavefield snapshots of the cross-correlation imaging process at three selected time steps: (a) before focusing, (b) at focusing, and (c) after focusing of the reverse wavefield has passed.[18] Left column: the forward wavefield is accessed backward in time in the order (a) 0.06 ms, (b) 0.032 ms, and (c) 0.014 ms. Middle colum: the reverse wavefield is accessed forward in time in the order (a) 0.045 ms, (b) 0.073 ms, and (c). 0.091 ms. In the forward and reverse wavefields, displacement field in x-direction is plotted.

its ability to reconstruct a quality image of possibly multiple flaws. Unlike solving an inverse problem, this wave-equation based imaging does not suffer from running long iterations and the risk of arriving at an unfaithful reconstruction (local-minimum solution). For engineered structures undergone non-destructive testing, such as the pipe, the wave velocity and density maps are often homogeneous and known, giving very good conditions for applying the presented imaging method. Concerning computational effort, using the supershot setup[18] greatly expedites the imaging procedure compared with using a number of single element shots. Imaging using no baseline data is also more easily performed on data resulting from the supershot setup and has been demonstrated in this work.

The reverse-time imaging by cross-correlating the forward and reverse wavefields of the guided waves is shown to be highly efficient in constructing a sharply focused reflectivity image of the defects in pipes. The presented implementation is advantageous in that it involves no algorithmic parameters to tune. If post-processing, such as smoothing, of the reconstructed image is required to emphasise specific defected features, the involved image processing can be quickly done given a few input parameters of the filtering kernel. This imaging procedure can be applied and extended to nondestructive evaluation of pipes of arbitrary geometries such as bend pipes or pipes of multiple layers and varying thicknesses and other 3D waveguide structures even in the case of wave interactions with the surrounding environment such as soil and water are present.

Acknowledgements

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

References

1. M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J.-L.Thomas, and F. Wu, *Time-reversed acoustics*, Reports Prog. Phys. **281**, 91–97, 1999.
2. E. H. Saenger, *Time reverse characterization of sources in heterogeneous media*, NDT E. Int. **44**, 751–759, 2011.
3. B. E. Anderson, M. Griffa, T. J. Ulrich, and P. A. Johnson, *Time reversal reconstruction of finite sized sources in elastic media*, J. Acoust. Soc. Am. **130**, EL219–EL225, 2011.
4. B. M. Harker and B. E. Anderson, *Optimization of the array mirror for time reversal techniques used in a halfspace environment*, J. Acoust. Soc. Am. **133**, EL351–EL357, 2013.
5. G. K. Kocur, E. H. Saenger, C. U. Grosse, and T. Vogel, *Time reverse modeling of acoustic emissions in a reinforced concrete beam*, Ultrasonics **65**, 96–104, 2016.
6. I. Lokmer, G. S. O'Brien, D. Stich, and C. J. Bean, *Time reversal imaging of synthetic volcanic tremor sources*, Geophys. Res. Lett. **36**, 1–6, 2009.
7. C. S. Larmat, R. A. Guyer, and P. A. Johnson, *Tremor source location using time reversal: Selecting the appropriate imaging field*, Geophys. Res. Lett. **36**, 2–7, 2009.

8. T. Horstmann, R. M. Harrington, and E. S. Cochran, *Using a modified time-reverse imaging technique to locate low-frequency earthquakes on the San Andreas Fault near Cholame, California*, Geophys. J. Int. **203**, 1207–1226, 2015.
9. B. Steiner, E. H. Saenger, and S. M. Schmalholz, *Time reverse modeling of low-frequency microtremors: Application to hydrocarbon reservoir localization*, Geophys. Res. Lett. **35**, L03307, 2008.
10. C. S. Larmat, J. P. Montagner, M. Fink, Y. Capdeville, A. Tourin, and E. Clévédé, *Time-reversal imaging of seismic sources and application to the great Sumatra earthquake*, Geophys. Res. Lett. **33**, L19312, 2006.
11. C. S. Larmat, J. Tromp, Q. Liu, and J. P Montagner, *Time reversal location of glacial earthquakes*, J. Geophys. Res.-Solid **113**, 1–9, 2008.
12. C. Werner and E. H. Saenger, *Obtaining reliable source locations with time reverse imaging: limits to array design, velocity models and signal-to-noise ratios*, Solid Earth **9**, 1487–1505, 2018.
13. X. Zeng, C. H. Thurber, D. R. Shelly, R. M. Harrington, E. S. Cochran, N. L. Bennington, D. Peterson, B. Guo, and K. McClement, *3-D P- and S-wave velocity structure and low-frequency earthquake locations in the Parkfield, California region*, Geophys. J. Int. **206**, 1574–1585, 2016.
14. H. Lamb, *On waves in an elastic plate, Proceedings of the Royal Society of London A: Mathematical*, Physical and Engineering Sciences **93**, 114–128, 1917.
15. D. N. Alleyne and P. Cawley, *Optimization of Lamb wave inspection techniques*, NDT E Int. **25**, 11–22, 1992.
16. P. D. Wilcox, *A rapid signal processing technique to remove the effect of dispersion from guided wave signals*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **50**, 419–427, 2003.
17. J. Zemanek Jr., *An experimental and theoretical investigation of elastic wave propagation in a cylinder*, J. Acoust. Soc. Am. **51**, 265–283, 1972.
18. L. T. Nguyen, G. K. Kocur, and E. H. Saenger, *Defect mapping in pipes by ultrasonic wavefield cross-correlation*, Ultrasonics **90**, 153–165, 2018.
19. J. Zemanek, *Defect detection in pipes using guided waves*, Ultrasonics **36**, 147–154, 1998.
20. B. Vogelaar and M. Golombok, *Quantification and localization of internal pipe damage*, Mech. Syst. Signal Process. **78**, 107–117, 2016.
21. K. R. Leonard and M. K. Hinders, *Guided wave helical ultrasonic tomography of pipes*, J. Acoust. Soc. Am. **114**, 767–774, 2003.
22. T. Gaul, L. Schubert, B. Weihnacht, and B. Frankenstein, *Localization of defects in pipes using guided waves and synthetic aperture focussing technique (SAFT)*, in EWSHM-7th European Workshop on Structural Health Monitoring, 2014.
23. Q. Wang, M. Hong, and Z. Su, *A sparse sensor network topologized for cylindrical wave- based identification of damage in pipeline structures*, Smart Mater. Struct. **25**, 075015, 2014.
24. E. Dehghan-Niri and S. Salamone, *A multi-helical ultrasonic imaging approach for the structural health monitoring of cylindrical structures*, Struct. Health Monit. **14**, 73–85, 2015.
25. T. Hayashi, *Non-contact imaging of pipe thinning using elastic guided waves generated and detected by lasers*, Int. J. Press. Vessels Pip. **153**, 26–31, 2017.

26. A. Bagheri, P. Rizzo, and K. Li, *Ultrasonic imaging algorithm for the health monitoring of pipes*, J. Civil Struct. Health Monit. **7**, 99–121, 2017.
27. M. Fink, *Time reversal of ultrasonic fields. I. Basic principles*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 555–566, 1992.
28. R. K. Ing and M. Fink, *Time-reversed Lamb waves*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **45**, 1032–1043, 1998.
29. B. Xu and V. Giurgiutiu, *Single mode tuning effects on Lamb wave time reversal with piezoelectric wafer active sensors for structural health monitoring*, J. Nondestr. Eval. **26**, 123–134, 2007.
30. H. W. Park, S. B. Kim, and H. Sohn, *Understanding a time reversal process in Lamb wave propagation*, Wave Motion **46**, 451–467, 2009.
31. T. Hayashi and M. Murase, *Defect imaging with guided waves in a pipe*, J. Acoust. Soc. Am. **117**, 2134–2140, 2005.
32. J. Davies and P. Cawley, *The application of synthetic focusing for imaging crack-like defects in pipelines using guided waves*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **56**, 759–771, 2009.
33. H. Nishino, S. Takashina, F. Uchida, M. Takemoto, and K. Ono, *Modal analysis of hollow cylindrical guided waves and applications*, Jpn. J. Appl. Phys. **40**, 364, 2001.
34. A. Velichko and P. D. Wilcox, *Excitation and scattering of guided waves: Relationships between solutions for plates and pipes*, J. Acoust. Soc. Am. **125**, 3623–3631, 2009.
35. D. Komatitsch and J. Tromp, *Introduction to the spectral element method for three-dimensional seismic wave propagation*, Geophys. J. Int. **139**, 806–822, 1999.
36. J. F. Claerbout, *Toward a unified theory of reflector mapping*, Geophysics **36**, 467–481, 1971.
37. M. Schickert, M. Krause, and W. Müller, *Ultrasonic imaging of concrete elements using reconstruction by synthetic aperture focusing technique*, J. Mater. Civil Eng. **15**, 235–246, 2003.
38. F. K. Gruber, E. A. Marengo, and A. J. Devaney, *Time-reversal imaging with multiple signal classification considering multiple scattering between the targets*, J. Acoust. Soc. Am. **115**, 3042–3047, 2004.
39. M. Grohmann, S. Müller, E. Niederleithinger, and S. Sieber, *Reverse time migration: introducing a new imaging technique for ultrasonic measurements in civil engineering*, Near Surface Geophys. **15**, 242–258, 2017.
40. L. T. Nguyen and R. T. Modrak, *Ultrasonic wavefield inversion and migration in complex heterogeneous structures: 2D numerical imaging and nondestructive testing experiments*, Ultrasonics **82**, 357–370, 2018.
41. L. Wang and F. G. Yuan, *Damage identification in a composite plate using prestack reverse-time migration technique*, Struct. Health Monit. **4**, 195–211, 2005.
42. J. He and F.-G. Yuan, *Damage identification for composite structures using a cross-correlation reverse-time migration technique*, Struct. Health Monit. **14**, 558–570, 2015.
43. L. T. Nguyen, G. K. Kocur, and E. H. Saenger, *Defect mapping in pipes by ultrasonic wavefield cross-correlation: A synthetic verification*, Ultrasonics **90**, 153–165, 2018.

Theoretical Condensed Matter

Theoretical Condensed Matter

Kurt Binder[1] **and Frithjof B. Anders**[2]

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

[2] Lehrstuhl für theoretische Physik II, Fakultät Physik,
Technische Universität Dortmund, 44221 Dortmund, Germany
E-mail: frithjof.anders@tu-dortmund.de

Since the foundation of the John von Neumann Institute for Computing (NIC), or its preceding institution "Höchstleistungsrechenzentrum (HLRZ)", theoretical condensed matter clearly was one of the areas of physics which profited substantially from the access to top-class supercomputers. Significant progress could be obtained via large scale simulations for problems such as spin glasses, glass transition of simple fluids, dynamics of colloidal suspensions *etc.*

In the last decade, the focus of research has shifted somewhat from these problems, which could be studied with methods of computational classical statistical mechanics, to problems where the quantum nature of condensed matter is the central aspect. This shift of scientific interests coincides with the trend towards the development of quantum computing. This is not a completely new vision on the horizon - already twenty years ago D. Loss and D. DiVincenzo proposed that the spin of a single electron in a semiconductor quantum dot could form a physical realisation of "qubit".[1] Developing quantum computers is a very hard task, but given the facts that Prof. DiVincenzo is one of the directors of the Institute for Advanced Simulation at the Forschungszentrum Jülich, and that the Jülich Supercomputing Centre (JSC) has launched the JUNIQ initiative, which will give researchers access to quantum computing platforms as they will become available, it is welcome that basic research related to this general field is present at the NIC. Understanding the relaxation times of qubits in suitably produced arrays is a hot topic of current research (see *e. g.* the brief review in Rev. 2).

The first article of the present chapter, "Optical Pumping of Electron Spins in Quantum Dot Ensembles" by I. Kleinjohann *et al.*, is an impressive example for the insight that one can gain by targeted simulations that are inspired by closely related experiments, to which the results can be compared. The starting point is the "central spin model", where the spin of an electron in an external magnetic field is coupled to an ensemble of nuclear spins that provide a heat bath. The simulation addresses the optical pumping experiments, and can account for the magnetic field dependence of the observed dephasing time very well, and clarify the nature of the resulting effective spin-spin interactions.

Also the second article in this chapter, "Benchmarking Supercomputers with the Jülich Universal Quantum Computer Simulator" by D. Willsch *et al.*, deals with aspects of quantum computing. Since it is clear that in the near future quantum computers with large numbers of qubits will not be available, and if they become available their use for practical computations will not be straightforward it is of interest already now to simulate how they will perform, using the existing standard digital computers. Such simulations are

particularly useful and even necessary for the validation of real world realisations of quantum computers, and one of the authors contributed to the collaboration on the quantum supremacy using a programmable superconducting processor.[3] This is a nontrivial task either, since adding a single qubit requires a doubling of the available computer memory. So simulating a universal quantum computer with 45 qubits already will require more than 5×10^{14} bytes of memory! Only extremely powerful computers can provide such a memory together with the also required powerful network connecting the compute nodes. Therefore Willsch *et al.* use several massively parallel machines such as JUQUEEN, JURECA and JUWELS at the JSC, as well as the K computer in Japan and the Taihulight computer in China, to run the Jülich universal quantum computer simulator (JUQCS). The number of Qubits that can then be included range from 43 (JURECA) to 48 (K computer, Taihulight). Gratifying, it could be demonstrated that the normalised elapsed time per operation did increase only slightly with the number of qubits used. It is argued that this approach has the additional benefit to provide a new type of benchmarks for the standard high-performance computers.

The last article, "Numerical Simulations of Strongly Correlated Electron Systems" by F. F. Assaad, deals with the simulation of many fermion-systems using auxiliary field Quantum Monte Carlo algorithms. The focus is on the description of the so-called ALF library ("Algorithms for Lattice Fermions"), developed by F. F. Assaad and his group. This library is suitable for a broad class of models, that are of physical interest. The article also presents an analysis and results for the so-called Falicow-Kimball model as a typical example and case study. It is shown that the phase diagram in the plane of variables T (temperature) and U (strength of the Hubbard term) shows a phase transition (of second order) from an insulating phase to a Fermi-liquid (for small U) or metal (for large U), while the transition from Fermi-liquid to the metal is not a sharp phase transition but only a gradual crossover.

Thus, there is also steady progress with Quantum Monte Carlo methods, although due to the well-known "minus sign problem" this does not include all cases of interest. It may be hoped that quantum computers, when available, will allow the direct simulation of the time evolution of such particularly difficult problems.

References

1. D. Loss and D. P. DiVincenzo, *Quantum computation with quantum dots*, Phys. Rev. A **57**, 120, 1998.
2. L. M. K. Vandersypen and M. A. Eriksson, *Quantum computing with semiconductor spins*, Phys. Today **72**, 38, 2019.
3. F. Arute, K. Arya, R. Babbush *et al.*, *Quantum supremacy using a programmable superconducting processor*, Nature **574**, 505, 2019.

Optical Pumping of Electron Spins in Quantum Dot Ensembles

Iris Kleinjohann, Andreas Fischer, Natalie Jäschke, and Frithjof B. Anders

Lehrstuhl für theoretische Physik II, Fakultät Physik,
Technische Universität Dortmund, 44221 Dortmund, Germany
E-mail: {iris.kleinjohann,frithjof.anders}@tu-dortmund.de

We use the central spin model to analyse optical pumping of singly charged quantum dots. An exact quantum mechanical approach is employed to calculate the impact of several million laser pulses. The periodical excitation imprints on the system and the electron spin polarisation refocuses directly before each pulse. Our calculations link the non-monotonic behaviour of the revival amplitude as function of the external magnetic field found in experiments to the nuclear Larmor precession. Furthermore, we extend our model to a cluster of coupled central spin models to investigate long-range spin-spin interactions in quantum dot ensembles. A semiclassical approach handles the large amount of spins. By means of the cluster approach, we provide an understanding of the electron spin dephasing at different external magnetic fields and spectral widths of the optical excitation. Our model reveals the counterbalancing effects generating a constant dephasing time in dependence of the spectral laser width in experiment.

1 Introduction

Coherent control of electron spins in quantum dot (QD) ensembles appears to be a promising root for realisations of quantum information processing. The confined electron in a singly charged semiconductor QD has limited overlap with the environment leading to a long spin life-time. The major source of decoherence of the electron spin results from the Overhauser field generated by the hyperfine coupling to the surrounding nuclear spins.

A substantial revival amplitude of electron spin polarisation[1–5] arises during the periodic application of laser pump pulses onto an ensemble of QDs subject to a constant magnetic field perpendicular to the optical axis. A resonance condition for the electron spin precession imposes a peaked non-equilibrium distribution of the Overhauser field component parallel to the external magnetic field that emerges from the initial Gaussian distribution. This imprint on the nuclear spins is robust against decoherence mechanisms and thus, is still accessible after minutes without optical driving[5, 6] opening the door to long memory times.

The dephasing of the electron spin polarisation generated by optical excitation depends on various counterbalancing mechanism. One of the dominant effects, aside of the hyperfine interaction and varying electron g factors, is the effective coupling between electron spins of different QDs on the sample. This inter QD interaction was indicated in two-colour pump-probe experiments[7] and presents the opportunity to correlate the spin dynamics of different QDs via tailored optical excitation.

2 Modelling of the System

The Fermi contact hyperfine interaction between the electron spin localised in a self-assembled QD and the surrounding nuclear spin bath provides the largest contribution

to electron spin dephasing[8] in a single semiconductor QD. Other interactions such as the nuclear dipole-dipole interactions are several orders of magnitude smaller and, therefore, will be neglected in the following.[8–10] The Hamiltonian of the central spin model[11] (CSM) accounts for the effect of the external magnetic field $\vec{B}_{\rm ext}$ on electron $\vec{S}$ and nuclear spins $\vec{I}_k$, as well as for the hyperfine interaction (see Fig. 1 (a) for a sketch)

$$\hat{H}_{\rm CSM} = g_e \mu_B \vec{B}_{\rm ext} \hat{\vec{S}} + \sum_{k=1}^{N} g_N \mu_N \vec{B}_{\rm ext} \hat{\vec{I}}_k + \sum_{k=1}^{N} A_k \hat{\vec{I}}_k \hat{\vec{S}} \tag{1}$$

The ratio between nuclear and electronic Zeemann energy, which on average has a value $z = g_N \mu_N / (g_e \mu_B) = 1/800$ in InGaAs QDs, depends on the electron and the nuclear g factor, g_e and g_N, as well as on the respective magnetic momenta, μ_B and μ_N. In a nuclear spin bath of size N, we label the hyperfine coupling constant for the k-th nuclear spin A_k. The nuclear spins weighted by these coupling constants generate the Overhauser field,

$$\hat{\vec{B}}_N = (g_e \mu_B)^{-1} \sum_k A_k \hat{\vec{I}}_k \tag{2}$$

an effective magnetic field acting on the electron spin. In combination with the external magnetic field, it determines the electron Larmor precession. The fluctuations of the Overhauser field, $\langle \vec{B}_{\rm N}^2 \rangle$, in absence of an external magnetic field define a characteristic time scale of the system $(T_N^*)^{-2} = \sum_k A_k^2 \langle \vec{I}_k^2 \rangle$ governing the short-time electron spin dephasing which is typically a few ns in real systems.

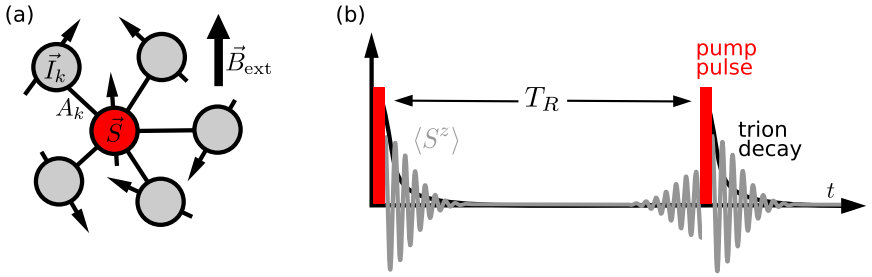


Figure 1. (a) In the central spin model (CSM) the electron spin $\vec{S}$ couples to a bath of nuclear spins $\vec{I}_k$. The external magnetic field $\vec{B}_{\rm ext}$ acts on electron and nuclear spins. (b) Pump pulses periodically generate electron spin polarisation $\langle S^z \rangle$ via trion excitation. The generated polarisation dephases while oscillating due to the magnetic field. As a result of synchronisation with the excitation periodicity, a revival of electron polarisation occurs before each laser pulse.

2.1 Optical Pumping

In optical pump-probe experiments, additional to the electron an electronic particle-hole excitation is created. The bound three particle state is called a trion, and its radiative decay is included in a Lindblad equation for the density operator $\hat{\rho}$ of the electron-nuclear spin system,

$$\dot{\hat{\rho}}(t) = -i \left[\hat{H}, \hat{\rho} \right] - \gamma \left(\hat{\rho} \hat{s} \hat{s}^\dagger + \hat{s} \hat{s}^\dagger \hat{\rho} - 2 \hat{s}^\dagger \hat{\rho} \hat{s} \right) \tag{3}$$

The operators $\hat{s} = |T\rangle\langle\uparrow|$, $\hat{s}^\dagger = |\uparrow\rangle\langle T|$ incorporate the transitions from the spin $|\uparrow\rangle$ to the trion state $|T\rangle$ and *vice versa*. The corresponding transition rate is typically $\gamma = 10\,\mathrm{ns}^{-1}$ in real systems. In the Hamiltonian $\hat{H}$ we include the CSM and the trion state. A scheme of the time scales in the periodically driven system with short optically pulses is shown in Fig. 1 (b). This dynamics can be treated in two ways, either with an exact quantum mechanical description or a semiclassical approach.

When focusing on quantum mechanical effects, we restrict ourselves to a small nuclear spin bath and calculate[12] the exact time evolution of the density operator. To compensate for the relatively small number of nuclear spins ($N = 6 - 8$), the hyperfine coupling constants A_k are drawn from their distribution function in large systems, and the results of N_C different realisations are averaged at the end. Typically about $N_C = 100$ configurations are used which allows for massive parallelisation of the calculations. The calculation of the time dependent density operator $\hat{\rho}(t)$, which consists of mainly matrix multiplications, is parallelised over all cores on one node and N_C nodes are used for a run. With that approach we can apply about 20 million laser pulses to the system with a repetition time of $T_R = 13.2$ ns and span 11 orders of magnitude in time, from the pulse duration of about 4 ps to 0.24 s in real time.

2.2 Semiclassical Approach

We also developed a hybrid approach[3] to combine the quantum description required to capture the effect of the periodic laser pulses and the consecutive trion decay. This semiclassical approach enables the treatment of larger nuclear spin baths[13] than in the quantum mechanical approach.

To this end, we incorporate[3] the quantum mechanical laser pulse by a mapping $\vec{S}_{\mathrm{prior}} \rightarrow \vec{S}_{\mathrm{after}}$ using a unitary time evolution operator and treat the Lindblad equation in the hybrid equation of motion (EOM). We extended[14] the EOM in our semiclassical approach to a coupled cluster theory comprising N_{QD} QDs labelled $i = 1, \ldots, N_{QD}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{S}^{(i)}(t) = \left(\vec{b}_T^{(i)} + g_e^{(i)}\vec{B}_{\mathrm{ext}}\right) \times \vec{S}^{(i)}(t) + \gamma P_T^{(i)}(0)\vec{\mathrm{e}}_z e^{-2\gamma t} \tag{4a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{I}_k^{(i)} = \left(A_k^{(i)}\vec{S}^{(i)} + z\vec{B}_{\mathrm{ext}}\right) \times \vec{I}_k^{(i)} \tag{4b}$$

$$\vec{b}_T^{(i)} = \sum_{k=1}^{N} A_k^{(i)}\vec{I}_k^{(i)} + \sum_{j\neq i} J_{ij}\vec{S}_j(t) \tag{4c}$$

with classical spin vectors and the probability of trion occupation P_T. The coupling between electron spins of different QDs, i and j, is named J_{ij}. Furthermore, we take into account different electronic g factors as well as off-resonant pumping caused by the mismatch between the trion excitation energy and the laser frequency.

In the semiclassical simulations, averaging results of various initial classical spin states mimics the quantum mechanical expectation value according to a saddle-point approximation.[13] Up to $N_C = 10^6$ configurations provide highly accurate numerical results. Each computing core handles 32 of these configurations simultaneously to enable vectorisation in 99 % of the program loops. The distribution of the independent configurations over several thousand cores is realised via MPI parallelisation.

3 Signatures of Long-Range Spin-Spin Interactions in an (In,Ga)As Quantum Dot Ensemble

In a two-colour spin pumping experiment[7] evidence was compiled that different QDs are coupled by a long-range spin-spin interaction of unknown origin. The experimental data are consistent with the assumption of a time independent Heisenberg term between pairs of electron spins localised in distinct QDs.

We benchmarked our approach by the simulation of a two size cluster ($N_{QD} = 2$) targeting the understanding of a two colour laser pumping experiment[7] that demonstrated the dependency of the decay time and the phase shifts on the delay time between pump pulses of different colours, acting on different QDs. We find that the theory is able to excellently reproduce the experimental results and establish that the effective spin-spin interaction is long-ranged and time constant. Our results disputes the original interpretation of the experimental paper[7] speculating about an optical Ruderman-Kittel-Kasuya-Yosida interaction (RKKY).

Recently, we extended the calculations to a cluster of up to $N_{QD} = 10$ QDs representing different sub-ensembles with comparable trion excitation energies. Laser pulses modelled by a Gaussian shape with spectral width ΔE have been used. Furthermore, we took into account a Gaussian distribution of electronic g factors and trion excitation energies whose properties have been determined by experiments using the photoluminescence spectra[15] of the QD ensemble. When investigating the dephasing of the electron spin polarisation after a laser pulse (see Fig. 2 (a)) we find three contributing mechanism. The dephasing arising due to the hyperfine coupling is governed by the dephasing time T_N^* encountered before. Additionally, the electron g factor distribution produces varying Larmor frequencies in the QD ensemble. The third contribution is ascribed to the interaction of an electron spin with surrounding unpolarised electron spins in other QDs. Combining these mechanism, the ensemble averaged dephasing time T^* of the electron spin polarisation, as measured in experiments, is obtained and analysed.

In Fig. 2 (b), we compare the experimental data with the calculated values of T^* for an averaged Heisenberg coupling between electron spins $\overline{J} = 0.4\,\mathrm{ns}^{-1}$. The value $\overline{J}$ reproducing the experimental features best is almost a factor 4 smaller than the value $J \approx 1\,\mu\mathrm{eV} \approx 1.5\,\mathrm{ns}^{-1}$ reported in Ref. 7. This smaller value of $\overline{J}$ is a consequence of modelling the ensemble by a cluster of ten QDs instead of a pair only.

The magnetic field dependence of the dephasing time T^* at a fixed spectral width $\Delta E = 1.5\,\mathrm{meV}$ is depicted in Fig. 2 (c). The numerical calculations with $\overline{J} = 0.4\,\mathrm{ns}^{-1}$ (triangles) as well as the experimental measurements (x markers) display a decrease of the dephasing time with increasing external magnetic field. For our choice of parameters, the numerical values of T^* are slightly bigger than in experiment, but show the same curvature. In Fig. 2 (d), we supplement a power fit $T^* \propto B_{\mathrm{ext}}^{-\alpha}$ to the numerical data. Like in the experimental results, the best agreement was achieved for $\alpha = 0.7$. Deviations from the power law with $\alpha = 1$ can be attributed to the importance of the hyperfine interaction and the electronic spin-spin interaction at lower magnetic fields.

These results will open the door for a new quest to determine the microscopic origin of the effective interaction. While a direct exchange interaction or a dipole-dipole interaction are too weak due to the typical QD distance of 100 nm, an indirect RKKY interaction mediated by carriers in the wetting layer is a potential candidate. Additional information

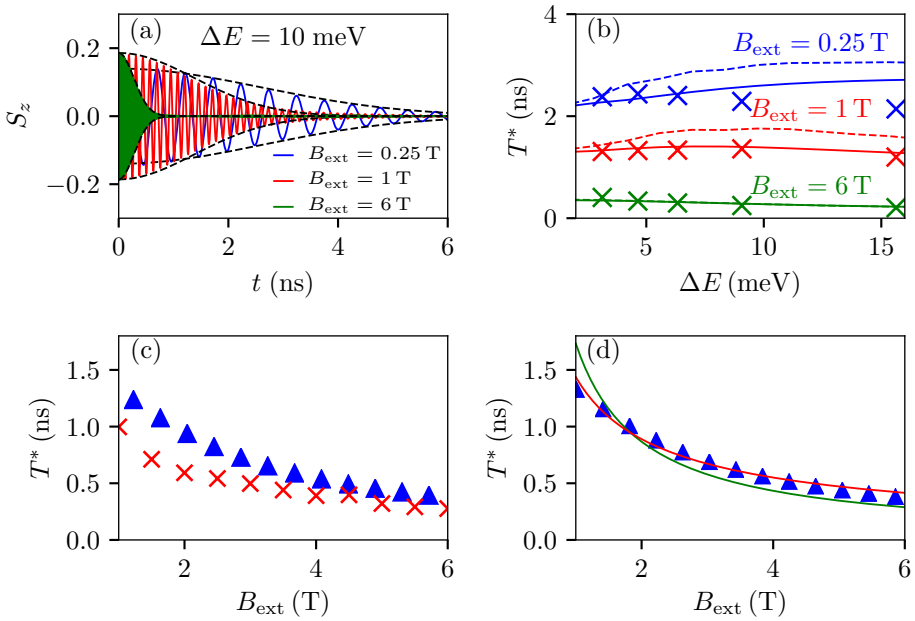


Figure 2. (a) Dephasing of the electron spin polarisation at different magnetic fields. (b) Dephasing time T^* as function of the spectral laser width ΔE. In addition to the numerical results after a single pulse (solid lines) and experimental data (x markers), we added the calculations after 100 pulses (dashed lines). (c) Dependence of T^* on the external magnetic field B_{ext}. Experimental data (x markers) and numerical results (triangles) are presented. (d) The data points of the numerical calculation (triangles) are compared to a dependency $\propto 1/B_{\mathrm{ext}}^{\alpha}$ with $\alpha = 1$ (green curve) and $\alpha = 0.7$ (red curve).

about the inter QD interaction might enable the design of pulse sequences in multi-colon setups that generate tailored quantum states and pave the way to quantum functionality.

4 Magnetic Field Dependence of the Electron Spin Revival Amplitude in Periodically Pulsed Quantum Dots

During a periodic pulse sequence with repetition time T_R, first the electron spin synchronises with the optical excitation and a purely electronic revival with an amplitude of approx. 0.077 evolves (see Fig. 3 (a) orange curve). As a consequence of the continuous periodic pulsing, the nuclear spins start to align generating a non-equilibrium Overhauser field distribution. The initially Gaussian distribution (dashed blue line in Fig. 3 (b)) transforms into a peaked structure (red curve). In combination with the external magnetic field, the Overhauser field enables either an integer or a half-integer number of electron revolutions between two pump pulses. In case of the external magnetic field $B_{\mathrm{ext}} = 1.95$ T for example, the peaks in the Overhauser field distribution correspond to an integer number of electron revolutions (indicated by gray dashed vertical lines). As a result of the non-equilibrium Overhauser field distribution, the electron spin revival amplitude either increases or de-

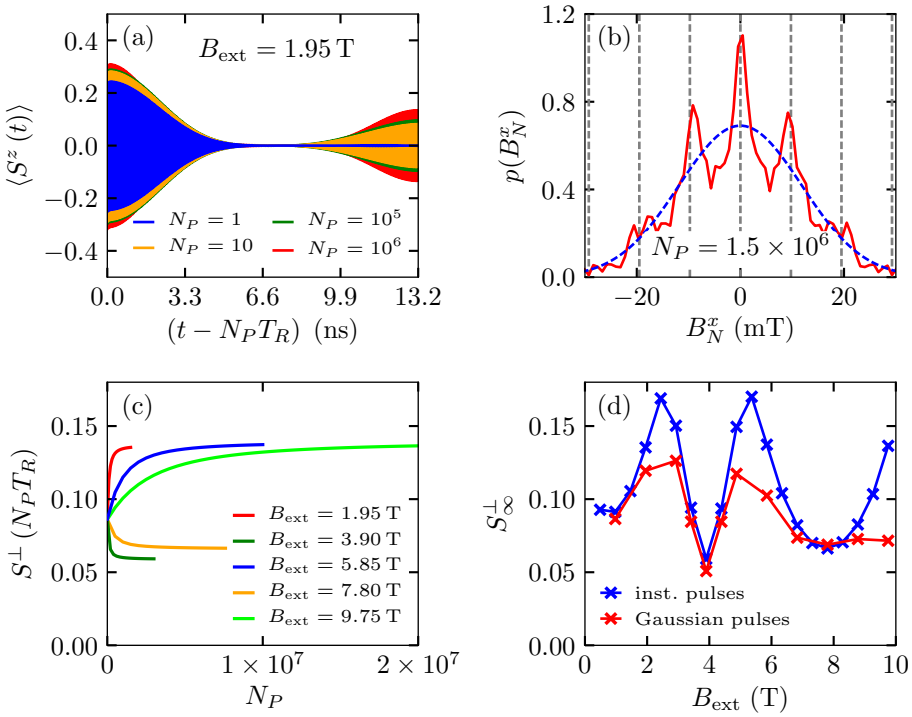


Figure 3. Evolution of electron spin polarisation $\langle S^z \rangle$ between two laser pulses. For $B_{\rm ext} = 1.95\,{\rm T}$ the revival grows due to the aligning of nuclear spins during the pulse sequence. (b) The Overhauser field distribution $p(B_N^x)$ along the magnetic field $B_{\rm ext} = 1.95\,{\rm T}$ leads to an integer number of electron revolutions during T_R. (c) The evolution of the revival amplitude $S^\perp(N_P T_R)$ with increasing number N_P of instantaneous optical pulses depends on $B_{\rm ext}$ and the peak structure evolving in $p(B_N^x)$. (d) Steady-state revival amplitude as function of the external magnetic field. The blue/red data points correspond to instantaneous/Gaussian laser pulses, respectively.

creases starting from the initial electronic value, as depicted in Fig. 3 (c). The speed of revival amplitude modification depends on the number of nuclear spin revolutions during the period T_R. At $B_{\rm ext} = 1.95$ T, for instance, the nuclear spins perform a quarter turn in between laser pulses. Due to the different resonance conditions, the steady-state amplitude reached after several million pump pulses in Fig. 3 (d) displays a non-monotonic behaviour as function of the external magnetic field.[12]

For different types of laser pulses employed in the simulation, instantaneous or Gaussian resonant π pulses, the electron revival is more or less efficiently generated. In case of instantaneous pulses the electron spin up state population is completely transferred to the trion state and thereupon decays. The longer the pulse duration becomes (or the stronger the external magnetic field is), the more revolutions the electron spin performs during the pulse which renders the pumping inefficient. This mechanism also yields a smaller revival amplitude for Gaussian pulses with 6 ps width especially for strong magnetic fields (see Fig. 3(d)).

The simulations with our quantum mechanical approach suited for long pulse sequences bring comprehension of experimental observations[12] that the revival amplitude displays a pronounced minimum at 4 T. The magnetic field dependence is even more complex in experiment due to different nuclear species with distinct resonance conditions that lead to a superimposed behaviour for the electron revival.

5 Concluding Remarks

We reported on our recent studies on singly-charged semiconductor quantum dots under optical excitation. We tackle the problem from two angles of view to analyse the nuclear induced frequency focusing effect that significantly enhances the spin coherence time as well as the long-range spin-spin interaction that enables information transfer between quantum dots. A numerically exact quantum approach is used to simulate several million laser pulses. By analysis of the resonance condition in the periodically driven system we explain the non-monotonic magnetic-field dependence of the revival amplitude observed in recent experiments.

Since the quantum approach is limited to a small number of spins, we utilise a semiclassical approach to simulate coupled clusters of quantum dots. The additional electronic spin-spin interaction that is incorporated in our cluster approach explains the dephasing time that is independent of the spectral laser width in experiment on quantum dot ensembles. Further investigations of the long-range spin-spin interaction might reveal the physical mechanism behind the effective coupling in the cluster approach.

Acknowledgements

This project was supported by the Deutsche Forschungsgemeinschaft through the transregio TRR 160 within the Projects No. A4 and No. A7. The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided for project HDO09 on the supercomputer JURECA at Jülich Supercomputing Centre (JSC). Furthermore, the authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project by providing computing time through the NIC on the GCS Supercomputers JUQUEEN and JUWELS at JSC.

References

1. W. Beugeling, G. S. Uhrig, and F. B. Anders, *Quantum model for mode locking in pulsed semiconductor quantum dots*, Phys. Rev. B **94**, 245308, 2016.
2. W. Beugeling, G. S. Uhrig, and F. B. Anders, *Influence of the nuclear Zeeman effect on mode locking in pulsed semiconductor quantum dots*, Phys. Rev. B **96**, 115303, 2017.
3. N. Jäschke, A. Fischer, E. Evers, V. V. Belykh, A. Greilich, M. Bayer, and F. B. Anders, *Nonequilibrium nuclear spin distribution function in quantum dots subject to periodic pulses*, Phys. Rev. B **96**, 205419, 2017.

4. A. Greilich, D. R. Yakovlev, A. Shabaev, Al. L. Efros, I. A. Yugova, R. Oulton, V. Stavarache, D. Reuter, A. Wieck, and M. Bayer, *Mode Locking of Electron Spin Coherences in Singly Charged Quantum Dots*, Science **313**, 341–345, 2006.
5. A. Greilich, A. Shabaev, D. R. Yakovlev, Al. L. Efros, I. A. Yugova, D. Reuter, A. D. Wieck, and M. Bayer, *Nuclei-Induced Frequency Focusing of Electron Spin Coherence*, Science **317**, 1896–1899, 2007.
6. N. Jäschke, F. B. Anders, and M. M. Glazov, *Electron spin noise under the conditions of nuclei-induced frequency focusing*, Phys. Rev. B **98**, 045307, 2018.
7. S. Spatzek, A. Greilich, Sophia E. Economou, S. Varwig, A. Schwan, D. R. Yakovlev, D. Reuter, A. D. Wieck, T. L. Reinecke, and M. Bayer, *Optical Control of Coherent Interactions between Electron Spins in InGaAs Quantum Dots*, Phys. Rev. Lett. **107**, 137402, 2011.
8. R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen, *Spins in few-electron quantum dots*, Rev. Mod. Phys. **79**, 1217–1265, 2007.
9. I. A. Merkulov, Al. L. Efros, and M. Rosen, *Electron spin relaxation by nuclei in semiconductor quantum dots*, Phys. Rev. B **65**, 205309, 2002.
10. M. I. Dyakonov, *Spin Physics in Semiconductors*, Springer Series in Solid-State Sciences **157**, Springer-Verlag Berlin Heidelberg, 2008.
11. M. Gaudin, *Diagonalisation d'une classe d'hamiltoniens de spin*, J. Physique **37**, 1087, 1976.
12. I. Kleinjohann, E. Evers, P. Schering, A. Greilich, G. S. Uhrig, M. Bayer, and F. B. Anders, *Magnetic field dependence of the electron spin revival amplitude in periodically pulsed quantum dots*, Phys. Rev. B **98**, 155318, 2018.
13. G. Chen, D. L. Bergman, and L. Balents, *Semiclassical dynamics and long-time asymptotics of the central-spin problem in a quantum dot*, Phys. Rev. B **76**, 045312, 2007.
14. A. Fischer, E. Evers, S. Varwig, A. Greilich, M. Bayer, and F. B. Anders, *Signatures of long-range spin-spin interactions in an (In,Ga)As quantum dot ensemble*, Phys. Rev. B **98**, 205308, 2018.
15. A. Greilich, D. R. Yakovlev, and M. Bayer, *Optical tailoring of electron spin coherence in quantum dots*, Solid State Communications **49**, 1466–1471, 2009.

Benchmarking Supercomputers with the Jülich Universal Quantum Computer Simulator

Dennis Willsch[1,2], **Hannes Lagemann**[1,2], **Madita Willsch**[1,2], **Fengping Jin**[1], **Hans De Raedt**[3], **and Kristel Michielsen**[1,2]

[1] Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {d.willsch, h.lagemann, m.willsch, f.jin, k.michielsen}@fz-juelich.de

[2] RWTH Aachen University, 52056 Aachen, Germany

[3] Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
E-mail: h.a.de.raedt@rug.nl

We use a massively parallel simulator of a universal quantum computer to benchmark some of the most powerful supercomputers in the world. We find nearly ideal scaling behaviour on the Sunway TaihuLight, the K computer, the IBM BlueGene/Q JUQUEEN, and the Intel Xeon based clusters JURECA and JUWELS. On the Sunway TaihuLight and the K computer, universal quantum computers with up to 48 qubits can be simulated by means of an adaptive two-byte encoding to reduce the memory requirements by a factor of eight. Additionally, we discuss an alternative approach to alleviate the memory bottleneck by decomposing entangling gates such that low-depth circuits with a much larger number of qubits can be simulated.

1 Introduction

The simulation of large universal quantum computers on digital computers is a difficult task since every increase in the number of simulated qubits by one corresponds to a multiplication of the required amount of memory by a factor of two. A simulation of a universal quantum computer with 45 qubits requires slightly more than 0.5 PB (0.5×10^{15} bytes) of memory. There exist only a few digital computers with this amount of memory and a powerful network connecting a large number of compute nodes.[1] Benchmarking such systems requires simulation software that can efficiently utilise the architecture of present day supercomputers. We present benchmarks of some of the most powerful supercomputers using the Jülich universal quantum computer simulator (JUQCS). A survey of JUQCS including its instruction set as well as benchmarks for Shor's algorithm[2] and the adder circuit[3] can be found in Ref. 4.

In this article, we use the term "universal quantum computer" for the theoretical pen-and-paper version of a gate-based quantum computer,[5] in which the operation of the device is defined in terms of simple, sparse unitary matrices, representing the quantum gates, acting on state vectors, without any reference to the real time evolution of physical systems. In particular, we use the word "universal" to refer to a simulation of the full state vector, independent of the particular quantum circuit, representing an algorithm in terms of a sequence of gates operated on the qubits. In addition, we explore a complimentary simulation approach that allows for an efficient simulation of a much larger number of qubits for low-depth circuits if only a subset of all amplitudes is required (see also Refs. 6–11).

Since the first massively parallel version of JUQCS was presented in 2007,[12] supercomputers have evolved significantly. Therefore, we have considered it timely to review and improve the computationally critical parts of the software and use it to benchmark some of the most powerful supercomputers that are operational today. The characteristics of the supercomputers that we have used for our benchmarks are summarised in Tab. 1.

	JUQUEEN	K computer	Sunway TaihuLight	JURECA	JUWELS
CPU	IBM PowerPC A2	eight-core SPARC64 VIIIfx	SW26010 manycore 64-bit RISC	Intel Xeon E5-2680 v3	Dual Intel Xeon Platinum 8168
clock frequency	1.6 GHz	2.0 Ghz	1.45 GHz	2.5 GHz	2.7 GHz
memory/node	16 GB	16 GB	32 GB	128 GB	96 GB
# threads/core used	1 – 2	8	1	1 – 2	1 – 2
# cores used	1 – 262144	2 – 65536	1 – 131072	1 – 6144	1 – 98304
# nodes used	1 – 16384	2 – 65536	1 – 32768	1 – 256	1 – 2048
# MPI processes used	1 – 524288	2 – 65536	1 – 131072	1 – 1024	1 – 2048
# qubits	46 (43)	48 (45)	48 (45)	43 (40)	46 (43)

Table 1. Overview of the computer systems used for benchmarking. The IBM Blue Gene/Q JUQUEEN[13] (decommissioned), JURECA,[14] and JUWELS[15] are located at the Jülich Supercomputing Centre in Germany, the K computer at the RIKEN Center for Computational Science in Kobe, Japan, and the Sunway TaihuLight[16] at the National Supercomputer Center in Wuxi, China. The row "# qubits" gives the maximum number of qubits that can be simulated with JUQCS–A (JUQCS–E). On JUWELS, the maximum number of qubits was limited to 43 (40) at the time of running the benchmarks.

JUQCS is portable software for digital computers and is based on Fortran 2003 code that was developed in 2007.[12] The revised version includes new operations to implement error-correction schemes and a parser for circuits specified in new quantum assembly flavours such as OpenQASM.[17, 18] JUQCS converts a quantum circuit into a form that is suitable for the simulation of the real-time dynamics of physical qubit models, such as NMR quantum computing[19] or quantum computer hardware based on superconducting circuits.[20]

The present version of JUQCS comes in two flavours. The first version, denoted by JUQCS–E, uses double precision (8-byte) floating point arithmetic and can be considered numerically exact (indicated by the E in the acronym). It has been used to simulate universal quantum computers with up to 45 qubits. The 45 qubit limit is only set by the amount of RAM memory available on the supercomputers listed in Tab. 1 (a universal simulation of N qubits using JUQCS–E requires slightly more than 2^{N+4} bytes).

A second version, denoted by JUQCS–A, trades memory for additional CPU time and has been used to simulate a universal quantum computer with up to 48 qubits. JUQCS–A uses adaptive coding to represent each amplitude of the quantum state with only two bytes.[4] This effectively reduces the memory requirements by a factor of eight relative to the one of JUQCS–E. The price to pay is a slightly longer execution time and a somewhat reduced numerical precision. However, we have found that the reduced precision (about 3 digits) is sufficiently accurate for all standard quantum circuits.[4]

We use the acronym JUQCS to refer to both versions of the software, while JUQCS–E and JUQCS–A are used to specifically refer to the numerically exact version and the version using adaptive coding, respectively. Since portability was an important design objective, we have so far refrained from using machine-specific programming. A separate

version of JUQCS–E utilising the potential of GPUs is under development.

The memory bottleneck can be alleviated by decomposing entangling two-qubit gates (entangling gates).[4] This trick can be used to great advantage if the number of entangling gates is not too large and if only a few of the coefficients of the final state vector need to be computed. The same idea has proven to be very useful in quantum Monte Carlo simulations.[21] Similar approaches have also been explored by other groups to simulate large random circuits with low depth.[6, 7, 9–11]

2 Simulating Universal Quantum Computers with JUQCS

In quantum theory, the state of a single qubit is represented by two complex numbers a_0 and a_1 which are normalised such that $|a_0|^2 + |a_1|^2 = 1$. A gate operation on the qubit changes these numbers according to

$$\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \leftarrow U \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \tag{1}$$

where U is a unitary 2×2 matrix. The state of N qubits is represented by a vector of 2^N complex numbers. Gate operations involving N qubits correspond to matrix-vector multiplications involving $2^N \times 2^N$ unitary matrices (see Ref. 5 for more information).

Typically, an N-qubit gate is expressed in terms of single-qubit gates (*i. e.* 2×2 matrices such as U in Eq. 1) or two-qubit gates, which only act on a subset of the qubits. Therefore, the matrices required to simulate N-qubit circuits are extremely sparse. In other words, a quantum gate circuit for a universal quantum computer is, in essence, a representation of several extremely sparse matrix-vector operations. This means that only a few arithmetic operations are required to update each of the 2^N coefficients of the state vector. Therefore, in principle, simulating universal quantum computers is rather simple as long as there is no need to use distributed memory and the access to the shared memory is sufficiently fast.[22–25] In practice, the speed to perform such operations is mainly limited by the bandwidth to (cache) memory.

For a large number of qubits, however, the only viable way to keep track of the 2^N complex coefficients is to use distributed memory, which comes at the expense of overhead due to communication between nodes, each of which can have several cores that share the memory (as is the case for all machines listed in Tab. 1). This makes up the "complicated" part of JUQCS that implements the MPI communication scheme. As described in Ref. 12, JUQCS reduces the communication overhead by minimising the transfer of data between nodes. As JUQCS can be configured to use a combination of OpenMP and MPI to especially tax the processors, the memory, the communication network, or any combination of these, it provides a practical framework to benchmark high-performance computers.

3 Validation and Benchmarking

To validate the operation of JUQCS, we have executed standard quantum algorithms as well as random circuits using all gates from the JUQCS instruction set for $N \leq 30$ qubits on both Windows and Linux systems. Validating the operation of JUQCS–E (JUQCS–A) when N reaches the limits $N = 45$ ($N = 48$), set by the amount of RAM available

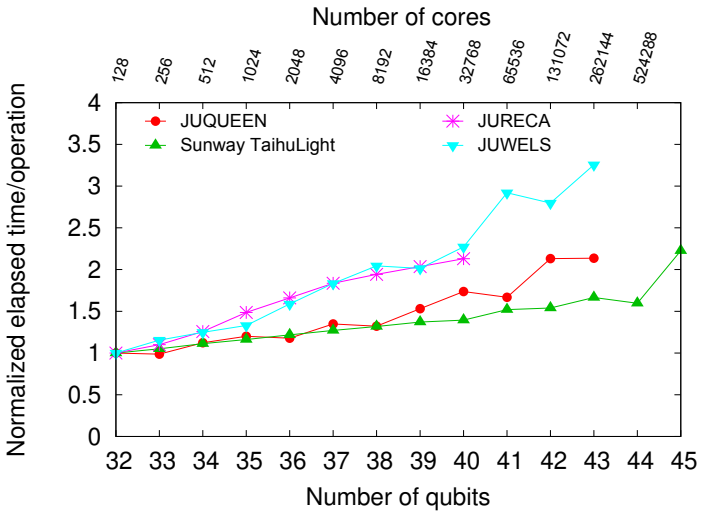


Figure 1. Weak scaling plot of JUQCS–E executing a Hadamard operation on qubit 0 and the sequence (CNOT 0 1), (CNOT 1 2), ..., (CNOT N-2, N-1), followed by a measurement of the expectation values of all qubits. Shown is the elapsed time per gate operation as a function of the number of qubits (normalised by the values corresponding to $N = 32$, *i. e.* 1.0 s (JUQUEEN), 5.1 s (Sunway TaihuLight), 1.4 s (JURECA), and 0.9 s (JUWELS)).

(see Tab. 1), is less trivial because of the requirement to use both MPI and OpenMP on distributed memory systems. For this reason, we made use of quantum circuits for which the exact outcome is known. In this article, we only present results for two particular representatives of such circuits since they are well suited to the purpose of benchmarking supercomputers. A more extensive discussion of the algorithms that were used to validate the operation and study the weak scaling behaviour of JUQCS is given elsewhere.[4]

The first circuit that we consider is the circuit that creates a uniform superposition over all 2^N basis states by performing a Hadamard gate on each of the N qubits. Since the number of states is too large to be verified by sampling, a practical method to check the result is to compute the single-qubit expectation values $\langle Q_\alpha(i)\rangle = (1 - \sigma_i^\alpha)/2$ where σ_i^α for $\alpha = x, y, z$ denotes the Pauli matrix on qubit i, the exact values being 0 for $\alpha = x$ and $1/2$ for $\alpha = y, z$. In this case, also JUQCS–A can compute the exact result since the encoding scheme is capable of representing the required amplitudes exactly.[4]

The second circuit is designed to create the maximally entangled state $(|0\ldots0\rangle + |1\ldots1\rangle)/\sqrt{2}$ by performing a Hadamard operation on the first qubit and a sequence of successive CNOT gates on qubits i and $i+1$ for $i = 0, 1, \ldots, N-2$. Weak scaling results of executing this circuit on the supercomputer systems listed in Tab. 1 are shown in Fig. 1 (JUQCS–E) and Fig. 2 (JUQCS–A).

We see that JUQCS beats the exponential increase in runtime with the number of qubits by doubling the computational resources with each added qubit. The weak scaling behaviour on the Sunway TaihuLight, the K computer, and JUQUEEN is close-to-ideal. However, the weak scaling behaviour on JURECA and JUWELS is not as good as the ones on the other supercomputers. Since the arithmetic work required for the Hadamard gate and the CNOT gate is rather low, the performance is mainly limited by the memory band-

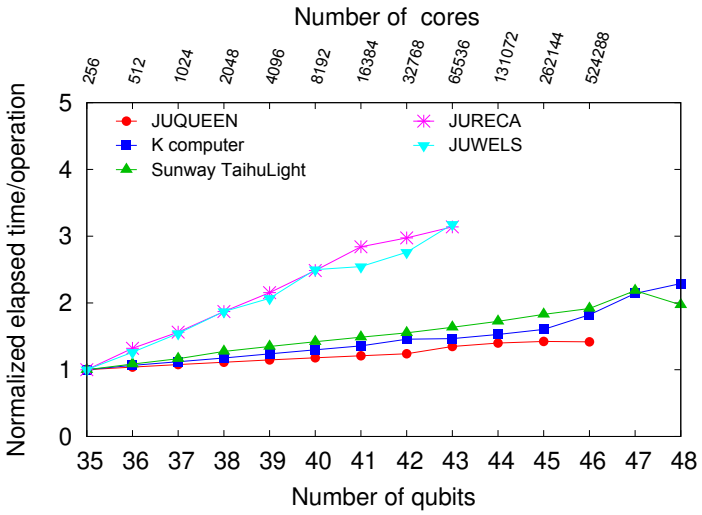


Figure 2. Weak scaling plot of JUQCS–A executing a Hadamard operation on qubit 0 and the sequence (CNOT 0 1), (CNOT 1 2), ..., (CNOT N-2, N-1), followed by a measurement of the expectation values of all qubits. Shown is the elapsed time per gate operation as a function of the number of qubits (normalised by the values corresponding to $N = 35$, *i. e.* 2.7 s (JUQUEEN), 3.8 s (K) , 19.9 s (Sunway TaihuLight), 2.4 s (JURECA), and 2.2 s (JUWELS)).

width. This suggests that there may be some limitations in the bandwidth to the memory and network on JURECA and JUWELS, compared to the other systems used in our benchmark. The best absolute run time is observed on JUWELS, closely followed by the other systems except the Sunway TaihuLight, which takes approximately four times longer.

Comparing Figs. 1 and 2, we also find that the computation time for JUQCS–E and JUQCS–A only differs by a factor of 1–4. The additional time for JUQCS–A is due to the encoding-decoding operation, which in turn affects the ratio between computation and communication. However, the additional time depends on the type of quantum gate. It can range, *e. g.* from almost zero for the CNOT gate to a factor of 2–3 for the Hadamard gate. As a result, comparing the computation times of JUQCS–A and JUQCS–E only makes sense for the same quantum circuit and even then, because of the difference in the number of qubits and the memory usage, interpreting differences in the elapsed times is not straightforward.

4 Memory Reduction by Decomposing Entangling Gates

There are many ways to alleviate the memory bottleneck. They are typically based on tensor-network contractions and (Schmidt) decompositions of two-qubit gates (cf. Refs. 6, 7, 9–11). In this section, we adopt an approach based on the decomposition of entangling gates using a discrete version of the Hubbard-Stratonovich transformation that has been used to great advantage in quantum Monte Carlo simulations.[21]

We start by expressing all entangling gates of an arbitrary circuit C in terms of single-qubit gates and the two-qubit CZ gate, which is always possible.[5] The action of the CZ gate on qubit i and qubit j (denoted by $\mathrm{CZ}_{i,j}$) is defined as a sign flip of all coefficients

with both qubits i and j in state $|1\rangle$. This action corresponds to the diagonal matrix $e^{i\pi(1+\sigma_i^z\sigma_j^z-\sigma_i^z-\sigma_j^z)/4}$, which can be decomposed into a sum of single-qubit operations according to

$$\mathrm{CZ}_{i,j} = \frac{1}{2} \sum_{s\in\{-1,1\}} e^{i(\sigma_i^z+\sigma_j^z)(xs-\pi/4)} \tag{2}$$

where x is a solution of $\cos(2x) = i$.

We partition all qubits $j = 0, \ldots, N-1$ into P mutually exclusive subsets labelled by $p = 0, \ldots, P-1$. The dimension of the corresponding subspace is denoted by $2 \leq D_p \leq 2^N$. If we decompose all CZ gates between different partitions p according to Eq. 2, we can express the circuit C as a sum of smaller subcircuits, $C = \sum_{\mathbf{s}\in\{-1,1\}^S} \bigotimes_{p=0}^{P-1} W_p(\mathbf{s})$, where $S \in \mathbb{N}$ denotes the total number of decomposed CZ gates, and $W_p(\mathbf{s})$ is a subcircuit that only acts on qubits in the partition p. Each subcircuit (a.k.a. simulation path) can be simulated independently with a quantum computer simulator such as JUQCS–E or JUQCS–A. Note that in the extreme case where all CZ gates are decomposed, we have $P = N$, $D_p = 2$, and $W_p(\mathbf{s})$ is a product of single-qubit gates on qubit $j = p$ only.

The quantum computer is initialised in the state $|\mathbf{0}\rangle$ of the computational basis. Consequently, the expression for the coefficient corresponding to the bit string $\mathbf{z}$ of the final state vector is given by

$$\langle \mathbf{z}|C|\mathbf{0}\rangle = \sum_{\mathbf{s}\in\{-1,1\}^S} \prod_{p=0}^{P-1} \langle \mathbf{z}_p|W_p(\mathbf{s})|\mathbf{0}_p\rangle \tag{3}$$

where $\mathbf{z}_p$ denotes the part of $\mathbf{z}$ belonging to the qubits contained in the partition p.

From Eq. 3 it follows that memory reduction can be achieved by separating the circuit into a sum of P subcircuits acting only on a D_p-dimensional subspace of the large 2^N-dimensional space. The computation of $\langle \mathbf{z}_p|W_p(\mathbf{s})|\mathbf{0}_p\rangle$ requires memory bound by the dimension D_p of this subspace. Obviously, reducing the memory requirements increases the number S of decomposed entangling gates and, consequently, also the computation time. This increase can be controlled by the choice of the P partitions.

A salient feature of the algorithm is that each term of the sum in Eq. 3 is independent. We have parallelised this sum using MPI to reduce the elapsed time by distributing the work on more cores. Furthermore, the evaluation of all $W_p(\mathbf{s})$'s is parallelised with OpenMP. If more than one coefficient of the final state vector is requested, a list of M coefficients $\{\langle \mathbf{z}|C|\mathbf{0}\rangle\}$ can be computed in a single run. The reason for this is that computations for different coefficients only differ by the index of the coefficient extracted from the vector $W_p(\mathbf{s})|\mathbf{0}_p\rangle$. Thus $W_p(\mathbf{s})|\mathbf{0}_p\rangle$ needs to be computed once for each partition p and M determines the number of coefficients extracted from this result.

In summary, this algorithm has a worst-case time complexity of $\mathcal{O}(2^S P \max\{D_p, M\}/RT)$ and a worst-case space complexity of $\mathcal{O}(R \max\{D_p, M\})$, where R denotes the number of MPI processes, T denotes the number of OpenMP threads, and M is the desired number of coefficients from the final state vector.

In Fig. 3, we present scaling results for the circuit used for Figs. 1 and 2, up to a maximum of $N = 128$ qubits and $M = 2^{27}$ extracted coefficients. The two coefficients corresponding to the states $|\mathbf{0}\ldots\mathbf{0}\rangle$ and $|\mathbf{1}\ldots\mathbf{1}\rangle$ are $1/\sqrt{2}$, as expected. Qualitatively, the

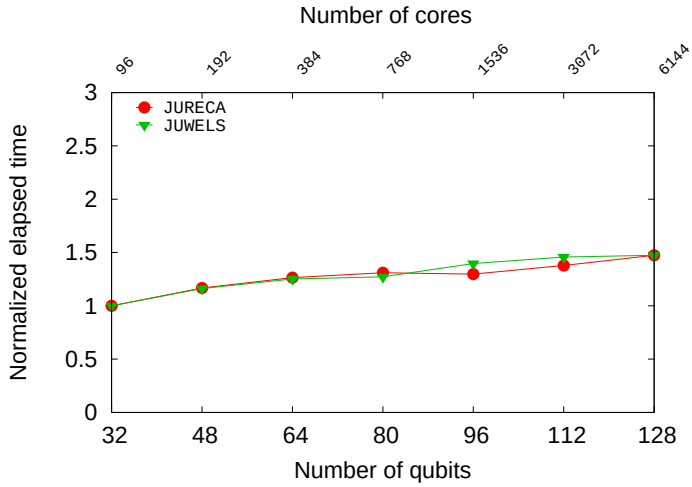


Figure 3. Scaling plot of a program implementing the memory reduction scheme according to Eq. 3, executing the same circuit as in Figs. 1 and 2. Shown is the elapsed time per gate operation as a function of the number of qubits (normalised by the values corresponding to $N = 32$, *i. e.* 2.12 s (JURECA) and 2.10 s (JUWELS)).

scaling behaviour of the algorithm on JURECA and JUWELS is nearly identical. Only the data points for the qubit numbers 80, 96, and 112 show small differences. Since all normalised run times are between 1 and 1.5 s, both JURECA and JUWELS show almost ideal scaling, suggesting that the previously observed limitations do not play a role for this kind of problem.

As the results presented in Fig. 3 are based on a circuit for which only a small number of entangling gates need to be decomposed, we also present results for a set of random quantum circuits, generated according to the procedure given by Boixo *et al.*[26] The circuits are characterised by their depth, which is the maximum number of layers (also called clock cycles or circuit moments) when all consecutive gates on different qubits are grouped into a single layer. These circuits pose a more difficult problem for the decomposition algorithm because the control and target qubits of the CZ gates are distributed in such a way that one cannot partition the circuit into smaller subcircuits without at least doubling the size S of decomposed entangling gates. Furthermore, the density of single-qubit gates is high. Consequently, at some point, increasing the number of MPI processes will not prevent the elapsed time from growing exponentially. In this respect, all algorithms based on a memory reduction using similar ideas[6, 7, 9–11] differ from the universal simulator JUQCS in that the scaling observed for JUQCS is almost independent of the particular simulated circuit.

Fig. 4 shows the normalised elapsed run times per operation for random circuit simulations as a function of the circuit depth. The number of simulated qubits is 42. We find that for circuits with a low depth, the run time is quasi constant because enough computational resources are available to distribute the work. For circuits with depth 13 and larger, we always use the same amount of computational resources, namely 512 MPI processes and 48 OpenMP threads per process. Consequently, we see a step-like increase in the run time

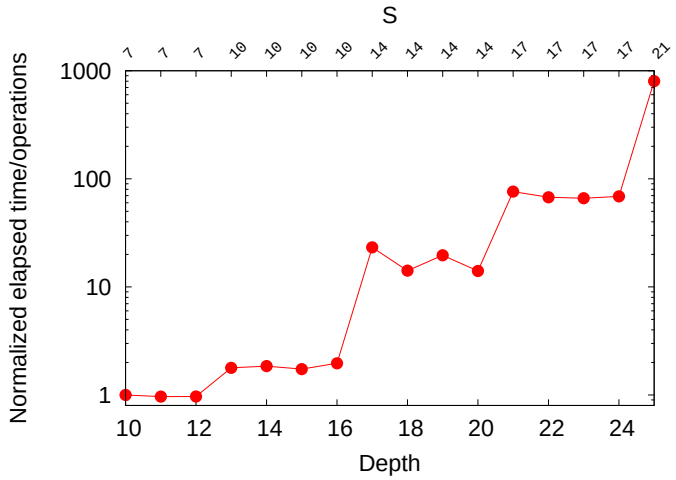


Figure 4. Scaling plot of a program implementing the memory reduction scheme according to Eq. 3, executing a set of random circuits for 42 qubits as a function of the circuit depth on JURECA. Shown is the elapsed time per gate operation, normalised by 0.0027 s corresponding to circuit depth 10. The axis on top of the figure shows the number S of decomposed entangling gates. The total number of gates ranges from 317 (depth 10) to 708 (depth 25). The circuits have been partitioned into two subcircuits of 21 qubits. The number of extracted coefficients is $M = 2^{20}$.

at depth 17, 21, and 25. These results show that, within reasonable fluctuations, the run time scales according to the time complexity discussed above when the parameters D_p, R, and T are held constant and only S changes.

5 Conclusion

The massively parallel quantum computer simulator JUQCS has been used to benchmark the Sunway TaihuLight,[16] the K computer, the IBM BlueGene/Q JUQUEEN,[13] and the Intel Xeon based clusters JURECA[14] and JUWELS[15] by simulating quantum circuits with up to $N = 48$ qubits. We observed close-to-linear scaling of the elapsed time as a function of the number of qubits on all tested supercomputers. The scaling performance on JUQUEEN, the Sunway TaihuLight, and the K computer tends to be better than on JURECA and JUWELS, suggesting some limitations in the bandwidth of the latter. The absolute execution times were best on JUWELS, closely followed by JUQUEEN, JURECA, and the K computer. The simulation on the Sunway TaihuLight was approximately four times slower.

Two methods to circumvent the memory problem associated with the simulation of quantum systems on a digital computer have been explored. The first uses an adaptive coding scheme to represent the quantum state in terms of 2-byte instead of 16-byte numbers. We observed that the reduction in memory has no significant impact on the accuracy of the outcomes (see also Ref. 4).

The second method uses a technique known from Quantum Monte Carlo simulations to express two-qubit gates in terms of sums of single-qubit gates. Using this technique,

we observed nearly ideal scaling on JURECA and JUWELS for a maximally entangling circuit with up to $N = 128$ qubits. Additionally, we used the method to simulate random 42-qubit circuits up to depth 25.

As the new generation of high-performance computers relies on accelerators or GPUs to deliver even more FLOPS, we have started to develop a CUDA-based version of JUQCS to explore and benchmark the potential of using GPUs for simulating universal quantum computers.

Since JUQCS can easily be configured to put a heavy burden on the processors, the memory, the communication network, or any combination of them, it may be a useful addition to the suite of benchmarks for high-performance computers.

Acknowledgements

We thank Koen De Raedt for his help in improving the JUQCS code. The authors acknowledge the computing time granted by the JARA-HPC Vergabegremium and provided on the JARA-HPC Partition part of the supercomputers JURECA[14] and JUQUEEN[13] at the Forschungszentrum Jülich. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project by providing computing time on the GCS Supercomputer JUWELS[15] at Jülich Supercomputing Centre (JSC). D. W. is supported by the Initiative and Networking Fund of the Helmholtz Association through the Strategic Future Field of Research project "Scalable solid state quantum computing (ZT-0013)". Part of the simulations reported in this paper were carried out on the K computer at RIKEN Center for Computational Science in Kobe, Japan, and the Sunway TaihuLight[16] at the National Supercomputer Center in Wuxi, China.

References

1. TOP500: `https://www.top500.org`
2. P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM Review **41**, 303, 1999.
3. T. G. Draper, *Addition on a Quantum Computer*, 2000, arXiv:quant-ph/0008033.
4. H. De Raedt, F. Jin, D. Willsch, M. Willsch, N. Yoshioka, N. Ito, S. Yuan, and K. Michielsen, *Massively parallel quantum computer simulator, eleven years later*, Comp. Phys. Comm. **237**, 47–61, 2019.
5. M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 10th anniversary edition, 2010.
6. E. Pednault, J. A. Gunnels, G. Nannicini, L. Horesh, T. Magerlein, E. Solomonik, and R. Wisnieff, *Breaking the 49-Qubit Barrier in the Simulation of Quantum Circuits*, 2017, arXiv:1710.05867 [quant-ph].
7. S. Boixo, S. V. Isakov, V. N. Smelyanskiy, and H. Neven, *Simulation of low-depth quantum circuits as complex undirected graphical models*, 2017, arXiv:1712.05384 [quant-ph].
8. J. Chen, F. Zhang, C. Huang, M. Newman, and Y. Shi, *Classical Simulation of Intermediate-Size Quantum Circuits*, 2018, arXiv:1805.01450 [quant-ph].
9. Z. Chen, Q. Zhou, C. Xue, X. Yang, G. Guo, and G. Guo, *64-qubit quantum circuit simulation*, Sci. Bull. **63**, 964–971, 2018.

10. I. L. Markov, A. Fatima, S. V. Isakov, and S. Boixo, *Quantum Supremacy Is Both Closer and Farther than It Appears*, 2018, arXiv:1807.10749 [quant-ph].
11. B. Villalonga, D. Lyakh, S. Boixo, H. Neven, T. S. Humble, R. Biswas, E. G. Rieffel, A. Ho, and S. Mandrà, *Establishing the Quantum Supremacy Frontier with a 281 Pflop/s Simulation*, 2019, arXiv:1905.00444 [quant-ph].
12. K. De Raedt, K. Michielsen, H. De Raedt, B. Trieu, G. Arnold, M. Richter, Th. Lippert, H. Watanabe, and N. Ito, *Massively parallel quantum computer simulator*, Comp. Phys. Comm. **176**, 121–136, 2007.
13. M. Stephan and J. Doctor, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, J. of Large-Scale Res. Facil. **1**, A1, 2015.
14. D. Krause and P. Thörnig, *JURECA: Modular supercomputer at Jülich Supercomputing Centre*, J. of Large-Scale Res. Facil. **4**, A132, 2018.
15. D. Krause, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, J. of Large-Scale Res. Facil. **5**, A135, 2019.
16. W. Zhang, J. Lin, W. Xu, H. Fu, and G. Yang, *SCStore: Managing Scientific Computing Packages for Hybrid System with Containers*, Tsinghua Science and Technology **22**, 675–681, 2017.
17. IBM Q, "Quantum experience", `https://www.research.ibm.com/ibm-q/`.
18. A. W. Cross, L. S. Bishop, J. A. Smolin, and J. M. Gambetta, *Open Quantum Assembly Language*, 2017, arXiv:1707.03429 [quant-ph].
19. H. De Raedt, K. Michielsen, A. Hams, S. Miyashita, and K. Saito, *Quantum Spin Dynamics as a Model for Quantum Computer Operation*, Eur. Phys. J. **B27**, 15–28, 2002.
20. D. Willsch, M. Nocon, F. Jin, H. De Raedt, and K. Michielsen, *Gate-error analysis in simulations of quantum computers with transmon qubits*, Phys. Rev. A **96**, 062302, 2017.
21. J. E. Hirsch, *Discrete Hubbard-Stratonovich transformation for fermion lattice models*, Phys. Rev. B **28**, 4059–4061, 1983.
22. H. De Raedt and K. Michielsen, *Computational Methods for Simulating Quantum Computers*, in Handbook of Theoretical and Computational Nanotechnology, M. Rieth and W. Schommers (Editors), American Scientific Publishers, Los Angeles, 2–48, 2006.
23. A. Aspuru-Guzik M. Smelyanskiy, and N. P. D. Sawaya, *qHiPSTER: The Quantum High Performance Software Testing Environment*, 2016, arXiv:1601.07195 [quant-ph].
24. N. Khammassi, I. Ashraf, X. Fu, C. G. Almudever, and K. Bertels, *QX: A high-performance quantum computer simulation platform*, in Design, Automation Test in Europe Conference Exhibition, 464–469, 2017.
25. Th. Häner and D. S. Steiger, *0.5 Petabyte Simulation of a 45-qubit Quantum Circuit*, in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC'17, ACM, 33:1–33:10, 2017.
26. S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Characterizing quantum supremacy in near-term devices*, Nat. Phys. **14**, 595–600, 2018.

Numerical Simulations of Strongly Correlated Electron Systems

Fakher F. Assaad

Institut für Theoretische Physik und Astrophysik and
Würzburg-Dresden Cluster of Excellence ct.qmat,
Universität Würzburg, Am Hubland, 97074 Würzburg, Germany
E-mail: assaad@physik.uni-wuerzburg.de

The richness of emergent phenomena that stem from the fundamental laws of quantum mechanics is astonishing. Topology, inherent to the integer Hall effect and Chern insulators, allows us to understand why a dirty two-dimensional electron gas can provide the most precise determination of fundamental constants. Electron correlations lead to the notion of fractionalisation and associated emergent lattice gauge theories widely studied in high energy physics. Finally, quantum engineering leads to amazing possibilities for designing novel materials and nanostructures that may very well define the building blocks of information technologies beyond silicon. Given this fascinating richness of phenomena, the natural question to ask for a numerically oriented researcher is: can one develop a flexible and efficient program package that allows one to define and simulate, at minimal programming cost, a wide set of model Hamiltonians? We have recently written an open source library, coined **A**lgorithms for **L**attice **F**ermions (ALF) that allows us to study a large variety of designer and realistic models. In this article, we will summarise aspects of the ALF-library, demonstrate its range of application and then concentrate on the case study of fractionalisation in a Falicov-Kimball model.

1 Introduction

Consider a single spin-1/2 degree of freedom defined in a two dimensional Hilbert space. The quantum mechanical state of N spin-1/2 degrees of freedom on a graph, that for instance provides a minimal model for undoped high temperature cuprate superconductors, corresponds to a vector in a 2^N dimensional Hilbert space. The temporal evolution of this state requires diagonalisation of a $2^N \times 2^N$ Hamiltonian matrix $\hat{H}$. Since for a macroscopic cuprate sample, $N = 10^{23}$ and the task is generically out of reach. How is it that we can nevertheless solve so called non-frustrated spin models in thermodynamic equilibrium on arbitrarily large lattices? The answer lies in the formulation of the partition function in terms of a Feynman path integral:

$$Z = \mathrm{Tr} e^{-\hat{H}/k_B T} = \int d\phi(\boldsymbol{x}, \tau) e^{-S(\phi)} \tag{1}$$

Here, T is the temperature, k_B the Boltzmann constant, $\boldsymbol{x}$ a site of the graph, τ the imaginary time running from 0 to $1/k_B T$ and finally S the action that is a functional of ϕ. For fermion systems, the action generically takes the form $S = S_0(\phi) - \log\det M(\phi)$ where S_0 is the action of the bosonic field and the determinant arises from the coupling of the scalar field to the fermionic degrees of freedom. The above equation is the basis for auxiliary field QMC[1] algorithms used in the lattice gauge and solid state communities. The key point is that for a given configuration of fields, one can calculate the action in polynomial time. In our implementation, we compute explicitly the determinant, and the computation

cost to evaluate the action for a single field configuration scales as $N^3/(k_BT)$. The functional integral can be computed stochastically. This can can be achieved in polynomial time only in special cases.

- **The sign problem.** To use Monte Carlo sampling we need a positive definite cost function. Hence the action has to be real. The path integral is by no means unique and one of the key questions is to ask if it is possible to find a formulation with real action that one can compute for a given field in polynomial time. This requirement defines a class of so called negative sign free problems. In this domain, tremendous progress has been achieved in terms of symmetry properties of the action that lead to negative sign free formulations.[2–5] Furthermore, and in the spirit of universality, the sign problem can be avoided by cleverly defining models that capture the relevant – negative sign free – physics and omits the irrelevant interactions that can generate negative sign problems. Such a designer model approach to study universal properties at criticality or to understand properties of specific phases is very much en vogue in the solid state community. A recent biased selection includes the following Refs. 6–10. If one is not able to find a negative sign free formulation, then one can use reweighing techniques to nevertheless formulate a Monte Carlo sampling. For a given precision the computational time will scale as $e^{\Delta N/k_BT}$ where Δ is a formulation dependent positive constant. There is a body of research that aims to find formulations of quantum many body problems that ease the negative sign problem by minimising the value of Δ. Here one can mention efforts based on the representation of the path integral in terms of a Lefschetz thimble decomposition.[11, 12] On each thimble the fermion sign is constant, but the problem is caused by the number of thimbles. The number of thimbles is however formulation dependent and optimal representations have recently been proposed.[13, 14]

- **Sampling.** The sign problem is only one of many issues. The action can be real, but the distribution can have fat tails that inhibit the very notion of Monte Carlo sampling.[15] Even for well defined distributions, critical slowing down remains an issue. For example there are many electron-phonon problems that are free of the negative sign problem but that suffer from very long autocorrelation times that inhibit precise calculations on large lattices. In the present version of the ALF library,[16] we are using discrete fields and global updating methods such as Langevin dynamics or Hybrid Monte Carlo are not applicable. We are planing to implement these updating schemes in a future release. We note that machine learning as a tool to propose global moves is presently en vogue.[17]

In the above, we have highlighted the challenges. There is nevertheless a number of model Hamiltonians that we can simulate efficiently. In what follows, we will briefly summarise the workings of the ALF-Library[16] and then concentrate on a case study of fractionalisation in a Falicov-Kimball model.

2 The ALF

The ALF-library is programmed following Fortran-2003 standard and comes with an MPI-implementation. We are only dependent on BLAS and LAPACK libraries and

all the heavy numerical calculations are done through calls to these libraries. Hence, if they are well optimised on the supercomputers we use, our code will perform well. For the program development, we have established a git software repository at `http://alf.physik.uni-wuerzburg.de` and a small steering committee meets regularly so as to discuss future developments and merge requests of feature branches into the master branch. On top of this, at each commit we run a series of tests, so as to at best track down bugs. For each project, we fork the master branch and if required implement the novel Hamiltonian. With this organisation, each run can be associated to a commit hash so that one can keep track of the version of the program that has produced the results. This organisation allows us to discuss and develop codes collaboratively, and has proven to be very efficient.

Being a Monte Carlo method, the ALF dumps the bins for a given observable on the disc. Although we have a feature branch aiming at using HDF5 file formats, the open source branch still prints out the data in scientific format. The ALF-library, comes with an independent error analysis suite that reads in the bins and produces final results. The path integral is formulated in imaginary time, so that to make contact with spectroscopic experiments, we have to carry out an analytical continuation. The present ALF-library comes with an implementation of the stochastic Maximum Entropy method[18] to carry out the continuation. This allows us to produce data that can be compared directly with experiment.

It is beyond the scope of this article to provide a detailed account of the ALF project, and the interested reader is referred to the documentation[16] and to our git instance at `http://alf.physik.uni-wuerzburg.de`. Here we will summarise one of the key points, namely the definition of the Hamiltonian on which the program package is based. Of course, the aim is to define a general Hamiltonian which can accommodate a large class of models. Our approach is to express the model as a sum of one-body terms, a sum of two-body terms each written as a perfect square of a one body term, as well as a one-body term coupled to an Ising field with dynamics to be specified by the user. The form of the interaction in terms of sums of perfect squares allows us to use generic forms of discrete approximations to the Hubbard Stratonovitch (HS) transformation.[19, 20] Symmetry considerations are imperative to enhance the speed of the code. We therefore include a *colour* index reflecting an underlying $SU(N)$ colour symmetry as well as a flavour index reflecting the fact that after the HS transformation, the fermionic determinant is block diagonal in this index.

The class of solvable models includes Hamiltonians $\hat{\mathcal{H}}$ that have the following general form:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_T + \hat{\mathcal{H}}_V + \hat{\mathcal{H}}_I + \hat{\mathcal{H}}_{0,I}, \text{ where} \tag{2}$$

$$\hat{\mathcal{H}}_T = \sum_{k=1}^{M_T} \sum_{\sigma=1}^{N_{\mathrm{col}}} \sum_{s=1}^{N_{\mathrm{fl}}} \sum_{x,y}^{N_{\mathrm{dim}}} \hat{c}^\dagger_{x\sigma s} T^{(ks)}_{xy} \hat{c}_{y\sigma s} \equiv \sum_{k=1}^{M_T} \hat{T}^{(k)} \tag{3}$$

$$\hat{\mathcal{H}}_V = \sum_{k=1}^{M_V} U_k \left\{ \sum_{\sigma=1}^{N_{\mathrm{col}}} \sum_{s=1}^{N_{\mathrm{fl}}} \left[\left(\sum_{x,y}^{N_{\mathrm{dim}}} \hat{c}^\dagger_{x\sigma s} V^{(ks)}_{xy} \hat{c}_{y\sigma s} \right) + \alpha_{ks} \right] \right\}^2 \equiv \sum_{k=1}^{M_V} U_k \left(\hat{V}^{(k)} \right)^2 \tag{4}$$

$$\hat{\mathcal{H}}_I = \sum_{k=1}^{M_I} \hat{Z}_k \left(\sum_{\sigma=1}^{N_{\mathrm{col}}} \sum_{s=1}^{N_{\mathrm{fl}}} \sum_{x,y}^{N_{\mathrm{dim}}} \hat{c}^\dagger_{x\sigma s} I^{(ks)}_{xy} \hat{c}_{y\sigma s} \right) \equiv \sum_{k=1}^{M_I} \hat{Z}_k \hat{I}^{(k)} \tag{5}$$

The indices and symbols have the following meaning:

- The number of fermion *flavours* is set by N_{fl}. After the HS transformation, the action will be block diagonal in the flavour index.

- The number of fermion *colours* is set by N_{col}. The Hamiltonian is invariant under $SU(N_{\mathrm{col}})$ rotations.

- N_{dim} is the total number of spacial vertices: $N_{\mathrm{dim}} = N_{\text{unit cell}} N_{\mathrm{orbital}}$, where $N_{\text{unit cell}}$ is the number of unit cells of the underlying Bravais lattice and N_{orbital} is the number of (spacial) orbitals per unit cell.

- The indices x and y label lattice sites where $x, y = 1, \cdots, N_{\mathrm{dim}}$.

- Therefore, the matrices $\boldsymbol{T}^{(ks)}$, $\boldsymbol{V}^{(ks)}$ and $\boldsymbol{I}^{(ks)}$ are of dimension $N_{\mathrm{dim}} \times N_{\mathrm{dim}}$.

- The number of interaction terms is labelled by M_V and M_I. $M_T > 1$ would allow for a checkerboard decomposition.

- $\hat{c}^{\dagger}_{y\sigma s}$ is a second quantised operator that creates an electron in a Wannier state centred around lattice site y, with colour σ, and flavour index s. The operators satisfy the anti-commutation relations:

$$\left\{\hat{c}^{\dagger}_{y\sigma s}, \hat{c}_{y'\sigma' s'}\right\} = \delta_{x,x'}\delta_{s,s'}\delta_{\sigma,\sigma'}, \quad \text{and} \quad \left\{\hat{c}_{y\sigma s}, \hat{c}_{y'\sigma' s'}\right\} = 0 \tag{6}$$

The Ising part of the general Hamiltonian (Eq. 2) is $\hat{\mathcal{H}}_{0,I} + \hat{\mathcal{H}}_I$ and has the following properties:

- $\hat{Z}_k$ is an Ising spin operator which corresponds to the Pauli matrix $\hat{\sigma}_z$. It couples to a general one-body term.

- The dynamics of the Ising spins is given by $\hat{\mathcal{H}}_{0,I}$. This term is not specified here; it has to be specified by the user and becomes relevant when the Monte Carlo update probability is computed in the code.

Note that the matrices $\boldsymbol{T}^{(ks)}$, $\boldsymbol{V}^{(ks)}$ and $\boldsymbol{I}^{(ks)}$ explicitly depend on the flavour index s but not on the colour index σ. The colour index σ only appears in the second quantised operators such that the Hamiltonian is manifestly $SU(N_{\mathrm{col}})$ symmetric. We also require the matrices $\boldsymbol{T}^{(ks)}$, $\boldsymbol{V}^{(ks)}$ and $\boldsymbol{I}^{(ks)}$ to be Hermitian.

We are continuously developing the library. Although the open source public fork only contains single spin flip updates, our git repository for development, includes parallel tempering schemes, global updates, as well as continuous fields. Continuous fields were important so as to implement the long range Coulomb repulsion required to provide a realistic modelling of graphene.[21] Future developments include the implementation of hybrid molecular dynamics, Langevin dynamics as well as imaginary time dependent Hamiltonians.

3 A Case Study: Fractionalisation in a Falicov-Kimball Model

Here we will summarise work that was carried out in collaboration in M. Hohenadler and that was published in Refs. 9, 22. The Falicov-Kimball model is a simplification of the SU(2) Hubbard model in which one of the two fermion flavours acquires an infinite mass. One can generalise this notion to the case of the SU(3) Hubbard model, and the Hamiltonian reads:

$$\hat{H}^{cQ} = -t \sum_{\langle ij \rangle, \sigma} \left(\hat{c}^{\dagger}_{i\sigma} \hat{c}_{j\sigma} + \text{H.c.} \right) - U \sum_{i} \left(\hat{n}_{i\uparrow} - \tfrac{1}{2} \right) \left(\hat{n}_{i\downarrow} - \tfrac{1}{2} \right) \hat{Q}_i \tag{7}$$

We use the same notation as above, σ runs over the two spin components, and $\hat{Q}_i = \pm 1$ is an Ising variable. Importantly, $\hat{Q}_i$ is a locally conserved quantity $\left[\hat{H}^{cQ}, \hat{Q}_i \right] = 0$ and can be understood as a modulation of the sign of the Hubbard term depending upon the site occupation of the infinite mass third fermion. As it stands, it is highly non-trivial to see that this model has phases where the electron degree of freedom fractionalises into two entities:

$$\hat{c}^{\dagger}_{i\sigma} \mapsto \hat{f}^{\dagger}_{i\sigma} \hat{s}^{z}_{i} \,, \quad \hat{c}_{i\sigma} \mapsto \hat{f}_{i\sigma} \hat{s}^{z}_{i} \tag{8}$$

Here, the so called orthogonal fermion, $\hat{f}^{\dagger}_{i\sigma}$, carries an electron charge as well as a Z_2 charge and the Ising degree of freedom, $\hat{s}^{z}_{i}$, only a Z_2 charge. The composite object, the fermion, possesses the usual quantum numbers. Since the Hilbert space per site consists of a fermion degree of freedom, $\hat{c}^{\dagger}_{i\sigma}$, and the Ising field, $\hat{Q}_i$, the above substitution does not expand the Hilbert space so that no constraint is required. In the orthogonal fermion formulation, we can write the Ising variable as:

$$\hat{Q}_i = \hat{s}^{x}_{i} (-1)^{\sum_\sigma \hat{f}^{\dagger}_{i\sigma} \hat{f}_{i\sigma}} = \hat{s}^{x}_{i} (-1)^{\sum_\sigma \hat{c}^{\dagger}_{i\sigma} \hat{c}_{i\sigma}} \tag{9}$$

where $\hat{f}^{\dagger}_{i\sigma} \hat{f}_{i\sigma} = \hat{c}^{\dagger}_{i\sigma} \hat{c}_{i\sigma}$ follows from Eq. 8. Here, $\hat{s}^{z}_{i}$ and $\hat{s}^{x}_{i}$ are represented by Pauli matrices acting on the Ising spin at site i. Using the identity

$$(-1)^{\sum_\sigma \hat{f}^{\dagger}_{i\sigma} \hat{f}_{i\sigma}} = \prod_{\sigma} (2 \hat{n}_{i\sigma} - 1) \tag{10}$$

we can rewrite Eq. 7 as

$$\hat{H}^{fs} = -t \sum_{\langle ij \rangle, \sigma} \left(\hat{f}^{\dagger}_{i\sigma} \hat{f}_{j\sigma} \hat{s}^{z}_{i} \hat{s}^{z}_{j} + \text{H.c.} \right) - \frac{U}{4} \sum_{i} \hat{s}^{x}_{i} \tag{11}$$

Finally, one can explicitly check that $\left[\hat{H}^{fs}, \hat{Q}_i \right] = 0$ so that $\hat{H}^{fs}$ and $\hat{H}^{cQ}$ are dual. The above Hamiltonian is not amenable to simulations with the ALF library and a further transformation has to be carried out. Consider the bond variable

$$\hat{s}^{z}_{i} \hat{s}^{z}_{j} \mapsto \hat{Z}_{ij} \tag{12}$$

Because a spin flip on a single site i under the action of $\hat{s}^{x}_{i}$ affects all four bond variables, the dual representation involves a so-called star operator,

$$\hat{s}^{x}_{i} \mapsto \hat{X}_{i,i+x} \hat{X}^{x}_{i,i-x} \hat{X}^{x}_{i,i+y} \hat{X}^{x}_{i,i-y} \tag{13}$$

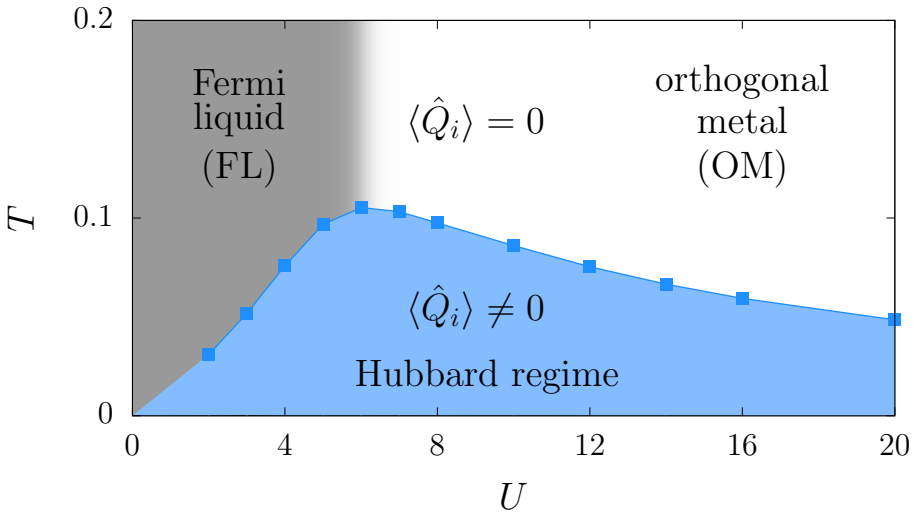


Figure 1. Phase diagram of the dual Hamiltonians (Eqs. 7, 11, and 14). A phase transition of the Ising variables $\hat{Q}_i$ at T_Q separates the low-temperature Hubbard regime from the high-temperature phase. Here, T_Q was determined from data for $L = 8$, see Ref. 22. The two metallic regimes at $T > T_Q$ appear to be separated by a crossover, as indicated by the colour gradient. Adapted from Ref. 9.

where $i \pm \alpha$ is a compact notation for the site at $\boldsymbol{r}_i \pm \hat{e}_\alpha$ and $\hat{X}$, $\hat{Z}$ corresponds to the x- and y-Pauli matrices acting on the bond variables. These steps lead to the Hamiltonian

$$\hat{H}^{fZ} = -t \sum_{\langle ij \rangle, \sigma} \left(\hat{f}^\dagger_{i\sigma} \hat{f}_{j\sigma} + \text{H.c.} \right) \hat{Z}_{ij} - \frac{U}{4} \sum_i \hat{Z}_{i,i+x} \hat{Z}_{i,i-x} \hat{Z}_{i,i+y} \hat{Z}_{i,i-y} \tag{14}$$

This Hamiltonian (Eq. 14) with bond Ising variables takes the form familiar of Ising lattice gauge theories coupled to matter. Here, there are important differences that lead to the unique properties of the phase diagram. Firstly, in Ising lattice gauge theories the Gauss law $\hat{Q}_i = \pm 1$ is imposed. Secondly, since $\hat{Z}_{i,i+y} \mapsto \hat{s}^z_i \hat{s}^z_j$ visons (a π-fluxes of the Z_2 field) are absent. In the above formulation, the model can be simulated with the ALF-library without encountering the negative sign problem.

The phase diagram of the model is plotted in Fig. 1 and contains three phases that one can understand very transparently when considering the slave spin formulation. Since $\hat{Q}_i$ commutes with the Hamiltonian, we expect this Ising variable to show a finite temperature transition akin to the two dimensional Ising universality class. Below the transition temperature $\hat{Q}_i$ orders ferromagnetically, such that at $T = 0$ we can replace $\hat{Q}_i$ by unity in Eq. 7 and recover the physics of the generic half-filled Hubbard model. At $U = 0$, the Ising field $\hat{Z}_{ij}$ has no quantum fluctuations, and with the vison-less constraint, $\hat{Z}_{i,i+a_x} \hat{Z}_{i+a_x,i+a_x+a_y} \hat{Z}_{i+a_x+a_y,i+a_y} \hat{Z}_{i+a_y,i} = 1$, one can show with a gauge transformation that any Ising bond configuration will be equivalent to the choice $\hat{Z}_{i,j} = 1$, or equivalently to the choice $\hat{s}^z_i = 1$. The model then reduces to a simple Fermi liquid. At finite values of U above Ising transition temperature, one cannot solve the model exactly,

but one can carry out educated guesses based on a mean-field analysis that breaks the local Z_2 symmetry inherent to the model.[23] At weak coupling, we can set $\langle \hat{s}_i^z \rangle = s > 0$ and $\langle \hat{s}_i^x \rangle = 0$ such that we recover the Fermi liquid state. On the other hand, at strong coupling, temporal fluctuations of the Ising field will be quick such that we expect the Ising correlations to decay exponentially in time and space, but the short ranged ones to remain finite. In a mean field picture we would hence set $\langle \hat{s}_i^z \rangle = 0$ but $\langle \hat{s}_i^z \hat{s}_{i+a_x}^z \rangle > 0$. Fractionalisation of the electron into an orthogonal fermion and a Z_2 field means that it can propagate coherently – in this mean-field vertex free picture – only if both constituents have a coherent propagation. Hence, in the disordered state of the Ising variable, the physical electron is localised and the single particle spectral shows no Fermi surface. Consider on the other hand a charge degree of freedom. The propagation of such an electron-hole pair does not depend on the coherent motion of the Z_2 field. Hence, charge fluctuations are gapless.

Our numerical calculations published in Refs. 9, 22 confirm the above mean field picture of the orthogonal metallic state in the Falicov-Kimball model. Seen from the perspective of particle-hole quantities the OM is a metallic (not superconducting) state. Seen from the perspective of the single particle, it is an insulator. Showing that this state of matter can emerge in a Falikov-Kimball model is one of the highlights of the work we recently carried out using the computational resources at the NIC.

4 Conclusions

The ALF implementation of the auxiliary field quantum Monte Carlo method provides us with a very flexible tool to simulate a variety of correlated electron systems. In the future, we will further develop this library and place emphasis on alternative sampling methods such as Hybrid Monte Carlo and Langevin dynamics as well as on time dependent Hamiltonians. The ALF-library and the associated git repository provide a documented and collaborative program development. It also allows to easily reproduce data since the output of each run includes the commit hash. As it stands, we have no efficient collaborative means to analyse the data of runs, and this is taken care of at the user level. Clearly each user stores data on tapes, but the access becomes very slow. Each run produces GBs of data in the forms of Monte Carlo bins such that a project cumulates to several TBs. Disc space is in principle not an issue but the challenge lies in defining the meta data as well as efficient tools that will allow us to navigate and search data bases and reanalyse, if required, the Monte Carlo time series. Such efforts will be pursued in the future. The pool of applications is very big. It ranges from toy models that capture interesting concepts such as fractionalisation, as described above, to realistic modelling of many body systems. In fact within the ALF-library we are able to consider long range Coulomb interactions,[21] a necessity for an accurate description of graphene. Furthermore, we are able to simulate very general spin systems embedded in a metallic environment.[24] This allows us to provide numerical support for a variety of magnetic nano-systems grown on metallic surfaces.[25, 26]

Acknowledgements

This work is a group effort. I would like to thank Martin Bercx, Florian Goth, Johannes Hofmann, and Jonas Schwab for the development of the ALF-Library and discussions, as

well as Stefan Beyl, Bimla Danu, Martin Hohenadler, Emilie Huffman, Yuhai Liu, Marcin Raczkowski, Toshihiro Sato, Francesco Parisen Toldin, Maksim Ulybyshev, and Zhenjiu Wang for many applications. We gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC). We thank the DFG for financial support through the FOR 1807, SFB 1170 as well as through the Würzburg-Dresden Cluster of Excellence on Complexity and Topology in Quantum Matter - ct.qmat (EXC 2147, project-id 39085490). Finally we thank KONWIRH (Competence Network for Technical, Scientific High Performance Computing in Bavaria) for providing seed funding that enabled us to start the ALF project and secure further support from the DFG Scientific Library Services and Information Systems (LIS) program under the grant number AS 120/14-1.

References

1. R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, *Monte Carlo calculations of coupled boson-fermion systems.*, Phys. Rev. D **24**, 2278–2286, 1981.
2. C. Wu and S.-C. Zhang, *Sufficient condition for absence of the sign problem in the fermionic quantum Monte Carlo algorithm*, Phys. Rev. B **71**, 155115, 2005.
3. E. Fulton Huffman and S. Chandrasekharan, *Solution to sign problems in half-filled spin-polarized electronic systems*, Phys. Rev. B **89**, 111101, 2014.
4. Z.-X. Li, Y.-F. Jiang, and H. Yao, *Majorana-Time-Reversal Symmetries: A Fundamental Principle for Sign-Problem-Free Quantum Monte Carlo Simulations*, Phys. Rev. Lett. **117**, 267002, 2016.
5. Z. C. Wei, C. Wu, Y. Li, S. Zhang, and T. Xiang, *Majorana Positivity and the Fermion Sign Problem of Quantum Monte Carlo Simulations*, Phys. Rev. Lett. **116**, 250601, 2016.
6. F. F. Assaad and T. Grover, *Simple Fermionic Model of Deconfined Phases and Phase Transitions*, Phys. Rev. X **6**, 041049, 2016.
7. T. Sato, M. Hohenadler, and F. F. Assaad, *Dirac Fermions with Competing Orders: Non-Landau Transition with Emergent Symmetry*, Phys. Rev. Lett. **119**, 197203, 2017.
8. S. Gazit, F. F. Assaad, S. Sachdev, A. Vishwanath, and C. Wang, *Confinement transition of Z2 gauge theories coupled to massless fermions: Emergent quantum chromodynamics and SO(5) symmetry*, Proceedings of the National Academy of Sciences **115(30)**, E6987-E6995, 2018.
9. M. Hohenadler and F. F. Assaad, *Fractionalized Metal in a Falicov-Kimball Model*, Phys. Rev. Lett. **121**, 086601, 2018.
10. Y. Liu, Z. Wang, T. Sato, M. Hohenadler, C. Wang, W. Guo, and F. F. Assaad, *Superconductivity from the condensation of topological defects in a quantum spin-Hall insulator*, Nature Communications **10**, 2658, 2019.
11. E. Witten, *A New Look At The Path Integral Of Quantum Mechanics*, 2010, arXiv:1009.6032 [hep-th].
12. E. Witten, *Analytic Continuation Of Chern-Simons Theory*, 2010, arXiv:1001.2933 [hep-th].

13. M. Ulybyshev, C. Winterowd, and S. Zafeiropoulos, *Taming the sign problem of the finite density Hubbard model via Lefschetz thimbles*, 2019, arXiv:1906.02726 [cond-mat.str-el].
14. M. Ulybyshev, C. Winterowd, and S. Zafeiropoulos, *Lefschetz thimbles decomposition for the Hubbard model on the hexagonal lattice*, 2019, arXiv:1906.07678 [cond-mat.str-el].
15. M. Bercx, J. S. Hofmann, F. F. Assaad, and T. C. Lang, *Spontaneous particle-hole symmetry breaking of correlated fermions on the Lieb lattice*, Phys. Rev. B **95**, 035108, 2017.
16. M. Bercx, F. Goth, J. S. Hofmann, and F. F. Assaad, *The ALF (Algorithms for Lattice Fermions) project release 1.0. Documentation for the auxiliary field quantum Monte Carlo code*, SciPost Phys. **3**, 013, 2017.
17. X. Y. Xu, Y. Qi, J. Liu, L. Fu, and Z. Y. Meng, *Self-learning quantum Monte Carlo method in interacting fermion systems*, Phys. Rev. B **96**, 041119, 2017.
18. K. S. D. Beach, *Identifying the maximum entropy method as a special limit of stochastic analytic continuation*, 2004, arXiv:cond-mat/0403055.
19. Y. Motome and M. Imada, *A Quantum Monte Carlo Method and Its Applications to Multi-Orbital Hubbard Models*, Journal of the Physical Society of Japan **66**, 1872–1875, 1997.
20. F. F. Assaad, M. Imada, and D. J. Scalapino, *Charge and spin structures of a $d_{x^2-y^2}$ superconductor in the proximity of an antiferromagnetic Mott insulator*, Phys. Rev. B **56**, 15001–15014, 1997.
21. H.-K. Tang, J. N. Leaw, J. N. B. Rodrigues, I. F. Herbut, P. Sengupta, F. F. Assaad, and S. Adam, *The role of electron-electron interactions in two-dimensional Dirac fermions*, Science **361**, 570–574, 2018.
22. M. Hohenadler and F. F. Assaad, *Orthogonal metal in the Hubbard model with liberated slave spins*, Phys. Rev. B **100**, 125133, 2019.
23. R. Nandkishore, M. A. Metlitski, and T. Senthil, *Orthogonal metals: The simplest non-Fermi liquids*, Phys. Rev. B **86**, 045128, 2012.
24. T. Sato, F. F. Assaad, and T. Grover, *Quantum Monte Carlo Simulation of Frustrated Kondo Lattice Models*, Phys. Rev. Lett. **120**, 107201, 2018.
25. M. Raczkowski and F. F. Assaad, *Emergent Coherent Lattice Behavior in Kondo Nanosystems*, Phys. Rev. Lett. **122**, 097203, 2019.
26. B. Danu, F. F. Assaad, and F. Mila, *Exploring the Kondo effect of an extended impurity with chains of Co adatoms in a magnetic field*, Phys. Rev. Lett. **123**, 176601, 2019.

Computational Soft Matter Science

Computational Soft Matter Science

Kurt Kremer

Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany
E-mail: kremer@mpip-mainz.mpg.de

Soft matter, in most cases collections of large organic molecules, whereby these molecules typically contain a huge number of intra-molecular degrees of freedom themselves, pose special challenges for experiment, theory and simulation. At the same time these materials are of universal importance, and this importance is even growing. Soft matter covers not only conventional plastics or polymer materials, which beyond simple packaging is an essential part of almost all technical products, medical devices *etc.* , it also is a key ingredient of novel functional materials. In addition, almost all biological materials can be characterised as soft matter, not to speak of our daily food. Thus, it is not surprising that soft matter simulations have a long tradition at NIC.

Simulating soft matter started with investigating simple polymer models, either walks on a lattice or bead spring chains in continuum. This went hand in hand with improved theoretical concepts as well as sophisticated experiments. As a result, basic concepts of macromolecular systems are rather well understood and one is at a stage to tackle more complicated and numerically more demanding questions. The latter being equally possible due to improved models and algorithms as well as due to improved hardware. Both contributions to soft matter simulations of this years volume are typical examples of this development.

Since the intrinsic dynamics of macromolecular systems is very slow, most materials never reach thermodynamic equilibrium during processing. Engineers or practitioners are used to such behaviour since ancient times, however we are nowadays getting to the point, where we can analyse such processes with molecular insight. Especially in polymer materials, the fact that the systems are not in equilibrium, but somewhere stuck or frozen makes up for many useful properties. Marcus Müller, Gaoyuan Wang, and Yongzhi Ren pick this up in their contribution *Do We Understand the Collective Short-Time Dynamics in Multicomponent Polymer Melts?*, where they present an initial study on the way to investigate dynamics of multicomponent polymer materials. These systems contain another specific challenge of modern materials, namely that the notion of a classical bulk material with interfaces, should be replaced by the notion of a bulk of interfaces. Thus, compositional and conformational fluctuations equally dominate the kinetics of structure and thus property development. Beginning with dynamical self consistent field theory (D-SCFT) they demonstrate the shortcomings this approach, which lacks a correct consideration of fluctuations. As a first test the decay of the collective structure factor of an ideal mixture of polymers is considered. Especially for short times D-SCFT fails to reproduce the simulation results. Unlike D-SCFT, simulations reveal a q-dependent decay rate. Switching to model B, originating from the theory of critical phenomena this is the variant with conserved density, significantly improves the description. This improvement can be taken as a starting point to analyse non ideal polymer mixtures.

The second contribution *Molecular Dynamics Simulations of Curcumin in the Interface Region of Triblock Copolymer Micelles* by Konstantin P. K. Luft and Stephan Gekle comes from the opposite side, namely atomistically detailed models. Beyond general generic concepts the variation of the local chemistry provides the second way of manipulating system properties. Both ways allow for a huge variation. While this offers new opportunities on the synthetic side, when it comes to biology related questions, one usually is dealing with pre-defined specific systems. The current simulations tackle a problem related to drug delivery, namely that many modern drugs are often not at all or only poorly water soluble, which poses special difficulties when it comes to drug formulation. To overcome this problem polymers, which are amphiphilic can be interesting candidates to package hydrophobic drug molecules and then enable the delivery in the watery environment of our bodies. The example investigated is a single curcumin molecule in water and in a polymeric ABA copolymer matrix. In simple terms curcumin consists of a short eight bond chain with two functionalised benzene rings at the ends. The current study focuses on the distance between these rings and their relative orientation of the two different curcumin tautomers in these different environments. It turns out that their distance in water is rather fixed and so is their parallel orientation. The molecule bends and the rings stack. In the polymer environment the situation is completely different. The amphiphilic nature of the chains allows for significant conformational freedom of the curcumin with significant weight on rather stretched conformations. This distance leads to a de-correlation of the orientation of the benzene rings, demonstrating the role of the local environment and giving hints for further development.

Though very different these two contributions reveal the power and the need of high level computer simulations for soft matter problems. Over the years many research projects covering the whole range of details of description benefited from support by NIC.

Do We Understand the Collective Short-Time Dynamics in Multicomponent Polymer Melts?

Marcus Müller[1], **Gaoyuan Wang**[1], **and Yongzhi Ren**[1,2]

[1] Institute for Theoretical Physics, Georg-August University, 37077 Göttingen, Germany
E-mail: mmueller@theorie.physik.uni-goettingen.de

[2] Key Lab of In-fiber Integrated Optics, Harbin Engineering University, Harbin, China

Phase separation in multicomponent polymer melts is a ubiquitous process in polymer engineering and consequently has also attracted abiding interest from simulation and theory. Whereas the equilibrium thermodynamics of macrophase separation and microphase separation in homopolymer blends and copolymers can be rather quantitatively described by Self-Consistent Field Theory (SCFT) or extensions that capture fluctuations, the kinetics of structure evolution poses challenges for a theoretical description. Examining simple, prototypical examples, we highlight the role of internal modes and indicate how Dynamic Self-Consistent Field Theory (D-SCFT) can be generalised to include the consequences of the subdiffusive single-chain dynamics for the collective kinetics on times comparable to the Rouse-relaxation time.

1 Introduction

Structure formation – macrophase separation or self-assembly into spatially modulated phases – in multicomponent polymer melts is ubiquitously employed in technical applications such as plastics toys made of rubber-toughened/high-impact polystyrene or soles of shoes and tire treads using poly(styrene-butadiene-styrene) (SBS) triblock copolymers. Blending two different polymers, one fabricates a composite with improved materials characteristics. By virtue of chain connectivity, long macromolecules have a small entropy of mixing, and a minuscule repulsion between the monomeric repeating units of different polymer species gives rise to domain formation. In these spatial regions, monomeric repeating units of one species are enriched and they are separated by interfaces from the domains of the other monomer species. The phase behaviour in equilibrium, the properties of interfaces between domains, and the self-assembly of block copolymers, where the two monomer species are covalently linked together into a macromolecule and macroscopic phase separation is avoided, are rather well understood. Since the long macromolecules in a melt strongly interdigitate, one molecule interacts with many neighbours and the equilibrium behaviour can be accurately described by SCFT.

In practical applications, however, thermodynamic equilibrium often is not achieved. For instance, in rubber-toughened polystyrene the hard, brittle polystyrene and the soft, deformable rubber do not macroscopically phase separate but rather form an assembly of interfaces with a characteristic domain size of micrometers. It is exactly this nonequilibrium morphology that gives rise to the improved impact resistance. Likewise, copolymers almost never form an ideal crystalline arrangement of domains. Instead, without external guiding fields, the morphology is riddled with defects and the perfecting of order is protracted.

Therefore it is important to understand the kinetics of structure formation in order to predict and tailor the structure evolution. The accurate knowledge of the equilibrium ther-

modynamics and the free energy of a phase-separated morphology is an excellent starting point. In the following, we consider an incompressible, two-component polymer melt, where $\phi_A(\mathbf{r},t)$ and $\phi_B(\mathbf{r},t) = 1 - \phi_A(\mathbf{r},t)$ denote the normalised number densities of monomer species, A and B. SCFT provides for each morphology, ϕ_A, an accurate value of the free energy, $\mathcal{F}[\phi_A]$, and of the exchange chemical potential, μ, assuming that the molecular conformations are in equilibrium with the instantaneous value of the density, $\phi_A(\mathbf{r},t)$. In a nonequilibrium situation, the chemical potential varies in space and its gradient gives rise to a current. Since the density is locally conserved, we obtain a model-B time evolution[1–3]

$$\frac{\partial \phi_{\mathbf{q},A}(t)}{\partial t} = -\mathbf{q}^2 \Lambda_{\mathbf{q}} \frac{\mu_{\mathbf{q}}[\phi_A]}{k_{\mathrm{B}}T} \tag{1}$$

where $\Lambda_{\mathbf{q}}$ denotes the Onsager coefficient that relates the gradient of the chemical potential to a current of the density. The wavevector-dependence of $\Lambda_{\mathbf{q}}$ signals the spatial nonlocality of the relation between chemical potential and current that stems from the chain connectivity. This kinetic equation forms the basis of D-SCFT.[4–12]

Consider an ideal mixture of two structurally and dynamically symmetric homopolymers with no repulsion between the two monomer species, $\chi_{\mathrm{FH}}N = 0$. Here, χ_{FH} and N denote the Flory-Huggins parameter and the chain length, respectively. Initially, at time $t = 0$, there is a small-amplitude, sinusoidal density modulation with wavevector $\mathbf{q}$. Within linear-response approximation, the chemical potential is linearly related to the amplitude of the density modulation, $\frac{\mu_{\mathbf{q}}}{k_{\mathrm{B}}T} = \frac{N}{S_{\mathbf{q}}}\phi_{\mathbf{q},A}$, where $S_{\mathbf{q}}$ denotes the static, collective structure factor. Inserting this expression into Eq. 1, we simply obtain

$$\phi_{\mathbf{q},A}(t) = \phi_{\mathbf{q},A}(0) \exp\left(-\mathbf{q}^2 \Lambda_{\mathbf{q}} \frac{N}{S_{\mathbf{q}}} t\right) \tag{2}$$

i. e. the density modulation exponentially decays with a wavevector-dependent relaxation time. By the same token, Eq. 1 also predicts that the collective, dynamic structure factor, $S_{\mathbf{q}}(t) \sim \langle \phi_{\mathbf{q},A}(t)\phi_{\mathbf{q},A}(0)\rangle$ decays exponentially in time at fixed $\mathbf{q}$.

On the other hand, since the mixture is ideal, the collective dynamic structure factor, $S_{\mathbf{q}}(t)$, is proportional to the dynamic form factor, $S^{(0)}_{\mathbf{q}}(t)$, of a single chain. Indeed, $S^{(0)}_{\mathbf{q}}(t)$ decays exponentially in time for $(\mathbf{q}R_{\mathrm{e}})^2 \to 0$ (with R_{e} being the root mean-squared end-to-end distance). In the opposite limit, $(\mathbf{q}R_{\mathrm{e}})^2 \gg 1$ and times, t, smaller than the Rouse time, τ_{R}, however, the subdiffusive behaviour of the monomer mean-squared displacement, $g_1(t)$, alters the decay. Qualitatively, one obtains[2]

$$\ln \frac{\phi_{\mathbf{q},A}(t)}{\phi_{\mathbf{q},A}(0)} = \ln \frac{S^{(0)}_{\mathbf{q}}(t)}{S^{(0)}_{\mathbf{q}}} \sim \mathbf{q}^2 g_1(t) \sim (\mathbf{q}R_{\mathrm{e}})^2 \sqrt{\frac{t}{\tau_{\mathrm{R}}}} \qquad \text{for} \quad t \ll \tau_{\mathrm{R}} \tag{3}$$

In Fig. 1 we present our simulation data using a dense melt, $\sqrt{\mathcal{N}} = 51\,200$, of finely discretised Gaussian chains, $N = 256$, within our soft, coarse-grained model.[14, 15] The simulations were performed with our parallel, GPU-accelerated program, SOMA.[16] Due to the softness of the potentials, macromolecular contours can cross each other in the course of their motion, and the single-chain dynamics is quantitatively described by the Rouse model.[17] Unlike Eq. 1, the simulations include thermal fluctuations but their effect is rather small because of the large invariant degree of polymerisation, $\bar{\mathcal{N}}$. Also the simulations do

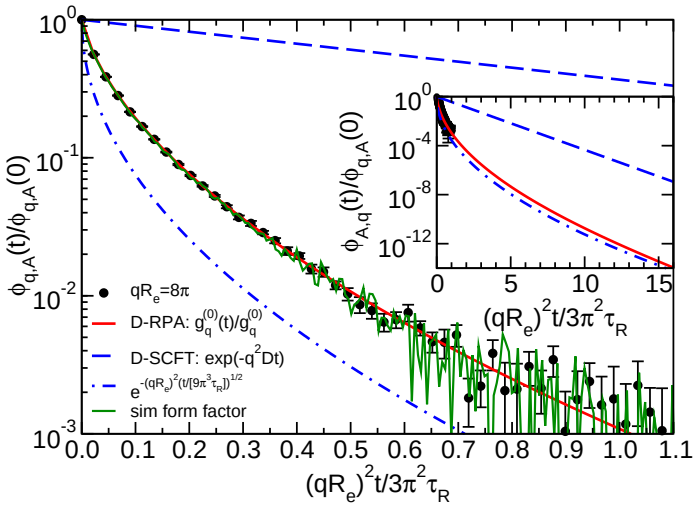


Figure 1. Relaxation of a density modulation with amplitude $\phi_{\mathbf{q},A}(t)$ at $qR_e = 8\pi$ in a symmetric homopolymer blend at $\chi_{FH}N = 0$. The decay of the amplitude in the simulation, (black circles), is compared with the dynamic, single-chain form factor, $S^{(0)}_{\mathbf{q}}(t)$ (red line – Rouse prediction, green line – simulation), and its limiting behaviours for long and short times (dashed and dashed-dotted blue lines, respectively). The inset illustrates the decay of the amplitude on longer time scales. Adapted from Ref. 13

not invoke the Random-Phase Approximation (RPA) approximation of a linear relation between density modulation and chemical potential.

The decay of the amplitude agrees quantitatively with the decay of the dynamic, single-chain form factor, and the semilogarithmic representation of the data highlights the nonexponential decay in time.

Thus, D-SCFT fails to account for the qualitative features of the collective kinetics of structure formation on short time scales. These scales, however, are important for a variety of processes and computational techniques:

- After a quench of a multicomponent melt from the homogeneous state deep into a (micro)phase-separated phase, initially, density fluctuations are strongly amplified, and the morphology obtained after this spinodal phase-separation dictates the subsequent ordering kinetics. The importance of the early-time dynamics holds, *a fortiori*, in the case that the structure formation is directed by external guiding fields, *e. g.* surface-directed spinodal decomposition or Directed Self Assembly (DSA).

- Process-Directed Self Assembly (PDSA) of copolymers tailors changes of thermodynamic control parameters, *e. g.* pressure or temperature, to reproducibly fabricate well-defined unstable states. The thermodynamic process is constructed such that this so-constructed unstable state relaxes into a new, metastable state with desired properties, *e. g.* a bicontinuous network morphology.[18, 19] This relaxation typically requires

changes of density on the scale of the unit cell of the spatially modulated phase. Since the periodicity is on the order R_e, the concomitant time scale is on the order of the Rouse time, τ_R.

- Advancements in parallel computing enable the investigation of ever larger systems, however, relevant problems of larger systems often also entail longer time scales. Thus, computational techniques to extend the time scale of particle-based simulations have attracted abiding interest. For instance, the Heterogeneous Multiscale Method (HMM)[20–22] aims at predicting the time evolution by parameterising a dynamic continuum model on-the-fly, using *e. g.* a description similar to D-SCFT. In order to project the short-time kinetics into the future, the dynamic continuum model must be able to accurately describe both, the short-time and the long-time collective kinetics.

In the following, we outline a generalisation of Eq. 1 that accounts for the subdiffusive, single-chain dynamics by an Onsager coefficient that not only depends on the wavevector but, additionally, on the frequency, ω. This is equivalent to introducing memory effects into Eq. 1, which are qualitatively expected when coarse-graining a microscopic particle-based model into a continuum description that only retains the slow, locally conserved density fields.

2 Generalised Model-B Dynamics

The strong bonded forces do not only give rise to a wavevector-dependence of $\Lambda_\mathbf{q}$ in Eq. 1 but, additionally, result in a subdiffusive single-chain dynamics according to the Rouse model.[17] In order to account for the single-chain dynamics in the collective kinetics, Semenov[23] proposed the following generalisation of Eq. 1

$$\frac{\partial \phi_{\mathbf{q},A}(t)}{\partial t} = -\mathbf{q}^2 \int_{-\infty}^{t} \mathrm{d}\tau \; \Lambda_\mathbf{q}(t-\tau) \frac{\mu_\mathbf{q}[\phi_{\mathbf{q},A}(\tau)]}{k_\mathrm{B}T} \tag{4}$$

with a time-dependent Onsager coefficient, $\Lambda_\mathbf{q}(t)$. Recently, using the Dynamic Random-Phase Approximation (D-RPA),[2, 23–25] we have derived an explicit expression for the Fourier-Laplace transform, $\tilde{\Lambda}_\mathbf{q}(\omega)$ of this generalised Onsager coefficient,[13]

$$\frac{i\omega}{\mathbf{q}^2 N \tilde{\Lambda}_\mathbf{q}(\omega)} = \frac{\tilde{\chi}_{\mathbf{q},AA} + 2\tilde{\chi}_{\mathbf{q},AB} + \tilde{\chi}_{\mathbf{q},BB}}{\tilde{\chi}_{\mathbf{q},AA}\tilde{\chi}_{\mathbf{q},BB} - \left(\tilde{\chi}_{\mathbf{q},AB}\right)^2} - \frac{1}{S_\mathbf{q}} \tag{5}$$

in terms of the dynamic susceptibilities

$$\tilde{\chi}_{\mathbf{q},\alpha\beta}(\omega) \equiv -\int_0^\infty \mathrm{d}t \; \frac{\partial S^{(0)}_{\mathbf{q},\alpha\beta}(t)}{\partial t} e^{-i\omega t} = S^{(0)}_{\mathbf{q},\alpha\beta}(0) - i\omega \tilde{S}^{(0)}_{\mathbf{q},\alpha\beta}(\omega) \tag{6}$$

with $S^{(0)}_{\mathbf{q},\alpha\beta}(t)$ being the concentration-weighted, dynamic, single-chain form factors of monomer species, α and β. $S_\mathbf{q}$ is the static, collective structure factor at $\chi_\mathrm{FH}N = 0$.

Notably, this Onsager coefficient does not depend on the strength of the pairwise interaction, $\chi_\mathrm{FH}N$. Thus, Eq. 4 conserves the separation between the thermodynamic force, $\nabla\mu$, and the kinetic coefficient that relates this thermodynamic force to a current.

To illustrate the consequences of the time-dependence of the Onsager coefficient, we consider an ideal, symmetric homopolymer blend with $\chi_{\rm FH}N = 0$. Eq. 5 simplifies to

$$\tilde{\Lambda}_{\mathbf{q}}(\omega) = \bar{\phi}_A(1-\bar{\phi}_A)\frac{g_{\mathbf{q},1}}{\mathbf{q}^2}\left[\frac{g_{\mathbf{q},1}}{\tilde{g}_{\mathbf{q},1}(\omega)} - i\omega\right] = \frac{S_{\mathbf{q}}}{N\mathbf{q}^2\tau_{\rm R}}\left[\frac{g_{\mathbf{q},1}\tau_{\rm R}}{\tilde{g}_{\mathbf{q},1}(\omega)} - i\omega\tau_{\rm R}\right] \qquad (7)$$

where $g_{\mathbf{q},f}$ and $\tilde{g}_{\mathbf{q},f}(\omega)$ denote the Debye function that characterises the static, single-chain form factor of the fraction, f, of a Gaussian chain and the Laplace transform of the dynamic, single-chain form factor, respectively. $S_{\mathbf{q}} = \bar{\phi}_A(1-\bar{\phi}_A)Ng_{\mathbf{q},1}$ is the collective structure factor of a symmetric blend at $\chi_{\rm FH}N = 0$.

Inserting this frequency-dependent Onsager coefficient in the Laplace transform of Eq. 4,

$$i\omega\tilde{\phi}_{\mathbf{q},A}(\omega) - \phi_{\mathbf{q},A}(0) = -\mathbf{q}^2\tilde{\Lambda}_{\mathbf{q}}\frac{\tilde{\mu}_{\mathbf{q}}}{k_{\rm B}T} \qquad (8)$$

where $\phi_{\mathbf{q},A}(0)$ denotes the starting morphology at $t = 0$, we obtain

$$\phi_{\mathbf{q},A}(0) = \frac{g_{\mathbf{q},1}}{\tilde{g}_{\mathbf{q},1}(\omega)}\tilde{\phi}_{\mathbf{q},A} \qquad \text{or} \qquad \phi_{\mathbf{q},A}(t) = \phi_{\mathbf{q},A}(0)\frac{g_{\mathbf{q},1}(t)}{g_{\mathbf{q},1}} = \phi_{\mathbf{q},A}(0)\frac{S^{(0)}_{\mathbf{q}}(t)}{S^{(0)}_{\mathbf{q}}} \qquad (9)$$

in accord with Eq. 3 and the simulation data in Fig. 1.

In the case that the generalised Onsager coefficient, $\tilde{\Lambda}_{\mathbf{q}}(\omega)$, does not depend on ω, we recover Eq. 1. To this end, the following sequence of approximations is often employed:

1. We define $\varphi_{\mathbf{q},\alpha\beta}(t) \equiv S^{(0)}_{\mathbf{q},\alpha\beta}(t)/S^{(0)}_{\mathbf{q},\alpha\beta}(0)$ in order to separate the equilibrium thermodynamic properties, characterised by the static, single-chain form factor, $S^{(0)}_{\mathbf{q},\alpha\beta}(0)$, from the dynamics. $\varphi_{\mathbf{q},\alpha\beta}(0) = 1$ and $\varphi_{\mathbf{q},\alpha\beta}(t) \to 0$ for $t \to \infty$.

2. Ignoring the details of the temporal relaxation, we can crudely approximate

$$\varphi_{\mathbf{q},\alpha\beta}(t) \approx e^{-t/\tau_{\mathbf{q},\alpha\beta}} \quad \text{and} \quad \tilde{\chi}_{\mathbf{q},\alpha\beta} \approx \frac{S^{(0)}_{\mathbf{q},\alpha\beta}}{1+i\omega\tau_{\mathbf{q},\alpha\beta}} \qquad (10)$$

   by an exponential decay with a wavevector-dependent timescale, $\tau_{\mathbf{q},\alpha\beta}$. Such an approximation is appropriate for symmetric homopolymer blends on large length scales, but it fails for copolymers (as shown below), and it is inappropriate for short length scales, *c.f.* Fig. 1.

3. Assuming that the relaxation times of the different dynamic, single-chain form factors were identical, $\tau_{\mathbf{q},\alpha\beta} \approx \tau_{\mathbf{q}}$, which is *e. g.* appropriate for symmetric homopolymer blends or the scale-free Porod limit, $(\mathbf{q}R_{\rm e})^2 \gg 1$ and $t \ll \tau_{\rm R}$, in symmetric copolymers, we obtain a frequency-independent Onsager coefficient

$$\tilde{\Lambda}_{\mathbf{q}}(\omega) \approx \frac{S_{\mathbf{q}}}{N\mathbf{q}^2\tau_{\mathbf{q}}} \qquad (11)$$

   where $S_{\mathbf{q}}$ is the collective structure factor of the ideal mixture, $\chi_{\rm FH}N = 0$. This gives rise to an Onsager coefficient, $\Lambda_{\mathbf{q}}(t) \approx \frac{S_{\mathbf{q}}}{N\mathbf{q}^2\tau_{\mathbf{q}}}\delta(t)$, that is local in time.

4. Assuming a diffusive relaxation, $\tau_{\mathbf{q}} \approx \frac{3\pi^2}{(\mathbf{q}R_{\mathrm{e}})^2}\tau_{\mathrm{R}} = \frac{1}{\mathbf{q}^2 D}$ with $D = \frac{R_{\mathrm{e}}^2}{3\pi^2\tau_{\mathrm{R}}}$ within the Rouse model being the single-chain self-diffusion coefficient, one arrives at the Onsager coefficient of D-SCFT,

$$\Lambda_{\mathbf{q}}^{\mathrm{D-SCFT}} \approx \frac{S_{\mathbf{q}}}{N} D \tag{12}$$

Specifically, for an ideal mixture – homopolymer blend or diblock copolymer melt – with $\chi_{\mathrm{FH}}N = 0$, small-amplitude density modulations give rise to $\frac{\mu_{\mathbf{q}}}{k_{\mathrm{B}}T} = \frac{N}{S_{\mathbf{q}}}\phi_{\mathbf{q},A}$. Thus Eqs. 1 and 12 yield the simple, general D-SCFT prediction

$$\phi_{\mathbf{q},A}(t) = \phi_{\mathbf{q},A}(0)\exp(-\mathbf{q}^2 D t) \qquad \text{for} \quad \chi_{\mathrm{FH}}N = 0 \tag{13}$$

i. e. density modulations decay *via* the diffusion of the macromolecules' centre of mass. This prediction agrees with the decay of the dynamic, single-chain form factor on large scales but fails to describe the simulation data for $(\mathbf{q}R_{\mathrm{e}})^2 \gg 1$, as demonstrated in Fig. 1.

Using extensive computer simulations on JUWELS through the NIC, we can validate each of these approximations, illustrate their consequences, and quantify to what extent the generalised model-B dynamics, Eq. 4, improves the description of structure evolution.

3 Validation: Decay of a Large-Scale Density Modulation in a Symmetric Copolymer Melt

For a symmetric copolymer melt at $\chi_{\mathrm{FH}}N = 0$, the Onsager coefficient of D-SCFT takes the simple form $\Lambda_{\mathbf{q}}^{\mathrm{D-SCFT}} = \left(g_{\mathbf{q},1/2} - \frac{1}{4}g_{\mathbf{q},1}\right) D$, and D-SCFT predicts that density modulations, $\phi_{\mathbf{q},A}(t)$, exponentially decay in time with the relaxation time $\tau_{\mathbf{q}} = \frac{1}{\mathbf{q}^2 D}$, *i. e.* the larger the scale of the density modulation the slower is its temporal decay.

In Fig. 2 we present the simulation data for the decay of a density modulation in a melt of symmetric diblock copolymers with $\chi_{\mathrm{FH}}N = 0$. Two wavevectors, $qR_{\mathrm{e}} = 2\pi/3$ and $2\pi/6$, are considered. The data are compared to the predictions of D-SCFT and two marked discrepancies are observed: (i) The decay of the density modulation in the particle-based simulation is much faster than predicted by D-SCFT, and (ii) the relaxation time in the simulation is largely independent from the wavevector.

We again expect that the density modulation decays like the dynamic, collective structure factor, $S_{\mathbf{q}}(t)$, that can be expressed in terms of dynamic, single-chain form factors by D-RPA. For $\chi_{\mathrm{FH}}N = 0$ we obtain

$$\frac{\phi_{\mathbf{q},A}(t)}{\phi_{\mathbf{q},A}(0)} = \frac{S_{\mathbf{q}}(t)}{S_{\mathbf{q}}} \overset{\text{D-RPA}}{\approx} \frac{g_{\mathbf{q},1/2}(t) - \frac{1}{4}g_{\mathbf{q},1}(t)}{g_{\mathbf{q},1/2} - \frac{1}{4}g_{\mathbf{q},1}} \tag{14}$$

To make progress, we note that, on large length scales, $(\mathbf{q}R_{\mathrm{e}})^2 \ll 1$, the dynamic form factor, $g_{\mathbf{q},f}(t)$, of a fraction, f, of a Rouse chain takes the form

$$g_{\mathbf{q},f}(t) \approx g_{\mathbf{q},f} e^{-\mathbf{q}^2 D t - (\mathbf{q}R_{\mathrm{e}})^2 h_f(t)} \tag{15}$$

where the function $h_f(t)$ characterises the deviation from the diffusive relaxation due to internal modes for $t \ll \tau_{\mathrm{R}}$. $h_f(t)$ monotonically increases from $h_f(0) = 0$ to a plateau value,

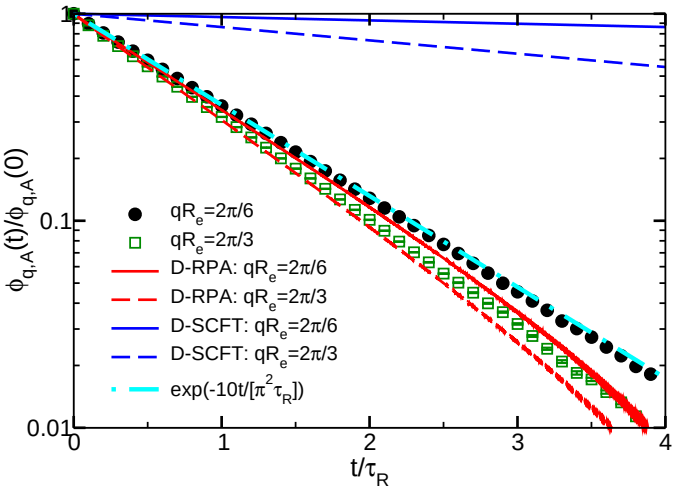


Figure 2. Relaxation of large-wavelength density fluctuation with $qR_{\rm e} = 2\pi/6$ and $2\pi/3$ in a disordered symmetric copolymer melt, $\chi_{\rm FH} = 0$. The simulation data (circles, squares) are compared to the prediction of D-RPA (red lines), D-SCFT (blue lines), and the approximation of the generalised model B (cyan line) according to Eq. 19 for $(\mathbf{q}R_{\rm e})^2 \to 0$ and $\omega\tau_{\rm R} \to 0$. Adapted from Ref. 13.

$h_f(\infty)$, for $t \gg \tau_{\rm R}$. Note that this function identically vanishes, when we consider an entire Rouse chain, *i. e.* $h_1(t) \equiv 0$, and $h_{1/2}(t)$ tends to its plateau value, $h_{1/2}(\infty) = 1/36$, on the time scale $\tau_{\rm R}$. Using Eq. 15 in Eq. 14, we obtain[13]

$$\frac{\phi_{\mathbf{q},A}(t)}{\phi_{\mathbf{q},A}(0)} \approx \left[1 - \frac{h_{1/2}(t)}{h_{1/2}(\infty)}\right] e^{-\mathbf{q}^2 Dt} + \mathcal{O}\left((\mathbf{q}R_{\rm e})^2\right) \tag{16}$$

Thus, the decay of the large-scale density modulation arises from the first factor, which approaches zero on the time scale, $\tau_{\rm R}$, that is independent from the wavevector, $\mathbf{q}$ – in agreement with the simulation data in Fig. 2.

In the case of symmetric copolymer melts, Eq. 5, adopts the simple form

$$\tilde{\Lambda}_{\mathbf{q}}(\omega) = \left(g_{\mathbf{q},1/2} - \frac{1}{4}g_{\mathbf{q},1}\right) \frac{1}{\mathbf{q}^2\tau_{\rm R}} \left[\frac{\left(g_{\mathbf{q},1/2} - \frac{1}{4}g_{\mathbf{q},1}\right)\tau_{\rm R}}{\tilde{g}_{\mathbf{q},1/2}(\omega) - \frac{1}{4}\tilde{g}_{\mathbf{q},1}(\omega)} - i\omega\tau_{\rm R}\right] \tag{17}$$

On large length scales, $g_{\mathbf{q},1/2} - \frac{1}{4}g_{\mathbf{q},1} = \frac{(\mathbf{q}R_{\rm e})^2}{144} + \mathcal{O}\left((\mathbf{q}R_{\rm e})^4\right)$ and, in the same limit, we use the Rouse model to compute

$$\tilde{g}_{\mathbf{q},1/2}(\omega) - \frac{1}{4}\tilde{g}_{\mathbf{q},1}(\omega) \approx \frac{2(\mathbf{q}R_{\rm e})^2}{3\pi^4} \sum_{p\ \rm odd} \frac{1}{p^4(\mathbf{q}^2 D + \frac{p^2}{\tau_{\rm R}} + i\omega)} \tag{18}$$

and obtain the frequency-dependent Onsager coefficient

$$\tilde{\Lambda}_{\mathbf{q}}(\omega) \approx \frac{\pi^6 D}{4608}\left[\frac{1}{\sum_{p\ \mathrm{odd}}\frac{1}{p^4(\mathbf{q}^2 D\tau_{\mathrm{R}}+p^2+i\omega\tau_{\mathrm{R}})}} - i\omega\tau_{\mathrm{R}}\right] \tag{19}$$

In marked contrast to $\Lambda_{\mathbf{q}}^{\mathrm{D-SCFT}}$, this generalised Onsager coefficient remains finite for $(\mathbf{q}R)^2 \to 0$, indicating that the relaxation stems from internal modes.[24, 25] Specifically, for $(\mathbf{q}R_{\mathrm{e}})^2 \to 0$ and $\omega\tau_{\mathrm{R}} \to 0$, we find $\tilde{\Lambda}_{\mathbf{q}}(\omega) \to 5D/24$, corresponding to a time-independent Onsager coefficient, $\Lambda_{\mathbf{q}} \to 5D/24$, that qualitatively differs from the D-SCFT prediction. This approximate Onsager coefficient yields a single-exponential decay, $\phi_{\mathbf{q},A}(t) = \phi_{\mathbf{q},A}(0)\exp(-10t/[\pi^2\tau_{\mathrm{R}}])$ that is in excellent agreement with the simulation data, see Fig. 2.

4 Concluding Remarks

Using computer simulations of a soft, coarse-grained model[14, 15] in conjunction with the parallel, GPU-accelerated program SOMA,[16] we have investigated the short-time dynamics of multicomponent polymer melts.[13] Even for the most basic kinetic process – the decay of a density modulation in an ideally miscible system, $\chi_{\mathrm{FH}}N = 0$ – we have found significant deviations from D-SCFT.[4–12] In the case of a miscible homopolymer blend, we observed that the amplitude of a short-scale density modulation does not exponentially decay in time and is much faster than predicted by D-SCFT. We trace this failure of D-SCFT back to the subdiffusive mean-squared monomer displacement for times smaller than τ_{R}. Likewise, for the temporal decay of a large-scale density modulation in symmetric diblock copolymers, we observe that the relaxation time scale is independent from the wavevector in the limit $(\mathbf{q}R_{\mathrm{e}})^2 \to 0$, whereas D-SCFT predicts $\tau_{\mathbf{q}} = 1/(\mathbf{q}^2 D)$.

We have demonstrated that these deviations can be accounted for by a generalised model B with a wavevector- and frequency-dependent Onsager coefficient.[23] Using D-RPA, we have obtained an explicit expression for the Onsager coefficient in terms of the dynamic, single-chain form factors.[13] Thereby a direct connection between the nontrivial single-chain dynamics and the generalised Onsager coefficient is established.

Combining this generalised Onsager coefficient with the accurate chemical potential of SCFT we expect that our approach can also be extended to the nonlinear regime, where D-RPA breaks down but SCFT still provides an accurate estimate for the chemical potential.

Our present validation of the generalised model B has been limited to all but the simplest, prototypical cases. For practical applications, multicomponent systems with a repulsion between different monomer species, $\chi_{\mathrm{FH}}N > 0$, are important as well as systems with structural and dynamic asymmetries. Moreover, our studies have been restricted to unentangled polymers because our soft, coarse-grained particle-based model does not enforce noncrossability of molecular contours and analytic predictions for the dynamic, single-chain form factors are available within the Rouse model.[17] Additionally, it is interesting to investigate more complex chain architectures, such as *e. g.* multiblock copolymers, star or branched polymers, and multicomponent polymer networks. In these examples the regime of subdiffusive dynamics is significantly more extended in time.

Given that the phase separation and self-assembly of complex multicomponent molecules have attracted abiding interest and that experimentally observed structures do

not always correspond to equilibrium phases but, instead, are dictated by the kinetics of structure formation, we expect that the outlined techniques will find ample applications to explore the kinetics of structure formation in these fascinating materials.

Acknowledgements

We thank Friederike Schmid, Jörg Rottler, and Matthias Krüger for stimulating discussions. Financial support by the Deutsche Forschungsgemeinschaft (DFG) within the CRC 937/TP07 is gratefully acknowledged. We thank the NIC for access to the supercomputer JUWELS at JSC.

References

1. P. C. Hohenberg and B. I. Halperin, *Theory of dynamic critical phenomena*, Rev. Mod. Phys. **49**, 435–479, 1977.
2. F. Brochard and P. G. de Gennes, *Dynamics of compatible polymer mixtures*, Physica A **118**, 289–299, 1983.
3. K. Binder, *Collective Diffusion, Nucleation, and Spinodal Decomposition in Polymer Mixtures*, J. Chem. Phys. **79**, 6387–6409, 1983.
4. J. G. E. M. Fraaije, B. A. C. van Vlimmeren, N. M. Maurits, M. Postma, O. A. Evers, C. Hoffmann, P. Altevogt, and G. Goldbeckwood, *The Dynamic Mean-Field Density Functional Method and its Application to the Mesoscopic Dynamics of Quenched Block Copolymer Melts*, J. Chem. Phys. **106**, 4260–4269, 1997.
5. T. Kawakatsu, M. Doi, and R. Hasegawa, *Dynamic density functional approach to phase separation dynamics of polymer systems*, Int. J. Mod. Phys. C **10**, 1531–1540, 1999.
6. G. J. A. Sevink, A. V. Zvelindovsky, B. A. C. van Vlimmeren, N. M. Maurits, and J. G. E. M. Fraaije, *Dynamics of surface directed mesophase formation in block copolymer melts*, J. Chem. Phys. **110**, 2250–2256, 1999.
7. E. Reister, M. Müller, and K. Binder, *Spinodal Decomposition in a Binary Polymer Mixture: Dynamic Self-Consistent-Field Theory and Monte Carlo Simulations*, Phys. Rev. E **64**, 041804, 2001.
8. H. Morita, T. Kawakatsu, and M. Doi, *Dynamic Density Functional Study on the Structure of Thin Polymer Blend Films with a Free Surface*, Macromolecules **34**, 8777–8783, 2001.
9. J. G. E. M. Fraaije, A. V. Zvelindovsky, and G. J. A. Sevink, *Computational Soft Nanotechnology with Mesodyn*, Mol. Sim. **30**, 225–238, 2004.
10. G. H. Fredrickson, *Dynamics and Rheology of Inhomogeneous Polymeric Fluids: a Complex Langevin Approach*, J. Chem. Phys. **117**, 6810–6820, 2002.
11. M. Müller and F. Schmid, *Incorporating fluctuations and dynamics in self-consistent field theories for polymer blends*, Adv. Polym. Sci **185**, 1–58, 2005.
12. S. Qi and F. Schmid, *Dynamic Density Functional Theories for Inhomogeneous Polymer Systems Compared to Brownian Dynamics Simulations*, Macromolecules **50**, 9831–9845, 2017.
13. G. Wang, Y. Ren, and M Müller, *Collective short-time dynamics in multicomponent polymer melts*, Macromolecules **52**, 7704–7720, 2019.

14. K. Ch. 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.
15. M. Müller, *Studying amphiphilic self-assembly with soft coarse-grained models*, J. Stat. Phys. **145**, 967, 2011.
16. L. Schneider and M. Müller, *Multi-architecture Monte-Carlo (MC) simulation of soft coarse-grained polymeric materials: SOft coarse grained Monte-Carlo Acceleration (SOMA)*, Comp. Phys. Comm. **235**, 463–476, 2019.
17. P. E. Rouse, *A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers*, J. Chem. Phys. **21**, 1272–1280, 1953.
18. 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.
19. D.-W. Sun and M. Müller, *Process-Accessible States of Block Copolymers*, Phys. Rev. Lett. **118**, 067801, 2017.
20. E. Vanden-Eijnden, *On HMM-Like Integrators and Projective Integration Methods for Systems with Multiple Time Scales*, Comm. Math. Sci. **5**, 495–505, 2007.
21. W. E. B. Engquist, X. T. Li, W. Q. Ren, and E. Vanden-Eijnden, *Heterogeneous Multiscale Methods: A Review*, Comm. Comput. Phys. **2**, 367–450, 2007.
22. M. Müller and K. C. Daoulas, *Speeding up intrinsically slow collective processes in particle simulations by concurrent coupling to a continuum description*, Phys. Rev. Lett. **107**, 227801, 2011.
23. A. N. Semenov, *Relaxation of long-wavelength density fluctuations in a concentrated polymer solution*, JETP **63**, 1230, 1986.
24. A. Z. Akcasu, M. Benmouna, and H. Benoit, *Application of random phase approximation to the dynamics of polymer blends and copolymers*, Polymer **27**, 1935–1942, 1986.
25. M. Benmouna, H. Benoit, R. Borsali, and M. Duval, *Theory of dynamic scattering from copolymer solutions using the random phase approximation*, Macromolecules **20**, 2620–2624, 1987.

Molecular Dynamics Simulations of Curcumin in the Interface Region of Triblock Copolymer Micelles

Konstantin P. K. Luft and Stephan Gekle

Biofluid Simulation and Modeling, Theoretische Physik VI, Universität Bayreuth,
95440 Bayreuth, Germany
E-mail: {konstantin.luft, stephan.gekle}@uni-bayreuth.de

In this work we present the results of an orientation analysis of a single tautomeric curcumin molecule in water and polymeric melts of poly(2-oxazoline) and poly(2-oxazine) based triblock copolymers. The results yield stacking of the aromatic phenyl rings for curcumin solvated in water, while showing that the molecule is mostly prolate when solvated in the polymeric melts. The comparison of curcumin in the two different triblock copolymers however showed no significant difference in the two observed reaction coordinates d and θ.

1 Introduction

Efficient delivery of drug molecules to patients is a major problem. In most cases, drug transportation cannot simply be achieved in an aqueous solution, since the majority of modern drugs have very poor water solubility, but requires some sort of carrier molecule which needs to have both, hydrophilic and hydrophobic parts, to be able to dissolve the drug molecules. Two examples of such hydrophobic drugs are paclitaxel (PTX) and curcumin (CUR), with water solubilities of $5.56\,\mathrm{mg/L}$[9] and $5.75\,\mathrm{mg/L}$,[10] respectively.

Lübtow *et al.*[1] showed a comparison of the maximal drug concentration of these drugs achievable in ABA triblock copolymers based on a hydrophilic part **A**, consisting of poly(2-methyl-2-oxazoline) (pMeOx), and a hydrophobic part **B**, consisting of poly(2-oxazoline) (pOx) or poly(2-oxazine) based polymers (pOzi), namely poly(2-n-propyl-2-oxazoline) (pPrOx), poly(2-butyl-2-oxazoline) (pBuOx), poly(2-n-propyl-2-oxazine) (pPrOzi) and poly(2-n-butyl-2-oxazine) (pBuOzi), of which two can be seen in Fig. 1. Surprisingly, they found that the structural isomers **ApBuOxA** and **ApPrOziA** showed the biggest difference in solubility for the different drugs.

While **ApBuOxA** achieved the best PTX drug loading of $\approx\ 48\,\mathrm{wt\,\%}$, **ApPrOziA** showed the second worst drug loading of approximately $25\,\mathrm{wt\,\%}$, corresponding to a drug concentration of $\rho_{\mathrm{PTX}} = 9.04 \pm 0.61\,\mathrm{g/L}$ and $\rho_{\mathrm{PTX}} = 3.27 \pm 0.01\,\mathrm{g/L}$, respectively, for a polymer concentration of $\rho_p = 10\,\mathrm{g/L}$. The same test for CUR displayed a significantly different behaviour, as it showed the highest drug loading of $\approx 54\mathrm{wt\,\%}$ with the polymer **ApPrOziA** and yielded the worst drug loading of $\approx 24\,\mathrm{wt\,\%}$ in combination with **ApBuOxA**, again with a polymer concentration of $\rho_p = 10\,\mathrm{g/L}$ and corresponding to a curcumin concentration of $\rho_{\mathrm{CUR}} = 11.91 \pm 0.59\,\mathrm{g/L}$ and $\rho_{\mathrm{CUR}} = 3.23 \pm 0.18\,\mathrm{g/L}$, respectively.[1]

Schulz *et al.*[2] did research on the morphological properties of similar copolymers with PTX loading, which yielded that the polymer-drug-macromolecular structure forms spherical micelles, with a size of 10 to 25 nm, even at small drug loadings in aqueous solutions. The structural model proposed by Pöppler *et al.*[3] gives a deeper insight into the

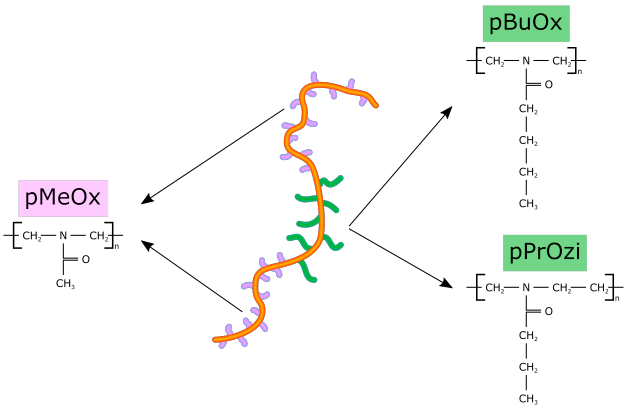


Figure 1. Schematic representation of two of the polymers used by Lübtow *et al.*[1]

properties within the micelles incorporating CUR molecules. The model states that at low drug loadings all drug molecules are present in the hydrophobic core, while the hydrophilic shell of the micelle is filled with water. As the drug loading increases, the core fills up and the drug molecules settle in the hydrophilic part, extruding water from the shell.

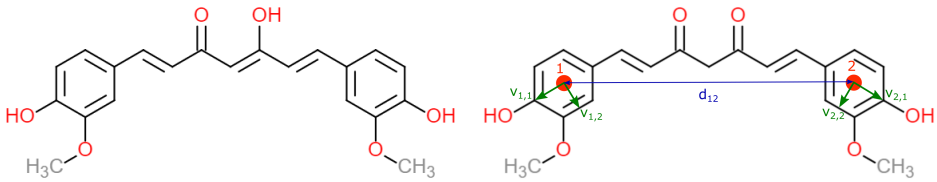
2 Simulation Methods and Systems

System Construction

Our particular interest is the difference in drug loading of the two structural isomers (ApPrOziA and ApBuOxA).[1] To get an idea of the origin of this difference, we used Molecular Dynamics to simulate a small sample system using **ApPrOziA** and **ApBuOxA** polymers and one curcumin molecule in a cubic volume with $a \approx 3.7\,\mathrm{nm}$ edge length. The small system size was chosen to represent a small volume within the interface region between the hydrophilic and hydrophobic parts of the polymer micelle. Due to the lack of water in this region as proposed by Pöppler *et al.*,[3] a polymeric melt was chosen to approximate the polymer movement within the micelles.

The copolymers used in the experiments of Pöppler *et al.*[3] had repeating units of $38{-}23{-}38$ for A-B-A, respectively. To generate the needed *GROMOS 54A7* forcefields, ATb[4, 11–14] was used. The restrictions for ATb's semi-empirical forcefields only allowed molecules with a maximum size of 500 atoms, thus the repeating units had to be adjusted. The hydrophilic/hydrophobic ratio of the experimentally used polymers is $\eta_{\mathrm{exp}} \approx 1.65$. To make most use of the limitation, repeating units of $11{-}7{-}11$ for A-B-A, respectively, were chosen for the simulations, resulting in a ratio of $\eta_{\mathrm{sim}} \approx 1.57$.

In nature, curcumin is a tautomeric molecule, existing in a diketo and in a keto-enol form (see Fig. 2). Manolova *et al.*[5] showed, that the keto-enol form of curcumin is the common form in organic solvents, while the diketo form is rare in this environment. With an increasing percentage of water in a mixture of organic solvents and water, the diketo



(a) Keto-enol

(b) Diketo

Figure 2. Illustration of the two tautomeric forms of curcumin, (a) showing the keto-enol form, (b) shows a schematic representation of the values obtained using the diketo from of curcumin as an example. The COM's, symbolised by the red dots, are used to calculate the distance d and also to calculate the vectors $\vec{v}_{i,j}$, which span the plane of the carbon rings.

form appears more frequent. We decided to simulate both forms of curcumin to achieve a better understanding of what happens within the micelles.

To characterise the system we looked at two reaction coordinates, namely the COM (centre of mass) distance d of the two carbon rings and the angle θ_{12} between the normal vectors of the planes created by the phenyl rings (see Fig. 2). The COM was calculated by using only the six carbon atoms per ring and neglecting the hydroxy and OMe rest, while the angle was calculated by

$$\theta_{12} = \frac{\vec{n}_1 \cdot \vec{n}_2}{|\vec{n}_1| \cdot |\vec{n}_2|} \tag{1}$$

with $\vec{n}_i = \vec{v}_{i,1} \times \vec{v}_{i,2}$.

Due to the OMe group of the phenyl ring, a difference in direction of $\vec{n}_i$ can be determined. Using both parameters one can determine the geometrical configuration of curcumin, which can then be used to generate a 2-dimensional configuration space.

Simulation Sequence

All simulations were done using *GROMACS 2018.1*[6–8] and analysed using the GROMACS analysis template. Each simulation used a time step of $dt = 1\,\text{fs}$, the trajectories were written every $\Delta t = 1\,\text{ps}$.

We built one aqueous system with a single curcumin molecule in diketo form and water using the *solvate* command in *GROMACS* to fill a box with edge length $a = 3\,\text{nm}$. This system was initialised using energy minimisation and a NPT-Ensemble run at a temperature of $T = 400\,\text{K}$ for $t = 10\,\text{ns}$. After the initialisation we obtained 300 start configurations from the trajectory, which were equally distributed within the configurations space (see Fig 4(a)).

The polymer systems were built using the *GROMACS* function *insert-molecules*, by inserting 12 polymer molecules and one of the two curcumin forms in a cubic box with edge length $a = 7\,\text{nm}$, thus creating four systems **PrOzi-Keto**, **BuOx-Keto**, **PrOzi-Enol**

and **BuOx-Enol**. A flow chart of the simulations used to obtain all needed results of the polymer systems can be seen in Fig. 3.

The start configurations were first simulated in an initialisation run consisting of an energy minimisation, followed by a run in a NVT-Ensemble, with a temperature of $T = 700\,\mathrm{K}$ for $t = 500\,\mathrm{ps}$, and a NPT-Ensemble simulation at the same temperature for $t = 30\,\mathrm{ns}$. The high temperature was chosen to allow the long polymer molecules to move faster, the NVT-Ensemble simulation allows the polymers move freely, while the simulation in the NPT-Ensemble lets them compress and entangle, thus achieving a better equilibration and greater structural difference from the start configuration of the polymers, compared to a NPT-run at $300\,\mathrm{K}$ for the same duration. The trajectory obtained by the NPT-run of the initialisation was used to search for different geometrical configurations of curcumin, equally distributed within the configuration space, which resulted in about 80 to 85 start configurations per initialisation run. For all systems the start configurations were then again minimised and used as start configuration in NPT-Ensemble simulations at $T = 300\,\mathrm{K}$ with a simulation time of $30\,\mathrm{ns}$ and $10\,\mathrm{ns}$ for the polymer melts and water, respectively. The resulting trajectories were then analysed and for each time frame the

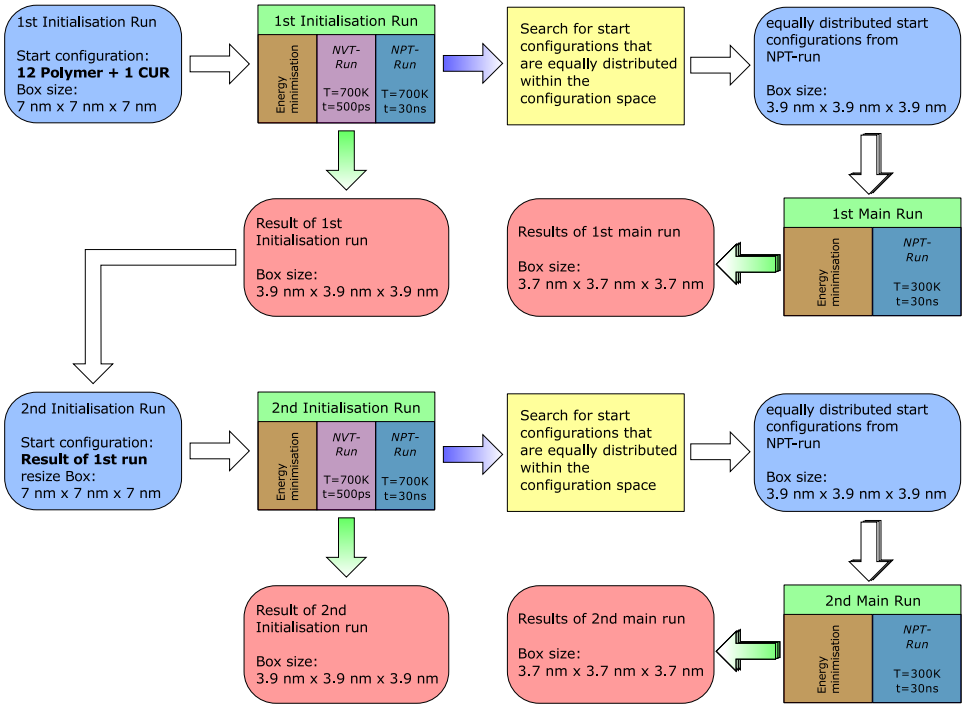


Figure 3. Flow chart of the simulation sequence. The blue rectangles with rounded corners represent the start configurations used, the red rectangles with rounded corners represent the results obtained. The rectangles with green title are showing the separate simulations used within one run (from left to right), while the yellow rectangles represent the program used to find start configurations. The green gradient arrows show that the result is obtained at the end of the run, while the blue gradient arrows show, that the trajectory obtained by the NPT-run was used in the next step.

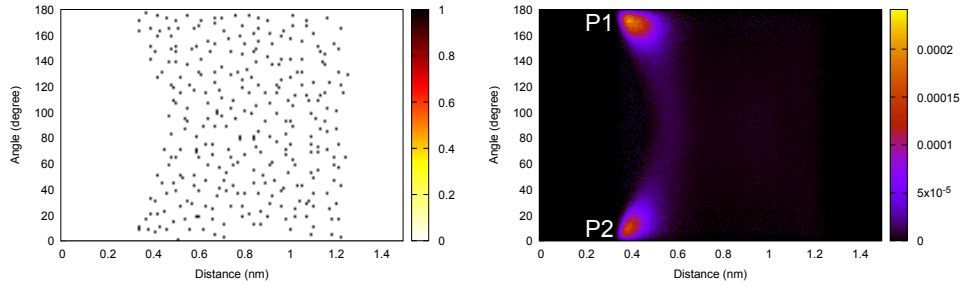
configuration was saved in a two dimensional histogram with bin sizes $B_{\text{angle}} = 0.5°$ and $B_{\text{distance}} = 2.5\,\text{pm}$ (see Figs. 4 and 5).

We used this pattern two times per system, using the result of the first initialisation run to start the second initialisation. To rule out the difference of the polymer configurations in the original initialisation of the systems, when comparing the diketo and the keto-enol forms of curcumin, a program was used to change one form into the other. This way we got four different starting points per system, where the only difference of the *Polymer*-Keto and *Polymer*-Enol systems is the form of curcumin and not the orientation of the polymers.

3 Results and Discussion

The results obtained from the simulation of the aqueous system displayed in a two dimensional configuration space plot can be seen in Fig. 4(b). One can see two prominent regions (P1 and P2) in configuration space, where P1 is at an angle of $\theta_{1,aq} \approx 170°$ and a distance of $d_{1,aq} \approx 0.4\,\text{nm}$ and P2 at $\theta_{2,aq} \approx 10°$ and $d_{2,aq} \approx 0.4\,\text{nm}$. By comparing the two regions one can see that P1 is slightly more prominent than P2. These two configurations correspond to two $\pi - \pi-$stacking forms of CUR's phenyl rings, where the P2 configuration is less frequent, due to the OMe groups of the two rings being next to each other, which can result in a separation of the rings when the OMe groups turn towards one another. In P1 on the other hand, the OMe groups are on opposite sites of the stacked rings and can thus not separate the rings easily, resulting in a higher stability of the P1 configuration. We also observed flipping from one state into the other, resulting in the fine, slightly bent line between the two peaks. The bent line is due to the flipping mostly involving the OMe group of one ring stacking onto the other ring, thus creating a bigger distance between the two rings centre of masses.

The results obtained from the four polymer systems however display a completely different behaviour of CUR compared to the aqueous system. Instead of two precise peaks in the configuration space plots, one can observe a much broader variety of configurations (see Fig. 5). The purple areas in Fig. 5(a)-(d) show the less often found configurations,

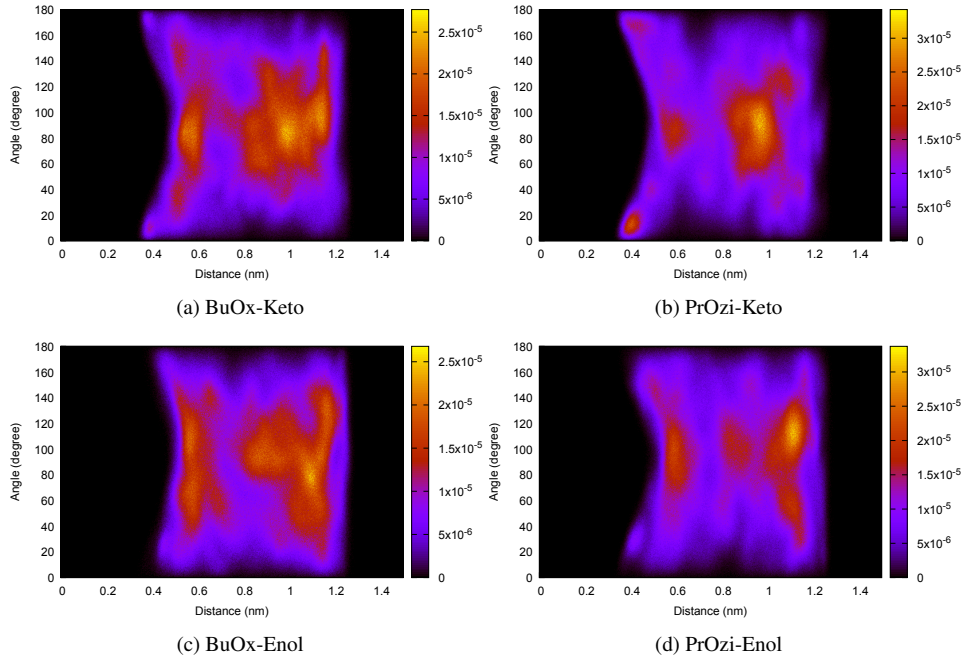


(a) Start configurations (b) Configuration Space

Figure 4. Configuration space plots of the water system, (a) shows the distribution of 300 start configurations within the configuration space, (b) shows the result of the start configurations after 10 ns of NPT-simulation, displaying two peaks P1 (upper peak) and P2 (lower peak).

which seem to be mostly found in CUR molecules moving towards a more prominent configuration. These configurations are displayed by two prominent regions, one at a distance of $d_{1,p} \approx 0.6\,\text{nm}$, the other at distances between $0.8\,nm$ and $1.2\,nm$. The system PrOzi-Keto additionally shows smaller peaks in the same positions as the P1 and P2 peaks in the aqueous system, as can be seen in Fig. 5(b), which again represent $\pi - \pi-$stacking of curcumin's phenyl rings. Further, when comparing the BuOx systems (Fig. 5(a) and (c)) to the PrOzi systems (Fig. 5(b) and (d)), one can see that the PrOzi system displays more confined areas in the configuration space.

However, the difference of the prominent configurations between the two polymers is not significant enough to be the sole reason for better solubility of CUR in ApPrOziA. A comparison of the diketo and keto-enol forms, both within the same polymer, yields that the prominent peaks of the keto-enol forms exhibit a more stretched form in angle direction, while the diketo form displays broader stretching in direction of the distance variable. This can be explained by the difference in bonding of both forms. While the keto-enol form is more planar due to the C-C double bond in the molecules middle part between the oxygen atoms (see Fig. 2(a)), and thus being less able to fold but moving by torsion. The diketo form on the other hand does not have a C-C double bond in the middle part, allowing it to fold more freely (see Fig. 2(b)).



(a) BuOx-Keto

(b) PrOzi-Keto

(c) BuOx-Enol

(d) PrOzi-Enol

Figure 5. Configuration space plots of the polymer systems, (a) shows curcumin in diketo form dissolved in ApBuOxA, (b) the diketo form dissolved in ApPrOziA, (c) the keto-enol form dissolved in ApBuOxA, and (d) the keto-enol form dissolved in ApPrOziA.

4 Conclusion

The simulations done in this work give an impression of how a single curcumin molecule is incorporated within the interface region of triblock copolymer micelles and its orientation in different solutions. While the results of our simulations were able to show a difference in orientation for curcumin solvated in water (stacking of the aromatic phenyl rings) and curcumin in polymeric melt (mostly showing a prolated form), it could not show significant differences when comparing the orientation of curcumin in ApBuOxA and ApPrOziA, thus not confirming the findings of Lübtow *et al.*[1]

A possible reason for this could be that the diffusion time of the polymeric melt is too large to get viable results by simulations as short as 30 ns. By using single curcumin molecules instead of multiple ones we also stray far from the concentration achieved by Lübtow *et al.*[1] and therefore neglect possible curcumin-curcumin interactions, which could also stabilise the micelle and increase drug loading. Taking the structural model for these specific polymer micelles proposed by Pöppler *et al.*[3] into account, simulations of the core and corona regions and comparisons of the different regions could also yield important informations for understanding the high drug loading ability of ApPrOziA, which will be investigated in the future.

Another way to find differences in the two polymer systems could be the radial distribution of different functional groups of the polymers in respect to functional groups of curcumin, since the keto-enol form exhibits an additional hydrogen bond donor, that could possibly alter the way in which the functional groups of the polymers move.

Acknowledgements

The authors gratefully acknowledge the computing time granted by the Jülich Supercomputing Centre (JSC).

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 326998133 - TRR 225 (subproject B07). Special thanks belong to Robert Luxenhofer for useful discussions.

References

1. M. M. Lübtow, L. Hahn, M. S. Haider, and R. Luxenhofer, *Drug Specificity, Synergy and Antagonism in Ultrahigh Capacity Poly(2-oxazoline)/Poly(2-oxazine) based Formulations*, J. Am. Chem. Soc. **139**, 10980–10983, 2017.
2. A. Schulz, S. Jaksch, R. Schubel, E. Wegener, Z. Di, Y. Han, A. Meister, J. Kressler, A. V. Kabanov, R. Luxenhofer, C. M. Papadakis, and R. Jordan, *Drug-Induced Morphology Switch in Drug Delivery Systems Based on Poly(2-oxazoline)s*, ACS Nano **8**, 2686–2696, 2014.
3. A.-C. Pöppler, M. M. Lübtow, J. Schlauersbach, J. Wiest, L. Meinel, and R. Luxenhofer, *Loading Dependent Structural Model of Polymeric Micelles by Solid-State NMR*, Preprint,2019, doi:10.26434/chemrxiv.8943251.v1.
4. A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink, and A. E. Mark, *An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0*, Journal of Chemical Theory and Computation **7**, 4026–4037, 2011.

5. Y. Manolova, V. Deneva, L. Antonov, E. Drakalska, D. Momekova, and N. Lambov, *The effect of the water on the curcumin tautomerism: A quantitative approach*, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **132**, 815–820, 2014.
6. E. Lindahl, B. Hess, and D. van der Spoel, *GROMACS 3.0: A package for molecular simulation and trajectory analysis*, J. Mol. Mod. **7**, 306–317, 2001.
7. S. Páll, M. J. Abraham, C. Kutzner, B. Hess, and E. Lindahl, *Tackling exascale software challenges in molecular dynamics simulati ons with GROMACS*, Solving Software Challenges for Exascale **8759**, 3–27, 2015.
8. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers*, SoftwareX **1**, 19–25, 2015.
9. `https://www.drugbank.ca/drugs/DB01229`
10. `https://www.drugbank.ca/drugs/DB11672`
11. `https://atb.uq.edu.au/molecule.py?molid=275982`
12. `https://atb.uq.edu.au/molecule.py?molid=8102`
13. `https://atb.uq.edu.au/molecule.py?molid=359238`
14. `https://atb.uq.edu.au/molecule.py?molid=350170`

Earth and Environment

Earth and Environment

Thomas Jung

Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research,
Bussestraße 24, 27570 Bremerhaven, Germany
E-mail: Thomas.Jung@awi.de

Department of Physics and Electrical Engineering, University of Bremen,
28334 Bremen, Germany

The field of Earth and Environment is an important research area that tackles a number of grand challenges of high societal relevance. Prominent examples include global and climate change, biodiversity loss and geo-risks. Simulation, one of the four pillars of science (augmenting theory, experiment and more recently data science), has a long tradition. In fact, weather forecasting has been one of the applications that have been run on the first computer by John von Neumann and his team. Nowadays, modelling is routinely being used in Earth and Environment, with prominent examples including numerical experimentation (enhanced system understanding), predictions (statements about possible future trajectories of systems), and data assimilation (advanced state estimation by optimally combining models and observations). Having access to the largest high-performance computing (HPC) systems, like the ones made available through NIC, is critical for advancing science, since large computational resources usually allow for a high degree of realism when it comes to simulation systems numerically. The three projects described in this issue – two relating to hydrological and terrestrial systems and one tackling inverse problems in geosciences – demonstrate this point very clearly.

Harrie-Jan Hendricks Franssen and colleagues use the Terrestrial Systems Modelling Platform (TSMP) along with the Parallel Data Assimilation Framework (PDAF) to get improved estimates of terrestrial systems for two different situations, namely a small area with hillslope and continental scale applications at high spatial resolution. Carrying out data assimilation experiments for such high-resolution coupled systems is absolutely cutting-edge, especially in the light of the excellent scalability characteristics demonstrated on different HPC systems at Jülich Supercomputing Centre (JSC).

In the second study, Prabrakhar Shrestha and colleagues make extensive use of the supercomputer JURECA at JSC to advance our understanding of the terrestrial water cycle by addressing a number of different aspects, using the Terrestrial Systems Modelling Platform (TSMP). In general, it is found that increased realism of the system increased our understanding. At the same time, however, it is pointed out that uncertainty may also increase due to poorly constrained aspects of the model, calling for further advances in computational resources along with ensemble approaches to quantify inevitable uncertainty.

Finally, the contribution from Bohlen and colleagues deals with the development and application of a new technique for the inversion of shallow-seismic and ultrasound data using full-waveform inversion, with the latter providing very realistic results but requiring large computational resources. It is shown that this technique allows to map structures at very high resolution compared to other methods used more commonly in previous studies.

Data Assimilation with the Integrated Terrestrial System Model TSMP-PDAF

Harrie-Jan Hendricks Franssen[1], **Wolfgang Kurtz**[2], **Bibi Naz**[1], **and Sebastian Gebler**[1]

[1] IBG-3 Agrosphere, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {h.hendricks-franssen, b.naz, s.gebler}@fz-juelich.de

[2] Leibniz Supercomputing Centre, Boltzmannstraße 1, 85748 Garching, Germany
E-mail: wolfgang.kurtz@lrz.de

This paper discusses predictions with integrated terrestrial system models, which model water and energy cycles from the deep subsurface to the upper atmosphere. Such predictions can be improved by data assimilation, making use of supercomputing. Two examples on very different scales (hillslope and continental) illustrate the strengths and remaining uncertainties of the approach.

1 Introduction

Traditionally, in hydrology, soil science and ecology compartment-specific simulation models are employed. For example, groundwater models were developed able to simulate groundwater flow in (heterogeneous) aquifers, and soil hydrological models able to simulate water flow in the unsaturated zone including soil evaporation and plant transpiration. In the last two decades integrated hydrological modelling and integrated terrestrial systems modelling have gained importance. In these integrated models multiple compartments of the terrestrial system are simulated jointly, including their two-way feedbacks. The development of integrated models is driven by the fact that the different compartments of the terrestrial system show strong non-linear interactions. In addition, increase in compute power made integrated model simulation with many unknowns feasible, first on large computing clusters, and later also on PC's. Examples for integrated hydrological models, which simulate water flow in soil, aquifer and streams in a coupled fashion are ParFlow[1] and HydroGeoSphere.[2]

Land surface models simulate the exchange of water and energy – the most recent model generation(s) also carbon and nitrogen – between the land and the atmosphere. Land surface models incorporated over time increasingly sophisticated representations of soil and vegetation to better simulate these exchange processes. Land surface models were therefore in an early stage already integrated models. Nevertheless, land surface models required a more mechanistic based representation of subsurface flow and heat transport, which gave rise to the development of integrated terrestrial system models, which couple land surface models and integrated hydrological models, like for example CLM-ParFlow.[3] More recently, such models were also coupled to atmospheric circulation models, as is the case of the Terrestrial System Modelling Platform (TSMP) that we use in this work.[4]

Although models like TSMP provide a more mechanistic and complete representation of the water and energy cycles in terrestrial systems than classical compartment-specific models, simulations are affected by large uncertainties related to uncertain initial and

boundary conditions and input parameters. Data assimilation approaches allow to estimate and reduce simulation uncertainties. Data assimilation can constrain initial conditions and parameters by correcting model simulations with measurements, which is based on the optimal weighting of simulated values on the one hand and measured values on the other hand. Ensemble based data assimilation, which characterises model uncertainty with a larger number of model runs, is used in this work. Examples of ensemble-based data assimilation algorithms are the Ensemble Kalman Filter[5] and the Particle Filter.[6]

This paper discusses first the Terrestrial Systems Modelling Platform, and the data assimilation framework coupled to it. Afterwards, two application examples are presented.

2 The Terrestrial Systems Modelling Platform TSMP

The Terrestrial System Modelling Platform (TSMP)[4] is a modular Earth system model which combines already pre-existing parallel compartment models for the atmosphere (COSMO[7]), the land surface (CLM version 3.5[8]) and the subsurface (ParFlow[1, 9, 10]). COSMO is a convection-permitting atmospheric model which is used as the operational forecast model of the German weather service. CLM is a land surface model that simulates the transfer and partitioning of energy, momentum, water and carbon fluxes between the atmosphere, vegetation and the subsurface. ParFlow calculates variably-saturated subsurface flow and surface water routing (using the kinematic wave approximation) in an integrated approach.[1]

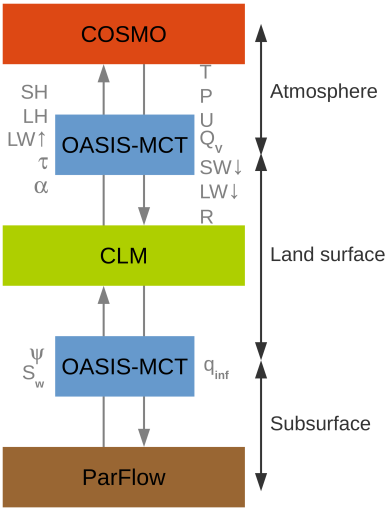


Figure 1. Coupling of the TSMP component models ParFlow (subsurface), CLM (land surface) and COSMO (atmosphere) by OASIS-MCT. The exchanged fluxes and state variables are: ψ (subsurface pressure), S_w (subsurface saturation), q_{inf} (net infiltration flux), SH (sensible heat flux), LH (latent heat flux), LW $\uparrow$ (outgoing long wave radiation), τ (momentum flux), α (albedo), P (air pressure), T (air temperature), U (wind velocity), SW $\downarrow$ (incoming short wave radiation), LW $\downarrow$ (incoming long wave radiation), Q_{V} (specific humidity) and R (precipitation).

The coupling library OASIS-MCT[11, 12] is used to connect the three component models of TSMP by exchanging information on fluxes and state variables at the conceptual boundaries of the respective compartment models (see Fig. 1) meaning that each compartment model acts as a lower or upper boundary condition for the adjacent compartment model. The data exchange between models via OASIS-MCT is managed by one global communicator (`MPI_COMM_WORLD`) that is shared between the involved compartment models. Information on exchanged variables and corresponding interpolation and scaling options as well as the required communication patterns are defined in the initialisation phase of the compartment models via OASIS-MCT library calls. The models are then dynamically coupled at runtime through information exchange at predefined time intervals. Key features of TSMP are its integral view on terrestrial processes, its improved physical process description (as compared to single compartment models), its modularity (different model combinations can be chosen) and a good scalability on HPC platforms.[12]

3 The Data Assimilation Framework TSMP-PDAF

Data assimilation with TSMP is enabled by coupling TSMP to the Parallel Data Assimilation Framework (PDAF).[13] PDAF[14] is a generic data assimilation library that provides a variety of parallel implementations of ensemble filters that are commonly used for data assimilation. In addition, PDAF provides functionality for data exchange between ensemble members and the parallel data assimilation algorithms. For the coupling of TSMP with PDAF, a set of functions needed to be defined, that provide PDAF with information and data for the assimilation. This includes, for example, functions to extract the relevant state variables of the TSMP models and functions for providing information on the observation data.

In the initialisation phase of TSMP-PDAF (before TSMP models are initialised), PDAF defines three MPI communicators for facilitating data exchange for the assimilation and the ensemble propagation (Fig. 2). The "model communicator" is defined for each model realisation and is used for the model forward integration. The "filter communicator" is used for the parallel filter algorithm and is only active for the processors of the first model realisation. The "coupling communicator" is used to collect/distribute information on model states across different model realisations to/from the filter communicator. Note that each rank in the model and coupling communicators usually only holds a part of the global model state vector. For using PDAF with the coupled TSMP model, modifications of the OASIS-MCT library were necessary to allow the ensemble integration. Specifically, for the data exchange between different compartment models of TSMP with OASIS-MCT, the model communicators are used instead of `MPI_COMM_WORLD` (see Sec. 2).

The PDAF filter algorithms are called during the time integration of the TSMP models at predefined assimilation intervals when measurements become available. In a first step, the relevant model states of the ensemble members are collected to the filter communicator by the coupling communicator. Then the filter algorithm is executed within the filter communicator using the available observation data. The updated model states are then distributed to the ensemble members via the coupling communicator and the TSMP models are integrated to the next assimilation step.

Similarly to TSMP, the data assimilation framework TSMP-PDAF is modular and thus allows to perform data assimilation with different model combinations as well as stand-

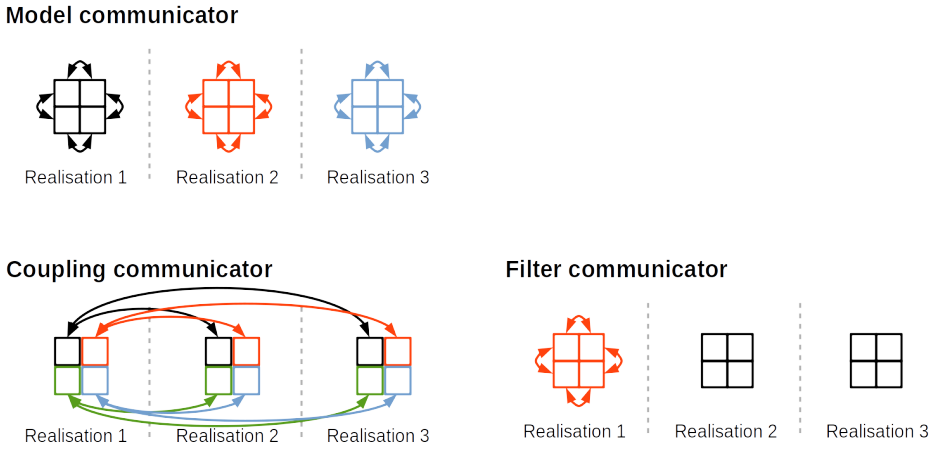


Figure 2. Communicators in PDAF for a parallel setup with 3 ensemble members and 4 processors per ensemble member. Colours indicate the membership of the respective processors and arrows exemplify the parallel communication between the different processors.

alone models of TSMP. Currently supported observation data are point-scale soil moisture measurements (for ParFlow and CLM), river stages and groundwater levels (both for ParFlow). TSMP-PDAF also allows to update several model parameters (hydraulic conductivity, Manning's coefficients and soil texture) with the mentioned observations.

The scalability of TSMP-PDAF was tested with a synthetic data assimilation experiment with a ParFlow-CLM model configuration on JUQUEEN (Fig. 3) using different scenarios regarding model output and ensemble quality. An ideal setup with identical ensemble members was used to determine the computational overhead of the ensemble integration and the filtering step. Results showed that the overhead starts to become relevant with more than 8192 processors. When using a realistic ensemble with varying model parameters, the scalability decreases due to the load imbalance caused by different run times of the individual ensemble members. The model output scenarios showed low variations of the total run time for the realistic ensemble meaning that the load imbalance had a more prominent effect for the utilised data assimilation setup.

4 Example of Data Assimilation for Soil Moisture

4.1 Hillslope Scale High-Resolution Integrated Land Surface-Subsurface Modelling

The CLM and ParFlow components of the TSMP-model were used for land surface-subsurface modelling of the highly equipped Rollesbroich site (0.38 km^2) in the Eifel hills in North-Rhine-Westphalia in Germany. At this site, information from a soil moisture sensor network measuring soil moisture content at 61 locations (5, 20 and 50 cm depth) was

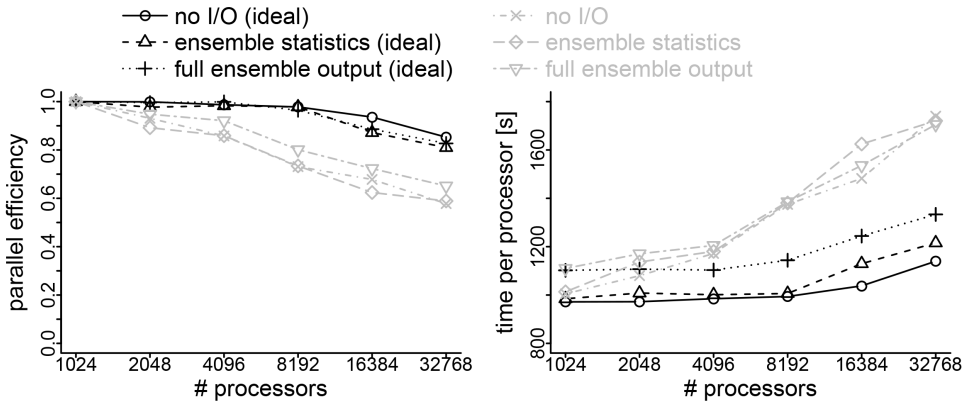


Figure 3. Scaling behaviour (left) and timing information (right) for TSMP-PDAF for a weak scaling test on JUQUEEN. Black lines show results for an idealised test case (identical ensemble members) and grey lines show results for a heterogeneous ensemble. The number of ensemble members is increased from 8 to 256. Each ensemble member used 32 processors for CLM and 96 processors for ParFlow. Different lines refer to experiments with different model output.

available. The research question was, whether for this hillslope research site the combination of physically based modelling with the TSMP-model and the dense observation network, allow to make an accurate characterisation of hydrological states and fluxes including discharge. In other works, it is often found that predicting discharge at the hillslope scale is particularly challenging.

This question was addressed with TSMP-PDAF and for the year 2011, where the Ensemble Kalman Filter was used for assimilating the data from the soil moisture sensor network. An ensemble of 128 (or 256) realisations with different 3D-heterogeneous fields of Mualem-van Genuchten parameters was generated. Also atmospheric forcings were considered uncertain and perturbed for each of the 128 (or 256) realisations. Different assimilation scenarios were run, including state updating alone, and joint state and parameter updating. Simulations were also carried out with a synthetic test case mimicking the Rollesbroich site, to get more insight in the role of model structural errors.

The combination of joint updating of model states and hydraulic conductivity was more efficient for soil water content (SWC) characterisation than state updating alone for the real-world case. On average, the root mean square error (RMSE), which measures the difference between measured and simulated SWC at the sensor locations, was reduced by 14% if states and parameters were updated jointly, but discharge estimation was not improved significantly. See also Fig. 4. Synthetic simulations showed much better results with an overall RMSE reduction by 55% at independent verification locations in case of daily SWC data assimilation including parameter estimation. Individual synthetic data assimilation scenarios with parameter estimation showed an increase of the Nash-Sutcliffe-Efficiency for discharge from -0.04 for the open loop run to 0.61. This shows that data assimilation in combination with high-resolution physically-based models can potentially strongly improve soil moisture and discharge estimation at the hillslope scale. Additional simulation experiments were performed to understand the difference between the real-world

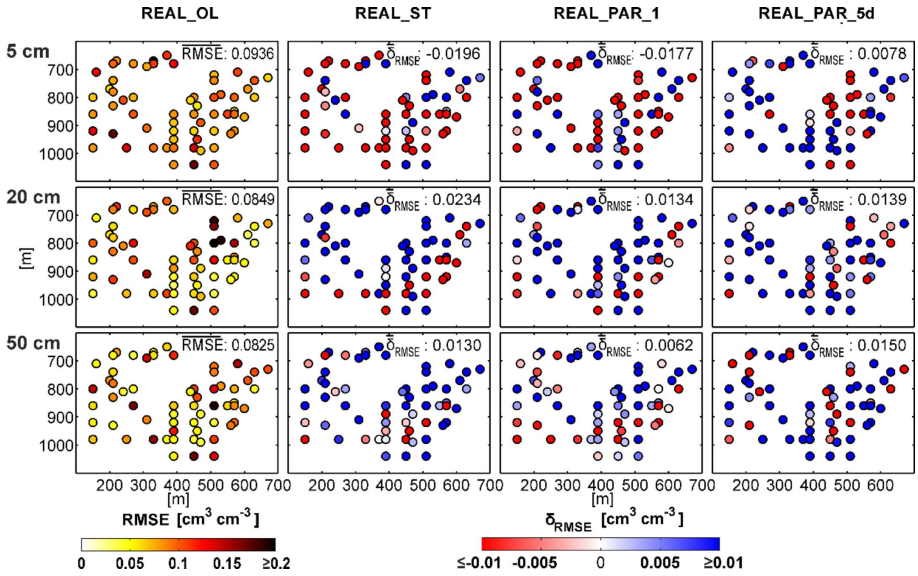


Figure 4. RMSE of SWC at individual locations for the open loop runs (left column) and changes in RMSE (increase implies improvement and decrease implies impoverishment) for three data assimilation scenarios (state updating alone (ST), daily joint state and parameter updating (PAR1) and joint state and parameter updating every five days (PAR5d) (three columns on the right)) of the real-world case for 2011.

case and the synthetic case. It was found that erroneous prior values of the geostatistical parameters are for example already able to explain a considerable part of the difference in performance between the real-world and synthetic case. On the other hand, if only the saturated hydraulic conductivity was unknown (and other soil hydraulic parameters were known), the performance hardly improved compared to the case that all soil hydraulic parameters were unknown. In summary, the large performance difference between synthetic and real-world experiments indicates the limits of such an approach associated with model structural errors like errors in the prior geostatistical parameters.

4.2 Continental-Scale High-Resolution Land Surface Data Assimilation System

Soil moisture is a key state variable which controls the exchange of water, energy and carbon fluxes between the land surface and atmosphere.[15] As a result, it plays an important role in many regional–scale applications, including meteorology, hydrology, flood forecasting, drought monitoring, agriculture and climate change impact studies.[16] Because of its high spatiotemporal variability, it is difficult to monitor soil moisture at large spatial scales. The knowledge of soil moisture at large scale, with reasonable temporal and spatial resolution, is therefore needed to provide locally representative information of soil moisture for regional hydrologic and agriculture applications. However, soil moisture is a difficult variable to obtain because there are no high-resolution soil moisture observations available at the continental scale, and observations from measurements or remote sensing products are sparse and temporally and spatially discontinuous. While land surface

models can provide high resolution large scale high resolution soil moisture estimates with complete spatial and temporal coverage, the soil moisture fields from models often suffer substantial errors owing to errors either in model forcing (such as precipitation, temperature and radiation) or inadequate model physics. We used the land surface data assimilation system CLM-PDAF to utilise the coarse resolution satellite soil moisture data to update the soil moisture estimates from the land surface model. A key capability of CLM-PDAF is the support for data assimilation that combines land surface processes with satellite and in-situ observations for the estimation of optimal land surface states. The data assimilation structure in CLM-PDAF allows to directly ingest remotely sensed high resolution observations of land surface conditions to produce accurate, spatially and temporally consistent fields of land surface states, with reduced associated error. The CLM-PDAF uses an Ensemble Kalman Filter (EnKF) algorithm to generate assimilated or reanalysis products. To effectively simulate the background-error covariances, a large enough ensemble size needs to be maintained in the data assimilation process, which linearly increases the computational resource requirements. Hence, we implemented the CLM-PDAF over Europe to provide downscaled estimates of the soil moisture with complete spatiotemporal coverage by combining historical satellite soil moisture (SM) observations with a high resolution land surface model (LSM) using data assimilation techniques. Using the CLM-PDAF, the satellite based soil moisture dataset ESA CCI (the European Space Agency Climate Change Initiative[17]) was assimilated into CLM using the EnKF[5] producing a high-resolution European surface soil moisture (SSM) reanalysis (called ESSMRA hereafter) dataset. This product overcomes the shortcomings of sparse spatial and temporal datasets and provides a better estimate of SM than obtained only by modelling or by sparse observations alone.

4.2.1 Work Flow for Generation of ESSMRA

The 3 km ESSMRA is generated by first implementing the regional land surface model setup coupled with the data assimilation framework as shown in Fig. 5a. In the second step the ESA CCI satellite-based data is assimilated into the CLM-PDAF setup to generate the daily 3 km ESSMRA product over Europe for the 2000 – 2015 time period (Fig. 5b). The ESSMRA dataset is also compared with other global soil moisture reanalysis products from the European Centre for Medium-Range Weather Forecasts Reanalysis 5[18](ERA5), the Global Land Data Assimilation System[19] (GLDAS) and the Global Land Evaporation Amsterdam Model[20] (GLEAM) data available at 0.25° resolution and hourly temporal resolution.

To assess the ability of ESSMRA to capture short term soil moisture variability in comparison to other existing reanalysis products (such as ERA5, GLDAS and GLEAM), summer standardised anomalies of SM were calculated ($SMA = \frac{SM_{jja} - \overline{SM}}{\sigma_{SM}}$) for a dry year (2003), wet year (2007) and normal year (2011) using average SM values over June, July and August (JJA) relative to the mean JJA SM for the 2000–2015 period. The spatial distribution shows patterns of positive and negative SM anomalies over Europe across all datasets for dry, wet and normal years (Fig. 6). For dry year 2003 (a record heat wave over Europe), CLM-DA (CLM estimated soil moisture with data assimilation) shows a similar area extent of negative anomalies as the ESACCI dataset and ERA5, whereas CLM-OL (CLM estimated soil moisture without data assimilation), GLDAS and GLEAM exhibits much stronger negative anomalies over central Europe. The SM anomaly from CLM-DA for the wet and normal years (2007, 2011) has a better match with ESACCI and other

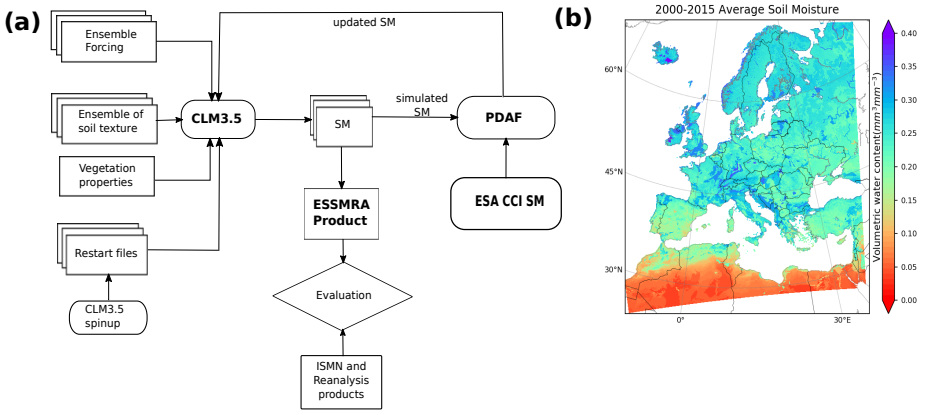


Figure 5. (a) Schematic of CLM-PDAF workflow adopted to generate high resolution ESSMRA product. (b) Average soil moisture (mm^3 mm^{-3}) for 2000 – 2015 time period.

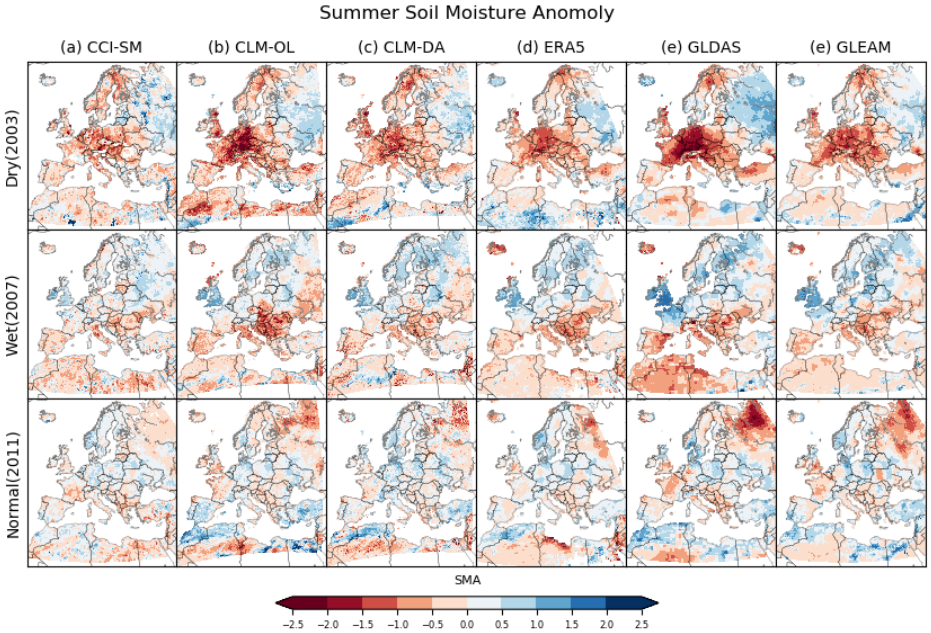


Figure 6. Spatial distribution of the standardised summer (JJA) soil moisture anomaly for the year 2003 (dry year), 2007 (wet year) and 2011 (normal year) and compared with existing reanalysis products. The summer anomaly is calculated for (a) satellite (ESACCI), (b) CLM-OL (c) CLM-DA, (d) ERA5, (e) GLDAS and (f) GLEAM.

reanalysis datasets except GLDAS which shows much stronger wet and dry anomalies than others.

5 Concluding Remarks

The parallel data assimilation framework (PDAF) has been coupled to the terrestrial modelling platform (TSMP) in order to condition integrated terrestrial system simulations to measurements. The framework shows a very good scalability on the Jülich supercomputers. The performance of TSMP-PDAF has been tested in synthetic and real-world cases, and at different spatial scales. Here we illustrate that the data assimilation performed well for simulations at the hillslope scale and continental scale.

Acknowledgements

The authors gratefully acknowledge the computing time granted through JARA, GCS and ESM on the supercomputers JURECA, JUQUEEN and JUWELS at Jülich Supercomuting Centre (JSC).

References

1. S. J. Kollet and R. M. Maxwell, *Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model*, Advances in Water Resources **29**, 945–958, 2006.
2. R. Therrien and E. Sudicky, *Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media*, Journal of Contaminant Hydrology **23**, 1–44, 1996.
3. S. Kollet and R. M. Maxwell, *Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model*, Water Resources Research **44**, W02402, 2008.
4. P. Shrestha, M. Sulis, M. Masbou, S. Kollet, and C. Simmer, *A Scale-Consistent Terrestrial Systems Modeling Platform Based on COSMO, CLM, and ParFlow*, Monthly Weather Review **142**, 3466–3483, 2014.
5. G. Burgers, P. J. van Leeuwen, and G. Evensen, *Analysis scheme in the ensemble Kalman filter*, Monthly Weather Review **126**, 1719–1724, 1998.
6. N. J. Gordon, D. J. Salmond, and A. F. M. Smith, *Novel-approach to nonlinear non-Gaussian bayesian state estimation*, IEEE Proceedings-F Radar and Signal Processing **140**, 107–113, 1993.
7. M. Baldauf, A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, *Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities*, Monthly Weather Review **139**, 3887–3905, 2011.
8. K. W. Oleson, G.-Y. Niu, Z.-L. Yang, D. M. Lawrence, P. E. Thornton, P. J. Lawrence, R. Stöckli, R. E. Dickinson, G. B. Bonan, S. Levis *et al.*, *Improvements to the Community Land Model and their impact on the hydrological cycle*, Journal of Geophysical Research **113**, G01021, 2008.
9. S. F. Ashby and R. D. Falgout, *A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations*, Nuclear Science and Engineering **124**, 145–159, 1996.

10. R. M. Maxwell, *A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling*, Advances in Water Resources **53**, 109–117, 2013.
11. S. Valcke, *The OASIS3 coupler: a European climate modelling community software*, Geoscientific Model Development **6**, 373–388, 2013.
12. F. Gasper, K. Goergen, P. Shrestha, M. Sulis, J. Rihani, M. Geimer, and S. Kollet, *Implementation and scaling of the fully coupled Terrestrial Systems Modeling Platform (TerrSysMP v1.0) in a massively parallel supercomputing environment – a case study on JUQUEEN (IBM Blue Gene/Q)*, Geoscientific Model Development **7**, 2531–2543, 2014.
13. W. Kurtz, G. He, S. J. Kollet, R. M. Maxwell, H. Vereecken, and H. J. Hendricks Franssen, *TerrSysMP-PDAF (version 1.0): a modular high-performance data assimilation framework for an integrated land surface-subsurface model*, Geoscientific Model Development **9**, 1341–1360, 2016.
14. L. Nerger and W. Hiller, *Software for ensemble-based data assimilation systems-Implementation strategies and scalability*, Computers & Geosciences **55**, 110–118, 2013.
15. S. I. Seneviratne, T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, *Investigating soil moisture-climate interactions in a changing climate: A review*, Earth-Science Reviews **99**, 125–161, 2010.
16. L. Brocca, F. Melone, T. Moramarco, and R. Morbidelli, *Spatial-temporal variability of soil moisture and its estimation across scales*, Water Resources Research **46**, W02516, 2010.
17. W. Wagner, W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, and M. Ertl, *Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture*, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Annals) **7**, 315–321, 2012.
18. "Copernicus Climate Change Service (C3s) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change Service Climate Data Store (CDS)".
19. M. Rodell, P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, *The Global Land Data Assimilation System*, Bulletin of the American Meteorological Society **85**, 381–394, 2004.
20. B. Martens, D. G. Miralles, H. Lievens, R. van der Schalie, R. A. M. de Jeu, D. Fernández-Prieto, H. E. Beck, W. A. Dorigo, and N. E. C. Verhoest, *GLEAM v3: satellite-based land evaporation and root-zone soil moisture*, Geoscientific Model Development **10**, 1903–1925, 2017.

Uncertainty in Terrestrial Water Cycle Simulations

Prabhakar Shrestha[1], Mauro Sulis[2], Stefan Poll[1], Tim Hoar[3], Harrie-Jan Hendricks Franssen[4,5], Clemens Simmer[1,5], and Stefan Kollet[4,5]

[1] Institute for Geosciences, Department of Meteorology, University of Bonn, Bonn, Germany
E-mail: {pshrestha, stefan.poll, csimmer}@uni-bonn.de

[2] Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg
E-mail: mauro.sulis@list.lu

[3] National Center for Atmospheric Research, Boulder, Colorado, USA
E-mail: thoar@ucar.edu

[4] Institute of Bio- and Geosciences IBG-3 (Agrosphere), Forschungszentrum Jülich, Jülich, Germany
E-mail: {h.hendricks-franssen, s.kollet}@fz-juelich.de

[5] Centre for High-Performance Scientific Computing in Terrestrial Systems, Geoverbund ABC/J, Jülich, Germany

TerrSysMP has been exploited to advance our understanding of terrestrial water cycle, by conducting km-scale simulations from field scale to continental scale at the massively parallel supercomputing environment of the Jülich Supercomputing Centre (JSC). The numerical simulations have led to quantification of uncertainties in the simulated terrestrial water cycle in terms of grid-scale representation of heterogeneity and bio-geophysical parameterisations. Ensemble simulations are thus prerequisite to quantify the uncertainty in the terrestrial water cycle, which then could also be utilised for data assimilation to improve prediction.

1 Introduction

The Terrestrial Systems Modelling Platform (TerrSysMP or TSMP)[1] with upgrade to OASIS3-MCT coupler for km-scale continental simulations in massively parallel supercomputing environments[2] has aided in advancing our understanding of the terrestrial water cycle including process related to *e. g.* vegetation and groundwater atmosphere connections.[1, 3–6] In an additional effort, Uebel *et al.*[7] extended TerrSysMP to simulate the terrestrial carbon cycle by including biogenic and anthropogenic CO_2 sources, however its use is currently limited due to non-readily available initial and lateral boundary conditions for CO_2. In the case of the terrestrial water cycle, Shrestha *et al.*[1] showed the presence of strong linkages between integrated surface water and groundwater dynamics (*e. g.* surface runoff with redistribution, groundwater abstraction) and atmospheric boundary layer (ABL) evolution. Their study further demonstrated the importance of hyper-resolution (< 1 km) for coupled simulation with the inclusion of groundwater models. Rahman *et al.*[3] demonstrated that groundwater table dynamics do affect atmospheric boundary layer height, convective available potential energy, and precipitation via the coupling with land surface soil moisture and energy fluxes for convective events over Western Germany. Keune *et al.*[4] also showed that the inclusion of a physically-based groundwater model significantly impacts the simulation of the land surface-atmosphere processes at the continental scale over Europe. However, uncertainty in the modelled groundwater-vegetation-

atmosphere feedbacks might persist when different atmospheric and land surface models are used, due to differences in modelled surface energy flux partitioning and precipitation patterns as shown by Sulis *et al.*[5] Besides, the inclusion of new physical processes does not often result in the expected improvement in predictions,[6] which is partly due to uncertainty in vegetation parameters[8] and the relative wet bias introduced with the inclusion of groundwater due to the lack of adequate resolution (often limited by computational burden over large regional domains[9]).

The modularity of TSMP has also allowed some flexibility to conduct land-atmosphere feedback experiments without including the explicit groundwater model[7, 8, 10, 11] or catchment hydrology simulations with offline atmospheric forcing.[9, 12, 13] The above studies have further advanced our understanding in terms of modelled pathways of evapotranspiration, grid resolution dependency of boundary layer mixing and surface energy flux partitioning, constituting roadmaps where further work is essential to improve the predictions by applying more physically based processes in our models.

In addition, the development of TSMP interface with modular data assimilation (DA) tools (*i. e.* applicable to any selected component models) has opened doors to improve model predictions, generate re-analysis data and investigate joint state parameter updates and model structural errors.[14–17] Kurtz *et al.*[14] implemented the DA interface for the hydrological component of TSMP with the Parallel Data Assimilation Framework (PDAF) which facilitates DA simulations at km-scale resolutions over regional domains, making efficient use of massively parallel supercomputing environments. Shrestha *et al.*[15] exploited PDAF for a DA of soil moisture (using 256 ensemble members) with the soil parameters to tune the rooting depth distribution and improve the simulated soil moisture variance and surface energy flux partitioning for a grassland site in Germany. Naz *et al.*[17] explored the potential of assimilating satellite soil moisture observations to produce downscaled and improved high-resolution soil moisture and runoff simulations at the continental scale. Their study showed general improvements in runoff over the continental scale, although also for some regions degraded performance was observed.

In the last 5 years, TSMP has evolved into a state-of-the-art terrestrial model to predict and quantify uncertainties in the terrestrial water cycle in terms of grid-scale representation of heterogeneity and bio-geophysical parameterisations from field scale to regional domains, exploiting the high-performance computing resources at JSC. In this work, we briefly report on simulations with TSMP performed on JURECA over western Germany to investigate:

1. groundwater modulation of the atmosphere at seasonal scales
2. effect of grid resolution on pathways of evapotranspiration
3. atmospheric boundary layer (ABL) schemes for hyper-resolution runs
4. weakly coupled data assimilation using the newly implemented DA interface with DART (Data Assimilation Research Testbed)[18] for all component models

2 Groundwater Modulation of Atmosphere under Seasonal Scales

In this study we evaluate the impact of increasing the complexity of subsurface-land surface physical processes on the performance of the numerical weather prediction model

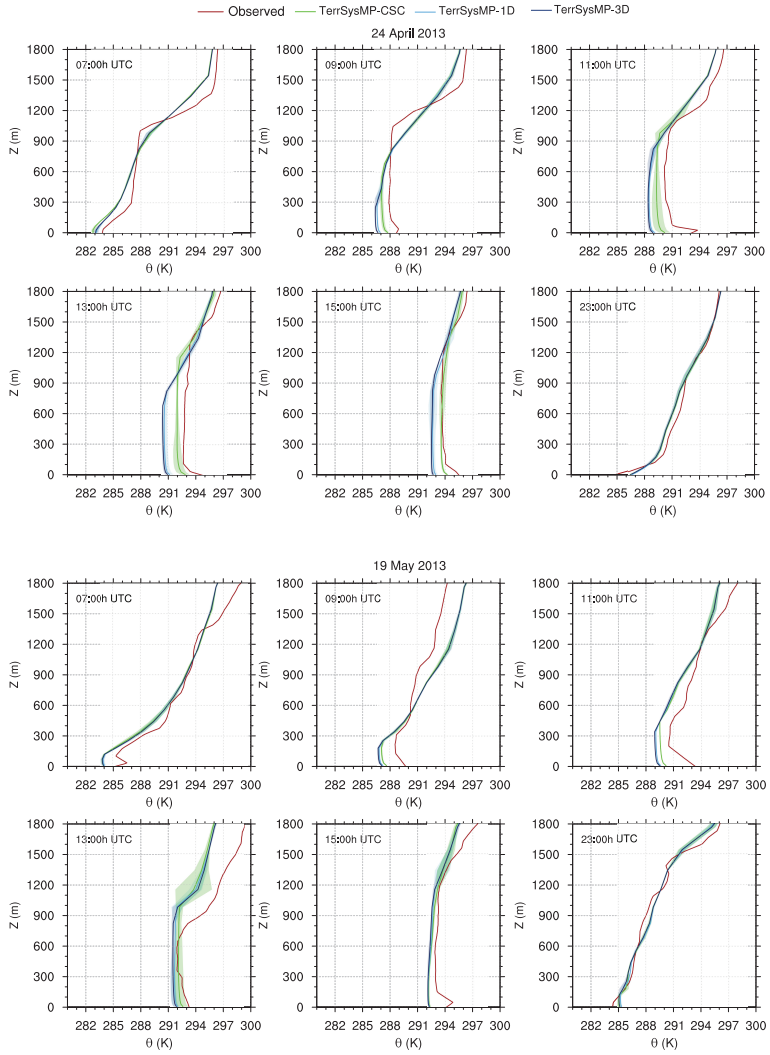


Figure 1. Observed and simulated vertical profiles of potential temperature at KITcube observatory. Shaded areas indicate the spatial variability ($\pm 2\sigma$) of model simulations around (± 4 km) the measurement point (Source: Sulis *et al.*, 2018[6]).

COSMO.[19] To this aim, mesoscale simulation experiments with three configurations of COSMO embedded in TSMP are carried out for a two month period over a domain located in western Germany. The configurations include the operational standalone COSMO model (TSMP-CSC), COSMO coupled with a more complex land surface parameterisation linking carbon fluxes with photosynthesis and stomatal resistance and with a 1D hydrological model (TSMP-1D), and the latter configuration augmented with a 3D groundwater model (TSMP-3D). The evaluation is performed using a wide array of atmospheric, land surface energy balance, and subsurface observations during the High Definition Clouds

and Precipitation for advancing Climate Prediction (HD(CP)^2) Observational Prototype Experiment (HOPE) between April and May 2013.[20]

Fig. 1 shows the comparison of the vertical profiles of potential temperature for the three model configurations and observations for different times of the day during two intense observation periods (April 24 and May 19, 2013). On April 24 the simulated profiles are cooler than the observations especially in the TSMP-1D and -3D configurations by a few degrees between 09:00 UTC and 13:00 UTC. Around 15:00 UTC, when the land surface and the atmosphere are strongly coupled, TSMP-CSC matches the well-mixed observed temperature profile, while both TSMP-1D and -3D simulations show a cold bias of about 2 K. On May 19, the three configurations simulate the observed contrasting boundary layer evolution between dry and wet conditions, with the latter characterised by a slower temperature response and a shallower inversion between 09:00 and 13:00 UTC. At 15:00 UTC all models simulate similar vertical profiles. The contrasting response between the three model configurations can be explained as follows. Under relatively dry conditions (April 24, 2013), the warmer vertical profiles of atmospheric temperature, and hence higher ABL, simulated by TerrSysMP-CSC are controlled by the larger sensible heat contribution at the land surface, which is due to the higher soil temperature (results not shown). On the contrary, under relatively wet conditions (May 19, 2013), differences between the three model configurations are largely attenuated in the simulation of soil temperature, which again explains the similar energy partitioning at the land surface and diurnal evolution of the ABL height.

3 Grid Resolution Effects on Pathways of Evapotranspiration

Evaporation (E) and plant transpiration (T) contributes to the total terrestrial evapotranspiration (ET). In this study, we use the hydrological component of TSMP for a central European mid-latitude climate regime (Inde catchment, Germany) to examine the impact of the horizontal grid resolution on the model simulated water fluxes with a particular emphasis on the uncertainty of T/ET model estimates.[15] The simulations are conducted for a relatively wet and dry year at four different grid resolutions (120, 240, 480 and 960 m).

Coarsening of the grid resolution shifts the frequency distribution of groundwater table depth for both years towards shallower depths. For both years, T does not change much (not shown here) while E increases with the increase in available soil moisture for evaporation at the surface due to grid coarsening. This increases ET and thereby decreases T/ET ratio, with the amplitude depending upon the local vegetation cover and the dry/wet year (decrease in T/ET with grid-coarsening being higher for the wet year). Fig. 2 shows the scaling behaviour of T/ET for different vegetation types with coarsening of the grid resolution. For the wet year T/ET decreases between 10 to 14 % for crops and 10 to 11 % for trees; for the dry year, its decrease is around 10 % for the crops and 4 to 6 % for trees. In terms of domain average, T/ET decreases by around 5 % and 8 % for the dry and wet year respectively.[15]

4 ABL Schemes for Hyper-Resolution Runs

An ensemble of 32 idealised and 17 real data simulations were conducted with TSMP over diurnal scales. The study investigates convectively induced secondary circulations in

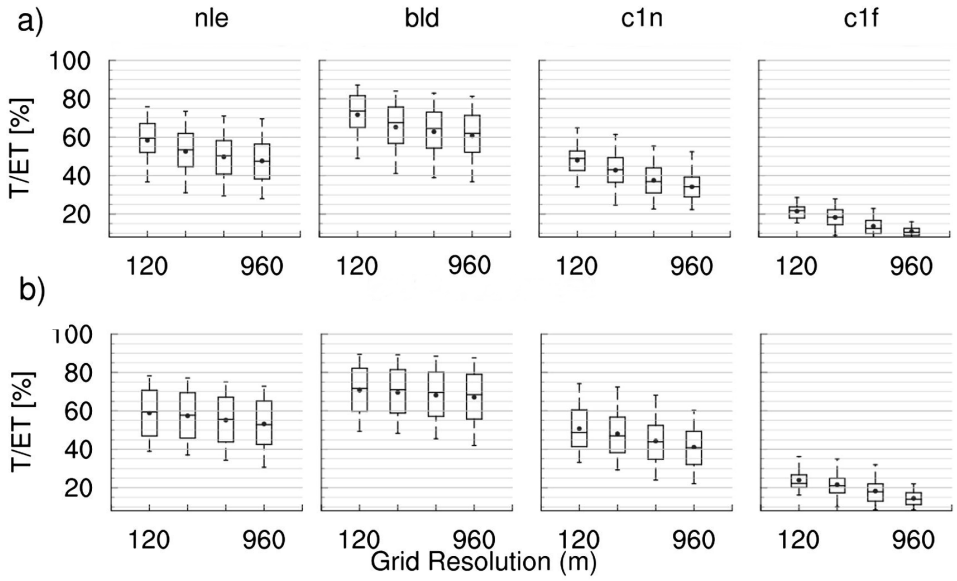


Figure 2. Scaling behaviour of T/ET with grid resolution for a) wet year (2009) and b) dry year (2011). The columns show the Plant Functional Types (PFTs) agricultural crops (c1f), agricultural crops with constant low LAI as substitute for urban areas (c1n), broadleaf deciduous tree (bld), and evergreen needle-leaf trees (nle). The box plot and the whiskers show the mean (solid markers), median, 5, 25, 75 and 95th percentile (Source: Shrestha *et al.*[15]).

response to varying grid spacings (Δx) within the turbulent grey zone (O(1 km)) for atmospheric boundary layer (ABL) schemes typically used in the numerical weather prediction.

The setup used large eddy simulation as benchmark runs for sensitivity simulations with varying turbulent mixing length scale and grid resolution, all located within the turbulent grey zone. The grid spacing of the land surface model, and thus its heterogeneity, is kept at $\Delta_x = 200$ m for all runs to remove any uncertainty in the simulated results due to land surface heterogeneity at different spatial scales.

The model run with the finest resolution $\Delta_x = 200$ m, $l_\infty = 70$ m is able to produce a vertical profile similar to the reference LES run (Fig. 3). While coarser simulations ($\Delta_x \geq 1000$ m) for the same asymptotic turbulent mixing lengthscale result in a stronger superadiabatic layer, as the model is not able to effectively remove heat from the surface by the ABL parameterisation. Due to the super-adibatic layer, artificial secondary circulation is triggered in the model. Increasing the asymptotic turbulent mixing lengthscale l_∞ to much larger values for the coarser grid spacings effectively removes the energy from the surface and attenuates the artificial secondary circulations.

Thus, the turbulent mixing length scales used in the ABL scheme can be tuned to suppress model generated convectively induced secondary circulations, while the non-resolved turbulence dealt with by the ABL scheme effectively propagates the surface fluxes into the ABL and sustains reasonable ABL profiles.

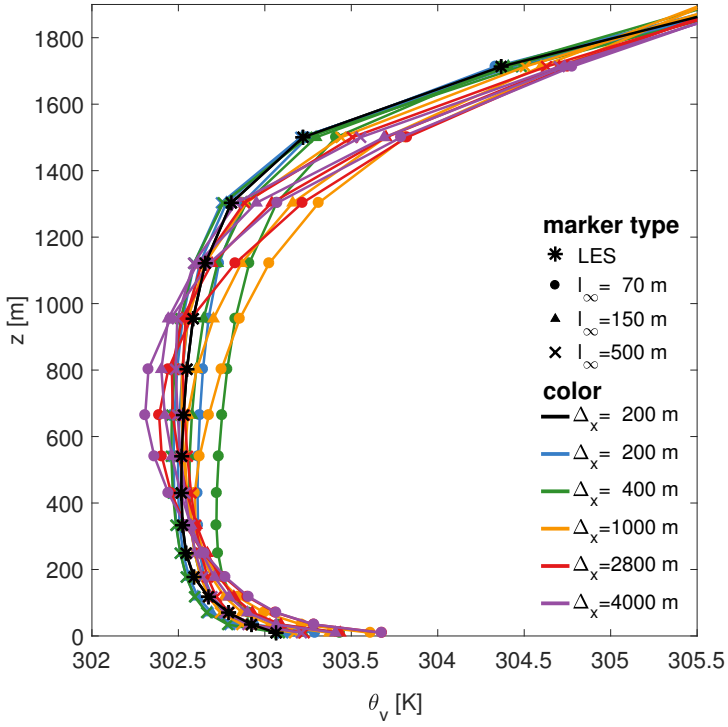


Figure 3. Domainaveraged vertical profile of virtual potential temperature on 25 July 2012 at 13:00 UTC for different grid spacings (Δ_x) and asymptotic turbulent mixing lengthscales (l_∞) (Source: Poll *et al.*[10]).

5 Weakly Coupled Data Assimilation (WCDA) with TerrSysMP-DART Interface

In a weakly coupled data assimilation experiment (WCDA), the models are tightly coupled and the DA only impacts the system that is directly responsible for the observations (*e. g.* assimilation of air temperature does not impact soil moisture directly or *vice versa*). WCDA is conducted for a semi-idealised setup with a horizontally homogeneous vegetated land surface and soil texture using the TerrSysMP-DART interface. DART is an open source ensemble data assimilation framework developed by the National Center for Atmospheric Research. DART provides data assimilation capability without changing the model code and already supports assimilation with COSMO and CLM. Assimilation support for ParFlow with DART was added along with software infrastructure for all component models required by TSMP.

The soil moisture is initialised horizontally homogeneous with a relative soil moisture content, $S_w = 0.11$ for the root zone, with ground water table at 3 m depth. Five types of simulations are performed with this idealised setup, which are summarised in Tab. 1. 48 ensemble members were generated by perturbing atmospheric boundary layer (ABL) temperature (below 850 hPa) and leaf area index (LAI) with a random uniform distribution.

1. Perfect Model Run to produce synthetic observations:	PM
2. Open Loop Run	OL
3. Weakly Coupled Data Assimilation with COSMO (ABL temperature)	$WCDA_{cos}$
4. Weakly Coupled Data Assimilation with CLM (Soil Temperature)	$WCDA_{clm}$
5. Weakly Coupled Data Assimilation with ParFlow (Soil Moisture)	$WCDA_{pfl}$

Table 1. Data Assimilation Experiments. OL and $WCDA$ runs are conducted with 48 ensemble members.

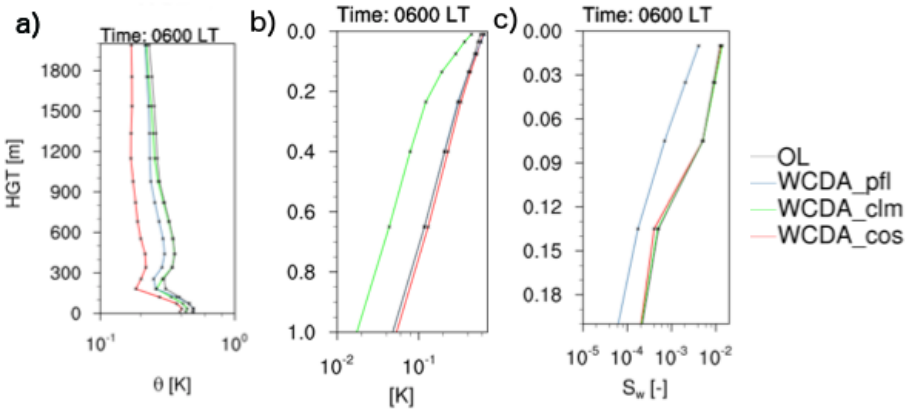


Figure 4. Diurnal evolution of the mean absolute error (MAE) of a) ABL potential temperature, b) soil temperature and c) soil moisture for the ensemble means. The vertical profiles at 0600 LT are temporally averaged for the entire period of simulation.

14 day simulations were conducted with DA at 00:00 UTC everyday for 10 different spatial locations at multiple heights, to examine its impact on diurnal scale evolution of the land surface states and the boundary layer.

Fig. 4 shows the time-averaged response of mean absolute error (MAE) for the ABL temperature, soil temperature and soil moisture with 6 hour lead forecast. Preliminary results indicate that the direct assimilation of boundary layer temperature results in the least MAE for $WCDA_{cos}$, whereas the assimilation of soil temperature and soil moisture also indirectly reduces the MAE for the $WCDA_{clm}$ and $WCDA_{pfl}$ relative to the open loop run. The direct assimilation of soil temperature and soil humidity also improves the MAE for $WCDA_{clm}$ and $WCDA_{pfl}$ respectively. But, the assimilation of soil moisture does not show any improvement in soil temperature and *vice versa*. Also, the assimilation of the ABL temperature shows no improvement in the surface soil moisture or temperature.

6 Concluding Remarks

While the inclusion of physically based groundwater in the integrated terrestrial systems model is essential to further improve our understanding of the terrestrial water cycle, the

increased complexity also introduces uncertainty associated with poor understanding of subsurface soil texture and grid scale dependence of simulated soil moisture. Such grid scale dependence was also found for boundary layer mixing parameterisations. Increased computational resources available at the Jülich Supercomputing Centre (JSC) will partly contribute to mitigate these issues by using hyper-resolutions, along with simultaneous efforts on optimising the convergence of groundwater model codes for larger domains. Besides, the model representation of vegetation with static plant physiological information also contributes to the uncertainty in the pathways of evapotranspiration. Ensemble simulations are thus a prerequisite to quantify the uncertainty in the terrestrial water cycle simulations, which then could also be utilised for data assimilation to improve prediction.

Acknowledgements

The study was conducted with support from SFB/TR32 (`www.tr32.de`) "Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modeling, and Data-Assimilation" funded by the Deutsche Forschungsgemeinschaft (DFG) and HD(CP)^2 - High definition clouds and precipitation for advancing climate prediction -research program funded by Ministry of Education and Research (BMBF). We also gratefully acknowledge the computing time (project HBN33) granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC) .

References

1. P. Shrestha, M. Sulis, M. Masbou, S. Kollet, and C. Simmer, *A Scale-Consistent Terrestrial Systems Modeling Platform Based on COSMO, CLM and ParFlow*, Mon. Weather. Rev. **142**, 3466–3483, 2014.
2. F. Gasper, K. Goergen, P. Shrestha, M. Sulis, J. Rihani, M. Geimer, and S. Kollet, *Implementation and scaling of the fully coupled Terrestrial Systems Modeling Platform (TerrSysMP v1.0) in a massively parallel supercomputing environment – a case study on JUQUEEN (IBM Blue Gene/Q)*, Geosci. Model Dev. **7**, 2531–2543, 2014.
3. M. Rahman, M. Sulis, and S. J. Kollet, *The subsurface-land surface-atmosphere connection under convective conditions*, Advances in Water Resour. **83**, 240–249, 2015.
4. J. Keune, F. Gasper, K. Goergen, A. Hense, P. Shrestha, M. Sulis, and S. Kollet, *Studying the influence of groundwater representations on land surface-atmosphere feedbacks during the European heat wave in 2003*, J. Geophys. Res. Atmos. **121**, 301–325, 2016.
5. M. Sulis, J. L. Williams, P. Shrestha, M. Diederich, C. Simmer, S. Kollet, and R. M. Maxwell, *Coupling groundwater, vegetation, and atmospheric processes: a comparison of two integrated models*, J. of Hydrometeorology **18**, 1489–1511, 2017.
6. M. Sulis, J. Keune, P. Shrestha, C. Simmer, and S. J. Kollet, *Quantifying the impact of subsurfaceland surface physical processes on the predictive skill of subseasonal mesoscale atmospheric simulations*, J. of Geophys. Res.: Atmos. **123**, 9131–9151, 2018.
7. M. Uebel, M. Herbst, and A. Bott, *Mesoscale simulations of atmospheric CO2 variations using a highresolution model system with processbased CO2 fluxes. Quarterly Journal of the Royal Meteorological Society*, QJRMS **143**, 1860–1876, 2017.

8. M. Sulis, M. Langensiepen, P. Shrestha, A. Schickling, C. Simmer, and S. J. Kollet, *Evaluating the influence of plant-specific physiological parameterizations on the partitioning of land surface energy fluxes*, J. of Hydrometeorol. **16**, 517–533, 2015.
9. P. Shrestha, M. Sulis, C. Simmer, and S. Kollet, *Impacts of grid resolution on surface energy fluxes simulated with an integrated surface-groundwater flow model*, HESS **19**, 4317–4326, 2015.
10. S. Poll, P. Shrestha, and C. Simmer, *Modelling convectively induced secondary circulations in the terra incognita with TerrSysMP*, QJRMS **143**, 2352-2361, 2017.
11. M. Uebel, and A. Bott, *Influence of complex terrain and anthropogenic emissions on atmospheric CO2 patterns – a highresolution numerical analysis*, QJRMS **144**, 34–47, 2018.
12. P. Shrestha, M. Sulis, C. Simmer, and S. Kollet, *Effects of horizontal grid resolution on evapotranspiration partitioning using TerrSysMP*, J. of Hydrol. **557**, 910–915, 2018.
13. S. Gebler, H.-J. H.Franssen, S. J. Kollet, W. Qu, and H. Vereecken, *High resolution modelling of soil moisture patterns with TerrSysMP: A comparison with sensor network data*, J. of Hydrol. **547**, 309–331, 2017.
14. W. Kurtz, G. He, S. J. Kollet, R. M. Maxwell, H. Vereecken, and H.-J. H. Franssen *TerrSysMP-PDAF (version 1.0): a modular high-performance data assimilation framework for an integrated land surface-subsurface model*, GMD **9**, 1341–1360, 2016.
15. P. Shrestha, W. Kurtz, G. Vogel, J. P. Schulz, M. Sulis, H. J. Hendricks Franssen, S. Kollet and C. Simmer, *Connection Between Root Zone Soil Moisture and Surface Energy Flux Partitioning Using Modeling, Observations, and Data Assimilation for a Temperate Grassland Site in Germany*, J. of Geophys. Res.: Biogeosciences **123**, 2839–2862, 2018.
16. J. Zhang, W. Kurtz, S. Kollet, H. Vereecken, and H. J. H. Franssen, *Comparison of different assimilation methodologies of groundwater levels to improve predictions of root zone soil moisture with an integrated terrestrial system model*, Adv. in Water Res. **111**, 224–238, 2018.
17. B. S. Naz, W. Kurtz, C. Montzka, W. Sharples, K. Goergen, J. Keune, H. Gao, A. Springer, H.-J. H. Franssen, and S. Kollet, *Improving soil moisture and runoff simulations at 3 km over Europe using land surface data assimilation*, HESS **23**, 277–301, 2019.
18. J. Anderson, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, *The data assimilation research testbed: A community facility*, Bull. Amer. Meteor. Soc. **90**, 1283–1296, 2009.
19. M. Baldauf, A. Seifert, J. Foerstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, *Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities*, Mon. Wea. Rev. **139**, 3887–3905, 2011.
20. A. Macke, P. Seifert, H. Baars, C. Beekmans, A. Behrendt, B. Bohn, J. Bühl, S. Crewell, T. Damian, H. Deneke, S. Düsing, A. Foth, P. Di Girolamo, E. Hammann, R. Heinze, A. Hirsikko, J. Kalisch, N. Kalthoff, S. Kinne, M. Kohler, U. Loehnert, B. L. Madhavan, V. Maurer, S. K. Muppa, J. Schween, I. Serikov, H. Siebert, C. Simmer, F. Spaeth, S. Steinke, K. Traeumner, S. Troemel, B. Wehner, A. Wieser, V. Wulfmeyer, and X. Xie, *The $HD(CP)^2$ Observational Prototype Experiment HOPe: An overview*, Atmos. Chem. Phys. **170**, 4887–4914, 2017.

Applications of Seismic Full-Waveform Inversion on Shallow-Seismic and Ultrasonic Data

Thomas Bohlen, Yudi Pan, and Jonas Müller

Geophysical Institute, Karlsruhe Institute of Technology, Karlsruhe, Germany
E-mail: thomas.bohlen@kit.edu

Conventional seismic imaging methods utilise a small portion of the information in the seismic data we obtain. Most methods analyse their arrival times or specific signal amplitudes only. In this report we further develop and apply a new seismic inversion and imaging technique that uses the full information content of the seismic recordings. Full waveform inversion (FWI) is an algorithm that accounts for the full seismic waveform. It iteratively retrieves multiparameter physical models of the material by numerically solving the wave equation and optimisation problem. FWI is currently a cutting-edge seismic inversion and imaging technique that enables to exploit the full information contained in the seismic waveforms over a broad range of frequencies and apertures for an improved estimation of physical parameters. It allows for a mapping of structures on spatial scales down to approximately half of the seismic wavelength, hence providing a tremendous improvement of resolution compared to travel-time tomography based on ray-theory. We especially focus on the applications of FWI on two different scales, including near-surface scale which extends from Earth's free surface to a depth around 10 meters and laboratory scale which targets on material from a few millimetre to less than one meter. Numerical tests and real-world applications are used to show the high resolution of FWI in reconstructing physical properties of the earth model and artificial material.

1 Introduction

The reconstruction of near-surface elastic-parameter models is of fundamental importance for near-surface geophysical and geotechnical studies. Surface waves dominate the shallow-seismic wavefield and are attractive for determining near-surface structures due to their relatively high signal-to-noise ratio in field recordings. With a rapid development in the theories of surface-wave methods, it has become increasingly popular over the last two decades to use surface waves as a non-invasive way to estimate near-surface structures. Those methods, however, only use part of the information contained in the seismic data and, therefore, lose resolution in the reconstructed models (Pan *et al.*[1]).

With a rapid increase in computational power, it has become feasible to use full-waveform inversion (Tarantola[2]) to resolve a subsurface model by fitting the observed waveform directly. Based on full-wavefield modelling, FWI is able to fully exploit the waveform information and is getting increasingly popular among shallow-seismic methods. Due to the existence of surface waves, the acoustic approximation, which is widely adopted in explorational seismics, is no longer valid in shallow seismics. The inclusion of surface waves in the wavefields also increases the nonlinearity of FWI. Shallow-seismic FWI is an ill-posed problem and could converge toward a local minimum especially when a poor initial model is provided. Therefore, special care should be paid when utilising FWI to resolve near-surface model (Groos *et al.*[3]).

Besides the application of FWI on shallow-seismic data, we also look at FWI in the context of non-destructive testing. The aim is to characterise small anomalies in building

materials with ultrasonic measurements to detect defects and inclusions which can significantly affect the integrity of building materials. Defects and intrusions in building materials can be build due to altering processes or an inadequate production. The advantage of ultrasonic measurements is, that they are non-destructive. Therefore, it can be used for almost every kind of testing of building materials, even for vulnerable structures like historic buildings. Thus, the results are of high practical interest. The project is done in cooperation with the Fraunhofer Institute for Nondestructive Testing (IZFP) in Saarbrücken.

The aim of full waveform inversion is the estimation of elastic material parameters and their composition to a subsurface model out of seismic observations. For most subsurface configurations there do not exist exact analytical solutions, so only approximated solutions can be found using numerical methods. Therefore, we iteratively minimise the misfit between the modelled and the observed data using time-domain staggered-grid finite differences. The model with the smallest misfit then describes the observed data best.

The big advantage of using FWI is, that it uses the whole information content of the observed data. In contrast, most methods only use the arrival time or specific signal amplitudes of the observations. Consequently, we have a large improvement in model resolution by using FWI compared to other methods such as traveltime tomography.

2 Methodology

2.1 Forward Modelling

Seismic modelling is the fundamental part of FWI and requires nearly all the computation time. In dependence of the field of application, the wave-propagation physics for an underlying subsurface model has to be described by an appropriate wave equation. On the one hand, this comprises the acoustic wave equation in this work. On the other hand, the problem has to be solved for two-dimensional or three-dimensional subsurface models (corresponding applications are referred to as 2D or 3D FWI). The numerical implementation of the wave equations consists of a time-domain finite-difference (FD) time-stepping method in the Cartesian coordinates. In this report, we use the FD scheme to solve the velocity-stress formulation of visco-elastic wave equation (Bohlen[4]). Due to finite model sizes, the wave equations are expanded by a perfectly matched layer to avoid artificial boundary reflections.

2.2 Inversion

Fig. 1 shows the general workflow of FWI. The solution of the inverse problem comprises several steps. The method is initialised by the choice of a 2D or 3D initial parameter model. Seismic velocities and/or mass density are assigned to the model m_0 at the first iteration. The initial model can be estimated from a prior information or computed by conventional methods. For each source of the acquisition geometry, seismic modelling is applied, *i. e.* the wavefield is emitted by the source and forward-propagates across the medium. A time series of this wavefield has to be stored in memory with respect to the whole volume. Synthetic seismic data is recorded at the receivers and the difference between the observed and the synthetic data is calculated, which results in residuals. For each source, the residual wavefield is backpropagated from the receivers to the source position.

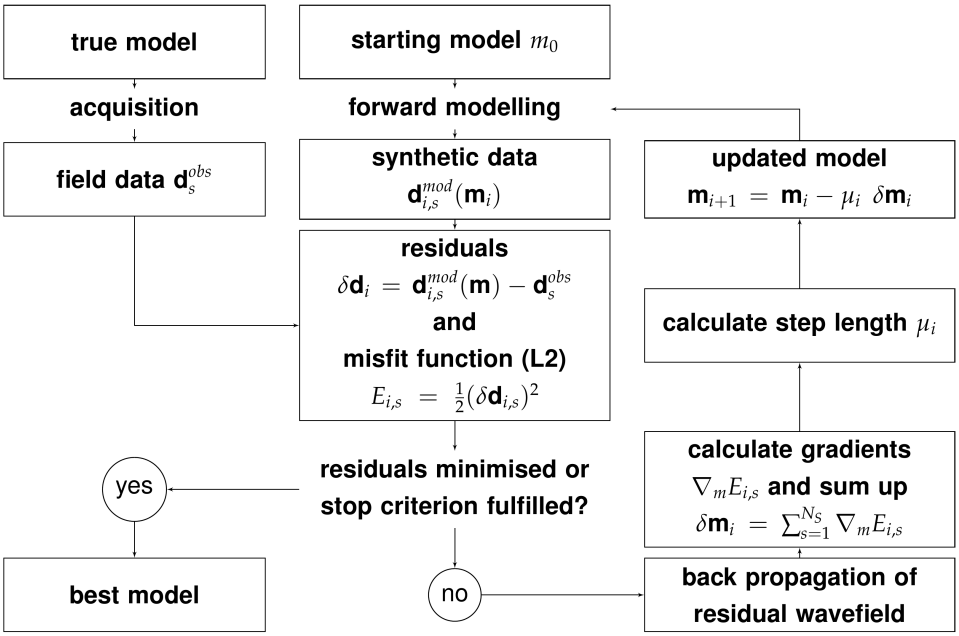


Figure 1. General workflow of FWI.

The cross-correlation of forward- and backpropagated wavefields yields shot-specific gradients. The computation of the global gradient δm for the entire acquisition geometry is given by the summation of all gradients $\delta m_{i,s}$. An optimised gradient δm_i is computed by subsequent preconditioning and application of the optimisation algorithm (*e. g.* conjugate gradient, L-BFGS). The update of the model parameter is the final step of an FWI iteration. The gradient δm_i has to be scaled by an optimal step length μ_i to get a proper model update at iteration step i. The estimation of μ_i is performed at each iteration with a line search algorithm and requires additional modellings.

3 Shallow-Seismic FWI

3.1 Synthetic Test

A checkerboard model is used to test the resolution of FWI. A homogeneous half-space model with an S-wave velocity of 200 m/s, a P-wave velocity of 500 m/s, and a density of 2000 kg/m^3 is used as the background model. The checkerboard is formed by blocks of 1 m $\times$ 1 m, and the S-wave velocity of each block is 220 or 180 m/s depending on its position (Fig. 2), *i. e.* $\pm$10 % perturbation in S-wave velocity is added to the background model. Twenty-four two-component geophones are placed along the free surface with an interval of 1 m (red dashed line in Fig. 2a). Six shot gathers are simulated with a vertical-force source spaced every 6 m (red stars in Fig. 2a). The nearest source-receiver offset for the leftmost source is 3 m. A shifted Ricker wavelet with a centre frequency of 40 Hz

is used as the source wavelet. The minimal wavelength of the observed surface wave is around 2 m, which is twice the size of the checkerboard block.

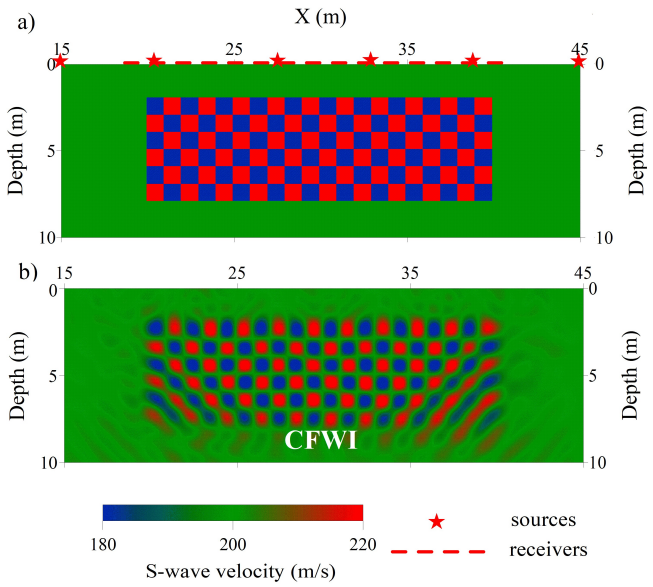


Figure 2. Checkerboard model test. a) True S-wave velocity model. Stars represent the locations of the sources, and the dashed line represents the geophone spread. b) Inversion results of Conventional FWI. The background without the checkerboard is used as the initial model.

We use the background model as the initial model and apply FWI on the observed data. The true source wavelet is used and only the S-wave velocity model is updated during the inversion. The conjugate gradient algorithm is used as the optimisation algorithm, and the gradient is calculated by the adjoint state algorithm. The checkerboard model can be reconstructed using FWI (Fig. 2b), which proves a relatively high resolution (approximately half a wavelength) of FWI. The lower left and right corners in the checkerboard are not reconstructed as well as the other areas, mainly due to a lower illumination in the lower corners.

3.2 Field Data Test

The field data set was acquired in Rheinstetten, Germany. The near-surface material at the test site was mainly composed of layered fluviatile sediments. A vertical source and forty-eight vertical-component geophones were placed along the survey line with a 1 m spacing (triangles in Fig. 3a). The source was placed every 4 m, and the first source position was located between the first and the second geophones. A total of 12 shot gathers are used in this example (asterisks in Fig. 3a).

We built a 1D depth-dependent S-wave velocity model as the initial model for FWI by using the conventional surface-wave method (Fig. 3a). The initial P-wave velocity was

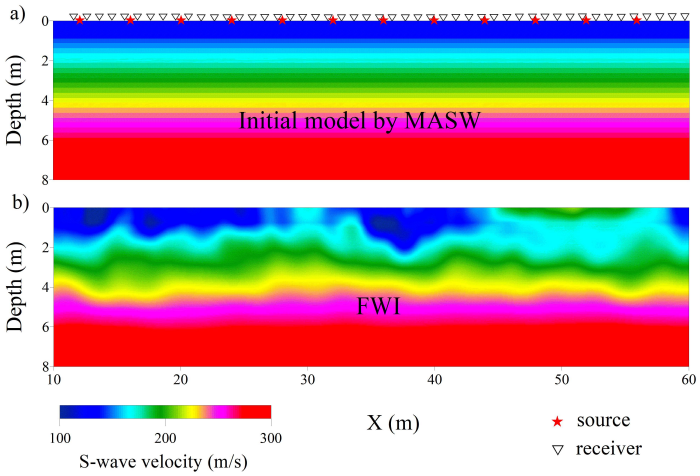


Figure 3. S-wave velocity model for the Rheinstetten data. a), b) and c) represent the initial model estimated by 1D MASW, the inversion result of MFWI after 52 iterations, and the pseudo-2D MASW result obtained by combining 16 1D profiles, respectively. Asterisks and inverted triangles represent the locations of the sources and the receivers, respectively.

built by inverting traveltimes of the first arrivals. We applied a 3D-to-2D transformation to the observed data since we used a 2D forward solver. We delayed the whole shot gathers by 0.03 s and killed the traces with a source-receiver offset shorter than 2 m since the signals in those traces are partly clipped. Since the acquired data contain relatively strong attenuation effects, we used constant Q models as passive information with a viscoelastic forward solver to mitigate the influence of viscosity. We started FWI by inverting a subset of data up to 10 Hz. The upper frequency limit of the low-pass filter was progressively increased to 60 Hz with a 5 Hz interval, with a minimum of 3 iterations in each stage. The inversion converged after a total of 52 iterations.

The inverted S-wave velocity model shows some 2D lateral variations of the subsurface (Fig. 3b). A ”V”-shaped low-velocity body can be identified at the central shallow part of the model, which corresponds to a refilled trench at the test site. The inversion result also shows a velocity inversion (high-velocity layer on top of a low-velocity layer) around 50 m and 0.2 m depth. It is worth mentioning that the low-velocity body is also identified at the same location in the reconstructed P-wave velocity model; however, its ”V” shape is less clear in the P-wave velocity model compared to the S-wave velocity model.

4 Ultrasonic FWI

4.1 Information about the Software

We develop an FWI software which is freely available under the GNU Affero General Public License. To parallelise the inverse problem, it is reformulated to a matrix-vector for-

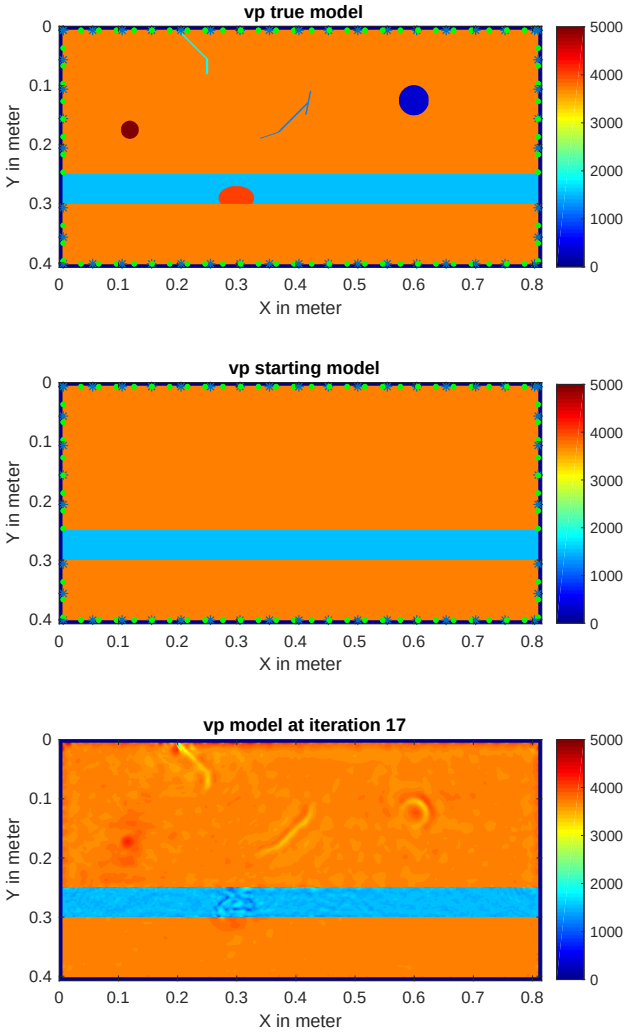


Figure 4. 2D elastic full waveform inversion of synthetic 2D data, exemplary shown for the parameter P-wave velocity (vp). Top: True model, middle: starting model, bottom: inversion result.

malism. The different implementations use explicit time domain finite difference forward modelling for the wave field calculations. The different FWI implementations consider 2D and 3D wave propagation and allow the reconstruction of acoustic, elastic, viscoacoustic and viscoelastic medium properties.

4.2 Synthetic Test

We use a 2D elastic FWI in the time domain and apply it to measured ultrasonic data at a concrete block, that is specifically build for this case. Inside the concrete block is a bore,

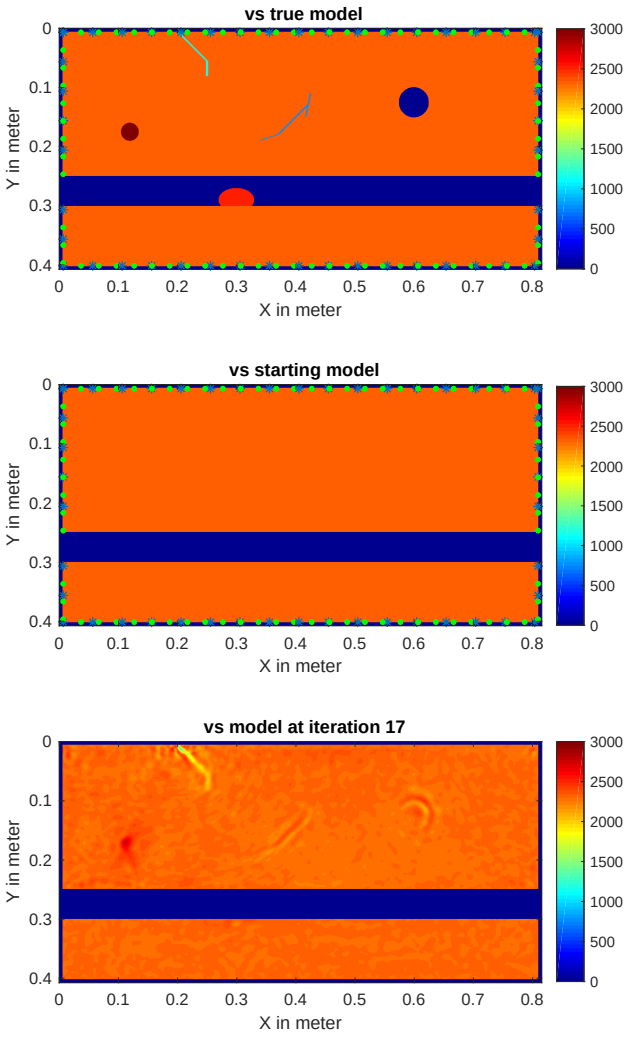


Figure 5. 2D elastic full waveform inversion of synthetic 2D data, exemplary shown for the parameter S-wave velocity (vs). Top: True model, middle: starting model, bottom: inversion result.

that can be filled with different materials with different properties. The block as well as the data acquisition is provided by the Fraunhofer Institute in Saarbrücken. The project mainly focuses on the reconstruction of the shape and properties (such as velocity and density) of anomalies.

The first step is to calculate some 2D synthetic tests to approximate a suitable quantity and positions of sources and receivers. After that we will perform the ultrasonic measurements on the concrete block and invert the field data. The small wavelengths of ultrasound are necessary to dissolve also small-scale anomalies. In our case the ultrasonic transmitter has a peak frequency of 100 kHz. Numerical seismic modelling is the fundamental part of

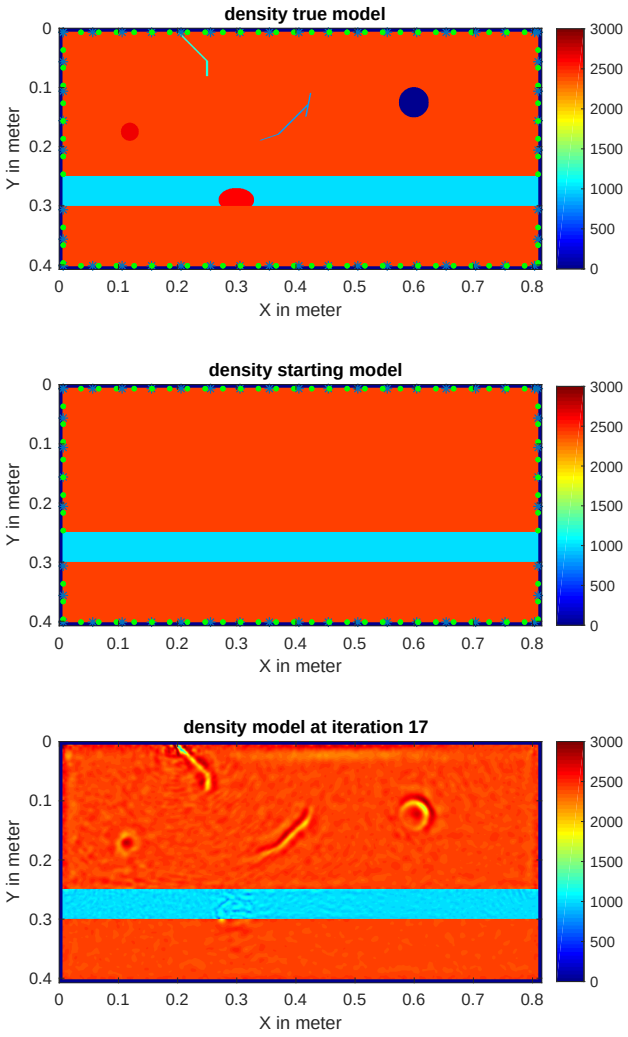


Figure 6. 2D elastic full waveform inversion of synthetic 2D data, exemplary shown for the density. Top: True model, middle: starting model, bottom: inversion result.

FWI, because usually there exist no exact analytical solution for a given subsurface configuration. The numerical modelling requires nearly all the computation time. The physics of wave propagation underlie the theory of continuum mechanics and can be described by an appropriate wave equation. The FD scheme solves the velocity-stress formulation of the elastic wave equation.

Figs. 4, 5 and 6 show the inversion results of 2D synthetic data for the P-wave velocity, S-wave velocity and density, respectively. The 2D synthetic data were obtained by forward modelling the wavefield in the true model (top model in Figs. 4, 5 and 6). The initial model contains a water-filled elongated pipe with a stone, two differently shaped cracks,

one high velocity anomaly and one air-filled low velocity anomaly. The starting model for the inversion is shown in the middle model in the figures. We follow a multi-scale approach to invert the data, that means we use multiple consecutive inversions containing different frequency ranges to prevent cycle skipping. We start with small frequencies creating a smooth model and go up step by step to higher frequencies. The final inversion result after 17 iterations in the last frequency range is shown in the bottom model of figures. As we can see, the shape of very small structures like cracks can be reconstructed quite good in all the three parameters. The inversion of the round high velocity anomaly on the left converges to the true velocity in the initial model. In contrast, the velocity of the air filled round low velocity anomaly on the right cannot be obtained correctly, as elastic waves do not propagate in air. Overall, it shows that FWI has a high resolution in reconstructing multiparameter physical models of artificial material and is promising to be used in ultrasonic testing in the future.

5 Concluding Remarks

We have applied full-waveform inversion to both shallow-seismic and ultrasonic data. Synthetic examples prove the high-resolution of FWI in reconstructing physical properties of the earth model and the artificial material. One real-world application proves the applicability of FWI. These results encourage us to use FWI in both shallow-seismic surveys and ultrasonic testings in the future.

Acknowledgements

We thank the NIC for access to the supercomputers JURECA and JUWELS at Jülich Supercomputing Centre (JSC).

References

1. Y. Pan, L. Gao, and T. Bohlen, *High-resolution characterization of near-surface structures by surface-wave inversions: from dispersion curve to full waveform*, Surveys in Geophysics **40**, 167–195, 2019.
2. A. Tarantola, *A strategy for nonlinear elastic inversion of seismic reflection data*, Geophysics **51**, 1893–1903, 1986.
3. L. Groos, M. Schäfer, T. Forbriger, and T. Bohlen, *Application of a complete workflow for 2D elastic full-waveform inversion to recorded shallow-seismic Rayleigh waves*, Geophysics **82**, R109–R117, 2017.
4. T. Bohlen, *Parallel 3-D viscoelastic finite difference seismic modelling*, Computers & Geosciences **28**, 887–899, 2002.

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 section we have two contributions:

Massively Parallel Multigrid with Direct Coarse Grid Solvers
by *M. Huber, N. Kohl, P. Leleux, U. Rüde, D. Thönnes, B. Wohlmuth*

and

BEAM-ME: Accelerating Linear Energy Systems Models by a Massively Parallel Interior Point Method
by *T. Breuer, M. Bussieck, K.-K. Cao, F. Fiand, H. C. Gils, A. Gleixner, D. Khabi, N. Kempke, T. Koch, D. Rehfeldt, M. Wetzel*

The first one concerns multigrid (MG) methods on parallel computers. In many cases, MG methods are asymptotically optimal and fast in the sense that their complexity grows only linearly with the number of unknowns. Nevertheless the efficient implementation of MG methods on parallel computers is a big challenge. It may happen that for MG solvers the parallel efficiency decreases when coarser grid levels are processed. To alleviate this trend, better load-balancing and a redistribution of the grids onto a subset of the processes, may become necessary to improve the performance. To reduce the overhead of the coarse grid solution phase, the authors develop an agglomeration algorithm such that the coarse grid solution can be executed on fewer processors. This reduces communication overhead and helps to improve the scaling behaviour for very large computations. In this contribution the authors consider a geophysical problem, which can be described by the Stokes problem on the spherical shell. In a previous work, the authors have compared different types of solvers and have found that a monolithic multigrid method with Uzawa type smoother combined for velocity and pressure achieves the best in time-to-solution and has the lowest memory requirements. It was also shown that the multigrid method delivers an excellent fine grid performance but that the used coarse grid Krylov type solver inhibits the overall parallel efficiency for very large scale applications.

This work was partly supported by the German Research Foundation through a Priority Programme.

The second contribution concerns Energy System Models (ESM). ESM are key to enable studies regarding policy design, analysis of technology pathways, or the analysis of future energy systems. One of the most successful modelling and solution methods for modern ESMs is linear programming (LP). LPs can be used to formulate and compute optimal states of an energy system and its extensions.

Many LPs, that were considered intractable two or three decades ago, can now be solved within seconds. Still, for large-scale LPs with hundreds of millions of variables and constraints such as those arising from high-resolution energy system models, even the best commercial solvers can take prohibitively long to find an optimal solution. On the other hand, the increasing availability of distributed-memory parallel computers offers a huge potential both for reducing solution time and avoiding memory bottlenecks. Unfortunately, general state-of-the-art LP solvers cannot (efficiently) run on distributed-memory systems. Therefore, one way forward is to develop more specialised algorithms for a distributed parallel solution, exploiting structures commonly found in ESMs.

For general LPs the two major algorithmic classes are simplex and interior-point methods. Interior-point methods are often more successful for large problems. While most of the steps in the algorithm can be parallelised efficiently, the formation of the actual Schur complement and the solution of the resulting linear system form a bottleneck and the authors of this contribution have put great effort into improving this part of the algorithm. Note that none of the commercial solvers allow for the distributed parallel solution of LPs. To achieve this goal on a broad scale, the block structure in the constraint matrix will help to improve the parallel efficiency.

The ELMOD ESMs, which are considered in this contribution, are used to analyse the impact of growing shares of renewable energies on the operation of the European transmission grid and the dispatch of conventional power plants.

The described research activities were funded by the German Federal Ministry for Economic Affairs and Energy (BMWi).

Massively Parallel Multigrid with Direct Coarse Grid Solvers

Markus Huber[1], **Nils Kohl**[2], **Philippe Leleux**[3], **Ulrich Rüde**[2,3], **Dominik Thönnes**[2], **and Barbara Wohlmuth**[1]

[1] Lehrstuhl für Numerische Mathematik, Technische Universität München, Garching, Germany
E-mail: {wohlmuth, huber}@ma.tum.de

[2] Lehrstuhl für Systemsimulation, Friedrich-Alexander Universität, Erlangen, Germany
E-mail: {ulrich.ruede, dominik.thoennes, nils.kohl}@fau.de

[3] Parallel Algorithms Team, CERFACS, Toulouse, France
E-mail: leleux@cerfacs.fr

Multigrid methods play an important role in the numerical approximation of partial differential equations. As long as only a moderate number of processors is used, many alternatives can be used as solver for the coarsest grid. However, when the number of processors increases, then standard coarsening will stop while the problem is still large and the communication overhead for solving the corresponding coarsest grid problem may dominate. In this case, the coarsest grid must be agglomerated to only a subset of the processors. This article studies the use of sparse direct methods for solving the coarsest grid problem as it arises in a multigrid hierarchy. We use as test case a Stokes-type model and solve algebraic saddle point systems with up to $\mathcal{O}(10^{11})$ degrees of freedom on a current peta-scale supercomputer. We compare the sparse direct solver with a preconditioned minimal residual iteration and show that the sparse direct method can exhibit better parallel efficiency.

1 Introduction

The numerical approximation of partial differential equations as obtained from models in science or engineering results in large sparse systems of algebraic equations. The solution of such systems requires techniques that can be efficiently executed on high-performance computing systems. In many cases, multigrid (MG) methods are asymptotically optimal and fast in the sense that their complexity grows only linearly with the number of unknowns and the constant of proportionality is only moderate.

However, the scaling of MG methods on modern supercomputers presents challenges and requires well-designed software structures and a performance aware implementation. In MG solvers it can be problematic that the parallel efficiency deteriorates when coarser grid levels are processed. When proceeding to successively coarser grids both computation and communication decrease, however, computation decreases faster than communication. As a consequence the communication overhead on coarser grids is larger despite the fact that the coarser grids demand progressively less work. To alleviate this trend, better load-balancing and a redistribution of the grids onto a subset of the processes, may become necessary to improve the performance.[1] The process of collecting the grid data and the associated linear system to a smaller number of processes is called agglomeration.

In this article, we apply heuristic strategies in combination with the hierarchical hybrids grids (HHG)[2] framework that realises a hybrid meshing approach. The scalability of the geometric MG solver (GMG) in the HHG framework was analysed and shown capable

to solve problems with more than 10^{13} degrees of freedom (DOF),[3] *i. e.* a problem with a solution vector that is 80 TByte large. It was also shown that the multigrid method delivers an excellent fine grid performance but that the used coarse grid Krylov type solver inhibits the overall parallel efficiency for very large scale applications. In this article we will present agglomeration strategies in combination with a direct coarse grid solver to compensate for this shortcoming.

The rest of this article is structured as follows: In Sec. 2, we introduce the model, the finite element discretisation and the general solver setup. Then, Sec. 3.1 details the geophysical setting and Sec. 3.2 analyses the deteriorating scalability of the standard coarse level solver of the HHG framework that is based on Krylov techniques. In Sec. 4.1, we describe the conversion routines from he HHG data-structures to standard sparse matrix data formats so that external solver frameworks can be employed. In Sec. 4.2, we introduce a master-slave agglomeration technique. The MUMPS framework is presented in Sec. 4.3. Sec. 5.1 considers MUMPS as standalone solver in several scaling experiments and Sec. 5.2 analyses the weak scaling behaviour of the combined method.

2 Model Problem, Discretisation and Solver

Let $\Omega \subset \mathbb{R}^3$ be an open and bounded domain. We consider the Stokes-type problem

$$\begin{aligned} -\operatorname{div}\left(\frac{\nu}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^\top)\right) + \nabla p &= \mathbf{f} \quad \text{in } \Omega \\ \operatorname{div}(\mathbf{u}) &= 0 \quad \text{in } \Omega \\ \mathbf{u} &= \mathbf{g} \quad \text{on } \partial\Omega \end{aligned} \tag{1}$$

where $\mathbf{u}$ denotes the velocity, p stands for the pressure, and $\mathbf{f}$ is a given forcing term. For simplicity of notation we restrict ourselves to Dirichlet boundary conditions $\mathbf{g}$ and the viscosity ν is assumed to be uniformly positive definite and scalar but has possibly large jumps. Problems of this structure can be found in many applications.

We use an initial tetrahedral mesh $\mathcal{T}_0$ and construct by uniform mesh refinement a hierarchy of meshes $\mathcal{T}_0 = \{\mathcal{T}_\ell, \ell = 0, \ldots, L\}$, $L > 0$ for the MG scheme, linked with linear interpolation for each component as prolongators and their adjoint as restrictions. For the moment we assume that Ω is resolved by the initial triangulation but generalisations are possible.[4] For the discretisation, we apply the lowest equal-order finite elements method stabilised with the pressure stabilisation Petrov-Galerkin (PSPG) technique.[5] Using (component-wise) nodal basis functions for velocity and pressure, we obtain a hierarchy of 2×2-block structured algebraic systems.

In a previous work,[3] we have compared different types of solvers and have found that a monolithic multigrid method with Uzawa type smoother combined for velocity and pressure achieves the best in time-to-solution and has the lowest memory requirements. In the following studies, we will use a geometric multigrid solver in form of a mildly variable V–cycle, *i. e.* we add for each coarser level 2 additional smoothing steps. This is denoted by V_{var}–cycle. The performance of the Uzawa smoother has also been theoretically and numerically investigated.[6]

3 Hierarchical Hybrid Grids (HHG)

Geometric multigrid methods are often considered as difficult to implement efficiently in parallel because of the reduced computational load on coarser grids. For our study the HHG framework[2] will be used. It provides parallel data structures and algorithms. In particular it features matrix-free concepts that enable the implementation of efficient parallel geometric multigrid methods.[7, 8]

3.1 Geophysical Setting

We study a problem that is motivated by geophysical simulations.[4] This leads to the Stokes problem Eq. 1 on the spherical shell $\Omega = \{\mathbf{x} \in \mathbb{R}^3 \colon 0.55 < \|\mathbf{x}\|_2 < 1\}$ and force term $\mathbf{f} = \mathrm{Ra}\,\tau \frac{x}{\|x\|}$, where $\mathrm{Ra} = 3.49649 \cdot 10^4$ is the dimensionless Rayleigh number and τ the normalised Earth's mantle temperature, as obtained from real-world measurements.[9] The mesh hierarchy has depth $L = 6$, plus 2 refinement levels used to get a structured coarse grid.

We consider either the iso-viscous case ($\nu(\mathbf{x}, T) \equiv 1$) or a viscosity profile given by lateral and radial variations:[10]

$$\nu(\mathbf{x}, T) = \exp\left(2.99 \frac{1 - \|\mathbf{x}\|_2}{1 - r_{\text{cmb}}} - 4.61T\right) \begin{cases} \frac{1}{10} \cdot 6.371^3 d_a^3 & \text{for } \|\mathbf{x}\|_2 > 1 - d_a \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

where d_a is the relative thickness of the asthenosphere. Thus, the Earth mantle is assumed to have layers with different viscosity characteristics. In particular, the asthenosphere, *i. e.* the outermost layer is assumed to be mechanically weaker. In the geophysics community, determining its size is still a field of active research.[4, 10] Here, we choose a thickness of 410 km for the asthenosphere and a corresponding viscosity jump by about a factor of 145. To close the system, we apply Dirichlet boundary conditions: on the surface we use plate velocity data,[11] and at the core-mantle boundary we apply no-slip conditions. The problem can now be solved by the monolithic Uzawa multigrid method by iterating until a residual reduction by five orders of magnitude has been reached. Note that for the simulations of systems subject to several types of error, *e. g.* model or measurement error, the specified tolerance is suitable to obtain a solution with an appropriate accuracy.

3.2 Standard Iterative HHG Coarse Level Solver

Initially, we employ the standard Krylov iterative method provided by the HHG package, *i. e.* a block-preconditioned minimal residual (*pminres*) iteration. This coarsest grid iteration is executed within each V-cycle with a stopping criterion enforcing that the preconditioned residual is reduced by three orders of magnitude.

The block-preconditioner here consists of a *velocity preconditioner*. For this the velocity block is approximately solved by a Jacobi-preconditioned conjugate gradient (PCG) method and we apply a lumped mass-matrix preconditioner for the pressure. The accuracy of the PCG method is specified by a relative residual reduction of two orders of magnitude. The use of this method is motivated by the fact that Krylov space methods can be easily implemented and parallelised. However, the error reduction depends on the condition number of the system matrix that deteriorates with the mesh size tending to zero. In

extreme scale computing, even the coarsest mesh can be large and have a small mesh size, when the number of processors is very large. Consequently, an increasing number of iterations become necessary to solve the problem on the coarsest grid with sufficient accuracy. Additionally, the number of iterations can depend sensitively on the viscosity variations.

We carry out our experiments on JUWELS a peta-scale supercomputers at the Jülich Supercomputing Centre ranked on position 30 of the TOP500[a] list (June 2019). The compute nodes are Intel Xeon Platinum 8168 processors with 2-sockets, each of 24 cores. The nodes have 96 GB memory and are linked with an Infiniband interconnection system.

In Tab. 1, we present the run-times (in seconds) averaged over the iterations of the MG scheme for the scenarios *iso-viscous* and the asthenosphere scenario with 410 km thickness, denoted by *jump-410*. We observe that the scalability is overall acceptable, but somewhat better in the iso-viscous case than for variable viscosity when the efficiency decreases to less than 75 %. For a more detailed analysis, we differentiate between the fine and the coarse grid average compute times. We observe then that the average run-times for the fine grids behave robustly across the scaling to large processor numbers, but the average timings for the coarse grid solver deteriorates. Note that this is expected, since we are using a sub-optimal algorithm to solve the system on the coarsest level.

The number of iterations required to reduce the residual by five orders of magnitude depends also on the shape of the elements in the triangulation of the input mesh. Thus, the iteration number is not constant when refining the mesh due to the viscosity variation and when unfavourable element shapes are created. For the iso-viscous case, we observe that 8 iterations are needed for a moderately sized problem and that even only 4 iterations are needed for the largest problems. For the jump-410 scenario the number of iterations is around 15 iterations in all cases. Using this as the baseline performance, we now turn to study an alternative robust coarse level solver.

scenarios			iso-viscous					jump-410				
proc.	DOF	DOF coarse	iter	total	fine	coarse	par. eff.	iter	total	fine	coarse	par. eff.
1 920	$2.1 \cdot 10^{10}$	$9.22 \cdot 10^{4}$	8	58.6	57.6	1.0	1.00	15	61.0	57.9	3.1	1.00
15 360	$4.3 \cdot 10^{10}$	$6.96 \cdot 10^{5}$	4	66.1	63.2	2.9	0.89	13	83.0	62.0	21.0	0.73
43 200	$1.7 \cdot 10^{11}$	$1.94 \cdot 10^{6}$	4	68.7	65.3	3.4	0.85	14	82.0	63.7	18.3	0.74

Table 1. Average time (in seconds) over the iterations of the mildly variable V–cycle: total, fine and coarse grid timings for iso-viscous and 410 asthenosphere scenario.

4 Coarse Level Strategies

To reduce the overhead of the coarse grid solution phase, we first develop an agglomeration algorithm such that the coarse grid solution can be executed on fewer processors. This reduces communication overhead and helps to improve the scaling behaviour for very large computations.

[a] `https://www.top500.org`

Our solution strategy on the coarse level consists of the following three steps:

1. Convert HHG format to sparse matrix data-format and agglomeration,
2. Solve the coarse level problem with an external library,
3. Redistribute and convert the approximation into HHG format.

4.1 Interfacing the Coarse Level Solver

One of the major advantages of the HHG framework is the highly-efficient data format combined with matrix-free techniques, which allows treating very large systems. On the other hand, these data-structures are not naturally compatible with software packages like HYPRE, MUMPS, or PETSc. An interoperability of HHG with these packages would be useful to benefit from the functionality, variability, and efficiency of such packages. To flexibly use these different software packages, the data from HHG must be converted in a classical sparse matrix data format using arrays for indices and values, *e. g.* the coordinate list (COO) format as used by the MUMPS solver. Essential for these formats is a unique global numbering of the DOFs across the processes.[2] We proceed by an ascending order per process and number first all DOFs of one process and then continue with the next one. By these identifications, the HHG matrices can be easily converted per process into array-like data structures.

4.2 Master-Slave Agglomeration

In a multigrid iteration the number of DOFs per process decreases when the algorithm proceeds to coarser grid levels. In the case of HHG, the load of the whole multigrid hierarchy is completely distributed based on the input mesh. Hence, the balance between computation and communication worsens on coarser grid levels. We will address this problem here by executing the coarse grid subroutines on fewer processes while letting the other processes remain idle.

To achieve this, we propose a *master-slave agglomeration*,[1] where data from several *slave* processes is accumulated to few *master* processes. Here we define the reduction factor $r \in \mathbb{N}_{\geq 1}$ as the size by which the number of computing processes $|\mathcal{P}|$ is divided such that we run on: $m = |\mathcal{P}|/r$ master processes with the accumulated data and the slave processes stay idle. We assume for simplicity that the reduction factor r is a divisor of $|\mathcal{P}|$. A case with $|\mathcal{P}| = 5$ and a reduction factor $r = 3$ and $m = 2$ is shown in Fig. 1. Once the

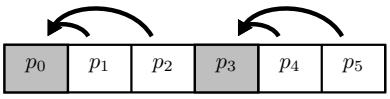


Figure 1. Master-slave agglomeration with reduction factor $r = 3$.

master processes have computed the results, they re-distribute them to the slave processes. The application to sparse matrix data formats involves array-like data-structures such as

C++ vectors. This simplifies the implementation of the agglomeration algorithm, since we only need to concatenate vectors.

Note that the selection of a suitable r is challenging and should respect the parallel architecture. In our case, we have chosen to agglomerate all the data from the same node. This makes it possible to perform the agglomeration with minimal communication overhead, but may then put extra communication burden on the parallel coarse grid solver. For our problems the agglomeration of the system matrix is performed only once, and is then kept in memory on the master processes. Note that for problems with evolving matrices, *e. g.* with time dependence or changing viscosity, the agglomerated matrix may have to be updated in each iteration.

4.3 MUMPS: A Parallel Sparse Direct Solver

MUMPS (MUltifrontal Massively Parallel direct Solver)[b] is a package for solving sparse systems of linear equations with symmetric (positive-definite or indefinite) or unsymmetric matrices. It is based on a direct method where the matrix is factorised into the product of triangular matrices which are then used to compute the solution using forward and backward substitution. The package is being developed by the MUMPS consortium. In the MG context, MUMPS has the advantage of being more robust than iterative methods, both in terms of run-time and accuracy. Also, with this solver the most consuming computations have to be performed only once for the whole MG scheme, when the factorisation of the matrix on the coarsest level can be reused.

MUMPS is a solver based on the multifrontal scheme where the factorisation of the input sparse matrix is achieved through a sequence of operations on relatively small dense matrices called "fronts".[12] Like most direct solvers, MUMPS achieves the solution of a system in three steps. Starting with a preprocessing followed by a symbolic factorisation (Analysis step), MUMPS then performs the actual factorisation of the fronts (Factorisation step). Most of the computational cost is spent in these 2 phases which are performed only once for a specific matrix. Finally, for each actual right hand side, the forward elimination and backward substitution operations are performed (Solve step).

Parallelism in MUMPS is implemented through a hybrid MPI/OpenMP model which makes the solver suited to modern distributed memory machines equipped with multicore processors.

5 Scaling Behaviour

In this section, we study the performance of the monolithic Uzawa multigrid method when a direct method is used as the coarsest level solver. The combined algorithm is implemented in the HHG framework, see Sec. 3, and the MUMPS library (version 5.1.2) is used for the coarsest grid problem, see Sec. 4.3. As example problem we study the discretised Stokes equation 1 with the settings described in Sec. 3.2. We use agglomeration techniques of 4.2 to run MUMPS on fewer processes. The computations are executed on the peta-scale supercomputer JUWELS. We start with a performance evaluation of the MUMPS solver in standalone mode for the coarsest level system of Sec. 5.1. This study is used to determine the best configuration to be used within the multigrid method of Sec. 5.2.

[b] `http://MUMPS-solver.org/`

5.1 Performance of the Direct Solver

In Tab. 2, we display results for a strong scaling study for different problem sizes of Sec. 3.2 and focus on the jump-410 problem that was found to be the most challenging in Sec. 3.2. Note that this strong scaling experiment is reversed in the sense that we start with the number of processes used for the finest grid problem such that the linear system is spread across all processors. In this mode, the coarsest grid computation suffers from too small granularity. Then we *reduce* the number of processes by an increasing reduction factor r, see equation in Sec. 4.2. We therefore find the best choice for the number of processes m that can be used for the coarsest grid problem, in order to minimise the run-time. Without agglomeration, the set of unknowns on each process is very small (less than 152 DOFs) leading to large communication overhead. This is clearly observed for the smallest problem for which the total timing with $r = 1$ is 30 times higher than with a reduction by $r = 48$. Thus using the reduced number of processes will be beneficial even if it will lead to idle processes when used within HHG. For a more detailed analysis, we also report the timings for analysis, factorisation, and solve phase separately.

To obtain minimal run-times for the problems with 1 920, 15 360 and 43 200 processes, the best choice is $r = 48$, $r = 96$ and $r = 192$, respectively. While all timings increase with the problem size, the solve step is always much faster than the factorisation and analysis, so that in cases where the factorisation can be re-used, this will lead to significantly better efficiency. Thus, when the cost of analysis and factorisation can be amortised over several multigrid iterations, the overall compute times will become competitive with using pminres as coarsest grid solver.

r	1920				15 360				43 200			
	proc.	ana.	fac.	solve	proc.	ana.	fac.	solve	proc.	ana.	fac.	solve
1	1 920	4.97	25.69	31.84	15 360	–	–	–	43 200	–	–	–
24	80	1.52	1.26	0.02	640	13.89	26.08	0.23	1 800	–	–	–
48	40	1.41	0.67	0.02	320	12.63	17.55	0.21	900	48.70	127.90	1.40
96	20	1.44	0.80	0.02	160	12.47	15.32	0.18	450	40.96	116.00	1.04
192	10	1.49	1.28	0.03	80	12.88	16.81	0.18	225	38.63	97.72	1.02
576	-	-	-	-	-	-	-	-	75	38.49	98.29	0.97

Table 2. Strong scaling study of the MUMPS sparse direct solver separated into analysis, factorisation and solve.

In a next set of experiments, we compare the efficiency of the direct solver for several configurations of the Stokes problem. In Tab. 3, we consider three different mesh resolutions of the problem for the iso-viscous and jump-410 scenarios. We use the agglomeration factors as previously obtained for each problem size. Different from the behaviour of pminres, we now observe that for a fixed problem size, the run-times remain stable for both scenarios. This is expected with an elimination based method whose run-time behaviour depends primarily only on the matrix nonzero structure. We finally observe that the direct solver produces results with an accuracy up to 10^{-18} in all cases, contrary to the the iterative pminres method. In practice, such an exact solution is not needed. Techniques to approximate the solution while decreasing the computational and memory complexity of

proc.	DOF coarse	scenario	analysis (s)	factorisation (s)	solve (s)	scaled residual
40	$9.22 \cdot 10^4$	iso-viscous	1.47	0.69	0.02	$1.8 \cdot 10^{-18}$
		jump-410	1.42	0.68	0.02	$1.9 \cdot 10^{-17}$
160	$6.96 \cdot 10^5$	iso-viscous	12.46	15.18	0.19	$5.7 \cdot 10^{-19}$
		jump-410	12.47	15.32	0.18	$1.2 \cdot 10^{-18}$
225	$1.94 \cdot 10^6$	iso-viscous	37.04	117.1	0.50	$5.31 \cdot 10^{-19}$
		jump-410	37.13	125.4	0.47	$1.27 \cdot 10^{-18}$

Table 3. Study of the influence of the viscosity scenario on the accuracy and run-time of the direct solver. Run-times are separated in analysis, factorisation and solve step. Each process runs on a separate node.

the method could be considered, such as single precision computation[13] or a suitable low rank matrix approximation.[14]

5.2 Performance Multigrid with Coarse Level Direct Solver

In this section, we apply the MUMPS sparse direct solver for the coarsest level problem in the HHG multigrid framework. We focus again on the jump-410 scenario using the two different coarsest level solvers, *i. e.* pminres from Sec. 3.2 and MUMPS as described in Sec. 4.3. In the following we will compare the average run-times of the V_{var}–cycle in a weak-scaling scenario.

In Tab. 4, we present the result of experiments with up to 43 200 processes. For a more detailed analysis, we display he fine grid and the coarse grid (*i. e.* MUMPS solve phase) average times separately for the iterations, and we sum the total time of analysis and factorisation for MUMPS. Additionally, we measure the total time for data transfer, that is, the time to perform the agglomeration and conversion between HHG and MUMPS data. According to the results from Tab. 2, we use the $r = 48, 96$ and 192, respectively, for agglomeration.

proc.	DOF		iter	time (s)					par. eff.
	fine	coarse		total	fine	coarse	ana. & fac.	trans.	
1 920	$2.1 \cdot 10^{10}$	$9.22 \cdot 10^4$	15	60.91	60.73	0.02	2.20	0.04	1.00
15 360	$4.3 \cdot 10^{10}$	$6.96 \cdot 10^5$	13	69.90	67.28	0.20	31.11	0.25	0.87
43 200	$1.7 \cdot 10^{11}$	$1.94 \cdot 10^6$	14	80.06	69.25	1.02	136.36	0.65	0.76

Table 4. Weak scaling of the V_{var}–cycle with a sparse direct and a simple Krylov coarse level solver. The run-times for total, fine and coarse are averages over the iterations. The run-times for analysis, factorisation and data transfer are the total timing.

In the scaling experiment we observe only a moderate increase in the average run-time of the fine grids, and the total time for the data transfer is neglectable in the overall timing. Moreover, the total coarse grid solves consume a fraction of less than 1.5 % of the overall time to solution for the biggest test case. However, the analysis and factorisation are much more costly, and thus these times may be a concern. Their timing grows up to $136.36\,s$

for the biggest test case which becomes $9.74\,s$ per iteration, when considering the whole solution process. For a fair comparison, this must be added to the $1.02\,s$ of the coarse level solve.

Using MUMPS decreases the times to process the coarsest grid by 40 % (resp. 80 %) as compared to the pminres solver for the largest (resp. the middle) test case for the scenario jump-410. Overall, the parallel efficiency of the solver has then been improved by 2 % points for the biggest case and 13 % points for the middle case problem.

Finally, we note that solving the coarsest grid problem very accurately with MUMPS has no effect on the multigrid convergence rate. This emphasises again the fact that such a high accuracy would not be needed and it would be beneficial to reduce the cost of the direct solver using appropriate approximation techniques.

6 Conclusion

Hierarchical hybrid grids multigrid iterative solvers in combination with an agglomeration of the processors on the coarse grid show an excellent parallel performance for large scale problems. Using a direct solver based on the MUMPS library results in an overall parallel efficiency of 76 %. As a large test case, a Stokes type system with up to $\mathcal{O}(10^{11})$ degrees of freedom has been solved with to $43\,200$ processes on the peta-scale supercomputer JUWELS.

Acknowledgements

The authors thank Patrick Amestoy, Jean-Yves L'excellent and Daniel Ruiz for their supporting discussion on efficient usage of the MUMPS framework. This work was partly supported by the German Research Foundation through the Priority Programme 1648 "Software for Exascale Computing" (SPPEXA) and WO671/11-1. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

References

1. D. A. May, P. Sanan, K. Rupp, M. G. Knepley, and B. F. Smith, *Extreme-Scale Multigrid Components within PETSc*, in Proceedings of the Platform for Advanced Scientific Computing Conference, PASC 2016, Lausanne, Switzerland, Article No. 5, 2016.
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. B. Gmeiner, M. Huber, L. John, U. Rüde, and B. Wohlmuth, *A quantitative performance study for Stokes solvers at the extreme scale*, J. Comput. Sci. **17**, 509–521, 2016.
4. S. Bauer, M. Huber, S. Ghelichkhan, M. Mohr, U. Rüde, and B. Wohlmuth, *Large-scale simulation of mantle convection based on a new matrix-free approach*, J. Comput. Sci. **31**, 60–76, 2019.

5. T. J. R. Hughes, L. P. Franca, and M. Balestra, *A New Finite Element Formulation for Computational Fluid Dynamics: V. Circumventing the Babuška-Brezzi Condition: A Stable Petrov-Galerkin Formulation of the stokes problem accommodating equal-order interpolations*, Comput. Methods Appl. Mech. Eng. **59**, 85–99, 1986.
6. W. Zulehner, *Analysis of Iterative Methods for Saddle Point Problems: A Unified Approach*, Math. Comput. **71**, 479–505, 2002.
7. F. Hülsemann, M. Kowarschik, M. Mohr, and U. Rüde, *Parallel Geometric Multigrid*, in Numerical Solution of Partial Differential Equations on Parallel Computers, A. M. Bruaset and A. Tveito (Editors), Springer, Lecture Notes in Computational Science and Engineering **51**, 165–208, 2005.
8. B. Gmeiner, U. Rüde, H. Stengel, C. Waluga, and B. Wohlmuth, *Towards textbook efficiency for parallel multigrid*, Numer. Math. Theory Methods Appl. **8**, 22–46, 2015.
9. N. A. Simmons, S. C. Myers, G. Johannesson, E. Matzel, and S. P. Grand, *Evidence for long-lived subduction of an ancient tectonic plate beneath the southern Indian Ocean*, Geophys. Res. Lett. **42**, 9270–9278, 2015.
10. D. Rhodri Davies, S. Goes, J. H. Davies, B. S. A. Schuberth, H.-P. Bunge, and J. Ritsema, *Reconciling dynamic and seismic models of Earth's lower mantle: The dominant role of thermal heterogeneity*, Earth and Planetary Science Letters **353-354**, 253–269, 2012.
11. R. D. Müller, M. Sdrolias, C. Gaina, and W. R. Roest, *Age, spreading rates, and spreading asymmetry of the world's ocean crust*, Geochem. Geophys. Geosyst. **9**, 1525–2027, 2008.
12. P. Amestoy, A. Buttari, J.-Y. L'Excellent, and T. Mary, *On the Complexity of the Block Low-Rank Multifrontal Factorization*, SIAM Journal on Scientific Computing **39**, A1710–A1740, 2017.
13. M. Baboulin, A. Buttari, J. Dongarra, J. Kurzak, J. Langou, J. Langou, P. Luszczek, and S. Tomov, *Accelerating scientific computations with mixed precision algorithms*, Computer Physics Communications **180**, 2526–2533, 2009.
14. M. Bebendorf and W. Hackbusch, *Existence of $\mathcal{H}$-matrix approximants to the inverse FE-matrix of elliptic operators with L_∞-coefficients*, Numerische Mathematik **95**, 1–28, 2003.

BEAM-ME: Accelerating Linear Energy Systems Models by a Massively Parallel Interior Point Method

Thomas Breuer[2], Michael Bussieck[3], Karl-Kiên Cao[4], Frederik Fiand[3], Hans Christian Gils[4], Ambros Gleixner[1], Dmitry Khabi[5], Nils Kempke[1], Thorsten Koch[1], Daniel Rehfeldt[1], and Manuel Wetzel[4]

[1] Zuse Institute Berlin/Technische Universität Berlin, Berlin, Germany
E-mail: {gleixner, kempke, koch, rehfeldt}@zib.de

[2] Jülich Supercomputing Centre (JSC), Forschungszentrum Jülich, Jülich, Germany
E-mail: t.breuer@fz-juelich.de

[3] GAMS Software GmbH, Frechen, Germany
E-mail: {mbussieck, ffiand}@gams.com

[4] Department of Energy Systems Analysis, Institute of Engineering Thermodynamics, German Aerospace Center (DLR), Stuttgart, Germany
E-mail: {karl-kien.cao, hans.gils, manuel.wetzel}@dlr.de

[5] High-Performance Computing Center Stuttgart (HLRS), Stuttgart, Germany
E-mail: khabi@hlrs.de

1 Introduction

The German energy revolution requires radical rethinking in business and science. A wide range of measures have been adopted by politicians in the course of the gradual switch to renewable energies: the last nuclear power plant is to be taken off the grid in 2022. By 2025, 40 to 45 percent of the electricity consumed nationwide will be generated from renewable energies. According to the coal commission's recommendation, the last coal-fired power plant should be shut down by 2038 at the latest. These and other changes will gradually lead to a fundamental restructuring of the energy system. Wind turbines and solar cells will be installed, power highways built and smart grids set up, energy infrastructures connected and new electricity storage technologies researched.

As concrete as this may sound, many of the factors in the planning game are uncertain. The development of the energy market is difficult to predict, and the climate outlook also raises questions. How will the sun shine in 50 years? When and where will the wind blow? Each eventuality that is fed into one of the analytical models for the energy systems of the future as a measurement, assumption or probability produces a more complex picture. This complexity is due to gigantic amounts of data that quickly push conventional computers and algorithms to the limits of their capacity and performance. The challenges of the energy system transformation are to a large extent challenges of information technology.

A widely used means in research and industry to analyse today's and future energy systems are Energy System Models (ESMs). Such models implement for instance the hourly dispatch of power stations, power transfer, and long-term investment decisions,

including sector coupling with transport and heating. As such, ESMs are key to enable studies regarding policy design, analysis of technology pathways, or the analysis of future energy systems. One of the most successful modelling and solution methods for modern ESMs is linear programming. Linear programs (LPs) can be used to formulate and compute cost-optimal states of an energy system and its extensions.

There has been tremendous progress in general LP solvers during the last 30 years.[1, 2] Many LPs that were considered intractable two or three decades ago can now be solved within seconds. Still, for large-scale LPs with hundreds of millions of variables and constraints such as those arising from high-resolution energy system models, even the best commercial solvers can take prohibitively long to find an optimal solution. Moreover, such large-scale problems might not even fit into the main memory of a modern desktop machine. On the other hand, the increasing availability of distributed-memory parallel computers offers a huge potential both for reducing solution time and avoiding memory bottlenecks. Unfortunately, general state-of-the-art LP solvers cannot (efficiently) run on distributed-memory systems. Therefore, one way forward is to develop more specialised algorithms for a distributed parallel solution, exploiting structures commonly found in ESMs.

Against this backdrop, the BEAM-ME project was initiated: an interdisciplinary research project carried out between 2015 and 2019 as part of the sixth energy research program of the German federal government (funding code: 03ET4023A-F). The participants were the Energy Systems Analysis Department of the German Aerospace Center, the Jülich Supercomputing Centre at Forschungszentrum Jülich, the High-Performance Computing Centre Stuttgart, the Mathematical Optimisation and Scientific Information Department at Zuse Institute Berlin, the Institute of Mathematics at Technische Universität Berlin and GAMS Software GmbH. This article describes the development and implementation of massively parallel algorithms for linear ESMs within the BEAM-ME project. Also, computational results for several large-scale ESMs will be presented.

2 Exhibiting Block Structure in Energy Systems Models

A typical structure observed in linear energy system models is the so-called *general arrowhead* or *doubly bordered block-diagonal* form:

$$
\begin{aligned}
\min \quad & \sum_{i=0}^{N} c_i^T x_i \\
\text{s.t.} \quad & T_0 x_0 & & & & & & = h_0 && (1)\\
& T_1 x_0 + W_1 x_1 & & & & & & = h_1 && (2)\\
& T_2 x_0 + & & W_2 x_2 & & & & = h_2 && (3)\\
& \vdots & & & \ddots & & \vdots & && \\
& T_N x_0 + & & & & & W_N x_N & = h_N && (4)\\
& U_0 x_0 + U_1 x_1 + & & U_2 x_2 & & \cdots & U_N x_N & = h_{N+1} && (5)\\
& x_0, & x_1, & & x_2, & \cdots & x_N & \geq 0 && (6)
\end{aligned}
$$

where $c_i, x_i \in \mathbb{R}^{n_i}, T_i \in \mathbb{R}^{m_i \times n_0}, U_i \in \mathbb{R}^{m_{N+1} \times n_i}$ for $i \in \{0, 1, \ldots, N\}$ and $W_i \in \mathbb{R}^{m_i \times n_i}$ for $i \in \{1, \ldots, N\}$, with $n_i, m_i \in \mathbb{N}_0$ for $i \in \{0, 1, \ldots, N\}$ and $m_{N+1} \in \mathbb{N}_0$. Note that also inequalities and variables bounds are allowed, but are not presented here for the sake of simplicity. This LP structure is quite general and can also be found in many applications beyond energy system modelling. Often, in an ESM each $i \in N$ corresponds to a point of time in a (discretised) time horizon. The *linking variables* x_0 may describe investment decisions. The *linking constraints* Eq. 5 usually include energy storage between two consecutive time steps.

For a typical ESM, the number of blocks that can be defined in the model depends on the motivation behind the research question. If, for example, a calendar year is described using time steps, a structure with 365 blocks based on the number of days can parallelise the solution process. Or exactly twice as many blocks are defined and thus days and nights are mapped separately. Not every block structure is equally suitable for the underlying algorithms. Therefore, there are different application variants. The modelling language GAMS was extended during the BEAM-ME project in such a way that the modeller can easily determine the block structure(s) of existing models by annotating the decision variables. Fig. 1 visualises several such annotations for small-scale problems.

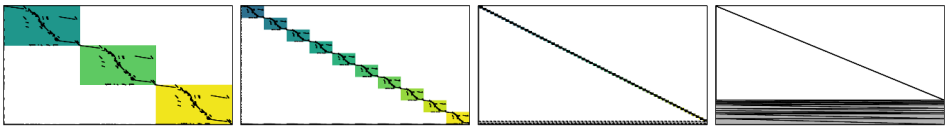


Figure 1. Non-zero plots of different block diagonal structures of a model instance. A challenge lies in finding an annotation that defines a sufficient number of blocks for parallelisation without the gray blocks of the linking variables and constraints becoming dominant.

3 A Parallel Interior Point Method for Linear Programs with General Arrowhead Form

For general LPs the two major algorithmic classes are simplex and interior-point methods.[3] Interior-point methods are often more successful for large problems, and they offer more potential for parallelisation since the main computational effort usually goes into factorising matrices. For a particular class of arrowhead LPs, two-stage stochastic optimisation problems, there exist already several distributed algorithms.[4–6] Two-stage stochastic optimisation problems do not have linking constraints and the diagonal blocks are intrinsically related. Both assumptions do not hold for typical ESMs. Moreover, a common feature of ESMs is a huge amount of linking constraints and variables (often more than 100 000). This large number of linking constraints and variables renders straightforward extensions of previous work on stochastic optimisation problems prohibitive.

To allow for the parallel solution of large-scale ESMs, we have developed and implemented several algorithms that are embedded within an interior-point method. Roughly speaking, an interior point method iteratively moves towards an optimal solution to the

given LP. The major part of the parallelisation takes place within the solution of the linear system, arising in each step of the interior-point algorithm. By suitable projection and permutation this linear system can be reduced to the following form:

$$\begin{bmatrix} K_1 & & & B_1 \\ & \ddots & & \vdots \\ & & K_N & B_N \\ B_1^T & \cdots & B_N^T & K_0 \end{bmatrix} \begin{bmatrix} \Delta z_1 \\ \vdots \\ \Delta z_N \\ \Delta z_0 \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_N \\ b_0 \end{bmatrix} \tag{7}$$

where

$$K_i = \begin{bmatrix} D_i & W_i^T \\ W_i & 0 \end{bmatrix}, \; K_0 = \begin{bmatrix} D_0 & T_0^T & U_0^T \\ T_0 & 0 & 0 \\ U_0 & 0 & 0 \end{bmatrix}, \; B_i = \begin{bmatrix} 0 & 0 & U_i^T \\ T_i & 0 & 0 \end{bmatrix} \tag{8}$$

The parallelisation is realised as specialised Schur complement decomposition, consisting of the following steps:

1. Multiply each row $i = 1, ..., N$ of Eq. 7 by $-B_i^T K_i^{-1}$.
2. Sum up all rows.
3. Solve $\left(K_0 - \sum_{i=1}^{N} B_i^T K_i^{-1} B_i\right) \Delta z_0 = b_0 - \sum_{i=1}^{N} B_i^T K_i^{-1} b_i$.
4. For each row $i = 1, ..., N$ insert Δz_0 and compute Δz_i.

The matrix

$$C = K_0 - \sum_{i=1}^{N} B_i^T K_i^{-1} B_i$$

is called the Schur complement of the above system. While most of the above steps can be parallelised efficiently, the formation of the actual Schur complement and the solution of the resulting linear system form a bottleneck and we have put great effort into improving this part of the algorithm.

The implementation builds on the existing solver PIPS-IPM,[7, 8] a solver for two-stage stochastic quadratic programs. The first step was the extension of this solver to handle linking constraints. Also, an interface to GAMS was implemented. However, a large number of further changes were necessary, to handle for example the strong linkage between individual blocks in typical ESM LPs, or the bad condition of these LPs. Also, considerable work had to be put into bringing the sequential run time closer to that of the leading commercial LP solvers (which have been developed over decades by the leading experts in the field). Most important among the newly implemented (and designed) methods are

- multiple-corrector interior-point algorithms,
- parallel LP scaling methods,
- parallel presolving methods,

- parallel preconditioning of the Schur complement.

More details on the preconditioned Schur complement approach can be found in Ref. 9. The parallel presolving algorithms are described in Ref. 10. Overall, more than 60 000 lines of PIPS-IPM code have been modified or added, which constitutes more than half of the entire active code base of the solver. Because of several requests from the energy system community, the new solver will be made publicly available at the end of 2019. In the remainder of this article, we will refer to the new solver as PIPS-IPM++, which is also the current working name.

4 Numerical Results

The numerical results reported in this section were obtained on the supercomputer JUWELS at the Jülich Supercomputer Centre. JUWELS is equipped with 2271 standard compute nodes (Dual Intel Xeon Platinum 8168), each with 96 GB memory and 2x24 cores at 2.7 GHz. JUWELS also includes 240 large memory compute nodes with 192 GB each (and otherwise the same configuration). The nodes are connected via a Mellanox EDR InfiniBand high-speed network. Only the standard compute nodes were used for PIPS-IPM++.

Within the BEAM-ME project modelling experts from GAMS together with energy system modellers from the German Aerospace Center developed a simplified, but representative, energy system model, called SIMPLE. While SIMPLE lacks many details that are considered in full-fledged energy system models, it is easily adaptable in size (*e. g.* the number of variables or diagonal blocks) and thus highly useful for testing new solution approaches. Fig. 2 shows the scaling behaviour of PIPS-IPM++ on a relatively small SIMPLE instance (5 144 806 constraints and 5 681 482 variables) on up to 512 MPI processes. Furthermore, the figure shows the scaling behaviour of a reduced version (denoted by *New solver reduced*) of PIPS-IPM++ that just uses the Schur complement decomposition described in Sec. 3, but none of the additionally implemented algorithms such as (distributed) preconditioned, iterative solution, LP scaling, and presolving. For each MPI configuration of PIPS-IPM++ (including the reduced version) two OpenMP threads were used. The figure also shows results for the leading four commercial LP solvers CPLEX 12.8[a], Gurobi 8.1[b], MOSEK 8.1[c], and Xpress 8.4.7[d] – which can only run on shared-memory. Fig. 2 shows that the commercial solvers using one thread are considerably faster than PIPS-IPM++ using one MPI process (and two OpenMP threads) – by a factor of 5 or more. However, the commercial solvers do not exhibit a good scaling behaviour. Three of the commercial solvers show a moderate speed-up up until 12 threads, but with more threads the run-time even deteriorates. PIPS-IPM++ on the other hand scales reasonably well, and requires merely 23 seconds with the maximum number of 512 MPI processes – whereas no commercial solver achieves a run-time below 400 seconds. Furthermore, the results of the reduced version of the new solver demonstrate that the straightforward application

[a] `https://www.ibm.com/products/ilog-cplex-optimization-studio`
[b] `https://www.gurobi.com`
[c] `https://www.mosek.com/`
[d] `https://www.fico.com/en/products/fico-xpress-optimization`

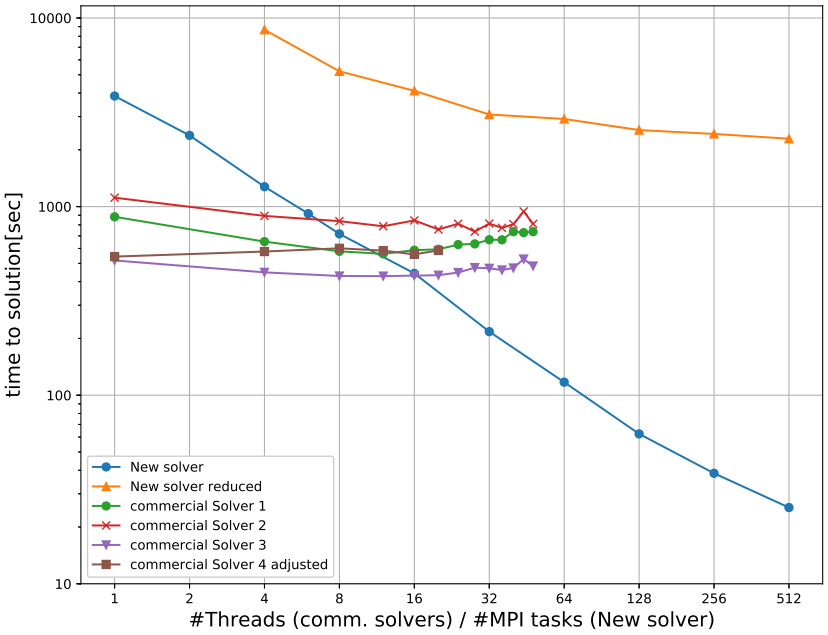


Figure 2. Scaling results of the leading commercial LP solvers (*Solver 1*,...,*Solver 4*) and PIPS-IPM++ (*new solver*) on a SIMPLE instance. PIPS-IPM++ was run with 2 OpenMP threads per MPI process.

of the Schur complement decomposition without any additional techniques is not competitive. Another notable observation is that PIPS-IPM++ is for this instance already on a shared-memory architecture highly competitive with the commercial solvers: On 32 cores (16 MPI processes, with two OpenMP threads each), PIPS-IPM++ is as fast as any commercial solver – even if one looks at their best performance among all number of threads used.

Further results are given in the following for the energy system models Premix[11] and ELMOD.[12] REMix has been developed (primarily) for the analysis of long-term energy scenarios with high shares of variable renewable energies. The ELMOD ESMs that we consider are used to analyse the impact of growing shares of renewable energies on the operation of the European transmission grid and the dispatch of conventional power plants. The first two ELMOD instances (CWE) comprise the entire transmission grid of the Central Western European region. For the other two instances (EU) this configuration was further extended to a set of 19 European countries with detailed transmission grid representation. This is the largest configuration currently possible for the ELMOD model and no instance of this configuration could previously be solved. For comparison we use the four state-of-the-art (commercial) LP solvers detailed above. We use 16 threads for each

of the commercial solvers – with more threads usually a performance degradation occurs. Note that none of the commercial solvers allow for the distributed parallel solution of LPs. Tab. 1 shows results on JUWELS of PIPS-IPM++ (with 2 OpenMP threads per MPI process) and the respective best commercial solver. ML signifies that the commercial solvers ran out of memory, even on the large nodes of JUWELS. As to the instances solvable on JUWELS, one notes that the run time per computing resources of PIPS-IPM++ is better than the best result of the commercial solvers (19 compute nodes versus 1 compute node).

	Size			PIPS-IPM++ resources		Run time (seconds)	
Instance	Variables	Constraints	Non-zeroes	MPI processes	Nodes	PIPS-IPM++	Best commercial
YSSP_exp	107 555 441	93 902 968	286 477 612	250	25	2311	ML
YSSP_disp	94 965 730	81 314 846	225 032 421	350	35	931	ML
ELMOD_CWE15	85 646 554	98 646 274	271 875 064	438	19	181	6 201
ELMOD_CWE16	85 883 074	98 909 074	272 602 144	438	19	216	6 111
ELMOD_EU15	224 677 686	254 304 961	712 452 541	876	38	1 245	ML
ELMOD_EU16	226 061 766	256 284 723	717 436 984	876	38	1 119	ML
SIMPLE_LARGE	227 060 381	206 036 266	818 449 005	1024	64	546	ML

Table 1. Computational results for large-scale instances.

5 Conclusion

The overall goal of the BEAM-ME project was to overcome the current performance limitations of linear energy systems models by exploiting high-performance computing resources. To achieve this goal on a broad scale, it was necessary to identify a common property that could be exploited algorithmically and is shared by a large majority of ESMs: block structure in the constraint matrix. The presence of this structure made it possible to develop a distributed-memory interior-point solver that showed good scalability for ESMs from different sources. This result stood out particularly in comparison to the poor scalability of today's state-of-the-art shared-memory solvers.

There are several paths for further development. One is the implementation of additional presolving methods. To improve robustness of the algorithm, we also plan to implement the homogeneous self-dual interior-point method.[13] Yet another extension, that is currently being implemented, is a hierarchical approach, which splits the Schur complement decomposition (and thus also the Schur complement) in several layers – with the aim to handle energy system models with even stronger linkage than the ones considered here.

Although HPC resources are certainly not ubiquitous, we hope that this generality and usability will help transfer the algorithmic progress achieved during the BEAM-ME project into the regular practice of the energy systems community.

Acknowledgements

The described research activities were funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) within the project BEAM-ME (grant number 03ET4023A-F). The authors gratefully acknowledge the Gauss Centre for Supercomputing

e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

References

1. R. E. Bixby, *Solving Real-World Linear Programs: A Decade and More of Progress*, Operations Research **50**, 3–15, 2002.
2. J. Gondzio, *Interior point methods 25 years later*, European Journal of Operational Research **218**, 587–601, 2012.
3. R. J. Vanderbei, *Linear Programming: Foundations and Extensions*, International Series in Operations Research & Management Science **196**, Springer, Boston, 2014.
4. J. Barnett, J.-P. Watson, and D. L. Woodruff, *BBPH: Using progressive hedging within branch and bound to solve multi-stage stochastic mixed integer programs*, Oper. Res. Lett. **45**, 34–39, 2017.
5. M. Colombo, J. Gondzio, and A. Grothey, *A warm-start approach for large-scale stochastic linear programs*, Math. Program. **127**, 371–397, 2011.
6. C. G. Petra, O. Schenk, and M. Anitescu, *Real-Time Stochastic Optimization of Complex Energy Systems on High-Performance Computers*, Computing in Science & Engineering **16**, 32–42, 2014.
7. M. Lubin, C. G. Petra, M. Anitescu, and V. Zavala, *Scalable stochastic optimization of complex energy systems*, in Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis (SC'11), Article No. 64, 2011.
8. C. G. Petra, O. Schenk, M. Lubin, and K. Gärtner, *An Augmented Incomplete Factorization Approach for Computing the Schur Complement in Stochastic Optimization*, SIAM J. Scientific Computing **36**, C139–C162, 2014.
9. D. Rehfeldt, H. Hobbie, D. Schönheit, A. Gleixner, T. Koch, and D. Möst, *A massively parallel interior-point solver for linear energy system models with block structure*, Zuse Institute Berlin (ZIB) Rep. **19-41**, 2019.
10. A. Gleixner, N.-C. Kempke, T. Koch, D. Rehfeldt, and S. Uslu, *First Experiments with Structure-Aware Presolving for a Parallel Interior-Point Method*, in Operations Research Proceedings, under review, 2019.
11. K. Cao, J. Metzdorf, and S. Birbalta, *Incorporating Power Transmission Bottlenecks into Aggregated Energy System Models*, Sustainability **10**, 1–32, 2018.
12. F. Hinz and D. Möst, *Techno-Economic Evaluation of 110 kV Grid Reactive Power Support for the Transmission Grid*, IEEE Transactions on Power Systems **33**, 4809–4818, 2018.
13. Robert J. Vanderbei, *The Homogeneous Self-Dual Method*, in Linear Programming, International Series in Operations Research & Management Science **114**, Springer, Boston, 361–381, 2008.

Fluid Mechanics and Engineering

Fluid Mechanics and Engineering

Christian Stemmer

Lehrstuhl für Aerodynamik und Strömungsmechanik, Technische Universität München,
Boltzmannstraße 15, 85748 Garching, Germany
E-mail: Christian.Stemmer@tum.de

Several outstanding research topics from fundamental fluid dynamics research processed at the Jülich supercomputing facility are presented in this section. Present on HPC facilities for many years, the unsolved problem of turbulence and its description in a lower dimensional space draws the attention of many research groups. The topics in this volume span from turbulence in Rayleigh-Bénard-convection, scale-adaptive turbulence modelling in centrifugal pumps, deep learning in the frame of sub-grid scale modelling for Large-Eddy Simulations in reactive flows to the simulations in the post-processing of experimental data to identify statistical properties of a large number of nanoparticles. All projects in common, the HPC-aspect is stressed and described, enabling the projects in the first place. Highly efficient methods are programmed and verified to attain the highest level of accuracy possible and necessary to describe the critical details of the wide variety of the presented flow regimes.

A cooperative project presented by Robert Kräuter, Dmitry Krasnov, Ambrish Pandey, Christiane Schneide, Kathrin Padberg-Gehle, Dimitrios Giannakis, Katepalli R. Sreenivasan, and Jörg Schumacher from TU Illmenau, Leuphana University Lüneburg, and the New York City University addresses a new approach on HPC-platforms with the help of machine learning (ML) algorithms. Turbulent flow data from Rayleigh-Bénard-convection cases are analysed via ML algorithms identifying potential to reduce data complexity.

Investigations on scale-adaptive turbulence models is in the focus of the work presented by Markus Hundshagen, Nicolas Casimir, Andreas Pesch, and Romuald Skoda from the Ruhr University Bochum. An improved method to include turbulence in the simulation of centrifugal pumps is presented based on pre-existing knowledge of turbulence in the modelling stage.

A working group at the combustion institute at the RWTH Aachen University (Mathis Bode, Dominik Denker, Jenia Jitsev, and Heinz Pitsch) presents a large-scale simulation exceeding 1 billion grid points on turbulent reactive planar temporally non-premixed jets. Employing deep-learning algorithms, the predictive capabilities of mixing statistics for under-resolved flow regions are explored and the additions effort measured. DNS data for verification purposes are generated and applied.

Maximilian J. Uttinger, Simon E. Wawra, Johannes Walter, and Wolfgang Peukert from the FAU in Erlangen contribute a data-analysis project where post-processing on HPC-systems provide a much more accurate insight into the particle densities, sizes, *etc.* of polysperse nanoparticle distribution. The high-level post-processing algorithms enhance the measurement accuracy considerably contributing to a steadily increasing measurement capability for small-scales in flows down to the nano-scale.

Machine Learning Applications in Convective Turbulence

Robert Kräuter[1], Dmitry Krasnov[1], Ambrish Pandey[1], Christiane Schneide[2], Kathrin Padberg-Gehle[2], Dimitrios Giannakis[3], Katepalli R. Sreenivasan[3,4], and Jörg Schumacher[1,4]

[1] Institut für Thermo- und Fluiddynamik,
Technische Universität Ilmenau, 98684 Ilmenau, Germany
E-mail: {ambrish.pandey, robert.kraeuter, dmitry.krasnov, joerg.schumacher}@tu-ilmenau.de

[2] Institut für Mathematik und ihre Didaktik,
Leuphana Universität Lüneburg, 21335 Lüneburg, Germany
E-mail: {christiane.schneide, kathrin.padberg-gehle}@leuphana.de

[3] Courant Institute of Mathematical Sciences,
New York University, New York City, NY 10012, USA
E-mail: dimitris@cims.nyu.edu

[4] Tandon School of Engineering, New York University, New York City, NY 11201, USA
E-mail: katepalli.sreenivsan@nyu.edu

Turbulent convection flows are ubiquitous in natural systems such as in the atmosphere or in stellar interiors as well as in technological applications such as cooling or energy storage devices. Their physical complexity and vast number of degrees of freedom prevents often an access by direct numerical simulations that resolve all flow scales from the smallest to the largest plumes and vortices in the system and requires a simplified modelling of the flow itself and the resulting turbulent transport behaviour. The following article summarises some examples that aim at a reduction of the flow complexity and thus of the number of degrees of freedom of convective turbulence by machine learning approaches. We therefore apply unsupervised and supervised machine learning methods to direct numerical simulation data of a Rayleigh-Bénard convection flow which serves as a paradigm of the examples mentioned at the beginning.

1 Introduction

Machine learning algorithms,[5] particularly in the form of multilayered deep neural networks, have lately found various new applications in the Big Data domain that increasingly affects many fields of our daily life, *e. g.* by speech recognition tools in cellular phones. These machine learning methods are yet about to find an established place in the modelling and analysis of turbulent flows despite a few promising attempts.[2] The comprehensive data records needed for the algorithms come from full-scale supercomputer simulations of the equations of fluid motion or laboratory experiments applying optical measurement techniques. A proper assimilation and processing of this vast amount of data requires a change of paradigms in data processing, structure recognition and subgrid modelling.

Machine learning can be classified into two big categories, supervised and unsupervised machine learning. Supervision refers to a training of the algorithm with labelled input-output examples. In the following, we review briefly our own machine learning analyses with examples from both categories. The flow of interest is turbulent Rayleigh-Bénard convection (RBC) which serves as a paradigm to many turbulent flows in nature and technology that are driven and sustained by temperature differences.[1] On the one hand, we will

apply unsupervised machine learning for the reconstruction of the large-scale circulation (LSC) in a cubical closed convection cell by means of data-based Koopman eigenfunctions.[4] Unsupervised machine learning is also applied to study clusters of Lagrangian particle trajectories in a large-aspect-ratio turbulent convection flow.[12] On the other hand, supervised machine learning algorithms in form of deep neural networks are taken to reduce the three-dimensional and fully resolved turbulent superstructures in the temperature field of a Rayleigh-Bénard flow,[9] to a two-dimensional planar temporal network[6] of maximal and minimal temperature ridges with point defects.[3]

In all examples, we solve the three-dimensional Boussinesq equations of turbulent convection by means of direct numerical simulations that resolve all relevant turbulence structures in the flow and provide the data base. They couple the velocity field $\boldsymbol{u}(\boldsymbol{x},t)$ with the temperature field $T(\boldsymbol{x},t)$. The equations are made dimensionless by using the height of the layer or cell H, the free-fall velocity U_f and the imposed temperature difference ΔT between bottom and top. This implies a natural convective time unit, the free-fall time $T_f = H/U_f$. The Boussinesq system contains the three control parameters: the Rayleigh number Ra, the Prandtl number Pr and the aspect ratio $\Gamma = L/H$ with the cell length $L = L_x = L_y$ that is set to values of $\Gamma = 1$, 16 or 25 in this work. The equations are given by

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \tag{1}$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \boldsymbol{\nabla})\boldsymbol{u} = -\boldsymbol{\nabla} p + \sqrt{\frac{Pr}{Ra}}\,\boldsymbol{\nabla}^2 \boldsymbol{u} + T\boldsymbol{e}_z \tag{2}$$

$$\frac{\partial T}{\partial t} + (\boldsymbol{u} \cdot \boldsymbol{\nabla})T = \frac{1}{\sqrt{RaPr}}\boldsymbol{\nabla}^2 T \tag{3}$$

No-slip boundary conditions for the fluid are applied at all walls, *i. e.* $u_i = 0$. The side walls are thermally insulated, *i. e.* $\partial T/\partial n = 0$ with n being the normal direction. At the top and bottom a constant dimensionless temperature of $T = 0$ and 1 is maintained, respectively. The equations are numerically solved by the Nek5000 spectral element method package.[8, 11] In addition, N_p massless Lagrangian tracer particles are advanced corresponding to

$$\frac{d\boldsymbol{X}_j}{dt} = \boldsymbol{u}(\boldsymbol{X}_j, t) \tag{4}$$

with $j = 1 \ldots N_p$. The Lagrangian particles are advected by a 3-step explicit Adams-Bashforth scheme. The interpolation of the velocity field to the particle position is done spectrally. The turbulent convection flow can be considered as a nonlinear dynamical system in a high-dimensional phase space $\mathcal{M}$. The state vector of the RBC is given by $\phi_m(t) = (\boldsymbol{u}(\boldsymbol{x}_i,t), T(\boldsymbol{x}_i,t))$ where $\boldsymbol{x}_i$ are the coordinates of the grid points, $i = 1, \ldots, N$. With a total number of grid points of N it follows that $\dim(\mathcal{M}) = 4N$. The dynamical evolution of the turbulent flow describes a trajectory in $\mathcal{M}$ which is determined by $\phi_m(t_{k+1}) = F_m(\phi_n(t_k), t_k)$ where F_m is a nonlinear vector valued function with $m, n = 1, \ldots, 4N$. Here $k = 1, 2, \ldots, N_t$ with the number of time steps N_t.

2 Unsupervised machine learning applications

Unsupervised learning extracts features in (high-dimensional) data sets without pre-existing labelled training data. These techniques are partly already well-established and well-known. They comprise for example clustering and dimensionality-reduction techniques, such as the well-known Dynamic Mode Decomposition.[15] We will describe two specific examples more detailed in the following.

Giannakis *et al.*[4] discuss the long-time evolution of a three-dimensional turbulent RBC flow in a closed cubic cell via the eigenfunctions of the linear Koopman operator $\mathcal{K}_t$ that defines a new dynamical system which rather governs the evolution of observables on $\mathcal{M}$ than the state vector $\phi_m(t)$ itself.[15] The space of observables and thus $\mathcal{K}_t$ are infinite-dimensional even though the associated nonlinear dynamical system is finite-dimensional. In practice, one seeks a finite-dimensional representation of $\mathcal{K}_t$; more detailed, a data-driven basis is calculated here to represent a regularised generator V of the unitary Koopman group $\mathcal{U}_t = \exp(tV)$ in the sense of a Galerkin approximation. This is exactly the point where unsupervised machine learning enters the scene. The goal is to find (or better to *geometrically learn*) a possibly complicated curved submanifold $\mathcal{N} \subset \mathcal{M}$ with $\dim(\mathcal{N}) \ll \dim(\mathcal{M})$ on which the few relevant degrees of freedom evolve that describe the large-scale flow of the RBC setting effectively. This is done by a diffusion process on the data snapshots, $\phi_i(t_k)$, that quantifies the local connectivity of two system states i, j

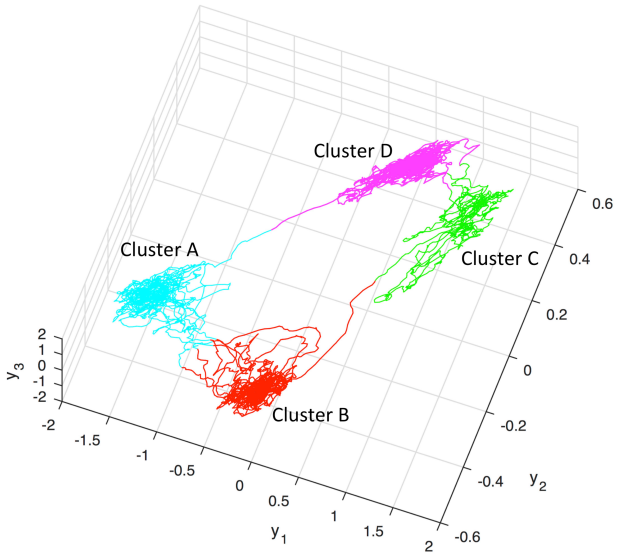


Figure 1. Dynamical trajectory of a turbulent RBC flow in a closed cubical cell at $\mathrm{Ra} = 10^7$. Shown is a three-dimensional subspace with coordinates which are constructed from the primary Koopman eigenfunctions. The trajectory is sampled over 10^4 convective free-fall time units T_f. The large-scale flow dynamics is found to cluster in four macro-states A, B, C, and D that correspond to large-scale circulation rolls in the box across the diagonals (two configurations times clockwise/counterclockwise spin). The figure is taken from Ref. 4.

and the resulting transition probability (or the Markov operator) with a Gaussian kernel which follows to

$$P_{ij} = \frac{1}{C} \exp\left(-\frac{1}{\epsilon Q} \sum_{q=0}^{Q-1} \|\boldsymbol{u}_{i-q} - \boldsymbol{u}_{j-q}\|_{L^2}^2 \right) \tag{5}$$

Here, C is a normalisation constant; we use the velocity field data only. The sum over q stands for an additional time delay-averaging. The eigenvectors of P_{ij} form the data-driven basis in which we can compute the Koopman eigenfunctions afterwards.

The resulting Koopman eigenfunctions can be grouped into subsets in accordance with the discrete symmetries in a cubic box. In particular, a projection of the velocity field onto the first group of eigenfunctions reveals the four stable large-scale circulation (LSC) states in the convection cell which are shown in Fig. 1. We recapture the preferential circulation

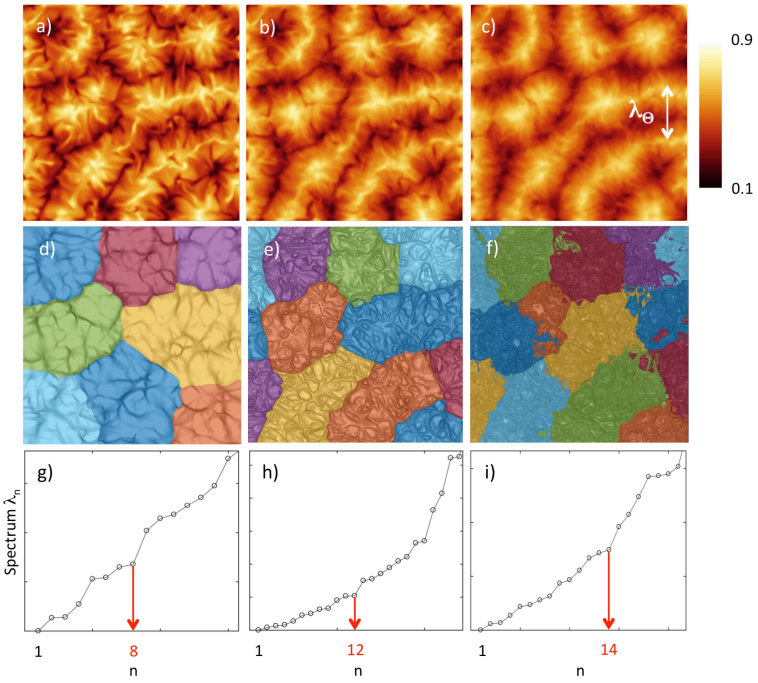


Figure 2. Lagrangian trajectory-based analysis of three-dimensional turbulent large-scale patterns in a RBC flow. We advected 512^2 individual Lagrangian tracers that started inside the thermal boundary layer at the bottom plate. (a–c) Time-averaged temperature field in the midplane. Progressing averaging times are 2.6, 10.4, and 30.1 free-fall convective time units, respectively. In panel (c), the characteristic pattern scale $\lambda_\Theta \approx 4H$ of the turbulent superstructures is indicated by a double arrow. (d–f) Lagrangian trajectory clusters obtained from the leading eigenvectors of the graph Laplacian matrix. Particles that belong to the same spectral cluster at the corresponding time are coloured equally. The background contours are the ridges of the maximum finite-time Lyapunov exponent. Ridges and clusters are indicated with respect to the initial Lagrangian particle position. (g–i) Corresponding eigenvalue spectra of the graph Laplacian matrix. The spectral gap between eigenvalues no. 8-9, 12-13, and 14-15 is used to detect k = 8, 12, and 14 trajectory clusters by a standard k-means clustering algorithm, respectively. The parameters are $(Ra, Pr, \Gamma) = (10^5, 0.7, 1)$ for this simulation. The figure is taken from Ref. 10.

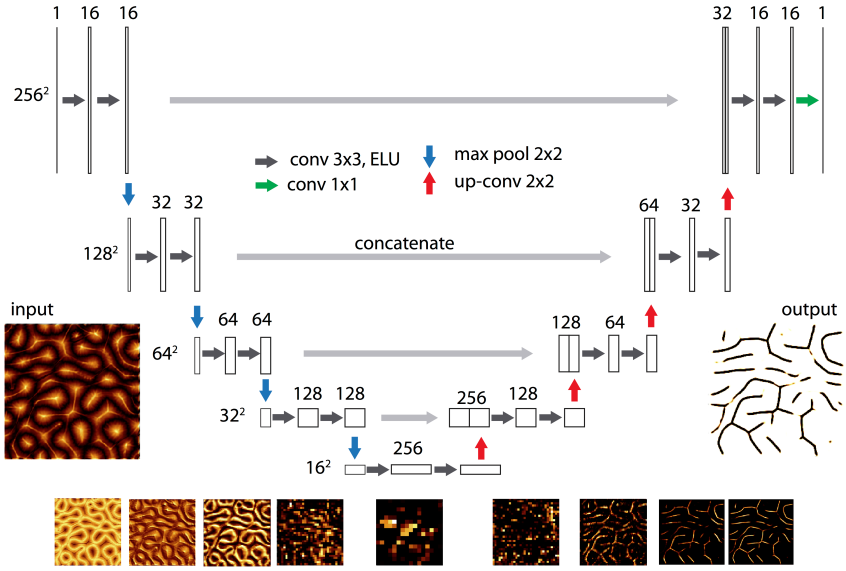


Figure 3. Workflow within the U-shaped deep neural network. The resolution of the image (or feature map) and the number of images are indicated at each step. At the bottom of the figure, we display some levels of the data processing within the hidden layers. The thumbnail contour plots are centred with respect to the corresponding hidden layer. Convolution and max-pooling operations are indicated in the legend. There are a total of 28 layers and 1 940 817 parameters. Figure is taken from the supplementary information of Ref. 3.

rolls in diagonal corners and the short-term switching through roll states parallel to the side faces which have also been seen in other simulations and experiments. The diagonal macroscopic flow states can last as long as 1000 convective free-fall time units.

Schneide *et al.*[12] applied two unsupervised machine learning techniques to a three-dimensional RBC flow with aspect ratio $\Gamma = 16$ – spectral clustering and density-based spatial clustering applications with noise (DBSCAN) – in order to analyse the time evolution of a Lagrangian particle ensemble in the RBC flow as a whole. To this end, the particle trajectories are composed into a network and its connection to the large-scale organisation of an extended three-dimensional turbulent convection flow – termed turbulent superstructures of convection – was studied. The set of N_p individually advected Lagrangian particles at a time t forms a set of vertices $\{v_1, \ldots, v_{N_p}\} \in V$ of a weighted and undirected graph $G = (V, E, w)$. The vertices are connected by edges $\{e_1, \ldots, e_M\} \in E$. The dynamical history of the particle ensemble up to time t can then be encoded into the weights w which are assigned to each edge. These weights are calculated as the inverse of a time-averaged distance of mutual tracer trajectories and set to zero if this distance exceeds a threshold value.[12] The latter is the graph sparsification step. By solving a balanced cut problem via an equivalent generalised eigenvalue problem of the Laplacian matrix of the graph[13] and a subsequent k-means clustering of the leading eigenvectors, the network is decomposed into k subgraphs or clusters. The obtained clusters of the graph could be related to large-scale ridge patterns of temperature as seen in Fig. 2. The dispersion of the Lagrangian particles

that start their evolution in close distance to each other is an immanent property of a turbulent flow and can be quantified by the largest Lyapunov exponent which is also shown as a background texture in the middle row panels of Fig. 2. Therefore, such a spectral clustering analysis is applicable for the shorter-time evolution only. For larger evolution times, a density-based clustering of pseudo-trajectories, which are obtained as time averages of Lagrangian trajectories, was applied to probe the long-living patterns in the convection flow. The latter analysis step gives us the most coherent subset of trajectories which are trapped for a long time in the core of the circulation rolls.

3 Supervised machine learning application

Supervised machine learning makes use of the fact that it is often easier to train a system, such as a deep convolutional neural network in the present case, with a number of labelled examples of an intended input-output, than to develop a specific computer program to provide the correct answer for all possible input data.

In Fonda *et al.* we discuss the application of a deep convolutional neural network (DCNN) to quantify the role of turbulent superstructures in the turbulent heat transport. We analysed data records that were obtained in a domain with a large aspect ratio of 25. The Prandtl number of the data sets was $Pr = 7$ and the Rayleigh numbers were $Ra = 10^5, 10^6$, and 10^7. As already stated in the introduction, the superstructures manifest as a temporal skeleton of ridges of hot upwelling and cold downwelling fluid with defects where the ridges merge or end. We trained a DCNN to reduce the three-dimensional temperature field to a temporal planar network in the midplane of the layer – a data compression by more than 5 orders of magnitude at the highest Rayleigh number. The resulting network quantified the turbulent heat transferred by the superstructure. Defect points that disappear and form anew with time are found to be "hot spots", which are points of locally enhanced heat flux (see Fig. 5). It was shown that the fraction of heat carried by the superstructure decreases as the Rayleigh number increases. We also found that an increasing amount of heat is carried by the small-scale background turbulence from the bottom to the top rather than by the coherent large-scale superstructure of the convection flow.

The convolutional deep neural network which we applied is U-shaped as seen in Fig. 3, denoted as U-net for short.[10] The U-net combines a contraction path of a standard convolutional neural network[5] with a subsequent expansion path of concatenations and upconvolutions that finally creates a detailed segmentation map. This U-net has been applied successfully to image segmentation of touching objects, *e. g.* of neuronal cell structures in electron microscopy data. The specific architecture of the U-net is essential for the present usage in a large-aspect-ratio RBC flow as it requires small sets of manually annotated data for the training only. This is in contrast to standard DCNNs for such images. The slow evolution of the superstructures in the large-aspect-ratio domain would otherwise require extremely long simulations of the full RBC flow over a few tens of thousands of convective time units to obtain an appropriate amount of independent training data, as discussed by Pandey *et al.*[9]

The U-net algorithm[10] was used within the Tensorflow software environment[14] in combination with the application programming interface Keras.[7] The temperature data in the midplane were originally obtained at a 1024^2 resolution by a spectral interpolation onto a Cartesian uniform mesh. We verified that a reduction of the resolution to 256^2 grid

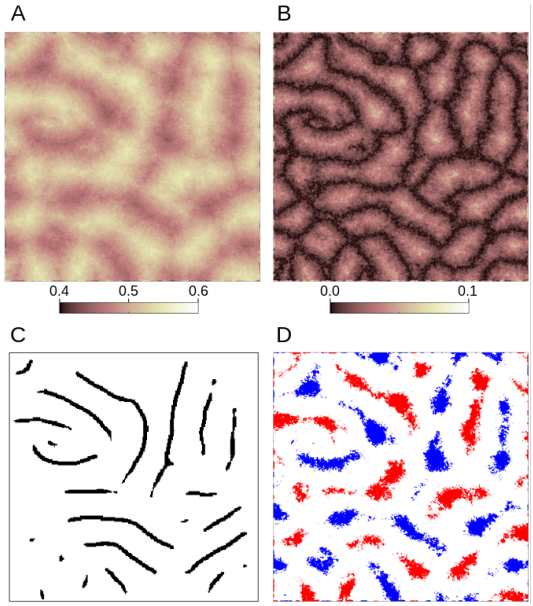


Figure 4. Input and output of the U-net. (A) Original two-dimensional and time-averaged temperature slice for $Ra = 10^7$ and $Pr = 7$. (B) Corresponding symmetrised temperature field $\Theta(x, y)$ that enters the U-net. (C) Ridge pattern as output of the U-net. (D) Comparison with a simple thresholding set. Blue pixels are for cold downflows and red pixels for hot upflows. These ridges have a strongly varying thickness and are partly interrupted when a simple thresholding is conducted. The figure is taken from Ref. 3.

points that cover the whole plane is sufficient to capture the ridge pattern. The training of the neural network for 1000 epochs took about 15 minutes on a cluster with NVIDIA GeForce GTX 1080. Note, however, that even 100 epochs were enough in our case. Fig. 4 demonstrates the extraction of the ridge pattern by the deep neural network, even though the data were for the highest Rayleigh number case very noisy. Input to the U-net is the symmetrised temperature field which is given by $\Theta(x, y) = |T(x, y, z = 1/2) - 1/2|$ and shown in panel B of the figure.

Since supervised machine learning is highly sensitive to the quality of the training data, the annotation of the samples is a key task. Different Rayleigh and Prandtl numbers cause changes in the sharpness, the size, and the amount of noise of the underlying ridge network of temperature field superstructures. In Fonda *et al.*[3] the trends with Rayleigh number at $Pr = 7$ were investigated. Here, we will take a closer look at the Prandtl number dependence at fixed Ra that affects the superstructures in a different way as discussed in Pandey *et al.*[9] in detail. Two scenarios were tested. In the first case, we created training data for the U-net based on low-noise snapshots with a clearly visible ridge network (see Fig. 5 (A, E)). The samples were highly augmented by several geometric and photometric manipulations including artificial blurring, addition of noise and distortion. In the second case, examples from three different Prandtl numbers, $Pr = 7, 0.7$ and 0.021, were combined for the training of the U-net leaving the corresponding manual annotations of the ridges much less detailed. It is found that the training data that contained many details and were highly

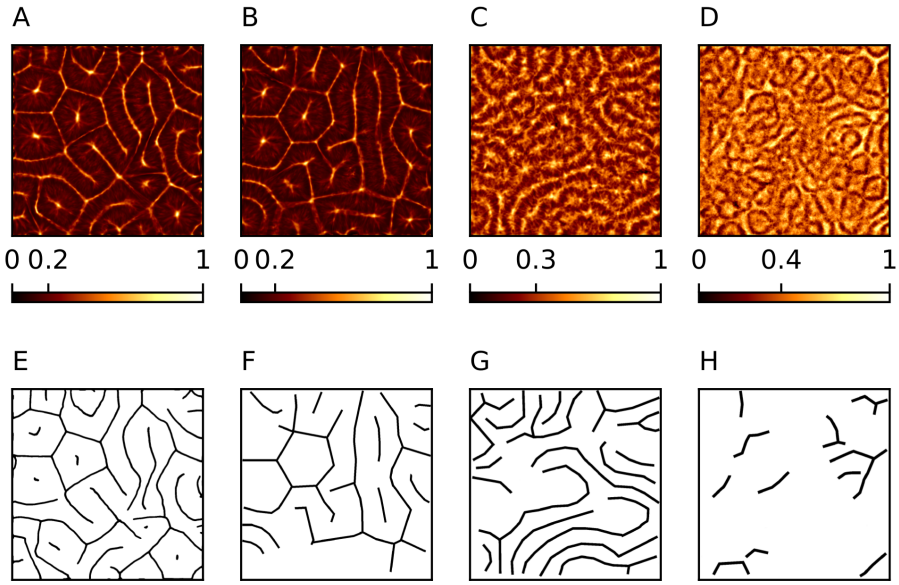


Figure 5. Different training data annotations of the symmetrised temperature field $\Theta(x, y)$ all for a Rayleigh number $Ra = 10^5$. (A, E) sample and annotation for $Pr = 7$. (B, F) again for $Pr = 7$. (C, G) for $Pr = 0.7$, and (D, H) for $Pr = 0.021$. All input data are normalised such that they vary between 0 to 1 and can be combined. The colour bars below the top row of panels also indicate the corresponding mean value.

augmented (case 1) slow down the convergence of the DCNN significantly. The corresponding U-net gives very detailed image segmentations. In contrast, the U-net which was trained with examples from all three simulation records (case 2) converged faster and produced segmentations that show the coarsest ridges only. The results of the ridge extraction with both U-nets are summarised in Fig. 6 for two different Prandtl numbers. Depending on the subsequent use of the output, one either obtains a fine-scale segmentation or a much coarser ridge pattern. Our present results show clearly that the application of the U-net becomes more challenging as the Prandtl number decreases. In these cases, the ridge patterns of the symmetrised temperature are more washed out due to the more vigorous fluid turbulence and the ridges appear to be more disconnected. Clearly, even the U-net approaches here its limitations.

The extracted ridge networks, as obtained in Fig. 6 (C, G), were subsequently used to quantify the heat transport across the convection layer. The dimensionless number that measures the turbulent heat transport in RBC is the Nusselt number, which is given by $Nu = 1 + \sqrt{RaPr}\langle u_z T\rangle_{V,t}$ with the vertical velocity component u_z, the temperature T, and the combined volume-time average $\langle\cdot\rangle_{V,t}$. In Fig. 7 (A) we compare Nu with a Nusselt number that measures the transport across the ridge network (and thus due to the turbulent superstructure (TSS)). It is given by $Nu^{\mathrm{TSS}} = \sqrt{RaPr}\, w_e \langle u_z \Theta\rangle_r$. Here, w_e is a weight factor that represents the area fraction of the network compared to the area content of the horizontal plane and r a segment of unit length in the ridge network. Fig. 7 (B-D) show the transport capability of the ridge network for three Pr. The heat flux by the

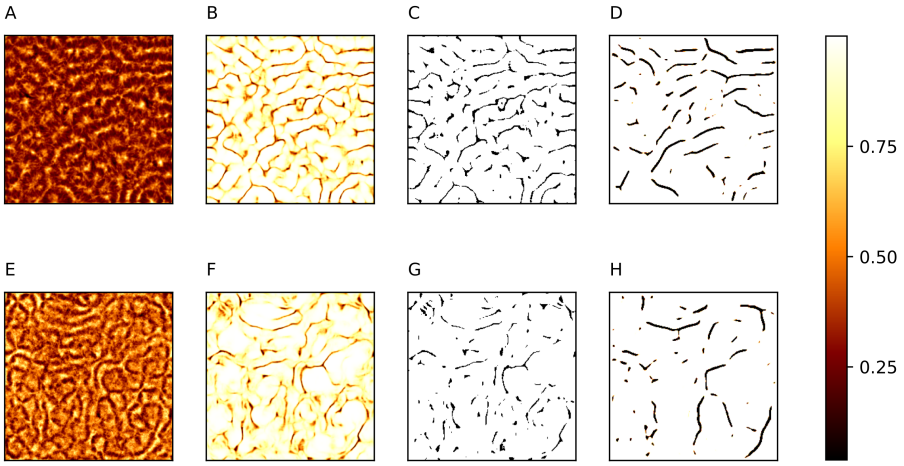


Figure 6. Output of the U-net for $Pr = 0.7$ (top row) and $Pr = 0.021$ (bottom row), both at $Ra = 10^5$. (A, E) Normalised input data of symmetrised temperature. (B, F) U-net output when trained with detailed annotation and high augmentation rate. (C, G) Conversion of (B, F) into a binary data set with a threshold of 0.5 is shown. (D, H) Direct output result of the U-net trained with less detailed annotations.

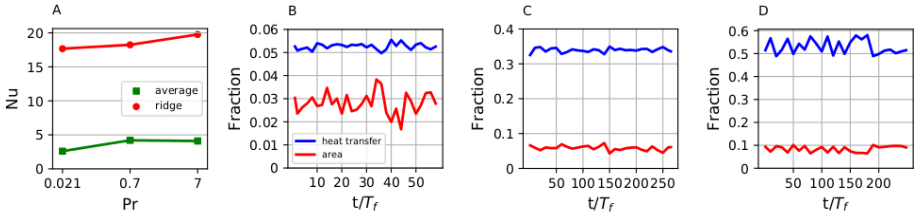


Figure 7. Turbulent heat transport analysis across the ridge network. (A) Mean convective heat transport across the whole plane (green) and transport due to the ridge network (red). (B-D) Amount of heat transport across the network *vs.* time and area fraction of the network *vs.* time. Data are for at $Ra = 10^5$ and $Pr = 0.021$ in B, $Pr = 0.7$ in C and $Pr = 7$ in D.

network increases faster than the area occupied by it with increasing Pr, suggesting that the transport by the ridge network becomes more efficient.

4 Concluding remarks

We have discussed a number of applications of unsupervised and supervised machine learning in turbulent convection. They were used to extract coherent patterns in these flows and to reduce the big amount of three-dimensional simulation data effectively. As in many other applications, we showed that the success of supervised machine learning with DCNNs is tightly related to the training methods. Future tasks in this field are related to predictions of turbulent flow behaviour itself, *e. g.* by the use of recurrent neural networks

and thus to construct effective small-scale parametrisations. These efforts are currently in progress and will be reported elsewhere.

Acknowledgements

The work of D. K. is supported by the Deutsche Forschungsgemeinschaft (DFG) and the work of A. P. and C. S. by the Priority Programme DFG-SPP 1881 *Turbulent Superstructures*. JS thanks the Tandon School of Engineering at New York University for financial support. We want to thank the John von Neumann-Institute for Computing for the steady support with supercomputing resources that make our research work in turbulent convection possible. The reported work was and is supported by the project HIL12.

References

1. F. Chillà and J. Schumacher, *New perspectives in turbulent Rayleigh-Bénard convection*, Eur. Phys. J. E **35**, 58 (25 pages), 2012.
2. K. Duraisamy , G. Iccarino, and H. Xiao, *Turbulence modeling in the age of dat*a, Annu. Rev. Fluid Mech. **51**, 357–377, 2019.
3. E. Fonda, A. Pandey, J. Schumacher, and K. R. Sreenivasan, *Deep learning in turbulent convection networks*, Proc. Natl. Acad. Sci. USA **116**, 8667–8672, 2019.
4. D. Giannakis, A. Kolchinskaja, D. Krasnov, and J. Schumacher, *Koopman analysis of the long-term evolution in a turbulent convection cell*, J. Fluid. Mech. **847**, 735–767, 2018.
5. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, The MIT Press, Cambridge, MA, 2017.
6. P. Holme and J. Saramäki, *Temporal networks*, Phys. Rep. **519**, 97–125, 2012.
7. F. Chollet *et al.*, *Keras*, `https://keras.io`, 2015.
8. `https://nek5000.mcs.anl.gov/`
9. A. Pandey, J. D. Scheel, and J. Schumacher, *Turbulent superstructures in Rayleigh-Bénard convection*, Nat. Commun. **9**, 2118 (11 pages), 2018.
10. O. Ronneberger, P. Fischer, and T. Brix, *U-Net: Convolutional Networks for Biomedical Image Segmentation*, Lecture Notes in Computational Science **9351**, 234–241, 2015.
11. 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.
12. C. Schneide, A. Pandey, K. Padberg-Gehle, and J. Schumacher, *Probing turbulent superstructures in Rayleigh-Bénard convection by Lagrangian trajectory clusters*, Phys. Rev. Fluids **3**, 113501 (15 pages), 2018.
13. J. Shi and J. Malik, *Normalized Cuts and image segmentation*, IEEE Trans. Pattern Anal. Mach. Intell. **22**, 888–905, 2000.
14. M. Abadi *et al.*, *TensorFlow: Large-scale machine learning on heterogeneous systems*, `https://www.tensorflow.org`, 2015.
15. M. O. Williams, I. O. Kevrekidis, and C. W. Rowley, *A data-driven approximation of the Koopman operator: Extending Dynamic Mode Decomposition*, J. Nonlinear Sci. **25**, 1307–1346, 2015.

High-Performance Flow Simulation and Scale-Adaptive Turbulence Modelling of Centrifugal Pumps

Markus Hundshagen, Nicolas Casimir, Andreas Pesch, and Romuald Skoda

Chair of Hydraulic Fluid Machinery, Ruhr University Bochum, 44801 Bochum, Germany
E-mail: romuald.skoda@ruhr-uni-bochum.de

While for the design point operation of centrifugal pumps, where an essentially steady flow field is present and statistical turbulence models yield an appropriate prediction of the characteristics, the flow field gets increasingly unsteady towards off-design operation. Special designs as *e. g.* sewage pumps are characterised by a single-blade impeller and show significantly unsteady characteristics even in the design point. For such highly-unsteady and turbulent flow fields, statistical models tend to fail. On the other hand, Large-Eddy Simulation models, where the large-vortex part of the turbulent spectrum is directly resolved, show a much better flow prediction. However, the spatial resolution and thus computational effort are too high for engineering real pump applications. Therefore, we provide an assessment of scale-adaptive turbulence simulation (SAS) models that recover a statistical flow solution in regions of low unsteadiness and – like Large-Eddy Simulation – resolve a part of the turbulent spectrum down to the available grid resolution for highly unsteady flow regions. After a thorough validation on standard turbulence test cases *e. g.* the periodic hill case, it is shown that with a moderately higher computational effort than statistical models, the SAS yields a considerable improvement of the prediction of the turbulence field in part load operation of a centrifugal pump while the mean flow field could be well predicted even with a well-established statistical model.

1 Introduction

Increasing demands on centrifugal pumps, *e. g.* by legislature, require higher efficiencies even at off-design operation, *i. e.* part load and overload, which are characterised by a highly unsteady and turbulent flow field due to flow separation and impeller-stator interaction. CFD methods are increasingly integrated in the design and optimisation process of pumps. Statistical (*i. e.* URANS) eddy-viscosity turbulence models are widely used for pump flow simulations. In this class of turbulence models, a significant simplification is introduced by an *a priori* time-average of the turbulent fluctuations. Thus, the resulting Reynolds-stress tensor has lost any spectral information of the turbulence field. Although statistical models may yield a good prediction of pump characteristics at close-to-design operation, they may increasingly fail towards off-design pump operation. In a previous study, a comparison of spatially and temporally high resolved measurement data to URANS simulations with the k-ω-Shear Stress Transport (SST) turbulence model[1] (referred to as SST in what follows) revealed that, although time-averaged head and ensemble-averaged flow angle were in good agreement for design and part load operation, turbulence intensity (TI) is over-predicted especially in part load operation near the volute tongue.[2] There are several other studies that show limitations of statistical models and benefits of scale-resolving Large-Eddy Simulation models (LES).[3–12] However, the computational effort of LES is tremendous at high Reynolds numbers encountered in centrifugal pumps because at least 80 % of the spectral energy must be resolved[13, 14] and a soundly resolved LES demands therefore an extensively high number of computational cells and very small

time steps. In most LES studies on centrifugal pumps, it remains unclear if a sufficient amount of the spectral energy is resolved to fulfil the LES criteria.

An interesting approach is the Scale-Adaptive Simulation method,[15] which is an improvement of statistical models with the ability to resolve the turbulent spectrum down to the available grid limit. This is achieved by a reduction of turbulent viscosity depending on the von Karman length scale and the integral length scale. A variant of this model is the k-ω-SST-SAS model (referred to as SAS).[16] Especially the fall-back of the SAS to a SST solution in regions of low spatial and temporal resolution is a convenient way to avoid uncertainties in spectral energy resolution.[16] Several studies show the improvement of prediction accuracy of this type of model in highly transient flow compared to statistical turbulence models.[5, 8, 12, 17, 18] However, in these studies it remains unclear what the "critical" spatial grid resolution is to activate the scale-resolving capability of SAS and thus to achieve an improvement over SST solutions. Therefore, we want to assess the capability of the SAS model with focus on grid dependence for centrifugal pump flow.

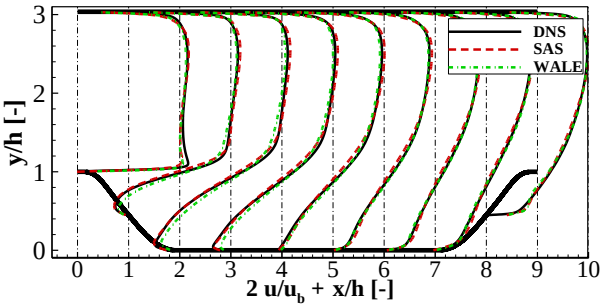
2 Numerical Method and Validation

The incompressible Navier-Stokes equations in their unsteady Reynolds-averaged form (URANS) are solved by the open source software foam-extend version 4.0.[a] The software foam-extend is a C++ toolbox for the development of customised solvers. A conservative finite volume co-located unstructured spatial discretisation is chosen. We utilise central differences in combination with second order total variation diminishing (TVD) differencing schemes for convective fluxes for scale resolving turbulence models. An implicit pressure based solver is utilised, *i. e.* a combined PISO[19] and SIMPLE[20] algorithm for incompressible unsteady flows since it is the most efficient choice for relatively small time steps with a Courant number ~ 1. We combine this solver with moving mesh capabilities. For the evaluation of convergence of the non-linear iterative PISO-SIMPLE algorithm, preliminary investigations have shown that a drop of the non-linear non-dimensional residual sum norm of each equation below a value of 10^{-5} is a suitable convergence criterion. While the statistical two-equation eddy-viscosity SST turbulence model by Menter[1] has been already available in the standard package of the software, we have implemented an SAS extension by Menter.[16] The statistical (SST) and scale-adaptive (SAS) variant only deviate by an additional source term in the dissipation scale (ω) equation so that the computational effort is essentially the same if the temporal and spatial resolution is the same.

To show the performance of the SAS, we have investigated several standard test cases that have been optimised for turbulence model assessment. *E. g.* in the channel and duct flow as well as asymmetric diffusor case (not shown here) the SAS model is not able to switch from URANS to scale-resolving simulations without an explicit introduction of synthetic turbulence and thus falls back to the statistical SST solution as has been observed also in other studies, *e. g.* Mehdizadeh *et al.*[21] Thus, we focus on the "periodic hill" test case where the periodic in- and outflow conditions support an amplification of model inaccuracies and is thus particularly challenging for turbulence models. In fact, the statistical SST variant completely fails to predict the separation and re-attachment location of the pronounced separation downstream of the hills (not shown here). On the other hand, we

[a] `http://foam-extend.fsb.hr/`

a) Streamwise velocity



b) Cross-streamwise velocity

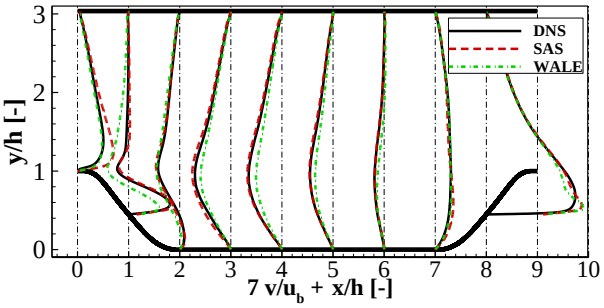


Figure 1. Mean velocity field of the periodic hill test case at bulk Reynolds number $\mathrm{Re_b} = 10.595$. DNS data by Krank *et al.*[23]

observe a distinct scale-resolving mode of the SAS model. In Fig. 1 a comparison of the SAS and a classical LES solution by the WALE model[22] is provided. We have performed the simulations on different computational grid resolutions (one simulation run on the finest grid demands an effort of about 60 thousand core hours) and show only the coarsest grid solution here to demonstrate the SAS model performance with "engineering" resources in terms of relatively coarse grids. In addition, benchmark data from Direct Numerical Simulations (DNS) by Krank *et al.*[23] are included in Fig. 1. The SAS matches the benchmark data and thus provides an essential improvement to the statistical SST solution. While for the stream-wise velocity component, *c. f.* Fig. 1(a), neither significant differences between SAS and LES nor between the grids (not shown) are discernible, it is interesting to see that for the cross-stream-wise component, *c. f.* Fig. 1(b), the LES for the coarser grid shows a significant deviation from data that diminishes with grid refinement (not shown here). The SAS solution matches the data well even on the coarse grid. Thus, this exemplary observation underlines the benefit of the SAS to yield appropriate results even on grids that do not fulfil the tremendous spatial and temporal resolution demands of a LES and where the LES may yield inaccurate results due to an under-resolution of the spectral energy cascade. This observation will be further explored in a centrifugal pump in what follows.

3 Centrifugal Pump Test Case

A single stage volute casing radial centrifugal pump (specific speed $n_S = 26\ \mathrm{min}^{-1}$) is investigated at part load operation. The pump is operated with air in order to enable hotwire anemometry measurements for a high temporal and spatial resolution of the velocity field near the impeller discharge (measurement plane depicted in Fig. 2(b)) and thus enables a comparison with turbulence statistics from the simulation. Since the maximum Mach number is evaluated to be below 0.3, it is assumed that compressibility effects on the velocity field are of minor relevance. Pump details and experimental data are documented by Hergt *et al.*,[24] Meschkat and Stoffel[25] and Casimir *et al.*[2] and are briefly summarised here.

The head characteristics is discussed by Casimir *et al.*,[2] where a good agreement between measurement and simulation has been obtained even with the statistical SST model. However, deviations of the local turbulence field in part load could be observed. Thus in the present study, we focus our attention to a discussion of the local velocity field and not on the integral characteristics. Mean flow angle and standard deviation of velocity (RMS) are presented in the measurement plane as functions of the axial position y/b, *c. f.* Fig. 2(a), the circumferential coordinate ε, *c. f.* Fig. 2(b) and the time coordinate x/t, which corresponds to the time range when one blade spacing passes the measurement position, *c. f.* Fig. 2(c). The local flow-angle α is evaluated with Eq. 1

$$\alpha = \frac{|c_\mathrm{r}|}{c_\mathrm{r}} acos\left(\frac{c_\mathrm{u}}{c_\mathrm{abs}}\right) \tag{1}$$

where c_r is the radial velocity component, c_u is the circumferential velocity component and c_abs is the magnitude of the absolute velocity. $\alpha < 90°$ means backflow and $\alpha > 90°$ means outflow against the rotational direction. Measured TI is evaluated by the RMS of the absolute velocity magnitude according to Eq. 2 and includes turbulent as well as non-

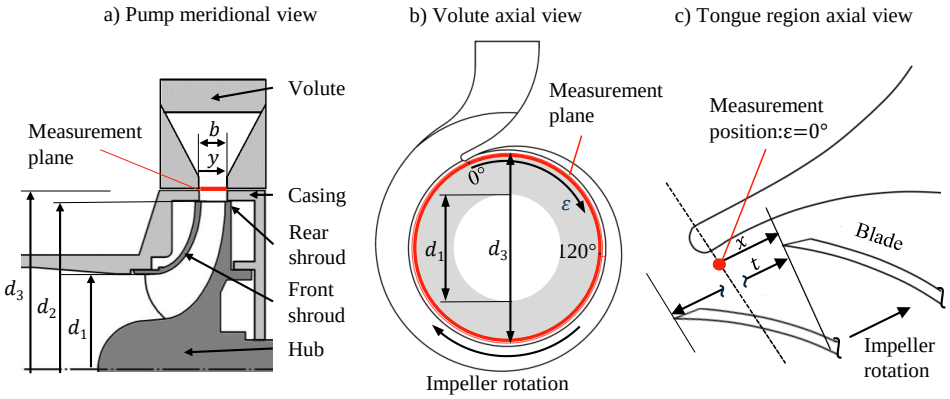


Figure 2. (a) Meridional contour; (b) Measurement plane; (c) Illustration of coordinate x, which corresponds to the time range when one blade spacing passes the measurement position.

turbulent fluctuations.

$$\mathrm{TI_{EXP}} = \frac{\mathrm{RMS}(c_{\mathrm{abs}})}{u_2} \tag{2}$$

u_2 is the circumferential impeller trailing edge velocity. For the simulations, a TI comparable to measurements is evaluated by taking into account temporally resolved URANS-velocity fluctuations (Eq. 3) and ensemble-averaged turbulent kinetic energy $\overline{k}$ (Eq. 4):

$$\mathrm{TI_{RMS}} = \frac{\sqrt{\mathrm{RMS}^2(c_{\mathrm{abs,Sim}})}}{u_2} \tag{3}$$

$$\mathrm{TI_{TKE}} = \frac{\sqrt{2\overline{k}_{\mathrm{Sim}}}}{u_2} \tag{4}$$

k corresponds to *a priori* time-averaged fluctuation in the URANS approach. Thus, total TI is evaluated according to Eq. 5 in the simulation.

$$\mathrm{TI_{SIM}} = \frac{\sqrt{(\mathrm{RMS}^2(c_{\mathrm{abs,Sim}}) + 2\overline{k}_{\mathrm{Sim}})}}{u_2} \tag{5}$$

For further information on the evaluation methods, we refer to Casimir *et al.*[2] and Pesch *et al.*[26]

4 Numerical Setup

The computational domain contains the impeller, volute casing, side chambers and the suction and pressure pipe, *c. f.* Fig. 3(c). The suction pipe has been elongated to a length of $5 \times \mathrm{D}$ to avoid an impact of boundary conditions on the part load vortex. A block structured hexahedral grid as shown in Fig. 3(b) with approximately 3 million nodes (named G1) is generated. For a grid study, the grid G1 is successively refined to approximately 24 million nodes (named G2) by bisection of node distances in each direction. For G1, average and maximum y^+ values equal about 10 and 60, and for G2 about 5 and 30, respectively. The fluid is incompressible air with a kinematic viscosity of $\nu = 10^{-5}\mathrm{m}^2/\mathrm{s}$.

A Dirichlet inlet boundary condition is set for velocity according to flow rates from experiment, together with a Neumann (zero-gradient) condition for static pressure. At the volute discharge, a Neumann boundary conditions is set for velocity (zero gradient) and a Dirichlet condition for static pressure. Temporal and spatial second order discretisation methods are used. Tabs. 1 and 2 list the settings for SST and for SAS simulations, respectively. As the time step size needs to be small with the SAS model ($\mathrm{CFL_{max}} < 1$), the solver is run in PISO mode, *i. e.* only one outer corrector step is performed for SAS simulations.

The rotating domain includes the suction pipe, impeller and side chambers, and the stationary domain includes the volute-casing and discharge pipe, *c. f.* Fig. 3(c). The governing equations are solved in the absolute frame of reference so that a moving grid is applied for the rotating domain. Thus, the rotor and the stator grid need to be coupled at each time step. Transient coupling of rotor and stator domain is achieved by a General Grid Interface (GGI) algorithm by Beaudoin and Jasak[30] that allows a non-conformal grid, *i. e.* non-matching grid nodes on both, rotor and stator side of the interface. As described by

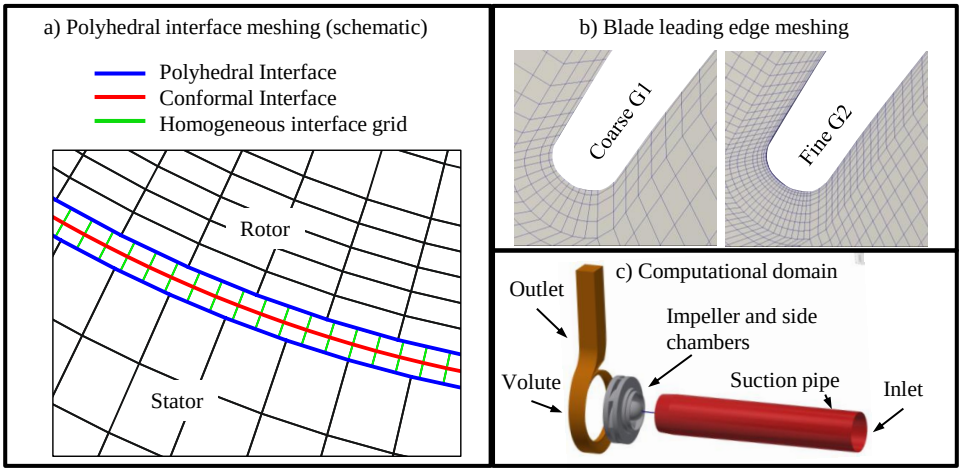


Figure 3. (a) Schematic illustration of the rotor-stator interface grid; (b) Computational grid at the blade leading edge; (c) Computational domain.

Casimir *et al.*[2] and Pesch *et al.*,[26] the algorithm yields spurious oscillations for non-planar interfaces due to local non-conservative flux transmission and thus needs to be modified for the impeller-stator sliding interface. We introduced a polyhedral interface meshing, *c. f.* Fig. 3(a) and a thin homogeneous interface grid whose circumferential node distribution matches the circumferential rotor movement within each time step, yielding matching interface nodes and thus a conservative interface at each time step, referred to as conformal interface. Details on the rotor-stator interface can be found in Casimir *et al.*[2] and Pesch *et al.*[26]

Parameter	**Description**
Turbulence model	SST[1]
Wall functions	k: kqrWallFunction ν_t: nutWallFunction ω: omegaWallFunction
Time discretisation	Backward Euler
Discretisation (div U)	LinearUpwind[27]
Discretisation (k and ω)	vanLeer[28]
Linear solver for p	BiCGStab
Linear solver (U, k, ω)	smoothSolver
Time step size	$1\,^\circ \rightarrow$ CFL ≈ 30
nCorrectors (PISO loops)	2
nOuterCorrectors (PIMPLE loops)	max. 50 with convergence control

Table 1. Summary of the numerical setup for SST simulations.

Parameter	Description
Turbulence model	SAS[16]
Wall functions	see Tab. 1
Time discretisation	Backward Euler
Discretisation (div U)	DESHybrid[29]
Discretisation (k and ω)	vanLeer[28]
Linear solver for p	BiCGStab
Linear solver for U, k, ω	smoothSolver
Time step size	1/32 $^\circ \rightarrow$ CFL < 1
nCorrectors (PISO loops)	2
nOuterCorrectors (PIMPLE loops)	1

Table 2. Summary of the numerical setup for SAS simulations.

For implicit pressure based solvers, a large part of the computational resources in terms of CPU-time are utilised for solving the pressure equation. Preconditioned conjugated gradient solvers (PCG) for the solution of the linear equation system have been found to be most efficient. The re-calculation of the flux assignments at the rotor-stator interface after each mesh motion step also contributes considerably to the resources. The code is parallelised with OpenMPI. For the benchmark, SAS simulations on a grid of 24 million cells with up to 20 computational nodes (960 Cores) are performed on the JUWELS cluster. A constant number of pressure iterations per time step is specified for the scaling assessment. A total amount of 4 time steps is calculated and one write operation of the solution is considered for the benchmark. The resulting scaling characteristics are shown in Fig. 4. It is noticeable, that the scaling behaviour is nearly ideal up to 12 nodes ($42\,000$ cells per core), which is mainly related to the excellent scalability of the PCG solver. The used amount of memory per node is approximately 16 GB.

Regarding data storage, large write operations are done for full-3D backup results with a size of $\sim$5 GB and a temporal interval of 5 minutes on 960 cores, while more frequent write operations are performed for monitoring results every 4 seconds with $\sim$20 MB of data.

5 Results

Ensemble-averaged mean flow angle α and TI are evaluated in the measurement plane in the impeller wake and compared between measurement and simulation for part load, *i. e.* at 40 % of the nominal pump load. In this off-design operation, turbulent flow unsteadiness is highly distinctive. Statistical convergence of the measurements is achieved by an ensemble-average over more than 100 impeller revolutions. For ensemble-averaging, 7 values per revolution are available for the 7-blade impeller. For SAS simulations, less revolutions are considered for ensemble-averaging due to the limit of computational resources, what results in a significantly smaller amount of samples than for the measurements.

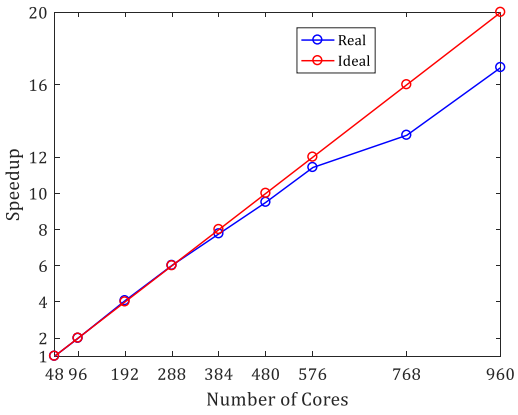


Figure 4. Scaling characteristics on JUWELS with a problem size of 24 million cells.

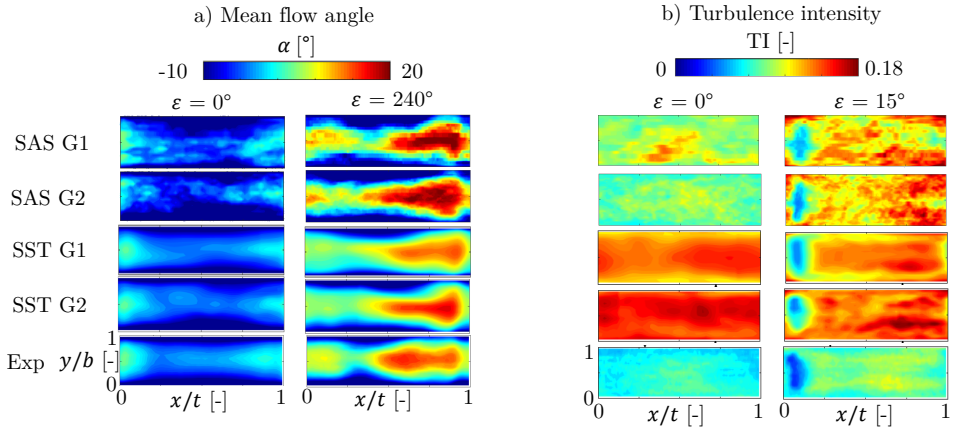


Figure 5. Ensemble-average flow angle (a) and turbulence intensity (b) for different circumferential volute positions ε.

Fig. 5(a) shows a comparison of an ensemble-averaged flow angle between SST, SAS and measurement data for two different volute positions near $\varepsilon = 0°$ and far off the volute tongue at $\varepsilon = 240°$. A noisy appearance of the contour plots shows that statistical convergence has not been achieved completely yet for the SAS simulations. Nevertheless, since preliminary evaluations of different numbers of ensembles did not show a significant change of the relevant pattern, a sufficiently reliable assessment of the SAS flow angle in comparison to both SST results and measurement data is assumed to be feasible. No significant mesh dependence is observed for the flow angle. The flow at the volute tongue at $\varepsilon = 0°$ is dominated by an outflow around the blade trailing edge (near x/t=0 and x/t=1) and a secondary back flow near the sidewalls (y/b=0 and y/b=1) at mid-channel (x/t=0.5). This behaviour can be seen on both grids for SAS and SST. Thus, SAS and SST results can be considered equally close to measured flow angle in terms of spatially resolved and ensemble-averaged flow angle.

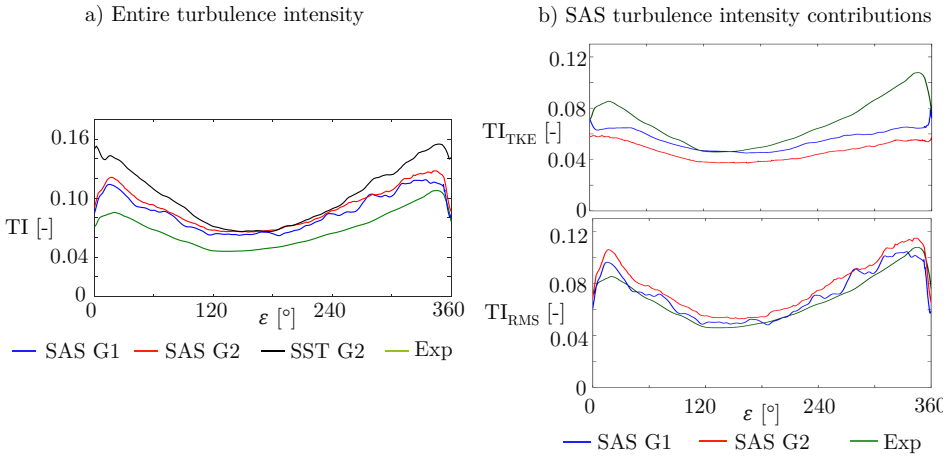


Figure 6. Axially and time-averaged TI for simulation and measurement data (a) and contributions to TI for the SAS model (b).

Regarding the TI results, for the SST model a significant overestimation can be seen in the immediate vicinity of the volute tongue at $\varepsilon = 0°$ (*c. f.* Fig. 5(b), left column) that gets even more pronounced with grid refinement. The TI is reduced significantly with the SAS model and thus matches the experimental data much better, particularly for the finer grid G2, what is a distinctive improvement of TI prediction. In the other volute regions in circumferential direction, here represented by the $\varepsilon = 15°$ position, TI is over-predicted systematically by both models.

Fig. 6(a) shows the axially and time-averaged TI distribution *vs.* circumferential volute position ε, where this behaviour is reflected over the entire spectrum of ε. In Fig. 6(a) TI for the simulation has been evaluated by Eq. 2. In order to provide a more comprehensive illustration of TI by the SAS model, Fig. 6(b) shows the two contributions to TI according to Eqs. 3 and 4. The amount of TI_{RMS} is higher for G2 which shows the scale-adaptive behaviour of the model, as with grid refinement the amount of resolved turbulent fluctuations increases. TI_{TKE} is lower for grid G2. Interestingly, the discrepancy of total TI to measurement data cannot be reduced by grid refinement. We assume that, since G1 and G2 are still in the range of wall functions, this may be a consequence of the impeller wake flow, and a wall-adjacent grid refinement towards low-Reynolds wall treatment in the impeller may improve the results, which will be the subject of further investigations.

6 Conclusions

Deviations of the ensemble-averaged flow angle to measurement data are small for both turbulence models, so that the SST model is found to be sufficiently accurate for mean flow prediction, *e. g.* mean flow angle. A different conclusion can be drawn for turbulent quantities, such as TI. A significant improvement of TI prediction near the volute tongue is found for the SAS model compared to SST results. Afar from the volute tongue, total TI is overestimated with both turbulence models. In particular the fact that a grid refinement

does not reduce the entire TI level (TI_{RMS} and TI_{TKE}) for the SAS model will be the subject of further studies.

The results presented on the 7 blade specific speed $n_S = 26$ min^{-1} pump are an extract of our pump simulation efforts and are representative of our turbulence modelling approach. A highly unsteady flow field that demands accurate turbulence predictions is also present in extremely low specific speed pumps ($n_S < 12$ min^{-1}) and special designs, such as single-blade pumps or positive displacement pumps. The findings presented here will also be transferred to these pump types.

Although we could show that the SAS model in its present form shows an improvement over the SST model even on a moderate grid resolution, a further improvement to obtain also a quantitative agreement to measurement data will be done in further studies by a thorough exemplary assessment *vs.* reference simulation data in terms of well-resolved LES results.

Acknowledgements

The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputers JURECA and JUWELS at Jülich Supercomputing Centre (JSC). The authors also gratefully acknowledge the financial support by the "Forschungskuratorium Maschinenbau e.V.". This project is funded by the Federal Ministry for Economic Affairs (BMWi).

References

1. F. R. Menter, M. Kuntz, and R. Langtry, *Ten Years of Industrial Experience with the SST Turbulence Model*, in Proceedings of the 4th International Symposium on Turbulence, Heat and Mass Transfer, Begell House, 625–632, 2003.
2. N. Casimir, Z. Xiangyuan, G. Ludwig, and R. Skoda, *Assessment of statistical eddy-viscosity turbulence models for unsteady flow at part and overload operation of centrifugal pumps*, in Proceedings of 13th European Conference on Turbomachinery Fluid dynamics & Thermodynamics (ETC13), Lausanne, ETC2019-047, 2019.
3. R. K. Byskov, C. B. Jacobsen, and N. Pedersen, *Flow in a Centrifugal Pump Impeller at Design and Off-Design Conditions – Part II : Large Eddy Simulations*, Journal of Fluids Engineering **125**, 73–83, 2003.
4. C. Kato, H. Mukai, and A. Manabe, *Large-Eddy Simulation of Unsteady Flow in a Mixed-Flow Pump*, International Journal of Rotating Machinery **9**, 345–351, 2003.
5. A. Posa, A. Lippolis, and E. Balaras, *Large-eddy simulation of a mixed-flow pump at off-design conditions*, Journal of Fluids Engineering **137**, 2015.
6. A. Posa, A. Lippolis, R. Verzicco, and E. Balaras, *Large-eddy simulations in mixed-flow pumps using an immersed-boundary method*, Computers & Fluids **47**, 33–43, 2011.
7. J. Shen, Y. Li, Z. Liu, and X. Tang, *Turbulent flow and pressure fluctuation prediction of the impeller in an axial-flow pump based on LES*, 6th Int. Conf. on Pumps and Fans with Compressors and Wind Turbines; IOP Conference Series: Material Science and Engineering **52**, 032015, 2013.

8. Q. Si, J. Yuan, S. Yuan, W. Wang, L. Zhu, and G. Bois, *Numerical Investigation of Pressure Fluctuation in Centrifugal Pump Volute Based on SAS Model and Experimental Validation*, Advances in Mechanical Engineering **6**, 972081, 2014.
9. T. Tokay and S. Constantinescu, *Validation of a large eddy simulation model to simulate flow in pump intakes of realistic geometry*, Journal of Hydraulic Engineering **132**, 1303–1315, 2006.
10. W. Wang and Y. Wang, *Analysis of inner flow in low specific speed centrifugal pump based on LES*, Journal of Mechanical Science and Technology **27**, 1619–1626, 2013.
11. Y. Wu, X. Tang, F. Wang, *An improved large eddy simulation of two-phase flows in a pump impeller*, Acta Mech. Sin. **23**, 635, 2007.
12. W. Zhang, Y. Yu, and H. Chen, *Numerical Simulation of Unsteady Flow in Centrifugal Pump Impeller at Off-Design Condition by Hybrid RANS/LES Approaches*, in High Performance Computing and Applications, Lecture Notes in Computer Science **5938**, Springer, 571–578, 2010.
13. J. Fröhlich, *Large Eddy Simulation turbulenter Strömungen: mit 14 Tabellen*, Lehrbuch Maschinenbau, Teubner, Wiesbaden, 1. Edition, 2006.
14. S. B. Pope, *Turbulent Flows*, Cambridge University Press, 2000.
15. J. C. Rotta, *Turbulente Strömungen: eine Einführung in die Theorie und ihre Anwendung*, B. G. Teubner, Stuttgart, 1972.
16. F. R. Menter and Y. Egorov, *The Scale-Adaptive Simulation Method for Unsteady Turbulent Flow Predictions. Part 1: Theory and Model Description*, Flow, Turbulence and Combustion **85**, 113–138, 2010.
17. A. Lucius and G. Brenner, *Unsteady CFD simulations of a pump in part load conditions using scale-adaptive simulation*, International Journal of Heat and Fluid Flow **31**, 1113–1118, 2010.
18. J. Schiffer, C. Bodner, H. Jaberg, S. Korupp, and L. Runte, *Performance analysis of a single-blade impeller pump based on unsteady 3D numerical simulation*, in Proceedings of the 3rd International Rotating Equipment Conference, Düsseldorf, 193–203, 2016.
19. R. I. Issa, *Solution of the implicitly discretised fluid flow equations by operator-splitting*, Journal of Computational Physics **62**, 40–65, 1986.
20. S. V. Patankar and D. B. Spalding, *A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows*, in Numerical Prediction of Flow, Heat Transfer, Turbulence and Combustion, Elsevier, 54–73, 1983.
21. A. Mehdizadeh, H. Foroutan, G. Vijayakumar, and A. Sadiki, *A new formulation of scale-adaptive simulation approach to predict complex wall-bounded shear flows*, Journal of Turbulence **15**, 629–649, 2014.
22. F. Nicoud and F. Ducros, *Subgrid-Scale Stress Modelling Based on the Square of the Velocity Gradient Tensor*, Flow, Turbulence and Combustion **62**, 183–200, 1999.
23. B. Krank, M. Kronbichler, and W. A. Wall, *Direct Numerical Simulation of Flow over Periodic Hills up to Re_H = 10,595*, Flow, Turbulence and Combustion **101**, 521–551, 2018.
24. P. Hergt, S. Meschkat, and B. Stoffel, *The flow and head distribution within the volute of a centrifugal pump in comparison with the characteristics of the impeller without casing*, Journal of Computational and Applied Mechanics **5**, 275–285, 2004.

25. S. Meschkat and B. Stoffel, *The local specific head at different circumferential positions in a volute casing centrifugal pump in comparison to the characteristic curve of the single rotor*, in Proceedings of the Hydraulic Machinery Systems 21st IAHR Symposium, Lausanne, 2002.
26. A. Pesch, N. Casimir, M. Hundshagen, and R. Skoda, *Assessment of Scale Adaptive Turbulence Modeling in Centrifugal Pump Simulations*, 14th OpenFOAM Workshop, Duisburg, 2019.
27. R. F. Warming and R. M. Beam, *Upwind Second-Order Difference Schemes and Applications in Aerodynamic Flows*, AIAA Journal **14**, 1241, 1976.
28. B. Van Leer, *Towards the Ultimate Conservative Difference Scheme. V. A Second-Order Sequel to Godunov's Method*, Journal of Computational Physics **32**, 101–136, 1979.
29. M. Strelets, *Detached eddy simulation of massively separated flows*, in 39th Aerospace Sciences Meeting and Exhibit, 2001.
30. M. Beaudoin and H. Jasak, *Development of a Generalized Grid Mesh Interface for Turbomachinery simulations with OpenFOAM*, Open source CFD International conference, Berlin, 2008.

Sub-Grid Scale Modelling at Scale with Deep Learning and up to 60 Billion Degrees of Freedom

Mathis Bode[1], **Dominik Denker**[1], **Jenia Jitsev**[2], **and Heinz Pitsch**[1]

[1] Institute for Combustion Technology, RWTH Aachen University, 52056 Aachen, Germany
E-mail: {m.bode, d.denker, h.pitsch}@itv.rwth-aachen.de

[2] Jülich Supercomputing Centre, Institute for Advanced Simulation, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: j.jitsev@fz-juelich.de

This work presents fully resolved direct numerical simulations (DNSs) of a turbulent reactive planar temporally non-premixed jet configuration with up to 60 billion degrees of freedom. As scalar mixing is of utmost importance for this kind of configuration, a novel deep learning (DL) approach in the context of large-eddy simulation is presented which results in predictive mixing statistics on underresolved grids. The usability of the mixing model is approved by applying it to the DNS data. Furthermore, node performance measurements for the training of the DL networks are shown for different computing clusters.

1 Introduction

Turbulent fluid flows pose some of the most difficult and fundamental problems in classical physics as turbulence is a complex, strongly non-linear, multi-scale phenomenon.[1] Consequently, the prediction of the statistics of fluctuating velocity and scalar fields and the development of models for a precise statistical prediction of these fields even in scale-resolved simulations is one of the main challenges in turbulence research.[2, 3] An important example is the prediction of scalar mixing in turbulent flows as it controls many processes of industrial relevance.

One often used modelling approach for turbulent flows is large-eddy simulation (LES). LES solves for the larger, flow-dependent scales of the flow and models all scales below a particular filter width.[4, 5] By assuming that the smaller, unresolved scales reveal certain universal features and decouple from the larger non-universal scales, models for LES can be built from relatively simple, semi-empirical algebraic relations. A systematic approach for developing such models is to perform fully resolved direct numerical simulations (DNSs), filter the resulting data with a given filter kernel, and find relations between the filtered data and the original DNS results. The objective of the present work is to develop a LES model for scalar mixing in turbulent flows based on deep learning (DL). This data-driven approach reconstructs the subfilter statistics, which can be used to close the filtered Navier-Stokes equations.

DL is one of the most emerging research topics in the age of big data. Examples of DL applications include image processing,[6–9] voice recognition,[10] or website customisation.[11] One reason for this is the continued growth of computational power (especially GPUs). Another reason is the availability of exceptionally large labelled experimental datasets. Data-driven methods and DL have also become more popular over the last years in the field of fluid mechanics including turbulence research.[12–15] However, simple networks or small and artificial datasets have often limited these studies in the past.

The remainder of this article is organised as follows. Sec. 2 describes the used reactive datasets. In Sec. 3, details about the data-driven reconstruction methodology are given, and results for inert data are discussed. The scalar mixing LES model is applied to the reactive data in Sec. 4. The paper finishes with conclusions.

2 Reactive DNS

Combustion in turbulent flows can be classified in several regimes based on the level of interaction of the turbulent flow field and the combustion chemistry. These combustion regimes have severe implications to the choice of adequate combustion models employed in the simulation of real world engineering applications. However, in the limit of intense scale interaction between turbulence and flame scales, no satisfactory combustion model exists. In this combustion regime, the well defined and understood flame structure is disintegrated by turbulent eddies. Unfortunately, the lack of suitable combustion models in this regime limits the quality of the prediction of LESs of novel and promising combustion processes, which rely on low temperatures and consequently feature slow combustion time scales. As both turbulence and chemistry scales must be fully resolved for an appropriate investigation of these combustion conditions, the conduction of high fidelity direct numerical simulations (DNS) is inevitable. To this end, a series of five DNSs of planar temporally evolving non-premixed jet flames was performed on the supercomputers JURECA, JUQUEEN, and JUWELS. The DNSs feature methane as fuel and air as oxidiser. Different levels of dilution of the fuel, as well as increasing turbulence intensities, allow for a precise determination of combustion regime boundaries. The configurations of the DNSs, as well as important non-dimensional numbers are summarised in Tab. 1. As simulating real world engineering configurations is currently impossible due to the immense computational resource requirements, the idealised configuration of the planar temporally evolving jet was chosen for maximising the flame surface as well as for the ease of obtaining statistics. In this configuration, the fuel stream is situated in stream-wise centre slab of the domain and the oxidiser is moving in the opposing direction at the upper and lower part of the domain. Two highly turbulent flame fronts form in the shear layers between the fuel and oxidiser streams.

	Low Re low dilution case	Low Re high dilution case	Intermediate Re case	High Da case	High Re case
$\mathrm{Re}_{\mathrm{jet},0}$	4500	4500	6000	6000	10 000
Da_τ	0.125	0.150	0.150	0.450	0.150
Z_{st}	0.20	0.45	0.45	0.45	0.45
$n_{\mathrm{gridpoints}}\,[10^9]$	0.4	0.3	0.6	1.6	1.2
DOF $[10^9]$	15	11	23	60	45

Table 1. Important non-dimensional numbers and numerical parameters of the DNSs of planar temporally evolving non-premixed jet flames conducted on the supercomputers JUQUEEN, JURECA, and JUWELS. The initial jet Reynolds number $\mathrm{Re}_{\mathrm{jet},0}$ indicates the ratio of the largest to the smallest turbulent length scales, the Damköler number Da_τ indicates the inverse level of flame extinction and the stoichiometric mixture fraction Z_{st} reflects the level of the dilution of the fuel stream. DOF indicates the degrees of freedom of the system of equations solved in each time step.

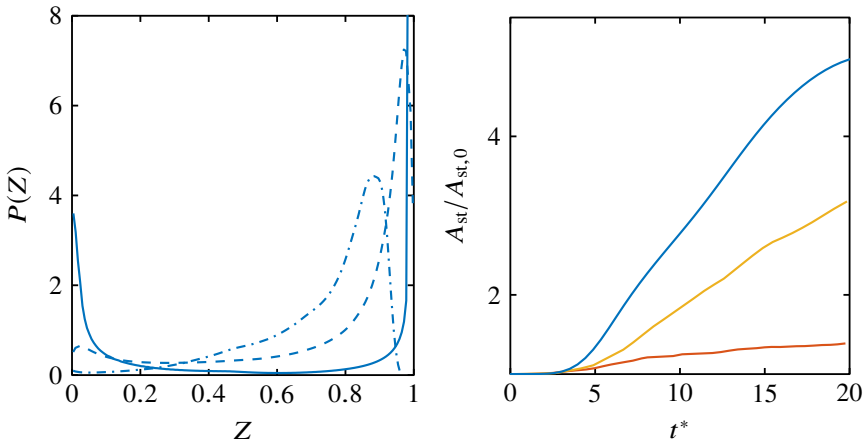


Figure 1. Left: PDF of the mixture fraction $P(Z)$ in the centre slab of the high Re case for three different time steps (solid line: early time step, dashed line: intermediate time step, dashed-dotted line: last time step). Right: temporal evolution of the normalised flame surface area $A_{\mathrm{st}}/A_{\mathrm{st},0}$ (red line: Low dilution case, yellow line: intermediate Re case, blue line: high Re case).

As fuel and oxidiser are initially separated and a combustable mixture must be formed prior to combustion, applications using non-premixed combustion rely heavily on turbulent mixing to achieve high efficiencies. Therefore, knowledge of the turbulent mixing process is of upmost importance for the reliable modelling of non-premixed combustion. The turbulent mixing process is illustrated by the example of the probability density function (PDF) of the mixture fraction $P(Z)$ in the centre slab of the domain of the High Re case in Fig. 1 (left). In the early time step, $P(Z)$ resembles a double-δ PDF with peaks at $Z = 0$ (pure oxidiser) and $Z = 1$ (pure fuel), which highlights the initial separation of the two streams. In later time steps, the PDF changes shape as turbulent transport continuously causes fuel and oxidiser to mix. In the final time step of the simulation, neither pure fuel nor pure oxidiser is present in the central slab.

The consequences of increased turbulent mixing for the combustion is demonstrated by the temporal evolution of the normalised flame surface area $A_{\mathrm{st}}/A_{\mathrm{st},0}$, with the initial flame surface area $A_{\mathrm{st},0}$, in Fig. 1 (right). A_{st} quickly rises with higher Reynolds numbers and, consequently, the overall consumption of fuel in the domain is increased significantly as higher turbulence intensities are reached.

The configuration of the DNS and the highly turbulent nature of the combustion is illustrated in Fig. 2. The exceedingly corrugated iso-surface of the stoichiometric mixture fraction is shown for the highest turbulence intensity case. The stoichiometric mixture fraction indicates the position of the reaction zone in non-premixed flames. Additionally, the local value of the temperature is displayed to show spots of local extinction.

The DNS were performed solving the reactive, unsteady Navier-Stokes equations in the low Mach number limit using the in-house solver CIAO.[16, 17] The momentum equations are spatially discretised with a fourth-order scheme and time advancement is performed using a semi-implicit Crank-Nicholson time integration.[18] The transport of species mass fractions is described using the Hirschfelder and Curtiss approximation to the diffusive

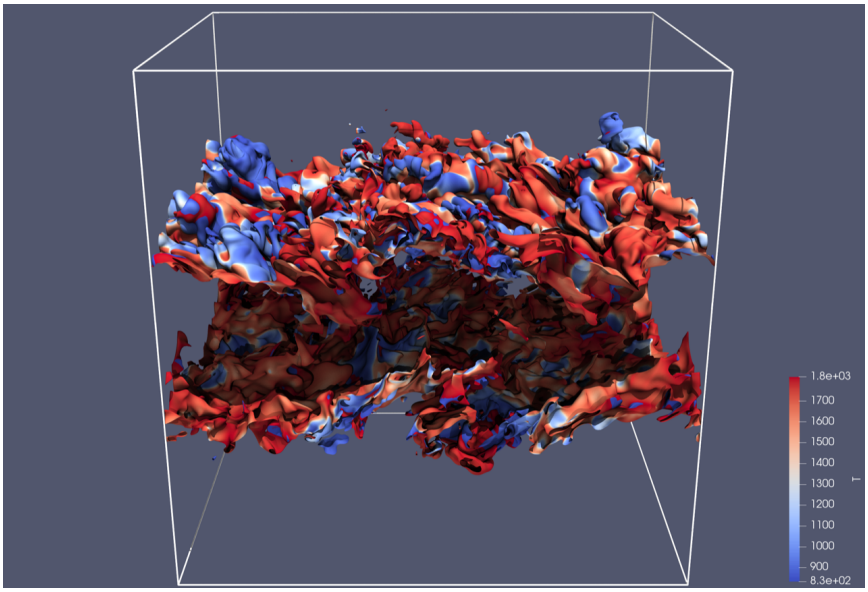


Figure 2. Iso-surface of the stoichiometric mixture fraction in the High Re case. The colour of the iso-surface indicates the local value of the temperature.

fluxes together with a velocity-correction approach for mass conservation. The temperature and species equations are advanced by introducing the symmetric operator split of Strang.[19] The chemistry operator uses a time-implicit backward difference method.[20] The chemistry is included using finite rate chemistry employing a mechanism for the oxidation of methane comprising 30 species and 102 reactions capable of capturing the anticipated effects of localised extinction and re-ignition.

To resolve all scales, from the detailed flame structure to the integral scales of the turbulent flow fields, computational meshes containing 0.3×10^9 to 1.6×10^9 grid points were employed. This results in a system of equations with up to 60 billion degrees of freedom which must be solved for each of the approximately $11\,000$ time steps needed for the transition from an initial laminar solution to a fully turbulent flame. The largest case in this series, the High Da case, consumed a total of 42 million CPU-h on the supercomputer JUQUEEN.

3 Modelling

The objective of this work is to move beyond conventional models for LES and develop a data-driven approach, which is able to predict instantaneous, fully-resolved three-dimensional flow realisations including scalar fields, such as the mixture fraction, by knowing solely the coarse-grained data fields. This approach generalises deconvolution techniques as no explicit knowledge of the filter kernel is required.

A neural network called turbulence reconstruction generative adversarial network (TR-GAN) is used in this work and predominantly motivated by the previous single image

super-resolution (SISR) works. A pioneer work for SISR was the super-resolution convolutional neural network (SRCNN) designed by Dong *et al.*[6] Great efforts towards SISR had been made through the community since SRCNN first emerged. Yet most of these SISR networks showed the deficiency of producing over-smoothed results, owing to focusing on the peak-signal-to-noise-ratio (PSNR) as the primary metrics. Resultantly, the high frequency details occurring in the ground truth could not be adequately reconstructed. This makes such networks especially fatal for turbulence applications, as the small-scale (high frequency) structures at high wave numbers are exactly what is intended to be restored. Amelioration regarding the PSNR-based SISR networks was elucidated by Johnson *et al.*,[21] who presented the concept "perceptual loss", *i. e.* a criterion to optimise the performance in the VGG-features space that substitutes the PSNR which optimised in the pixel space.

The other fundamental framework named generative adversarial network (GAN) was presented by Ref. 22. A GAN is composed of two models, a generator that captures the data distribution and generates new data, and a discriminator that learns to distinguish whether a sample stems from the original data distribution (genuine) or the generator (fake). During training, the generator learns to produce samples that are indistinguishable for the discriminator, while the discriminator learns to more accurately judge the genuineness.

For this work, the state-of-art ESRGAN was adapted to three-dimensional (3-D) turbulence data and a problem-specific, physically-based loss functions was introduced. The resulting network is called turbulence reconstruction GAN (TRGAN). The TRGAN is supposed to recuperate the fully-resolved DNS results from coarse turbulence data, such as LES data.

The perceptual loss proposed for the ESRGAN based on VGG-feature space is apparently not as suitable for the turbulence data as the geometrical features from VGG19 are not representative for turbulent flows. Hence, a new formulation for the cost function was developed inspired by physical flow constraints.

Before training the TSRGAN as a combined model, the generator is pretrained with root-mean-square error (RMSE) due to the complexity of the residual-in-residual-dense-block (RRDB). For the combined model, the loss function for reconstructing velocity fields is proposed as

$$l = \beta_1 l_{\mathrm{RADG}} + \beta_2 l_{\mathrm{pixel}} + \beta_3 l_{\mathrm{gradient}} + \beta_4 l_{\mathrm{continuity}} \tag{1}$$

with $\beta_1, \beta_2, \beta_3$, and β_4 being coefficients weighting the different loss term contributions. l_{RADG} is the 'realistic average' discriminator/generator loss, which is the accuracy feedback between discriminator and generator as given by Wang *et al.*[7] For a scalar ϕ, the pixel loss l_{pixel} is defined as

$$l_{\mathrm{pixel}} = \mathrm{MSE}(\phi^{\mathrm{predicted}}, \phi^{\mathrm{DNS}}) \tag{2}$$

The mean-square error (MSE) operator is given by

$$\mathrm{MSE}(\{\cdot\}_1, \{\cdot\}_2) = \frac{1}{N_{\mathrm{samples}}} \sum_{i=1}^{N_{\mathrm{samples}}} (\{\cdot\}_1^i - \{\cdot\}_2^i)^2 \tag{3}$$

with N_{samples} as the number of all samples, *i. e.* the total number of grid points of the reconstructed field. If the MSE operator is applied on tensors including vectors, it is applied

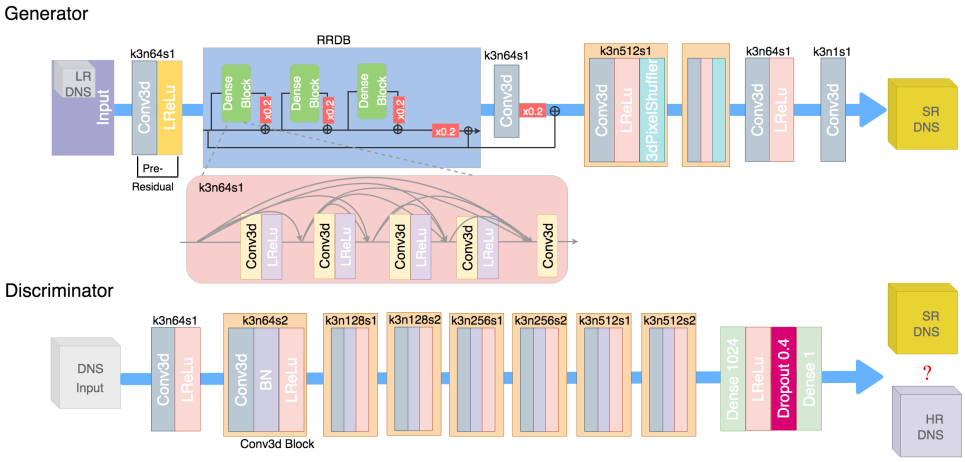


Figure 3. TRGAN network structure. Image from Bode *et al.*[13]

to all elements separately. Afterwards the resulting tensor is mapped into a scalar using the 1-norm. The gradient loss l_{gradient} is defined as

$$l_{\text{gradient}} = \text{MSE}(\nabla\phi^{\text{predicted}}, \nabla\phi^{\text{DNS}}) \tag{4}$$

with ∇ as del operator. $l_{\text{continuity}}$ is the continuity loss, which enforces the continuity equation in the reconstructed field and reads

$$l_{\text{continuity}} = \text{MSE}(\nabla \cdot \mathbf{u}^{\text{predicted}}, \mathbf{0}) \tag{5}$$

with $\mathbf{u}$ as velocity vector. The TRGAN network is graphically presented in Fig. 3. Note that differently from the state-of-art SR networks, the TRGAN does not involve up-/ or downsampling layers. The high-/ and low-resolution turbulence data refer to the energy spectral property of the discussed turbulence, not the trivial image resolution.

The TRGAN was implemented by using Keras API with TensorFlow backend, and the training was performed on the JURECA GPU nodes due to the large size of the DNS dataset. The results for a reconstructed velocity and scalar field are shown in Figs. 4 and 5 for an inert case. Homogeneous isotropic turbulence data were used, and the quality of the DL closure can be seen.

Training with large data, such as the reactive DNS dataset used in this work, is not possible without computing clusters. The GPU performance of JURECA, JUWELS (both at Jülich Supercomputing Centre, Forschungszentrum Jülich), and Claix18 (at IT Center, RWTH Aachen University) and their scaling performance are depicted in Fig. 6. Here, the numerical values give the average number of subboxes with dimension $16 \times 16 \times 16$, which were trained to the TRGAN within one minute. The better performance of the Tesla V100 GPUs (JUWELS, CLAIX18) compared to the Tesla K80 GPUs (JURECA) is obvious.

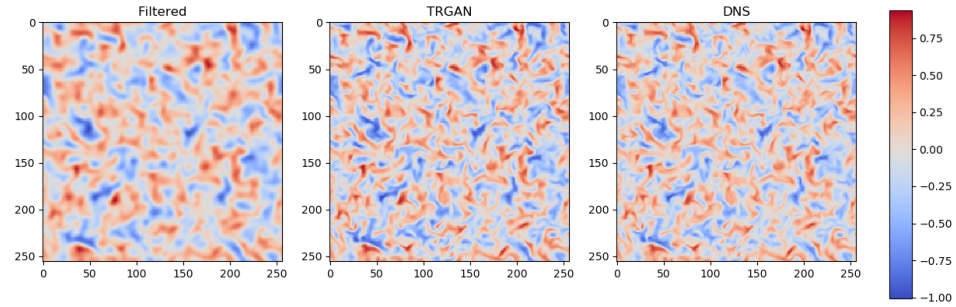


Figure 4. Velocity reconstruction.

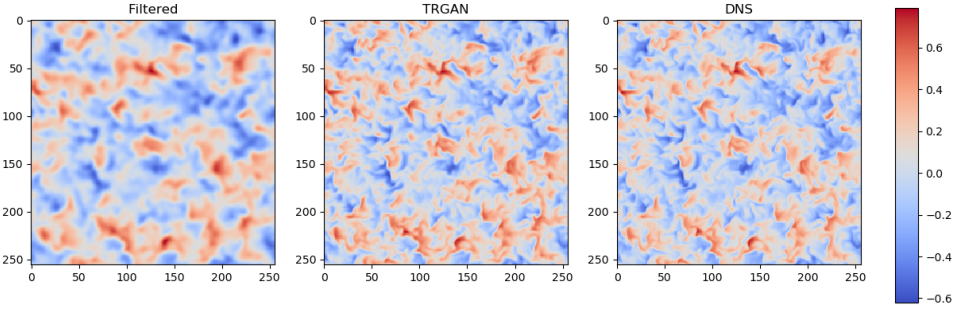


Figure 5. Passive scalar reconstruction.

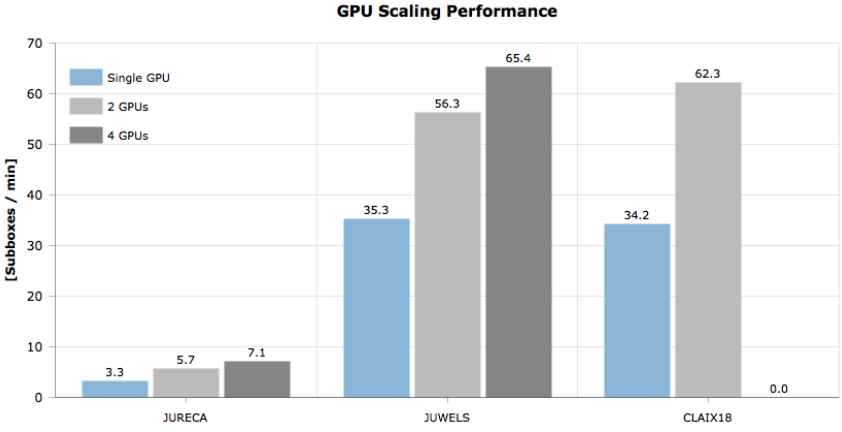


Figure 6. GPU performance.

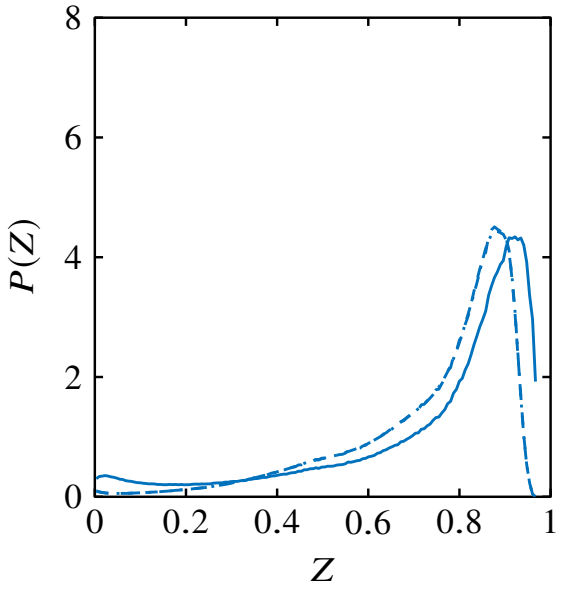


Figure 7. PDF of the mixture fraction $P(Z)$ in the centre slab of the high Re case for the last time step from Fig. 1 (solid line: no model, dashed line: DL model, dashed-dotted line: DNS).

4 Application

In order to validate the data-driven DL modelling approach presented in Sec. 3, the high Reynolds number DNS data of an early time step are filtered and interpolated to a coarser mesh with only one third of the number of points per direction compared to the corresponding DNS. Using this coarser mesh, the mixing results shown in Fig. 1 are recomputed with and without the DL model. The result is compared to the DNS result. As can be seen in Fig. 7, the DNS and modelled results match very well while the underresolved simulation without model mixes too slow.

5 Conclusions

Fully resolved DNSs of a turbulent reactive non-premixed planar temporal jet configuration and a data-driven DL modelling approach are presented in this work. It is shown how the combination of computationally expensive simulations and novel LES modelling methods can be used to accurately predict scalar mixing even on underresolved grids. An estimation for training cost is given.

Acknowledgements

The authors gratefully acknowledge the computing time granted for the projects JHPC09, JHPC22, and JHPC55 by the JARA Vergabegremium and provided on the JARA Partition part of the supercomputers JUQUEEN and JURECA at Forschungszentrum Jülich. In

addition, the authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding project JHPC09 by providing computing time on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC). Furthermore, the authors want to thank Michael Gauding, Zeyu Lian, and Antonio Attili for their help.

References

1. D. Ruelle and F. Takens, *On the nature of turbulence*, Les rencontres physiciens-mathématiciens de Strasbourg-RCP25 **12**, 1–44, 1971.
2. U. Piomelli, *Large-eddy simulation: achievements and challenges*, Progress in Aerospace Sciences **35**, 335–362, 1999.
3. A. R. Kerstein, *Turbulence in combustion processes: Modeling challenges*, Proceedings of the Combustion Institute **29**, 1763–1773, 2002.
4. M. Germano, U. Piomelli, P. Moin, and W. H. Cabot, *A dynamic subgrid-scale eddy viscosity model*, Physics of Fluids A: Fluid Dynamics **3**, 1760–1765, 1991.
5. A. Leonard, *Energy cascade in large-eddy simulations of turbulent fluid flows*, in Turbulent Diffusion in Environmental Pollution, Advances in Geophysics **18**, 237–248, 1975.
6. C. Dong, C. C. Loy, K. He, and X. Tang, *Learning a deep convolutional network for image super-resolution*, in Computer Vision, Lecture Notes in Computer Science **8692**, Springer, 184–199, 2014.
7. X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao, and X. Tang, *ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks*, Lecture Notes in Computer Science **11133**, 63–79, 2019.
8. H. Greenspan, B. Van Ginneken, and R. M. Summers, *Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique*, IEEE Transactions on Medical Imaging **35**, 1153–1159, 2016.
9. N. Wang and D.-Y. Yeung, *Learning a deep compact image representation for visual tracking*, in Advances in Neural Information Processing Systems **26**, 809–817, 2013.
10. G. Hinton, L. Deng, D. Yu, G. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, B. Kingsbury *et al.*, *Deep neural networks for acoustic modeling in speech recognition*, IEEE Signal processing magazine **29**, 82–97, 2012.
11. M. Langheinrich, A. Nakamura, N. Abe, T. Kamba, and Y. Koseki, *Unintrusive customization techniques for web advertising*, Computer Networks **31**, 1259–1272, 1999.
12. M. Bode, M. Gauding, J. H. Göbbert, B. Liao, J. Jitsev, and H. Pitsch, *Towards Prediction of Turbulent Flows at High Reynolds Numbers Using High Performance Computing Data and Deep Learning*, Lecture Notes in Computer Science **11203**, 614–623, 2018.
13. M. Bode, M. Gauding, K. Kleinheinz, and H. Pitsch, *Deep learning at scale for subgrid modeling in turbulent flows*, 2019, arXiv:1910.00928 [physics.comp-ph].
14. J. N. Kutz, *Deep learning in fluid dynamics*, Journal of Fluid Mechanics **814**, 1–4, 2017.
15. R. Maulik and O. San, *A neural network approach for the blind deconvolution of turbulent flows*, Journal of Fluid Mechanics **831**, 151–181, 2017.

16. M. Bode, M. Davidovic, and H. Pitsch, *Towards Clean Propulsion with Synthetic Fuels: Computational Aspects and Analysis*, in High Performance Computing in Science and Engineering '18, W. Nagel W., D. Kröner D., M. Resch M. (Editors), Springer, 185–207, 2018.
17. M. Bode, N. Collier, F. Bisetti, and H. Pitsch, *Adaptive chemistry lookup tables for combustion simulations using optimal B-spline interpolants*, Combustion Theory and Modelling **23**, 674–699, 2019.
18. O. Desjardins, G. Blanquart, G. Balarac, and H. Pitsch, *High Order Conservative Finite Difference Scheme for Variable Density Low Mach Number Turbulent Flows*, J. Comp. Phys. **227**, 7125–7159, 2008.
19. G. Strang, *On the construction and comparison of difference schemes*, SIAM J. Numer. Anal. **5**, 506–517, 1968.
20. A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward, *SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers*, ACM Transactions on Mathematical Software (TOMS) **31**, 363–396, 2005.
21. J. Johnson, A. Alahi, and L. Fei-Fei, *Perceptual losses for real-time style transfer and super-resolution*, in European Conference on Computer Vision, Springer, 694–711, 2016.
22. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, *Generative adversarial nets*, in Advances in Neural Information Processing Systems **27**, 2672–2680, 2014.

Size, Shape, Density and Spectral Analysis of Nanoparticles by Analytical Ultracentrifugation Using UltraScan

Maximilian J. Uttinger, Simon E. Wawra, Johannes Walter, and Wolfgang Peukert

Institute of Particle Technology, Interdisciplinary Center for Functional Particle Systems,
Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
E-mail: wolfgang.peukert@fau.de

1 Introduction

Experimental data of analytical ultracentrifugation (AUC) are governed by sedimentation and diffusion of (macro-)molecules and particles in solution. While the sedimentation properties define the radial position of the measured sedimentation boundaries, information on the diffusion is included in the boundary broadening.[1–3] Both effects can be tracked via different optical systems, including interference, absorption and fluorescence detection. The radial and temporal evolution of the optical signal, which is related to the solute concentration, can then be analysed with well-established finite element solutions of Lamm's equation. The results from this multidimensional characterisation provide the possibility of analysing molecular mass distributions, size distributions or information on core-shell structures.

UltraScanIII provides the highest resolution analysis possible today for AUC experiments, in part due to the high-performance computing (HPC)-enabled methodology employed in the optimisation algorithms. This capability offers unrivalled resolution and allows the investigator to more accurately identify shape and molecular weight parameters for all solutes present in a sample. The efforts to date have resulted in the development of three software components to provide high-resolution data analysis of AUC experiments:

1. A multi-platform graphical user interface desktop module;

2. A highly parallelised and scalable computational backend implemented on supercomputers. This module tackles finite element simulations and computationally demanding parallel optimisation algorithms and high-resolution modelling methods to solve the inverse problem of extracting molecular parameters from experimental data, and

3. The UltraScan Laboratory Information Management Systems (USLIMS) web interface for managing experimental data and submitting and managing computing jobs on high-performance computational backend resources.

The algorithms and performance characteristics of UltraScan codes are discussed in published manuscripts detailing the methods: Solutions of the Lamm equation for non-interacting systems and for systems involving arbitrary reaction mechanisms,[4–6] optimisation methods involving the 2-dimensional spectrum analysis (2DSA)[7] and genetic algorithms (GA).[8] Global analysis is possible in both optimisation approaches as documented

in Ref. 9. Furthermore, Monte Carlo (MC) analysis can be employed to statistically evaluate the confidence intervals for the determined parameters. Recently, the parametrically spectrum analysis (PCSA) has been introduced as a further versatile tool for AUC data evaluation.[10] The accuracy and efficiency of the parallel implementation of these approaches has been documented and compared to traditional methods and has been determined to be far superior.[11]

In 2014, a major upgrade to the multiwavelength absorption (MWL) analysis infrastructure was implemented[12] by Borries Demeler and co-workers. In 2015, the algorithms were perfected[13] for maximum efficiency. The message-passing interface (MPI)-based UltraScan analysis software and the USLIMS are both deployed at the Jülich Supercomputing Centre, and the former is deployed on the JURECA supercomputer. The software is maintained remotely by the UltraScan development team at the University of Lethbridge at Lethbridge, Canada. Job submission is accomplished via the UNICORE middleware. The excellent assistance from the staff at Jülich significantly improved the stability and the reliability of submitting and managing large parallel analysis computing jobs via the USLIMS web interface. The function of the high-performance computation resources for large parallel analysis jobs on the JURECA cluster have now reached maturity and will be an excellent tool for research efforts going forward. Complimenting these HPC analysis resource developments, hardware advances have also increased the data quality from MWL-AUC. For a long time, the data quality of the MWL detectors has been insufficient for high-resolution evaluation in the UV, the region of the spectrum most amenable for analysing biopolymers, synthetic polymers and many semiconductor nanoparticles. Improvements of the UV performance of the detection systems now allows for high quality evaluations of biopolymers[14] and nanoparticles.[14, 15] The evaluation of biopolymers is the most widespread use of AUC but there is a significantly increasing demand for application in nanoparticle science. The capability of the MWL hardware to measure biopolymer systems and nanoparticles absorbing in the deep UV makes the increasing use of this technology highly promising.

2 Analysis of Polydisperse Core-Shell Nanoparticles

An example for the application of high resolution analysis in the field of particle technology is the comprehensive characterisation of core-shell nanoparticles. Unlike biopolymers, particle systems exhibit considerable polydispersity in terms of size. Especially for nanoparticles this also leads to dispersity in terms of density, because solvation and bound ligand or stabiliser molecules considerably change the particle's density at the lower nanoscale, which in turn affect its hydrodynamic properties. In order to obtain reliable size information on the particle core, it is therefore crucial to know the effective particle density as well as the size of the individual fractions of the core-shell system. With that information being available, the underlying core size distribution can be reliably determined. Therefore, several methodologies were developed in the past to address the size dependent change in density when measuring sedimentation and diffusion coefficients in AUC.[4, 5] Recently, the 2DSA was extended in such way that it also allows for variations in the density of particles of known shape by spanning a custom 2-dimensional grid (CG-MC) of sedimentation coefficients and effective partial specific volumes. The latter is equivalent to the inverse of the effective density of a core-shell particle.[15] In general, this allows determination of the

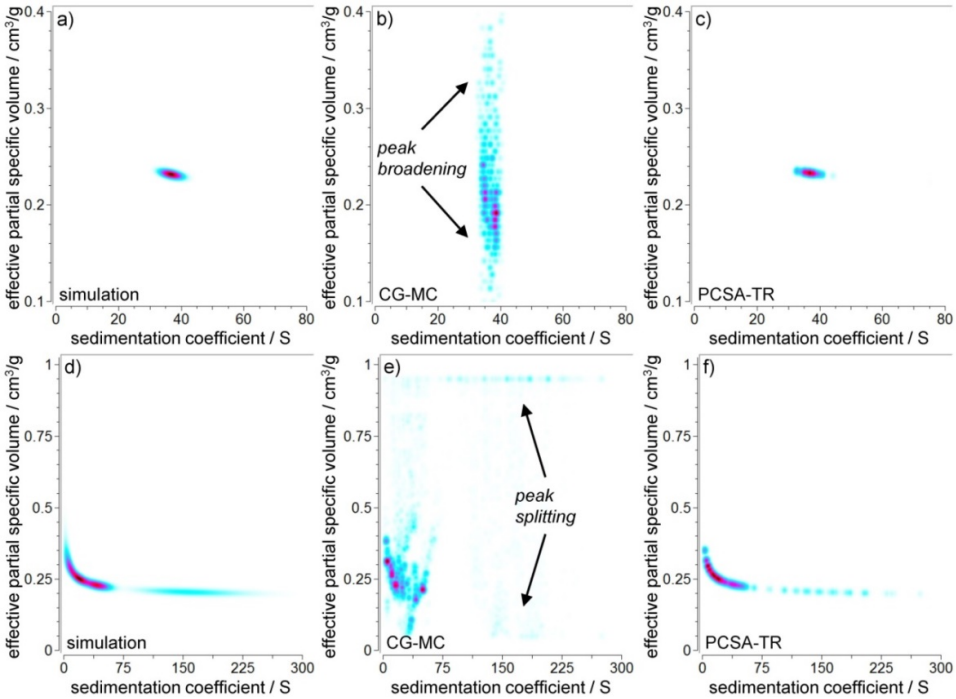


Figure 1. Simulation model for narrow particle size distributions (a-c) and simulation model for polydisperse (d-f) particle size distributions together with 2D evaluations using CG-MC (b, e) and the PCSA-TR with a 2nd order power law parameterisation (c, f). Model in (a-c) has been simulated with a random noise level of 2 % and a rotor speed of 40 krpm . Model in (d-f) has been simulated with a random noise level of 0.5 % and a rotor speed of 20 krpm (d-f). Higher concentration is indicated by a more reddish colour. Reproduced from Ref. 5 with permission from The Royal Society of Chemistry.

size and density distribution of nanoparticles via a single AUC experiment.[7]

However, our studies on polydisperse nanoparticles revealed that broadly distributed systems pose some significant challenges for current state-of-the-art data evaluation methods. It was shown that existing methods provide either insufficient resolution or do not correctly reproduce the core-shell properties. Therefore, the performance of different data evaluation algorithms was investigated first by means of simulated data. Exemplary evaluation results taken from the publication for two different models (mono- and polydisperse) are depicted in Fig. 1. Peak splitting occurred for the 2DSA-CG-MC as also seen in preliminary experimental studies.

Most importantly, it could be shown that a novel data analysis methodology for PCSA provides much better results and allowed retrieval of the originally simulated model with excellent agreement.[9] In the revised PCSA, the model is constrained in such way that only parametrised combinations of sedimentation coefficients and partial specific volumes are allowed. Varying these parameters significantly reduces the degree of freedom during the analysis and prevents peak splitting on the second dimension (partial specific volume). A 2nd order power law was found to provide the best results for core-shell nanoparticles

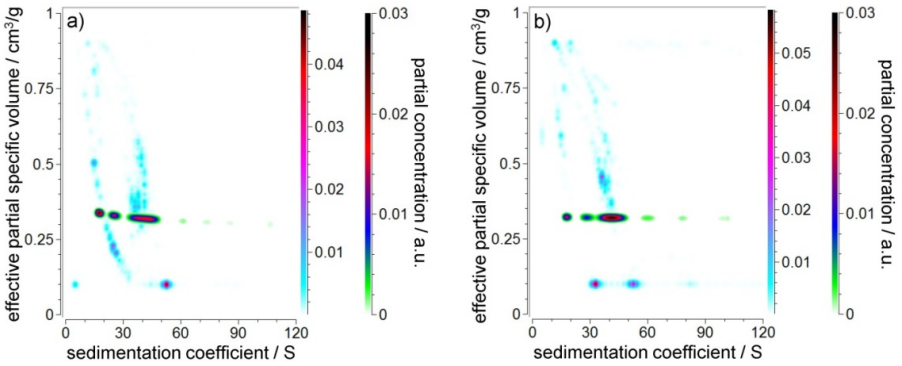


Figure 2. Results of the 2DSA-CG-MC (blue-purple-red) and PCSA-TR (green-blue-red) of ZnO nanoparticles ripened in a batch synthesis for a) 3 h and b) 4 h at 35°C. Evaluation was performed for SV-data at 270 nm. Reproduced from Ref. 5 with permission from The Royal Society of Chemistry.

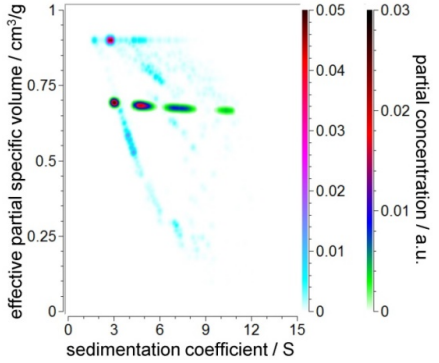


Figure 3. Results of the 2DSA-CG-MC (blue-purple-red) and PCSA-TR (green-blue-red) of $CuInS_2$ nanoparticles ripened in a batch synthesis for 1 h. The root-mean-square deviation values were 0.00871 OD and 0.00880 OD for the CG-MC and PCSA-TR, respectively. Evaluation was performed for SV-data at 350 nm. Reproduced from Ref. 5 with permission from The Royal Society of Chemistry.

with size dependent densities. The parametrisation variables are determined during the fitting process via the criterion of the lowest root-mean-square deviation. Different levels of random noise proportional to the loading and local concentration were further compared to assess the accuracy, robustness and limitations of the PCSA approach.

Finally, 2nd order power law PCSA was applied to experimental data. Two different ZnO quantum dots samples, ripened for 3 h and 4 h, respectively, were investigated by means of 2DSA-CG-MC and PCSA (Fig 2). The latter analyses provided excellent results for the sedimentation coefficients and partial specific volumes, whereas the 2DSA failed due to heavy peak splitting. $CuInS_2$ nanoparticles prepared via batch synthesis were used to obtain additional experimental evidence to test the capabilities of the new methodology (Fig. 3).

In summary, the studies revealed that the PCSA-TR with the 2nd order power law pa-

rameterisation excels at a wide range of noise levels and rotor speeds where previous algorithms failed to produce satisfactory results. This finally permits the determination of core size distributions of polydisperse core-shell nanoparticles from the obtained sedimentation coefficients and partial specific volumes. Nevertheless, the level of random noise should not be too high ($< 2\,\%$) and the applied rotor speed should be reasonable for the investigated distributions to obtain accurate results.

3 Hard- and Software Developments for Multiwavelength AUC

In a collaborative work together with the University of Konstanz, significant advancements in detector hardware and software for MWL-AUC experiments could be achieved, demonstrating improvement in both the spectral performance and UV capabilities of the instrument. The hardware is an extension of the Open AUC MWL detector developed in academia and industry, which was first introduced in 2006 by Bhattacharyya *et al.*[10] Additional modifications as well as new analytical methods available for MWL data have since been reported. New and continuing improvements have been added to the MWL detector, including mirror source and imaging optics, UV sensitive acquisition modes and revised data acquisition software.[14] The marked improvement of experimental data promises to provide access to increasingly complex systems, especially semiconductor nanoparticles, synthetic polymers, biopolymers, and other chromophores absorbing in the UV. Details of the detection system and components were examined to reveal the influences on data quality and to guide further developments. The benchmark comparisons of data quality across platforms shown can further serve as a reference guide for evaluation of commercial absorbance optics.

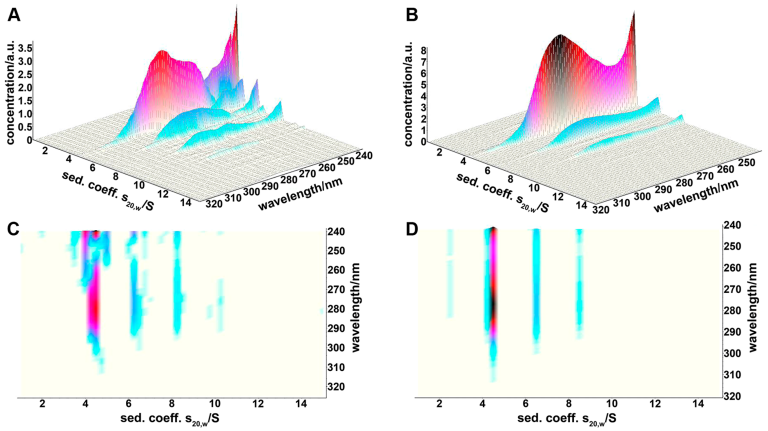


Figure 4. Sedimentation velocity data evaluation results of BSA recorded with the lens focused illumination-lens imaging optics (A and C) and with mirror focused illumination-mirror imaging optics (B and D). Panels A and B show 3D plots of sedimentation coefficient versus wavelength as obtained by parallel processing of the individual wavelength data sets with the 2DSA-MC analysis method in UltraScanIII. Panels C and D are pseudo 3D projection views of the same results. Higher concentration is indicated by darker colour. Reprinted with permission from Ref. 14. Copyright (2018) American Chemical Society.

The spectral analysis of MWL-AUC data is especially demanding for data evaluation tools, as the radial and temporal evolution of optical signal is monitored at several hundreds of wavelengths. The vast amount of data makes high-resolution spectral and hydrodynamic characterisation on desktop PCs impossible. Fig. 4 depicts results of MWL analyses performed using UltraScanIII and HPC. The protein bovine serum albumin (BSA) was measured in the old and the revised experimental setup and it was found that a significant increase in wavelength resolution and reproducibility could be achieved. This development allows users to improve the coupling of hydrodynamic and optical properties of macromolecules and particles and to study size, shape or composition dependent optical properties with high accuracy.

4 Analysis of Biological Systems via Multiwavelength AUC

Quantitative characterisation of the interaction between biological macromolecules is essential for the understanding of many natural biological processes. With the pioneering work of Schachman and co-workers,[16] it is possible to study macromolecular components via SV AUC experiments. It is well known that global MWL analysis makes it possible to discriminate different protein complexes, which is advantageous when studying protein-protein interactions. Thus, Balbo *et al.* proposed to use spectral information in order to overcome limitations for mixtures containing several macromolecular components.[17] As an example for the capabilities of MWL-AUC in this field, the protein myoglobin was studied individually and in a mixture with BSA.[14] The results of the analyses are depicted in Fig. 5.

It gets apparent that AUC in combination with MWL analysis and HPC resources can be used to study the sedimentation and spectral properties of macromolecules with high resolution. No interaction between the two different biological systems is observable as can be seen from the spectrum.

In a further study, AUC was used in order to investigate a virus system. Information on the aggregation behaviour of the viral protein was of particular importance since it affects biological behaviour within the cell nucleus. The PCSA-MC evaluation for one exemplary dataset is shown in Fig. 6. The functional relation between frictional ratio and sedimentation coefficient was fitted. When the two parameters are known, the molar masses can be calculated accordingly. Consequently, the oligomeric state of the viral protein under investigation can be determined. The existence of monomers, dimers and trimers was revealed. The information from this MWL-AUC experiment helps to understand the formation and the behaviour of the viral protein. Further studies involving this system are currently in preparation.

5 Analysis of Simulated MWL-AUC Data

Brownian Dynamics (BD) algorithms are based on a microscopic description of particle motion tracked for individual particles along their trajectory. For a certain time step, each particle moves according to a deterministic sedimentation drift and a stochastic diffusional drift.[18] Mie's scattering theory was included into the BD algorithm to simulate extinction weighted data as one would obtain from MWL-AUC experiments.[11] In order to study

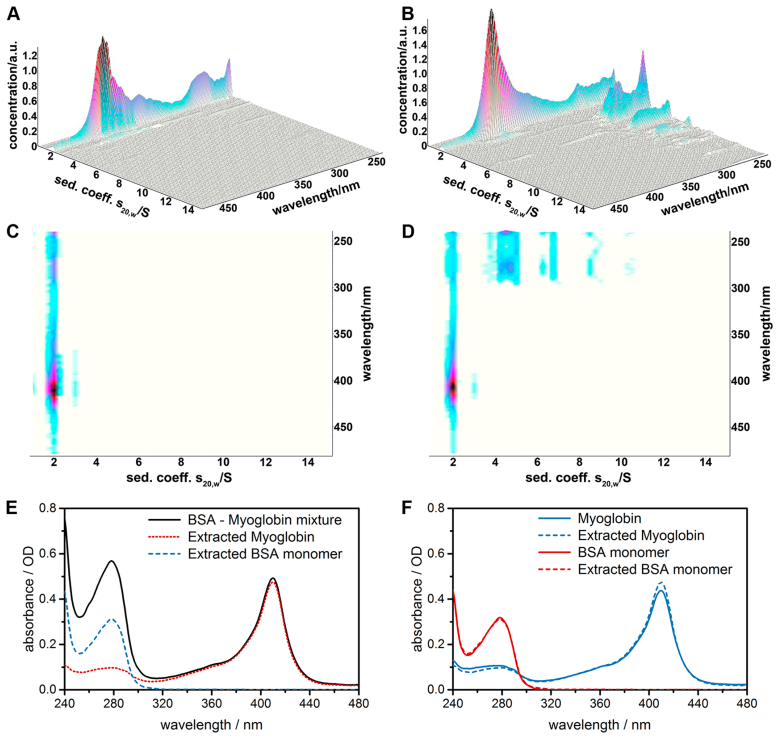


Figure 5. SV data evaluation results of myoglobin (panels A and C) and a BSA/myoglobin mixture (panels B and D) recorded with the MWL-AUC. Panels A and B show 3D plots of sedimentation coefficient versus wavelength as obtained by parallel processing of the individual wavelength data sets with the 2DSA-MC analysis method in UltraScanIII. Panels C and D are pseudo 3D projection views of the same results. Panel E shows the absorbance spectra of the mixture, and the spectra of the AUC resolved solutes extracted from the 2DSA data. Panel F shows the extracted spectra along with the corresponding spectra of the pure samples collected in a normal spectrophotometer. Higher concentration is indicated by darker colour. Reprinted with permission from Ref. 14. Copyright (2018) American Chemical Society.

nanoparticles in solution with this modelling approach, the presented algorithm had to be validated. This was achieved by reproducing the input to the simulation by the respective analysis tools. The validation of the data simulations via UltraScanIII is visualised in Fig. 7A for TiO_2 nanoparticles with a mean diameter of 5 nm and a standard deviation of 0.5 nm. The plot shows the extinction weighted sedimentation coefficient distribution, which served as input to the simulation for a given wavelength of 295 nm along with the 2DSA analysis in UltraScanIII. TR with a specified regularisation parameter of 0.95 was used for the analysis in order to smooth the obtained distributions. 220 wavelengths from 280 nm to 600 nm were evaluated using HPC in order to extract the maximum information content of the simulation. The input to the simulation can be well reproduced as shown exemplarily for one wavelength in Fig. 7A.

The extraction of the optical properties is visualised in Fig. 5B. It shows the extinction profile as input to the simulation as a function of the wavelength and the extracted spec-

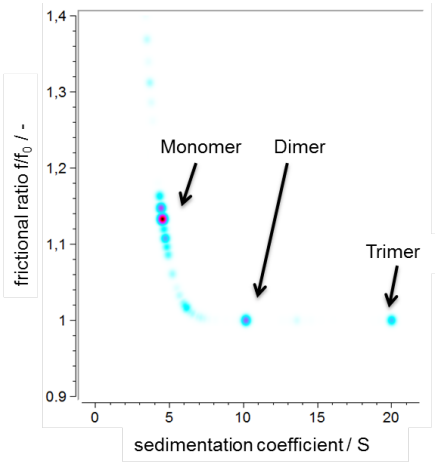


Figure 6. Extracted sedimentation coefficients and frictional ratios from a PCSA analysis in UltraScanIII with 100 Monte-Carlo iterations. The biological sample clearly shows three species, monomer, dimer and trimer. Higher concentration is indicated by a more reddish colour.

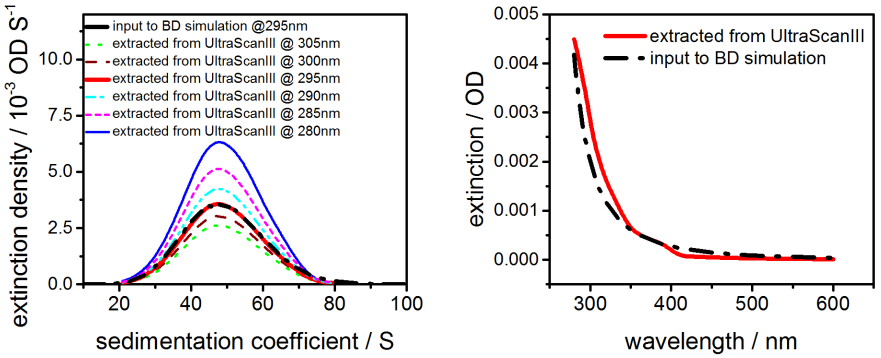


Figure 7. (A) Retrieved extinction weighted size distribution from UltraScanIII along with the input to the simulation. (B) Extinction profile as a function of the wavelength as input to the Brownian Dynamics simulation and as extracted from UltraScanIII.

trum from UltraScanIII. Besides hydrodynamics, the spectral properties can be reproduced. Moreover, it was shown that it is possible to include hydrodynamic and thermodynamic non-ideality in the BD algorithm.[19] Non-ideality directly translates into a change in the shape of the sedimentation boundaries. In particular, hydrodynamic and thermodynamic non-ideality phenomena lead to concentration-related self-sharpening effects of the sedimentation boundaries. In terms of data analysis, non-ideality phenomena are challenging especially in the case of polydisperse samples such as synthetic polymers and nanoparticles.[20] In order to correctly account for all effects influencing the sedimentation boundary simultaneously, high-resolution data analysis is indispensable. The manifold possibilities

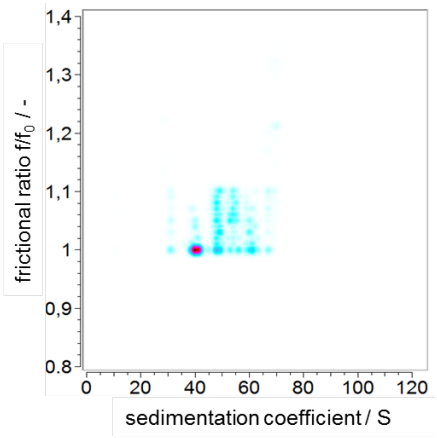


Figure 8. Frictional ratio *vs.* sedimentation coefficient as extracted from a 2DSA-MC analysis in UltraScanIII for simulated non-ideal sedimentation data. The non-ideal sedimentation behaviour results in an apparent heterogeneity in shape anisotropy. Higher concentrations are indicated by a more reddish colour.

and flexibility of UltraScanIII to evaluate sedimentation data with different models together with the computational capabilities of HPC make it a promising tool for the in-depth study of non-ideality of polydisperse samples. For illustrative purposes of the effect of non-ideal sedimentation, titania nanoparticles (mean diameter = 5 nm, width = 0.5 nm) were modelled at a concentration of 1 g/L using BD. Their sedimentation behaviour exhibits hydrodynamic non-ideality (ks = 100 mL/g).

As tools taking into account non-ideal sedimentation for particle size distributions are not yet available, a 2DSA analysis with 100 Monte-Carlo iterations was performed. The results for one out of 220 wavelengths are exemplary shown in Fig. 8. As can be seen in Fig. 8, the data analysis detects one species, which is in line with the expected value s_{mean} = 40 S. The influence of hydrodynamic non-ideality results in an increased broadening of the peak, which could be misinterpreted when considering the results. Thus, we believe that future developments in terms of data analysis should focus on taking non-ideality phenomena into account when analysing polydisperse nanoparticulate systems. Respective developments are currently ongoing. We expect to gain more insight from studying interparticulate interactions via our theoretical approach and the presented analysis method. UltraScanIII with HPC capabilities can certainly contribute, as it offers high flexibility together with high resolution and throughput.

6 Summary

UltraScanIII provides the possibility to perform multidimensional analysis of AUC experiments with highest resolution in order to give insight into *e. g.* molecular mass distributions, size distributions or information on core-shell structures. This is possible due to the HPC-enabled methodology employed in the optimisation algorithms of UltraScanIII. This capability offers excellent resolution and allows the investigator to more accurately identify shape and molecular weight parameters for all solutes present in a sample. In

this contribution, it is demonstrated how these possibilities can be used to analyse polydisperse core-shell nanoparticles with respect to size and density within a single MWL-AUC experiment. Moreover, advances in the spectral performance and UV capabilities of the MWL instrument together with high-resolution data analysis enable the study of complex systems such as polymers and semiconductor nanoparticles. Finally, the analysis of simulated MWL data is presented. Brownian Dynamics serves as a powerful forward model that simulates MWL AUC experiments with high precision, in this way enabling a comprehensive comparison of results from different data analysis. Furthermore, it is able to deal with non-ideality phenomena which play a crucial role for systems of high concentrations. This is especially challenging for polydisperse nanoparticles. Future studies should thus point towards accounting for hydrodynamic and thermodynamic non-ideality throughout data analysis to gain more insight into the interactions of complex systems.

Acknowledgements

The authors want to thank all collaborators for the studies shown, in particular Borries Demeler from University of Lethbridge, Canada, Gary Gorbet from University of Texas, USA, as well as Joe Pearson and Helmut Cölfen from University of Konstanz, Germany. 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. W. Mächtle and L. Börger, *Analytical Ultracentrifugation of Polymers and Nanoparticles*, Springer Laboratory, Springer, 2006.
2. K. L. Planken and H. Cölfen, *Analytical ultracentrifugation of colloids*, Nanoscale **2**, 1849, 2010.
3. P. Schuck, *Analytical Ultracentrifugation as a Tool for Studying Protein Interactions*, Biophys. Rev. **5**, 159–171, 2013.
4. B. Demeler, T.-L. Nguyen, G. E. Gorbet, V. Schirf, E. H. Brookes, P. Mulvaney, A.'a O. El-Ballouli, J. Pan, O. M. Bakr, A. K. Demeler, B. I. Hernandez Uribe, N. Bhattarai and R. L. Whetten, *Characterization of size, anisotropy, and density heterogeneity of nanoparticles by sedimentation velocity*, Anal. Chem. **86**, 7688–7695, 2014.
5. B. Demeler and H. Saber, *Determination of Molecular Parameters by Fitting Sedimentation Data to Finite-Element Solutions of the Lamm Equation*, Biophys. J. **74**, 444–454, 1998.
6. W. Cao and B. Demeler, *Modeling analytical ultracentrifugation experiments with an adaptive space-time finite element solution of the Lamm equation*, Biophys. J. **89**, 1589–1602, 2005.
7. E. Brookes, W. Cao, and B. Demeler, *A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape*, Eur. Biophys. J. **39**, 405–414, 2010.

8. E. Brookes and B. Demeler, *Genetic Algorithm Optimization for Obtaining Accurate Molecular Weight Distributions from Sedimentation Velocity Experiments*, in Analytical Ultracentrifugation VIII, C. Wandrey and H. Cölfen (Editors), Springer, 33–40, 2006.
9. B. Demeler, *Methods for the design and analysis of sedimentation velocity and sedimentation equilibrium experiments with proteins*, Current protocols in protein science **60**, 7.13.1–7.13.24, 2010.
10. G. Gorbet, T. Devlin, B. I. H. Uribe, A. K. Demeler, Z. L. Lindsey, S. Ganji, S. Breton, L. Weise-Cross, E. M. Lafer, E. H. Brookes, B. I. Hernandez Uribe, and B. Demeler, *A parametrically constrained optimization method for fitting sedimentation velocity experiments*, Biophys. J. **106**, 1741–1750, 2014.
11. E. Brookes and B. Demeler, *Parallel computational techniques for the analysis of sedimentation velocity experiments in UltraScan*, Colloid Polym. Sci. **286**, 139–148, 2008.
12. G. E. Gorbet, J. Z. Pearson, A. K. Demeler, H. Cölfen, and B. Demeler, *Next-Generation AUC: Analysis of Multiwavelength Analytical Ultracentrifugation Data*, Methods Enzymol. **562**, 27–47, 2015.
13. H. Kim, E. H. Brookes, and B. Demeler, *A performance predictor for UltraScan supercomputer calculations*, in Proceedings of the 2015 XSEDE Conference on Scientific Advancements Enabled by Enhanced Cyberinfrastructure (XSEDE '15), G. D. Peterson (Editor), ACM Press, New York, Article No. 42, 2015.
14. J. Pearson, J. Walter, W. Peukert, and H. Cölfen, *Advanced Multiwavelength Detection in Analytical Ultracentrifugation*, Analytical chemistry **90**, 1280–1291, 2018.
15. J. Walter, G. Gorbet, T. Akdas, D. Segets, B. Demeler, and W. Peukert, *2D analysis of polydisperse core-shell nanoparticles using analytical ultracentrifugation*, Analyst **142**, 206–217, 2016.
16. S. J. Edelstein and H. K. Schachman, *The simultaneous determination of partial specific volumes and molecular weights with microgram quantities*, J. Biol. Chem. **242**, 306–311, 1967.
17. A. Balbo, K. H. Minor, C. A. Velikovsky, R. A. Mariuzza, C. B. Peterson, and P. Schuck, *Studying multiprotein complexes by multisignal sedimentation velocity analytical ultracentrifugation*, Proc. Natl. Acad. Sci. USA **102**, 81–86, 2005.
18. M. J. Uttinger, S. E. Wawra, T. Guckeisen, J. Walter, A. Bear, T. Thajudeen, P. J. Sherwood, A. Smith, A. M. Wagemans, W. F. Stafford, and W. Peukert, *A Comprehensive Brownian Dynamics Approach for the Determination of Non-ideality Parameters from Analytical Ultracentrifugation*, Langmuir **35**, 11491-11502, 2019.
19. T. Thajudeen, J. Walter, M. Uttinger, and W. Peukert, *A Comprehensive Brownian Dynamics-Based Forward Model for Analytical (Ultra)Centrifugation*, Part. Part. Syst. Charact. **34**, 1600229, 2017.
20. M. J. Uttinger, J. Walter, T. Thajudeen, S. E. Wawra, and W. Peukert, *Brownian dynamics simulations of analytical ultracentrifugation experiments exhibiting hydrodynamic and thermodynamic non-ideality*, Nanoscale **2**, 200, 2017.

Computational Plasma Physics

Computational Plasma Physics: Physics with High-Intensity, Ultra-Short Optical Pulses

Karl-Heinz Spatschek

Institut für Theoretische Physik, Heinrich-Heine-Universität Düsseldorf,
40225 Düsseldorf, Germany
E-mail: khs@tp1.uni-duesseldorf.de

The Nobel Prize in Physics 2018 was awarded "for groundbreaking inventions in the field of laser physics" with one half to Arthur Ashkin "for the optical tweezers and their application to biological systems", the other half jointly to Gérard Mourou and Donna Strickland "for their method of generating high-intensity, ultra-short optical pulses". The Chirped Pulse Amplification (CPA) technique, invented by Strickland and Mourou,[1] opened an entirely new area of research for short pulse laser-matter interaction. The following two papers show very interesting developments in fundamental physics[2] and applications to plasma accelerators,[3] respectively, being provided by the availability of high-intensity laser pulses.

High-intensity laser interaction with matter is a broad field with an explosive scientific development during the last three decades. Since the invention of the Chirped Pulse Amplification (CPA) technique,[1] short pulse laser-matter interaction lead to an entirely new area of research, namely what is called now relativistic optics. Fig. 1 gives an overview over the highest intensity laser pulses being available at time. At lot of new physics will become accessible in the coming years.

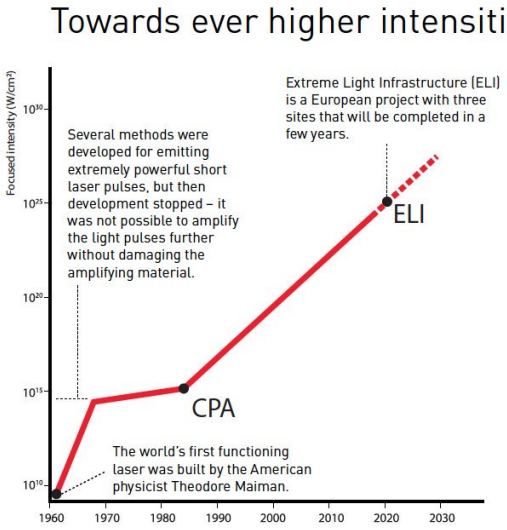


Figure 1. Overview over records in focused laser intensity in dependence on time, © Johan Jarnestad/The Royal Swedish Academy of Sciences.

Before the CPA-area, the possibility of amplifying pulses directly was limited since intensities of some GW/cm^2 would damage the gain medium. Stretching the pulse in a controlled (reversible) manner by adding a chirp (temporal variation of the frequency), is the first stage in CPA. The chirped pulse is then amplified. For that usually one or more conventional laser amplifier stages can be used to increase the pulse energy by a factor of 10^7–10^9. Finally, a compressor performs the exact optical inverse of the stretcher to deliver an amplified pulse of the same duration as the oscillator at the beginning. Optical Parametric Amplification (OPCPA) in a nonlinear crystal promises even higher intensities than conventional CPA.

It is convenient to characterise the laser field by the dimensionless parameter $a_0 = eA_0/m_e c$, where A_0 is the peak vector potential of the laser pulse. Since a_0 corresponds to the normalised (by $m_e c$) maximum transverse momentum of an electron in the laser field, $a_0 \geq 1$ is often referred to as the entry into the relativistic optics. Even university facilities like the ARCTURUS at the University of Düsseldorf are able to produce $a_0 \sim 10$. A next big step will be reached at $a_0 \sim 1000$ when also ions will be accelerated to relativistic velocities.[4] For higher intensities it will be possible to even study electron-positron pair creation and other quantum-electrodynamic (QED) effects.

The first paper by Liseykina *et al.*[2] discusses radiation friction which will become important at intensities $\geq 10^{23}$ W/cm^2. It provides essential tools, both to interpret current experiments and to give directions for the design and performance of facilities under development. The radiation-dominated regime will certainly bestow us with a variety of new insights into fundamental physics.

The second paper by Zeng *et al.*[3] deals with laser and plasma wakefield accelerators. For a laser wakefield accelerator usually a moderately relativistic laser intensity of the order of 10^{18} W/cm^2 at focus is sufficient. Two schemes are demonstrated and tested by numerical simulations. The first one is a hybrid laser-plasma wakefield accelerator, whereas the second one is the laser-plasma telescope system (using a plasma eyepiece) which largely reduces the focal space required in laser wakefield accelerators driven by a petaWatt-class laser.

References

1. D. Strickland and G. Mourou, *Compression of amplified chirped optical pulses*, Opt. Commun. **55**, 447, 1985.
2. T. Liseykina, D. Bauer, S. Popruzhenko, and A. Macchi, *Collective Effect of Radiation Friction in Laser-Driven Hole Boring of Dense Plasma Targets*, this volume.
3. M. Zeng, A. Martinez de la Ossa, and J. Osterhoff, *Particle-in-Cell Simulation Studies for Hybrid Laser-Plasma Accelerators and Plasma Eyepieces*, this volume.
4. A. Frank, J. Fuchs, L. Lancia, G. Lehmann, J. Marques, G. Mourou, C. Riconda, K. H. Spatschek, T. Toncian, L. Vassura, S. Weber, and O. Willi, *Amplification of ultra-short light pulses by ion collective modes in plasmas*, Eur. Phys. J. Special Topics **223**, 1153, 2014.

Collective Effect of Radiation Friction in Laser-Driven Hole Boring of Dense Plasma Targets

Tatyana Liseykina[1,2], Dieter Bauer[1], Sergey Popruzhenko[3], and Andrea Macchi[4]

[1] Institut of Physics, University of Rostock, 18051 Rostock, Germany
E-mail: {tatyana.liseykina, dieter.bauer}@uni-rostock.de

[2] Institute of Computational Mathematics and Mathematical Geophysics SD RAS, Lavrentjev ave. 6 , 630090, Novosibirsk, Russia

[3] Prokhorov General Physics Institute RAS, Vavilova str. 38, 119991 Moscow, Russia
E-mail: sergey.popruzhenko@gmail.com

[4] CNR, National Institute of Optics, Adriano Gozzini research unit, Pisa, Italy
E-mail: andrea.macchi@ino.it

We aim at numerically modelling classical and quantum dynamics of charged particles and electromagnetic fields under conditions realised in the interaction of superintense laser pulses with various types of massive targets. The focus of our research is put on an interaction regime where the particles' dynamics and their radiation emission exert a strong mutual influence on each other. This interaction regime is referred to in the modern research literature as the radiation-dominated regime and attracts a constantly growing interest in connection with the soon expected increase of the record intensities available in laser laboratories. Laser sources of the next generation including the Extreme Light Infrastructure (ELI) pillars in the Czech Republic, Hungary and Romania, the Apollon laser in France and several other facilities, providing laser powers up to 10 Petawatt are expected to raise the available intensity by two orders of magnitude putting first laboratory experiments in the radiation-dominated regime within reach.

1 Introduction

Advanced, predictive numerical simulations are an essential tool both to interpret current experiments and to give directions for the design and performance of facilities under development. In a strong connection with the proposal, design and interpretation of experiments, our project focuses on the newly emerged topic of laser-plasmas in the radiation-dominated regime (RDR). Here we refer to the RDR having in mind any interaction setup where radiation of quanta by elementary particles significantly affects their individual or/and collective dynamics. Apparently this distortion of particle's dynamics influences the process of radiation itself, making the whole problem self-consistent and, due to the extreme intensities of the applied laser fields, highly nonlinear and even nonperturbative. For many years, the theoretical studies of elementary particles and plasma strongly coupled to their radiation were separated from possible experimental verifications by a huge gap in laser intensity and therefore remained on the level of academic theory. Nowadays, with the new generation of laser sources expected to come into operation in the nearest years[1, 2] or in a foreseeable future[3] this research became of topical interest and already resulted in a number of new fundamental findings.

A high, especially in the last decade, interest in effects of *radiation friction* (RF), *i. e.* of the back-action on the electron of the electromagnetic field emitted by the electron itself,

in the physics associated with ultra-intense laser pulses interacting with matter is justified, among others, by the fact that some qualitative effects owe their very existence to the radiation friction force. Such phenomena, if observed experimentally, can provide an unambiguous proof of the crucial role the radiation friction plays at high intensities of laser fields and high energies of plasma electrons. A specific realisation[4] of the inverse Faraday effect (IFE), *i. e.* the generation of magnetic fields due to absorption of angular momentum by the plasma, in the interaction of extremely strong, with intensity $10^{23} - 10^{24}$ W/cm^2, *circularly polarised* (CP) infrared laser pulse with *dense optically thick* plasma target, is such an effect. In this particular setup no other mechanism but radiation friction can provide the transfer of angular momentum from the laser field to plasma.[a] The results of three-dimensional particle-in-cell simulations (PIC) with the radiation reaction force included within the framework of classical physics, have demonstrated that the generated magnetic field could achieve the value of several Gigagauss at the laser field intensity around 10^{24} W/cm^2. This finding calls for two extensions. First, signatures of an IFE-generated magnetic field can be detected already at 10^{23} W/cm^2, *i. e.* at intensities expected at the laser facilities on the near future including ELI and Apollon.[1, 2] In this prospect our calculations[4] require further exploration, which will aim at identifying laser and plasma parameters optimal for an experimental search. Second, extending our results into the domain of even higher intensities[b] one needs to take into account effects of quantum electrodynamics on the plasma motion. We expect that our analytic and numerical investigations will allow answering the question on the existence of principal limitations to the value of the magnetic field generated in the interaction of superintense laser radiation with plasma.

The interaction of intense and super-intense coherent radiation with plasmas offers some of the most challenging problems for numerical physics. This is because the most important and interesting phenomena are multiscale and cannot be described hydrodynamically. The first feature imposes the use of very large numerical grids since the spatial size and/or evolution time of the system is much larger than the smallest spatial and temporal scales to be resolved. The second one implies that hydrodynamic description in real space must be abandoned and much more demanding kinetic equations in phase space have to be solved. This leads to an estimate that a "realistic" simulation would require a three-dimensional (3D) spatial grid with thousands of grid-points in each spatial direction, resulting in a total number of grid-points exceeding some tens of billions. In addition, while PIC method allows a great saving of memory in allocating the momentum space with respect to an Eulerian approach, thousands of particles per cell may be needed to properly resolve high energy tails in the distribution as well as sharp density gradients which always occur in intense laser-plasma interactions. Thus, the progress in this area is directly related to the possibility to access larger high performance computer resources.

[a] A number of different laser-plasma setups provide conditions necessary for the generation of superstrong quasistatic magnetic fields via the excitation of high electron currents. Recently reported records of the experimentally achieved magnetic field strengths approach the value of 1 Gigagauss, while the theory suggests several schemes for generating of even higher fields.

[b] This work currently remains in the realm of theoretical speculations, which could also become of practical interest with the next generation of laser sources[3] coming into play.

2 Description of Methods and Algorithms

All simulations presented in this paper were performed with the JURECA Supercomputer at the John von Neumann Institute for Computing (NIC) using the 3D electromagnetic PIC code UMKA,[5] developed for studying laser-plasma interactions in the strong-field regime. Generally speaking, a PIC code implements a particle-grid numerical method, where a collisionless plasma is sampled by a large number of computational particles and electromagnetic fields are discretised on a grid. The computational particles move across the grid cells and are accelerated by the electromagnetic fields defined at the grid points, which are computed by reconstructing the source current density from the particle positions and velocities. UMKA employs state-of-art, widely used numerical algorithms such as an Yee Cartesian lattice for electromagnetic fields, the Boris pusher to integrate the Lorentz force and the current reconstruction scheme[6] to satisfy the continuity equation. UMKA results have been benchmarked with other codes, compared with both analytical modelling and experimental results and reported in several publications.

At optical laser intensities exceeding $I \sim 10^{23}$ W/cm^2 the plasma electrons become ultrarelativistic within a fraction of the laser wave period. They experience a very strong acceleration and therefore emit relatively large amounts of electromagnetic radiation, causing radiation friction effects, which arise from the back-action on the accelerated electron of the electromagnetic field generated by the electron itself,[7] more and more important as the laser intensity increases. First PIC simulation that included RF[8] in the frame of the Lorentz–Abraham–Dirac (LAD) equation with the RF force that is exact for a point particle,[9] showed that RF effects become important at intensities exceeding 5×10^{22} W/cm^2, and are thus expected to play a crucial role in experiments. It is therefore important to incorporate RF in PIC simulations of laser-plasma by an appropriate modelling, keeping the essential RF effects into account while retaining at the same time the capability to perform large-scale simulations. In order to take RF effects self-consistently into account one should solve the LAD equation. This equation, however, suffers from inconsistencies such as the existence of the so-called "runaway" solutions in which the electron momentum grows exponentially in the absence of external fields. In the realm of classical electrodynamics, *i. e.* neglecting quantum effects, the LAD equation

$$mc\frac{du^\mu}{d\tau} = eF^{\mu\nu}u_\nu - e\tau_0\left(\frac{d^2u^\mu}{d\tau^2} + u^\mu u^\nu \frac{d^2u_\nu}{d\tau^2}\right),\ \tau_0 = 2e^2/(3mc^3) \tag{1}$$

can be consistently approximated by the Landau–Lifshitz (LL) equation

$$mc\frac{du^\mu}{d\tau} = eF^{\mu\nu}u_\nu + e\tau_0\left[u_\nu u^\alpha \partial_\alpha F^{\mu\nu} + \frac{e}{mc}F^{\mu\nu}F_{\nu\alpha}u^\alpha + \frac{e}{mc}\left(F^{\nu\beta}u_\beta F_{\nu\alpha}u^\alpha\right)u^\mu\right]$$

which is obtained by inserting the unperturbed Lorentz acceleration in Eq. 1.[10] This automatically removes the third order derivative of the position and thus runaway solutions, and brings us back to a more conventional phase space description. In 3D notation the above equation reads

$$\frac{d\mathbf{p}}{dt} = e\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right) + e\tau_0\gamma\left[\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\mathbf{E} + \frac{\mathbf{v}}{c}\times\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\mathbf{B}\right]$$
$$+ \tau_0\frac{e^2}{mc}\left\{\left[\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right)\times\mathbf{B} + \left(\frac{\mathbf{v}}{c}\cdot\mathbf{E}\right)\mathbf{E}\right] - \gamma^2\left[\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right)^2 - \left(\frac{\mathbf{v}}{c}\cdot\mathbf{E}\right)^2\right]\frac{\mathbf{v}}{c}\right\}$$

As long as a classical description is adequate, RF effects are relevant and quantum effects are subdominant, one can safely use the RF force given in textbooks[10]

$$\mathbf{F}_{\rm RF} \simeq -\frac{2}{3} r_c^2 \left[\gamma^2 \left(\mathbf{F}_{\rm L}^2 - \left(\frac{\mathbf{v}}{c} \cdot \mathbf{E} \right)^2 \right) \frac{\mathbf{v}}{c} - \mathbf{F}_{\rm L} \times \mathbf{B} - (\frac{\mathbf{v}}{c} \cdot \mathbf{E}) \mathbf{E} \right] \tag{2}$$

Here, $r_c \equiv e^2/mc^2 \simeq 2.8 \times 10^{-9}$ μm is the classical electron radius, $\mathbf{F}_{\rm L} = \mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}$, and we dropped the small terms containing the temporal derivatives of the fields.

To account self-consistently for the effect of radiation emission on the *electron dynamics* we implemented in UMKA the radiation reaction via the LL approach, using a numerical scheme,[11] based on the assumption that the acceleration of particles is dominated by the Lorentz force, with the RF force giving a smaller, albeit non-negligible contribution.[c] The numerical implementation[11] allows the addition of RF effects to any PIC code based on the standard Boris pusher algorithm for the acceleration of the particles at a small computational cost. We emphasise that the inclusion of the radiation loss as a dissipative process via the RF force requires the following assumptions: (i) the dominant frequencies in the escaping radiation are much higher than the highest frequency that can be resolved on the numerical grid, (ii) the radiation at such frequencies is incoherent, (iii) the plasma is transparent to such frequencies.

In the UMKA code, parallelisation is implemented by a decomposition in the (y, z)-discretised coordinates, perpendicular to the x-propagation of the laser pulse. This choice provides an initially balanced partition of the particles for problems where the particle density is initially uniform in the (y, z)-plane. A dynamical load-balancing keeps the data partition balanced as particles move across different domains.

3 Simulations

Based on the equations of macroscopic electrodynamics and conservation laws, a description of IFE in the field of an intense laser pulse[4] predicts the maximal amplitude of the quasistatic longitudinal magnetic field excited on the axis of a laser beam to be linear with respect to the laser magnetic field amplitude $B_{\rm L}$ and to the fraction of the laser energy η associated with the irreversible transfer of angular momentum from the laser field to the plasma

$$B_{xm} = C\eta a_0 B_0 \equiv C\eta B_{\rm L} \tag{3}$$

Here, $a_0 = eE_{\rm L}/m_e c\omega \equiv E_{\rm L}/B_0$, $B_0 = m_e c\omega/e$ are the dimensionless laser field amplitude and the characteristic magnetic field, respectively, $E_{\mathbf{L}} = B_{\rm L}$ and ω is the laser field frequency. For a laser pulse with $\lambda = 2\pi c/\omega = 0.8\,\mu$m wavelength $B_0 = 1.34 \times 10^8$ G. A dimensionless coefficient C is determined by the shape of the laser pulse envelope and has typical values $C \simeq 0.1 - 0.2$. The structure of Eq. 3 is consistent with the general theory of IFE.[12] The coefficient η reads

$$\eta = \frac{\omega L_{\rm abs}}{U_{\rm L}} \tag{4}$$

[c] For a benchmark with other approaches see M. Vranic, J. L. Martins, R. A. Fonseca, L. O. Silva, *Classical Radiation Reaction in Particle-In-Cell Simulations*, Computer Physics Communications **204**, 141-151, 2016.

where $L_{\rm abs}$ is the angular momentum absorbed by the plasma and $U_{\rm L} = \mathcal{A}\lambda^3 a_0^2 B_0^2$ is the energy stored in the laser pulse. The dimensionless coefficient $\mathcal{A}$ is determined by the pulse duration, time envelope and focusing. Eqs. 3-4 are insensitive to a particular physical mechanism of the angular momentum transfer. In particular, Eq. 3 applies independently of the impact of quantum effects on the plasma dynamics. In the high-field regime, radiation of plasma electrons is the only mechanism for energy absorption, thus $\omega L_{\rm abs} = U_{\rm rad}$ where $U_{\rm rad}$ is the radiation energy emitted by the electrons, and Eq. 4 reads

$$\eta \equiv \eta_{\rm rad} = \frac{U_{\rm rad}}{U_{\rm L}} \equiv \frac{\int d^3r \int dt P_{\rm rad}(\mathbf{r},t) n_e(\mathbf{r},t)}{U_{\rm L}} \leq 1 \tag{5}$$

Here, $P_{\rm rad}$ is the emission power for a single electron moving under the action of the local electromagnetic field with a time- and space-dependent envelope and n_e is the electron concentration. The *conversion efficiency* η is the key value for the determination of the magnetic field amplitude.

Below we present and discuss typical results from 3D PIC simulations of the interaction of an intense, tightly focused CP laser pulse with an optically opaque plasma target of thickness $D > 10\lambda$, where $\lambda = 0.8\,\mu$m corresponding to a Ti:Sapphire laser. The initial plasma density is $n_0 = 90 n_c = 1.55 \times 10^{23}\,{\rm cm}^{-3}$. The charge-to-mass ratio for ions is taken Z/A = 1/2. The supergaussian laser pulse is introduced via the time-dependent boundary condition at the plasma surface as

$$\vec{a}(r, x=0, t) = a_0 \left(\vec{e}_y \cos\omega t + \vec{e}_z \sin\omega t\right) \exp\left(-(r/r_0)^4 - (ct/r_L)^4\right) \tag{6}$$

with $r = \sqrt{y^2+z^2}$, $r_0 = 3.8\lambda, r_L = 3.0\lambda$ and duration (FWHM of the intensity profile) 14.6 fs. In our PIC calculations we varied the laser amplitude in the interval $a_0 = 300 - 750$, which corresponds to the peak intensities $(3.8 - 23.7) \times 10^{23}$ W/cm^2 and the total pulse energy at $1.08 \div 6.71$ kJ. The numerical box had a $[30 \times 25 \times 25]\lambda^3$ size, with 40 grid cells per λ in each direction and 125 particles per cell for each species. The simulations were performed on $5000 \div 10\,000$ cores of the JURECA Cluster Module.

Radiation Friction Losses in Laser-Driven "Hole Boring" of Dense Plasma Targets

In the considered laser-plasma interaction scenario the ponderomotive force of the laser field pushes and piles up electrons in the skin layer creating a static field that acts on the ions, so that effectively the radiation pressure is exerted on the whole plasma target. In a thick target, ions are thus accelerated at the front surface, causing steepening of the density profile and "hole boring" through the plasma, Fig. 1.

Analysis of the 3D distribution functions of the radiation power density $\mathcal{P}(x, r, v_x)$, calculated as $\mathcal{P} = n_e \mathbf{v} \cdot \mathbf{F}_{\rm RF}$, and of the electron and ion density $n_{e,i}(x, r, v_x)$ extracted from the simulation shows that most of the emitted radiation comes from electrons having velocities $v_x > 0$, and located close to the receding front of the ion density. This is illustrated for $a_0 = 500$ in Fig. 2 where space-time plots in the (x, t) plane are shown for the radiation power and the particle densities along the $x-$axis and at $r = 1\lambda$ distance from the axis. The density fronts move in the forward direction with average velocity $0.41c$ (see also Fig. 1), in fair agreement with the value of the hole boring velocity[13] $v_{\rm HB} = 0.47c$. Small oscillations in the front position are visible in correspondence with the generation of plasma bunches in the forward direction. The power density plot shows that most of

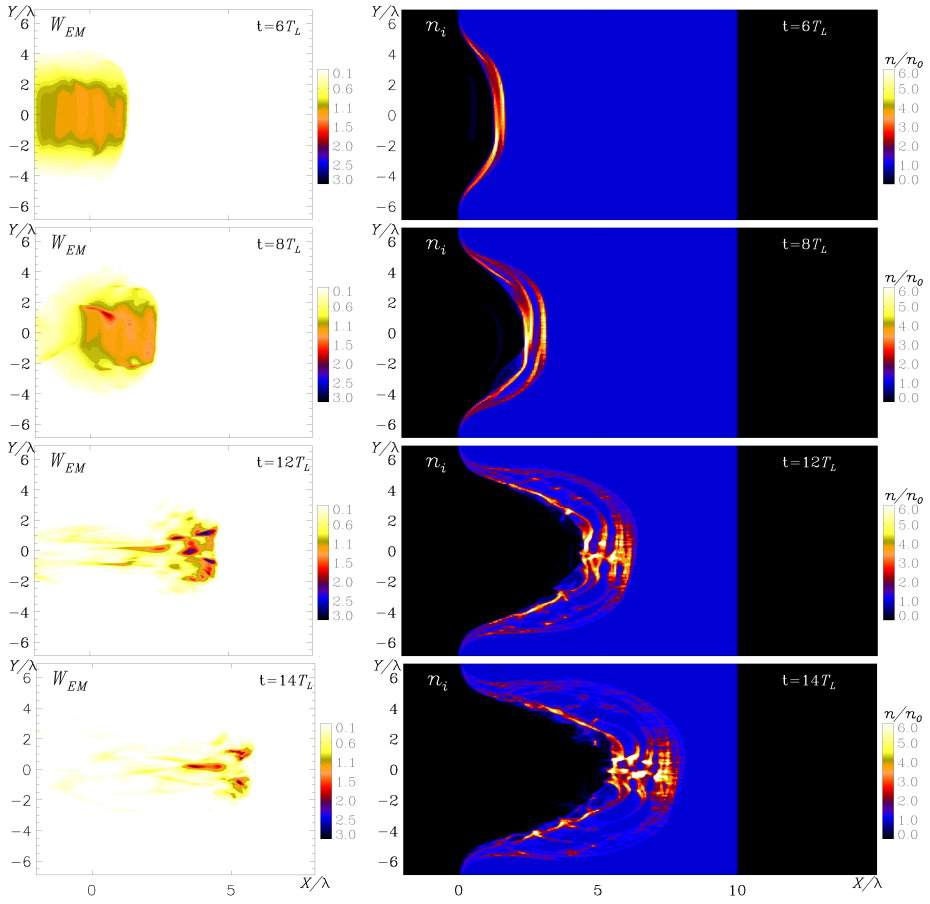


Figure 1. (left) Distribution of the ion density and (right) electromagnetic energy density $W_{\rm EM}/a_0^2$ at $t = 6, 8, 12, 14 \cdot (2\pi/\omega)$. The plasma target of an initial density $n_0 = 1.55 \times 10^{23}\,{\rm cm}^{-3}$ is irradiated by a short, tightly focused laser pulse with $a_0 = 600$. See text for further details.

the emission originates close to the hole boring front. Emission due to returning electrons with velocity $v \simeq -c$ is visible after $t = 11 \cdot 2\pi/\omega$, but its contribution to the total emitted power is small, presumably because of the low density in the returning jets, as seen in Fig. 2 of the $n_e(x,t)$ plot. This is also confirmed by the time-velocity plot of the radiation power density $\mathcal{P}(v_x,t)$, shown in Fig. 3.

Quantum Effects on Radiation Friction Driven Magnetic Field Generation

Overall, the results of our simulations demonstrated a good qualitative agreement to the predictions of the theoretical model,[14] which provides an accurate estimate of the radiation friction losses η for CP fields under the assumption that the classical RF regime is retained. In particular, the model[14] reproduces results of 3D PIC simulations with a 20 % accuracy and shows that longitudinal acceleration of the plasma and attenuation of the laser field on

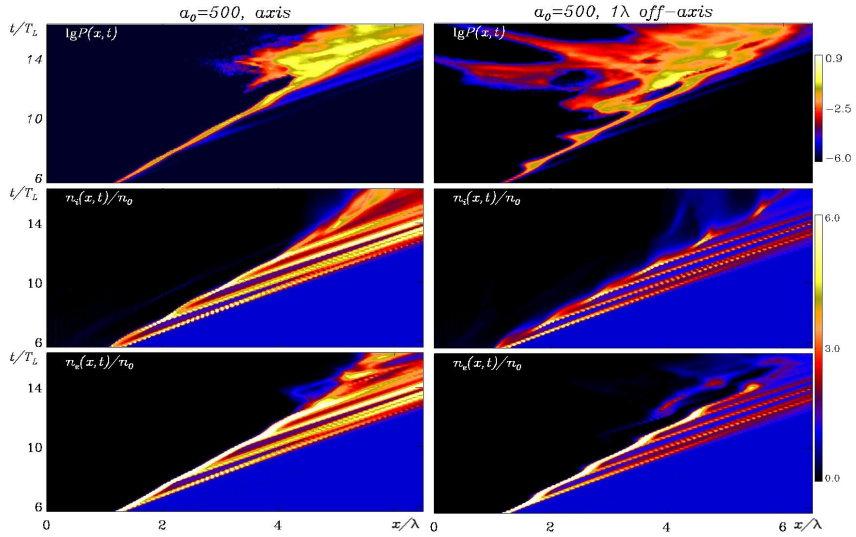


Figure 2. Space-time plots of the radiation power density $\mathcal{P}(x,t)$ (top, logarithmic scale, arbitrary units), ion density $n_i(x,t)$ (middle) and electron density $n_e(x,t)$ (bottom) all evaluated along the $x-$axis (left) and at $r = 1\lambda$ distance from the $x-$axis (right). The plasma target is irradiated by a short, tightly focused laser pulse with $a_0 = 500$.

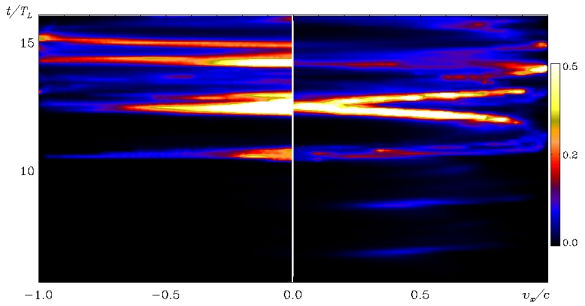


Figure 3. The distribution of the radiation power density $\mathcal{P}(v_x,t)$ (linear scale, arbitrary units) evaluated at $r = 1\lambda$ distance from the $x-$axis. Note, that the values of $\mathcal{P}$ for $v_x < 0$ are magnified by 20.

a small evanescence length inside the plasma, together with effects of time-space averaging, lead to a considerable suppression in the conversion efficiency so that, ultimately, it appears on the level of $\eta \approx 0.2$ for $a_0 = 600$, leading to the upper limit of the magnetic field $B_{\rm xm} \approx 3.2 \times 10^9$ G at the laser intensity 7.5×10^{23} W/cm^2. However, both the above presented simulations and the theoretical approach[14] completely discard the effect of quantum recoil on the spectrum of emitted radiation. On the one hand the parameter $\chi = \dfrac{e\hbar}{m^3c^4}\sqrt{(-(F_{\mu\nu}p^\nu)^2} = \dfrac{E'_{\rm L}}{E_{\rm cr}}$ which determines the significance of quantum effects and equals the ratio of the external (laser) electric field in the electron rest frame $E'_{\rm L}$ to

the critical field of quantum electrodynamics, $E_{\rm cr} = m^3c^3/e\hbar$, remains smaller than unity up to intensities $\sim 10^{25}$ W/cm^2, making a classical description of dynamics and radiation of electrons applicable at least on the qualitative level. On the other hand the spectrum of emitted photons appears considerably modified by quantum effects already[15, 16] for $\chi \approx 0.1$. For the considered parameters this is achieved already at $a_0 \simeq 200$ and therefore for $I \simeq 1.9 \times 10^{23}$ W/cm^2, so that quantum corrections to the power of radiation may become numerically important. We therefore account for the suppression of the radiation power due to the off-set in the emission spectrum by introducing the quantum factor $g(\chi)$.[d] To that end we have further modified our code by introducing the factor χ in the expression for the radiation friction force $\tilde{\mathbf{F}}_{\rm RF} = g(\chi)\mathbf{F}_{\rm RF}$. The quantum parameter χ is calculated at each time step by taking the values of electric and magnetic fields at the position of each electron.

The distribution of the electron concentration, of the g-factor and of the radiation power extracted from the simulations are shown in Fig. 4 (a-c) and the comparison of the space distribution of the generated axial component of the magnetic field for the simulations with $g = 1$ and $g(\chi) \neq 1$ is presented in Fig. 4 (c, d). The value of the magnetic field for $g = g(\chi)$ is approximately two times smaller than for $g = 1$. This is in agreement with the corresponding decrease of the value of the radiation friction losses. Notwithstanding, the qualitative behaviour of both values hardly changes, and the scaling laws[14] found within the fully classical model remain valid in the semiclassical domain. The space-time structure of the magnetic field appears considerably modified by the quantum effect. This implies that the axial magnetic field remains of the same order of magnitude but is quite sensitive to the quantum modification of the radiation friction force.

4 Conclusion

The collective laser-plasma dynamics can play a crucial role in determining the amount of radiation friction losses.[4, 14, 18] Our fully dimensional relativistic PIC simulations with the radiation reaction implemented via the LL force show that in the interaction of a CP pulse with a thick plasma slab of overcritical initial density the collective effects greatly reduce radiation losses with respect to electrons driven by the same laser pulse in vacuum. The important consequence of this result[14] is that the reliability of classical calculations can be shifted up to intensities $\sim 10^{24}$ W/cm^2. The simulation results show a good agreement of the calculated conversion efficiency of the laser energy into incoherent radiation with the predictions of analytic calculations, which account for the influence of radiation losses on the single electron trajectory, the global "hole boring" motion of the laser-plasma interaction region under the action of radiation pressure, and the inhomogeneity of the laser field in both longitudinal and transverse direction.

Radiation losses in the interaction of superintense CP laser pulses with high-density plasmas can lead to the generation of strong quasistatic magnetic fields via absorption of the photon angular momentum. Magnetic field strengths of several Gigagauss can be achieved at laser intensities 10^{24} W/cm^2 which proceed at the border between the classical and the quantum interaction regimes. Such huge magnetic fields, besides affecting the plasma dynamics, provide an unambiguous signature of RF effects and may be measured by polarimetry methods. We improve the classical modelling of the laser interaction with

[d] General formulas for the quantum factor $g(\chi)$ can be found in the literature.[15] For practical calculations we use a fit[17] $g(\chi) = (1 + 12\chi + 31\chi^2 + 3.7\chi^3)^{-4/9}$.

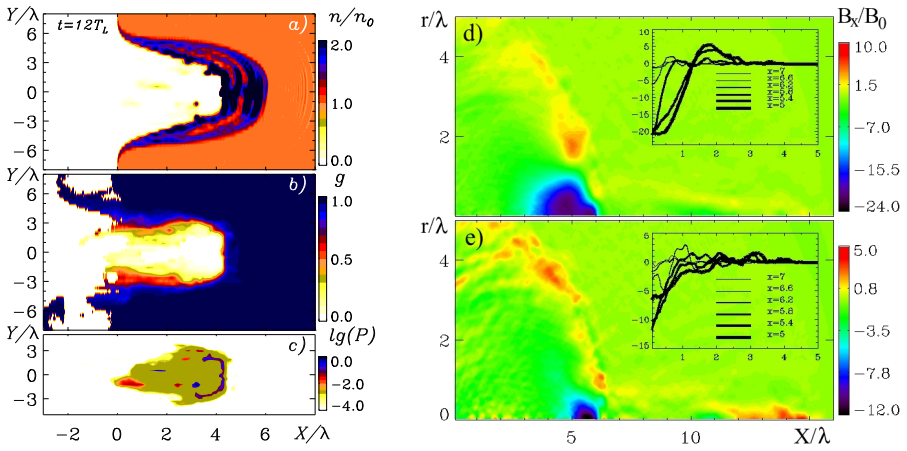


Figure 4. (left) Distributions of the electron concentration (a) of the g-factor (b) and of the radiation power (c) in (x, y) plane, extracted from the simulation for $a_0 = 500$ at $t = 10 \cdot (2\pi/\omega)$ when the total radiation power is close to its maximum. (right) The magnetic field component B_x along the laser pulse propagation direction for g=1 (c) and $g = g(\chi) \neq 1$ (d) for $a_0 = 400$ is recorded at $t = 26 \cdot (2\pi/\omega)$, *i. e.* after the laser pulse is already reflected from the plasma. The onsets in (c) and (d) show the distributions of the magnetic field along the radial direction for fixed values of x.

overcritical plasma in the "hole boring" regime by using a modified radiation friction force accounting for quantum recoil and spectral cut-off at high energies. The results of analytic calculations and full dimensional PIC simulations show that in foreseeable scenarios the quantum effects may lead to a 50 % decrease of the magnetic field amplitude, possibly making the effect a suitable diagnostics for benchmarking of the radiation friction theories.

Acknowledgements

Numerical simulations were performed using the computing resources granted by the John von Neumann Institute for Computing under the project HRO04. T. V. L. acknowledges support by the Russian Science Foundation (research grant 19-71-20026).

References

1. S. Kühn, M. Dumergue, S. Kahaly *et al.*, *The ELI-ALPS facility: the next generation of attosecond sources*, J. Phys. B: At. Mol. Opt. Phys. **50**, 132002, 2017, `https://www.eli-beams.eu`, `http://www.eli-np.ro`.
2. J. P. Zou, C. Le Blanc, D. N. Papadopoulos *et al.*, *Design and current progress of the Apollon 10 PW project*, High Power Laser Sci. Eng. **3**, E2, 2015;
   D. N. Papadopoulos, J. P. Zou, C. Le Blanc *et al.*, *First commissioning results of the Apollon laser on the 1 PW beam line*, in Conference on Lasers and Electro-Optics (CLEO), IEEE, paper STu3E.4, 2019.
3. A. V. Bashinov, A. A. Gonoskov, A. V. Kim *et al.*, *New horizons for extreme light physics with mega-science project XCELS*, Eur. Phys. J. ST **223**, 1105, 2014;

J. Kawanaka, K. Tsubakimoto, H. Yoshida *et al.*, *Conceptual design of sub-exawatt system by using optical parametric chirped pulse amplification*, J. Phys. Conf. Ser. **688**, 012044, 2016.
4. T. V. Liseykina, S. V. Popruzhenko, and A. Macchi, *Inverse Faraday effect driven by radiation friction*, New J. Phys. **18**, 072001, 2016.
5. V. A. Vshivkov, N. M. Naumova, F. Pegoraro, and S. V. Bulanov, *Nonlinear electrodynamics of the interaction of ultra-intense laser pulses with a thin foil*, Phys. Plasmas **5**, 2727, 1998.
6. J. Villasenor and O. Buneman, *Rigorous change conservation for electromagnetic field solvers*, Comp. Phys. Comm. **69**, 306-316, 1992;
T. Zh. Esirkepov, *Exact charge conservation scheme for Particle-In-Cell simulations with arbitrary form-factor*, Comp. Phys. Comm. **135**, 144-153, 2001.
7. J. D. Jackson, *Classical Electrodynamics, 3rd Edition*, Wiley, 1998.
8. A. Zhidkov, J. Koga, A. Sasaki *et al.*, *Radiation damping effects on the interaction of ultraintense laser pulses with an overdense plasma*, Phys. Rev. Lett. **88**, 185002, 2002.
9. F. Rohrlich, *The correct equation of motion of a classical point charge*, Phys. Lett. A **283**, 276-278, 2001.
10. L. D. Landau and E.M. Lifshitz, *The classical theory of fields, 2nd Edition*, Elsevier, 1975.
11. M. Tamburini, F. Pegoraro, A. Di Piazza *et al.*, *Radiation reaction effects on radiation pressure acceleration*, New J. Phys. **12**, 123005, 2010.
12. L. P. Pitaevskii, *Electric forces in a transparent dispersive medium*, Sov. Phys. JETP **12**, 1008, 1961;
M. G. Haines, *Generation of an axial magnetic field from photon spin*, Phys. Rev. Lett. **87**, 135005, 2001.
13. A. P. L. Robinson, P. Gibbon, M. Zepf *et al.*, *Relativistically correct hole-boring and ion acceleration by circularly polarized laser pulses*, Plasma Phys. Control. Fusion **51**, 024004, 2009.
14. S. V. Popruzhenko, T. V. Liseykina, and A. Macchi, *Efficiency of radiation friction losses in laser-driven "hole boring" of dense targets*, New J. Phys. **21**, 033009, 2019.
15. V. I. Ritus, *Quantum effects on the interaction of elementary particles with an intense electromagnetic field*, J. Russ. Laser Res. **6**, 497, 1985;
J. G. Kirk, A. R. Bell, and I. Arka, *Pair production in counter-propagating laser beams plasma*, Plasma Phys. Contr. Fusion. **51**, 085008, 2009.
16. F. Niel, C. Riconda, F. Amiranoff *et al.*, *From quantum to classical modeling of radiation reaction: a focus on the radiation spectrum*, Plasma Phys. Contr. Fusion **60**, 094002, 2018.
17. A. G. R. Thomas, C. P. Ridgers, S. S. Bulanov, B. J. Griffin, and S. P. D. Mangles, *Strong radiation-damping effects in a gamma-ray source generated by the interaction of a high-intensity laser with a wakefield-accelerated electron beam*, Phys. Rev. X **2**, 041004, 2012.
18. E. N. Nerush and I. Y. Kostyukov *Laser-driven hole boring and gamma-ray emission in high-density plasmas*, Plasma Phys. Control. Fusion **57**, 035007, 2015;
D. Del Sorbo, D. R. Blackman, R. Capdessus *et al.*, *Efficient ion acceleration and dense electron positron plasma creation in ultra-high intensity laser-solid interactions*, New J. Phys. **20**, 033014, 2018.

Particle-in-Cell Simulation Studies for Hybrid Laser-Plasma Accelerators and Plasma Eyepieces

Ming Zeng, Alberto Martinez de la Ossa, and Jens Osterhoff

Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany
E-mail: jens.osterhoff@desy.de

Plasma wakefield accelerators driven by either laser or electron beams have shown great potential for future applications. Output beam quality from plasma has improved tremendously over the past decade. This, to a large extend, was enabled by progress in high-performance computing and numerical techniques based on particle-in-cell simulations. In this proceedings paper, we present two recent simulation studies, on hybrid plasma accelerators and on plasma-based laser focusing, opening new avenues in the application of compact accelerators and for the generation of high-brightness electron beams.

1 Introduction

Laser-driven (LWFAs)[1] and beam-driven (PWFAs) plasma wakefield accelerators[2] can generate accelerating electric fields of 10 to 100 GV/m, many orders of magnitude higher than achievable with conventional radio frequency (RF) technology. This enables a path towards greatly miniaturised and, therefore, cost-effective particle accelerators with a significant societal impact potential. Such a development may allow for the proliferation of high-energy accelerator technology far beyond national-lab-scale research centres and could multiply access opportunities to, for example, accelerator-based cutting-edge photon science machines such as free electron lasers. Furthermore, it will notably lower the financial entrance bar for developing countries to engage in accelerator research.

Great progress over the past decades has led to the successful demonstration of GeV-range beams accelerated on only centimetre scales in plasma[3] and improved control and stability.[4] However, achieving the required beam quality for the envisioned disruptive application of novel accelerators in photon science, material science, medicine, and particle physics remains a challenge. Due to the micrometre spatial scale and femtosecond duration of the acceleration structure, the experimental diagnostics and optimisation of LWFAs and PWFAs are challenging. Thus, computer simulation algorithms, in particular Particle-in-Cell (PIC) techniques, have proven successful and are a cornerstone of predictive plasma accelerator research.[5, 6]

In PIC algorithms, fields (electric force, magnetic force, currents, *etc.*) are deposited onto grids (or cells) while particles are modelled as clouds (or shaped macro-particles) with the volume of one or a few adjacent cells. One macro-particle commonly contains a large number of real particles such that the number of particles for computation purposes is largely reduced. Time steps in PIC are chosen such that charged macro particles and fields only propagate to nearby cells making massive parallelisation possible.

In this proceeding, we demonstrate progress in our simulation studies of laser and plasma wakefield accelerators. In Sec. 2 we show a pioneering study on hybrid Laser-Plasma Accelerators (LPWFA), which deploy an LWFA-generated electron beam to drive

a PWFA stage. In Sec. 3 we present our concept of a laser-plasma telescope system with a plasma eyepiece to flexibly adjust the laser spot sizes.

2 Simulation Studies on Hybrid Laser-Plasma Accelerators

2.1 Motivation

PWFAs are considered to offer improved control over the process of injection and acceleration of a witness beam compared to LWFAs. One of the reasons lies in the way that electron-beam drivers propagate through plasma, being immune to the dephasing and diffraction effects affecting the performance of LWFAs. Thus, in principle, PWFAs may allow for longer and more stable acceleration lengths, resulting in higher energy and potentially higher quality beams. However, a major limitation for the widespread use of PWFA technology is the requirement of a large-scale accelerator facility providing the drive beams. In contrast, with a comparably compact size, high-power laser facilities for LWFAs have proliferated over the past decade and electron beams with several hundreds of picocoulomb charge are nowadays routinely generated by LWFAs.[7, 8] In a hybrid LWFA-driven PWFA (LPWFA), the LWFA-generated high-current electron beam is used as driver

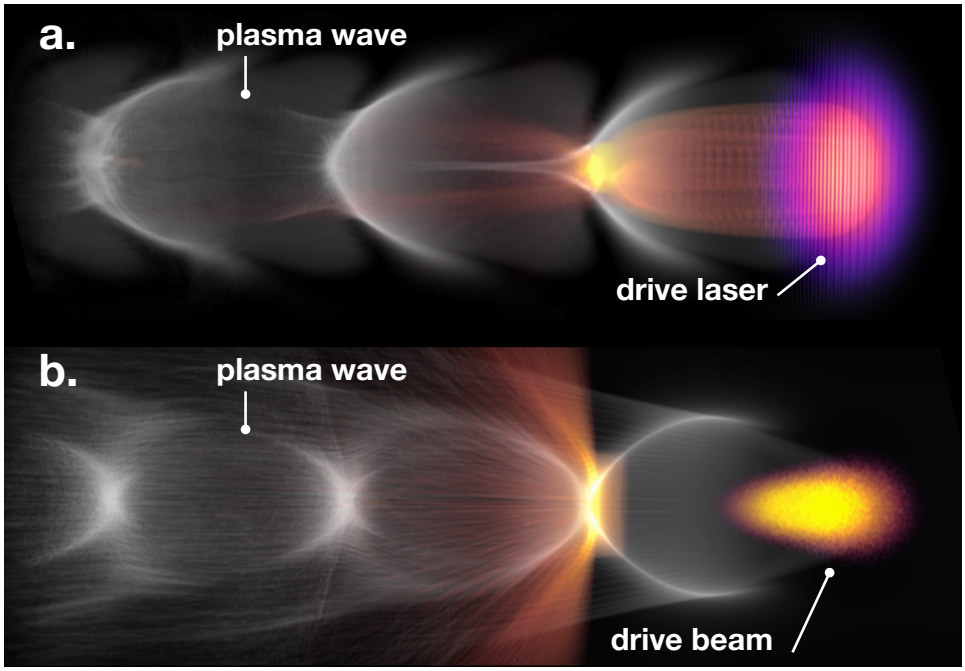


Figure 1. OSIRIS PIC simulations for an LWFA (**a.**) and a PWFA (**b.**). The figures show a laser and electron-beam driving a plasma-electron density wave. The density of electrons being injected is also shown. In both simulations injection is triggered by means of induced ionisation. In a hybrid Laser-Plasma Accelerator (LPWFA) both plasma acceleration principles are combined for the production of superior quality beams from a particularly compact setup.

of a subsequent PWFA stage. In the PWFA stage, a new witness beam with largely improved quality is to be generated and accelerated to higher energies.[9–12] In essence, the PWFA stage operates as a beam brightness and energy transformer of the LWFA output, aiming to reach the demanding beam quality requirements of *e. g.* a free-electron laser (FEL), without sacrificing the small spatial footprint and the relatively low cost offered by LWFAs.

2.2 PIC Simulations for Conceptual Design

In order to assess the feasibility of the hybrid LPWFA concept, detailed 3D PIC simulations were performed with the code OSIRIS[13] for the LWFA and PWFA stages. As illustrated in Fig. 1, the OSIRIS code is capable to reproduce with high fidelity the physics of the process of injection and acceleration of witness beams for both kind of plasma accelerators. Supported by large-scale PIC simulations performed on the JUWELS supercomputer at JSC, a conceptual study for hybrid LPWFAs acting as energy and brightness transformers was recently published.[14] Our results highlight that, by utilising electron beams produced in current LWFA laboratories as drivers of a PWFA, a new generation of ultrashort (femtosecond), high-energy and high-brightness electron beams could be produced. The improvement with respect to the original LWFA beam can be a factor two in energy and about a factor 100 000 in brightness, potentially enabling beam-quality demanding applications such as future compact FELs.

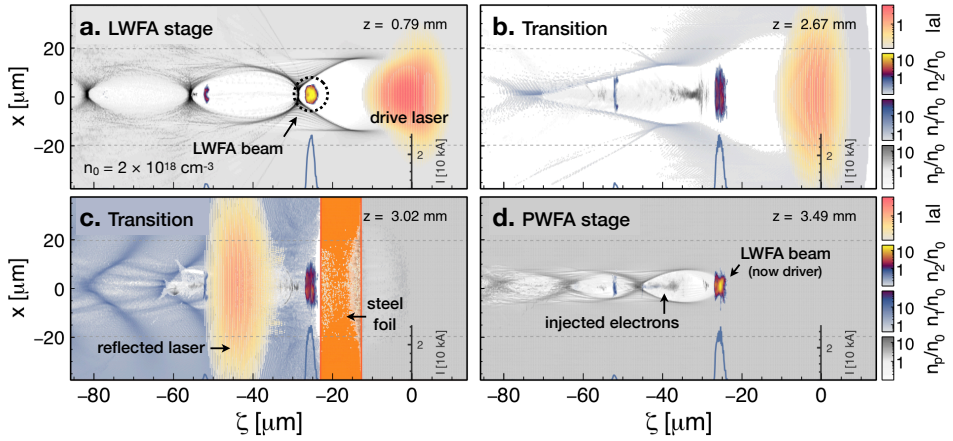


Figure 2. Start-to-end simulation demonstrating the feasibility of a LWFA-to-PWFA transition in the hybrid LPWFA experiment at HZDR. The process is initiated with a laser driving a plasma accelerator in a first gas jet and producing a high-current electron beam (**a.**). Both the laser and the produced beam exit the first stage and enter a second gas jet (**b.**). A 12.5 μm thin steel foil is placed in the beginning of the second jet to reflect the laser, while letting the electron beams go through (**c.**). With the laser removed from second stage, the LWFA-created beam drives a plasma accelerator in which background electrons may be trapped and form a new higher quality witness beam (**d.**).

2.3 PIC Simulations for Proof-of-Concept Experiment

Motivated by the promising simulations results shown in Ref. 14, we have explored the feasibility of this concept through dedicated experiments at Helmholtz-Zentrum Dresden-Rossendorf (HZDR),[15] using the DRACO laser system[16] for the LWFA stage and the thereby produced electron beam as driver for a subsequent PWFA stage. First experimental results are outstanding, demonstrating for the first time that LWFA beams can drive a high-gradient PWFA by themselves at a high plasma density ($\sim 4 \times 10^{18}$ cm^{-3}), where accelerating gradients surpassing 100 GV/m are expected. In conjunction with the direct observation of beam-driven plasma waves, evidence of the acceleration of a new witness beam in the PWFA stage has been observed. Dedicated PIC simulations reproducing the concatenation of both stages have been performed, exhibiting a remarkable qualitative similarity with the experiments, and demonstrating accelerating gradients higher than 100 GV/m in the dedicated PWFA stage. This kind of start-to-end simulations (*c. f.* Fig. 2), performed on JUWELS at JSC, represent a milestone by themselves as they address for the first time key aspects of the LWFA-to-PWFA transition process.

An article containing full details is being reviewed for publication.[17]

3 Simulation Studies on Plasma Eyepieces for Petawatt-Class LWFAs

3.1 Motivation

Nowadays petawatt lasers have become high-priority tools for studying the intrinsic properties of the microscopic physical world.[18] Manipulation of such powerful lasers is a big challenge due to the lack of high damage-threshold optical materials. The current solution is to use large beam apertures so that the laser power is spread across a large area of the optical element to prevent damage. As a consequence the diameter of the final focusing mirror needs to be of the order of a considerable fraction of a metre to prevent breakdown of the optic. This leads to the focusing system being costly and difficult to replace. One has to prepare a few focusing systems with different focal lengths for different applications, which makes the difficulty more intractable. For LWFA applications usually a fixed moderately relativistic laser intensity of the order of 10^{18} $\mathrm{W/cm}^{-3}$ at focus is required, which means increasing focal spot sizes and focal lengths with the increasing laser power. This may lead to extremely long focal lengths of the order of 100 (1000) metres for future 10 (100) petawatt lasers.

We have introduced a telescope system to solve these difficulties.[19] As illustrated in Fig. 3, a laser beam is pre-focused by conventional optics with the focal length of f_0 to position z_0 in vacuum with the waist radius of w_0 and normalised peak vector potential amplitude of $a_{0\mathrm{peak}}$. Then it propagates in vacuum for a distance d and touches the vacuum-plasma interface at z_1. Afterwards the laser undergoes self-focusing[20] and is refocused at z_2 with the spot size radius of w_2 and normalised peak vector potential amplitude of $a_{2\mathrm{peak}}$, where w_2 is the first local maximum of laser spot size radius in the plasma. The thickness of the plasma required $l = z_2 - z_1$ is referred to as the thickness of the plasma eyepiece. Finally, the total function of this setup is to focus the laser beam to an effective spot size of w_2 within a distance of $f = f_0 + d + l$. Due to the strong plasma response to the laser beam, d and l are usually much smaller than f_0 so that $f \approx f_0$. Thus, the advantage of such a setup is obvious: a conventional focusing optic for reaching the focal spot

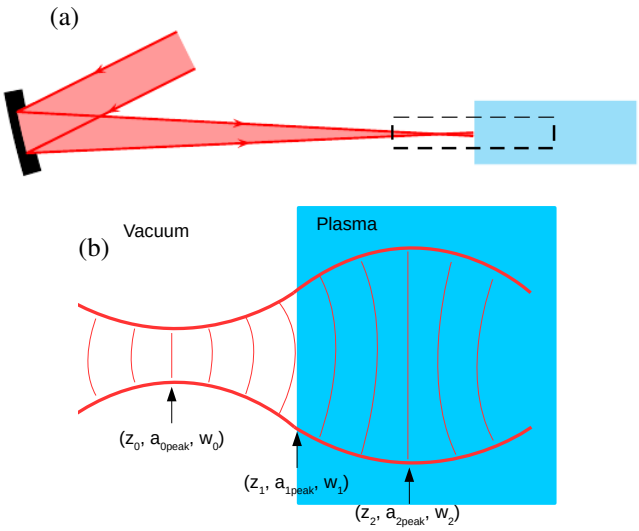


Figure 3. Schematic view of the laser profile (thicker red curves) evolution in the plasma eyepiece region. The laser is pre-focused in vacuum by conventional optics as shown in (a), with the dashed box region shown in (b). The focal plane is at z_0 and laser peak normalised vector potential amplitude is $a_{0\mathrm{peak}}$ and spot size is w_0. After a distance $d = z_1 - z_0$ the laser enters the plasma region start at z_1. Due to the self-focusing effect of lasers in plasmas, the laser wavefront (thinner red curves) is bent concavely and becomes flat again at z_2. As a consequence the laser spot size reaches a local maximum of w_2.

size of w_2 would have the focal length of fw_2/w_1 which can be several times longer than our setup. Moreover, one can easily adjust the effective focal spot size w_2 by changing the plasma density and d with the telescope system, while in a conventional focusing system one has to replace the large and expensive focal optics for different focal spot sizes.

We show one simulation example performed by OSIRIS[13] in Fig. 4. The simulation is initialised ($t = 0$) at around $z_1 = 0$ with the laser focal plane at $k_p z_0 = -60$, where $k_p = \sqrt{4\pi r_e n_p}$ is the plasma wave number, n_p is the background plasma density and $r_e \approx 2.82 \times 10^{-15}$ m is the classical electron radius. The laser focal waist radius is $k_p w_0 = 4$, thus the laser spot size at the vacuum-plasma interface is $k_p w_1 = 5$. The simulation grid information is the same as the simulations shown in Sec. 3.2. After propagating a distance $k_p l = 76$ ($\omega_p t = 76$), the laser spot size reaches a local maximum of $k_p w_2 = 6.2$ due to self-focusing effect.

3.2 Scan of a 4-Dimensional Parameter Space

We have done an analytical study of the plasma eyepiece which is based on non-perturbed plasma assumptions.[19] In the blowout regime the ponderomotive force of the laser is so strong that almost all of the plasma electrons are driven out of the laser axial region,[21, 22] thus the non-perturbed plasma model reaches its limits and PIC simulations are required.

We have performed full 3-dimensional PIC simulations on JUWELS using the code OSIRIS[13] to scan the 4-dimensional parameter space (k/k_p, $k_p w_0$, $a_{0\mathrm{peak}}$, $k_p d$). The longitudinal simulation box size is $10k_p^{-1}$, and the lengths of the two transverse dimensions

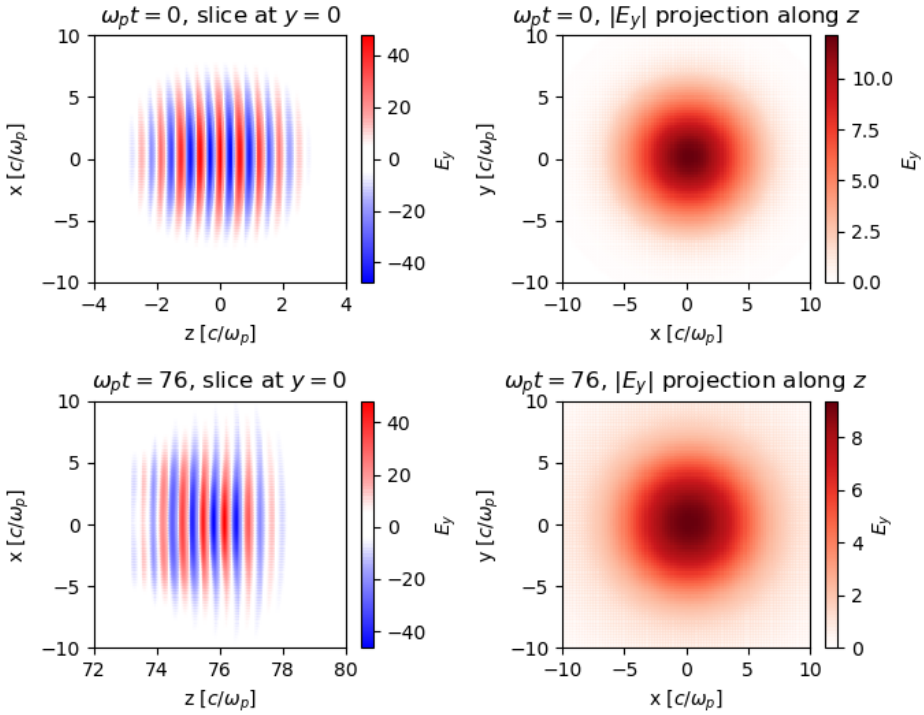


Figure 4. An example simulation of the laser evolution in the plasma eyepiece region. The parameters are $k/k_p = 10$, $k_p w_0 = 4$, $k_p d = 60$, $a_{0\mathrm{peak}} = 10$ and laser FWHM duration is $\omega_p \tau = 4$. The plasma region starts at $z_1 = 0$ and is half-infinitely long. The simulation started at the vacuum-plasma interface $z_1 = 0$ at $t = 0$ (top two subplots). At $k_p z_2 = 76$ (bottom two subplots) the laser spot size reaches a local maximum of $k_p w_2 \approx 6.2$. The left column subplots are the slices of the laser electric field side view, and the right column subplots are the projections of the absolute value of the laser electric field to the x-y plane (front view).

k/k_p	Low resolution	High resolution
10	256×256^2	512×512^2
20	512×256^2	1024×512^2
30	512×256^2	1024×512^2
40	1024×256^2	2048×512^2

Table 1. Longitudinal and transverse number of cells used for different values of k/k_p in the simulations. Two resolutions are used to confirm the convergence of the simulations.

are taken as 10 times of the estimated value of w_2 (to be discussed below). Two resolutions have been adopted to confirm the convergence of the simulations according to Tab. 1, while the higher resolution results are taken as the final results. The full-width-at-half-maximum (FWHM) laser pulse duration is fixed as $\omega \tau_{\mathrm{FWHM}} = 4$. The simulations are initialised at $t = 0$ with laser centre at $z = z_1 = 0$, while the laser focal planes are at $-k_p d$.

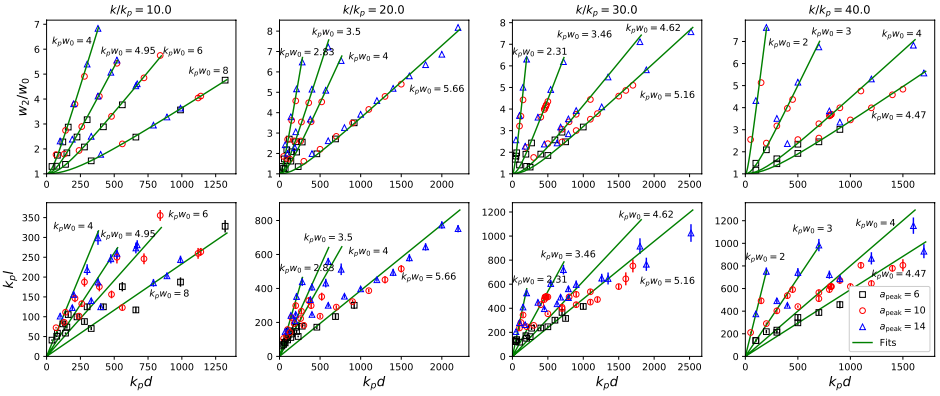


Figure 5. Simulation results for w_2/w_0 (row 1) and l (row 2) *vs.* d for different laser and plasma parameters (k/k_p, k_pw_0, $a_{0\mathrm{peak}}$). In each subplot, the results for $a_{0\mathrm{peak}} = 6$, 10 and 14 are marked as black squares, red circles and blue triangles, respectively. The green lines are the fits of the data with the same k/k_p and k_pw_0 by Eq. 1, with k_pw_0 near the green lines. The laser wave number parameter normalised to the plasma wave number k/k_p takes the values 10, 20, 30 and 40 as shown on the top of each column.

In each diagnostic point of the simulations, the absolute value of the laser electric field is projected to the $x-y$ plane (front-view), and two-dimensional Gaussian fit (with the function $E = E_0 \exp\left[-(x-x_0)^2/w^2 - (y-y_0)^2/w^2\right]$) is performed for the laser radius w. The values of w_2 are measured by taking the local maximum of w *vs.* z and the values of plasma eyepiece thickness $l = z_2 - z_1$, where z_2 is the location of the local maximum of w.

Simulation results for w_2 and l are collected and plotted in Fig. 5. We have found that w_2 can be estimated by a function

$$\frac{w_2}{w_0} = \sqrt{1 + \frac{d^2}{\zeta^2}} \tag{1}$$

where

$$k_p\zeta \approx (0.950 - 0.0027k/k_p)k_pz_R - 1.17k/k_p - 12.6 \tag{2}$$

From Fig. 5 we have also found l almost linear depends on d, *i. e.*

$$l = \chi d \tag{3}$$

where

$$\chi \approx 21.0\,(k_pw_0)^{-2.08} \tag{4}$$

Eqs. 1-4 predict excellent scalability of the telescope system including a plasma eyepiece. In a petawatt-class laser system one can largely reduce the space required for the focus system. We show two exemplary simulations in Fig. 6.

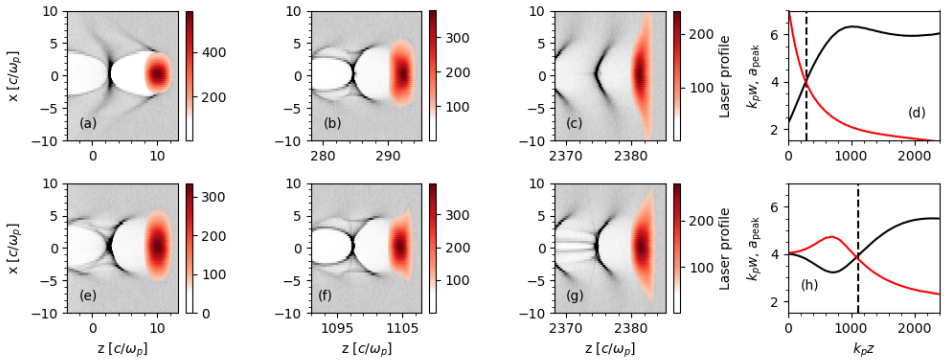


Figure 6. Comparative simulations with (a to d) and without (e to h) a plasma eyepiece. The effective laser spot size adjusted by the plasma eyepiece in (a) - (d) is approximately the laser spot size in the later case, but the former case has a much reduced focal length of the pre-focusing conventional optics system.

4 Conclusion

We have demonstrated two schemes towards future laser and plasma wakefield accelerators. The first is the hybrid laser-plasma wakefield accelerator, which largely reduces the setup scale of a conventional PWFA and promises the possibility to perform PWFA studies in a small size laser laboratory. The second is the laser-plasma telescope system using a plasma eyepiece, which largely reduces the focal space required in a LWFA driven by a petawatt-class laser, and enables the scan of laser spot size parameters. Our studies using large-scale parallelled PIC simulations demonstrate the potential of supercomputing for calculating complex systems such as laser-plasmas.

Acknowledgements

We thank the OSIRIS consortium (IST/UCLA) for access to the OSIRIS code. We also gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC). This work is supported by the Helmholtz MT ARD scheme and the Helmholtz ZT-0009 project.

References

1. T. Tajima and J. M. Dawson, *Laser Electron Accelerator*, Phys. Rev. Lett. **43**, 267–270, 1979.
2. P. Chen, J. M. Dawson, R. W. Huff, and T. Katsouleas, *Acceleration of Electrons by the Interaction of a Bunched Electron Beam with a Plasma*, Phys. Rev. Lett. **54**, 693–696, 1985.

3. W. P. Leemans, B. Nagler, A. J. Gonsalves, Cs. Tóth, K. Nakamura, C. G. R. Geddes, E. Esarey, C. B. Schroeder, and S. M. Hooker, *GeV electron beams from a centimetre-scale accelerator*, Nat. Phys. **2**, 696, 2006.
4. J. Osterhoff, A. Popp, Zs. Major, B. Marx, T. P. Rowlands-Rees, M. Fuchs, M. Geissler, R. Hörlein, B. Hidding, S. Becker, E. A. Peralta, U. Schramm, F. Grüner, D. Habs, F. Krausz, S. M. Hooker, and S. Karsch, *Generation of Stable, Low-Divergence Electron Beams by Laser-Wakefield Acceleration in a Steady-State-Flow Gas Cell*, Phys. Rev. Lett. **101**, 085002, 2008.
5. C. K. Birdsall and D. Fuss, *Clouds-in-clouds, clouds-in-cells physics for many-body plasma simulation*, Journal of Computational Physics **3**, 494–511, 1969.
6. J. M. Dawson, *Particle simulation of plasmas*, Rev. Mod. Phys. **55**, 403–447, 1983.
7. J. P. Couperus, R. Pausch, A. Köhler, O. Zarini, J. M. Krämer, M. Garten, A. Huebl, R. Gebhardt, U. Helbig, S. Bock, K. Zeil, A. Debus, M. Bussmann, U. Schramm, and A. Irman, *Demonstration of a beam loaded nanocoulomb-class laser wakefield accelerator*, Nature Communications **8**, 487, 2017.
8. M. Zeng, M. Chen, Z.-M. Sheng, W. B. Mori, and J. Zhang, *Self-truncated ionization injection and consequent monoenergetic electron bunches in laser wakefield acceleration*, Physics of Plasmas **21**, 030701, 2014.
9. B. Hidding, G. Pretzler, J. B. Rosenzweig, T. Königstein, D. Schiller, and D. L. Bruhwiler, *Ultracold Electron Bunch Generation via Plasma Photocathode Emission and Acceleration in a Beam-Driven Plasma Blowout*, Phys. Rev. Lett. **108**, 35001, 2012.
10. A. Martinez de la Ossa, J. Grebenyuk, T. J. Mehrling, L. Schaper, and J. Osterhoff, *High-Quality Electron Beams from Beam-Driven Plasma Accelerators by Wakefield-Induced Ionization Injection*, Phys. Rev. Lett. **111**, 245003, 2013.
11. G. Wittig, O. Karger, A. Knetsch, Y. Xi, A. Deng, J. B. Rosenzweig, D. L. Bruhwiler, J. Smith, G. G. Manahan, Z.-M. Sheng, D. A. Jaroszynski, and B. Hidding, *Optical plasma torch electron bunch generation in plasma wakefield accelerators*, Phys. Rev. ST Accel. Beams **18**, 081304, 2015.
12. A. Martinez de la Ossa, Z. Hu, M. J. V. Streeter, T. J. Mehrling, O. Kononenko, B. Sheeran, and J. Osterhoff, *Optimizing density down-ramp injection for beam-driven plasma wakefield accelerators*, Phys. Rev. Accel. Beams **20**, 091301, 2017.
13. 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*, in Computational Science – ICCS 2002, P. M. A. Sloot, A. G. Hoekstra, C. J. K. Tan, and J. J. Dongarra (Editors), Springer, 342–351, 2002.
14. A. Martinez de la Ossa, R. W. Assmann, M. Bussmann, S. Corde, J. P. Couperus Cabadağ, A. Debus, A. Döpp, A. Ferran Pousa, M. F. Gilljohann, T. Heinemann, B. Hidding, A. Irman, S. Karsch, O. Kononenko, T. Kurz, J. Osterhoff, R. Pausch, S. Schöbel, and U. Schramm, *Hybrid LWFA–PWFA staging as a beam energy and brightness transformer: conceptual design and simulations*, Phil. Trans. R. Soc. A **377**, 20180175, 2019.
15. T. Heinemann, R. W. Assmann, J. P. Couperus, B. Hidding, A. Knetsch, O. Kononenko, A. Köhler, T. Kurz, U. Schramm, O. Zarini, A. Irman, and A. Martinez de la Ossa, *Investigating the Key Parameters of a Staged Laser- and Particle Driven Plasma Wakefield Accelerator Experiment*, in Proceedings of International Particle Accelerator Conference (IPAC '17), 1703–1706, 2017.

16. U. Schramm, M. Bussmann, A. Irman, M. Siebold, K. Zeil, D. Albach, C. Bernert, S. Bock, F. Brack, J. Branco, J. P. Couperus, T. E. Cowan, A. Debus, C. Eisenmann, M. Garten, R. Gebhardt, S. Grams, U. Helbig, A. Huebl, T. Kluge, A. Köhler, J. M. Krämer, S. Kraft, F. Kroll, M. Kuntzsch, U. Lehnert, M. Loeser, J. Metzkes, P. Michel, L. Obst, R. Pausch, M. Rehwald, R. Sauerbrey, H. P. Schlenvoigt, K. Steiniger, and O. Zarini, *First results with the novel petawatt laser acceleration facility in Dresden*, Journal of Physics: Conference Series **874**, 012028, 2017.
17. T. Kurz, T. Heinemann, M. F. Gilljohann, Y. Y. Chang, J. P. Couperus Cabadağ, A. Debus, O. Kononenko, R. Pausch, S. Schöbel, R. W. Assmann, M. Bussmann, H. Ding, J. Götzfried, A. Köhler, G. Raj, S. Schindler, K. Steiniger, O. Zarini, S. Corde, A. Döpp, B. Hidding, S. Karsch, U. Schramm, A. Martinez de la Ossa, and A. Irman, *Demonstration of a compact plasma accelerator powered by laser-accelerated electron beams*, 2019, arXiv:1909.06676 [physics.acc-ph].
18. E. Cartlidge, *Physicists are planning to build lasers so powerful they could rip apart empty space*, Science, 2018, doi:10.1126/science.aat0939.
19. M. Zeng, A. Martinez de la Ossa, K. Poder, and J. Osterhoff, *Plasma Lenses for Relativistic Laser Beams in Laser Wakefield Accelerators*, 2019, arXiv:1901.07974 [physics.plasm-ph].
20. C. E. Max, J. Arons, and A. B. Langdon, *Self-Modulation and Self-Focusing of Electromagnetic Waves in Plasmas*, Phys. Rev. Lett. **33**, 209–212, 1974.
21. A. Pukhov and J. Meyer-ter Vehn, *Laser wake field acceleration: the highly nonlinear broken-wave regime*, Appl. Phys. B **74**, 355–361, 2002.
22. W. Lu, M. Tzoufras, C. Joshi, F. S. Tsung, W. B. Mori, J. Vieira, R. A. Fonseca, and L. O. Silva, *Generating multi-GeV electron bunches using single stage laser wakefield acceleration in a 3D nonlinear regime*, Phys. Rev. ST Accel. Beams **10**, 061301, 2007.

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. Bücker, 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. Bücker, 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, 418 pages
ISBN: 978-3-95806-109-5

NIC Series Volume 49
NIC Symposium 2018 – Proceedings
22 – 23 February 2018 | Jülich, Germany
edited by K. Binder, M. Müller, A. Trautmann (2018), vi, 448 pages
ISBN: 978-3-95806-285-6

NIC Series Volume 50
NIC Symposium 2020 – Proceedings
27 – 28 February 2020 | Jülich, Germany
edited by M. Müller, K. Binder, A. Trautmann (2020), v, 424 pages
ISBN: 978-3-95806-443-0

The John von Neumann Institute for Computing (NIC) was established in 1998 by Forschungszentrum Jülich and Deutsches Elektronen-Synchrotron DESY to support the supercomputeroriented 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 10th NIC Symposium are from projects that have been supported by the supercomputers JURECA, JUQUEEN and JUWELS in Jülich. They cover selected topics in the fields of Astrophysics, Biology and Biophysics, Elementary Particle Physics, Materials Science, Theoretical Condensed Matter, Soft Matter Science, Earth and Environment, Computer Science and Numerical Mathematics, Fluid Mechanics and Engineering, and Plasma Physics.

JÜLICH Forschungszentrum

GSI

Deutsches Elektronen-Synchrotron DESY

NIC Series Volume 50

ISBN 978-3-95806-443-0