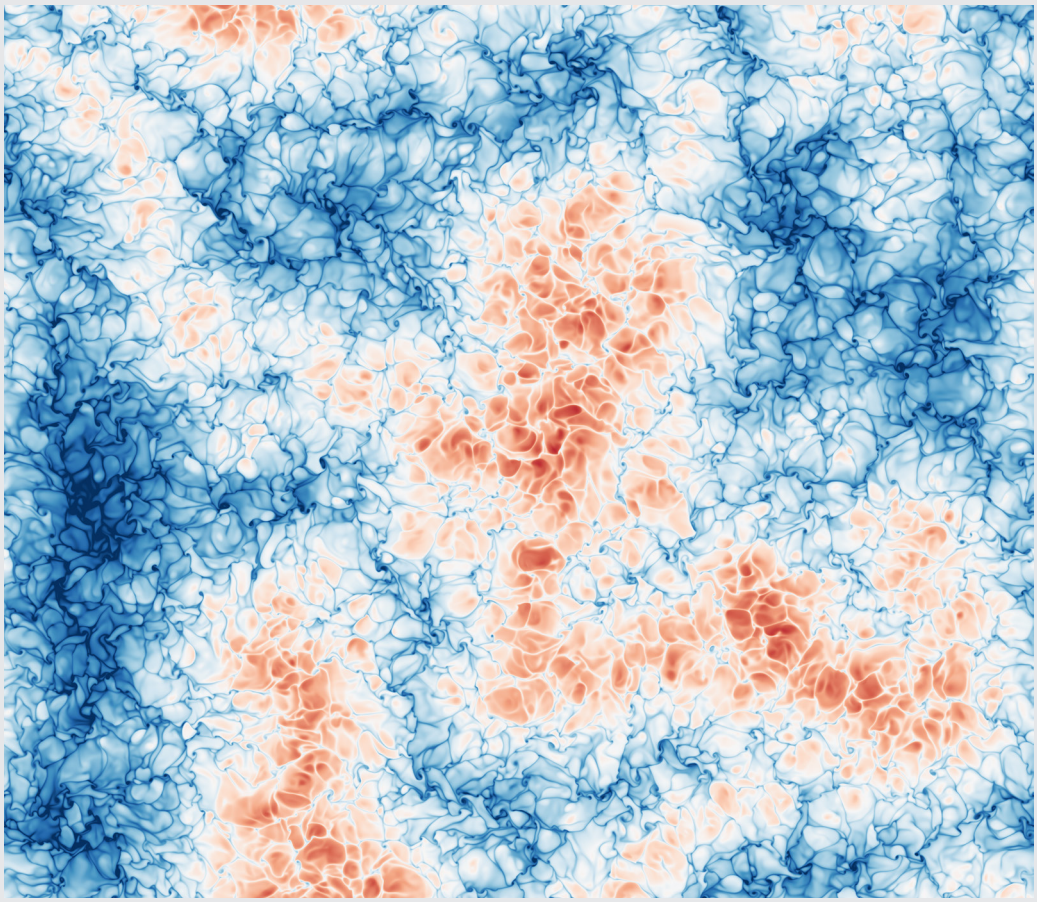
John von Neumann Institute for Computing

NIC Symposium 2022

29 – 30 September 2022 | Jülich, Germany

M. Müller, Ch. Peter, A. Trautmann (Editors)

Proceedings



NIC Series

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 51

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

NIC Symposium 2022

29 – 30 September 2022 I Jülich, Germany

M. Müller, Ch. Peter, A. Trautmann (Editors)

Proceedings

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 51

ISBN 978-3-95806-646-5

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 2022

We thank J. Schumacher and Ph. Vieweg (Technische Universität Ilmenau) for the image used in the cover design.

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 51

ISBN 978-3-95806-646-5

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

This is an Open Access publication distributed under the terms of the Creative Commons Attribution License 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preface

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

Christine Peter
Department of Chemistry, University of Konstanz, 78457 Konstanz, Germany
christine.peter@uni-konstanz.de

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

On September 29 and 30, 2022, computational scientists will convene in Jülich for the 11th NIC symposium to present their exciting research – illustrating the diverse range of modern, computational science at the John von Neumann Institute for Computing (NIC). For many of us, the previous NIC symposium in February 2020 was the last face-to-face event before the COVID-19 pandemic struck, and we have shifted the present symposium to the fall in order to allow for a lively in-person exchange. It has been a tradition that the biannual symposia are accompanied by proceedings that address both, computational scientists and practitioners as well as the general public, interested in the advancement of computational science and its applications in diverse, contemporary research fields. We are delighted that also this year it was possible to showcase the breadth of high-performance computing research, with contributions from elementary particle physics, astrophysics, and statistical physics of hard and soft condensed matter, computational chemistry and life sciences, as well as material science, fluid-dynamics engineering, and climate research.

Since its foundation 24 years ago, the NIC continuously provides the scientific community with essential high-performance computing resources, training, and technical support on the highest performance level. Within the Gauss Centre for Supercomputing (GCS) the Jülich Supercomputing Centre (JSC) provides a modular supercomputer, comprised of a versatile, easy-to-use, CPU-based cluster module that has been installed 2018 and a GPU-based booster module. The latter has been deployed in 2020 and features more than 3700 NVIDIA A100 GPUs. By virtue of the efficient, dedicated and much appreciated training and user support of the JSC, this new architecture has been swiftly adopted across the different scientific communities. This has been an important step in moving forward the frontiers of established simulation techniques e.g., in elementary particle physics or extremely large, particle-based simulations in materials physics and engineering and has also enabled new applications using big scientific data analysis and machine-learning strategies. For instance, in the MLPerf Training HPC competition in November 2021 researchers of the JSC and the Steinbuch Centre for Computing at the Karlsruhe Institute of Technology (Helmholtz AI) were able to perform the fastest ever and most computationally intensive AI calculations in Europe.

As of June 2022, the GPU-accelerated JUWELS Booster still ranks at place 11 of the TOP 500 list of the fastest supercomputers in the world (18 of the GREEN 500 list) but a continuous strong support remains essential to keep up with the fast pace of development in the U.S.A., China, and Japan, as well as other European countries in order to provide top-notch computational facilities to the German scientific community. In this context it is very exciting that the European High Performance Computing Joint Undertaking (EuroHPC JU) has selected the Forschungszentrum Jülich – a partner of GCS – to operate the first exascale supercomputer in Europe. The overall costs for the system amount to 500 million €, half of which is provided by EuroHPC JU and a further 250 million euros stem in equal parts from the German Federal Ministry of Education and Research (BMBF) and the Ministry of Culture and Science of the State of North Rhine-Westphalia (MKW NRW). The computer with the name JUPITER ("Joint Undertaking Pioneer for Innovative and Transformative Exascale Research") will be installed as of 2023.

From all projects at the NIC,[a] the contributions to the symposium and the proceedings have been selected to represent the diversity of outstanding research enabled by the NIC supercomputing resources. NIC computing time is granted by a stringent peer-review system that focuses on the scientific quality of the proposed research. The international pool of expert reviewers of the NIC peer-review board, headed by Prof. Dietrich Wolf, plays a vital role in sustaining the very high quality of the projects and in fairly and effectively allocating the valuable computational resources. Therefore it is a great pleasure to sincerely thank the distinguished reviewers for their engagement and time that they invest in fulfilling this essential task.

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 modern algorithms and computational strategies in different fields of natural science and engineering. An exciting step towards the future of information technology is the usage of quantum computers. In collaboration with the Canadian provider D-wave Systems, a quantum annealer with more than 5,000 qubits has been put into operation at the JSC on January 17, 2022. This quantum annealer is part of the Jülich UNified Infrastructure for Quantum computing (JUNIQ). In order to provide an introduction to these new computational opportunities for the NIC community, the proceedings feature an extra chapter on quantum computing, comprised of contributions from the JSC director, Prof. Thomas Lippert, Prof. Kristel Michielsen, and Prof. David DiVincenzo.

At the semiannually meetings, the NIC peer-review board, chaired by Dietrich Wolf, awards the title "NIC excellence project" [b] out of all proposals. These outstanding projects are highlighted in a special section, featuring the following projects:

- Frithjof Anders, Universität Dortmund, Coherent control of the electronic and nuclear spins in quantum dot ensembles
- Hartmut Wittig, Universität Mainz, Lattice QCD with Wilson Quarks at zero and non-zero Temperature

[a] https://www.john-von-neumann-institut.de/en/research/projects/project-distribution
[b] https://www.john-von-neumann-institut.de/en/research/nic-excellence-projects

- Ulf-G. Meißner, Universität Bonn und Forschungszentrum Jülich, Nuclear Lattice Simulations
- Markus Heyl, Max-Planck-Institut für Physik komplexer Systeme, Dresden, Quantum many-body dynamics from neural networks with GPU acceleration
- Szabolcs Borsanyi, Bergische Universität Wuppertal, The standard model under investigation
- Kurt Kremer, Max-Planck-Institut für Polymerforschung, Mainz, Germany, Bridging drastically coarse-grained and microsopic descriptions in hierarchical modelling of Soft Matter. Application: Non-linear viscoelasticity of polymer melts

The NIC also supports multiple 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 center, and develop and apply sophisticated simulation strategies for important research topics, ranging from biology to high-energy physics. Their recent results are highlighted in a special chapter in the proceedings.

We hope that you will be thrilled, 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.

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, Ms. Gina Bongartz and Ms. Janina Liebmann for their valuable help in organising the 11th NIC Symposium in Jülich.

Since 2021, Christine Peter (Universität Konstanz) is acting as deputy chairwoman of the NIC scientific council, following Kurt Binder (Johannes-Gutenberg Universität, Mainz). At this occasion we would like to take the opportunity to thank Kurt Binder for his leadership and continuous support of the NIC in various roles during the last two decades. It has been a great pleasure to collaborate with Kurt Binder, and we have benefited from his enormous experience, guidance, and enthusiasm.

Jülich, September 2022

Marcus Müller Christine Peter Alexander Trautmann

[c] https://www.john-von-neumann-institut.de/en/research/research-groups

Contents

Quantum Computing

Perspectives of Quantum Computing at the Jülich Supercomputing Centre
T. Lippert, K. Michielsen **3**

Quantum Computers in 2022
D. DiVincenzo **25**

Hybrid Quantum Classical Simulations
D. Willsch, M. Jattana, M. Willsch, S. Schulz, F. Jin, H. De Raedt, K. Michielsen **31**

The NIC Research Groups

TopSuite: Collection of Deep Learning-Based Metamethods to Predict and Evaluate Protein Structures and Properties
F. Koenig, D. Mulnaes, K. van der Weg, H. Gohlke **47**

The Columbia Plot of the QCD Phase Transition
F. Cuteri, A. D'Ambrosio, R. Kaiser, O. Philipsen, A. Schön, A. Sciarra **55**

Native-Contact Guided Replica Exchange Simulations of Structured RNA
A. Voronin, A. Schug **65**

The NIC Excellence Projects

Spin Dynamics in Quantum Dot Ensembles
I. Kleinjohann, A. Fischer, F. B. Anders **77**

Searching for the H Dibaryon in Lattice QCD
J. R. Green, A. D. Hanlon, P. M. Junnarkar, M. Padmanath, S. Paul, H. Wittig **87**

Nuclear Lattice Effective Field Theory: Status A.D. 2022
U.-G. Meißner **97**

Quantum Many-Body Dynamics from Neural Networks with GPU Acceleration
M. Schmitt, H. Burau, M. Heyl **111**

Cross-Correlators of Conserved Charges from Lattice QCD
S. Borsányi, Z. Fodor, J. N. Günther, P. Parotto, A. Pásztor **123**

Molecular Dynamics Study of the Glass Transition of Highly Entangled Polymer Melts
H.-P. Hsu, T. Stuehn, K. C. Daoulas, K. Kremer **135**

Astrophysics

Introduction
P. L. Biermann **147**

The Reionisation of the Local Universe in the HESTIA Suite
L. Conaboy, I. T. Iliev, N. I. Libeskind **151**

Growth of Seed Black Holes in Galactic Nuclei
R. Spurzem, F. Rizzuto, M. Arca Sedda, A. Kamlah, P. Berczik, Q. Shu, A. Tanikawa, T. Naab **159**

Theoretical Chemistry

Introduction
C. Peter **175**

Two-Dimensional Hydrogen Structures at Ultra-High Pressures
F. Calcavecchia, T. D. Kühne, M. Holzmann **177**

Molecular Origin of the Unusual Proton/Fluoride Stoichiometry of CLC-Type Fluoride Transporters
M. G. Chiariello, M. Alfonso-Prieto, E. Ippoliti **189**

Elementary Particle Physics

Introduction
S. Schaefer **201**

The Quark and Gluon Momentum Fractions in the Pion
C. Alexandrou, S. Bacchio, G. Bergner, J. Finkenrath, A. Gasbarro, K. Hadjiyiannakou, K. Jansen, B. Kostrzewa, K. Ottnad, M. Petschlies, F. Pittler, F. Steffens, C. Urbach, U. Wenger **203**

Baryons, their Flavour Structure and Low Energy Constants
G. Bali for the RQCD Collaboration **211**

Towards Reaching Physical Conditions: Lattice QCD Results for Hadron Structure and Quark Masses
K. Jansen for the Extended Twisted Mass Collaboration **221**

Materials Science

Introduction
G. Bihlmayer **231**

Data-Efficient and Physics-Inspired Machine-Learning Models to Treat Complex Catalyst Materials and Reaction Networks
W. Xu, K. Reuter, M. Andersen **233**

***Ab Initio* Simulations for Microelectronics: $ZnGa_2O_4$ and Hexagonal BN**
J. Dabrowski, F. Reichmann, M. Franck, M. Lukosius **245**

Atomic-Scale Insights into Friction from Molecular Dynamics Simulations: Fluorinated Carbon Coatings and PTFE
T. Reichenbach, L. Mayrhofer, G. Moras, M. Moseler **253**

Condensed Matter Theory

Introduction
F. B. Anders **265**

Optical Excitations of Two-Dimensional Semiconductors
M. Rohlfing **267**

Numerical Simulations of Strongly Correlated Electron Systems
F. F. Assaad, B. Danu, F. Goth, A. Götz, Z. Liu, F. Parisen Toldin, M. Raczkowski, G. Rein, T. Sato, J. Schwab, J. Stafusa E. Portela, M. Ulybyshev **279**

Massively Parallel Simulations of Correlated Materials
X.-J. Zhang, N. Rihaisamani, E. Adibi, E. Koch, E. Pavarini **289**

Computational Soft Matter Science

Introduction
K. Kremer **301**

Controlling the Orientation of Mesophases by Processing: Insights from Computer Simulation
N. Blagojevic, M. Müller **303**

Transport in Nanowires and Colloidal Systems and Crystal Growth
K. Beck, B. Gast, K. Hofmann, J. Holder, M. Isele, A. Lüders, F. Müller, M. Ring, R. Schmid, P. Stengele, E. Zander, P. Nielaba **315**

Earth and Environment

Introduction
T. Jung **329**

The Impact of Greenland Ice Sheet Melting on the North Atlantic Simulated with the FESOM Sea Ice-Ocean Model
S. Stolzenberger, R. Rietbroek, C. Wekerle, B. Uebbing, J. Kusche **331**

Molecular Modelling of Hydrothermal Fluids
S. Jahn, M. Schulze, R. Chattopadhyay **341**

Computer Science and Numerical Mathematics

Introduction
D. Kröner **353**

Massively Parallel Likelihood-Free Parameter Inference for Biological Multi-Scale Systems
E. Alamoudi, J. Starruß, N. Bundgaard, R. Müller, F. Reck, F. Graw, L. Brusch, J. Hasenauer, Y. Schälte **355**

Uncertainty Quantification and Resource-Demanding Computer Vision Applications of Deep Learning
J. Burghoff, R. Chan, H. Gottschalk, A. Mütze, T. Riedlinger, M. Rottmann, M. Schubert **367**

Fluid Mechanics

Introduction
C. Stemmer **383**

Simulation of Sound Wave Focusing by Variable Density Medium
S.-R. Koh, I. Zhukov, J.-H. Kim, W. Frings **385**

Oscillatory Control of Compressible Channel Flow Using Direct Numerical Simulations
M. Ruby, H. Foysi **395**

Analysis of the Large-Scale Order in Turbulent Mesoscale Convection
P. P. Vieweg, F. Heyder, J. P. John, J. D. Scheel, J. Schumacher **405**

Identification of Very-Large-Scale Motions in Turbulent Channel Flows by Direct Numerical Simulations and Data-Driven Decomposition
C. Chi, A. J. Smits, S. Wolligandt, H. Theisel, D. Thévenin **415**

Plasma Physics

Introduction
K.-H. Spatschek **427**

Recent Advancements in HPC Simulations of Dynamos and Multiphysics Plasmas
M. Wilbert, F. Allmann-Rahn, A. Giesecke, R. Grauer **429**

Interaction of Extremely Intense Flows of Electromagnetic Energy and QED Processes in Supercritical Fields
L. Reichwein, M. Filipovic, X. F. Shen, K. Jiang, C. Baumann, A. Pukhov **441**

Quantum Computing

Perspectives of Quantum Computing at the Jülich Supercomputing Centre

Thomas Lippert[1,2] **and Kristel Michielsen**[1,3]

[1] Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {th.lippert, k.michielsen}@fz-juelich.de

[2] Department of Computer Science and Mathematics, Goethe Universität Frankfurt, Germany

[3] Faculty for Mathematics, Informatics and Natural Sciences, RWTH Aachen University, Germany

The strategy of the Jülich Supercomputing Centre to apply quantum computers in the Jülich Quantum Computing Infrastructure JUNIQ is presented. Following the JUNIQ quantum computer emulator JUQCS, the D-Wave quantum annealer JUPSI and the soon-to-be-implemented Pasqal quantum simulator JURY, pilot systems based on gates as well as early-stage experimental systems, on their stony path to become universal, will be included. The integration strategy of JUNIQ into the HPC environment of the Jülich Supercomputing Centre (JSC) is outlined. An exemplary application is presented that was implemented on the emulator JUQCS, which emulates an ideal gate-based quantum computer, and enables comparison with the D-Wave quantum annealer JUPSI. Long-term perspectives are highlighted.

1 Introduction

Forty years have passed since Richard Feynman's landmark visions on quantum computing were published[a;1,2]. Today, many scientists are convinced that we are on the threshold of the practical exploitation of quantum effects in scientific and technical computing. Indeed, in early 2022, the Jülich Supercomputing Centre (JSC) could put into operation a special variant of a quantum computer, a so-called quantum annealer (QA), manufactured by the Canadian company D-Wave Systems Inc. ("D-Wave"), a system with 5640 qubits called "D-Wave Advantage" and nicknamed JUPSI[3,4], see Fig. 1.

Figure 1. D-Wave Advantage System JUPSI – JUelich Pioneer for Spin Interference.

[a]…"because nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy. …Can you do it with a new kind of computer – a quantum computer? …Now it turns out, as far as I can tell, that you can simulate this with a quantum system, with quantum computer elements. …For example, the spin waves in a spin lattice …imitating Bose-particles in the field theory. I therefore believe it's true that with a suitable class of quantum machines you could imitate any quantum system, including the physical world. But I don't know whether the general theory of this intersimulation of quantum systems has ever been worked out …Now, what kind of physics are we going to imitate? …the physical world is quantum mechanical, and therefore the proper problem is the simulation of quantum physics." Indeed, as early as 1959, he had spoken of digital computing with quantum systems in an address at Caltech, with the famous headline "There's Plenty of Room at the Bottom!"

A QA is probably closest to Feynman's original vision of quantum computing, as suggested by the quote given in footnote a. A quantum spin array is assigned couplings that encode the problem to be solved. From an initial Hamiltonian, which in the z-basis generates a complete superposition state from a superposition of all the possible 2^L basis states, L being the number of qubits, which as a result of a magnetic field in the x-direction takes the lowest energy state of this operator, the system is adiabatically transferred into the problem Hamiltonian; the quantum spins should remain in the lowest energy state – more precisely, the spin-system at all times takes the lowest eigenstate of the slowly-varying Hamiltonian, and thus, at the end of the process, we know with high probability that the system is in the ground state, and that the measurement of the qubits (spins) at this point of the process should not introduce any more perturbations in the ideal world. This certainty is based on the famous adiabatic theorem of Born and Fock[5].

A QA performs a continuous physical process, *i.e.* the time course of the system is continuous. In this sense, a QA can be considered an *analog* quantum computer.

This is contrasted with the *digital* quantum computer, usually called a *gate-based* or *gate-level* quantum computer (GBQC). A GBQC is also constructed from arrays of spins that are, in contrast to the QA, controllably coupled together. The same universal spin Hamiltonian that allows the QA to be modelled is also used for the GBQC as a description model. The GBQC can manipulate single qubits as well as gates of two or three qubits in such a way that quantum states of the involved qubits can be specifically transformed and influenced. This allows quantum circuits and quantum algorithms to be executed. Usually the qubits are brought into the superposition of 2^N states at the beginning of the operation, this is done here by local Hadamard transformations, leading to a state coherently superposed of all eigenstates. Using 1-qubit and 2-qubit operations (*e.g.*, a CNOT) one can tinker entangled states representing non-local correlations. The trick is to use such manipulations to get the system into the specific eigenstate at the end of the algorithm, that is supposed to solve the posed problem. Given this situation, one then knows with certainty that one will get an unambiguous result, since the measurement, just as in the case of the QA, no longer introduces perturbations. However, on the one hand, only a few algorithms are known that lead to such a maximum amplitude amplification, but fortunately there are some important ones that amplify only a few amplitudes, *i.e.*, suppress most of them, – Shor's algorithm is of this type – and thus allow for a statistical determination of the solution. On the other hand, the qubit and gate manipulations of today's GBQC are not yet at a level of fidelity that leads to generally useful results – here quantum error correction promises a way out[6].

The currently existing digital quantum computing systems (*e.g.*, from IBM, Google, Rigetti Computing, IonQ, AQT, IQM or eleQtron) are all at such a prototype stage, as they do not meet the precision requirements for a universal GBQC. Preskill has given this stage of development the somewhat euphemistic name NISQ, the Noisy Intermediate-Scale Quantum era[7], in the hope that it is already possible to perform useful operations on such systems, *e.g.* for simulations in physical chemistry. Furthermore, there are systems that are in a deliberately chosen preliminary stage to universal GBQC, such as the so-called quantum simulator (QS) of Pasqal[8], which is on its path to become a fully digital system in a few years.

In this sense, the JSC bases its strategy for quantum computing on three pillars: The first pillar is based on the classification of quantum computers as "analog" or "digital"

systems and their intermediate stages explained here, aiming at the provision of all these variants. The goal is to use systems with a high enough technological maturity as pilot production systems. For this purpose, we have introduced a classification scheme that follows the NASA technology classification[9], but in terms of Quantum Technology Readiness Levels (QTRL)[10]. Pilot production systems should show a QTRL of eight and above (QTRL8 to QTRL9). We also will implement or grant access to systems with a technological maturity of five to seven (QTRL5 to QTRL7) – this includes NISQ systems – that are considered as important test machines, and systems in an early development stage with QTRL of 3 to 4 by remote access to the respective lab for performing hardware-software co-design.

The second pillar of JSC's strategy is the tightest possible integration of breakthrough computing technologies into JSC's world-leading high performance computing (HPC) systems. For more than ten years, JSC has pursued the technical interconnection and deep integration of mutually autonomous supercomputer systems. Corresponding hardware and software have been developed in the DEEP projects funded by the European Commission and European High-Performance Computing Joint Undertaking (EuroHPC JU) since 2012[11]. The resulting concept of hardware modularity called Modular Supercomputing Architecture (MSA) allows functional parallelism at the system level[12]. The MSA concept, realised in software by ParTec's ParaStation Modulo middleware suite[13], is the perfect platform to integrate quantum computing into an HPC environment. By coordinating HPCQS[14], JSC is researching with European science, engineering and industry partners to extend the MSA integration technology to include quantum computing.

The third pillar of JSC's quantum computing strategy is the creation of a world-leading quantum computer user infrastructure. This includes the technical provision of the quantum computing systems (integrated into the HPC environment) – *nota bene* under European legislation –, the deployment of the systems via peer-reviewed calls for proposals – a rolling call for JUPSI is already running – the support of the users in simulation labs, algorithm development groups and cooperative research. These activities are carried out in the *Jülich UNified Infrastructure for Quantum computing* (JUNIQ)[15]. JUNIQ is the first infrastructure of its kind worldwide. And there is already a next level of infrastructure building under way: based on HPCQS[14], the JSC brings the principle of JUNIQ to the European level.

At JSC, we believe that our strategy will enable the earliest possible integration of breakthrough quantum computational technologies into an HPC environment. This is necessary insofar as the currently most promising algorithms for solving optimisation problems (*e.g.* the Quantum Approximate Optimisation Algorithm – QAOA) and finding solutions to eigenvalue problems in chemistry (*e.g.* the Variational Quantum Eigensolver – VQE) require the tightest coupling of classical with quantum algorithms.

This text is structured as follows: After a description of the existing and planned systems in Sec. 2 that JUNIQ aims integrating into its portfolio following our QTRL definition, a short report on an exemplary application on JUPSI as well as on JUQCS, the Jülich Universal Quantum Simulator, is given in Sec. 4. Then, the long-term perspectives of JUNIQ are outlined, followed by a summary and outlook in Sec. 6.

2 Existing and Planned Systems of JUNIQ

In planning its quantum computers, JSC focuses on both the applicability of the technology in practical computer simulations and the availability of systems at various maturity levels to bring the technology to users in a progressing manner as early as possible. Such planning requires the definition of a classification scheme of the technological maturity of the systems as a basis.

2.1 Quantum Technology Readiness Levels – QTRL

The QTRL scale is a metric introduced by the authors in 2017 to describe the maturity of quantum computing technology, see Fig. 2. The scale consists of nine technology readiness levels (QTRLs), with QTRL1 denoting the lowest and QTRL9 the highest. For simplicity, we evaluate both variants, analog (adiabatic/annealer) and digital (gate-based quantum computers), in one scheme, but also emphasise that the lack of error correction still pushes gate-based systems to a lower QTRL today.

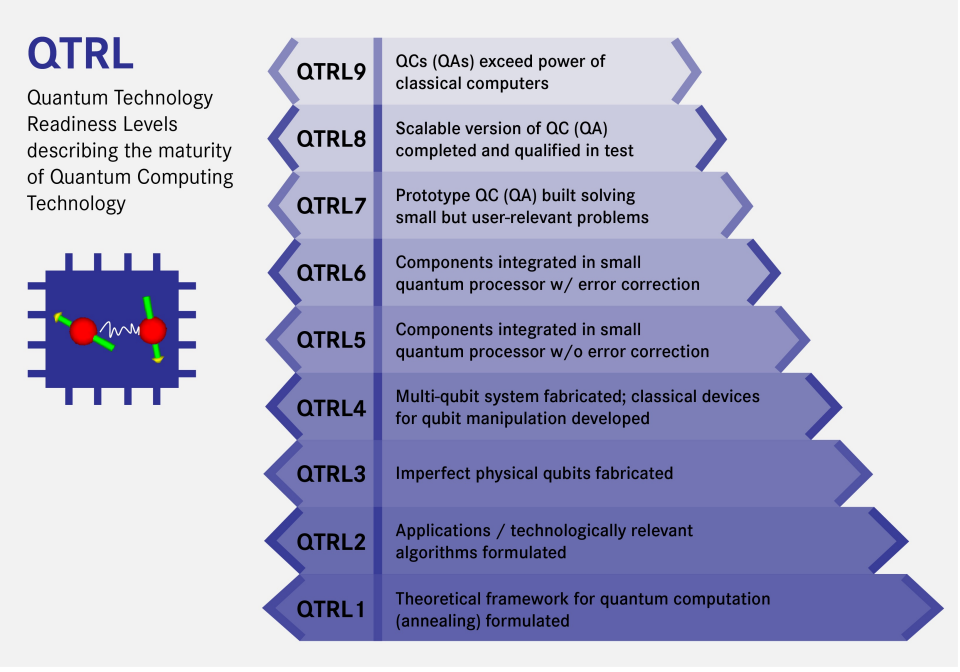


Figure 2. Definition of Quantum Technology Readiness Levels – QTRL.

1. A quantum computing technology is at QTRL of 1 when the theoretical framework for quantum computing has been formulated. Such theoretical studies of the basic properties of the quantum computing devices are moved towards applied research and development.

2. The quantum technology reaches QTRL of 2 once the basic device principles have been studied and applications or computationally relevant algorithms are formulated. QTRL2 quantum computing technology is speculative, as there are little to no experimental results yet supporting the theoretical studies.

3. Fabricated imperfect physical qubits, the basic building blocks of quantum computing devices, are defined to be at QTRL of 3. At this stage, laboratory studies aim to validate theoretical predictions of qubit properties. Theoretical and laboratory studies are required to determine whether these basic elements of the quantum computing technology are ready to proceed further through the development process.

4. At QTRL4, multi-qubit systems are fabricated and classical devices for qubit manipulation are developed. Both technical components of the quantum computer are tested together.

5. QTRL5 quantum computing technology comprises components integrated in a small quantum processor without error correction. At QTRL5, quantum computing devices are undergoing rigorous testing including running of various algorithms for benchmarking, still without error correction.

6. Components integrated in a small quantum processor with error correction are at a QTRL of 6. Rigorous testing and running algorithms is still the most important task for the QTRL6 quantum computing technology.

7. QTRL7 quantum computing technology is a prototype quantum computer (or annealer) solving small problems already relevant for applications. The prototype is demonstrated in a user environment.

8. A scalable version of a quantum computer (annealer) completed and qualified through test and demonstration is considered to be at QTRL8.

9. Once quantum computers or annealers exceed the computational power of classical computers for general (specific) problems the quantum computing technology is labelled with QTRL9 in our QTRL scheme.

We emphasise again that the scheme presented here includes QAs, QSs as well as GBQCs. This is for simplicity of classification, but it is not intended to compare the competitive aspects and prospects of the different approaches. It is obvious that quantum annealing, as the less complicated but presumably also less widely applicable technology, has already reached a higher level of maturity today. This is also a consequence of the fact that a GBQC requires high qubit fidelity to allow for efficient error correction. The lower the fidelity, the larger is the number of qubits required. At fidelity of $> 99.9\%$ the estimates range between 1,000 and 10,000 qubits per logical qubit. This reduces to a more reasonable estimate of 100 to 1,000 qubits per logical qubit for a fidelity of $> 99.99\%$, which could enable the realisation of a few logical qubits in the next years. A large number of logical error-corrected qubits required for true quantum supremacy seems not yet within reach, however.

2.2 The Jülich Universal Quantum Computing Simulator JUQCS – World's Most Capable Quantum Emulator

Most quantum computing systems can be described by a universal spin model Hamiltonian[16]:

$$H_{\text{universal}}(t) = -\sum_{j=1}^{L}\sum_{k=1}^{L}\sum_{\alpha=x,y,z} J_{j,k,\alpha}(t)\sigma_j^\alpha\sigma_k^\alpha - \sum_{j=1}^{L}\sum_{\alpha=x,y,z} h_{j,\alpha}(t)\sigma_j^\alpha, \qquad (1)$$

where σ^α, $\alpha = x, y, z$ are the Pauli matrices, $J_{j,k,\alpha}(t)$ is the strength of the interactions between the α-components of spins j and k, $h_{j,\alpha}(t)$ stands for a potentially time-dependent (external) field. The spin model can be adapted to most variants of today's and planned quantum computers like GBQCs, QSs and QAs for all varieties of technological realisation. The quantum dynamics of the model is controlled and/or probed by the static and/or time-dependent external fields $h_{j,\alpha}(t)$. Depending on the physical system, also the exchange parameters $J_{j,k,\alpha}(t)$ can be controlled.

The quantum computer emulator suite JUQCS[17,18], developed within a collaboration between JSC and the University of Groningen, simulates the execution of a quantum circuit on an ideal GBQC – a pen-&-paper quantum computer, including the quantum depolarising channel error model. The quantum computer emulator comes in two flavours: (i) A hybrid MPI/ OpenMP/GPU version simulating the ideal GBQC to machine-precision (JUQCS-E and JUQCS-G), and (ii) a hybrid MPI/OpenMP version that performs the same task with reduced precision but enabling the emulation of three more qubits relative to version (i) (JUQCS-A).

JUQCS runs on various kinds of hardware ranging from PCs to high-end supercomputers with distributed and/or shared memory and with CPUs and/or GPUs, easily ports to different software environments, and can be used to benchmark supercomputers[19]. In the course of its development, JUQCS has set a series of world records of simulating the largest ideal GBQC[20]. Currently, this is a 48-qubit GBQC simulated on the Japanese K computer and on the Chinese Sunway TaihuLight.

JUQCS allows us to run quantum computing algorithms on an ideal (!) GBQC with up to 43 qubits by means of JUWELS. In this way, we are able to validate quantum computing codes[17–19,21–23] and test the intrinsic scalability of algorithms not hindered by the shortcomings of current real GBQCs.

An interesting usage model for JUQCS is to leverage the modular capabilities of JUWELS, the leading HPC system at JSC. JUWELS consists of a CPU cluster, the JUWELS Cluster, and a GPU cluster, the JUWELS Booster. When the classical optimisation part of a hybrid classical-quantum computing code such as QAOA is running on JUWELS Cluster, JUWELS Booster can act as an emulator of an ideal GBQC realised by JUQCS. In this way JUQCS allows us to run real HPC-quantum computing codes on an ideal GBQC with up to 43 qubits on JUWELS.

In addition to the JUQCS suite, the JSC and University of Groningen team also developed software to simulate physical models respecting the real physical implementation of a quantum computer, including multi-level spin systems, as no real system strictly is a 2-level system[18,24]. For this purpose, the time-dependent Schrödinger equation for a time-dependent Hamiltonian describing the quantum computer hardware and its control has to be solved. This software is also suitable to simulate quantum-spin systems that are cou-

pled to a heath bath, enabling to study the effect of the environment on the operation of a quantum computer.

2.3 D-Wave Advantage System JUPSI – A Quantum Annealer

JUPSI, the Jülich Advantage™ quantum system of D-Wave has started operation within JUNIQ at JSC in January 2022[25]. A rolling call for quantum computer time proposals has started as early as in February 2022[15].
JUPSI, see Fig. 1, comprises the following components:

- The quantum processing unit (QPU)
- The QPU control system
- The QPU I/O system
- The cryogenic dilution refrigerator system
- The magnetic shielding system
- The radio-frequency (RF)-shielded enclosure

The D-Wave Advantage QPU is a lattice of tiny metal loops, each of which acts either as a qubit or as a coupler. Below their critical temperature, these loops become superconductors and exhibit quantum-mechanical effects. The system is operated at a temperature below 15 mK. The QPU of the Advantage system provides 5,640 qubits and 40,484 couplers. In order to reach this scale, the chip can make use of 1,030,000 Josephson junctions.

The qubits of the Advantage QPU are interconnected in a topology known as *Pegasus*. In the Pegasus graph each qubit is connected with 15 other qubits except for the qubits at the edges.

As mentioned, the QPU must be kept at a temperature near absolute zero and isolated from the surrounding environment in order to behave quantum mechanically. The system meets these requirements by

- operating the QPU at temperatures below 15 mK. These cryogenic temperatures are achieved using a closed-loop cryogenic dilution refrigerator system;
- shielding the QPU from electromagnetic interference, achieved using a radio frequency (RF)-shielded enclosure and a magnetic shielding subsystem.

The RF-shielded enclosure (an electromagnetic-tight box) houses the cryostat, the QPU, the I/O system, the rest of the QPU control system, and most of the magnetic shielding system.

The cryogenic dilution refrigerator system includes a cryostat, a nitrogen cold trap, and equipment racks. The cryostat, which is surrounded by a magnetic shielding system, is mounted within the RF-shielded enclosure; the cold trap and equipment racks sit outside of it. Communication between the equipment racks and the cryostat occurs through flexible hoses and cabling. The dilution refrigerator has an internal fluid path through which helium flows in a closed cycle. The cycle includes the cold-trap dewar, where impurities that might cause blockages in the circulation path are removed. The cold-trap dewar is maintained at a

cryogenic temperature by immersion in a liquid nitrogen dewar. The refrigeration system's nitrogen cold trap sits outside the RF-shielded enclosure and has flexible connections to the pumping control rack.

The server rack contains the data processing, system control, monitoring, network, and backup power resources for the system. The data processing resources provide the web user interface, web services, and the underlying communications infrastructure for users to interact with the D-Wave system. Control and monitoring servers run applications that control the system and provide remote monitoring capability. The monitoring server tracks diagnostic information from the various subsystems and alerts D-Wave support personnel if it detects abnormal events. The contents of the server rack connect flexibly to the refrigerator's instrumentation rack and to feed-through connections in the RF-shielded enclosure wall.

The gas handling control rack contains gas manifolds and pumps that circulate the helium and also contains pumps and specialised equipment used only during system service.

The pumping control rack houses the compressor for the cooling, additional pumps that circulate the helium and a control cabinet that houses the electrical and pneumatic control subsystems. This part consumes most of the required power of 25 kW.

The universal spin model Hamiltonian of Eq. 1 becomes much simpler when it is reduced to describe a D-Wave QA:

$$H_{\text{annealer}}(t) = -A(t)\sum_{j=1}^{L}\sigma_j^x - B(t)\left(\sum_{j=1}^{L}\sum_{k=1}^{L}J_{j,k}\sigma_j^z\sigma_k^z + \sum_{j=1}^{L}h_j\sigma_j^z\right). \tag{2}$$

Note that two new functions $A(t)$ and $B(t)$ have been introduced. They vary in opposite directions between zero and one during the annealing process. This allows to start the process governed by the first part of the Hamiltonian, which introduces the necessary quantum fluctuations in the ground state of the system, and further on to move the system adiabatically to the Hamiltonian of interest describing the problem to solve. The adiabatic theorem tells us that the system stays in the ground state, when the process is carried out slow enough. At the end of the process the system is in the ground state of the problem Hamiltonian, equivalent to the minimum energy of the spin model Hamiltonian [b]. In a sense one can compare the adiabatic process of the QA with the amplitude amplification in GBQC: at the end, the systems should be in a state where measurement is as unambiguous as possible.

2.4 Pasqal's Fresnel Analog Quantum Computer – A Quantum Simulator

The third system to be installed in JUNIQ – taking JUQCS on JUWELS as an ideal GBQC and validation basis as number one – is a QS with more than 100 Rydberg atoms. With preliminary working name JURY, the system will be manufactured by the French company Pasqal[8]. JURY, funded through the EuroHPC JU project HPCQS[14] in fact is a twin of a system located at CEA, Paris. JURY will be hosted by JUNIQ with installation in 2023, and will be tightly integrated with the JUWELS MSA at JSC. The QS technology grew

[b] This idealised process of course only works for ideal systems that are not affected by noise or other disturbing effects.

from 49 qubits in 2018 to more than 100 as of today and will presumably reach 1000 qubits in 2023.

The QPU of JURY is based on a configurable 2D or also 3D array of single neutral rubidium atoms. The array can be seen as a register, where each single atom plays the role of a qubit. Laser light is used to assemble and read out registers made of more than a hundred of qubits and to perform fully programmable quantum processing. In addition, electronic controls are applied to tune the light properties, apply instructions arising from the quantum software stack and extract information through atomic detection.

Figure 3. Pasqal quantum simulator. Picture credits: Pasqal.

A QS can, for example, be used to solve chemistry calculations for drug discovery, as well as for very hard optimisation problems like the maximum independent set (MIS) problem arising in a manifold of application fields. The MIS problem can be tackled by using an ensemble of interacting cold neutral atoms as the quantum resource, thus realising Feynman's dream of simulating a quantum system by another, more easy to handle quantum system.

2.5 JUNIQ's Plans for Gate-Based Quantum Computers – We are in the NISQ Era

In JSC's strategy, gate-based systems will complement the emulator JUQCS, D-Wave's QA JUPSI and Pasqal's forthcoming QS JURY such that all varieties, from an analog system (QA) to an analog-digital hybrid system (QS) to a fully digital system (GBQC [NISQ]), will be represented and provided in JUNIQ.

Not surprisingly, GBQCs still remain *in status nascendi* with QTRL ≤ 5. This is because such systems require individual control over qubits as well as their interactions, preferably with full connectivity. Furthermore, qubits and gates must fulfil quality criteria such as a high fidelity of qubits and gates as well as good crosstalk and readout properties in order to be able to perform reliable quantum computations. It is generally assumed that a fidelity of 99.99% is required for efficient error correction. As a consequence, all technologies used so far are limited in terms of scalability, which is also due to their analog control technology.

JUNIQ intends to offer three generations of GBQCs. This approach includes more than one prototype production system on-site based on different building principles with QTRL5 – we plan to connect such a system to the exascale supercomputer – and soon hopefully higher (as well as remote access to more such systems), remote access to experimental systems in the lab with QTRL4, and remote access to devices at QTRL3 for testing single qubits and their connections.

In Tab. 1, we present an overview over the development of quantum computer technologies including the Pasqal QS and the D-Wave QA. The most advanced technologies are cold neutral atoms (Pasqal), ion traps (eleQtron, AQT, IonQ), and superconducting qubits (IQM, as well as US companies like IBM, Rigetti Computing, Google). The characteristics of the different approaches are taken into account when selecting potential systems for JUNIQ and assessing their QTRL.

	Cold neutral atoms	**Superconducting qubits**	**Ion traps**	**Silicon photonics**	**Microwave photonics**	**Quantum dots in silicon**	**Nitrogen vacancy centres in diamond**	**Topological qubits**	**Quantum annealing**
Qubit	Trapped cold atoms	Josephson effect	Trapped ions	Single photons	Photons, superconductivity	Electron spins in semiconductors	Atomic nucleus spins	Anion pair, quasiparticles	Josephson effect
# qubits demonstrated	196 Pasqal	127 IBM	32 IonQ	A few	1	49 Intel	6 QDTI	0	5640 D-Wave
States	Atomic energy levels	Phase, energy, current direction	Ion energy levels	Polarisation	Polarisation	Electron spin direction	Energy level of cavity	Anyon direction	Current direction
Gates	Laser	Microwaves, Josephson	Laser	Polarising and dichroic filters	Microwaves	Microwaves	Laser	2D anyon inversion	Gates announced
PROS	Scalability, 3D connectivity, performance	Availability	Fidelity, Availability	Room temperature	Performance	Scalability	Theo. Performance	Theo. Performance	In production
CONS	No digital unit so far	Weak scalability	Weak scalability	Performance t.b.d.	No 2-qubits yet	Noise	Theoretical	Not existing yet	Not universal
QTRL	5	5	5	3	2	3	2	1	8

Table 1. Overview of quantum computing technologies.

2.5.1 eleQtron – A German Ion Trap NISQ GBQC Development

eleQtron, a spin-off of the University of Siegen, is planning to build an ion trap GBQC based on up to 320 $^{171}Yb^{+}$ ions. This is microwave-controlled ion trap technology. In this respect, it differs from systems with laser-based control, for example, and promises easier scalability. A prototype of an 8-qubit NISQ GBQC of this type already exists[26,27]. Features of the eleQtron technology are:

- The qubit control of eleQtron works with microwaves, therefore lasers are not needed for qubit manipulation.
- Operation is possible at room temperature.
- A vibration-damped base is not needed.
- eleQtron has a prototype of 8 qubits today; the company expects 20 to 30 qubits in less than 2 years and 50 to 60 in less than 3 years.
- The qubits are represented by Ytterbium ions.
- The qubit connectivity is all-to-all.
- The one-qubit gate fidelity is 99.9%, the two-qubit gate fidelity is 99.8%.
- The cross talk properties of eleQtron qubits are $< 10^{-4}$ in the Rabi regime and $< 10^{-8}$ in intensity.

With regard to alternative technologies, elecQtron offers a GBQC with proven qubit gates in the best quality currently available.

2.5.2 AQT – An Ion Trap NISQ GBQC from Austria

AQT is a spin-off company from the well known quantum computing group of Innsbruck university. AQT's quantum computer hardware is based on ion trap qubits[28]. AQT differs from eleQtron in that it uses laser-controlled qubits instead of microwaves. Recently, AQT unveiled a quantum computer demonstrator in two 19-inch racks based on $^{40}Ca^{+}$ qubits in a linear Paul trap, evolving from a laboratory-based proof-of-concept experiment to robust, integrated hardware for quantum information processing. It has been shown that the demonstrator can generate maximally entangled Greenberger-Horne-Zeilinger (GHZ) states with up to 24 qubits[29] and that as many as 50 ions can be trapped and cooled. Features of AQT's technology are:

- The qubit control of AQT works with lasers, which allow for faster operation than microwaves.
- Operation is possible at room temperature.
- A vibration-damped base is not needed.
- AQT has a prototype of 24 qubits today; the company expects a mid-term upgrade with 50 qubits and hopes that a quantum computer with more that 100 qubits is feasible based on the given architecture.

- The qubits are represented by calcium ions.
- The qubit connectivity is all-to-all.
- The one-qubit gate fidelity is 99,86%, the two-qubit gate fidelity is 97–99%.
- The resonant crosstalk properties of AQT qubits are $< 1\%$ and the non-resonant crosstalk $< 1.25 \times 10^{-4}$.

Among the ion-trap systems, AQT offers a gate-based system with fastest operation.

2.5.3 IQM – Superconducting NISQ GBQC from Finland

At the end of 2021, IQM installed its first quantum computer with five superconducting qubits. In 2023, IQM plans to deliver a 20-qubit GBQC to supercomputer centres to enable experiments with NISQ algorithms. In 2025, IQM aims to deliver a first commercial 1,000-qubit GBQC called "Prometheus", which IQM expects to achieve superior performance to solve real-world industrial problems. IQM is coordinating the two-year Horizon 2020 project of the very same name, Prometheus, which includes the development of this commercial GBQC. The processors developed in this project could then meet the scale required to run commercially viable quantum algorithms. The processor size also seems sufficient for the implementation of quantum accelerators, *i.e.*, quantum processors installed in large supercomputer centres.

Recently, IQM described its unimon qubit, a superconducting-qubit type, which combines the desired properties of high non-linearity, full insensitivity to DC charge noise, insensitivity to flux noise, and a simple structure consisting only of a single Josephson junction in a resonator[30]. According to IQM the unimon qubit might have the potential to break the 99.99% fidelity limit required for efficient quantum error correction and might achieve possible quantum advantage with noisy systems.

- The qubit control of IQM works with voltage pulses.
- Operation needs cryogenic temperatures.
- Probably, a vibration-damped base is required.
- IQM has announced to operate a prototype with 5 qubits today; the company expects 20 qubits in 2023 and 1000 qubits within 3 years.
- The qubits are represented by superconducting Josephson junctions, named unimon.
- The qubit connectivity is not known.
- The expectation of IQM is that the one-qubit gate fidelity reaches 99.99%.
- The cross talk properties of IQM qubits are not available.

With the target of a qubit fidelity of 99.99% IQM could become a promising candidate for first error corrected quantum computing. Given the required step in scale of more than a factor of 200, it remains to be seen if this goal can be achieved in a time frame of less than three years.

2.5.4 OpenSuperQ, DAQC and Qsolid Quantum Computers

JUNIQ's tasks also include access to quantum computer systems in early stages of development. These systems are mostly laboratory systems or prototype systems. Within this setting, JUNIQ will provide cloud access to experimental quantum computers developed in European and national research projects. The first experimental GBQC that can be accessed by the end of 2022 is a superconducting GBQC with 10 qubits from the quantum flagship project OpenSuperQ[31] launched in 2018.

The digital-analogue quantum computer (DAQC) of IQM that is being built in the eponymous BMBF (Federal Ministry of Education and Research) project DAQC combines the advantages of an analog quantum computer that is not very prone to errors with the flexibility of digital circuits. The DAQC is to be paired with supercomputers and will take over the task of a computing accelerator. Access to the DAQC with 54 superconducting qubits is planned for 2024.

The BMBF project QSolid[32] will develop an integrated, fully HPC-embedded quantum computing demonstration system based on cryogenic superconducting quantum processors integrated in a fully developed hardware system with control, readout and infrastructure down to the specific, optimised firmware and software. The demonstrator will house several processors. In the first phase of the project, QSolid will research high-quality components and demonstrate the technology of the medium-term device, a 10-qubit system in ladder geometry. JUNIQ cloud access to this system is planned for 2024. Based on the lessons learned from this system, three processors based on the same components and technologies will be developed. The goal is to build a GBQC with at least 30 superconducting qubits.

3 Integration Strategy

A central concern of JUNIQ is to integrate the locally installed quantum computers as closely as possible into the HPC environment of the JSC. This is done with respect to running hybrid programs, where parts are carried out on large HPC systems and other parts (will) require the functionality of quantum computers or QAs and QSs. Although this type of use, which is functional parallelism at the highest level of integration, is still in its infancy today, it is nevertheless to be expected that almost all practical applications of quantum computing will be based on this approach. Therefore, developing hardware and software systems as quickly as possible for tightly coupled low-latency use is a priority.

JUNIQ employs the concept of the MSA for the integration. It allows to enable the lowest latency integration of the QA, the imminent QS and the coming GBQCs [c]. Modular operation requires joint scheduling that will ensure an efficient exploitation of available resources.

The quantum computers are used as modules in the MSA, closely coupled with other specific modules such as the general purpose CPU and GPU acceleration systems or a tiered common high speed storage. The low-latency connection to other modules is

[c] Tight integration through the MSA will overcome the latency limitations of the cloud service-based approaches favoured so far by major manufacturers[33]. The danger of the strict cloud approach, however, lies not only in the latency limitation, but rather in the tendency of hyper-scalers to move towards the monopolisation of quantum computing and thus of hybrid HPC-QC services, to the detriment of science and subsequently industry.

achieved via an ultra-high-speed network like the one of Mellanox, but also alternatives exist.

The MSA integration will be enabled by extending ParaStation Modulo™ with an execution environment for hybrid HPC-quantum computing simulations and workflows on modular supercomputers. A portal solution provides convenient and flexible access via a web browser enabling a production-level JupyterHub Portal[34] (HPC-PaaS and QC-PaaS).

JUNIQ and HPCQS work on the integration of the QPU's system software and its front-end software system into the full management stack of the modular supercomputer including user and software management, storage access, and provisioning.

JUNIQ provides a full hybrid software stack including compilers, technical libraries and tools, that will allow users to translate use cases into quantum programs without having to deal with low-level instructions.

4 Exemplary Application: Simplified Tail Assignment Problem

To give a representative impression of the work being carried out at JSC as part of JUNIQ, we consider computational results for the simplified tail assignment problem, a potential application for both QAs and GBQCs investigated as part of the European Quantum Flagship's OpenSuperQ project[35,36] and of JUNIQ[18]. These considerations demonstrate the different degrees of maturity of QAs and GBQCs as well as the importance of the intrinsic scalability of quantum algorithms per se.

Airlines are faced with one of the largest and most challenging planning problems required to be solved regularly. For example, an airline might need to plan and operate over 1000 flights per day to over 150 cities in over 70 countries, using hundreds of aircraft of different types. The most significant expenses for an airline often consist of costs associated with aircraft and flight crew. Obtaining optimised plans for crew and aircraft while respecting regulations imposed by aviation authorities, airlines and unions is thus crucial and non-trivial. The problem called tail assignment (TAP) describes a mathematical optimisation model with constraints that when solved can provide airlines with efficient plans for how to use their aircraft. In the TAP, the task is to assign a set of airplane flights between airports to a set of routes, which are sets of flights operated in sequence by the same airplanes. The name "tail assignment" comes from the fact that each aircraft is identified by the registration number on its tail fin, its tail number. Quantum computers may provide new methods to help solving the TAP in the future.

In order to simplify the optimisation problem for studying it with real quantum computing devices Vikstål *et al.* and Willsch *et al.* considered the simplified TAP for which each flight has to be covered exactly once (no overlapping routes) and for which all flight costs are equal to zero[18,35,36]. This simplified version of the TAP is an exact cover problem, which in matrix form reads

$$\min_{x_r=0,1} \sum_{f=0}^{F-1} \left(\sum_{r=0}^{R-1} A_{rf} x_r - 1 \right)^2 , \tag{3}$$

where $r = 0, \ldots, R-1$ enumerates all routes, $f = 0, \ldots, F-1$ enumerates the flights, $x_r \in \{0, 1\}$ are the Boolean problem variables with $x_r = 1$ if route r is to be selected, $A_{rf} \in \{0, 1\}$ are the elements of the constraint matrix A with $A_{rf} = 1$ if flight f is

contained in route r. The QUBO and Ising formulations of the exact cover problem are given in Ref. 36.

On the D-Wave QA, Willsch *et al.* studied exact cover problems corresponding to a series of realistic airline problem instances obtained by random sampling from a real-world data set with up to $N = 40$ routes, each of which contains several out of 472 flights[36]. The exact cover problem is to find a selection of the N routes to carry out the 472 flights between the airports so that the routes do not overlap. On a quantum computer, each route is represented by a qubit. If a route is to be selected, the corresponding qubit ends up in the state $|1\rangle$ when measured. By construction, the ground state of each problem instance, an N-qubit problem, is unique and contains 9 qubits in state $|1\rangle$. This means that the solution consists of nine routes. Each route is assigned to an aircraft. All other states represent invalid solutions, in the sense that not all 472 flights are covered exactly once. For 40 routes the number of possible selections is $2^{40} \approx 10^{12}$. For 120 routes, corresponding to a 120 qubit-problem, the number of selections already grows to 10^{36}.

The particular problem instance for $N = 40$, illustrated in Fig. 4, translates into a quantum optimisation problem with 40 qubits that is almost fully connected (711 of 780 couplers $J_{i,j}$ are non-zero). Hence, solving this problem requires couplers that do not physically exist on the QPU of the D-Wave annealers DW2000Q and JUPSI, the Advantage 5.1 system. The 2000+ (5000+) physical qubits of the DW2000Q QPU (JUPSI) are connected in a Chimera (Pegasus) topology to 6 (15) other physical qubits on average. Hence, in order to solve the problem it has to be embedded on the D-Wave QPU. With the embedding, the effective connectivity between the qubits is increased by combining several physical qubits into a logical qubit. The physical qubits that form a logical qubit are called a chain. In order to ensure that the physical qubits within a chain operate as a single logical qubit, the couplers $J_{i,j}$ between the physical qubits are set to a reasonably large, negative value called the chain strength. The relative chain strength is defined as the ratio of the chain strength and the maximum strength, which is the maximum absolute value of all h_i and $J_{i,j}$, and thus takes values between zero and one.

Embeddings, mappings from each logical qubit to a chain of physical qubits, are not unique and may considerably affect the quality of the solution. Therefore, in benchmarking studies it is important to investigate different embeddings and relative chain strengths.

Fig. 5 shows results for exact cover problems that correspond to the simplified TAP with 472 flights described above, with 30, 36 and 40 routes (*i.e.*, logical qubits), respectively[36]. Because these problems have a unique ground state the success rate can be calculated. It is obtained by counting the number of samples with energy zero. For each problem 10 different embeddings were generated and the success rate was evaluated as a function of the relative chain strength. It can be seen that the results obtained on the D-Wave Advantage system are generally much better than those obtained on the D-Wave 2000Q, especially for the larger problems. The reason for this is that the problems are almost fully connected, which makes them more suitable to be embedded on the Pegasus topology of the QPU of the Advantage system.

In order to assess the performance of both QAs for larger but more sparsely connected problem instances Willsch *et al.* studied exact cover problems with 50 to 120 logical qubits[36]. For each problem size, they considered six problem instances. The exact cover problems correspond to a simplified TAP as described above, but with 535 flights. Also for these problems the ground state is unique and known. It has 40 qubits in the state $|1\rangle$.

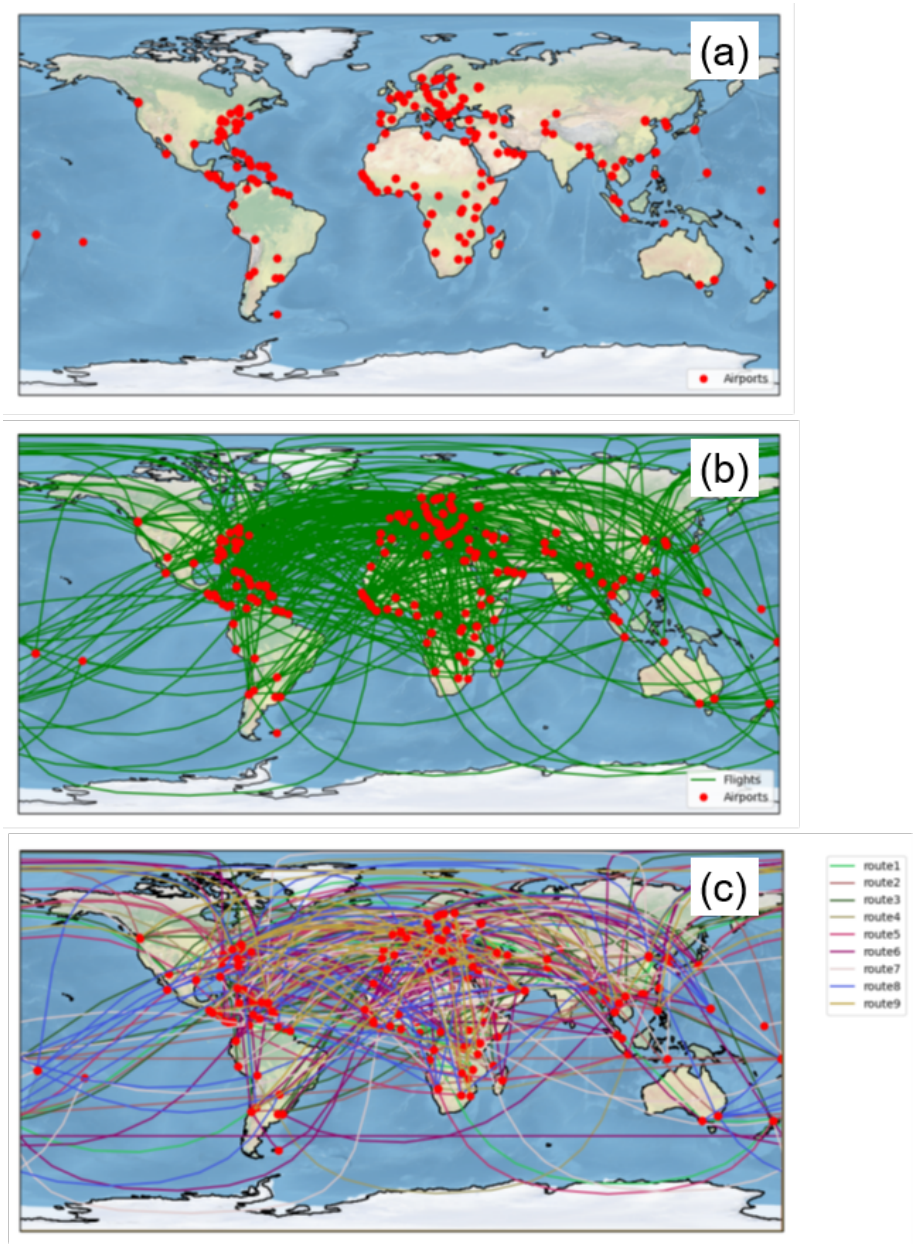


Figure 4. Problem instance of a TAP. (a) airports that are involved; (b) the 472 flights between the airports that have to be performed; (c) unique solution with 9 routes covering the 472 flights exactly once.

The success rates as a function of the relative chain strength are shown in Fig. 6. It is seen that in general for these larger problems the success rates are smaller and that the

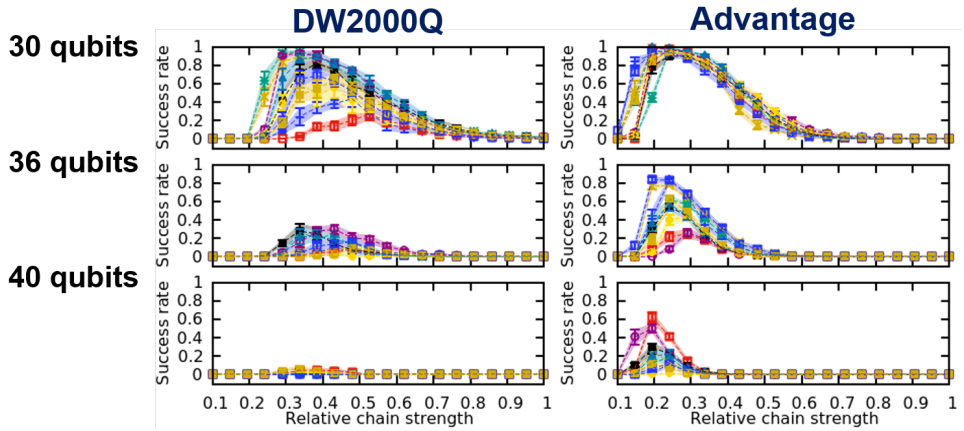


Figure 5. Success rates as a function of the relative chain strength for exact cover problems with 30, 36 and 40 logical qubits that have 90% of nonzero couplers, solved on a D-Wave 2000Q (DW_2000Q_VFYC_6) and an Advantage system (Advantage_system1.1) with the default annealing time 20μs. The scan of the relative chain strength is repeated for 10 different, randomly generated embeddings (represented by different colours) and with 10 repetitions each to gather statistics. Markers indicate the corresponding standard deviation above and below the mean. Filled areas between the markers are guides to the eye.

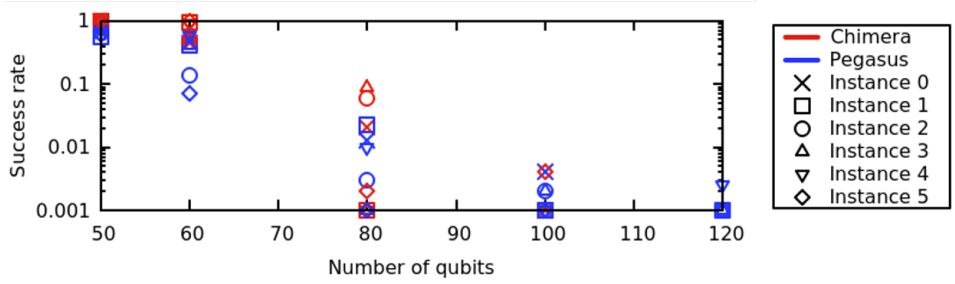


Figure 6. Success rates a function of the relative chain strength for exact cover problems with 50-120 logical qubits that have 20% of nonzero couplers, solved on a D-Wave 2000Q (DW _2000Q_VFYC_6) and an Advantage system (Advantage_system1.1). Different markers indicate different problem instances. The annealing times and the number of reads have been varied from 20μs to 2000μs and from 400 to 1000, respectively, in order to find a solution within the maximum run time of the systems (for more details, see Ref. 36).

D-Wave 2000Q is not able to solve the largest problems anymore.

For cross-platform benchmarking purposes, Willsch *et al.* also solved the simplified TAP with 40 qubits by means of QAOA and the approximate quantum annealing algorithm (AQA) for GBQCs. For this purpose they used JUQCS on JUWELS Booster. In this case JUWELS plays the role of an ideal QPU of a GBQC with all-to-all connectivity. Results of this study can be found in Ref. 18 and in this volume. The overall result of the cross-platform benchmarking study is that for this type of problems the D-Wave QA outperforms the ideal GBQC running QAOA in terms of the success rate for finding a solution.

5 Future of JUNIQ – A Visionary Roadmap

The currently observed and enjoyed acceptance of both its simulation service JUQCS and the D-Wave QA speaks strongly for JUNIQ's chosen strategy of bringing users as close as possible to the resources and as early as possible. While it is clear that active use of the systems in HPC production tasks cannot yet take place, we expect that the modular connection of the systems will happen in the next future, especially through the positive pressure of the HPCQS project and other activities at JSC, which are currently still under NDA and cannot yet be described here.

In mid-2022, JUNIQ plans to offer cloud access to the OpenSuperQ system at FZJ/PGI, and a Pasqal quantum simulator will be installed in the JUNIQ building in mid-2023. Cloud access to the digital-analog IQM DAQC system will be available in mid-2024.

As Fig. 7 demonstrates, it is intended to go into an upgrade of JUNIQ by the end of 2024 (a project called JUNIQ-Upgrade). Besides an upgrade of the D-Wave QA, the installation of a strong GBQC is planned, which will be connected directly with the coming exascale supercomputer JUPITER for execution of hybrid HPC-QC simulations

As the core of the Jülich HPCQS coordination activities, JUNIQ will thus be able to set the tone for the chosen European path towards an integrated and federated EuroQCS (European Quantum Computing and Simulation) infrastructure[33]. This will be done in parallel and in coordination with the expansion of the federation of European HPC systems. It is expected that EuroHPC JU will want to establish quantum computing systems at all pre-exascale and exascale systems, and possibly also at midrange sites.

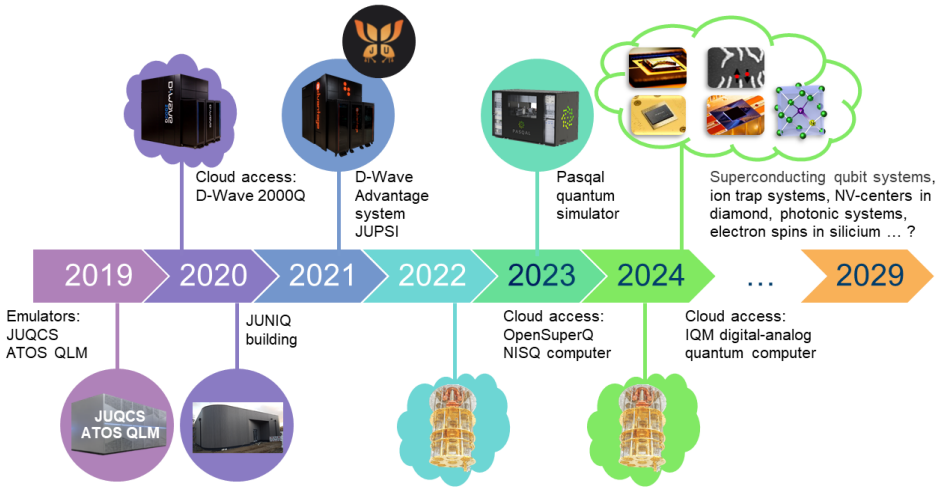


Figure 7. Roadmap JUNIQ.

6 Summary and Outlook

JUNIQ is a manufacturer-independent, comprehensive quantum computing user facility integrated in the Jülich Supercomputing Centre, with peer-reviewed access for European users from academia and industry. As such JUNIQ is a first step towards EuroQCS.

JUNIQ will offer four categories of computing systems that emulate or directly exploit quantum effects: the existing quantum computing emulator JUQCS can emulate ideal "pen-&-paper" quantum computers with up to 43 qubits on JUWELS, and the D-Wave quantum annealer JUPSI, which is in operation since begin of 2022, allows to treat optimisation problems, classification problems, and quantum simulations on more than 5000 qubits.

We have presented the tail assignment problem as an exemplary application to understand the ideal GBQC by means of the emulator JUQCS and to compare to the annealer.

The Pasqal quantum simulator JURY, to be implemented in the JUNIQ building in mid-2023, will enable the simulation of quantum spin systems and the Heisenberg model, as well as the QAOA and VQE algorithms. We have also presented three variants of gate-based quantum systems that are candidates in JUNIQ and have achieved a QTRL of 5. In addition, access will be provided to early stage experimental systems.

We have presented JUNIQ's unique integration strategy, where quantum computers become modules of the JSC HPC environment to realise functional parallelism at the highest parallel system level.

JUNIQ is the first infrastructure of its kind in the world to offer computing time on multiple systems on the basis of a call for user project submissions, in the same way as for HPC, with the difference that it is a rolling call for which submissions can be made at any time. Administrative support for the allocation is already provided by the Office for Computing Time Allocation at JSC, which carries out these tasks for the John-von-Neumann Institute and the Gauss Centre for Supercomputing (GCS).

We can expect the demand for hybrid QC-HPC computing to grow strongly in the next few years, as it is this type of application that allows the first productive use of quantum computers. With the realisation of the modular integration of quantum computers in JUNIQ into the HPC environment of the JSC, the allocation of QPU computing time is to be joined with the allocation processes of the NIC and GCS, in the same way as GPU resources have been put forward some years ago. There is a well-founded expectation that quantum computers can already play an important role in connection with the upcoming exascale supercomputer JUPITER.

Acknowledgements

We acknowledge support from the project Jülich UNified Infrastructure for Quantum computing (JUNIQ) that has received funding from the German Federal Ministry of Education and Research (BMBF) and the Ministry of Culture and Science of the State of North Rhine-Westphalia. We thank Hans De Raedt for much useful advice. K. M. and T. L. thank the members of their research groups for technical discussions.

References

1. R. P. Feynman, *There's Plenty of Room at the Bottom*, Engineering and Science **23**, 22–36, 1960.
2. R. P. Feynman, *Simulating Physics with Computers*, Int. J. Theor. Phys. **21**, 467–488, 1982, doi:10.1007/BF02650179.
3. https://www.dwavesys.com/resources/white-paper/the-d-wave-advantage-system-an-overview/
4. https://www.fz-juelich.de/en/ias/jsc/systems/quantum-computing/juniq-facility/juniq/d-wave-advantagetm-system-jupsi
5. M. Born, V. Fock, *Beweis des Adiabatensatzes*, Z. Physik **51**, 165–180, 1928, doi:10.1007/BF01343193.
6. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
7. J. Preskill, *Quantum Computing in the NISQ era and beyond*, Quantum **2**, 79, 2018, doi:10.22331/q-2018-08-06-79.
8. https://pasqal.io/
9. https://www.nasa.gov/directorates/heo/scan/engineering/technology/technology_readiness_level
10. https://www.fz-juelich.de/en/ias/jsc/about-us/structure/research-groups/qip/technology-readiness-level-of-quantum-computing-technology-qtrl
11. https://www.deep-projects.eu/
12. G. Hager and G. Wellein, *Introduction to High Performance Computing for Scientists and Engineers*, Chapman and Hall / CRC computational science series, CRC Press, 2011.
13. A. Kreuzer, E. Suarez, N. Eicker, and Th. Lippert (ed.) *Porting applications to a Modular Supercomputer – Experiences from the DEEP-EST project*, Forschungszentrum Jülich GmbH, Zentralbibliothek, Verlag, 2021.
14. https://www.hpcqs.eu/
15. https://www.fz-juelich.de/en/ias/jsc/systems/quantum-computing/juniq-facility
16. H. De Raedt et al., *Quantum Computer Emulator*, Comput. Phys. Commun. **132**, 1–20, 2000.
17. K. De Raedt et al., *Massively Parallel Quantum Computer Simulator* Comput. Phys. Commun. **176**, 121–136, 2007, doi:10.1016/j.cpc.2006.08.007.
    H. De Raedt et al., *Massively parallel quantum computer simulator, eleven years later*, Comput. Phys. Commun. **237**, 47–61, 2019, doi:10.1016/j.cpc.2018.11.005.
18. D. Willsch et al., *GPU-accelerated simulations of quantum annealing and the quantum approximate optimization algorithm*, Comput. Phys. Commun. **278**, 108411, 2022, doi:10.1016/j.cpc.2022.108411.
19. D. Willsch et al., *Benchmarking supercomputers with the Jülich Universal Quantum Computer Simulator*, NIC Symposium 2020, Publication Series of the John von Neumann Institute for Computing (NIC), NIC Series 50, 255–264, 2020.
20. https://www.hpcwire.com/2010/06/28/quantum_computer

_simulation_new_world_record_on_jugene/
https://www.hpcwire.com/2017/12/18/world-record-quantum-computer-46-qubits-simulated/
https://www.aachener-zeitung.de/lokales/juelich/juelicher-forscher-team-haelt-weltrekord-im-simulieren-eines-quantencomputers_aid-24383883
21. M. Willsch et al., *Benchmarking the quantum approximate optimization algorithm*, Quant. Inf. Proc. **19**, 197, 2020, doi:10.1007/s11128-020-02692-8.
22. F. Arute et al., *Quantum supremacy using a programmable superconducting processor*, Nature **574**, 505–510, 2019, doi:10.1038/s41586-019-1666-5.
23. M. S. Jattana et al., *Assessment of the variational quantum eigensolver: application to the Heisenberg model*, Front. Phys., doi: 10.3389/fphy.2022.907160.
24. H. De Raedt and K. Michielsen, *Computational Methods for Simulating Quantum Computers*, Handbook of Theoretical and Computational Nanotechnology, Vol. 3: Quantum and molecular computing, quantum simulations, Chapter 1, M. Rieth and W. Schommers eds., American Scientific Publisher, Los Angeles, 248, 2006.
D. Willsch et al., *Gate-error analysis in simulations of quantum computers with transmon qubits*, Phys. Rev. A **96**, 062302, 2017.
M. Willsch et al., *Real-time simulation of flux qubits used for quantum annealing*, Phys. Rev. A **101**, 012327, 2020.
H. Lagemann et al., *Numerical analysis of effective models for flux-tunable transmon systems*, arXiv:2201.02402.
25. https://www.hpcwire.com/off-the-wire/europes-first-quantum-computer-with-more-than-5k-qubits-launched-at-julich/
https://www.youtube.com/watch?v=oeFBhBMv0fk
26. C. Piltz et al., *A trapped-ion-based quantum byte with* 10^{-5} *next-neighbour cross-talk*, Nat. Commun. **5**, 4679, 2014, doi:10.1038/ncomms5679.
27. https://www.eleqtron.com/en/
28. I. Pogorelov et al., *Compact Ion-Trap Quantum Computing Demonstrator*, PRX QUANTUM **2**, 020343, 2021, doi:10.1103/PRXQuantum.2.020343.
29. D. M. Greenberger et al., *Going Beyond Bell's Theorem.* In: M. Kafatos (eds), Bell's Theorem, Quantum Theory and Conceptions of the Universe, Fundamental Theories of Physics, vol. 37, Springer, Dordrecht, 69–72, 1989, doi:10.1007/978-94-017-0849-4_10.
30. E. Hyyppä et al., *Unimon qubit*, arXiv:2203.05896.
31. https://opensuperq.eu/
32. https://www.q-solid.de/
33. https://qt.eu/about-quantum-flagship/newsroom/european-quantum-computing-simulation/
34. https://jupyter-jsc.fz-juelich.de/hub/login
35. P. Vikstål et al., *Applying the Quantum Approximate Optimization Algorithm to the Tail Assignment Problem*, Phys. Rev. Appl. **14**, 034009, 2020, doi:10.1103/PhysRevApplied.14.034009.
36. D. Willsch et al., *Benchmarking Advantage and D-Wave 2000Q quantum annealers with exact cover problems*, Quantum Inf. Process. **21**, 141, 2022, doi:10.1007/s11128-022-03476-y.

Quantum Computers in 2022

David P. DiVincenzo

Peter Grünberg Institut, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: d.divincenzo@fz-juelich.de

After observing and contributing for nearly thirty years, I offer some perspective on the current state of play in quantum computers.

1 25 Years A-Building – as Assessment

This contribution is only very peripherally about numerically intensive computing. It is rather about the quest to, someday, have a fundamentally new tool to do numerically intensive computing, namely a quantum computer. That someday is today in the form of JUPSI, the D-wave quantum computer at Forschungszentrum Jülich[a]. But my work has as its objective a more ambitious use of qubits, in the form of gate-based quantum-coherent computation. This potentially more powerful alternative has been envisioned for many decades now, but may still be some decades away from its full realisation.

While there is perhaps never a good time to assess a work in progress, except when it is actually completed, I will nevertheless give a small try at it here. Getting to a quantum computer is no small thing, and one should not take such an ultra-marathon without a few breaks in between. I make no pretense to be systematic, comprehensive, or precisely predictive.

First, I will reflect on how some perspectives have remained immovable in 25 years. Immediately after the first big assessment of the possibilities of physical implementation of quantum computers in (I would say) 1994, it was observed that ion traps had many of the right ingredients for quantum computation. A first paper on quantum gates came out very quickly; its speed reflected the fact that hardly anything new had to be done, merely a relabelling of an experiment that was already being done. But pretty soon, a roadmap of sorts was drawn up, with, in my opinion (then as now) pretty much all the right ingredients[b].

This approach has seen steady progress. But why, with such a big head start, has it not been the clear front-runner? I think that here, as elsewhere, the physics-to-engineering transition of mindset has been very slow to be achieved. Maybe one problem can be traced to the varying interpretations of "scalable". At one level, "scalable" just means being able to have more qubits to act on.

Perhaps unfortunately for subsequent developments, ion traps already had, in 1995, a modality to scalability: The linear ion trap could hold N atoms in a straight line. Even to this day, I see this discussed as an acceptable route to scalability, with even the advantage that any qubit can interact and be involved in gate operations with any other qubit. I think that this is a dangerous fantasy. First because all-to-all is not such a huge algorithmic

[a] See `https://juniq.fz-juelich.de/`
[b] See `https://qist.lanl.gov/`

advantage, but it is a huge pain when it comes to suppressing unwanted entanglements. Second, parallel operations throughout the qubit register are difficult or impossible in the linear arrangement. Third, the physics of taking N to infinity doesn't work anyway – the trap strength becomes weaker and weaker as N is increased.

I should of course make explicit my conflicts of interest, and also lack of foresight (but terrific hindsight) in making these critical comments. Right from the outset I was prejudiced towards realising quantum bits via the solid state, since that was my specialty circa 1994. This superficially matched with the orientation of my employer, IBM, which had been a leader in creating enormous infrastructures for constructing solid state (conventional) computers.

2 Quantum Computers and Solid State Physics

My first attempt at this, published in Ref. 1, is summarised in the table from that paper (Table 1). It contains some reasonable guesses, like electrons trapped on impurities or in quantum dots, and solid-state NMR. It was intentionally tongue-in-cheek in part (certainly in the case of the Mössbauer effect), and manifestly incomplete (electrons in gold).

Within a year, work on one of these possibilities really got going: NMR quantum computing, not in the solid state but in molecules in liquid solution. This initial splash pained me: I felt that there were some manifest no-gos in the way that this work got underway.

I perhaps had too snobbish an attitude then (as now) about what was the right way to go about this. Doing real experiments can always uncover unexpected opportunities and necessities for the future. The NMR effort to "factor 15" a couple of years later certainly showed the necessity of a systematic approach to efficient, high-precision pulse control, better than any abstract theoretical analysis of the problem ever could have done. But the violation of the prescription, elementary but perhaps not explicitly spoken yet in the later 1990s, that your computation should start with a cleared register (e.g., all 0s) was strongly violated by the nature of room-temperature NMR experiments. The clever expedient of the "effective pure state" somewhat obscured the unscalability of this approach. This, and the

TABLE I. Important times for various two-level systems in quantum mechanics, which might be used as quantum bits. $t_{\rm switch}$ is the minimum time required to execute one quantum gate; it is estimated as $\hbar/\Delta E$, where ΔE is the typical energy splitting in the two-level system. t_ϕ is the phase coherence time as seen experimentally. t_ϕ is the upper bound on the length of time over which a complete quantum computation can be executed accurately. The ratio of these two times gives the largest number of steps permitted in a quantum computation using these quantum bits.

Quantum system	$t_{\rm switch}$ (sec)	t_ϕ (sec)	Ratio
Mössbauer nucleus [35]	10^{-19}	10^{-10}	10^9
electrons-GaAs [36]	10^{-13}	10^{-10}	10^3
electrons-Au [37,7]	10^{-14}	10^{-8}	10^6
trapped ions–In [38]	10^{-14}	10^{-1}	10^{13}
electron-spin [9]	10^{-7}	10^{-3}	10^4
electron–quantum-dot [39]	10^{-6}	10^{-3}	10^3
nuclear spin [21]	10^{-3}	10^4	10^7

Figure 1. Table reprinted from Ref. 1.

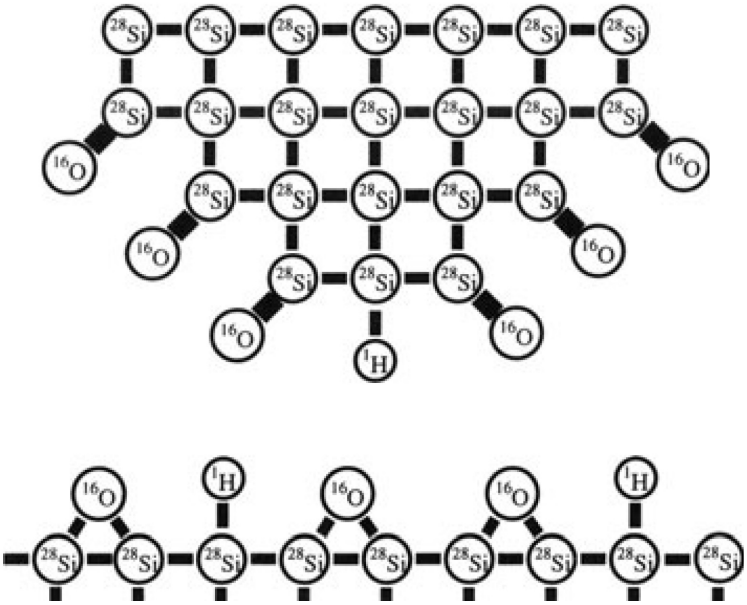


Figure 2. Concept atomic-force microscope quantum computer, reprinted from Ref. 2.

other unscalable aspect of living with an assembly of nuclear spins that could be present in a chemically stable molecule, meant that "factor 15" was not and could not be followed up by "factor 21", let alone "factor 637".

I sought to remedy this state of affairs in an initially scattershot way, by writing offhandedly in various papers some side-comments about what people should be focusing on. By the year 2000, I finally wrote out fully what I had been trying to say since 1994.

Not that I had the full story in 1994, or in 2000. To illustrate my lack of foresight, I would like to point out the best that I could offer in my 1994 paper. This is illustrated in Fig. 2, reproduced from Ref. 2. It posited some atom-scale engineering which made possible a scanning tip of an inert substance with a solitary hydrogen bonded at the end. "Inert" meant something with no unpaired spins, neither electronic nor nuclear, illustrated by a cartoon of silicon (with Si-29 intentionally left out). Thus, the spin of the single proton would be the clean qubit. The tip is to face a similarly constituted surface, with more proton qubits. Some rf pulsing for enabling operations is vaguely alluded to. Knowing that there had to be some idea for initialisation and measurement, I state some ideas for there being electronic paramagnetic centres on the surface, for the purpose both of swapping in a freshly polarised electronic state, and swapping out the nuclear state to make possible some kind of single-electron spin measurement. Scalability was fancifully dealt with by making the tips the ends of gear teeth, with a continually turning apparatus that bring the tips in contact with spin-containing grooves in the adjacent wheels.

3 Some Criteria

OK, enough of this Rube Goldberg (or perhaps Eric Drexler) nonsense. It is perhaps hard to say which aspect of this is the least realistic: structure fabrication is barely one iota closer to making this possible after the passage of 25 years, the mechanics of rapidly moving parts such as these will surely always remain unworkable. When continuing to defend myself five years later, I still felt it necessary to say that quantum computing is not science fiction. Any readers who had been paying attention to my writing would hopefully have come to this conclusion. I cannot blame the real nuclear-spin practitioners from cutting a lot of my fancy corners.

But I insisted that there had to be *some* answer to a minimal set of criteria. Since in the end I composed them as a set of carefully worded aphorisms, I will quote them now[4]:

1. A scalable physical system with well characterised qubits
2. The ability to initialise the state of the qubits to a simple fiducial state
3. Long relevant decoherence times
4. A "universal" set of quantum gates
5. A qubit-specific measurement capability

I am grateful to the anonymous Wikipedia author who quoted these exactly, with some abridgement (which I adopt here); I have seen them quoted far too freely, in ways that miss a lot of the meaning. Long exegeses could be written on the choice of the adjectives in these criteria, all of which were purposefully chosen.

In 1996 we thought (and published in 1998[3]) that the spins of individual electrons would be capable of addressing all these criteria. In our first paper we tried to say some thing on all five points. Our effort was quite influential, but it is clear that we were quite naive on several points. While we knew that nuclear spins would have to be discussed, we were not attuned to what a big role they would play in decoherence. We understood that holding electrons in quantum dots would be a good way to move the control problem to scales greater than the atomic, but we were not attuned to the issue that achieving entanglement over the "long" distance of, say, 200nm, would not leave enough room to fit in the control apparatus that is necessary for a multi-qubit scalable device. We now know that 20.000nm is a better target, and we have now spent many years of efforts to achieve this two order of magnitude increase of length scale in spin-qubit systems. This is, so far, without definitive success.

4 Onward with Superconductors

Superconducting devices did not look promising to me in the years 1994-96. I thought that such macroscopic structures would not have good enough quantum coherence. Events proved my understanding to be wrong, although I would say that it took until 2002 for that to be clearly the case. Many of us had not paid attention to the fact that, already in 1985, superconducting electric circuits had been shown to exhibit energy-level quantisation, confirming the concept of "macroscopic quantum coherence" which Leggett had

advanced (and questioned) at that time[5]. But these "artificial atoms" of 1985, and 1995, and 2000, had very inferior properties to those of real atoms, with very limited manifestation of quantum coherence.

But among experts, Devoret expressed confidence[6] that the designability of electric circuits would provide ways of overcoming the low coherence of the early experiments. The 2002 paper of his group was a watershed moment for this specialty, in which he proved that such design efforts would pay off[8]. The "quantronium" type qubit introduced in this paper did not survive long in the annals of superconducting qubits, but this is because the design paradigm was indeed fruitful and led to rapid subsequent improvement, rendering quantronium obsolete.

Actually, the design flexibility of superconducting quantum circuits had already produced a fertile shoot from the tree in a rather different direction. Following these lines of thinking independently, Hans Mooij had proposed the "flux qubit"[9]. This qubit was designed largely for its ease of measurement, for it has two eigenstates that are magnetically quite distinct, so that the measurement could be done by standard magnetometry. But this distinguishability also made the coherence difficult to improve, although it ultimately was possible many years later. But before this was done, the inherent robustness of the flux qubit, both for measurement and qubit-qubit coupling, convinced the founders of the D-wave company, early in the 2000s, that they should fix on this design for their adiabatic optimiser. Since the concept of this optimiser did not require high quantum coherence, this choice was appropriate for them and remained unchanged through their many cycles of development. The coherence improvements of the mid 2000s onward were not needed in the D-wave strategy, and were not adopted.

For gate-based quantum computing neither the flux qubit, nor quantronium, were adequate, but Devoret's vision of improvement by electrical optimisation proved true. The Yale group made two related, and crucial, discoveries. First, they understood that resonant circuits could boost the sensitivity of measurement, and make the measurement of otherwise hard-to-measure qubits feasible. Then, they designed and made a much better (but hard to measure) qubit by a modification of a very basic qubit circuit, the Cooper-pair box. The resulting qubit, the "transmon", has been the workhorse of many successive generations of coherent processors, from the late 2000s onward.

In truth, the processors that have resulted, especially the universally accessible cloud offerings of IBM from 2016 onward, have sometimes barely deserved the name of "computer"[7]. It seems that, as currently designed, they are guaranteed to crash every 20 seconds or so due to a cosmic ray event[10]. I do not see this as a sign that this is on a track for failure. There are severe growing pains, but there is growth. One can foresee the overcoming of all obstacles. I take it as serious that IBM makes a mention of a million-qubit processor in its public roadmap[c]. The coming years will continue to be very interesting in the development of new quantum computer hardware.

References

1. D. P. DiVincenzo, *Two-bit Gates are Universal for Quantum Computation*, Phys. Rev. A 51, 1015, 1995, arXiv:cond-mat/9407022.

[c] See `https://research.ibm.com/blog/ibm-quantum-roadmap`

2. D. P. DiVincenzo, *Quantum Computation*, Science 270, 255–261, 1995.
3. D. Loss and D. P. DiVincenzo, *Quantum computation with quantum dots*, Phys. Rev. A 57, 120–126, 1998, arXiv:cond-mat/9701055.
4. D. P. DiVincenzo, *The Physical Implementation of Quantum Computationi*, Fortschritte der Physik 48, 771–784, 2000 (Special issue on Experimental Proposals for Quantum Computation), arXiv:quant-ph/0002077. Reprinted in Scalable Quantum Computers - Paving the Way to Realization, Eds. S. L. Braunstein and H.-K. Lo, Wiley-VCH, 1–14, 2001.
5. A. J. Leggett, *Macroscopic Quantum Systems and the Quantum Theory of Measurement*, Progress of Theoretical Physics Supplement, Volume 69, 80–100, March 1980, doi:10.1143/PTP.69.80.
6. M. Devoret, *Quantum fluctuations in electrical circuits*, in Quantum Fluctuations / Les Houches, Eds. S. Reynaud, E. Giacobino, J. Zinn-Justin, Elsevier, 351–386, 1997.
7. K. Michielsen, M. Nocon, D. Willsch, F. Jin, T. Lippert, H. De Raedt, *Benchmarking gate-based quantum computers*, Computer Physics Communications, Volume 220, 44–55, 2017, doi:10.1016/j.cpc.2017.06.011.
8. D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret, *Manipulating the Quantum State of an Electrical Circuit*, Science, Vol 296, Issue 5569, 886–889, 3 May 2002, doi:10.1126/science.1069372.
9. T. P. Orlando, J. E. Mooij, L. Tian, C. H. van der Wal, L. S. Levitov, S. Lloyd, and J. J. Mazo, *Superconducting persistent-current qubit*, Phys. Rev. B 60, 15398, 1999.
10. L. Cardani, F. Valenti, N. Casali, et al., *Reducing the impact of radioactivity on quantum circuits in a deep-underground facility*, Nature Communications 12, 2733, 2021.

Hybrid Quantum Classical Simulations

Dennis Willsch[1], **Manpreet Jattana**[1,2], **Madita Willsch**[1,3], **Sebastian Schulz**[1,2], **Fengping Jin**[1], **Hans De Raedt**[1,4], **and Kristel Michielsen**[1,2,3]

[1] Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {d.willsch, m.jattana, m.willsch, se.schulz, f.jin, k.michielsen}@fz-juelich.de

[2] RWTH Aachen University, 52056 Aachen, Germany

[3] AIDAS, 52425 Jülich, Germany

[4] Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands
E-mail: deraedthans@gmail.com

We report on two major hybrid applications of quantum computing, namely, the quantum approximate optimisation algorithm (QAOA) and the variational quantum eigensolver (VQE). Both are hybrid quantum classical algorithms as they require incremental communication between a classical central processing unit and a quantum processing unit to solve a problem. We find that the QAOA scales much better to larger problems than random guessing, but requires significant computational resources. In contrast, a coarsely discretised version of quantum annealing called approximate quantum annealing (AQA) can reach the same promising scaling behaviour using much less computational resources. For the VQE, we find reasonable results in approximating the ground state energy of the Heisenberg model when suitable choices of initial states and parameters are used. Our design and implementation of a general quasi-dynamical evolution further improves these results.

1 Introduction

Quantum computing[1] is an emerging computer technology that uses quantum effects in the design of its computational model. There are two major paradigms in quantum computing, namely the gate-based quantum computer and the quantum annealer.

A gate-based quantum computer is inspired by the circuit model for classical, digital computers. This means that every program is defined in terms of a sequence of fundamental operations, the so-called *quantum gates*. Each quantum gate operates on the fundamental units of computation, the so-called *quantum bits* or *qubits*. The gate-based quantum computer executes these quantum gates step-by-step and thus updates its internal quantum state. At the end of the quantum gate sequence, the quantum state is "measured", which means that the quantum computer outputs, with a certain probability, one out of several classical bitstrings. Typically, the number of bits in this bitstring is given by the number of qubits in the quantum computer. At the time of writing, gate-based quantum computers with approximately one hundred qubits have been manufactured[2].

A quantum annealer, on the other hand, tries to harness the natural evolution of a quantum system, steered by external magnetic fields, to solve an optimisation problem. Internally, it also operates on a set of qubits, which are measured at the end of a quantum annealing run. It thus also produces a bitstring as output. At the time of writing, quantum annealers with more than 5000 qubits have been manufactured[3].

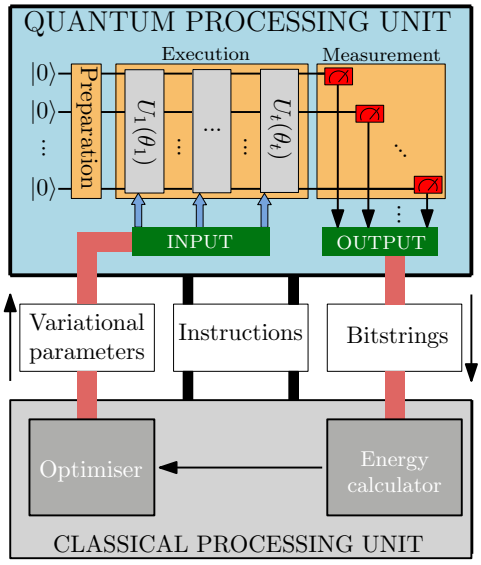


Figure 1. **Schematic diagram of hybrid quantum classical variational algorithms.** The tasks are divided between a QPU and a CPU. The CPU suggests certain parameters of a parametrised quantum circuit. This quantum circuit, along with other instructions, is then sent to and executed by the QPU. Following this, the measurement results from the QPU are sent back to the CPU, which in turn evaluates the energy to be optimised. From this result, the CPU obtains new parameters for the quantum circuit. This process continues until some convergence criterion is met.

Both types of quantum computers share the property that the same program can produce different bitstrings obtained after a run. The computational models are thus inherently probabilistic. This means that quantum computers can also be used as *samplers*. A program is often executed multiple times to obtain a representative distribution of bitstrings.

In this article, we consider simulations of hybrid quantum classical variational algorithms. Such algorithms obtain their name from the combined usage of gate-based quantum and classical computers in a single application. The working of such algorithms is visualised in Fig. 1. The optimiser in the classical central processing unit (CPU) sends the parameters placed in a circuit to be executed by the quantum processing unit (QPU). The QPU prepares a problem specific initial state, executes the circuit, and sends the results to the CPU after measurement. The outcomes of the measurements, called bitstrings, are used to calculate the energy which is given to the optimiser. The optimiser then decides what new parameters will lower the energy, and the cycle continues until convergence. The CPU also controls the QPU through other instructions, e.g. the number of measurements, microwave pulses, or the time between measurements.

This article is structured as follows. Sec. 2 contains an overview of the Jülich Universal Quantum Computer Simulator, which is used for the simulation of the following two hybrid quantum classical applications. In Sec. 3, we introduce the QAOA, discuss its relation to AQA, and compare both algorithms when applied to the tail assignment problem. In Sec. 4, we discuss the VQE and its application to the Heisenberg model. Sec. 5 contains our conclusions.

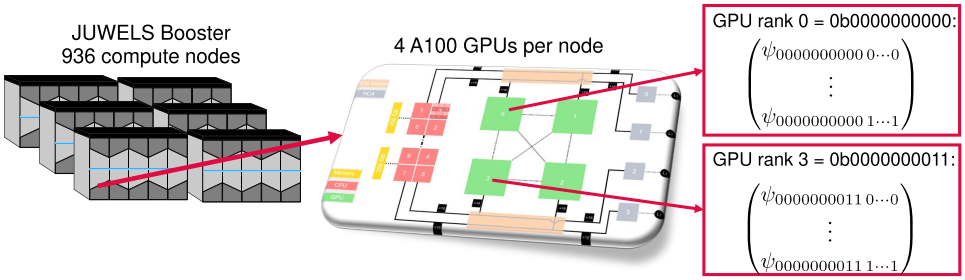


Figure 2. **Distribution of the coefficients of the state vector $|\psi\rangle$ across the GPUs and compute nodes of JUWELS Booster[7].** Each GPU is handled by a single MPI process. For each GPU, the global qubit indices of the coefficients represent its MPI rank. For the GPUs belonging to MPI rank 0 and 3 this is indicated on the right (the 10 leftmost indices are the global qubit indices). The complex coefficients of the local qubits are stored locally on the GPUs.

2 Jülich Universal Quantum Computer Simulator

In this section, we briefly describe the Jülich Universal Quantum Computer Simulator (JUQCS), and outline how a program for a gate-based quantum computer can be simulated using supercomputers. More details and in-depth descriptions of the implementation are given in Refs. 4–6.

The basic unit of computation for quantum computers is the qubit. Mathematically, a qubit is described by a unit vector of two complex numbers $|\psi\rangle = (\psi_0, \psi_1)$ fulfilling $\langle\psi|\psi\rangle = |\psi_0|^2 + |\psi_1|^2 = 1$. Usually, the basis states (i.e., orthonormal basis vectors) are denoted by $|0\rangle$ and $|1\rangle$. An N-qubit system is defined by 2^N complex numbers

$$|\psi\rangle = \psi_{0\ldots00}\,|0\ldots00\rangle + \psi_{0\ldots01}\,|0\ldots01\rangle + \ldots + \psi_{1\ldots11}\,|1\ldots11\rangle\,, \tag{1}$$

where $|0\ldots00\rangle, \ldots, |1\ldots11\rangle$ denote the computational basis states[1] and the complex coefficients $\psi_{0\ldots00}, \ldots, \psi_{1\ldots11}$ again fulfil $\langle\psi|\psi\rangle = 1$. The 2^N complex coefficients in the state $|\psi\rangle$ can be written as a rank-N tensor $\psi_{q_{N-1}\cdots q_1 q_0}$ where $q_j \in \{0, 1\}$ denote the indices.

Since the number of coefficients grows exponentially in the number of qubits N, the memory requirement grows exponentially too, which makes large-scale simulations of universal quantum computers only possible on supercomputers with enough (distributed) random access memory. To simulate for instance a quantum computer with $N = 42$ qubits, the tensor $\psi_{q_{N-1}\cdots q_1 q_0}$ requires $16 \times 2^{42}\,\mathrm{B} = 64\,\mathrm{TiB}$ of memory (using double precision floating-point numbers) while the simulation of $N = 21$ qubits only requires $16 \times 2^{21}\,\mathrm{B} = 32\,\mathrm{MiB}$ of memory.

The GPU version of JUQCS (JUQCS–G) distributes the complex coefficients over the memory of the participating GPUs. A sketch is shown in Fig. 2. Each GPU stores a power of two coefficients, say 2^M coefficients, of the state vector $|\psi\rangle$ in its local memory. Qubits j whose amplitudes $\psi_{q_{N-1}\ldots q_{j+1}0q_{j-1}\ldots q_0}$ and $\psi_{q_{N-1}\ldots q_{j+1}1q_{j-1}\ldots q_0}$ are stored on the same local memory are thus called *local* qubits. Qubits whose amplitudes are distributed over different GPUs are called *global* qubits. The total number of GPUs needed is $N_{\mathrm{GPU}} = 2^{N-M}$.

Since the state vector is distributed over several GPUs, potentially also over different compute nodes, it is necessary to transfer data from one GPU to another over the network, for instance if (part of) the data is needed for a computation on another GPU or node. We use the *Message Passing Interface* (MPI) to transfer the data between different GPUs. Each MPI process controls a GPU, and its rank $r \in \{0, \ldots, N_{\mathrm{GPU}} - 1\}$ is given by the global qubit indices in binary notation.

The application of a quantum gate is performed as the matrix-vector update $U|\psi\rangle$, where the matrix U is unitary and acts only on a few-qubit subspace. For instance, the matrix corresponding to a single-qubit gate is a 2×2 unitary matrix $U_1 = (u^{(1)}_{qq'})$. It acts only on a single-qubit subspace and thus transforms the coefficients of $|\psi\rangle$ in terms of 2-component updates: Acting on qubit j, U_1 transforms the state vector $|\psi\rangle$ according to

$$\psi_{q_{N-1}\cdots q_{j+1} q q_{j-1}\cdots q_0} \leftarrow \sum_{q'=0}^{1} u^{(1)}_{qq'} \psi_{q_{N-1}\cdots q_{j+1} q' q_{j-1}\cdots q_0}\,, \tag{2}$$

for $q = 0, 1$. Similarly, a two-qubit gate is a 4×4 unitary matrix that acts on a two-qubit subspace of $|\psi\rangle$, and so on. A suitable set of one- and two-qubit gates (called a universal gate set) is sufficient for the implementation of a universal quantum computer (simulator)[8,9]. The full gate set of JUQCS is documented in Ref. 5.

The implementation of JUQCS exploits the special structure of single-, two- and three-qubit operations. Large, dense matrices are never stored nor operated on. It is sufficient to loop over pairs of coefficients of the state vector and to perform the 2-component updates given in Eq. 2 for the same 2×2 matrix U_1 (how U_1 looks like depends on the particular gate). Depending on the index of the qubit that the gate operates on, the pairs of coefficients are grouped differently. Accordingly, for two-qubit gates, the loop goes over quadruples of coefficients of $|\psi\rangle$ and the 4-component updates are performed with a 4×4 matrix (whose entries are determined by the particular gate).

A quantum gate acting on a global qubit requires MPI communication, because coefficients of $|\psi\rangle$ that need to be combined in the update are stored on different GPUs. Thus, to perform a single-qubit gate on a global qubit, the transfer of $2^N/2$ coefficients (i.e., half of the state vector) between pairs of GPUs is required. The communication overhead is minimised by *relabelling* global and local qubits after the transfer happened once. Thus, the coefficients do not need to be transferred back again after the transformation. Each MPI rank has to keep track of the current labelling of global and local qubits. More details on this MPI communication scheme are given in Ref. 4.

3 Quantum Approximate Optimisation Algorithm

The QAOA[10] is an algorithm for gate-based quantum computers to solve combinatorial optimisation problems. Such problems can be mapped to binary optimisation problems, which are defined in terms of an N-bit binary string $\mathbf{x} = x_1 \ldots x_N$, with the goal of finding a bitstring $\mathbf{x}$ that minimises a classical objective function $C(\mathbf{x}) : \{0,1\}^N \to \mathbb{R}$.

For quantum computers, a natural way of encoding a binary optimisation problem is to express its solution as the ground state (i.e., the state corresponding to the lowest energy)

of the so-called Ising Hamiltonian

$$H_{\mathrm{C}} = \sum_{i=0}^{N-1} h_i \sigma_i^z + \sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z. \tag{3}$$

Here, $\{h_i\}$ and $\{J_{ij}\}$ denote real numbers that represent the optimisation problem instance, and $\sigma_i^z = I \otimes I \cdots I \otimes \mathrm{diag}(1,-1) \otimes I \otimes \cdots \otimes I$ (with I being the 2×2 identity matrix) is the Pauli z matrix acting on the i^{th} qubit subsystem.

The idea of the QAOA is to solve the optimisation problem by bringing the quantum computer into a variational state

$$|\beta, \gamma\rangle = \left(\prod_{k=p}^{1} \exp\left[-i\beta_k H_{\mathrm{D}}\right] \exp\left[-i\gamma_k H_{\mathrm{C}}\right] \right) |+\rangle^{\otimes N}, \tag{4}$$

where $\gamma = (\gamma_1, ..., \gamma_p)$ and $\beta = (\beta_1, ..., \beta_p)$ are the $2p$ variational parameters to be optimised, $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ is the uniform superposition of the qubit state $|0\rangle$ and $|1\rangle$, H_C is the cost Hamiltonian encoding the optimisation problem (see Eq. 3), and $H_{\mathrm{D}} = -\sum_{i=0}^{N-1} \sigma_i^x$ is a so-called driving Hamiltonian (i.e., a sum of Pauli x matrices). By finding optimal angles β and γ that minimise the energy expectation value $E_p(\beta, \gamma) = \langle \beta, \gamma | H_{\mathrm{C}} | \beta, \gamma \rangle$, one hopes to find the ground state of H_{C}. As $|\beta, \gamma\rangle$ is a linear combination of 2^N potential states, the idea is that optimising the $2p$ variational parameters may be much simpler than solving the original optimisation problem. However, note that in general, determining a set of optimal β and γ may require numerous queries to a quantum computer. A quantum circuit to bring the quantum computer into the variational state $|\beta, \gamma\rangle$ given by Eq. 4 is shown in Fig. 3.

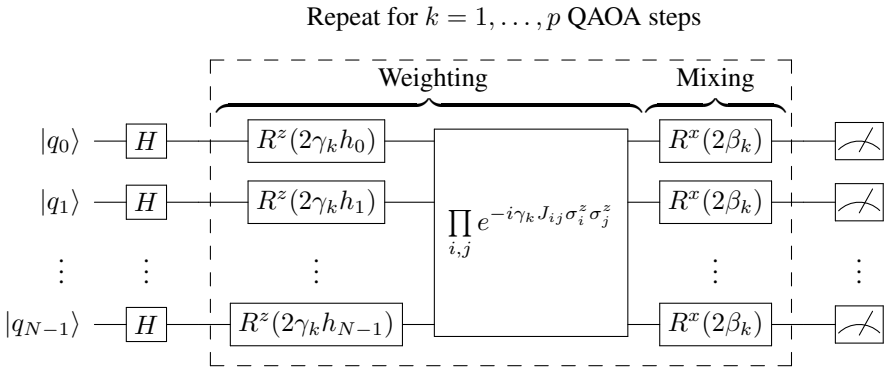


Figure 3. **General QAOA circuit.** Initially, we bring the qubits into a uniform superposition over all states using the Hadamard gate H. Then, we apply $k = 1, \ldots, p$ QAOA steps with parameters $\beta_1, \ldots, \beta_p$ and $\gamma_1, \ldots, \gamma_p$ (see Eq. 4). Each QAOA step k consists of a "weighting step" using single-qubit z rotation gates $R^z(\varphi) = \exp(-i\varphi\sigma^z/2)$ and the set of two-qubit gates indicated in the circuit, followed by a "mixing step" using single-qubit x rotation gates $R^x(\varphi) = \exp(-i\varphi\sigma^x/2)$. The result after the measurement gates is used to update the variational parameters until the energy is sufficiently low (cf. Fig. 1).

3.1 Emulation of the QAOA

To emulate the QAOA, JUQCS produces the variational state $|\beta, \gamma\rangle$ (see Eq. 4) by simulating the quantum circuit in Fig. 3 for given variational parameters β and γ. For the optimisation phase of the QAOA, the energy expectation value in the variational state $E_p(\beta, \gamma) = \langle\beta, \gamma| H_C |\beta, \gamma\rangle$ is computed by JUQCS and passed to the optimiser (cf. Fig. 1; we use optimisers from the `scipy` library[11]). The optimiser then returns new values for the variational parameters, which are in turn used for the next simulation of the QAOA circuit, and so on. In case the optimisation does not converge or if it converges too slowly, an additional stopping criterion of a maximum of 200 calls to JUQCS is used.

Note that in practice, the only quantity that can be estimated using a real quantum computer is the energy expectation value $E_p(\beta, \gamma)$, and not the success probability P_{success} (i.e., the probability of the ground state in the variational state $|\beta, \gamma\rangle$). Although we could in principle also optimise for the success probability, we consider the realistic situation and optimise for the energy only. However, the success probability is also evaluated by JUQCS to assess the quality of the solution (cf. Fig. 4 below).

3.2 Approximate Quantum Annealing

In this section, we first review how to simulate the time evolution of a quantum annealing process, and then we explain under which conditions we refer to "approximate" quantum annealing (AQA). Additionally, we point out the structural relation to the QAOA introduced above. A more detailed discussion of AQA and its comparison to the QAOA is given in Ref. 6.

In order to simulate a quantum annealing process, the time-dependent Schrödinger equation (TDSE) with $\hbar = 1$, $i\partial_t |\psi(t)\rangle = H(t) |\psi(t)\rangle$, for the quantum annealing Hamiltonian $H(t) = A(t)H_{\text{D}} + B(t)H_{\text{C}}$ has to be solved at times $0 \leq t \leq t_{\text{anneal}}$, where t_{anneal} is the annealing time. Here, H_D (H_C) is the driving (cost) Hamiltonian used in Eq. 4, and the so-called annealing functions A and B satisfy $A(0) \gg B(0)$ and $A(t_{\text{anneal}}) \ll B(t_{\text{anneal}})$. Therefore, we also call $H(0) \approx H_D$ and $H(t_{\text{anneal}}) \approx H_C$ the initial and final Hamiltonian, respectively.

The TDSE can be solved by time-stepping using the second-order Suzuki-Trotter product formula[12,13]

$$|\Psi((l+1)\tau)\rangle = \left\{ \exp\left[-\frac{i\tau}{2} A(l\tau) H_D\right] \times \exp[-i\tau B(l\tau) H_C] \right. \\ \left. \times \exp\left[-\frac{i\tau}{2} A(l\tau) H_D\right] \right\} |\Psi(l\tau)\rangle \,, \tag{5}$$

where τ denotes the time step and $l = 0, \ldots, n$ (such that $t_{\text{anneal}} = (n+1)\tau$). The action of the three matrix exponentials in Eq. 5 on any state vector can be computed analytically. The initial state is taken to be the ground state of H_{D}, namely $|\Psi(0)\rangle = |+\rangle^{\otimes N}$. Note that the structure of Eq. 5 and of the QAOA (cf. Eq. 4) are essentially the same.

The idea of AQA is now to use a time step τ that is too large and a number of time steps n that is too small to yield an accurate discretisation of the time evolution. Thus, we do not expect that the sequence of states $|\Psi((l+1)\tau)\rangle$ resembles the solution of the time evolution of a quantum annealing process. However, we hope that we can still generate a final state which has a relatively large overlap with the ground state of the final problem

Hamiltonian H_C by using only a small number of matrix-vector operations. This heuristic method is motivated by empirical findings[14–18] that the values for β and γ obtained from the optimisation were often found to approximately follow curves that are suitable as coarsely discretised quantum annealing functions A and B.

The amounts of single-qubit and two-qubit gates required to implement either one time step of AQA, or one layer of the QAOA, are the same. Thus, for fixed n, (i.e., $n+1$ AQA time steps) the computational effort is equivalent to a single QAOA circuit evaluation (i.e., no optimisation) where $p = n + 1$.

3.3 Tail Assignment Problem

The tail assignment problem[19] is a formulation of aircraft assignment problems that respects typical operational constraints such as connection times, preassigned activities, maintenance and airport curfews. We consider the constraint-only version of the tail assignment problem used in Ref. 20 (other models for the tail assignment problem can be found in Refs. 18, 21, 22). In this form, the problem can be formulated as an exact cover problem, which is an NP-complete problem[23] from the class of set covering and partitioning problems. The corresponding classical objective function to be minimised reads

$$C(\mathbf{x}) = \sum_{f=1}^{F} \left(\sum_{i=1}^{N} \mathbf{A}_{if} x_i - 1 \right)^2 . \tag{6}$$

Here, $\mathbf{A} \in \{0,1\}^{N \times F}$ is a Boolean matrix that defines the exact cover problem instance. In the context of aircraft assignment, each of the F columns of $\mathbf{A}$ represents one flight, and each of the N rows represents a given route. Thus, $\mathbf{A}_{if} = 1$ if and only if flight f is contained in route i. The solution $\mathbf{x} = x_1 \ldots x_N$ therefore selects a set of aircraft routes from $\mathbf{A}$ such that each flight is covered exactly once (see Ref. 20, which also contains the explicit expression for H_C in Eq. 3 as a function of $\mathbf{A}$). We consider the same $N = 30, \ldots, 40$ qubit problems with $F = 472$ flights that have been used in Ref. 6.

3.4 Results

In this section, we summarise some representative QAOA and AQA results. In Fig. 4, we plot the success probability P_{success} as a function of the problem size N for two versions of AQA and the QAOA. The AQA results are for $n = 5$ steps with time step $\tau = 0.1$ (green squares) and $\tau = 0.8$ (green crosses). The QAOA results are for standard QAOA (i.e., the variational parameters are optimised individually for each problem instance, initialised from AQA with $n = 6$ and $\tau = 0.4$; blue circles), and for pre-optimised QAOA (i.e., the variational parameters are optimised for a 30-qubit problem, initialised from AQA with $n = 5$ and $\tau = 0.2$, and these same values are reused for all larger problem instances; blue triangles). Further results and additional details can be found in Ref. 6.

Both QAOA and AQA obtain better success probabilities than can be found by random guessing (red dash-dot-dotted line), but for the QAOA they are overall higher than for AQA. This is reasonable since the variational QAOA parameters have been initialised with AQA results. More important, however, is the scaling behaviour as a function of the problem size N. For AQA with $\tau = 0.1$ (dash-dotted line), the scaling is equal to the scaling of random guessing ($1/2^N$). The scaling of the QAOA with pre-optimised variational

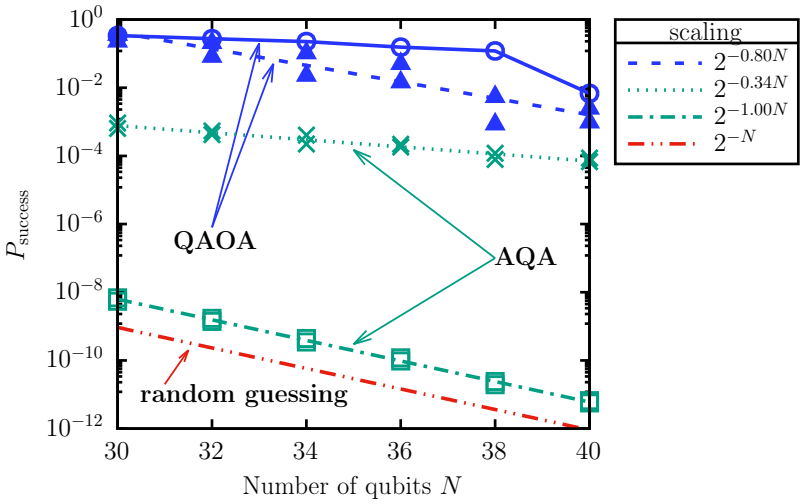


Figure 4. **Success probabilities P_{success} obtained from AQA and the QAOA, in comparison to random guessing, for different problem sizes N.** The success probabilities obtained with the QAOA are shown for a case in which the variational parameters are separately optimised for each problem size N (blue circles, $p = 7$), and for a case in which fixed variational parameters are taken from a previous optimisation of the $N = 30$ case (blue triangles, $p = 6$). The success probabilities obtained with AQA are shown for $\tau = 0.8$ (green crosses) and $\tau = 0.1$ (green squares). The dashed, dotted, and dash-dotted lines represent fits to the average success probabilities (in these cases, two instances per problem size have been simulated). The red dash-dot-dotted line represents the probability $1/2^N$ to find the solution by random guessing. The solid line is a guide to the eye.

parameters (dashed line) is only marginally better. However, both AQA with $\tau = 0.8$ (dotted line) and standard QAOA (solid line) show considerably good scaling behaviour (albeit still exponential).

Merely the particular QAOA data point for the largest problem size (the blue circle at $N = 40$) falls out of this scaling behaviour. Here, the maximum number of 200 JUQCS-calls (and thus the limit of the computational budget allocated for this job) was reached. This represents the fact that the QAOA optimisation procedure requires a significant amount of computational work on many nodes of the supercomputer.

This is a noteworthy advantage of AQA, as it can extend the good scaling also to $N = 40$ without requiring any involved optimisation procedure once a good value for τ is found. Interestingly, the regime of "good" values appears to be for large time steps τ, where AQA is actually not a genuine discretisation of quantum annealing anymore.

4 Variational Quantum Eigensolver

An interesting and common problem is to find the ground state energy of a Hamiltonian H. As this is a hard problem to solve on classical computers, quantum computers are considered as an alternative route to solve the problem. When the ground state $|\psi_0\rangle$ can be

prepared on a quantum computer, the expectation value

$$\langle H \rangle_0 = \frac{\langle \psi_0 | H | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} \tag{7}$$

gives the ground state energy. The task is to find this ground state energy by initialising the quantum computer in an arbitrary initial trial state $|\psi_t\rangle$. The variational theorem states that $\langle H \rangle_t$ for any $|\psi_t\rangle \neq |\psi_0\rangle$ is an upper bound to the ground state energy. This means that

$$\langle H \rangle_t \geq E_0, \tag{8}$$

where the equality sign holds if and only if $|\psi_t\rangle = |\psi_0\rangle$. In Eq. 8, E_0 is the lowest eigenvalue or the ground state energy of the physical system described by H.

The VQE is an implementation of the variational principle on a quantum computer. Its working is illustrated in Fig. 1. The VQE derives its name from the fact that it was used to solve for the lowest eigenvalue of a Hamiltonian[24].

4.1 Heisenberg Model

The Hamiltonian for the one-dimensional Heisenberg model is given by

$$H = \sum_{i>j}^{N} \Big(J_{ij}^{xx} \sigma_i^x \cdot \sigma_j^x + J_{ij}^{yy} \sigma_i^y \cdot \sigma_j^y + J_{ij}^{zz} \sigma_i^z \cdot \sigma_j^z \Big), \tag{9}$$

where N denotes the number of spins, σ^x, σ^y, and σ^z are the Pauli matrices, and J_{ij} is the coupling interaction between the i^{th} and j^{th} spins. We use units such that $\hbar = 1$ and the J's are dimensionless. If $J_{ij} = 1$ for all $i, j = 1, \ldots, N$, we call H the *isotropic antiferromagnetic* model Hamiltonian. We consider bipartite spin lattices in our simulations.

4.1.1 The Ansatz

The variational principle allows any ansatz to be used to solve the ground state energy problem. However, not every ansatz can express the ground state of H. An ansatz[25] that can find a reasonable approximation to the ground state energy is given by

$$U(\boldsymbol{\theta}) = \Big[\prod_{l=N-1}^{1} \prod_{k=N}^{l+1} U_{lk}(\theta_{lk}) \Big] \Big[\prod_{l=N-1}^{1} \prod_{k=N}^{l+1} U_{kl}(\theta_{kl}) \Big], \tag{10}$$

where

$$U_{pq}(\theta_{pq}) = \begin{cases} e^{-i\theta_{pq}\sigma_p^y \sigma_q^x} & \text{if } p = N \text{ or } q = N, \\ e^{-i\theta_{pq}\sigma_p^y \sigma_q^x \sigma_N^z} & \text{otherwise.} \end{cases} \tag{11}$$

If an ansatz can express the ground state, it is still challenging to optimise the parameters such that $|\psi_t(\boldsymbol{\Theta})\rangle = |\psi_0\rangle$, where $\boldsymbol{\Theta}$ are the optimised parameters. In addition to the ansatz, it is important to specify the initial state (say $|\Psi_0\rangle$) of the quantum computer. The ansatz then acts on the initial state, i.e. $\boldsymbol{U}(\boldsymbol{\theta}) |\Psi_0\rangle$, with a set of initial parameters to give the initial energy. The task of the VQE is to lower this energy until it equals the ground state energy. A better choice of initial energy can improve the performance of the VQE. To obtain a better (than random) initial energy for the Heisenberg model, it helps to set the

initial state to the Néel state and the initial parameters to zero. On a lattice with an even number of spins, the Néel state is prepared by setting adjacent spins anti-parallel. On a quantum computer, and for a one dimensional model, this is achieved by setting the qubits with the odd (even) indices to $|0\rangle$ ($|1\rangle$), or *vice versa*.

4.1.2 Results

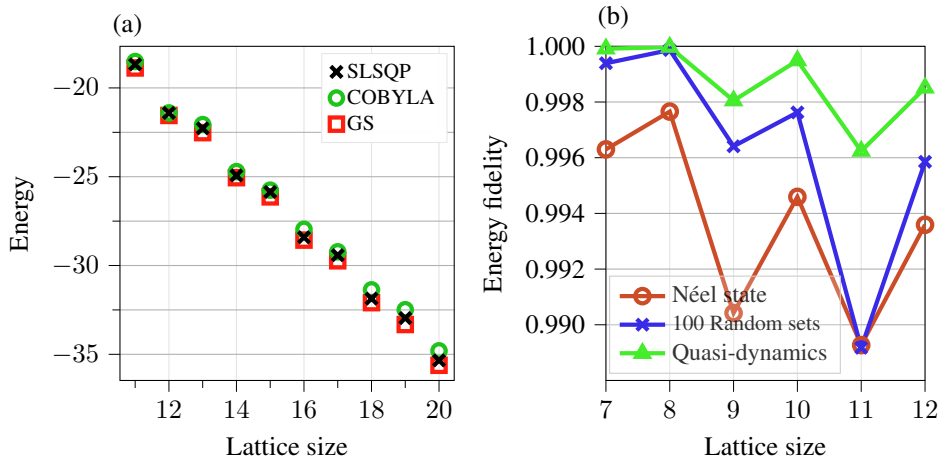


Figure 5. (a) Optimised variational energies obtained using gradient based (SLSQP) and gradient free (COBYLA) algorithms in comparison to the ground state energies for different lattice sizes. (b) Comparison between the energy fidelities obtained using the Néel initial state, best of 100 random sets of parameters, and quasi-dynamics.

We find the ground state energy of isotropic Heisenberg rings of length 11 to 20 using two different optimisers. We used the gradient based (free) optimiser SLSQP (COBYLA)[26,11,27]. The results are compared against the ground state energy found by the Lanczos algorithm and are shown in Fig. 5(a). We observe, as expected, that using different optimisers leads to different results. SLSQP performs better in all cases but the difference in obtained energies appears to grow with increasing lattice sizes.

4.2 VQE and Quasi-Dynamics

When using the VQE, the variational quantum state as a function of m parameters is written as

$$|\psi\rangle = U(\boldsymbol{\theta})\,|\Psi_0\rangle = U_m(\theta_m)\ldots U_1(\theta_1)\,|\Psi_0\rangle\,. \tag{12}$$

If $\mathbb{U}(\boldsymbol{\Theta})$ ($U(\boldsymbol{\theta})$) represents the unitary operators corresponding to the optimised (unoptimised) numeric values of the parameters, then the VQE can be said to perform the task $U(\boldsymbol{\theta}) \to \mathbb{U}(\boldsymbol{\Theta})$. After the optimiser signals convergence, the state obtained is

$$|\Psi_1\rangle = \mathbb{U}(\boldsymbol{\Theta})\,|\Psi_0\rangle\,. \tag{13}$$

The idea is to use this state as an initial state for another round of VQE optimisation. To do so, substitute Eq. 13 in Eq. 12 repeatedly (p times), such that the final state is

$$|\Psi_p\rangle = \mathbb{U}_p(\boldsymbol{\Theta}_p) \ldots \mathbb{U}_1(\boldsymbol{\Theta}_1) |\Psi_0\rangle . \tag{14}$$

The procedure represented by Eq. 14 is termed *quasi-dynamical evolution*[28]. A suitable choice of initial parameters for each $U(\boldsymbol{\theta})$ is $\boldsymbol{\theta} = [0, \ldots, 0]$. This step is necessary in order to preserve the progress made until then. By avoiding random initialisation we also avoid potential barren plateaus[29]. Similar to the QAOA, the performance of the quasi-dynamics is expected to improve as p is increased.

4.2.1 Results

We perform the quasi-dynamical evolution on isotropic rings of lengths 7 to 12. We set the threshold for the quasi-dynamics such that the process stops when the energy at step p can no longer be lowered below 10^{-4} compared to the $(p-1)^{\text{th}}$ value. The results are compared against the final energy obtained when (1) the Néel initial state and (2) one hundred random sets of initial parameters were used. We define the energy fidelity as the ratio of the variational energy obtained upon convergence of the optimiser to the ground state energy. For (2), we plot only the highest energy fidelity obtained from all cases. The results are shown in Fig. 5(b). We observe that the quasi-dynamical evolution can find a better energy fidelity.

In our results we used the same $\boldsymbol{U}(\boldsymbol{\theta})$ (from Eq. 10) for each of the p iterations in Eq. 14. While it improves the results for the cases we tested, such an elementary choice may not always converge to the ground state energy. Eq. 14 allows the use of any arbitrary $\boldsymbol{U}(\boldsymbol{\theta})$; however, it remains an open question which choices of $\boldsymbol{U}(\boldsymbol{\theta})$ will find the ground state energy for problems in general.

We observed that the number of random restart cases that found an energy better than the energy found when starting from the Néel state dropped as the lattice size is increased (data not shown). Since the ansatz has more parameters for larger lattices, one solution is to increase the number of restarts. However, this approach can be very resource expensive. Thus, random initialisation is not a suitable strategy. It is crucial to have a good combination of ansatz and initial parameters to have a reasonable chance of finding the ground state energy.

5 Conclusions

In this article, we have discussed two popular hybrid quantum classical algorithms, namely the QAOA and the VQE. Both have been simulated on JUWELS Booster using JUQCS–G.

For the QAOA, we found that the best scaling of the success probability can be obtained when the full variational optimisation procedure is carried out. However, due to tremendous computational resources required for simulations with $N = 40$ qubits, it was not possible to reach convergence for the largest problem size. Therefore, we considered AQA as an alternative algorithm with only a single parameter τ to solve this problem. Although slightly worse in terms of actual success probabilities, AQA could extend the same promising scaling behaviour also to the largest problems. Interestingly, the resulting

optimal τ was so large that it put AQA into a regime where it does not accurately describe a discretised quantum annealing process anymore.

For the VQE, we found that a suitable choice of an ansatz coupled with a good choice of initial parameters is critical for finding the ground state energies. For the Heisenberg model, one such choice is the Néel state with all variational parameters initialised to zero. However, which choices are helpful in general is an open question. We introduced and tested quasi-dynamical evolution, a technique which builds upon the VQE and improves its performance. An open question for the quasi-dynamics is the choice of $\boldsymbol{U}$ in each substitution which guarantees finding the ground state energy.

Acknowledgements

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 at Jülich Supercomputing Centre (JSC). We would like to thank M. Svensson for providing the exact cover problem instances. D.W. and M.W. acknowledge support from the project Jülich UNified Infrastructure for Quantum computing (JUNIQ) that has received funding from the German Federal Ministry of Education and Research (BMBF) and the Ministry of Culture and Science of the State of North Rhine-Westphalia. M. J. acknowledges support from the project OpenSuperQ (820363) of the EU Quantum Flagship.

References

1. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, Cambridge University Press, New York, 2010.
2. IBM Q team, "127-qubit backend: IBM Q 127 Washington backend specification V1.1.0", 2022.
3. C. McGeoch and P. Farré, *The D-Wave Advantage System: An Overview*, Tech. Rep., D-Wave Systems Inc, Burnaby, BC, Canada, 2020, D-Wave Technical Report Series 14-1049A-A.
4. K. De Raedt, K. Michielsen, H. De Raedt, B. Trieu, G. Arnold, M. Richter, T. Lippert, H. Watanabe, and N. Ito, *Massively parallel quantum computer simulator*, Comput. Phys. Commun., **176**, no. 2, 121, 2007.
5. 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*, Comput. Phys. Commun., **237**, 47–61, 2019.
6. D. Willsch, M. Willsch, F. Jin, K. Michielsen, and H. De Raedt, *GPU-accelerated simulations of quantum annealing and the quantum approximate optimization algorithm*, Comput. Phys. Commun., **278**, 108411, 2022.
7. S. Kesselheim, A. Herten, K. Krajsek, J. Ebert, J. Jitsev, M. Cherti, M. Langguth, B. Gong, S. Stadtler, A. Mozaffari, G. Cavallaro, R. Sedona, A. Schug, A. Strube, R. Kamath, M. G. Schultz, M. Riedel, and T. Lippert, "JUWELS Booster – A Supercomputer for Large-Scale AI Research", 2021.
8. D. Deutsch, A. Barenco, and A. Ekert, *Universality in Quantum Computation*, Proc. R. Soc. Lond. A, **449**, no. 1937, 669, 1995.

9. D. P. DiVincenzo, *Two-bit gates are universal for quantum computation*, Phys. Rev. A, **51**, 1015–1022, Feb 1995.
10. E. Farhi, J. Goldstone, and S. Gutmann, *A Quantum Approximate Optimization Algorithm*, 2014, arXiv:1411.4028.
11. E. Jones, T. Oliphant, P. Peterson, et al., "SciPy: Open source scientific tools for Python", 2001.
12. M. Suzuki, *General Decomposition Theory of Ordered Exponentials*, Proc. Japan Acad. B, **69**, no. 7, 161–166, 1993.
13. H. De Raedt and K. Michielsen, *Computational Methods for Simulating Quantum Computers*, in: Quantum and Molecular Computing, Quantum Simulations, M. Rieth and W. Schommers, (Eds.), vol. 3 of *Handbook of Theoretical and Computational Nanotechnology*, 1–48, American Scientific, Los Angeles, 2006.
14. G. E. Crooks, *Performance of the Quantum Approximate Optimization Algorithm on the Maximum Cut Problem*, arXiv:1811.08419, 2018.
15. F. G. S. L. Brandao, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, *For Fixed Control Parameters the Quantum Approximate Optimization Algorithm's Objective Function Value Concentrates for Typical Instances*, 2018, arXiv:1812.04170.
16. M. Willsch, D. Willsch, F. Jin, H. De Raedt, and K. Michielsen, *Benchmarking the quantum approximate optimization algorithm*, Quantum Inf. Process., **19**, 197, 2020.
17. L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, *Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices*, Phys. Rev. X, **10**, 021067, Jun 2020.
18. P. Vikstål, M. Grönkvist, M. Svensson, M. Andersson, G. Johansson, and G. Ferrini, *Applying the Quantum Approximate Optimization Algorithm to the Tail-Assignment Problem*, Phys. Rev. Applied, **14**, 034009, Sep 2020.
19. M. Grönkvist, *The Tail Assignment Problem*, PhD thesis, Chalmers University of Technology and Göteborg University, 2005.
20. D. Willsch, M. Willsch, C. D. Gonzalez Calaza, F. Jin, H. De Raedt, M. Svensson, and K. Michielsen, *Benchmarking Advantage and D-Wave 2000Q quantum annealers with exact cover problems*, Quantum Inf. Process., **21**, 141, 2022.
21. L. Martins, *Applying Quantum Annealing to the Tail Assignment Problem*, PhD thesis, University of Porto, 2020.
22. M. Svensson, M. Andersson, M. Grönkvist, P. Vikstål, D. Dubhashi, G. Ferrini, and G. Johansson, *A Heuristic Method to solve large-scale Integer Linear Programs by combining Branch-and-Price with a Quantum Algorithm*, arXiv:2103.15433, 2021.
23. R. M. Karp, *Reducibility among Combinatorial Problems*, 85–103, Springer US, Boston, MA, 1972.
24. A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, *A variational eigenvalue solver on a photonic quantum processor*, Nature Communications, **5**, no. 1, 4213, 2014.
25. M. S. Jattana, F. Jin, H. De Raedt, and K. Michielsen, *Assessment of the variational quantum eigensolver: application to the Heisenberg model*, Frontiers in Physics, **10**, May 2022, doi:10.3389/fphy.2022.907160.
26. M. J. D. Powell, "A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation", in: Advances in Optimization and Numerical Analysis, 51–67. Springer Netherlands, Dordrecht, 1994.

27. D. Kraft, *A software package for sequential quadratic programming*, DLR German Aerospace Center – Institute for Flight Mechanics, 1988.
28. M. S. Jattana, F. Jin, H. De Raedt, and K. Michielsen, *Improved variational quantum eigensolver via quasi-dynamical evolution*, 2022, arXiv: 2202.10130.
29. J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, *Barren plateaus in quantum neural network training landscapes*, Nature Communications, **9**, no. 1, 1–6, Dec 2018.

The NIC Research Groups

TopSuite: Collection of Deep Learning-Based Metamethods to Predict and Evaluate Protein Structures and Properties

Filip Koenig[1], Daniel Mulnaes[1], Karel van der Weg[2], and Holger Gohlke[1,2]

[1] Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
E-mail: {filip.koenig, daniel.mulnaes, gohlke}@hhu.de

[2] John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Institute of Biological Information Processing (IBI-7: Structural Biochemistry), and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {k.van.der.weg, h.gohlke}@fz-juelich.de

To understand the function of proteins, knowledge of various properties is necessary. While for many properties experimental determination protocols are available, these may be time-consuming and laborious. An alternative is given by computational predictions, building on recent advances in machine learning technology. Here, we present TopSuite, a collection of deep learning-based predictors. TopSuite contains programs for protein model quality assessment, template-based protein structure prediction, domain boundary prediction, secondary structure and membrane topology prediction, and protein-protein interface contact prediction.

1 Introduction

To develop new pharmaceutical and biotechnological products, it is of utmost importance to understand how proteins as targets or enzymes work. Various experimental procedures are available to determine different properties of proteins, with their three-dimensional structure being the most crucial property to gaining insight into protein function. While the experimental determination of protein structure allows for atomic resolution information, it is expensive, time-consuming, and comes without a guarantee of success in all cases[1]. An alternative to experimental determination is computational protein structure prediction. Having been a major scientific problem in computational biology for decades[2], recent advances in deep learning technology led to the development of tools like AlphaFold[3] and RoseTTAfold[4], that show drastic improvements in the accuracy of predicted protein structures, providing structural models of even complex proteins that can match experimental structures within experimental uncertainty.

Apart from pure structure prediction, a variety of additional properties are relevant for understanding the biological context of proteins, e.g., predicting the borders of protein domains or the interaction and orientation between proteins and membranes, i.e., membrane topology. Similar to structure prediction, these tasks have benefited from deep learning techniques[5,6].

A collection of deep learning-based metamethods to predict and evaluate protein structures and properties has been developed over the past seven years in our group, bundled under the name TopSuite. TopSuite contains software for protein model quality assess-

ment (TopScore), template-based protein structure prediction (TopModel), domain boundary prediction (TopDomain), secondary structure and membrane topology prediction (TopProperty), and, under development, protein-protein interface contact prediction (TopInterface)[7–10]. To make the methods available to the public, web interfaces to all programs were developed[11] [a].

2 Methods

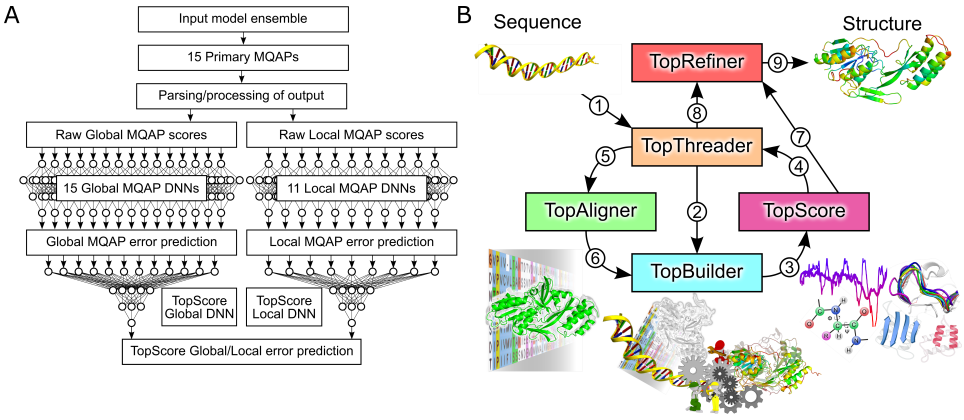


Figure 1. (A) The architecture of TopScore to assess protein model quality. 15 primary predictors are run on the input model ensemble. MQAP stands for Model Quality Assessment Program. The output is processed and fed to individual DNNs for each primary predictor, one set for global and one set for local scores. The prediction output is then used as input for the final DNNs, which predict a global or residue-wise TopScore value. (B) Workflow of TopModel for protein structure prediction. The user provides an input sequence to TopThreader (1). The resulting threader alignments are used by TopBuilder to build models (2), which are subsequently scored with TopScore (3). TopThreader then clusters and ranks the templates to remove false positives (4). TopAligner is used to build ensembles of aligned templates and subsequently consensus alignments (5), which are fed to TopBuilder (6). The models are again scored with TopScore (3). Subsequently, TopThreader takes these models along with primary threader scores and creates a ranking for templates according to the predicted similarity to the native structure (4). After the selection of the templates, TopAligner is used again to build an ensemble of multitemplate and pairwise alignments (5), which TopBuilder builds models from (3). The resulting models are scored with TopScore (4). Models from the multi-template ensemble (7) and single-template models (8) are chosen and used by TopRefiner to create a final model by combination and refinement of good structural parts (9). Figures were taken from Refs. 7 and 8.

All methods in TopSuite are metamethods and have a similar basic structure. In each case, a varying number of primary predictors is run on the user input to produce the raw data for the prediction. These predictions are normalised and set up as input features, potentially combined with other features extracted directly from the input data. One or multiple consecutive stages of deep neural networks (DNN) are then used to predict the respective properties. This kind of architecture is illustrated in Fig. 1A with the DNNs of TopScore as an example. Using multiple DNNs connected in series has the advantage

[a] https://cpclab.uni-duesseldorf.de/topsuite/

that one can present features in different forms to separate stages, allowing the network to learn specific patterns instead of overwhelming it by presenting all available information at once. The final output of the last-stage neural network is processed and, in case our web interfaces are used, is made available as a results webpage.

TopModel has a series of additional workflow modules, as it is not only predicting or classifying properties but builds complete structural models. A scheme of the architecture of TopModel is shown in Fig. 1B. The workflow is split into five separate modules: TopThreader, TopAligner, TopBuilder, TopScore, and TopRefiner. TopThreader uses twelve threading programs in combination with DNNs in a top-down consensus approach to select optimal templates. TopAligner uses eight alignment programs to build an ensemble of pairwise template-template alignments to ensure similar quality and fold of templates for the model building. TopBuilder combines output from Modeller9[12] and Rosetta[13] to build structural models based on alignments with template structures. TopScore is used at various points in the workflow to score models. TopRefiner selects and combines different models to replace regions with low TopScore values.

All programs of TopSuite are programmed with the Python programming language. TopScore and TopModel use a multilayer perceptron (MLP) regression model from SciKit-learn, while TopDomain, TopProperty, and TopInterface use Tensorflow and Keras in combination with ResNet[14–17]. For TopDomain, TopProperty, and TopInterface, it is relevant to encode the local protein neighbourhood to let the DNN capture the local context. Therefore, a sliding window approach is used in which residues surrounding the target residue are represented as the peripheral pixels in an image. This is realised by using convolutional neural networks (CNN), which are well established in the field of image recognition.

3 Results

3.1 TopScore

TopScore predicts a derivative of the lDDT score[18], a superposition-independent target score, called the "lDDT error", which is defined as 1-lDDT. It provides a global score for the model as a whole, as well as a local score for every residue of a model. To provide an evaluation scheme that is free of clustering information, a second metric besides TopScore is computed, TopScoreSingle. This is done because clustering-free methods are better in selecting the best model, especially if the model ensemble is heterogeneous[7]. In comparison, AlphaFold provides an intrinsic model accuracy estimate called pLDDT[3], calculated directly from the final stage of the single representation. While it also uses an MLP to calculate the score, its input features cannot be interpreted as obviously as the input values to TopScore, making a direct comparison difficult.

TopScore was evaluated with different metrics to benchmark its performance in four distinct tasks: The ability to separate good from wrong models, the accuracy of the predicted global lDDT error, the accuracy in ranking model ensembles, and how good it is at finding the best model. Overall, TopScore and TopScoreSingle significantly outperform all primary predictors in all categories. When calculating correlation scores with Pearson's $R_{\mathrm{All}}{}^2$, TopScore achieves a value of 0.93 for global and 0.78 for local error prediction. Fig. 2E illustrates the connection between global TopScore and the true lDDT error on a test system, indicating that the error prediction is more precise on very good and very wrong models.

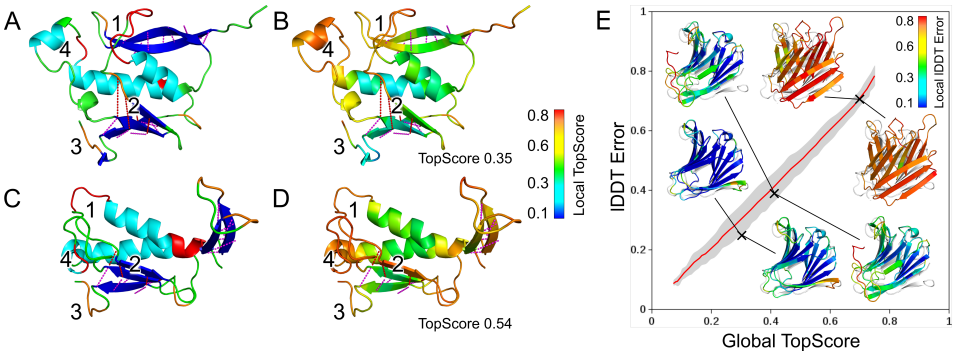


Figure 2. A-D Evaluation of the predicted structure of lipoprotein LipoP from *C. difficile* with experimental results from NMR. (A) Model predicted by TopModel. The model is coloured according to the agreement with NMR results. Blue: Residues of β-strands that agree with the NMR data. Orange: Residues that were found to be in β-strand conformation according to the NMR data but not in the TopModel prediction. Cyan: Residues of α-helices that agree with the NMR data. Red: Residues that were found to be in α-helical conformation according to the NMR data but not in the TopModel prediction. Magenta lines: NOE restraints of β-sheets that agree with the model. Red lines: NOE restraints of β-sheets that disagree with the model. (B) The TopModel model is coloured according to residue-wise TopScore. (C) Model structure predicted by the program dPPAS2 from the LOMETS server. The colour scheme is the same as in panel (A). (D) The model from dPPAS2 is coloured according to residue-wise TopScore. (E) Performance of TopScore. From the 3DRobot dataset[20], three random models from PDB entry 4BMB were selected and coloured according to TopScore (lower triangle) and lDDT error (upper triangle). Figures were taken from Refs. 7 and 8.

3.2 TopModel

TopModel uses five submodules with numerous deep neural networks to build accurate alignments that are used in homology modelling of protein structures. TopModel depends on identifying high-quality templates; therefore, the module TopThreader uses a multi-step DNN-assisted selection procedure to find optimal templates. AlphaFold, on the other hand, can use template information, but such information plays only a minor role in its performance[3]. By contrast, it mainly exploits information extracted from deep sequence alignments. We view the two methods as complementary to each other, with TopModel being a good option for providing high-confidence structural predictions if high-quality templates are available and AlphaFold being a valuable tool in the case no templates are available.

To validate the predictive power of TopModel, nuclear magnetic resonance (NMR) experiments were performed on the lipoprotein LipoP from *C. difficile*. The model from TopModel and the model from the best primary predictor, the dPPSA2 program of LOMETS[19], are compared with experimental results from NMR (Fig. 2A-D). The model from TopModel shows good agreement with the experimental assignments from NMR, with Matthews correlation coefficients (MCC) of 0.81 for β-strands, 0.68 for α-helices, and 0.66 for coils. The good model quality is also corroborated by a TopScore value of 0.35.

Additionally, the performance of TopModel and AlphaFold were compared in terms of TopScore values computed for models generated by either method (Fig. 3). AlphaFold predicts models with a better TopScore than TopModel, in the case of good models, whereas TopModel produces models with a better TopScore in the high-TopScore region.

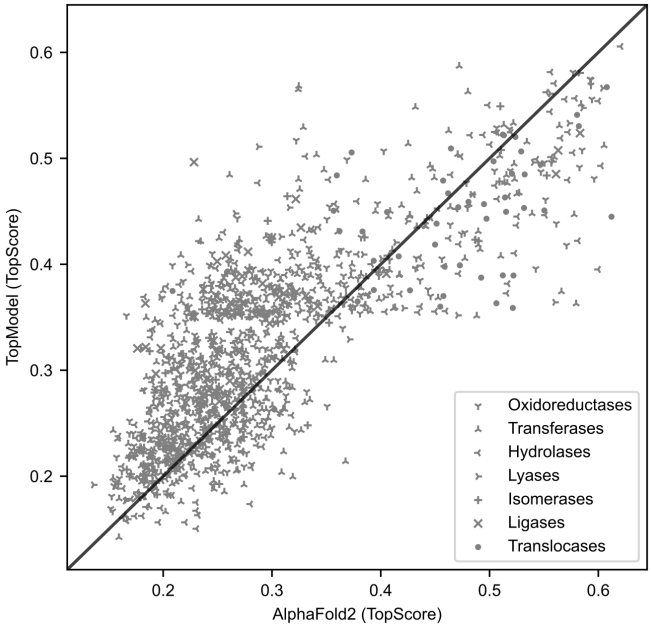


Figure 3. Comparison of TopModel and AlphaFold predictions scored by TopScore each. Lower TopScore values indicate better structural models.

3.3 TopProperty

TopProperty uses eleven primary predictors of secondary structure (SS) and solvent accessibility (SA) and sixteen primary predictors of transmembrane topology (TMT) and membrane exposure (ME) to produce features that are fed to two separate 1D CNNs with four respectively five differently sized sliding windows to perform a meta-prediction on these properties. While it is possible to use predicted structures from AlphaFold to deduce the TMT or ME directly, TopProperty can predict these properties from sequence alone and, therefore, is independent of structure availability.

TopProperty is benchmarked on the NOUMENON dataset[21] for evaluating the performance of SS and SA predictions and on the OPM dataset[22], consisting of transmembrane proteins, for evaluating the performance of SS and SA as well as TMT and ME predictions. TopProperty shows overall superior performance for predicting SS and SA compared to the respective primary predictors when the Q3 metric was used, both for the NOUMENON and the OPM dataset. The Q3 score is the percentage of residues with a correctly predicted 3-state secondary structure[23]. For TMT, TopProperty was benchmarked against methods that specifically predict either transmembrane α-helical bundles (TMH) or transmembrane β-barrels (TMB). TopProperty significantly outperforms all other primary predictors on the Q3 score. For ME, no competitor method provides any predictions for TMHs, therefore, no comparison to another method was possible. For TMBs, only one competitor method provides predictions but outperformed TopProperty on the MCC metric.

3.4 TopDomain

TopDomain applies 53 primary predictors of domain boundaries and feeds the output of these combined with homology- and sequence-based features into two stages of 1D CNNs with varying sliding window sizes. TopDomain contains workflows for sequence- (TopDomain, $\text{TopDomain}_{\text{Seq}}$) as well as structure-based ($\text{TopDomain}_{\text{Parse}}$) domain boundary predictions. Additionally, a binary predictor ($\text{TopDomain}_{\text{TMC}}$) predicts if domain parsing is necessary, e.g., to facilitate subsequent protein structure prediction. TopDomain is benchmarked on two datasets, the manually annotated TopDomain dataset and the CASP dataset based on annotations from the CASP organisers on CASP11-13 proteins. For both datasets, TopDomain, $\text{TopDomain}_{\text{Seq}}$, and $\text{TopDomain}_{\text{Parse}}$ significantly outperform their respective best primary predictor. TopDomain achieves an F1-score of 73.8 % for domain boundary prediction within ± 10 residues for the TopDomain dataset and a value of 42.8 % for the CASP dataset. Although, in general, tools such as AlphaFold and RoseTTAfold can predict structures for full-length sequences, often no fold can be predicted for parts of the structure. In these cases, it can be fruitful to split the respective sequence into separate domains and try to predict structures for the domains separately.

3.5 TopInterface

TopInterface uses four primary predictors of coevolutionary couplings together with structural, energetic, and sequence-based features as input to three stages of CNNs, leveraging 2D sliding windows to represent the local neighbourhood context. In comparison to AlphaFold, it uses additional ways of building paired multiple sequence alignments that are necessary for the calculation of coevolutionary couplings of protein pairs. Especially, TopInterface has a dedicated workflow to build paired alignments for protein pairs coming from two different species, allowing the modelling of inter-species protein-protein interactions (PPI). TopInterface provides a distance prediction together with an upper and

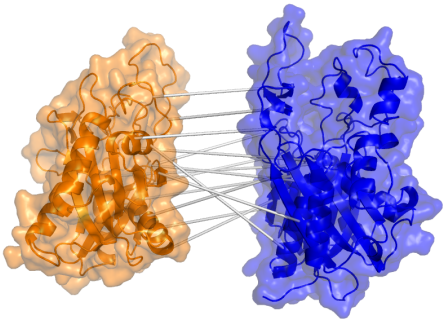


Figure 4. TopInterface distance restraints. TopInterface takes two protein structures as input, extracts various structural, energetic, and sequence-based features, and performs three consecutive stages of deep learning to predict distance restraints with upper and lower bounds for every inter-protein residue-residue pair. These restraints can be used in a docking protocol to create a structural model of the protein-protein complex. The distance restraints are illustrated as grey lines between proteins of PDB ID 5DWZ.

lower bound for every inter-protein residue-residue pair, which can be used in a distance-restrained docking, as presented in Fig. 4. It has workflows for homo- and heterodimers, as well as domain-domain interactions and inter-species PPIs.

4 Conclusion

TopSuite, a collection of metamethods for predicting protein structures and properties, was developed in our group. It consists of the programs TopScore, TopModel, TopDomain, TopProperty, and TopInterface. All programs utilise DNNs to perform their respective tasks, using carefully chosen sets of features and DNN architectures. As new protein structure prediction tools such as AlphaFold and RoseTTAfold have been developed, the accuracy of template-free structure predictions has reached new levels. While the TopSuite programs cannot provide predictions with similar quality without the use of templates, we view TopSuite as a complementary approach for properties that cannot be directly derived from predicted models from AlphaFold or RoseTTAfold, e.g., domain boundary predictions in the case of failing to build a multidomain protein directly from an end-to-end approach. Therefore, we expect TopSuite to be a valuable tool in providing such property predictions. In future developments, we seek to combine the power of sophisticated alignment pairing workflows in TopInterface with AlphaFold for building protein complexes of higher stochiometry.

Acknowledgements

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 267205415 - SFB 1208 (project A03 to HG), and by the Bundesministerium für Bildung und Forschung (BMBF) - Förderkennzeichen 031L0182 - InCelluloProtStruct to HG, and by the Helmholtz Association of German Research Centres (Helmholtz School for Data Science in Life, Earth and Energy (HDS-LEE) to HG). We are grateful for computational support and infrastructure provided by the "Zentrum für Informations- und Medientechnologie" (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) to HG on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user ID: HKF7, VSK33). Figs. 1 and 2 reprinted (adapted) with permission from Refs. 7 and 8; Copyright (2018, 2020), American Chemical Society.

References

1. H. Deng, Y. Jia, and Y. Zhang, *Protein structure prediction*, Int. J. Mod. Phys. B **32**, 18, 2018.
2. D. Petrey and B. Honig, *Protein structure prediction: inroads to biology*, Mol. Cell **20(6)**, 811-819, 2005.
3. J. Jumper et al., *Highly accurate protein structure prediction with AlphaFold*, Nature **596(7873)**, 583-589, 2021.
4. M. Baek et al., *Accurate prediction of protein structures and interactions using a three-track neural network*, Science **373(6557)**, 871-876, 2021.

5. Q. Shi et al., *DNN-Dom: predicting protein domain boundary from sequence alone by deep neural network*, Bioinformatics **35(24)**, 5128-5136, 2019.
6. Z. Liu, Y. Gong, Y. Bao, Y. Guo, H. Wang, and G.N. Lin, *TMPSS: A Deep Learning-Based Predictor for Secondary Structure and Topology Structure Prediction of Alpha-Helical Transmembrane Proteins*, Front. Bioeng. Biotechnol. **8**, 629937, 2020.
7. D. Mulnaes and H. Gohlke, *TopScore: Using Deep Neural Networks and Large Diverse Data Sets for Accurate Protein Model Quality Assessment*, J. Chem. Theory Comput. **14(11)**, 6117-6126, 2018.
8. D. Mulnaes, et al., *TopModel: template-based protein structure prediction at low sequence identity using top-down consensus and deep neural networks*, J. Chem. Theory Comput. **16(3)**, 1953-1967, 2020.
9. D. Mulnaes, P. Golchin, F. Koenig, and H. Gohlke, *TopDomain: Exhaustive Protein Domain Boundary Metaprediction Combining Multisource Information and Deep Learning*, J. Chem. Theory Comput. **17(7)**, 4599-4613, 2021.
10. D. Mulnaes, S. Schott-Verdugo, F. Koenig, and H. Gohlke, *TopProperty: Robust Metaprediction of Transmembrane and Globular Protein Features Using Deep Neural Networks*, J. Chem. Theory Comput. **17(11)**, 7281-7289, 2021.
11. D. Mulnaes, F. Koenig, and H. Gohlke, *TopSuite Web Server: A Meta-Suite for Deep-Learning-Based Protein Structure and Quality Prediction*, J. Chem. Inf. Model. **61(2)**, 548-553, 2021.
12. N. Eswar et al., *Comparative protein structure modeling using MODELLER*, Curr. Protoc. Bioinform. **15(1)**, 5-6, 2006.
13. C. A. Rohl et al., *Protein structure prediction using Rosetta*, Meth. Enzymol. **383**, 66-93, 2004.
14. F. Pedregosa et al., *Scikit-learn: Machine learning in Python*, J. Mach. Learn. Res. **12**, 2825-2830, 2011.
15. M. Abadi et al., *Tensorflow: A system for large-scale machine learning*, 12th USENIX Symposium on operating Systems Design and Implementation **OSDI 16**, 265-283, 2016.
16. F. Chollet, *Keras*, 2015.
17. K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 770-778, 2016.
18. V. Mariani, M. Biasini, A Barbato, and T Schwede, *lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests*, Bioinformatics **29(21)**, 2722-2728, 2013.
19. S. Wu and Y. Zhang., *LOMETS: a local meta-threading-server for protein structure prediction*, Nucleic Acids Res. **35(10)**, 3375-3382, 2007.
20. H. Deng, Y. Jia, and Y. Zhang, *3DRobot: automated generation of diverse and well-packed protein structure decoys*, Bioinformatics **32(3)**, 378-387, 2016.
21. G. Orlando, D. Raimondi, and W. F. Vranken, *Observation selection bias in contact prediction and its implications for structural bioinformatics*, Sci. Rep. **6**, 36679, 2016.
22. M. A. Lomize et al., *OPM: orientations of proteins in membranes database*, Bioinformatics **22(5)**, 623-625, 2006.
23. M. Spencer, J. Eickholt, and C. Jianlin, *A Deep Learning Network Approach to ab initio Protein Secondary Structure Prediction*, IEEE/ACM Trans. Comput. Biol. Bioinform. **12(1)**, 103-112, 2014.

The Columbia Plot of the QCD Phase Transition

Francesca Cuteri[1], Alfredo D'Ambrosio[1,2], Reinhold Kaiser[1,2], 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, sciarra}@itp.uni-frankfurt.de

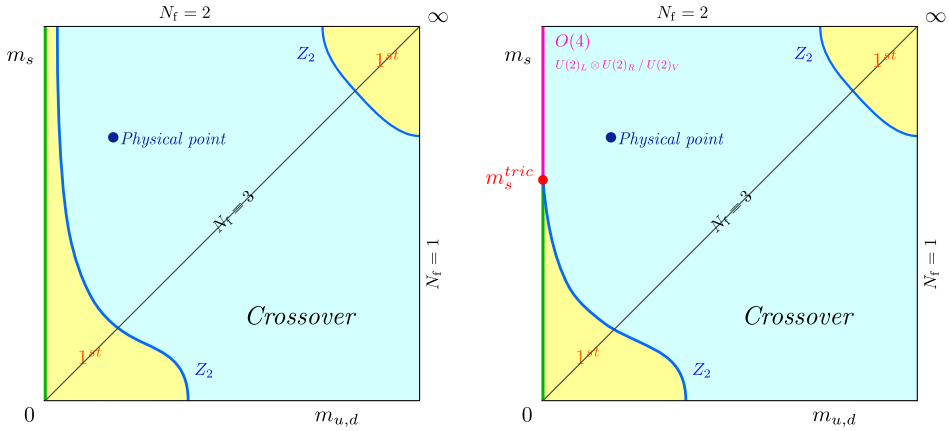
[2] John von Neumann Institute for Computing (NIC), GSI, Planckstr. 1, 64291 Darmstadt, Germany

Quantum Chromodynamics (QCD) is the fundamental theory of strongly interacting matter composed of quarks and gluons. The QCD phase diagram specifies in which form strongly interacting matter exists as a function of temperature and baryon density. Because of a severe sign problem, finite densities are not amenable to Monte Carlo simulations. It is then important to constrain the phase diagram by collecting information about parameter regions away from the physical point, where non-perturbative information is accessible. A particularly important role is played by the chiral limit of vanishing light quark masses, which cannot be simulated directly due to a singular fermion determinant. We report significant progress from a long term project to determine the nature of the thermal QCD transition as a function of the quark masses, the number of flavours and the lattice spacing. In particular, we were able to analyse the lattice chiral limit, and find QCD to display a second order chiral transition for $N_f = 2 - 6$ species of massless quarks in the continuum limit.

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 bound hadron states, among them the familiar pions, kaons, protons and neutrons. On the other hand, at large temperatures or densities, the average energy per particle is higher and the system 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. 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 lattice QCD, a reformulation of the theory on a discrete space-time grid.

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 has been known for a while 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]. Away from the physical point, the order of the finite



(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 transition as a function of quark masses.

temperature transition at zero density depends on the quark masses. Two possibilities, first predicted nearly 40 years ago by the epsilon expansion applied to linear sigma models with symmetry breaking patterns as QCD[2], are schematically shown in Fig. 1. The theories with $N_f = 1, 2, 3$ mass degenerate quark flavours correspond to the right and top boundary lines, and the diagonal, respectively. In the limits of zero u, d quark masses or infinite quark masses (left boundary and upper right corner), there are non-analytic phase transitions due to the spontaneous breaking of exact global symmetries, the chiral symmetry for massless quarks and the centre symmetry of the QCD gauge group for static quarks. In the pure gauge limit (upper right corner), the transition was established to be of first order in the continuum[3]. The chiral transition is much harder, as the massless limit cannot be simulated directly. Simulations at finite quark masses on coarse lattices appeared to confirm the predicted[2] first-order region for $N_f = 3$[4,5]. On the other hand, one observes an analytic crossover at intermediate quark masses, with second order boundary lines separating these regions, called chiral and deconfinement critical lines, respectively. Both lines have been shown to belong to the $Z(2)$ universality class of the 3d Ising model[6–9]. Over the years the chiral critical line was found to strongly recede with decreasing lattice spacing[10], which is parametrised by increasing temporal lattice extent N_τ (see below). It has remained open whether the chiral phase transition for two quark flavours is of first or second order. Clarifying this question is also important for real physics: the smallness of the u, d-quark masses in nature may put them within the scaling region of the critical line, so that the physical transition may show remnants of criticality. Over the last few years, by simulating an extended parameter space, our group has managed to obtain strong evidence for the continuum Columbia plot, as we will describe now.

2 Computational Strategy

The central object in statistical physics is the partition function $\mathcal{Z}$ of the system of interest. For lattice QCD it is expressed as a path integral over the gluon fields U, including a determinant of the Dirac operator for each species of quarks, which we always take to be degenerate. An expectation value of some observable $\mathcal{O}$ then reads

$$\langle \mathcal{O} \rangle = \mathcal{Z}^{-1} \int \mathcal{D}U \; O \; (\det D[U; am])^{N_f} \exp\{-S_g[U]\} \; , \tag{1}$$

where $S_g[U]$ is the discretised gauge action and $D[U]$ the fermion Dirac operator, which we employ in the standard unimproved staggered or Wilson discretisations (for detailed expressions see Ref. 11). The bare microscopic parameters of the partition function are the lattice gauge coupling $\beta = 6/g^2$, and the bare quark mass am, given in units of the lattice spacing a. The temperature T is specified by the inverse temporal lattice extent

$$T = (a(\beta) N_\tau)^{-1} \; . \tag{2}$$

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

For our numerical investigation, we use the standard unimproved Wilson gauge and staggered fermion actions. All numerical simulations have been performed using the publicly available OpenCL-based code CL^{2}QCD, which is optimised to run efficiently on AMD GPUs and contains an implementation of the RHMC algorithm for unimproved rooted staggered fermions. Version `v1.0`[12] has been employed for simulations on smaller N_τ on the L-CSC supercomputer at GSI, while version `v1.1`[13] has been run on the HLR supercomputer at Goethe University to run the most costly simulations. The thousands of necessary simulations were efficiently handled by the `BaHaMAS` software[14].

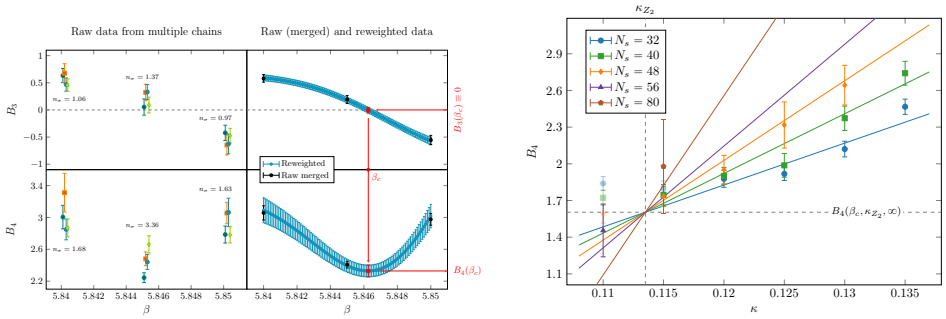
Phase transitions can be located and their nature be identified by finite size scaling analyses of standardised moments of the distribution of an appropriate order parameter,

$$B_n(\beta, m, N_\sigma) = \frac{\langle (\mathcal{O} - \langle \mathcal{O} \rangle)^n \rangle}{\left\langle (\mathcal{O} - \langle \mathcal{O} \rangle)^2 \right\rangle^{n/2}} . \tag{3}$$

For the deconfinement phase transition with heavy masses, the order parameter is the Polyakov loop, $\mathcal{O} = L$, for the chiral phase transition it is the chiral condensate, $\mathcal{O} = \overline{\psi}\psi$. We first locate the phase boundary β_c by the condition of vanishing skewness for the distribution of the observable $B_3(\beta_c, am, N_f, N_\tau, N_s) = 0$. The order of the transition can be determined by the kurtosis $B_4(\beta, m)$[16]. In the thermodynamic limit $N_\sigma \to \infty$, the kurtosis $B_4(\beta_c, m, N_s)$ takes the values of 1 for a first order transition and 3 for an analytic crossover, respectively, with a discontinuity when passing from a first order region to a crossover region via a second order point, where it takes the value[17] 1.604 for the 3D Ising universality class. On finite but sufficiently large volumes, it can be expanded about the critical point,

$$B_4(\beta_c, am, N_f, N_\tau, N_s) = 1.604 + B\,(\beta_c, N_f, N_\tau)\,(am - am_c) N_s^{1/\nu} + \dots \; , \tag{4}$$

through which it passes smoothly. As the volume is increased, the rate of the approach to the thermodynamic limit is governed by a 3D Ising critical exponent, $\nu = 0.6301$. Dots indicate additional terms that vanish in the infinite volume limit.



(a) Left: examples of B_3 for heavy quarks on a 6×36^3 lattice, obtained from different Markov chains, n_σ denotes the maximal difference between them in standard deviations. Merged raw and reweighted data for B_3 (top) and B_4 (bottom) are also shown, with the determination of β_c and $B_4(\beta_c)$ in red.

(b) The kurtosis on the phase boundary, $B_4(\beta_c)$, evaluated for different quark masses , parametrised by κ, on $N_\tau = 8$ lattices. Lines represent a common fit of all displayed volumes to the scaling formula Eq. 4.

Figure 2. The general procedure of our finite size scaling analysis, shown at the example of a deconfinement critical point in the heavy quark region[15].

For each parameter combination, we generally simulated four independent Monte Carlo chains until their $B_{3,4}$-values agreed to within three standard deviations or better, upon which they were merged. In order to tune precisely to the phase boundary, the multi-histogram method was used to interpolate between simulated β-values[18]. This is shown in Fig. 2(a). An example of a fit of our data to Eq. 4 is shown in Fig. 2(b). Altogether, the results presented in the following sections are based on $\sim$ 200 million Monte Carlo trajectories spread over $\sim$ 700 different parameter combinations, obtained over a span of several years.

3 The Heavy Mass Corner

Even though it is far from the physical point, an investigation of the heavy mass corner is very well motivated. Firstly, at the physical point there is still very little understanding about the interplay of chiral symmetry restoration and deconfinement, which are related to two different symmetries that are both explicitly broken. By contrast, the heavy mass corner has no remnant of chiral symmetry, and the deconfinement aspects of the thermal transition can be studied separately. Secondly, the fermion determinant has no singularities there, so the situation is fully controlled and can serve as benchmark for non-perturbative analytic approaches or effective theories of QCD.

Unfortunately, such simulations are also extremely compute intensive, because *i)* it turns out that larger physical volumes are required than for light quarks to approach the thermodynamic limit, and *ii)* extremely fine lattices are necessary for the Compton wave lengths of the heavy quarks to remain larger than the lattice spacing a.

Following the computational steps outlined in Sec. 2, we used $N_f = 2$ Wilson fermions to determine the critical hopping parameter κ (related to quark mass), representing the

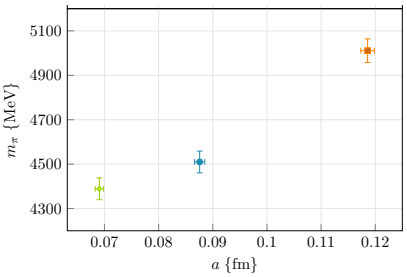


Figure 3. Pion masses evaluated on the $Z(2)$-endpoint of the $N_f = 2$ deconfinement transition for different lattice spacings with Wilson fermions[15].

β_c	am_π	a {fm}	m_π {GeV}	T_c {MeV}
5.9821	1.72039(7)	0.0888(10)	3.82(4)	278(3)
6.0129	1.98121(7)	0.0872(9)	4.48(5)	283(3)

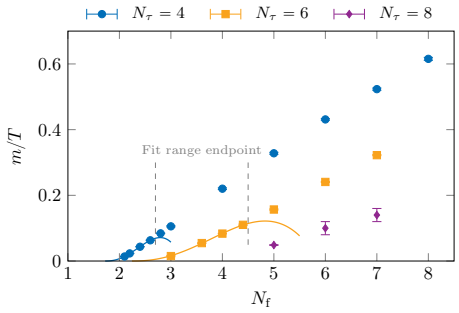
Table 1. Critical coupling, temperature and pion masses for the $Z(2)$ deconfinement critical point from staggered fermions[19].

$Z(2)$ boundary of the first-order transition in the heavy mass corner, for three different lattice spacings parametrised by $N_\tau = 6, 8, 10$. These values were converted into pseudo-scalar meson masses, which are shown in Fig. 3. Strong cutoff effects are still visible, and in particular we have $am_{PS} = am_\pi > 1$ which means that condition *ii)* above is not yet satisfied. To get a feeling for the size of the cutoff effects, we started another set of simulations with $N_f = 2$ staggered fermions, where so far only one lattice spacing corresponding to $N_\tau = 8$, is available[19]. The critical pseudo-scalar mass is bracketed by the values given in the Tab. 1. Once again $am_{PS} = am_\pi > 1$, but we see that the critical pseudo-scalar masses between the Wilson and staggered calculations are compatible at the $\sim 15\%$-level, which one may take as a rough estimate for the remaining systematic error. Additional simulations are running for $N_\tau = 12$ Wilson and $N_\tau = 10$ staggered, in order to make the continuum estimates more precise.

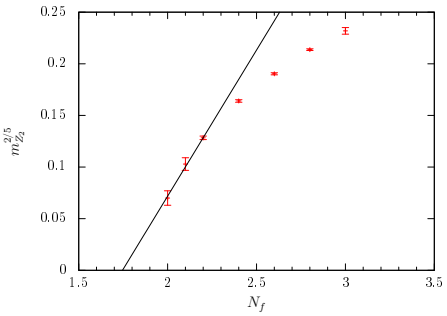
4 The Chiral Transition

Since the chiral limit invariably requires extrapolations, we have developed a novel way of analysing the lattice parameter space that significantly reduces the associated systematic errors. In particular, when working with degenerate fermions only, as in Eq. 1, one may also simulate this expression for non-integer values of N_f[20]. In this generalised parameter space, a change of a first-order chiral transition to second order one, as in Fig. 1(b), requires a tricritical value $N_f^{\rm tric}$. Since we see first-order chiral transitions on coarse lattices and there cannot be a chiral transition for $N_f = 1$, such tricritical points are guaranteed to exist in the lattice massless limit. Moreover, the functional shape, by which a critical boundary line of a first-order transition enters such a tricritical point, is dictated by tricritical exponents of the scaling field, which in our case is the quark mass, $(am)^{2/5}$.

This is shown in Fig. 4, which shows our results for the chiral critical line separating regions of first-order transitions (below) from crossover (above). In Fig. 4(a) one observes the first-order transition getting stronger with N_f and with increasing lattice spacing. The cutoff effects are growing with N_f, for $N_f = 7$ the critical bare quark mass is reduced by a stunning factor of five when the lattice spacing is reduced (N_τ is increased) by a factor of two! No continuum critical mass m_c is discernible yet for any N_f. On the



(a) Chiral critical line for three different lattice spacings[21].



(b) Chiral critical line for $N_\tau = 4$ only, with masses rescaled to represent the scaling field[20].

Figure 4. The chiral critical line for unimproved staggered fermions. Every point of the plane represents a phase boundary. Regions above the $Z(2)$-lines represent crossover, those below first-order transitions. Fits on the left show next-to-leading order tricritical scaling, on the right leading-order tricritical scaling.

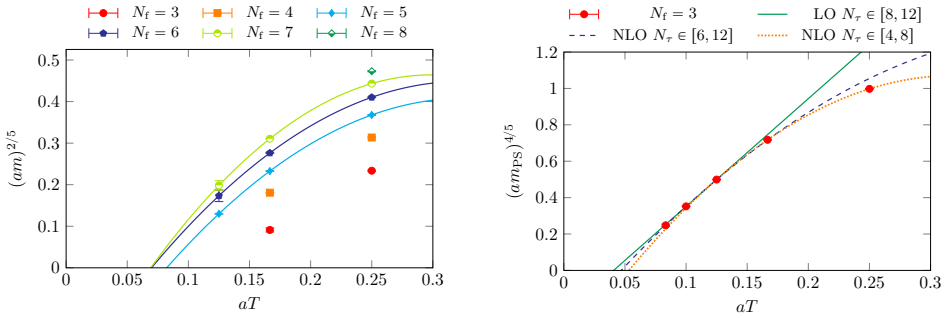
other hand, the intercepts of the critical lines at $m = 0$ constitute the sought-after tricritical line $N_f^{\mathrm{tric}}(N_\tau)$. Tricritical scaling can be appreciated in the rescaled Fig. 4(b), where the data approach a leading order scaling relation[20]. This allows for an extrapolation

$$am_c\big(N_f(N_\tau), N_\tau\big) = D(N_\tau)\big(N_f - N_f^{\mathrm{tric}}(N_\tau)\big)^{5/2} + \dots \,. \tag{5}$$

The scaling relation has been inverted here, because in simulations N_f is fixed whereas the critical quark mass is computed with statistical errors. Next-to-leading order scaling fits in the left figure predict $N_f^{\mathrm{tric}}(N_\tau = 4) \approx 1.71(3)$ and $N_f^{\mathrm{tric}}(N_\tau = 6) \approx 2.20(8)$. While no extrapolation for $N_\tau = 8$ is available yet, the critical line flattens and shifts to the right, implying a line $N_f^{\mathrm{tric}}(N_\tau)$. One concludes that, for unimproved staggered fermions, the $N_f = 2$ massless theory shows a first-order transition on $N_\tau = 4$, but a second-order transition on all finer lattices and in the continuum. The question now is what happens to the $N_f \geq 3$ theories.

This can be seen in a different variable pairing in Fig. 5(a), where the same critical bare quark masses are rescaled by the tricritical exponent, and plotted as a function of N_τ for different fixed N_f-values. Only a slight curvature is exhibited by those N_f-lines with three data points, which thus are compatible with next-to-leading order scaling and a tricritical point at some finite $N_\tau^{\mathrm{tric}}(N_f)$, obtained by the corresponding extrapolations. To interpret these results, we recall that the continuum limit for any theory, $a \to 0$, is located in the lower left corner of Fig. 5(a). Hence, the existence of a finite $N_\tau^{\mathrm{tric}}(N_f)$ implies that the first-order transitions seen at some finite lattice spacings are *not* continuously connected to the continuum limit, and therefore constitute a lattice artefact. Consequently, in all these cases the continuum chiral phase transition must be of second order.

The result we obtained is in plain contradiction of the first-order chiral transitions analytically predicted for $N_f \geq 3$[2]. It is therefore crucial to check it by a different lattice discretisation scheme. Here we use already published data for the chiral critical line of $N_f = 3$ $O(a)$-improved Wilson fermions, specified by the pseudo-critical meson mass m_{PS}[22]. To relate it to the scaling variable, we exploit the proportionality of its square



(a) Unimproved staggered fermions[21].

(b) $O(a)$-improved Wilson fermions[22].

Figure 5. The chiral critical surface projected onto the $(am, aT = N_\tau^{-1})$ plane for two different lattice discretisations, with tricritical scaling fits to both[21].

to the quark mass, known from chiral perturbation theory, $m_{PS}^2 \propto m$. The published data[22] are replotted in Fig. 5(b), with the vertical axis rescaled. Perfect tricritical scaling is observed, with extrapolations that are robust under variation between leading and next-to-leading order, as well as when using only the coarsest lattices $N_\tau = 4, 6, 8$. Thus, $O(a)$-improved Wilson fermions lead to the same conclusion for $N_f = 3$, namely that the chiral phase transition for $N_f = 3$ in the continuum is of second-order.

In order to get a complete overview of the lattice phase diagram, we are also investigating how Fig. 5 continues to the right, i.e. to *larger* lattice spacings. First results indicate that the chiral critical lines bend down also in this direction[23], so that the order of the chiral phase transition changes from second order on very coarse lattices, via first order on in between lattice spacings back to second order in the continuum. These investigations illustrate the importance of proper continuum extrapolations when addressing the order of the finite temperature phase transition, which can be qualitatively different when considered at finite lattice spacing.

5 Conclusions

After several years of lattice simulations in a large parameter space, combined with a new approach to exploit tricritical scaling for the purpose of chiral extrapolations, we have managed to obtain convincing evidence for the qualitative form of the Columbia plot in the continuum, as shown in Fig. 6. The chiral transition is of second order for all values of the strange quark masses and disappears immediately, when a finite u, d-quark mass is switched on. This is in contrast to the deconfinement transition for heavy quarks, which is of first order and ends in a $Z(2)$-critical line at some finite quark mass configurations.

Note that our analysis has nothing to say about the universality class of the chiral transition. Since the chiral symmetry is larger for $N_f = 3$ than for $N_f = 2$, one would expect the universality class and its associated critical exponents to change between the upper and lower left corners of the Columbia plot. A further open question concerns the size of the scaling regime around the chiral critical line. This is important to judge whether the

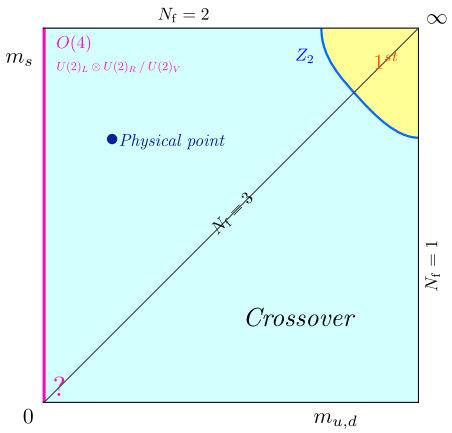


Figure 6. The Columbia plot according to the simulation results presented in this article[21].

physical point is sensitive to scaling from the chiral limit, as indicated in the introduction. Thus, there are still many simulations left to do until the Columbia plot is settled fully quantitatively.

Acknowledgements

The authors acknowledge support by the Deutsche Forschungsgemeinschaft (DFG) through the grant CRC-TR 211 “Strong-interaction matter under extreme conditions”. F.C. and O.P. in addition acknowledge support by the State of Hesse within the Research Cluster ELEMENTS (Project ID 500/10.006).

References

1. Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabo, *The Order of the quantum chromodynamics transition predicted by the standard model of particle physics*, Nature, **443**, 675–678, 2006.
2. R. D. Pisarski and F. Wilczek, *Remarks on the Chiral Phase Transition in Chromodynamics*, Phys. Rev., **D29**, 338–341, 1984.
3. G. Boyd, J. Engels, F. Karsch, E. Laermann, C. Legeland, M. Lutgemeier, and B. Petersson, *Thermodynamics of SU(3) lattice gauge theory*, Nucl. Phys. B, **469**, 419–444, 1996.
4. F. R. Brown, F. P. Butler, H. Chen, N. H. Christ, Z. Dong, W. Schaffer, L. I. Unger, and A. Vaccarino, *On the existence of a phase transition for QCD with three light quarks*, Phys. Rev. Lett., **65**, 2491–2494, 1990.
5. Y. Iwasaki, K. Kanaya, S. Kaya, S. Sakai, and T. Yoshie, *Finite temperature transitions in lattice QCD with Wilson quarks: Chiral transitions and the influence of the strange quark*, Phys. Rev. D, **54**, 7010–7031, 1996.

6. F. Karsch, E. Laermann, and C. Schmidt, *The Chiral critical point in three-flavor QCD*, Phys. Lett., **B520**, 41–49, 2001.
7. P. de Forcrand and O. Philipsen, *The QCD phase diagram for three degenerate flavors and small baryon density*, Nucl. Phys., **B673**, 170–186, 2003.
8. S. Kim, P. de Forcrand, S. Kratochvila, and T. Takaishi, *The 3-state Potts model as a heavy quark finite density laboratory*, PoS, **LAT2005**, 166, 2006.
9. A. Kiyohara, M. Kitazawa, S. Ejiri, and K. Kanaya, *Finite-size scaling around the critical point in the heavy quark region of QCD*, Phys. Rev. D, **104**, no. 11, 114509, 2021.
10. O. Philipsen, *Constraining the phase diagram of QCD at finite temperature and density*, PoS, **LATTICE2019**, 273, 2019.
11. C. Gattringer and C. B. Lang, *Quantum chromodynamics on the lattice*, vol. 788, Springer, Berlin, 2010.
12. C. Pinke, M. Bach, A. Sciarra, F. Cuteri, L. Zeidlewicz, C. Schäfer, T. Breitenfelder, C. Czaban, S. Lottini, and P. F. Depta, "`CL`^{2}`QCD - v1.0`", Sept. 2018.
13. A. Sciarra, C. Pinke, M. Bach, F. Cuteri, L. Zeidlewicz, C. Schäfer, T. Breitenfelder, C. Czaban, S. Lottini, and P. F. Depta, "`CL`^{2}`QCD - v1.1`", Feb. 2021.
14. A. Sciarra, "`BaHaMAS`", Feb. 2021.
15. F. Cuteri, O. Philipsen, A. Schön, and A. Sciarra, *Deconfinement critical point of lattice QCD with N_f=2 Wilson fermions*, Phys. Rev. D, **103**, no. 1, 014513, 2021.
16. K. Binder, *Finite size scaling analysis of Ising model block distribution functions*, Z. Phys., **B43**, 119–140, 1981.
17. A. Pelissetto and E. Vicari, *Critical phenomena and renormalization group theory*, Phys. Rept., **368**, 549–727, 2002.
18. A. M. Ferrenberg and R. H. Swendsen, *Optimized Monte Carlo analysis*, Phys. Rev. Lett., **63**, 1195–1198, 1989.
19. R. Kaiser, O. Philipsen, and A. Sciarra, *The QCD Deconfinement Critical Point for $N_f = 2$ Flavors of Staggered Fermions*, in: 38th International Symposium on Lattice Field Theory, 12 2021.
20. F. Cuteri, O. Philipsen, and A. Sciarra, *QCD chiral phase transition from noninteger numbers of flavors*, Phys. Rev. D, **97**, no. 11, 114511, 2018.
21. F. Cuteri, O. Philipsen, and A. Sciarra, *On the order of the QCD chiral phase transition for different numbers of quark flavours*, JHEP, **11**, 141, 2021.
22. Y. Kuramashi, Y. Nakamura, H. Ohno, and S. Takeda, *Nature of the phase transition for finite temperature $N_{\rm f} = 3$ QCD with nonperturbatively O(a) improved Wilson fermions at $N_{\rm t} = 12$*, Phys. Rev. D, **101**, no. 5, 054509, 2020.
23. F. Cuteri, A. D'Ambrosio, O. Philipsen, and A. Sciarra, *The chiral phase transition from strong to weak coupling*, in: 38th International Symposium on Lattice Field Theory, 12 2021.

Native-Contact Guided Replica Exchange Simulations of Structured RNA

Arthur Voronin[1] **and Alexander Schug**[2,3]

[1] Steinbuch Centre for Computing, Karlsruhe Institute of Technology,
76344 Eggenstein-Leopoldshafen,Germany
E-mail: arthur.voronin@kit.edu

[2] John von Neumann Institute for Computing, Jülich Supercomputing Centre,
Forschungszentrum Jülich, 52428 Jülich, Germany
E-mail: al.schug@fz-juelich.de

[3] Faculty of Biology, University of Duisburg-Essen, 45117 Essen, Germany

Proteins and other biomolecules such as RNA and their interactions determine life at the molecular level. Similar to proteins, RNA plays central roles in living organisms including genetic regulation or information transfer. RNA structure and function are closely coupled yet structurally resolving RNAs is experimentally challenging. *In silico* structure prediction techniques can complement experimental efforts to gain structural information. Here, we explore how replica-exchange molecular dynamics (REX) enhanced with contact data can be used to predict RNA tertiary structure. Such contact data can be obtained by statistical inference of spatial adjacency in the exponentially growing sequence databases via methods such as direct coupling analysis. Our computational approach can thus complement experimental efforts and predict biomolecular structure for a class of biomolecules, which is structurally less explored than proteins. We test this approach on three different structured RNA. We find that contact-guided REX is capable of improving models of a RNA structure and can deliver physically reasonable 3D structures.

1 Introduction

The incredible advances in sequencing techniques have lead to an exponential growth of RNA sequence data[1]. The development of improved statistical methods and alignment software allows building multiple sequence alignment data (MSA) based on inferring structural or even functional relations within and between phylogenetic trees[2–8]. Freely accessible databases such as UniProt, the protein family (Pfam) and RNA family databases (Rfam) offer a treasure trove of sequence information on both proteins and ribonucleic acids (RNA)[9–11]. In the last decade, direct coupling analysis (DCA) and similar methods have distilled this information to predict spatial adjacency based on co-evolution by statically identifying linked mutational patterns[12–18]. Similar to using experimental information as restraints[19–21], contact information can be used as structural restraints to build biomolecular models of, e.g., protein complexes[12] or conformational transitions[22] even on a large scale[23] or to redesign protein signalling[24]. More recently, data-driven machine-learning (ML) methods have shown to astonishingly enhance protein 3D structure prediction[25–27]. These ML approaches, however, rely on deep neural networks with an accordingly massive number of free parameters. Naturally, training these deep networks require equally very large datasets of structurally known 3D structures and accompanying MSAs. As only a limited number of available experimentally resolved structures exist for RNA, these methods cannot be directly transferred to them[28,29]. It is, however, possible to improve RNA

contact prediction[16,18] by shallow learning approaches[30] with limited amounts of parameters.

In the current study, we tested to what extend molecular dynamics can structurally predict RNA or stabilise its native state. We will compare approaches based on regular molecular dynamics (MD), replica exchange (REX) and contact enhanced molecular dynamics (REX bias). REX is a parallel implementation of MD by running multiple non-interacting copies of the system in parallel at different temperatures[31–33]. Switching between different temperatures allows overcoming entrapment in local minima. As REX maintains thermodynamically correct ensembles during the simulation, such simulations offer not only a structure prediction method but also the possibility to derive, e.g., partitions sums. As partition sums allow to derive basically arbitrary thermodynamic properties of a system, such simulations can also provide insight into the physical properties that drive structure adoption. The conducted study is thus similar to our prior work on proteins, where we compared MD vs. REX and carefully investigated optimal weight and bias potentials[34a]. We additionally compare our results to another prior study[30], where we used SimRNA[35] for RNA structure prediction. SimRNA is based on a Monte-Carlo approach for the sampling of the RNA conformational space and a statistical potential to evaluate the energy of the configuration.

2 Results

Method	Backbone RMSD (Å)		
	3d2g	4tzx	4yaz
MD ref	7.6	3.8	20.4
REX ref	7.4	3.6	20.3
REX biased†	5.0	4.0	15.8
SimRNA	15.7	7.5	16.1
SimRNA + mfDCA†	8.5	3.0	20.1
SimRNA + CoCoNet†	16.2	7.8	14.0

Table 1. **Lowest observed RMSD values of performed RNA MD simulations.** Listed are the method, PDB ids and backbone RMSDs with respect to the native fold. Values are obtained via own simulations (upper half) or other methods (lower half, obtained from SI of Ref. 30). Best cases are highlighted for each RNA target. †: Cases with L bias contacts (L: sequence length).

The tested RNA targets were the TPP riboswitch (PDB id: 3d2g[36]), Adenine riboswitch (PDB id: 4tzx[37]), and the 3',3'-cGAMP riboswitch (PDB id: 4yaz[38]). To mimic DCA-based contacts, we added L native contacts (L is the sequence length) into the biased simulations. These contacts were randomly chosen from the native contact map, as seen in Fig. 4. While "real" DCA data would also include erroneous contact data, using only native contacts serves as a first baseline to test the possible improvement. In Tab. 1 we

[a]It should be noted that biasing potentials can be corrected after simulation to regain unbiased thermodynamic equilibrium/properties.

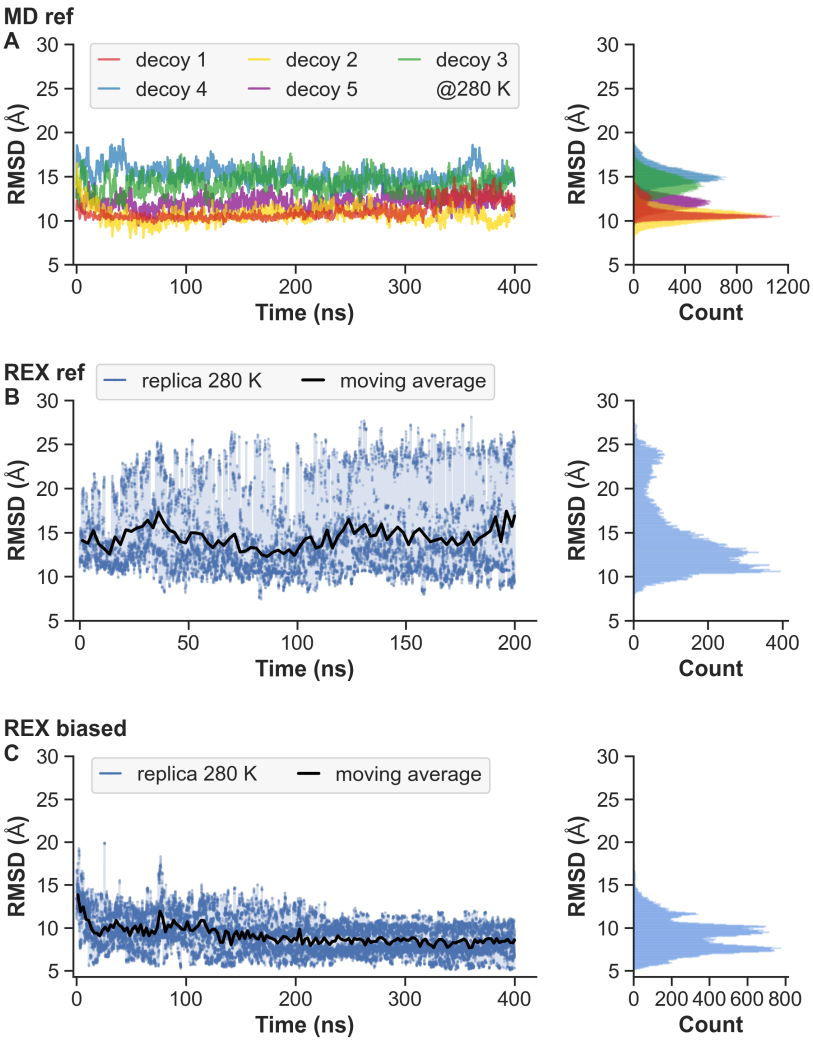


Figure 1. **RMSD comparison of simulations with TPP riboswitch (PDB id: 3d2g[36]).** Backbone RMSD of MD trajectories at 280 K. MD simulations were performed with five unique starting conformations (*decoys*). **(A)** MD without bias. **(B)** REX without bias. **(C)** REX with bias.

summarised the lowest observed RMSD values with respect to the native fold we found in our simulations. Figs. 1 to 3 show a comparison of MD and REX simulations.

In the case of TPP riboswitch (cf. Fig. 1) RMSD values reached approx. 5 Å during our simulations. Our REX bias strongly outperformed all other methods, whereas normal REX reached values of 7.4 Å and biased SimRNA with added mean field DCA[14] only 8.5 Å. In the case of the second RNA target, i.e. Adenine riboswitch (cf. Fig. 2), normal REX achieved 3.6 Å and performed slightly better than the biased variant with 4.0 Å.

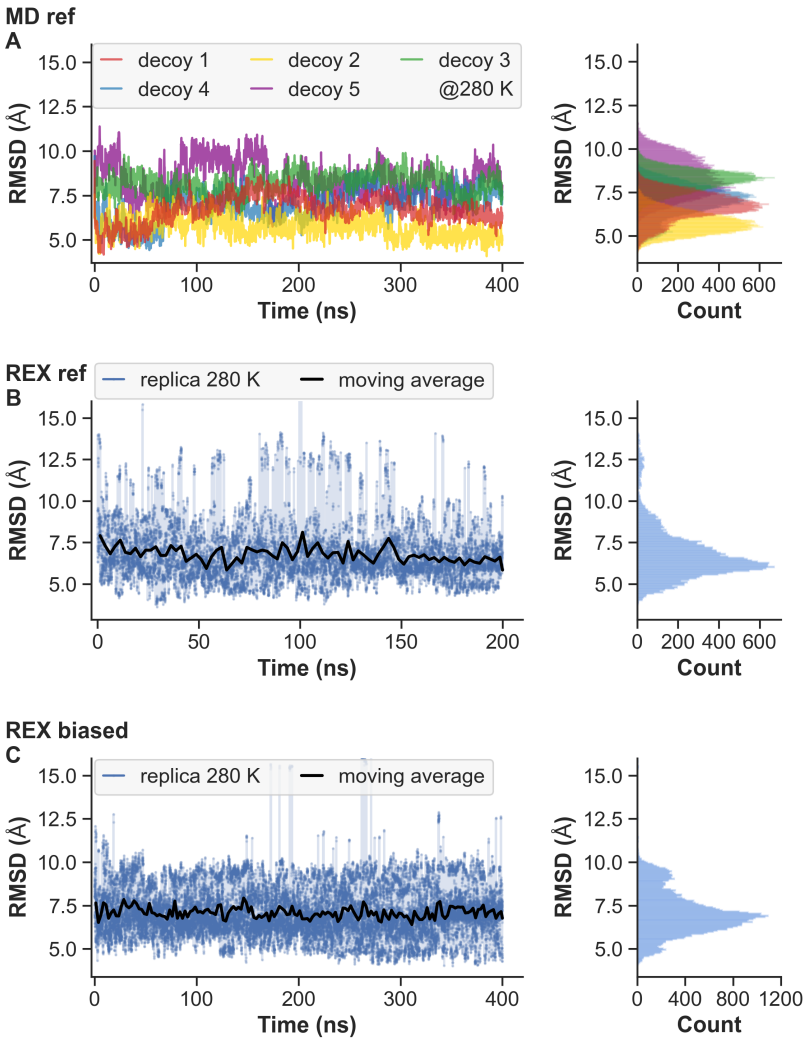


Figure 2. **RMSD comparison of simulations with Adenine riboswitch (PDB id: 4tzx[37]).** Backbone RMSD of MD trajectories at 280 K. MD simulations were performed with five unique starting conformations (*decoys*). **(A)** MD without bias. **(B)** REX without bias. **(C)** REX with bias.

SimRNA + mean field DCA yielded the best results with an RMSD of 3.0 Å. We want to emphasise that our applied force field seems for these cases to be very accurate and reliable, which is reflected in the good RMSD statistics of the performed MD simulations for both TPP riboswitch and Adenine riboswitch. Furthermore, the Adenine riboswitch seems to be very flexible by nature. The attractive force resulting from the applied bias potential was found to be too strong and resulted in a slight local bending of the RNA structure, which is reflected in the higher RMSD values as compared to the unbiased case. Lastly,

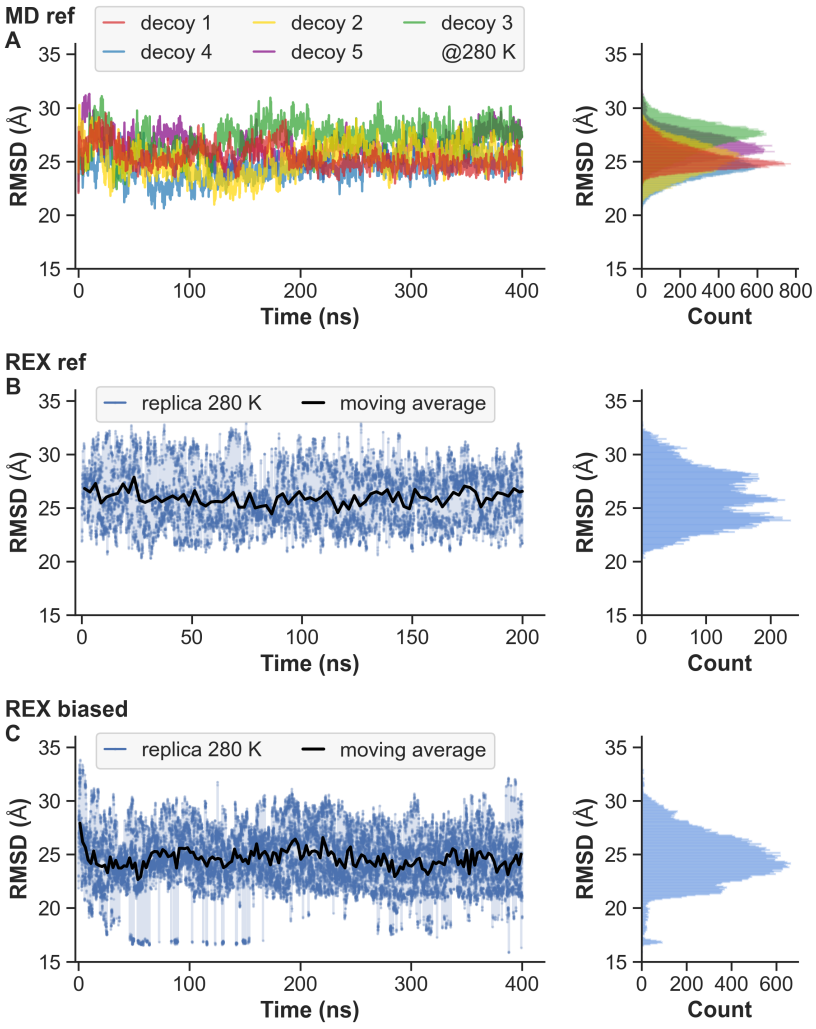


Figure 3. **RMSD comparison of simulations with 3',3'-cGAMP riboswitch (PDB id: 4yaz[38]).** Backbone RMSD of MD trajectories at 280 K. MD simulations were performed with five unique starting conformations (*decoys*). **(A)** MD without bias. **(B)** REX without bias. **(C)** REX with bias.

the comparison of the 3',3'-cGAMP riboswitch simulations are shown in Fig. 3. Our performed simulations yield the best results for the biased REX case, yet again. We want to stress, however, that the overall poor performance might have two courses. Either a) the force fields struggle with this particular example or b) the structure of the 3',3'-cGAMP riboswitch was experimentally measured as a dimer which is not reflected in the simulations, where we used only one molecule chain. Due to the missing counterpart and their physical interactions, the simulations would more likely to adopt non-native other struc-

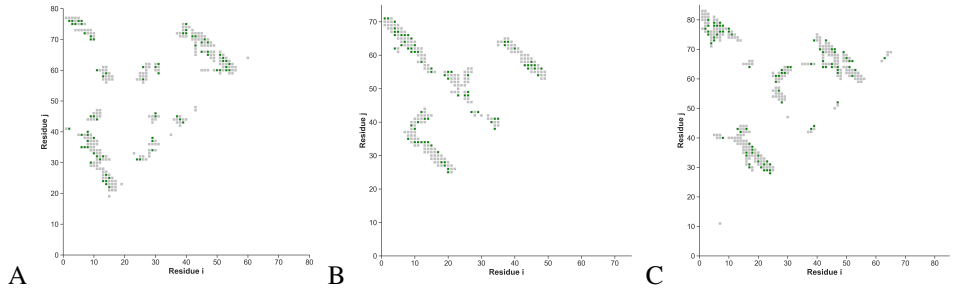


Figure 4. **Contact map of (A) TPP riboswitch (PDB id: 3d2g[36]), (B) Adenine riboswitch (PDB id: 4tzx[37]) and (C) 3',3'-cGAMP riboswitch (PDB id: 4yaz[38]).** Displayed are the native contacts (gray) and 75/70/80 randomly selected true-positive bias contacts (green) used in the respective REX-Bias simulations.

tures and therefore all methods led to relatively high RMSD values. While the biased REX simulations yielded values of 15.8 Å, SimRNA + CoCoNet was capable to achieve 14.0 Å. As all compared methods struggle, we tend to assume that b), i.e. the dimeric nature, is the source of the poor quality predictions.

3 Discussion and Summary

Overall, the application of contact-guided biased REX was able to provide low RMSD structures in two cases with one case only leading to poor quality structure. The poor performance of the third case can likely be attributed to its dimeric nature. There are several future avenues. First, DCA or related methods cannot reliably predict high numbers of error-free contacts. Future simulations should probe the methods robustness towards being fed erroneous contact information. Second the applied bias potential was optimised for proteins and not for RNAs. Hence, this method might or even should yield even better results once the bias potential has been optimised/ fine tuned specifically for RNA targets. Nevertheless, contact-guided REX provides good results which are on par with the best performances of the other mentioned methods, as shown in Tab. 1.

4 Methods

Each REX simulation had 60 replicas and a total of 20 unique starting conformations/ decoys (cf. Tab. 2). The temperatures were distributed covering a range of 280K to 370K. Time steps were set to 2 fs and exchange attempts each 500 MD steps. Exchange rates over the duration of the simulation were in the order of $10 - 15\%$. The all-atom simulations of this study utilised the OL15 nucleic force field[39,40] and TIP3P[41] explicit water model. All MD simulations are performed using GROMACS 2020[42] with cube-shaped systems boxes with periodic boundary conditions (PBC). For all simulations, we use a leap-frog integrator, V-rescale thermostat and Parrinello-Rahman barostat. For evaluation we used our toolkit pyrexMD[43].

Replica number	Decoy RMSD (Å) 3d2g	4tzx	4yaz	Replica number	Decoy RMSD (Å) 3d2g	4tzx	4yaz
1/21/41	10.96	9.37	21.45	11/31/51	12.64	7.67	23.35
2/22/42	13.55	7.49	23.31	12/32/52	14.13	6.57	22.89
3/23/43	12.41	6.23	23.28	13/33/53	12.60	7.88	22.58
4/24/44	13.95	9.68	21.56	14/34/54	13.20	8.65	21.54
5/25/45	12.45	7.47	22.48	15/35/55	14.32	9.37	23.57
6/26/46	13.03	11.24	24.68	16/36/56	11.15	7.78	22.53
7/27/47	13.08	8.49	20.86	17/37/57	10.97	7.78	24.90
8/28/48	10.94	8.90	24.58	18/38/58	12.58	10.57	24.35
9/29/49	13.21	9.41	22.35	19/39/59	13.58	9.19	21.29
10/30/50	11.85	10.30	21.80	20/40/60	12.79	9.19	21.45

Table 2. **Starting decoy accuracy of performed REX simulations.** Table shows the corresponding replica numbers, PDB ids of RNA targets and backbone root-mean-square-deviation (RMSD) before the simulation started.

Acknowledgements

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. B. Hwang, J-H. Lee, and D. Bang, *Single-cell RNA sequencing technologies and bioinformatics pipelines*, Experimental & Molecular Medicine, **50**, 1–14, 2018.
2. S. R. Eddy, *Hidden Markov models*, Curr. Opin. Struct. Biol., **6**, 361–365, 1996.
3. E. P. Nawrocki and S. R. Eddy, *Infernal 1.1: 100-fold faster RNA homology searches*, Bioinformatics, **29**, 2933–2935, 2013.
4. S. R. Eddy and R. Durbin, *RNA sequence analysis using covariance models*, Nucleic acids research, **22**, no. 11, 2079–2088, 1994.
5. J. D. Thompson, T. J. Gibson, and D. G. Higgins, *Multiple sequence alignment using ClustalW and ClustalX*, Curr. Protoc. Bioinf., **00**, 2.3.1–2.3.22, 2002.
6. H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley, and R. Lopez, *Analysis Tool Web Services from the EMBL-EBI*, Nucl. Acids Res., **41**, W597–W600, 2013.
7. D. J. Lipman and W. R. Pearson, *Rapid and sensitive protein similarity searches*, Science, **227**, 1435–1441, 1985.
8. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, *Basic local alignment search tool*, J. Mol. Biol., **215**, 403–410, 1990.
9. The UniProt Consortium, *UniProt: a worldwide hub of protein knowledge*, Nucl. Acids Res., **47**, D506–D515, 2019.
10. S. Griffiths-Jones, A. Bateman, M. Marshall, A. Khanna, and S. R. Eddy, *Rfam: an RNA family database*, Nucl. Acids Res., **31**, 439–441, 2003.

11. R. D. Finn, A. Bateman, J. Clements, P. Coggill, R. Y. Everhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. L. Sonnhammer, J. Tate, and M. Punta, *Pfam: the protein families database*, Nucl. Acids Res., **42**, D222–D230, 2014.
12. A. Schug, M. Weigt, J. N. Onuchic, T. Hwa, and H. Szurmant, *High-resolution protein complexes from integrating genomic information with molecular simulation*, Proceedings of the National Academy of Sciences, **106**, no. 52, 22124–22129, 2009.
13. M. Weigt, R. A. White, H. Szurmant, J. A. Hoch, and T. Hwa, *Identification of direct residue contacts in protein-protein interaction by message passing*, Proc. Natl. Acad. Sci. U.S.A., **106**, 67–72, 2009.
14. F. Morcos, A. Pagnani, B. Lunt, A. Bertolino, D. S. Marks, C. Sander, R. Zecchina, J. N. Onuchic, T. Hwa, and M. Weigt, *Direct-coupling analysis of residue coevolution captures native contacts across many protein families*, Proc. Natl. Acad. Sci. U.S.A., **108**, E1293–E1301, 2011.
15. D. S. Marks, T. A. Hopf, and C. Sander, *Protein structure prediction from sequence variation*, Nature Biotechnol., **30**, 1072–1080, 2012.
16. E. De Leonardis, B. Lutz, S. Ratz, S. Cocco, R. Monasson, A. Schug, and M. Weigt, *Direct-Coupling Analysis of nucleotide coevolution facilitates RNA secondary and tertiary structure prediction*, Nucleic acids research, **43**, no. 21, 10444–10455, 2015.
17. S. Cocco, C. Feinauer, M. Figliuzzi, R. Monasson, and M. Weigt, *Inverse statistical physics of protein sequences: a key issues review*, Rep. Prog. Phys., **81**, 032601, 2018.
18. M. B. Zerihun, F. Pucci, E. K. Peter, and A. Schug, *pydca v1. 0: a comprehensive software for Direct Coupling Analysis of RNA and Protein Sequences*, Bioinformatics, **36**, no. 7, 2264–2265, 2020.
19. M. Weiel, I. Reinartz, and A. Schug, *Rapid interpretation of small-angle X-ray scattering data*, PLoS Computational Biology, **15**, no. 3, e1006900, 2019.
20. I. Reinartz, C. Sinner, D. Nettels, B. Stucki-Buchli, F. Stockmar, P. T. Panek, C. R. Jacob, G. U. Nienhaus, B. Schuler, and A. Schug, *Simulation of FRET dyes allows quantitative comparison against experimental data*, The Journal of Chemical Physics, **148**, no. 12, 123321, 2018.
21. M. Weiel, M. Götz, A. Klein, D. Coquelin, R. Floca, and A. Schug, *Dynamic particle swarm optimization of biomolecular simulation parameters with flexible objective functions*, Nature Machine Intelligence, **3**, no. 8, 727–734, 2021.
22. A. E. Dago, A. Schug, A. Procaccini, J. A. Hoch, M. Weigt, and H. Szurmant, *Structural basis of histidine kinase autophosphorylation deduced by integrating genomics, molecular dynamics, and mutagenesis*, Proceedings of the National Academy of Sciences, **109**, no. 26, E1733–E1742, 2012.
23. G. Uguzzoni, S. J. Lovis, F. Oteri, A. Schug, H. Szurmant, and M. Weigt, *Large-scale identification of coevolution signals across homo-oligomeric protein interfaces by direct coupling analysis*, Proceedings of the National Academy of Sciences, **114**, no. 13, E2662–E2671, 2017.
24. R. R. Cheng, F. Morcos, H. Levine, and J. N. Onuchic, *Toward rationally redesigning bacterial two-component signaling systems using coevolutionary information*, Proceedings of the National Academy of Sciences USA, **111**, no. 5, E563–E571, 2014.

25. A. Kryshtafovych, T. Schwede, M. Topf, K. Fidelis, and J. Moult, *Critical assessment of methods of protein structure prediction (CASP)—Round XIII*, Proteins: Structure, Function, and Bioinformatics, **87**, no. 12, 1011–1020, 2019.
26. A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A. W. R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D. T. Jones, D. Silver, K. Kavukcuoglu, and D. Hassabis, *Improved protein structure prediction using potentials from deep learning*, Nature, **577**, no. 7792, 706–710, Jan 2020.
27. M. AlQuraishi, *End-to-End Differentiable Learning of Protein Structure*, Cell Systems, **8**, no. 4, 292–301.e3, 2019.
28. F. Pucci and A. Schug, *Shedding light on the dark matter of the biomolecular structural universe: Progress in RNA 3D structure prediction*, Methods, **162-163**, 68–73, 2019.
29. F. Pucci, M. B. Zerihun, E. K. Peter, and A. Schug, *Evaluating DCA-based method performances for RNA contact prediction by a well-curated dataset*, RNA, rna–073809, 2020.
30. M. B. Zerihun, F. Pucci, and A. Schug, *CoCoNet—boosting RNA contact prediction by convolutional neural networks*, Nucleic Acids Research, **49**, no. 22, 12661–12672, 2021.
31. U. H. E. Hansmann and Y. Okamoto, Curr. Opin. Struct. Biol., **9**, 177–183, 1999.
32. U. H. E. Hansmann and L. T. Wille, Phys. Rev. Lett., **88**, 068105, 2002.
33. A. Schug, W. Wenzel, and U. H. E. Hansmann, J. Chem. Phys., **122**, 194711, 2005.
34. A. Voronin, M. Weiel, and A. Schug, *Including residual contact information into replica-exchange MD simulations significantly enriches native-like conformations*, PloS one, **15**, no. 11, e0242072, 2020.
35. M. J. Boniecki, G. Lach, W. K. Dawson, K. Tomala, P. Lukasz, T. Soltysinski, K. M. Rother, and J. M. Bujnicki, *SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction*, Nucleic Acids Research, **44**, no. 7, e63–e63, 12 2015.
36. S. Thore, C. Frick, and N. Ban, *Structural basis of thiamine pyrophosphate analogues binding to the eukaryotic riboswitch*, Journal of the American Chemical Society, **130**, no. 26, 8116–8117, 2008.
37. J. Zhang and A. R. Ferré-D'Amaré, *Dramatic improvement of crystals of large RNAs by cation replacement and dehydration*, Structure, **22**, no. 9, 1363–1371, 2014.
38. A. Ren, X. C. Wang, C. A. Kellenberger, K. R. Rajashankar, R. A. Jones, M. C. Hammond, and D. J. Patel, *Structural Basis for Molecular Discrimination by a 3',3'-cGAMP Sensing Riboswitch*, Cell reports, **11**, no. 1, 1–12, 2015.
39. M. Zgarbová, M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T. E. Cheatham III, and P. Jurecka, *Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles*, Journal of Chemical Theory and Computation, **7**, no. 9, 2886–2902, 2011.
40. M. Zgarbová, J. Sponer, M. Otyepka, T. E. Cheatham III, R. Galindo-Murillo, and P. Jurecka, *Refinement of the sugar–phosphate backbone torsion beta for AMBER force fields improves the description of Z-and B-DNA*, Journal of chemical theory and computation, **11**, no. 12, 5723–5736, 2015.

41. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *Comparison of simple potential functions for simulating liquid water*, The Journal of chemical physics, **79**, no. 2, 926–935, 1983.
42. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, *GROMACS: fast, flexible, and free*, Journal of computational chemistry, **26**, no. 16, 1701–1718, 2005.
43. A. Voronin and A. Schug, *pyrexMD: Workflow-Orientated Python Package for Replica Exchange Molecular Dynamics*, Journal of Open Source Software, **6**, no. 68, 3325, 2021.

The NIC Excellence Projects

Spin Dynamics in Quantum Dot Ensembles

Iris Kleinjohann, Andreas Fischer, and Frithjof B. Anders

Condensed Matter Theory, Department of Physics, TU Dortmund University,
44221 Dortmund, Germany
E-mail: {iris.kleinjohann, frithjof.anders}@tu-dortmund.de

We investigate the spin dynamics of localised charge carriers, e.g., in semiconductor quantum dots, for which the hyperfine interaction with surrounding nuclear spins is strongly pronounced. Under optical cooling of nuclei, a strongly correlated nuclear-spin polaron state can form entailing long correlation times of electron and nuclear spin. Furthermore, inter quantum dot interactions in an ensemble of quantum dots produce entanglement of the electron spins in different singly charged dots visible in the cross-correlation spectrum. We developed versatile numerical methods customised to reveal the relevant dynamics of the respective scenario ranging from quantum mechanical treatments requiring customised diagonalisation routines to a semiclassical approach based on configuration averaging efficiently parallelised on massive parallel cluster architectures.

1 Introduction

Localised electron spins in quantum dot (QD) ensembles[1–4] have been discussed as implementation for qubits for quantum information processing. The advantage over superconducting qubits would be the straight forward integration in conventional semiconductor architectures with a potential coupling to optoelectronic elements that can provide coherent control.

The entanglement between the electron and nuclear spins is mediated by the hyperfine interaction[5,6] between the locally bound charge carrier spin and the nuclear spins in the volume of the localisation area. While the electron spin coherence time is limited by disordered nuclear spins, the formation of a nuclear-spin polaron state by the means of optical cooling reduces the dephasing and increases the spin coherence time: The nuclear spin noise is strongly suppressed, and the upper limit of the electron spin coherence time is determined by the time scale of additional spin-flip rates induced by the environment. This is based on the observation that the electron remains coupled to the lattice, whereas the very long lifetime of the nuclear spin polarisation up to several hours[7] indicates a strong decoupling of the nuclear spins from the environment.

Spin-spin interactions between pairs of electron spins in the QD ensemble have been demonstrated by two-colour pumping of singly charged semiconductor QDs[8]. We focus on spin-noise measurements of the ensemble in equilibrium in order to shed some light onto the potential microscopic origin of this inter-QD interaction and showed[9] that the two-colour cross-correlation spectrum can be used for extracting details on the inter-QD coupling distribution of so far unknown microscopic origin. However cross correlations have been much less studied theoretically[10] and have not been investigated by spin-noise spectroscopy in QDs yet.

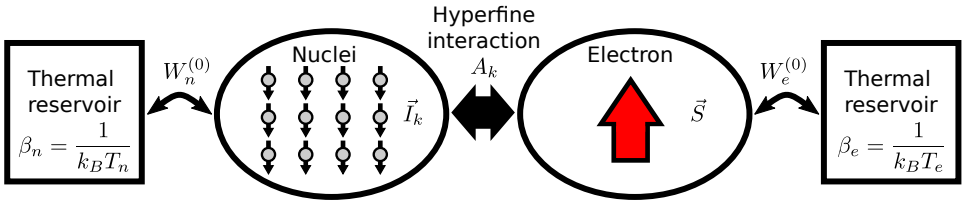


Figure 1. Modelling of the cooled electron-nuclear spin system. The hyperfine interaction between the electron spin $\vec{S}$ and the surrounding nuclear spins $\{\vec{I}_k\}$ is included in the Hamiltonian, whereas the distinct spin temperatures, T_e and T_n, are installed by the coupling to two separate thermal reservoirs which govern the spin and energy exchange of the system.

2 Model for Nuclear-Spin Polaron Formation

The localised electron spin in a singly charged semiconductor QD is coupled to the surrounding nuclear spins in the material via pronounced hyperfine interaction which constitutes a favourable ingredient for the realisation of strongly correlated spin states. Under optical cooling of the nuclear spins, the electron spin and the nuclear spins are expected to align oppositely (for positive coupling constants) forming a so called nuclear-spin polaron state[11] that minimises the hyperfine energy. The relevant Hamiltonian has the type of a central spin model (CSM),

$$H_{hf} = \sum_k A_k[\lambda(S^x I_k^x + S^y I_k^y) + S^z I_k^z], \tag{1}$$

with the electron spin operators S^x, S^y and S^z as well as the operators of an individual nuclear spin I_k^x, I_k^y and I_k^z. We label the nuclear spins with an index k and denote the respective hyperfine coupling constant by A_k. Furthermore, we include a uniaxial anisotropy via the parameter λ relevant for modelling localised hole spins. To account for the situation as it occurs under optical cooling of the nuclear spin system, we introduce two distinct temperatures of the electron spin and the nuclear spin thermal reservoirs, T_e and T_n. Experiments allow for the cooling of the nuclear system to much lower temperatures than the electronic temperature, $T_n < T_e$, since the nuclear spins strongly decouple from their environment whereas the electron spin is subject to relatively strong spin-lattice coupling and mostly retains the lattice temperature, typically $T_e \approx 4$ K. We enlarge the CSM to an open quantum system by coupling two separate thermal reservoirs with spin flips rates as sketched in Fig. 1. The spin and energy exchange of the system with the reservoirs is included in a quantum mechanic treatment via a kinetic rate equation formalism that captures the temporal evolution of the spin system as well as its steady state in thermal non-equilibrium for a broad range of temperatures T_e and T_n.

3 Numerical Results for the Cooled Spin System

For efficient calculation of the steady-state spin expectation values, the box model approximation of the hyperfine interaction in the Ising limit, $\lambda = 0$, is applied, i. e., we set $A_k = A_0$ for all nuclear spins in the system which allows for the treatment of $N = 10^5$

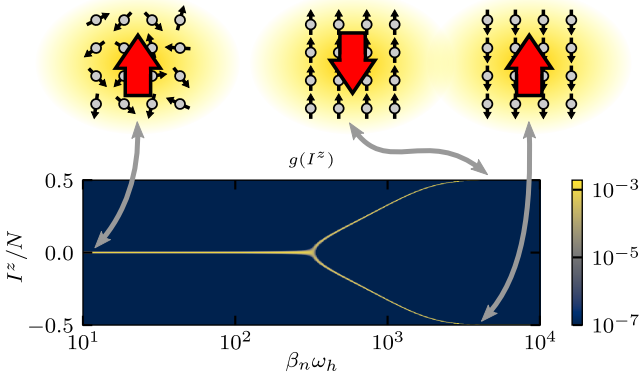


Figure 2. Nuclear distribution function $g(I^z)$ in dependence of the inverse nuclear spin temperature β_n for $\lambda = 0$. The inverse electron spin temperature is fixed at $\beta_e\omega_h = 0.05$. At low temperatures the formation of a strongly correlated nuclear-spin polaron formation is observed. The colour plot is taken from Ref. 12.

nuclear spins equalling to a realistic number of nuclei within the electron spin localisation volume. The system with a distribution of coupling constants beyond the box model approximation can be simulated by Monte Carlo methods[12]. The dephasing rate of the electron in the disordered nuclear bath, $\omega_h = \sqrt{\sum_k A_k^2} = \sqrt{N}A_0$, serves as the reference energy scale.

3.1 Distribution Function of the Nuclear Spin

The formation of the nuclear-spin polaron state with the decreasing temperature can be conveniently observed by means of the distribution function of the nuclear spins. For indistinguishable nuclei, it is useful to introduce the total nuclear spin $I^z = \sum_k I_k^z$. The related distribution function $g(I^z)$ which is obtained from the diagonal of the system's density operator is displayed in Fig. 2. At a fixed inverse electron spin temperature, $\beta_e\omega_h = \omega_h/k_BT_e = 0.05$, the distribution function changes its shape with decreasing the nuclear spin temperature, i. e. increasing the inverse $\beta_n = 1/k_BT_n$. At high temperatures, the spins are randomly aligned since the thermal energy considerably exceeds the hyperfine energy. As a result the distribution function $g(I^z)$ has a Gaussian shape with a maximum at $I^z = 0$. Lowering the temperature to a certain point, the single maximum of $g(I^z)$ splits into two maxima at $I^z \neq 0$ which move to larger absolute values of I^z with decreasing nuclear spin temperature. These two branches correspond to the two possible orientations of the nuclear-spin polaron state along the z axis for $\lambda = 0$.

For the description of a general anisotropic system, i. e. $\lambda \neq 0$, however the dynamics of the full density operator has to be incorporated. To this end we generalised our theory to a Lindblad formalism capturing the quantum system in contact with the external temperature-inducing reservoirs[13]. The steady-state density operator of the generalised Lindblad formalism can be numerically obtained from diagonalisation of the associated transition matrix with customised Krylov subspace methods. We found that in a spin system with arbitrary anisotropic hyperfine interaction the nuclear polaron state forms along

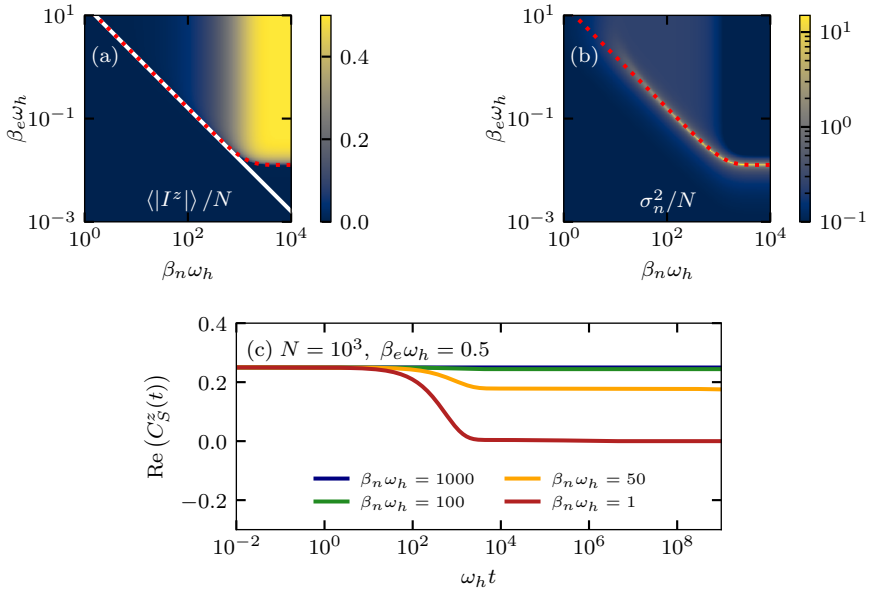


Figure 3. (a) Absolute value of the total nuclear spin z component, $\langle|I^z|\rangle$, and (b) its fluctuations, σ_n^2, in dependence of the inverse electron and nuclear spin temperatures, β_e and β_n, for a system with $N = 10^5$ individual nuclear spins. (c) Electron spin auto-correlation as a function of time for different nuclear spin temperatures. The plots calculated for $\lambda = 0$ are taken from Refs. 12, 13.

the axes of strongest hyperfine coupling respecting the symmetry of the respective Hamiltonian[13].

3.2 Nuclear-Spin Polaron Formation

For the investigation of a nuclear polaron state formation, the expectation value of the absolute value of the total nuclear spin z projection[14], $\langle|I^z|\rangle$, is examined over a broad temperature range, see Fig. 3(a). In the low temperature regime a notable orientation of the nuclear spins is observed which eventually approaches the maximum alignment $\langle|I^z|\rangle/N = 0.5$ (yellow colour coding). The quantum fluctuations of $\langle|I^z|\rangle$,

$$\sigma_n^2 = \left\langle (I^z)^2 \right\rangle - \langle|I^z|\rangle^2 , \tag{2}$$

provide a valuable quantity[14] for the examination of the crossover temperature below which a nuclear-spin polaron state is formed, see Fig. 3(b). The maximum line of σ_n^2 within the (β_n, β_e) plane clearly indicates where the nuclear spins start to align forming the polaronic state. Analytically we can predict the crossover temperatures[12] to obey the relation

$$\beta_n = \frac{4}{A_0} \mathrm{arctanh}\left(\frac{4}{(N+2)\beta_e A_0}\right) \tag{3}$$

indicated by the red dotted line. This analytic criterion is based on the bifurcation point of the nuclear distribution function and remarkably coincides with the numerical results.

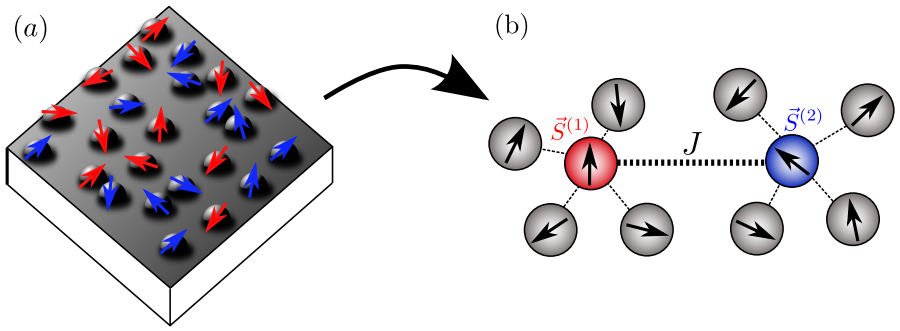


Figure 4. (a) Scheme of an ensemble of singly charged QDs. The red coloured electron spins are excited by one laser, the blue coloured spins are addressed by the second laser. (b) Simplified model with two representative QDs. The two coupled electron spins are subject to hyperfine interaction with separate nuclear spin baths.

The auto-correlation function of the electron spin, $C_S^z(t) = \langle S^z(0)S^z(t)\rangle$ as shown in Fig. 3(c), reveals that the polaron formation at low temperatures is accompanied by an increase of correlation time: The nuclear spin orientation prevents a loss of electron spin coherence yielding a significantly prolonged correlation time.

4 Extended Central Spin Model

For a localised electron spin in a QD, the predominant interaction is the hyperfine interaction with the surrounding nuclear spins captured by the CSM, Eq. 1. However at long time scales further interactions contribute to the spin dynamics. Experiments suggest[8,15] that in an ensemble of QDs the dynamics of the electron spins in the individual dots are correlated: The electron spins in different dots interact with each other via direct or indirect exchange mechanisms. While directly adjacent QDs may couple via direct exchange interaction, an indirect long-range spin exchange interaction mediated by charge carriers in the wetting layer is possible as well.

For potential application in quantum information processing, interactions between the QDs are essential. A useful observable for the examination of the dynamics of interacting QDs is the spin cross-correlation function which can be obtained in two-colour spin-noise experiments[10]. Each of the two lasers reads out the electron spin fluctuations of a distinct QD subensemble, see Fig. 4(a). By a correlation of the individual laser signals[10] one can obtain the cross-correlation function

$$C_2^{(\times)}(t) = \left\langle \tilde{S}_z^{(1)}(0)\tilde{S}_z^{(2)}(t)\right\rangle + \left\langle \tilde{S}_z^{(2)}(0)\tilde{S}_z^{(1)}(t)\right\rangle, \tag{4}$$

with the electron spin polarisation $\tilde{S}_z^{(i)}$ in subensemble i.

From a numerical point of view, the inter QD interaction leads to a complex network of interacting nuclear and electron spins: Each QD consists of an electron spin coupled to roughly 10^5 nuclear spins. Furthermore, there are $10^5 - 10^6$ QDs within the laser spot on a typical sample which interact with each other. To handle this complex system on a super computing cluster we mapped[9] an ensemble of 10^5 interacting QDs onto 10^5 pairs of representative QDs, where the representative QDs stand for the two subensembles excited

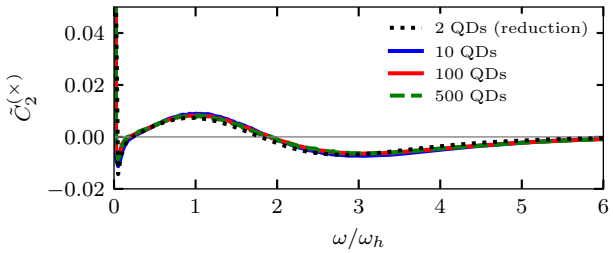


Figure 5. Cross-correlation spectra for clusters of different size are compared to a reduced two-QD model (without external magnetic field). Each QD contains $N = 100$ nuclear spins. The plot is taken from Ref. 9.

in two-colour spectroscopy as depicted in Fig. 4(b). The cross-correlation spectrum of the QD pairs can be computed independently, allowing for massive parallelisation in a semiclassical approach based on configuration averaging.

The Hamiltonian of an interacting QD pair, H_2, includes a Heisenberg coupling between the electron spins $\vec{S}^{(i)}$ with coupling strength J

$$H_2 = \sum_i H_1^{(i)} + J\vec{S}^{(1)}\vec{S}^{(2)} \tag{5}$$

where $H_1^{(i)}$ is the Hamiltonian of a single QD with index i and the effective inter QD coupling J is drawn from its distribution function. The Hamiltonian $H_1^{(i)}$ includes the isotropic hyperfine interaction according to Eq. 1 with $\lambda = 1$ as well as additional contributions such as an external magnetic field or nuclear quadrupolar interactions arising from the growth induced tensions in the self-assembled QDs. While quadrupolar interactions are negligible on short time scales, they lead to disorder in the nuclear spin bath on longer time scales and thereby affect the low frequency spectrum.

5 Numerical Results for Interacting Quantum Dots

The cross-correlation function $\tilde{C}_2^{(\times)}(t)$ is constructed in such a way, that it vanishes for non-interacting QDs but displays non-zero contributions for finite interaction J. In order to establish the quality of the mapping from a QD ensemble onto a configuration averaging of QD pairs with a random J, we plot the corresponding cross correlation spectrum $\tilde{C}_2^{(\times)}(\omega)$ in Fig. 5 for QD clusters with a different number of QDs and compare the results to a configuration averaged QD pair calculation. We use equal hyperfine couplings, $A_k = A_0 = \omega_h/\sqrt{N}$, and neglect quadrupolar interactions. These simplifications allow for the simulation of large QD ensembles on long times scales. The spectrum of a reduced QD pair agrees excellently with the complex simulation of an ensemble with 500 QDs.

For the analysis of the contributions in the cross-correlation spectrum, we refer to the sum rule

$$\int_{-\infty}^{\infty} \tilde{C}_2^{(\times)}(\omega)\,\mathrm{d}\omega = 0 \tag{6}$$

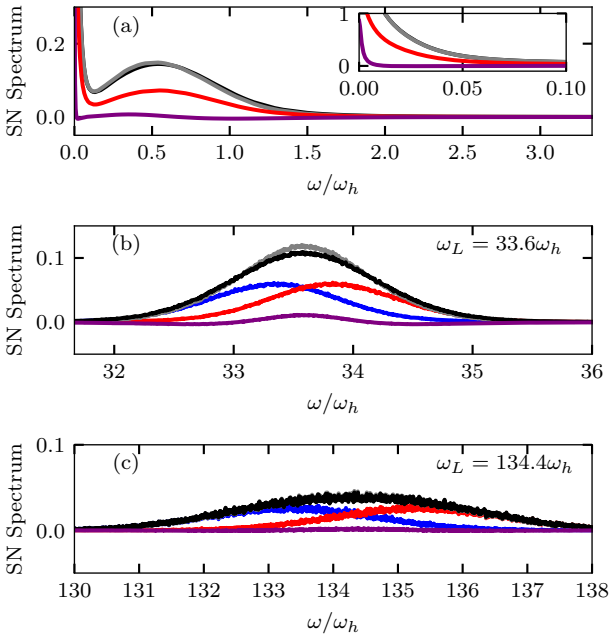


Figure 6. Correlation spectra for an external magnetic field of (a) 0T, (b) 1T, and (c) 4T. The autocorrelation spectra $\tilde{C}_2^{(1)}$ (red) and $\tilde{C}_2^{(2)}$ (blue), the sum $\tilde{C}_2^{(1)} + \tilde{C}_2^{(2)}$ (gray), the combined autocorrelation spectrum $\tilde{C}_2^{(1+2)}$ (black) and the cross-correlation spectra $\tilde{C}_2^{(\times)}$ (purple) are depicted as a function of the angular frequency ω. The plots are taken from Ref. 9.

valid for the high-temperature limit $\beta\omega_h > 1$ explored here. Each positive component in the spectrum is counterbalanced by negative contributions at other frequencies. We can clearly recover this relation in Fig. 5: The weights of the broad positive and negative peak placed around $\omega/\omega_h = 2$, which stem from fluctuations of the nuclear Overhauser field, cancel each other in magnitude. The narrow positive and negative peak in the low-frequency spectrum can be attributed to the inter-QD interaction which breaks the conservation of the total angular momentum of a single QD: Without inter-QD interactions, the total angular momentum of each individual QD is conserved. With interactions between the QDs, only the total angular momentum of the full ensemble is conserved which leads to additional slow dynamics reflected by the low-frequency spectrum.

For the spectra in Fig. 6, we utilised a realistic set of parameters[9] that account for a self-assembled QD ensemble. We include a distribution of the hyperfine couplings, nuclear quadrupolar interactions[16–18] and a dispersion of the electron g factor. The three panels show the cross-correlation spectrum (purple) for three different transversal magnetic fields, (a) $0\,\mathrm{T}$, (b) $1\,\mathrm{T}$ and (c) $4\,\mathrm{T}$. In addition to the cross-correlation spectrum, the auto-correlation spectra $\tilde{C}_2^{(i)}$ (red/blue), the sum of both (black) and the combined autocorrelation spectrum $\tilde{C}_2^{(1+2)}$ (grey) are displayed. By introducing quadrupolar interactions, the conservation of angular momentum for each single QD is further lifted even

for $J = 0$. In the cross-correlation spectrum, this fact is reflected by almost completely vanishing cross-correlations in the low-frequency range, see inset in Fig. 6(a).

A transversal external magnetic field shifts the correlation spectra to higher frequencies such that low frequency contributions disappear. In addition the g-factor distribution of the electron spins becomes relevant. In Fig. 6(b), the autocorrelation spectra of both QDs (red and blue) diverge as a result of the deviating mean electron g factors of the subensembles: the auto-correlation function is centred around the mean Larmor frequency of each subensemble. The cross-correlation spectrum is centred between the autocorrelation functions of both QDs. However, for very strong magnetic fields, see Fig. 6(c), the cross-correlation spectrum is strongly attenuated by the g factor dispersion.

Our analysis reveals that the cross-correlation spectrum is an important tool for understanding the dynamics of interacting QDs. Furthermore, other interactions interfere with the inter-QD interaction producing additional features at characteristic frequency ranges which match the energy scale of the specific interaction: Albeit quadrupolar interactions attenuate the cross-correlation spectrum at zero external magnetic field and the electron g factor dispersion leads to extinction of the cross correlations at strong external fields, in between a wide range of magnetic fields provides the opportunity for the examination of the characteristic of the inter-QD interaction.

6 Concluding Remarks

We reported on two aspects of our recent studies on singly-charged semiconductor QD ensembles under optical excitation: the nuclear polaron formation at low temperatures and the cross-correlation spectroscopy of interacting QDs. We used augmenting approaches: a fully quantum mechanical approach that is able to handle 10^3-10^5 nuclear spins in the box model approximation in the case of the nuclear-spin polaron formation and a representation of the quantum mechanical time evolution by a quaternionic treatment[19] for cross-correlation spectroscopy. Both approaches can be parallelised very efficiently and our code has been tailored to the JURECA Booster as well as to JUWELS.

Acknowledgements

This project was supported by the Deutsche Forschungsgemeinschaft through the transregio TRR 160 within the Projects No. A4 and No. A7. We thank Mikhail M. Glazov and Nikolai A. Sinitsyn for many fruitful discussions. Computation time was provided by the NIC at Jülich Supercomputing Centre under project No. HDO09.

References

1. 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**, no. 4, 1217–1265, Oct 2007.
2. M. I. Dyakonov, (Ed.), *Spin physics in semiconductors*, Springer Series in Solid-State Sciences 157. Springer International Publishing, 2nd ed. edition, 2017.

3. M. M. Glazov, *Electron & Nuclear Spin Dynamics in Semiconductor Nanostructures*, Series on Semiconductor Science and Technology. OUP Oxford, 2018.
4. D. Loss and D. P. DiVincenzo, *Quantum computation with quantum dots*, Phys. Rev. A, **57**, no. 1, 120–126, Jan 1998.
5. I. A. Merkulov, Al. L. Efros, and M. Rosen, *Electron spin relaxation by nuclei in semiconductor quantum dots*, Phys. Rev. B, **65**, 205309, Apr 2002.
6. W. A. Coish and D. Loss, *Hyperfine interaction in a quantum dot: Non-Markovian electron spin dynamics*, Phys. Rev. B, **70**, 195340, Nov 2004.
7. V. K. Kalevich, K. V. Kavokin, I. Merkulov, and M. R. Vladimirova, "Dynamic nuclear polarization and nuclear fields", in: Spin physics in semiconductors, M. I. Dyakonov, (Ed.), Springer Series in Solid-State Sciences 157, 309, Springer International Publishing, 2nd edition, 2017.
8. 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, Sep 2011.
9. A. Fischer, I. Kleinjohann, N. A. Sinitsyn, and F. B. Anders, *Cross-correlation spectra in interacting quantum dot systems*, Phys. Rev. B, **105**, 035303, Jan 2022.
10. D. Roy, L. Yang, S. A. Crooker, and N. A. Sinitsyn, *Cross-correlation spin noise spectroscopy of heterogeneous interacting spin systems*, Scientific Reports, **5**, no. 1, 9573, Apr 2015.
11. I. A. Merkulov, Al. L. Efros, and M. Rosen, *Electron spin relaxation by nuclei in semiconductor quantum dots*, Phys. Rev. B, **65**, 205309, Apr 2002.
12. A. Fischer, I. Kleinjohann, F. B. Anders, and M. M. Glazov, *Kinetic approach to nuclear-spin polaron formation*, Phys. Rev. B, **102**, 165309, Oct 2020.
13. I. Kleinjohann, A. Fischer, M. M. Glazov, and F. B. Anders, *Nuclear spin polaron-formation: anisotropy effects and quantum phase transition*, submitted to Phys. Rev. B, 2022, arXiv:2202.02105.
14. K. Binder, *Finite size scaling analysis of ising model block distribution functions*, Zeitschrift für Physik B Condensed Matter, **43**, no. 2, 119–140, Jun 1981.
15. 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, Nov 2018.
16. C. Bulutay, *Quadrupolar spectra of nuclear spins in strained InGaAs quantum dots*, Phys. Rev. B, **85**, 115313, Mar 2012.
17. N. A. Sinitsyn, Yan Li, S. A. Crooker, A. Saxena, and D. L. Smith, *Role of Nuclear Quadrupole Coupling on Decoherence and Relaxation of Central Spins in Quantum Dots*, Phys. Rev. Lett., **109**, 166605, Oct 2012.
18. J. Hackmann, Ph. Glasenapp, A. Greilich, M. Bayer, and F. B. Anders, *Influence of the Nuclear Electric Quadrupolar Interaction on the Coherence Time of Hole and Electron Spins Confined in Semiconductor Quantum Dots*, Phys. Rev. Lett., **115**, 207401, Nov 2015.
19. P. R. Girard, *The quaternion group and modern physics*, European Journal of Physics, **5**, no. 1, 25–32, Jan 1984.

Searching for the H Dibaryon in Lattice QCD

Jeremy R. Green[1], **Andrew D. Hanlon**[2], **Parikshit M. Junnarkar**[3], **Madanagopalan Padmanath**[4], **Srijit Paul**[5], **and Hartmut Wittig**[5,4]

[1] School of Mathematics and Hamilton Mathematics Institute, Trinity College, Dublin 2, Ireland

[2] Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA

[3] Institut für Kernphysik, Technische Universität Darmstadt, Schlossgartenstraße 2, 64289 Darmstadt, Germany

[4] Helmholtz Institute Mainz, GSI Helmholtz Centre for Heavy Ion Research, Staudingerweg 18, 55128 Mainz, Germany

[5] PRISMA$^+$ Cluster of Excellence and Institute for Nuclear Physics, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany
E-mail: hartmut.wittig@uni-mainz.de

Hadronic resonances play a central role for the understanding of the strong interaction and provide crucial information for hadron-hadron interactions and the formation of light nuclei. Here we investigate the possible existence of the H dibaryon, a conjectured state of two Λ hyperons which is an example of a double hypernucleus. We present results from lattice QCD calculations, employing the formalism of finite-volume quantisation which allows for a reliable determination of the binding energy from the pole of the scattering amplitude. Our results, obtained at six different lattice spacings in the limit of SU(3)-flavour symmetry demonstrate clear evidence for the existence of a bound H dibaryon with binding energy of 4.5 MeV. We discuss the extension of our calculations to the more realistic case of broken SU(3) symmetry and to different dibaryon systems.

1 Introduction

Lattice QCD has had enormous success in describing the hadron spectrum, both in the mesonic and baryonic sectors[1–4]. Masses of stable particles can be determined in lattice QCD from the exponential decay rate of the two-point correlation function of suitable interpolating operators. However, this ignores the fact that most hadronic states are resonances that decay into two or more lighter hadrons. Lüscher's method of finite-volume quantisation[5,6] is an elegant and versatile formalism that allows for the investigation of resonances and the determination of scattering amplitudes from which bound state and resonance poles can be extracted from lattice QCD data. This has paved the way for investigations of hadron-hadron interactions and bound states from first principles.

In this contribution we apply finite-volume quantisation to study the H dibaryon and related dibaryon systems. We have performed the first lattice calculation of the binding energy of the H dibaryon in the continuum limit of three-flavour QCD. Our findings imply that effects of non-zero lattice spacing are large and must be removed by a systematic extrapolation to the continuum limit before drawing any definite conclusions about the existence of such a bound state.

2 The H Dibaryon

The H dibaryon is a spinless flavour-singlet state, originally proposed by Jaffe[7] in 1977 as a deeply bound system of six quarks with flavour content $uuddss$ ("hexaquark"). While it is unclear whether such a state is realised in nature, it has been speculated that a deeply bound scalar $uuddss$ hexaquark might qualify as a dark matter candidate, provided that its binding energy is of the order of a few hundred MeV[8–15]. An alternative interpretation of a possible H dibaryon is that of a weakly bound state of two Λ hyperons. Clearly, in this case one expects the binding energy to be of order 10 MeV. As another motivation, studying the properties of a potential Λ-Λ bound state will help constrain hadronic (Λ-Λ) interactions which are relevant for the physics of double hyper-nuclei, neutron-rich matter and neutron stars. Recently, experimental data for two-particle correlations in p-p, p-Pb and Au-Au collisions[16–18] have been analysed to constrain the Λ-Λ interaction and provide model estimates for the binding energy of the H dibaryon. There is a long history of calculations studying whether the H dibaryon is a prediction of QCD. Results for the binding energy B_H from the more recent calculations with dynamical quarks[19–26] vary considerably, with estimates ranging from a few MeV up to 75 MeV, depending on the methodology and/or the value of the pion mass (see Fig. 3). Recently, employing near-physical pion and kaon masses, the HAL QCD Collaboration reported that the Λ-Λ interaction is only weakly attractive and does not sustain a bound or resonant dihyperon[25].

3 Finite-Volume Energy Levels

Our task is to determine the energy spectrum of a six-quark state with spin zero, which is also an SU(3)-flavour singlet. To this end we compute a hermitian correlator matrix $C_{ij} \equiv \langle O_i(t) O_j^\dagger(0) \rangle$ from a basis of interpolating operators O_i. The finite-volume energy spectrum $\{E_n\}$ determines the exponential fall-off of C_{ij}. Our basis of interpolators consists of products of two single-baryon operators, projected to momenta $\vec{p}_1$ and $\vec{p}_2$, with total momentum $\vec{P} = \vec{p}_1 + \vec{p}_2$. In order to realise the flavour content and symmetry of the H dibaryon, we take appropriate linear combinations of bilocal isospin-zero operators $\Lambda\Lambda$, $\Sigma\Sigma$ and $N\Xi$ with strangeness $S = -2$, which transform according to the singlet irreducible representation of SU(3) flavour. A key ingredient in our calculation is the use of the "distillation" technique[27] and Laplacian-Heaviside smearing[28]. This allows us to compute "timeslice-to-all" quark propagators which are used to compute hermitian correlator matrices C_{ij} of bilocal interpolating operators in all frames characterised by the total momentum $\vec{P}$. Finite-volume energy levels are then determined in each frame by solving a generalised eigenvalue problem[29–31] i.e.

$$C(\tau_D) v_n = \lambda_n \, C(\tau_0) v_n \tag{1}$$

for fixed timeslices τ_0 and τ_D and subsequently using the eigenvectors v_n to transform the correlator matrix $C(t)$ into an approximately diagonal correlator $\tilde{C}(t)$ according to

$$\tilde{C}_{nm}(t) \equiv v_n^\dagger C(t) v_m. \tag{2}$$

Asymptotically, i.e. for $t \gg 0$, the diagonal elements of $\tilde{C}(t)$ fall off exponentially with a rate proportional to the finite-volume energy levels for the given frame. In practice, we determine the shift ΔE of energy levels of the H dibaryon relative to the

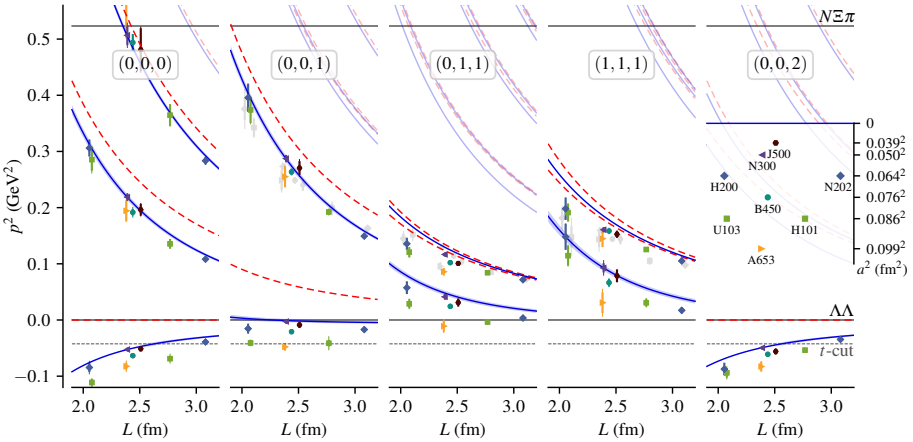


Figure 1. Finite-volume spectrum: centre-of-mass scattering momentum p^2 versus lattice extent L. Solid horizontal lines show the two- and three-particle thresholds, while dashed horizontal lines represent the t-channel cut of $-m_\pi^2/4$. The non-interacting spectrum is denoted by red dashed curves, and solid blue curves show the interacting spectrum determined in the continuum. The inset serves as a legend showing L and a^2 for the ensembles used in our calculation.

energy of two non-interacting Λ particles, by fitting the large-t behaviour of the ratio $R_n(t) \equiv \tilde{C}_{nn}(t)/[C_\Lambda^{\vec{p}_1}(t)\, C_\Lambda^{\vec{p}_2}(t)]$.

Our calculations have been performed on gauge ensembles employing O(a)-improved Wilson fermions, generated as part of the CLS effort[32]. For our study of the H dibaryon[26] we focused initially on the SU(3)-symmetric case, i.e. with degenerate up, down and strange quark masses, corresponding to $m_\pi = m_K \simeq 420\,\text{MeV}$. Correlator matrices were computed for six different values of the lattice spacing, allowing for a detailed scaling test of our results. In Fig. 1 we show the finite-volume energy levels expressed in terms of the squared scattering momentum, defined by

$$p^2 = \tfrac{1}{4}(E^2 - \vec{P}\cdot\vec{P}) - m_\Lambda^2\,, \tag{3}$$

where $\vec{P}$ denotes the frame momentum, and m_Λ is the mass of a single Λ hyperon. These data form the basis for the subsequent amplitude analysis based on the finite-volume quantisation condition.

4 Scattering Amplitude and Bound State

Two-particle finite-volume quantisation conditions can be written in the generic form[33]

$$\det\left[\tilde{K}^{-1}(p^2) - B(p^2)\right] = 0, \tag{4}$$

where $\tilde{K}$ contains the scattering amplitude and B depends on the volume, frame, and irrep. In finite volume, the symmetry group of rotations and inversions is finite, and the irreducible representations of the little group of $\vec{P}$ replace total angular momentum J as a conserved quantum number. Each irrep contains an infinite tower of different J. The

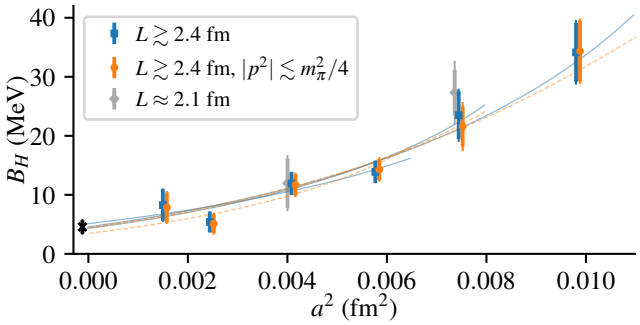


Figure 2. Continuum extrapolation of the binding energy. Blue, orange and grey data points represent results from different kinematical regimes and/or spatial volumes. The final estimate is shown as the black cross.

rows/columns of $\tilde{K}$ and B correspond to the different coupled channels and partial waves that can scatter as well as the different J.

For studying the H dibaryon, we truncate Eq. 4 to its simplest form, with a single channel and keeping only S-wave, neglecting D-wave and higher partial waves. The quantisation condition becomes

$$p \cot \delta_0(p^2) = \frac{2}{\sqrt{\pi} L \gamma} Z_{00}^{\vec{P} L/(2\pi)} \left(1, \left(\frac{pL}{2\pi}\right)^2\right), \tag{5}$$

where $Z_{00}^{\vec{D}}$ is a generalised zeta function. This can be used in two ways. Given a finite-volume energy level, we can obtain p^2, and then Eq. 5 provides the S-wave phase shift δ_0 at p^2. Conversely, given an *ansatz* for the phase shift as a function of p^2, we can find the solutions to Eq. 5 and compare them with the finite-volume energy levels: effectively, Eq. 5 can translate an *ansatz* for $\delta_0(p^2)$ into an *ansatz* for the spectrum.

Our primary analysis is based on performing global fits to the spectra from all of the ensembles with different lattice spacings and different volumes. The fit *ansatz* assumes that $p \cot \delta_0(p^2)$ can be described by a polynomial in p^2 with coefficients that are affine functions of the squared lattice spacing a^2:

$$p \cot \delta_0(p^2) = \sum_{i=0}^{N-1} c_i p^{2i}, \quad c_i = c_{i0} + c_{i1} a^2. \tag{6}$$

This fit, combined with the quantisation condition, is able to provide a good description of our finite-volume energy levels, with the exception of the excited states in frames $(0,1,1)$ and $(1,1,1)$, which are strongly affected by the neglected D wave. Sending $a \to 0$, we obtain the continuum finite-volume energy levels shown as the blue curves in Fig. 1.

Bound states correspond to poles on the physical sheet of the scattering amplitude below threshold, i.e. for $p = i\kappa$, $\kappa > 0$, and can be found as solutions to $p \cot \delta(p) = -\sqrt{-p^2}$. The dependence of the binding energy B_H of the H dibaryon on a^2 is shown in Fig. 2, where the curves correspond to global fits and the points correspond to fits to individual ensembles. There is a strong dependence on the lattice spacing: B_H on the coarsest ensemble is larger than its value in the continuum by a factor of about 7.5. Our

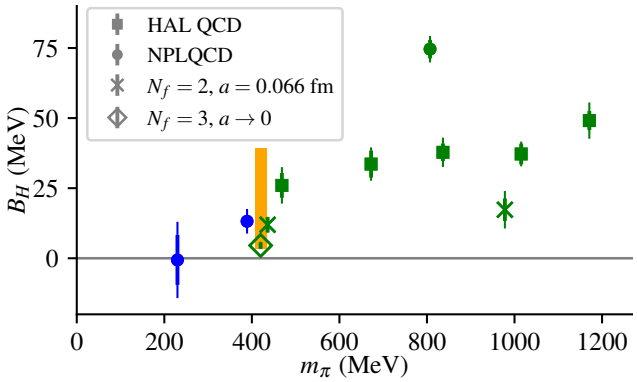


Figure 3. Binding energy of the H dibaryon plotted versus the squared pion mass. Our result in the continuum limit is shown as the diamond, with the orange vertical band representing the range of binding energies obtained at non-zero lattice spacings, as shown in Fig. 2. Our results in two-flavour QCD at non-zero lattice spacing[24] are shown as green crosses. The remaining data points were obtained by the NPLQCD[19–21] and HAL QCD[22,23] collaborations.

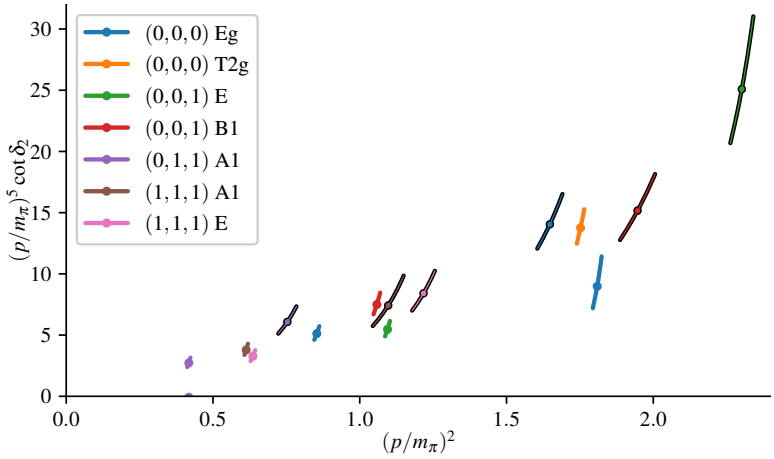


Figure 4. D-wave phase shift in the flavour singlet sector, shown as $p^5 \cot \delta_2(p^2)$ versus p^2, in units of the pion mass. Results are shown for two ensembles with the same lattice spacing but different volumes: N202 (solid colours) and H200 (with black outlines). See the inset in Fig. 1 for further details. Different colours correspond to different $\vec{P}$ and different irreps.

final result in the continuum limit is $B_H = 4.56 \pm 1.13 \pm 0.63$ MeV, where the second (systematic) uncertainty is obtained by varying the plateau region used to obtain the spectra and by performing cuts on a, L, and p^2.

Our result is shown alongside other lattice calculations in Fig. 3: the cleanest comparison is with the other green points, which are also based on SU(3)-symmetric setups. Every previous calculation was done using a single lattice spacing and, therefore, without control over discretisation effects. We obtain a significantly smaller binding energy, and

our observed discretisation effects are large enough to explain this discrepancy. This small binding energy, albeit at unphysical quark masses, is evidence against a deeply bound hexaquark state. If this is confirmed for physical light and strange quark masses, it will firmly rule out the scenario of a $uuddss$ bound state as a dark matter candidate.

The set of baryon-baryon correlation functions we computed at the SU(3) point is much larger than the dataset used for the H dibaryon. In particular, we have data in other SU(3) flavour sectors and in nontrivial little group irreps. The latter are essential for constraining higher partial waves. As an example, we show preliminary results for the flavour singlet D-wave phase shift, determined at a single nonzero lattice spacing, in Fig. 4. The data provide good constraints on this phase shift, with the main limitation being the lack of data at small p^2. Constraining higher partial waves near threshold will be difficult without larger volumes.

Analysis of nucleon-nucleon scattering is currently underway. In nature, there is a bound $J^P = 1^+$ deuteron, but the situation with a heavier-than-physical pion mass is unclear: some lattice calculations find no evidence for bound nucleon-nucleon states, whereas others report a deuteron more deeply bound than in nature, as well as a bound 0^+ state. The deuteron sector is particularly challenging to analyse, since it is spin one and the relevant scattering system has coupled S and D waves.

5 SU(3) Breaking

The calculations performed in our published work on the H dibaryon were carried out such that the u and d quarks are heavier than in nature, the s quark is lighter, and $m_u+m_d+m_s$ is set equal to its physical value. The resulting SU(3) symmetry greatly simplifies our study of the H dibaryon, as only one (flavour singlet) baryon-baryon scattering channel contributes. With broken SU(3), there are three relevant coupled channels: $\Lambda\Lambda$, $N\Xi$, and $\Sigma\Sigma$. For physical quark masses, the splittings between thresholds for these three channels are considerably larger than our computed H dibaryon binding energy at the SU(3) symmetric point; therefore, we expect that SU(3) breaking can play a significant role in determining whether there is a bound state. We have generated correlator data for several ensembles

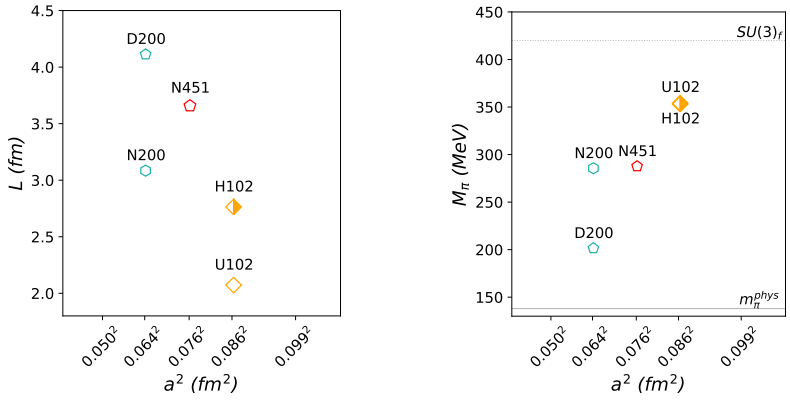


Figure 5. Ensembles with SU(3) breaking, on the (L, a^2) plane and on the (m_π, a^2) plane.

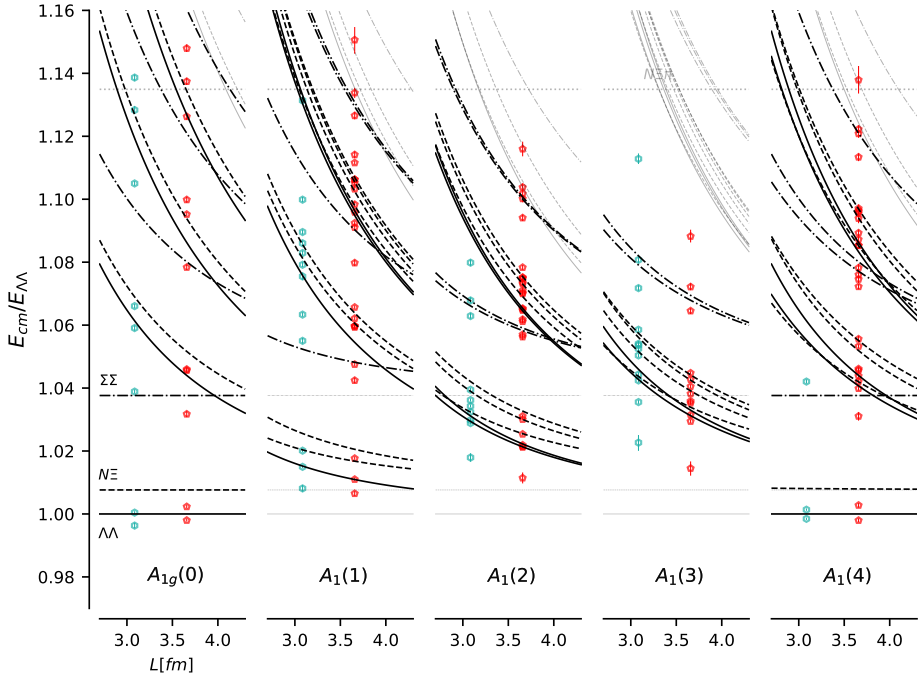


Figure 6. Energy spectrum in the H-dibaryon sector (isospin zero, strangeness -2) in the trivial finite-volume irreps on ensembles with $m_\pi \approx 280$ MeV. Red and cyan markers refer to ensembles N451 and N200, respectively. Solid, dashed, and dash-dotted curves show non-interacting $\Lambda\Lambda$, $N\Xi$, and $\Sigma\Sigma$ energy levels and thresholds.

with pion masses above the physical pion mass but below the SU(3) symmetric pion mass, shown in Fig. 5. This will allow us to study the dependence of baryon-baryon scattering observables on the amount of SU(3) breaking. We also employ multiple lattice spacings so that we can begin to study discretisation effects at a lighter pion mass.

We illustrate these data using the spectrum in the H dibaryon sector on the ensembles with a pion mass of roughly 280 MeV, shown in Fig. 6. The presence of three coupled channels causes the spectrum to be much denser, and some levels are nearly degenerate. This causes the analysis of the spectrum using quantisation conditions to be much more challenging than in the SU(3)-symmetric case; we are exploring methods such as eigenvalue decomposition of the quantisation condition[34].

6 Conclusions

Finite-volume quantisation is a powerful formalism that allows for the investigation of hadron resonances and hadron-hadron interactions from first principles, including the reliable determination of binding energies and scattering lengths. We have performed the first comprehensive investigation of the H dibaryon in the continuum limit of three-flavour QCD, finding evidence for a weakly bound state. Our study has shown that discretisation

effects are sizeable and must be taken into account to obtain a reliable estimate for the binding energy. The extension of our analysis to higher partial waves and to nucleon-nucleon scattering is under way. We have also obtained first results on finite-volume energy levels in the SU(3)-broken case, with the light $(u,\, d)$ and strange quark masses tuned closer to their physical values in nature. Future work will include charmed tetraquarks, such as the T_{cc} and charmonium-nucleon pentaquark states.

Acknowledgements

We thank Max Hansen, Ben Hörz and Daniel Mohler for many helpful discussions. Calculations for this project have been performed on JUQUEEN, JURECA and JUWELS at JSC as part of project CHMZ21. This research was partly supported by DFG through the Cluster of Excellence "Precision Physics, Fundamental Interactions and Structure of Matter" (PRISMA+ EXC 2118/1) funded within the German Excellence Strategy (Project ID 39083149), as well as the Collaborative Research Centers SFB 1044 "The low-energy frontier of the Standard Model" and CRC-TR 211 "Strong-interaction matter under extreme conditions" (Project ID 315477589). ADH is supported by: (i) The U.S. Department of Energy, Office of Science, Office of Nuclear Physics through the Contract No. DE-SC0012704 (S.M.); (ii) The U.S. Department of Energy, Office of Science, Office of Nuclear Physics and Office of Advanced Scientific Computing Research, within the framework of Scientific Discovery through Advance Computing (SciDAC) award Computing the Properties of Matter with Leadership Computing Resources. We are grateful to our colleagues within the CLS initiative for sharing ensembles.

References

1. S. Aoki *et al.*, Phys. Rev. **D79**, 034503, 2009, arXiv:0807.1661.
2. S. Dürr *et al.*, Science **322**, 1224, 2008, arXiv:0906.3599.
3. C. Alexandrou *et al.*, Phys. Rev. **D78**, 014509, 2008, arXiv:0803.3190.
4. A. Bazavov *et al.*, Rev. Mod. Phys. **82**, 1349, 2010, arXiv:0903.3598.
5. M. Lüscher, Nucl. Phys. **B354**, 531, 1991.
6. M. Lüscher, Nucl. Phys. **B364**, 237, 1991.
7. R. L. Jaffe, Phys. Rev. Lett. **38**, 195, 1977, [Erratum: Phys. Rev. Lett. 38, 617, 1977].
8. G. R. Farrar, 2017, arXiv:1708.08951.
9. C. Gross, A. Polosa, A. Strumia, A. Urbano, and W. Xue, Phys. Rev. D **98**, 063005, 2018, arXiv:1803.10242.
10. G. R. Farrar, 2018, arXiv:1805.03723.
11. E. W. Kolb and M. S. Turner, Phys. Rev. D **99**, 063519, 2019, arXiv:1809.06003.
12. S. D. McDermott, S. Reddy, and S. Sen, Phys. Rev. D **99**, 035013, 2019, arXiv:1809.06765.
13. The BaBar Collaboration, J. P. Lees *et al.*, Phys. Rev. Lett. **122**, 072002, 2019, arXiv:1810.04724.
14. K. Azizi, S. S. Agaev, and H. Sundu, J. Phys. G **47**, 095001, 2020, arXiv:1904.09913.
15. G. R. Farrar, Z. Wang, and X. Xu, 2020, arXiv:2007.10378.

16. L. Adamczyk *et al.*, Phys. Rev. Lett. **114**, 022301, 2015, arXiv:1408.4360.
17. S. Acharya *et al.*, Phys. Rev. C **99**, 024001, 2019, arXiv:1805.12455.
18. S. Acharya *et al.*, Phys. Lett. B **797**, 134822, 2019, arXiv:1905.07209.
19. S. R. Beane *et al.*, Phys. Rev. Lett. **106**, 162001, 2011, arXiv:1012.3812.
20. S. R. Beane *et al.*, Mod. Phys. Lett. A **26**, 2587, 2011, arXiv:1103.2821.
21. S. R. Beane *et al.*, Phys. Rev. D **87**, 034506, 2013, arXiv:1206.5219.
22. T. Inoue *et al.*, Phys. Rev. Lett. **106**, 162002, 2011, arXiv:1012.5928.
23. T. Inoue *et al.*, Nucl. Phys. A **881**, 28, 2012, arXiv:1112.5926.
24. A. Francis *et al.*, Phys. Rev. D **99**, 074505, 2019, arXiv:1805.03966.
25. K. Sasaki *et al.*, Nucl. Phys. A **998**, 121737, 2020, arXiv:1912.08630.
26. J. R. Green, A. D. Hanlon, P. M. Junnarkar, and H. Wittig, Phys. Rev. Lett. **127**, 242003, 2021, arXiv:2103.01054.
27. M. Peardon *et al.*, Phys. Rev. D **80**, 054506, 2009, arXiv:0905.2160.
28. C. Morningstar *et al.*, Phys. Rev. D **83**, 114505, 2011, arXiv:1104.3870.
29. C. Michael, Nucl. Phys. B **259**, 58, 1985.
30. M. Lüscher and U. Wolff, Nucl. Phys. B **339**, 222, 1990.
31. B. Blossier, M. Della Morte, G. von Hippel, T. Mendes, and R. Sommer, JHEP **04**, 094, 2009, arXiv:0902.1265.
32. M. Bruno *et al.*, JHEP **02**, 043, 2015, arXiv:1411.3982.
33. C. Morningstar *et al.*, Nucl. Phys. B **924**, 477, 2017, arXiv:1707.05817.
34. A. J. Woss, D. J. Wilson, and J. J. Dudek, Phys. Rev. D **101**, 114505, 2020, arXiv:2001.08474.

Nuclear Lattice Effective Field Theory: Status A.D. 2022

Ulf-G. Meißner[1,2,3]

[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) and Institut für Kernphysik (IKP-3), Forschungszentrum Jülich, 52425 Jülich, Germany

[3] Tbilisi State University, 0186 Tbilisi, Georgia

I discuss recent developments in nuclear lattice effective field theory, which is a premier tool in the theory of nuclear structure and reactions. I discuss a number of exciting results on topics such as nuclear thermodynamics, perturbative nuclear structure calculations beyond first order perturbation theory and a three-dimensional tomography of the carbon nucleus. I also discuss alpha-alpha scattering in the multiverse and give an outlook on upcoming results at N3LO precision.

1 Introduction

Understanding the formation of strongly interacting systems such as atomic nuclei from first principles calculations is still one of the biggest challenges within contemporary theoretical physics. While the theory of the strong interactions, Quantum Chromodynamics (QCD), is well tested in many processes, the matter that leads to life in our Universe is based on nuclei, which 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. Much progress in the understanding of the structure and dynamics of nuclei has been made in the context of Nuclear Lattice Effective Field Theory (NLEFT)[1], which combines the so successful low-energy chiral effective field theory of QCD with stochastic methods (Monte Carlo simulations). While direct calculations of nuclei based on quarks and gluons in the framework of lattice QCD are essentially impossible due to the severe sign problem, formulating the nuclear forces in terms of protons, neutrons and pions is not only more appropriate, but also comes with the added value of the approximate Wigner SU(4) (spin-isospin) symmetry of the underlying nuclear interactions. This symmetry in fact suppresses the sign oscillations strongly, and in the limit of an exact Wigner SU(4) symmetry, spin-isospin saturated nuclei like e.g. ^{4}He are free of any sign oscillation. In NLEFT simulations, Euclidean space-time is discretised on a torus of volume $L^3 \times L_t$, where L is the side length of the spatial dimension, and L_t denotes the extent of the Euclidean time dimension. The lattice spacing in the spatial (temporal) dimensions is a (a_t). The maximal momentum on the lattice is $p_{\rm max} \equiv \pi/a$, which serves as the UV regulator of the theory. Nucleons are point-like 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. Properties of multi-nucleon systems are computed by means of the projection Monte

Carlo (MC) method. Each nucleon is treated as a single particle propagating in a fluctuating background of pion and auxiliary fields. Both local and non-local smearings are applied to the nucleon creation and annihilation operators. Euclidean time projection is started from some initial state Ψ_A for Z protons and N neutrons (with $A = Z + N$). Then, a Wigner SU(4) symmetric Hamiltonian is used as a computationally inexpensive filter for the first few Euclidean time steps. This suppresses sign oscillations dramatically. Finally, the full LO chiral EFT Hamiltonian is applied and one calculates the ground state energy and other properties from the correlation function $Z(t) \equiv \langle \Psi_A | \exp(-tH) | \Psi_A \rangle = \mathrm{Tr}\{M^{L_t}\}$, in the limit of large Euclidean projection time t, with M the normal-ordered transfer-matrix operator and L_t the number of Euclidean time steps. Higher-order contributions are computed as perturbative corrections to the LO amplitude. The properties of excited states are obtained from a multi-channel projection MC method. Further, lattice improvement and smearing is applied to the momentum-dependent operators and operators are included, that allow to remove the artifacts from the breaking of rotational symmetry on the lattice. A much more detailed description is given in the recent monograph[1].

2 Nuclear Thermodynamics

The equation of state of strongly interacting matter is one of the central topics in contemporary nuclear physics, as it plays an important role in the early universe and the generation of gravitational waves in neutron star mergers. There are also important connections between the nuclear equation of state and heavy-ion collisions. It is well established that highly-excited nuclear states can be treated statistically as part of an equilibrium thermal distribution. The large density of states at high energies allows a treatment in terms of thermodynamical concepts, such as temperature, entropy, and free energy. Within NLEFT a new paradigm for calculating *ab initio* nuclear thermodynamics for lattice simulations was developed in Ref. 2. This is based on a new and efficient method, called the pinhole trace algorithm (PTA), for computing nuclear observables at nonzero temperature using a canonical ensemble with fixed numbers of protons and neutrons. For a canonical ensemble with fixed nucleon number A, volume $V = L^3$ and temperature T, the expectation value of any observable $\mathcal{O}$ is given by

$$\langle \mathcal{O} \rangle_\beta = \frac{Z_\mathcal{O}(\beta)}{Z(\beta)} = \frac{\mathrm{Tr}_A(e^{-\beta H}\mathcal{O})}{\mathrm{Tr}_A(e^{-\beta H})}, \tag{1}$$

with $Z(\beta)$ the partition function, $\beta = T^{-1}$ the inverse temperature, H the Hamiltonian, and Tr_A is the trace over the A-body Hilbert space. Throughout, canonical units with $\hbar = c = k_B = 1$ are utilised. The PTA can be used to efficiently compute $Z(\beta)$ and $Z_\mathcal{O}(\beta)$ on the lattice. The PTA is an extension of the pinhole algorithm introduced in Ref. 3 to sample the spatial positions and spin/isospin indices of the nucleons. The new feature is that a quantum mechanical trace over all possible states is performed in addition. The partition function $Z(\beta)$ can be written explicitly in the single particle basis as

$$Z(\beta) = \sum_{c_1,\cdots,c_A} \langle c_1, \cdots, c_A | \exp(-\beta H) | c_1, \cdots, c_A \rangle, \tag{2}$$

where the basis states are Slater determinants composed of point particles, $c_i = (\boldsymbol{n}_i, \sigma_i, \tau_i)$ which combine the lattice vector of the i-th particle, its spin σ_i i and isospin τ_i. The neutron

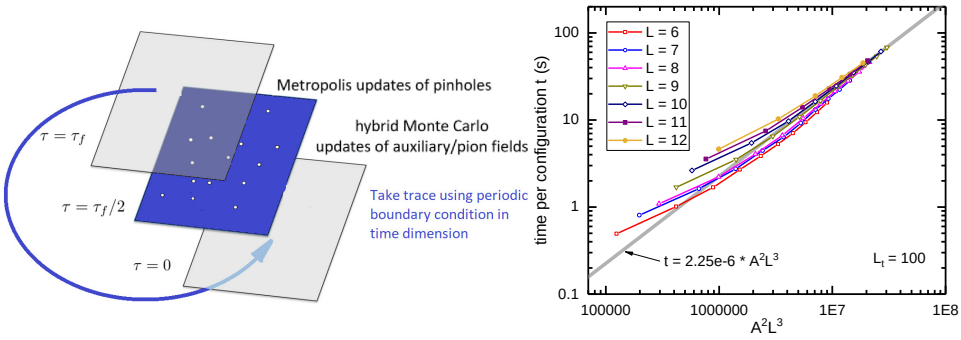


Figure 1. Left panel: The pinhole trace algorithm (PTA). At mid-times $t = L_t a_t/2$, a sheet with pinholes of spin-isospin indices $(i_n j_n)$ is inserted and the periodic boundary conditions impose the temperature $\beta = 1/T$. Right panel: Scaling of the PTA with nucleon number A and volume $V = L^3$.

number N and proton number Z are separately conserved, and the sum in Eq. 2 is limited to the subspace with the specified values for N and Z. By decomposing the interactions in H using auxiliary fields, one obtains the path-integral expression for Eq. 2

$$Z(\beta) = \sum_{c_1,\cdots,c_A} \int \mathcal{D}s_1 \cdots \mathcal{D}s_{L_t} \langle c_1, \cdots, c_A | \, M(s_{L_t}) \cdots M(s_1) | c_1, \cdots, c_A \rangle, \quad (3)$$

with the $M(s_{n_t})$ the normal-ordered transfer matrix for time step n_t, and s_{n_t} is a shorthand notation for all auxiliary fields at that time step. In the PTA one evaluates Eq. 3 using Monte Carlo methods, see the left panel of Fig. 1 for a graphical illustration and more details are given in Ref. 2. The arguably biggest advantage of the PTA is its time scaling $\sim A^2 V L_t$, see the right panel of Fig. 1. This is very different to commonly used algorithms based on the grand-canonical ensemble, like the well-known BSS method first described in Ref. 4. The BBS algorithm scales as $AV^2 L_t$, so the cost savings of the PTA is a factor of V/A. Thus, the speedup associated with the PTA can be as large as one thousand, depending on the lattice spacing and particle density. Another important technical aspect related to the calculation of thermodynamic quantities is the use of twisted boundary conditions (TBCs). The finite volume together with the chosen boundary condition will induce fictitious shell effects. The origin of the finite volume shell effects is the constraint imposed by the boundary conditions on the particle momenta. One way to overcome this is to use TBCs[5]. Here, extra phases are attached to the wave function when a particle cross the boundaries. Averaging over all possible twist angles provides an efficient way of approaching the infinite volume limit[6–9,2]. In Ref. 2 simulations to investigate the nuclear liquid-vapour phase transition were performed on $L^3 = 4^3, 5^3, 6^3$ cubic lattices with up to 144 nucleons and a spatial lattice spacing $a = 1.32$ fm, such that the corresponding momentum cutoff is $\Lambda = \pi/a \approx 471$ MeV. The temporal lattice spacing is taken to be $a_t = 1/2000$ MeV^{-1}. The temperatures considered spanned the range from 10-20 MeV and the densities varied from $n = 0.008$ fm^{-3} to $n = 0.20$ fm^{-3}. Remember that nuclear matter density is $n_0 = 0.16$ fm^{-3}. These settings allow one to explore the whole region relevant to the nuclear matter liquid-vapour phase transition. Throughout, symmetric nuclear matter with $N = Z$ and without the Coulomb interaction was considered. In Fig. 2 the calculated chemical potential and pressure isotherms are

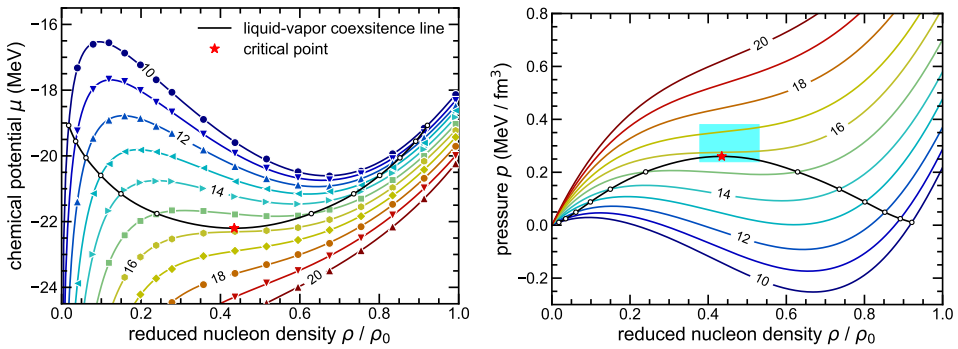


Figure 2. Left panel: The μ-ρ isotherms of symmetric nuclear matter in a volume $L^3 = 6^3$. The numbers on the lines are temperatures in MeV. The black line denotes the liquid-vapour coexistence line derived from Maxwell construction, and the red star marks the calculated critical point. Right panel: The p-ρ isotherms of symmetric nuclear matter for $L^3 = 6^3$. The black line denotes the liquid-vapour coexistence line, and the red star marks the calculated critical point. The cyan rectangle marks the empirical critical point extracted from heavy-ion collisions[10].

shown. Each point represents a separate simulation. All the MC errors for μ are smaller than 0.02 MeV and not shown explicitly. The critical point is then deduced from solving the equations $d\mu/d\rho = d^2\mu/d\rho^2 = 0$. The uncertainties in the critical values are estimated by propagating the simulation and interpolation errors. The critical temperature, density, and chemical potential are, in order: $T_c = 15.80(3)$ MeV, $\rho_c = 0.089(1)$ fm^{-3}, and $\mu_c = -22.20(1)$ MeV. The liquid-vapour coexistence line is determined through the Maxwell construction of each isotherm and depicted as a solid black line in Fig. 2. All the other bulk thermodynamic quantities can be reliably extracted based on the calculated high-precision energies and chemical potentials. In the right panel of Fig. 2 the deduced pressure-density isotherms are shown as well as the corresponding liquid-vapour coexistence line and the critical point. The calculated critical pressure is P_c =0.260(3) MeV/fm^3. For comparison, the critical point extracted by analysing the cluster distributions in heavy-ion collisions, $T_c = 17.9(4)$ MeV, $\rho_c = 0.06(1)$ fm^{-3} and $P_c = 0.31(7)$ MeV/fm^3 is also shown[10]. Note that these calculations employ a leading-order chiral interaction which slightly overestimates the nuclear matter saturation density ρ_0, which leads to a systematic error of about 10%. Thus, the quality of both ρ_0 and ρ_c calculations can be improved by including higher-order corrections. Note that the calculated isotherms follow exactly the pattern expected for a liquid-vapour phase transition in a finite system. Above T_c the system is in a supercritical state, while below T_c the pure liquid and vapour phases exist in the high- and low-density regime, respectively. For states encompassed by the two arms of the coexistence line, the system is a mixture of the liquid and vapour phases. In the future, these calculations will be improved by using more precise interactions.

3 Perturbative Quantum MC Method for Nuclear Physics

Quantum Monte Carlo (QMC) simulations are a powerful method for addressing quantum many-body problems in many area of physics. Often, the so-called "sign problem" leads

to exponential scaling of the computational effort with system size. Despite this, there are a number of cases where QMC algorithms can be applied without sign problems (such as the leading-order SU(4) symmetric nuclear interaction used in the preceding section). The realistic systems of physical interests, though, often deviate from these ideal models significantly and have a sign problem. In these cases, perturbation theory can be used to bridge the difference between the simplified and the realistic interaction. However, in QMC perturbation theory is mostly limited to the first order. Improving the quality of the perturbative calculations requires going to higher orders. In Rayleigh-Schrödinger perturbation theory, the second-order energy correction involves a summation over all quantum states that can be reached via the perturbing interaction. Such a calculation over all quantum states is not compatible with QMC, which targets only the lowest energy state. Thus, a different approach is needed. To solve this problem, a new computational framework called perturbative QMC (ptQMC), which allows for the efficient calculation of higher-order perturbative corrections within the Euclidean time formalism, was introduced in Ref. 11. To be specific, consider the nuclear Hamiltonian given $H = K+V_0+V_C$, with $K = -\nabla^2/2m$ the kinetic energy operator and m the nucleon mass. The interaction is split into a dominant term V_0 and a correction V_C. The perturbative expansion of the energy takes the form $E = E_0 + \delta E_1 + \delta E_2 + \ldots$, where the indices $0, 1, 2, ...$ denote the leading order, the first order perturbative correction and so. The partial energy contributions at each order are

$$\begin{aligned} E_0 &= \langle\Psi_0|(K+V_0)|\Psi_0\rangle/\langle\Psi_0|\Psi_0\rangle\ ,\ \ \delta E_1 = \langle\Psi_0|V_C|\Psi_0\rangle/\langle\Psi_0|\Psi_0\rangle, \\ \delta E_2 &= \mathrm{Re}(\langle\Psi_0|V_C|\delta\Psi_1\rangle - \delta E_1\langle\Psi_0|\delta\Psi_1\rangle)/\langle\Psi_0|\Psi_0\rangle, \end{aligned} \tag{4}$$

where the wave function Ψ has a similar perturbative expansion

$$\begin{aligned} |\Psi\rangle &= \lim_{L_t\to\infty} M^{L_t/2}|\Psi_T\rangle = |\Psi_0\rangle + |\delta\Psi_1\rangle + \mathcal{O}(V_C^2), \\ |\Psi_0\rangle &= \lim_{L_t\to\infty} M_0^{L_t/2}|\Psi_T\rangle,\ |\delta\Psi_1\rangle = \lim_{L_t\to\infty}\sum_{k=1}^{L_t/2} M_0^{L_t/2-k}(M-M_0)M_0^{k-1}|\Psi_T\rangle, \end{aligned} \tag{5}$$

with $M_0 =: \exp\{-a_t(K+V_0)\}$: the zeroth order transfer matrix and the $\mathcal{O}(a_t^2)$ terms are omitted. Consequently, all matrix elements and overlaps can be expressed with the amplitudes,

$$\mathcal{M}(O) = \langle\Psi_T|M_0^{L_t/2}OM_0^{L_t/2}|\Psi_T\rangle, \tag{6}$$

$$\mathcal{M}_k(O) = \langle\Psi_T|M_0^{L_t/2}OM_0^{L_t/2-k}MM_0^{k-1}|\Psi_T\rangle, \tag{7}$$

where $k = 1, 2, \cdots, L_t/2$. Here O is the operator inserted in the middle time step like $1, K+V_0$ or V_C. In $\mathcal{M}_k(O)$ the k-th copy of M_0 is replaced by the full transfer matrix M. The transfer matrices M_0 and M in these amplitudes are computed using the auxiliary field formalism. The energies E_0 and δE_1 are just the expectation values $\langle O\rangle = \mathcal{M}(O)/\mathcal{M}(1)$ with $O = K+V_0$ or V_C. These can be calculated by sampling the auxiliary fields s in M_0 with standard algorithms. For δE_2 we need to evaluate the quotient $\mathcal{M}_k(O)/\mathcal{M}(1)$. What remains to be calculated is an integral over the auxiliary field c from the inserted M in $\mathcal{M}_k(O)$. For every sample $\{s_1, s_2, \cdots, s_{L_t}\}$ we have

$$\mathcal{M}_k(O) = \int \mathcal{D}cP(c+\bar{c})\langle\cdots O\cdots M(s_k, c+\bar{c})\cdots\rangle_T, \tag{8}$$

where the ellipses denote the transfer matrices $M_0(s_t)$ with $t \neq k$, $\langle\rangle_T$ the expectation value in the state $|\Psi_T\rangle$ and $P(c)$ is the standard normal distribution. In Eq. 8 we have made a variable change $c \to \bar{c} + c$ with c real integral variables. Here, $\bar{c}(\boldsymbol{n})$ is a constant field

$$\bar{c}(\boldsymbol{n}) = \frac{\partial}{\partial c(\boldsymbol{n})} \ln\langle \cdots M(s_k, c) \cdots \rangle_T \Big|_{c=0} = \sqrt{-a_t C}\langle \cdots : M_0(s_k)\rho_c(\boldsymbol{n}) : \cdots \rangle_T / \mathcal{M}(1), \tag{9}$$

where the ellipses again represent the M_0's, C is the coupling constant for the V_C term. Generally, $\bar{c}$ is a complex field, e.g., for repulsive interactions such as Coulomb we have $C > 0$, and the square root in Eq. 9 introduces an imaginary factor i. In this case the integrand in Eq. 8 contains non-zero phases that may induce a severe sign problem. The variable change in Eq. 8 serves to alleviate this problem[12]. To see this, we take the logarithm of the integrand in Eq. 8, expand the result near $c = 0$ and apply Eq. 9. We find that the terms linear in c and $\bar{c}$ which cause the sign problem cancel exactly and the integrand can be factorised as

$$\mathcal{M}_k(O) = \mathcal{M}(s) \exp\left(\frac{\bar{c}^2}{2}\right) \int \mathcal{D}c \exp\left(-\frac{c^2}{2} + \epsilon\right), \tag{10}$$

where we omit the sum over lattice sites, and ϵ is a residual term containing quadratic and higher powers of c. Because in $\mathcal{M}_k(s, c)$ a common factor $\sqrt{a_t}$ is attached to every c variable, ϵ is a small number of the order $\mathcal{O}(a_t)$. For sufficiently small a_t, Eq. 10 means that the integrand in Eq. 8 is a product of a normal distribution and a slowly varying function $\exp(\epsilon)$. We can use stochastic methods to evaluate Eq. 8 by sampling the c field with a standard normal distribution. This evaluation is unbiased and its uncertainty is determined

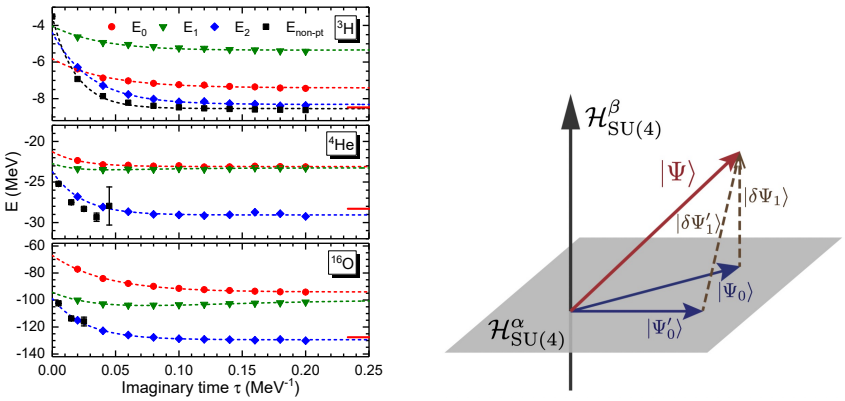


Figure 3. Left panel: ptQMC binding energies as functions of the projection time τ compared with non-perturbative results. The circles (red), down triangles (green) and diamonds (blue) denote the energies at the zeroth, first and second orders, respectively. The squares (black) represent the exact results calculated with sparse matrix multiplications for ^{3}H and full non-perturbative QMC for ^{4}He and ^{16}O, respectively. Each group of results are fitted with a sum of exponential functions (dashed lines). The red bars mark the experimental binding energies. Right panel: Schematic plot for a perturbative calculation. The zeroth order wave functions $|\Psi_0\rangle$ and $|\Psi'_0\rangle$ are confined in a subspace corresponding to an *irrep* of SU(4).

by the variation of $\exp(\epsilon)$. We benchmark the ptQMC using a realistic nuclear chiral force with two-body interactions up to N^{2}LO[13,14], for details see Ref. 11. In order to compute ground states of $H = K + V_{2\mathrm{N}} + V_{\mathrm{cou}}$ using ptQMC, the zeroth order Hamiltonian is $H_0 = K + V_0$ and one calculates the energy corrections with respect to $V_C = H - H_0$. V_0 is the non-locally smeared SU(4) interaction from Ref. 15, which is independent of the spin and isospin and captures the essential elements of the nuclear force. In the left panel of Fig. 3 we compare the results obtained using ptQMC with non-perturbative results. We use a periodic box of size $L = 10$ for ^{3}H and $L = 8$ for the other nuclei. The temporal step is $a_t = 1/1000$ MeV^{-1}. For ^{3}H, the system is small enough that we can use exact sparse matrix calculations. For larger nuclei this is not possible, and so we perform fully non-perturbative QMC calculations instead, which result in large error bars due to severe sign problems. For the ^{16}O nucleus, the sign problem sets in so quickly that we cannot find meaningful results for large τ to make a reliable extrapolation. However, the ptQMC calculations are free from sign problems. The corresponding statistical errors are smaller than the size of the symbols in the figure. A sum of decaying exponential functions is used to capture the residual effects of higher energy excitations and the results are extrapolated to $\tau \to \infty$. For all three nuclei, the second order energy corrections are large and essential in reproducing the data. While this might seem contrary to the normal expectation of the perturbative series, this is actually a consequence of the symmetry breaking. As the unperturbed Hamiltonian H_0 respects the SU(4) symmetry, the wave function $|\Psi_0\rangle$ must belong to one of its irreducible representations (*irreps*). The full Hamiltonian breaks the SU(4) symmetry, thus its ground state $|\Psi\rangle$ is a mixture of different SU(4) *irreps*. As is shown in the right panel of Fig. 3, the components of $|\Psi\rangle$ that mixes the SU(4) *irreps* can only be seen in $|\delta\Psi_1\rangle$ or δE_2. Note that this effect is strongest for the one-pion-exchange potential as it breaks both the Wigner-SU(4) and the spin SU(2) symmetries. Having found such a large second order correction, one must examine the perturbative series beyond second order. This can be done using the deuteron as an example. For the deuteron in a small periodic box, an explicit calculation shows that while the second order correction is sizable due to symmetry-breaking perturbations, higher orders beyond second order are small. This is consistent with the findings for heavier nuclei discussed above. For more general calculations, it is also possible that the higher orders have alternate signs and cancel with each other to give a small residual term. The method presented here is free from the sign problem and can be applied to QMC calculations for many-body systems in nuclear physics, condensed matter physics, ultracold atoms, and quantum chemistry.

4 Emergent Geometry of the Carbon Nucleus

The ^{12}C nucleus is under active investigations both experimentally and theoretically, but still there is little consensus about the nature of various excited states, like e.g. the famous Hoyle state, that was also studied in NLEFT[16,17]. There are two main impediments to reaching definitive conclusions about the structure of the low-lying ^{12}C states. The first is the inability to perform calculations that can handle strong multi-particle correlations. The second is the inability to measure the detailed spatial correlations required to determine the intrinsic structure of the twelve-particle wave function. In Ref. 18 both problems were addressed. The full twelve-particle correlations were calculated and a model-independent tomographic projection to determine the intrinsic three-dimensional structure of each nu-

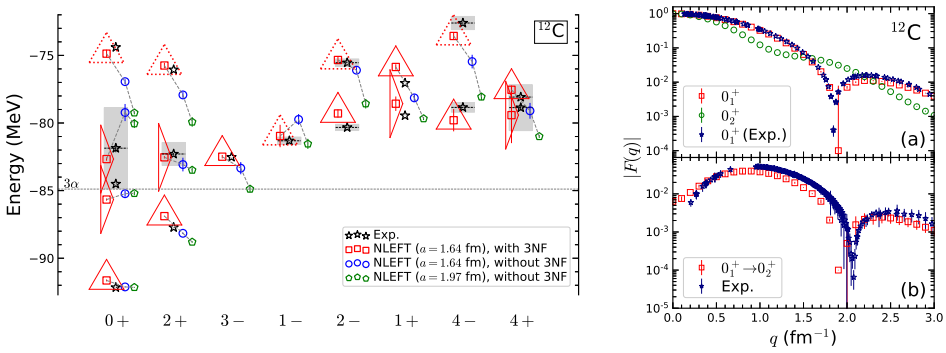


Figure 4. Left Panel: Spectrum of ^{12}C (red squares) in comparison with experimental (black stars). The error bars correspond to one standard deviation errors. The grey shaded regions indicate decay widths for cases where it has been measured. Earlier results for $a = 1.64$ fm (blue circles) and $a = 1.97$ fm (green pentagons) are also shown[20]. The triangular shapes indicate the intrinsic shape of each state, either equilateral or obtuse triangle arrangements of alpha clusters. The dotted lines for some equilateral triangles indicate significant distortions or large-amplitude displacements of the alpha clusters. Right Panel: The absolute value of the charge form factor $F(q)$ versus momentum transfer q. The top figure (a) shows the ground state (red squares) and Hoyle state (green circles), and the bottom figure (b) shows the transition from the ground state to the Hoyle state (red squares). The error bars correspond to one standard deviation errors. Experimental data (purple stars) are shown for comparison[22–25].

clear state was given. In these calculations, a simple SU(4) symmetric interaction was used, as in the description the ground state energies of light and medium-mass nuclei[15] and the thermodynamics of symmetric nuclear matter[2]. It also reproduces the low-energy spectrum of ^{12}C well[20]. This interaction involves four parameters, that are determined from a fit to the groundstate energies of ^{4}He and ^{12}C, the carbon charge radius and and to several electromagnetic transition rates. An assortment of different initial states for each state of ^{12}C is used to verify that the choice of initial state does not affect the final observables. The initial states considered include many different shell model states as well as different geometric configurations of alpha clusters. The pinhole algorithm is used to determine the probability distribution for the nucleon positions, spins, and isospins[3]. For each pinhole configuration, we know the positions of all A nucleons, and thus the position of each nucleon relative to the centre of mass is easily calculated. From this information, we can compute many nuclear observables like the charge density. The charge form factor $F(q)$ is then calculated by Fourier transforming the charge density. We have calculated the ^{12}C spectrum up to excitation energies of about 15 MeV for $a = 1.64$ fm. The results are plotted as red squares in the left panel of Fig. 4 for different values of the angular momentum and parity. For comparison we show the experimental data (black stars)[21] and earlier results, which were performed without three-nucleon forces[20]. Overall, the agreement with the empirical results is quite good. In the right panel of Fig. 4, the form factors for the ground state and the Hoyle state in the top figure (a), and the transition form factor from the ground state to the Hoyle state in the bottom figure (b) are shown compared to the corresponding experimental data. The agreement is fairly good. Using the pinhole algorithm, we can uniquely determine the three alpha clusters in each state. These three alpha clusters define a triangle

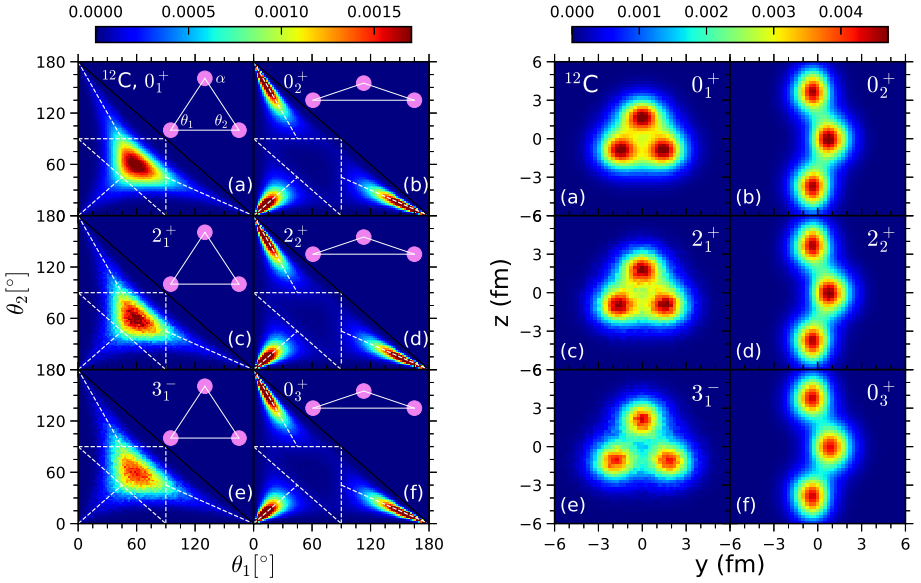


Figure 5. Nuclear density distributions for the (a) 0_1^+ ground state, (b) 0_2^+ Hoyle state, (c) 2_1^+, (d) 2_2^+, (e) 3_1^-, and (f) 0_3^+ states. The red (blue) colour signals a high (low) probability. Left Panel: Density distribution for the two inner angles of the triangle formed by the three alpha clusters. The two axes are for the two inner angles θ_1 and θ_2 measured in degrees. Right Panel: Tomographic projection of the nuclear density for different states of ^{12}C. In each case the orientation of the shortest root-mean-square direction is aligned with the x axis.

in three-dimensional space with interior angles θ_1, θ_2, and $180^\circ - \theta_1 - \theta_2$. In the left panel of Fig. 5, we show the probability distributions as a function of θ_1 and θ_2 for ground state an various excited states. The dashed white triangle formed by the line segments $\theta_1 = 90^\circ$, $\theta_2 = 90^\circ$, and $\theta_2 = 90^\circ - \theta_1$, represents cluster configurations that are right triangles. The interior region of the dashed white triangle corresponds to configurations that are acute triangles, and the exterior region corresponds to obtuse triangles. The other three white dashed line segments along the lines $\theta_1 = \theta_2$, $\theta_1 = \theta_3$, and $\theta_2 = \theta_3$ represent cluster configurations that are obtuse isosceles triangles. For the ground state, the probability distribution is strongly centred around an equilateral triangle, $\theta_1 = \theta_2 = \theta_3 = 60^\circ$. The 2_1^+ and 3_1^- states have similar equilateral triangular shapes. In contrast, the 0_2^+ Hoyle state corresponds to an obtuse isosceles triangle. This finding is consistent with older NLEFT studies[16,17]. The 2_2^+ and 0_3^+ states also have obtuse isosceles triangular shapes. We can also define a model-independent tomographic projection of the three-dimensional nuclear density for the states of ^{12}C, for details see Ref. 18 . In the right panel of Fig. 5, we show the density distribution of selected states of ^{12}C. The 0_1^+, 2_1^+, 3_1^-, 4_1^-, and 4_2^+ states have similar intrinsic equilateral triangular shapes, consistent with an interpretation as members of a rotational band built on top of the 0_1^+ state. The 0_2^+, 2_2^+, 4_1^+ states have similar intrinsic obtuse isosceles triangle shapes and are consistent with belonging to a rotational band built on top of the 0_2^+ state. These findings are consistent with previous studies based on group theoretical considerations[26]. We note that models where the Hoyle state has an equilateral

triangle symmetry predict an additional 3^- and 4^- state in the Hoyle state rotational band. The 0_3^+ state has been discussed as a breathing mode excitation of the Hoyle state[27–29], but its detailed structure remains a matter of debate. For example, in a recent work[27,28] the 0_3^+ and Hoyle states are suggested to have an equilateral triangular shape. A gas-like structure with a very large radius has also been predicted[29] for the 0_3^+ state. Our lattice findings suggest that the 0_3^+ state is a small-amplitude vibrational excitation of the Hoyle state. Our findings for the intrinsic shapes of the low-lying states of ^{12}C are summarised by the triangular shapes in the left panel of Fig. 4. Further studies using more precise interactions are planned.

5 Alpha-Alpha Scattering in the Multiverse

Alpha-alpha (α-α) scattering is one of the most fundamental reactions in nuclear (astro)physics. It features some fine-tuning, as the large near-threshold S-wave results from a state with $(J^P, I) = (0^+, 0)$ at an energy $E_R \simeq 0.1$ MeV above the threshold, see e.g. the review[30], with a tiny width of $\Gamma_R \simeq 6$ eV. It is precisely this small width (long lifetime) of the unstable ^{8}Be nucleus that allows for the reaction with the third α particle in the 3α reaction at sufficiently high temperatures and densities. Here, another fine-tuning appears, namely the closeness of the Hoyle state to the 3α threshold. The fine-tunings in these (and other) fundamental nuclear reactions together with other fine-tunings in particle physics and cosmology have led to the concept of the *Multiverse*, where our Universe with its observed values is part of a larger structure of universes featuring different sets of the fundamental constants. Related to this are anthropic considerations, which is the philosophical idea that the parameters governing our world should fit the intervals compatible with the existence of life on Earth. Coming back to nuclear physics, the closeness of the Hoyle state energy to the 3α threshold invites investigations about the stability of this resonance condition under changes of the fundamental parameters of the strong and the electromagnetic (em) interactions. This was already investigated in models, e.g. Ref. 31, and NLEFT[32–34]. The same chiral EFT at N2LO order combined with the so-called Adiabatic Projection Method (APM), that allows for *ab initio* calculations of nuclear reactions, as developed and applied in Refs. 35–39 can be used to study the sensitivity of the low-energy α-α phase shifts on variations in the light quark mass $\hat{m}$ and the em fine-structure constant $\alpha_{\rm EM}$[40]. While the investigation of the resonance enhancement in the 3α process due to the Hoyle state already sets rather stringent limits on the possible variations of the fundamental parameters, one has to be aware that these results are afflicted with some inherent uncertainties, as in the corresponding stellar simulations only the distance of the Hoyle state to the 3α-threshold is varied. Translating this into a dependence on, say, the light quark mass assumes that only the nuclei directly involved in the 3α process are subject to these changes, but of course one should perform the complete stellar simulations (reaction networks) with appropriately modified masses and reaction rates. At present, this is only possible for Big Bang Nucleosynthesis, but not for the whole nuclear reaction networks in stars. Therefore, the *ab initio* computation of the dependence of α-α scattering on the fundamental parameters of the Standard Model is not subject to such uncertainties and paves the way for more elaborate network calculations in the Multiverse. First, consider the physical values of the various coupling constants and masses (our Universe). Due to improvements in the APM compared to the pioneering study of α-α scattering in

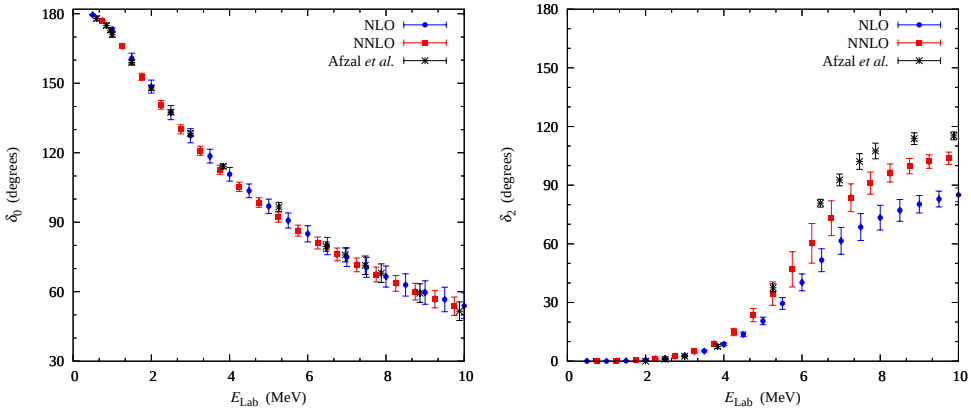


Figure 6. Left panel: S-wave α-α scattering phase shift δ_0 versus the energy in the laboratory system, $E_{\rm lab}$. Right panel: D-wave α-α scattering phase shift δ_2 versus the energy in the laboratory system, $E_{\rm lab}$. The blue circles and red squares represent the predictions at NLO and NNLO, respectively, while the data are given by the black crosses.

Ref. 38, we obtain a very good description of the S- and D-wave phase shifts up to energies $E_{\rm lab} \simeq 10\,$MeV at NNLO in the chiral expansion, see Fig. 6. For the study of the variations under changes of the pion mass with $|\delta M_\pi/M_\pi| \leq 10\%$, we rely on the pion mass dependent nuclear Hamiltonian worked out in Ref. 34. To this order, the ^{8}Be nucleus is slightly bound. In the S-wave phase shift, we find a dramatic effect (unbinding of the two-alpha system) for changes of -5% and -7% at NLO and NNLO, respectively. We have also considered the pion mass variation of the S-wave effective range function, which is less sensitive to the binding issue and shows an added repulsion for negative pion mass shifts. This additional repulsion will certainly impact the position and the lifetime of ^{8}Be. The pion mass variation on the D-wave is somewhat more pronounced, as seen by the effect on the corresponding resonance parameters and also by the D-wave effective range function. The dominant electromagnetic effect on the α-α scattering phase shifts is the long-ranged Coulomb potential that is included exactly by using a spherical wall with Coulomb boundary conditions. Taking this effect into account via the Coulomb-modified effective range expansion, we find very small effects of variations of $\alpha_{\rm EM}$ on the S- and D-wave phase shifts. We have further shown that up-to-and-including NLO in the chiral expansion, the dependence of the α-α scattering phase shifts on the QCD θ-angle is entirely given by the θ-dependence of the pion mass. In summary, we find that α-α scattering sets weaker constraints on the variation of the light quark masses and the fine-structure constant than that given by the closeness of the 3α threshold to the Hoyle state. However, as discussed in detail e.g. in Refs. 31, 34, this requires stellar modelling which introduces some model-dependence. In contrast to that, the investigation of α-α scattering discussed here is truly *ab initio* and not affected by such effects. Still, to further improve these calculations, a better determination of the pion mass dependence of the singlet and triplet NN scattering lengths from lattice QCD is mandatory.

6 Outlook: Towards N3LO Precision

At this time, NLEFT is a well-established *ab initio* framework at the forefront of the theory of nuclear structure and reactions. At present, high precision calculations at N3LO in the chiral counting are being performed. This is based on pinning down a precise leading order interaction as discussed in this contribution with SU(4) invariant interactions that are locally and non-locally smeared. In that way, the higher order corrections and higher few-nucleon forces can be treated in perturbation theory, as also discussed here. The underlying N3LO representation of the chiral two-nucleon force has already been worked out, see Ref. 41, and has been refined by including operators that restore Galilean invariance, see Ref. 42. First preliminary results for nuclei up to $A = 40$ have been reported e.g in Refs. 43, 44 and much more detailed and extended high-precision results not only for ground state energies, but also the spectrum of selected nuclei and the oxygen and calcium drip lines will be reported soon.

Acknowledgements

I thank my NLEFT colleagues for their contributions to the results presented here. This work is supported in part by the DFG (DFG Project ID 196253076 - TRR 110) 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), by VolkswagenStiftung (grant no. 93562), and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101018170, EXOTIC). The computational resources were provided by the Jülich Supercomputing Centre (JSC) at Forschungszentrum Jülich and by RWTH Aachen. Special resources on the JURECA-DC supercomputer at the JSC are particularly acknowledged.

References

1. T. A. Lähde and U.-G. Meißner, "Nuclear Lattice Effective Field Theory: An introduction," Lect. Notes Phys. **957**, 1-396, 2019.
2. B. N. Lu, N. Li, S. Elhatisari, D. Lee, J. E. Drut, T. A. Lähde, E. Epelbaum, and U.-G. Meißner, Phys. Rev. Lett. **125**, 192502, 2020.
3. S. Elhatisari *et al.*, Phys. Rev. Lett. **119**, 222505, 2017.
4. R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278, 1981.
5. N. Byers and C. N. Yang, Phys. Rev. Lett. **7**, 46-49, 1961.
6. E. Y. Loh, Jr. and D. K. Campbell, Synthetic Metals **27**, A499, 1988.
7. R. Valenti, C. Gros, P. J. Hirschfeld, and W. Stephan, Phys. Rev. B **44**, 13203, 1991.
8. G. Hagen, T. Papenbrock, A. Ekström, K. A. Wendt, G. Baardsen, S. Gandolfi, M. Hjorth-Jensen, and C. J. Horowitz, Phys. Rev. C **89**, 014319, 2014.
9. B. Schuetrumpf, W. Nazarewicz, and P. G. Reinhard, Phys. Rev. C **93**, 054304, 2016
10. J. B. Elliott, P. T. Lake, L. G. Moretto, and L. Phair, Phys. Rev. C **87**, 054622, 2013.
11. B. N. Lu, N. Li, S. Elhatisari, Y. Z. Ma, D. Lee, and U.-G. Meißner, [arXiv:2111.14191 [nucl-th]].

12. S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401, 2003.
13. E. Epelbaum, H. W. Hammer, and U.-G. Meißner, Rev. Mod. Phys. **81**, 1773, 2009.
14. P. Reinert, H. Krebs, and E. Epelbaum, Eur. Phys. J. A **54**, 86, 2018.
15. B. N. Lu, N. Li, S. Elhatisari, D. Lee, E. Epelbaum, and U.-G. Meißner, Phys. Lett. B **797**, 134863, 2019.
16. E. Epelbaum, H. Krebs, D. Lee, and U.-G. Meißner, Phys. Rev. Lett. **106**, 192501, 2011.
17. E. Epelbaum, H. Krebs, T. A. Lähde, D. Lee, and U.-G. Meißner, Phys. Rev. Lett. **109**, 252501, 2012.
18. S. Shen, T. A. Lähde, D. Lee, and U.-G. Meißner, [arXiv:2202.13596 [nucl-th]].
19. D. Lee, Prog. Part. Nucl. Phys. **63**, 117, 2009.
20. S. Shen, T. A. Lähde, D. Lee, and U.-G. Meißner, Eur. Phys. J. A **57**, 276, 2021.
21. J. H. Kelley, J. E. Purcell, and C. G. Sheu, Nucl. Phys. A **968**, 71, 2017.
22. I. Sick and J. S. Mccarthy, Nucl. Phys. A **150**, 631, 1970.
23. P. Strehl, Z. Phys. **234**, 416, 1970.
24. H. Crannell, X. Jiang, J. T. O'Brien, D. I. Sober, and E. Offermann, Nucl. Phys. A **758**, 399, 2005.
25. M. Chernykh, H. Feldmeier, T. Neff, P. von Neumann-Cosel, and A. Richter, Phys. Rev. Lett. **105**, 022501, 2010.
26. D. J. Marin-Lambarri, *et al.*, Phys. Rev. Lett. **113**, 012502, 2014.
27. K. C. W. Li, *et al.*, Phys. Rev. C **105**, 024308, 2022.
28. K. C. W. Li, *et al.* Phys. Lett. B **827**, 136928, 2022.
29. B. Zhou, A. Tohsaki, H. Horiuchi, and Z. Ren, Phys. Rev. C **94**, 044319, 2016.
30. S. A. Afzal, A. A. Z. Ahmad, and S. Ali, Rev. Mod. Phys. **41**, 273, 1969.
31. H. Oberhummer, A. Csoto, and H. Schlattl, Science **289**, 88, 2000.
32. E. Epelbaum, H. Krebs, T. A. Lähde, D. Lee, and U.-G. Meißner, Phys. Rev. Lett. **110**, 112502, 2013.
33. E. Epelbaum, H. Krebs, T. A. Lähde, D. Lee, and U.-G. Meißner, Eur. Phys. J. A **49**, 82, 2013.
34. T. A. Lähde, U.-G. Meißner, and E. Epelbaum, Eur. Phys. J. A **56**, 89, 2020.
35. M. Pine, D. Lee, and G. Rupak, Eur. Phys. J. A **49**, 151, 2013.
36. S. Elhatisari and D. Lee, Phys. Rev. C **90**, 064001, 2014.
37. A. Rokash, M. Pine, S. Elhatisari, D. Lee, E. Epelbaum, and H. Krebs, Phys. Rev. C **92**, 054612, 2015.
38. S. Elhatisari, D. Lee, G. Rupak, E. Epelbaum, H. Krebs, T. A. Lähde, T. Luu, and U.-G. Meißner, Nature **528**, 111, 2015.
39. S. Elhatisari, D. Lee, U.-G. Meißner, and G. Rupak, Eur. Phys. J. A **52**, 174, 2016.
40. S. Elhatisari, T. A. Lähde, D. Lee, U.-G. Meißner, and T. Vonk, JHEP **02**, 001, 2022.
41. N. Li, S. Elhatisari, E. Epelbaum, D. Lee, B. N. Lu, and U.-G. Meißner, Phys. Rev. C **98**, 044002, 2018.
42. N. Li, S. Elhatisari, E. Epelbaum, D. Lee, B. Lu, and U.-G. Meißner, Phys. Rev. C **99**, 064001, 2019.
43. N. Li, S. Elhatisari, E. Epelbaum, D. Lee, B. N. Lu, and U.-G. Meißner, PoS **CD2018**, 099, 2019.
44. S. Elhatisari, Springer Proc. Phys. **238**, 585, 2020.

Quantum Many-Body Dynamics from Neural Networks with GPU Acceleration

Markus Schmitt[1], **Heiko Burau**[2], **and Markus Heyl**[3]

[1] Institute for Theoretical Physics, University of Cologne, 50937 Köln, Germany
E-mail: markus.schmitt@uni-koeln.de

[2] Max-Planck-Institut für Physik komplexer Systeme,
Nöthnitzer Straße 38, 01187 Dresden, Germany
E-mail: burau@pks.mpg.de

[3] Theoretical Physics III, Center for Electronic Correlations and Magnetism, Institute of Physics,
University of Augsburg, 86135 Augsburg, Germany
E-mail: markus.heyl@uni-a.de

In recent years the field of nonequilibrium quantum many-body systems has seen a rapid development, in particular, by the remarkable progress in experiments. Today, so-called quantum simulators provide access to quantum matter with an unprecedented control, including especially the possibility to study the dynamics of interacting quantum many-body systems far from equilibrium. The theoretical description of the resulting real-time evolution is, however, facing a severe challenge. While powerful methods have been developed to capture many important aspects of quantum nonequilibrium dynamics in one- or high-dimensional systems, a substantial gap has remained in the physically relevant regime of dimensions two and three. The aforementioned experimental progress as well as a number of open theoretical questions render the development of versatile numerical tools applicable to intermediate dimensions a pressing issue. Here, we report our progress in this direction by means of a novel approach based on a neural network encoding of the quantum many-body wave function. We demonstrate that this approach can outperform alternative state-of-the-art methods in the reached time scales as well as system sizes. Moreover, we could successfully address the long-time dynamics of a strongly disordered spin model, revealing, e.g., new insights into the spatio-temporal build-up of many-body localised spin-glass order.

1 Introduction

Enormous advances in experimental techniques during the past two decades, like the development of quantum simulators of different kinds and ultrafast pump-probe techniques, allow unprecedented control and time-resolved measurements on quantum systems with many interacting degrees of freedom[1–10]. These developments turned the spotlight on a number of fundamental open theoretical questions: Under which conditions and in what sense will a closed system approach a thermal state when prepared far from equilibrium initially? How can the dynamics be characterised on a macroscopic level and what are characteristic time scales? Is there a notion of phases beyond the equilibrium paradigm? Addressing these questions theoretically is a demanding task: For investigations based on microscopic model systems one has to devise strategies to deal with the curse of dimensionality inherent to the quantum many-body problem. Particularly challenging is the development of efficient and versatile computational methods that serve as crucial link between experimental observations and theoretical models.

While powerful methods have been devised to capture many important aspects of quantum non-equilibrium dynamics in one- or high-dimensional systems, a substantial gap has remained in the physically relevant regime of dimensions two and three. Tensor network methods such as Matrix Product States have proven extremely efficient for moderately entangled wave functions, including, in particular, the low energy states in low dimensions[11–13]. However, simulating the non-equilibrium real-time dynamics becomes challenging, especially in two-dimensional systems, due to rapidly increasing entanglement. This bounds the maximally accessible time scales substantially. In the opposite limit of high spatial dimensions non-equilibrium dynamics can be accurately captured by time-dependent Dynamical Mean Field Theory[14–18] . Being exact for lattices with infinite connectivity, DMFT, however, only provides a limited spatial resolution constraining its reliable applicability to lower-dimensional systems.

We follow a novel route by employing artificial neural networks to encode the complex quantum many-body wave function[19]. Thereby, we aim to advance the theoretical description of non-equilibrium quantum matter addressing the challenging regimes of two spatial dimensions and of long-time dynamics. Importantly, we have introduced several algorithmic improvements of essential importance in order to push the frontier making our approach competitive and partially superior to other state-of-the art numerical approaches such as tensor networks[20,21].

2 Neural Quantum States

Their proven merit for pattern recognition and generalisation renders artificial neural networks (ANNs) a promising ansatz for variational wave functions. Wave functions that are constructed from ANNs are called neural quantum states (NQS). Importantly, ANNs become universal function approximators in the limit of large network size[22–24]. This means that any inductive bias that might be present in small ANNs can be systematically removed by increasing the network. Hence, in simulations based on NQS the network size serves as a control parameter and the accuracy can be asserted self-consistently by checking for convergence with increasing size of the NQS; in that sense, NQS simulations are numerically exact.

ANNs in their most basic form are composed as alternating sequences of affine-linear and non-linear transformations of the input data. In the schematic depiction in Fig. 1 the nodes in layer l correspond to a vector of activations $\vec{a}^{(l)}$. The affine-linear map

$$z_i^{(l+1)} = \sum_j W_{ij}^{(l+1)} a_j^{(l)} + b_i^{(l+1)} \tag{1}$$

produces the pre-activations $\vec{z}^{(l+1)}$ of the subsequent layer $l+1$. From these, the activations in layer $l+1$ are obtained through a non-linear function $\sigma(\cdot)$ that is applied element-wise to the pre-activations:

$$a_j^{(l+1)} = \sigma\big(z_i^{(l+1)}\big) \tag{2}$$

To construct a NQS $\psi_\theta(\vec{s})$, this iterative prescription is initialised with the computational basis configuration, $\vec{a}^{(0)} = \vec{s}$, and the final activation corresponds to the resulting logarithmic wave function coefficient, e.g.,

$$\log \psi_\theta(\vec{s}) = a_1^{(L)} + i a_2^{(L)} \ . \tag{3}$$

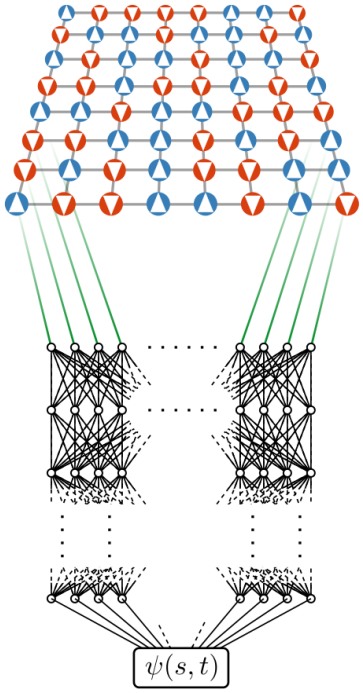


Figure 1. Schematic depiction of a neural quantum state. The artificial neural network maps the computational basis configuration to the corresponding wave function coefficient. Figure taken from Ref. 20.

The weights $W_{ij}^{(l)}$ and biases $b_i^{(l)}$ together are summarised as variational parameters θ. The details how the complex valued wave function coefficient is constructed from the output of the ANN can, however, vary.

In general, there is no restriction on the ANN architectures that can be used to construct NQS and many have been explored, including Restricted Boltzmann Machines, convolutional, recurrent, sparse, or graph neural networks[19,25–28]. Some of these choices incorporate additional structure that can be beneficial for the task, e.g., inherent translational invariance or an autoregressive property that allows for direct sampling of uncorrelated configurations.

3 Time-Dependent Variational Principle

Having established neural quantum states as a new class of variational wave functions, we now turn to the formulation of the variational problem leading to the algorithm that we employ to simulate the dynamics of isolated quantum many-body systems. Fig. 2 illustrates the core idea of the approach. The objective is to solve the time-dependent Schrödinger

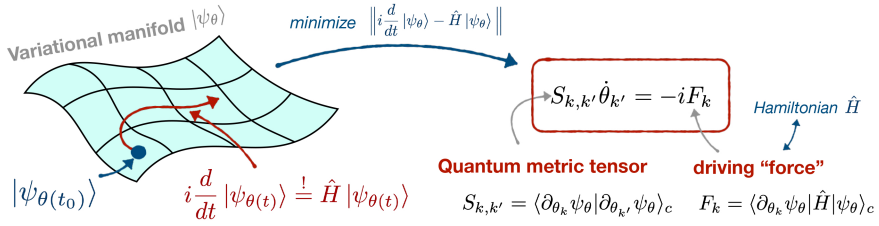


Figure 2. Schematic illustration of the time-dependent variational principle to simulate dynamics of isolated quantum many-body systems using neural quantum states.

equation

$$i\frac{d}{dt}|\psi\rangle = \hat{H}|\psi\rangle \ , \tag{4}$$

where $\hat{H}$ denotes the Hamiltonian that describes the physics of the system of interest. More precisely, we aim to find the time-dependent trajectory of the variational parameters $\theta(t)$ such that $|\psi_\theta\rangle$ best approximates Eq. 4 at each point in time. The minimisation of a distance between the change in the variational wave function and the time evolution generated by the Hamiltonian results in a first-order (non-linear) differential equation for the variational parameters,

$$S_{k,k'}\dot{\theta}_{k'} = -iF_k \ , \tag{5}$$

with the quantum metric tensor $S_{k,k'} = \langle\partial_{\theta_k}\psi_\theta|\partial_{\theta_{k'}}\psi_\theta\rangle$ and the force $F_k = \langle\partial_{\theta_k}\psi_\theta|\hat{H}|\psi_\theta\rangle$.

Given the wave function at time t, $|\psi_{\theta(t)}\rangle$, both $S_{k,k'}$ and F_k can be approximated by a Monte Carlo sampling of the corresponding Born distribution $|\psi_{\theta(t)}(\vec{s})|^2$. Hence, the central steps of the algorithm to compute the time evolution of a variational NQS wave function are

1. Obtain $S_{k,k'}$ and F_k by Monte Carlo sampling of $|\psi_{\theta(t)}(\vec{s})|^2$
2. Solve Eq. 5 for $\dot{\theta}_k$
3. Update the wave function parameters: $|\psi_{\theta(t+\tau)}\rangle = |\psi_{\theta(t)+\tau\dot{\theta}}\rangle$
4. Continue with step 1.

4 Parallel Implementation and Performance

4.1 Parallel Implementation

When time-evolving NQS, a majority of the computational effort is spent on the evaluation of the ANN for given input configurations. This task is, however, ideally suited for a hybrid parallelisation scheme that is depicted in Fig. 3.

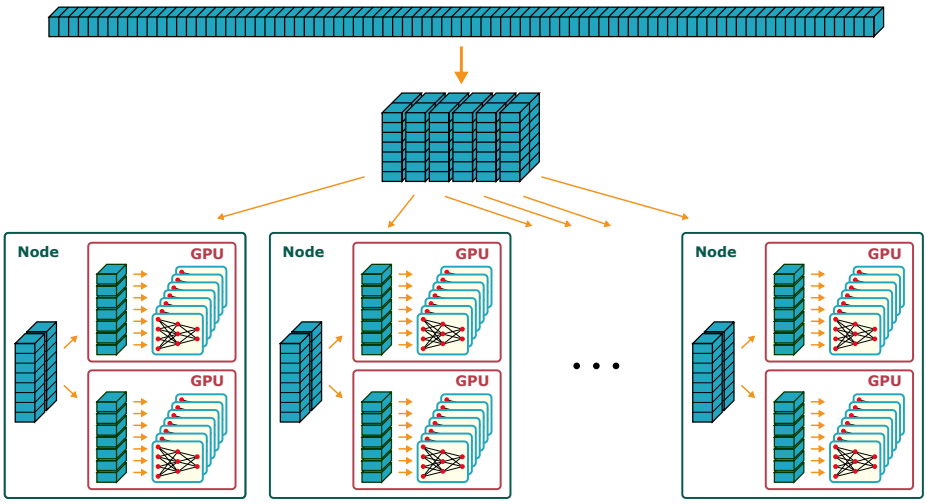


Figure 3. Parallelisation scheme. Each blue box represents the task of evaluating the ANN for a given input configuration. These tasks are distributed across the available nodes and on each node across the available GPUs. Furthermore, the batched evaluation on GPUs yields additional performance gains due to enhanced arithmetic intensity. Figure taken from Ref. 29.

On the one hand, the algorithm comprises large numbers of simultaneous and independent network evaluation tasks: Monte Carlo sampling is an embarrassingly parallel task and in addition the computation of the right hand side of Eq. 5 requires the wave function coefficients for all computational basis states that are connected to the sampled configurations by matrix elements of the Hamiltonian. On this level, the algorithm is amenable to distributed parallelisation.

On the other hand, individual network evaluations mean affine-linear transformations and element-wise application of non-linear functions. These operations are ideally suited for execution on array processors like GPUs. The performance benchmark in the following section will reveal that additional batching of network evaluations can greatly enhance the performance.

Our `jVMC` library[29] provides an efficient implementation of the common building blocks of NQS algorithms. It is based on the `JAX` library[30] for automatic differentiation and just-in-time compilation and it employs an MPI parallelisation for multi node computing.

4.2 Parallel Performance

We benchmarked the performance of the `jVMC` library on the `JUWELS Booster` module, as reported in Ref. 29. The setting for the benchmark is an individual optimisation step as prescribed by Eq. 5 for a transverse-field Ising chain consisting of 50 spin-1/2 degrees of freedom. The NQS architecture is a symmetrised recurrent neural network comprising 1510 variational parameters. For the expectation values $N_{MC} = 3 \times 10^5$ Monte Carlo samples are generated from the NQS by direct sampling.

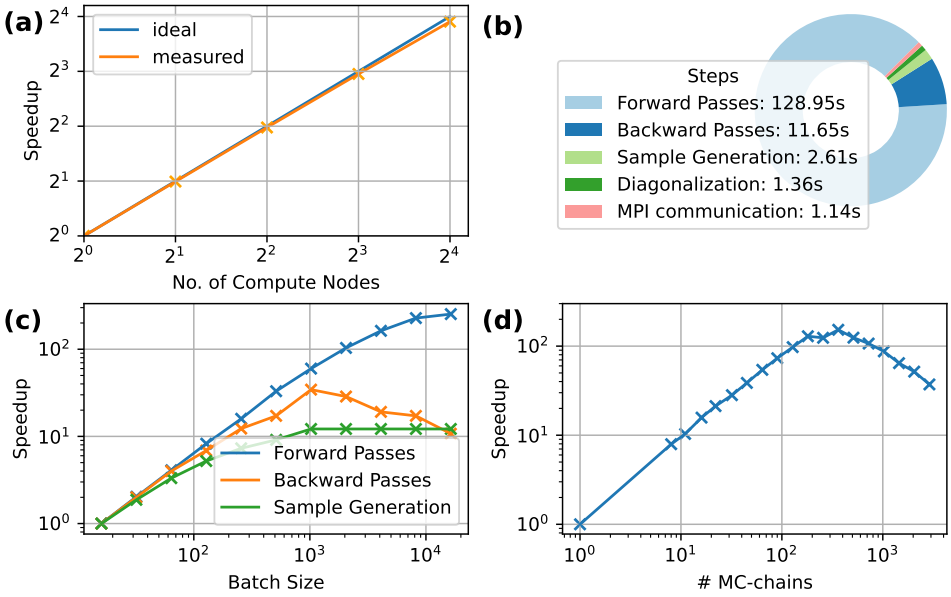


Figure 4. a) Speedup due to the MPI parallelisation across multiple nodes. b) Proportions of compute time contributed by individual parts of the algorithm. c) Effect of batched neural network evaluation. d) Speedup by running multiple Markov chains in parallel on each device for Monte Carlo sampling. Figure taken from Ref. 29.

Fig. 4a) shows the speedup achieved over single node performance as a function of the number of nodes. On `JUWELS Booster` each node has 4 `NVIDIA A100` GPUs, meaning that the maximal number of GPUs used concurrently in this benchmark is 64. The resulting speedup is close to ideal. The reason for this excellent performance of the parallel implementation is shown in Fig. 4b): by far the largest proportion of the total compute time corresponds to forward or backward evaluations of the neural network with no mutual dependencies.

Fig. 4c) shows additional speedup achieved on the level of parallel execution on individual GPUs by exposing additional parallelism through batching of evaluation tasks. The resulting enhancement in arithmetic intensity enables further speedups of up to $100\times$.

When direct sampling of the NQS is not possible, one has to resort to Markov chain Monte Carlo. In that case, multiple Monte Carlo chains can be run in parallel on each GPU. Fig. 4d) displays for a fixed total number of samples the speedup as a function of the number of independent Monte Carlo chains. The timings show a sweet spot where full occupation of the Streaming Multiprocessors is achieved and before computations start to be performed sequentially.

5 Applications

In order to illustrate the performance of the NQS for numerically calculating the real-time dynamics of quantum many-body systems, we will discuss two representative examples.

5.1 Unitary Long-Time Dynamics with NQS for Quantum Renormalisation Group Transformations

Let us first consider an application of NQS to effective descriptions of quantum matter by means of renormalisation group (RG) techniques. RG transformations are essential for the theoretical understanding of the emergent behaviour of many-body systems. By iteratively integrating out degrees of freedom, RG transformations generate effective descriptions for the remaining degrees of freedom. They are traditionally used for the description of the equilibrium properties. When it comes to their dynamics, so-called strong-disorder RGs generate effectively a so-called quantum circuit, which is a sequence of unitary transformations applied to the initial condition, see Fig. 5. Importantly, the calculation of the dynamical properties generated by such a quantum circuit has previously always been relying on further approximation which has limited the quantitative predictive power of RG methods so far. In Ref. 21 we have provide a solution to this challenge by means of NQS through encoding the dynamics generated by the quantum circuit into an artificial neural network. Concretely, we have applied our idea to the disordered quantum Ising model, which is one of the standard models for the strong-disorder RG:

$$H = \sum_{\langle ij \rangle} \left[J_{ij} \sigma_i^z \sigma_j^z + J_{ij}^{(x)} \sigma_i^x \sigma_j^x \right] + \sum_i h_i \sigma_i^x \,, \tag{6}$$

with $J_{ij} \in [-J, J]$, $J_{ij}^{(x)} \in [-J^{(x)}, J^{(x)}]$, and $h_i \in [-h, h]$ drawn randomly from uniform distributions. Here, σ_i^α with $\alpha = x, y, z$ denote the Pauli matrices on lattice site i. For sufficiently strong disorder in the spin coupling J_{ij}, i.e., $J \gg J^{(x)}, h$, this system is entering a so-called MBL spin glass ordered phase where each eigenstate $|E\rangle$ exhibits a nonvanishing Edwards-Anderson order parameter $\chi_{EA}(E) = N^{-2} \sum_{ij} |\langle E|\sigma_i^z \sigma_j^z|E\rangle|^2 > 0$. On a physical level the main result our work has been the theoretical calculation of the spatiotemporal buildup of such many-body localised spin-glass order (see Fig. 5), which has remained elusive up to this point.

Concretely, we were interested in the following physical question. Preparing the system in a symmetric initial condition, which doesn't break the $\mathbb{Z}_2$ symmetry of the MBL spin glass phase. How does the system build up the long-range order? We were addressing

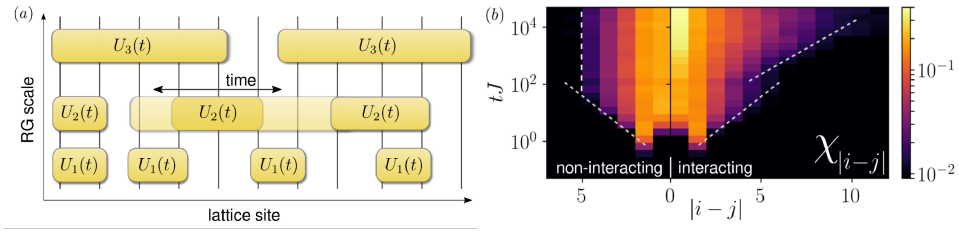


Figure 5. (a) Schematic illustration of the quantum circuit which results from a strong-disorder renormalisation group transformation. (b) Spatiotemporal buildup of long-range many-body localised spin-glass order ass quantified by $\chi_{|i-j|}$ comparing a noninteracting limit of the underlying model to the generic interacting case. Due to our machine learning approach we can reach very long times which allows us to identify a long-time regime where the buildup of MBL spin glass order comes to rest for the noninteracting case but further develops for the interacting model. Figure taken from Ref. 21.

this question by taking as an initial condition $|\psi_0\rangle = |+\rangle$, which is the eigenstate of all σ_l^x operators with eigenvalue $+1$. Studying the subsequent dynamics generated by the above Hamiltonian in the regime of J_{ij}, i.e., $J \gg J^{(x)}, h$ we observe a slow buildup of spatiotemporal order by means of our machine learning algorithm, see Fig. 5. While for a noninteracting limit $J^{(x)} = 0$ this buildup comes to stop at some point in time, for the full interacting system we find that this buildup continues in a second stage with a lightcone that is growing logarithmically.

5.2 Dynamical Universality in 2D Quantum Matter

A further highlight of the utilised NQS approach has been our contribution to unravel real-time universal behaviour of 2D quantum matter upon dynamically crossing a quantum phase transition. While the theory of the so-called quantum Kibble-Zurek mechanism (QKZM) predicts this universal behaviour on a general level, it has remained a theoretical challenge to verify and test the prediction in interacting quantum matter beyond one dimension. In a combined numerical effort encompassing various state-of-the art computational approaches including the NQS we have been able to show dynamical universality in line with the QKZM[31]. Concretely, we have considered the 2D quantum Ising model

$$H = -J \sum_{\langle ij \rangle} \sigma_i^z \sigma_j^z - h \sum_i \sigma_i^x \,, \tag{7}$$

on a 2D square lattice which is both paradigmatic on a theoretical level and of concrete experimental relevance in so-called Rydberg atom arrays. At equilibrium, the model exhibits

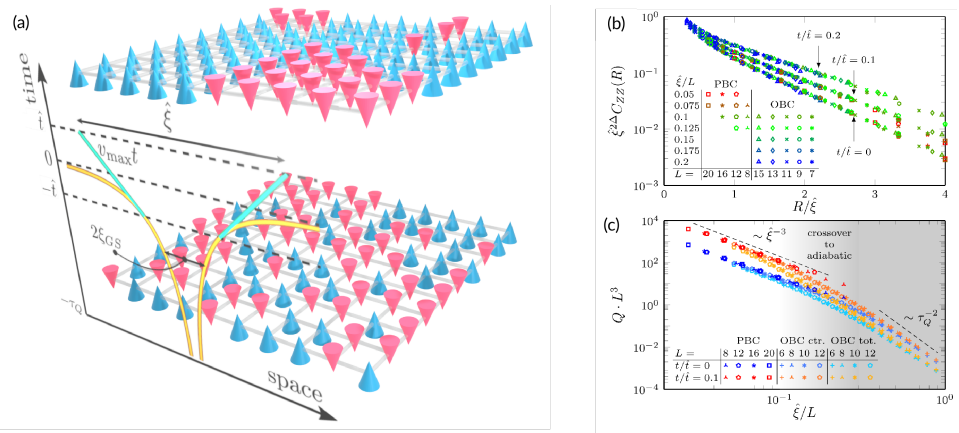


Figure 6. (a) Schematic illustration of the quantum Kibble-Zurek mechanism where as a function of time the system is sweeped across a phase transition. Due to a slow-down of the dynamics in the vicinity of the critical point the growth of the correlation length (indicated by the yellow line) is not capable to keep up with the maximum velocity (cyan). (b) Rescaled ferromagnetic spin correlation function in the vicinity of the quantum critical point exhibiting a compelling data collapse. The data labelled by PBC is obtained by our NQS approach and compared to other data by matrix-product states labelled by OBC. As one can see the NQS approach allows us to reach significantly larger system sizes L which is crucial for our key numerical observation of a power-law behaviour for the excitation energy density Q (c). Figure taken from Ref. 31.

a quantum phase transition in its ground state at $g_c/J = 3.04438$ separating a ferromagnetic phase for $g < g_c$ from a paramagnetic one for $g > g_c$.

Within the QKZM we are interested in dynamically "sweeping" across the quantum phase transition by making the couplings in the Hamiltonian explicitly time dependent

$$g(t)/g_c = 1 - \epsilon(t), \quad J(t)/J_c = 1 + \epsilon(t), \quad \epsilon(t) = \frac{t}{\tau_Q} \,, \tag{8}$$

with $-\tau_Q < t < \tau_Q$ and consequently τ_Q denoting the temporal duration of the sweep. Computing the dynamics with NQS but also with other methods we find that some physical properties obey scaling behaviour as predicted by the QKZM, see Fig. 6. This includes on the one hand the excitation energy density $Q = \frac{1}{L^2}\left(\langle H(t)\rangle - E_0(t)\right)$ where $E_0(t)$ refers to the instantaneous ground state energy during the sweep On the other hand we also consider the ferromagnetic spin correlation function $C^{zz}(t, R) = \langle\psi(t)|\sigma^z_m\sigma^z_n|\psi(t)\rangle$ where R denotes the real-space distance between the spins m, n. As one can see from Fig. 6 the data obtained for both quantities exhibit a compelling collapse after proper rescalings with system size L or the emergent correlation length $\hat{\xi}$ as predicted by the QKZM. Our work has for the first time demonstrated the QKZM in interacting 2D quantum matter.

6 Concluding Remarks

In this contribution we have summarised the developments of our machine learning approach to solve for the quantum many-body dynamics in interacting two-dimensional systems. We have described both the basic algorithmic idea of the technique as well as illustrated the power of the approach for two specific examples. In the future we aim for developing further the underlying codebase to make it generally usable and even more performant so as to be able to address even more challenging problems.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 853443), and M. H. further acknowledges support by the Deutsche Forschungsgemeinschaft via the Gottfried Wilhelm Leibniz Prize program. MS was supported through the Leopoldina Fellowship Programme of the German National Academy of Sciences Leopoldina (LPDS 2018-07). Parts of the numerical simulations were performed at the Max Planck Computing and Data Facility in Garching. Moreover, 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. M. Greiner, O. Mandel, T. W. Hansch, and I. Bloch, *Collapse and revival of the matter wave field of a Bose-Einstein condensate*, Nature, **419**, 51–54, 2002.

2. T. Kinoshita, T. Wenger, and D. S. Weiss, *A quantum Newton's cradle*, Nature, **440**, 900–903, 2006.
3. I. M. Georgescu, S. Ashhab, and F. Nori, *Quantum simulation*, Rev. Mod. Phys., **86**, 153–185, Mar 2014.
4. J.-y. Choi, S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, *Exploring the many-body localization transition in two dimensions*, Science, **352**, no. 6293, 1547–1552, 2016.
5. M. Mitrano, A. Cantaluppi, D. Nicoletti, S. Kaiser, A. Perucchi, S. Lupi, P. Di Pietro, D. Pontiroli, M. Riccò, S. R. Clark, D. Jaksch, and A. Cavalleri, *Possible light-induced superconductivity in K3C60 at high temperature*, Nature, **530**, 461–464, 2016.
6. P. Jurcevic, H. Shen, P. Hauke, C. Maier, T. Brydges, C. Hempel, B. P. Lanyon, M. Heyl, R. Blatt, and C. F. Roos, *Direct Observation of Dynamical Quantum Phase Transitions in an Interacting Many-Body System*, Phys. Rev. Lett., **119**, 080501, Aug 2017.
7. C. Gross and I. Bloch, *Quantum simulations with ultracold atoms in optical lattices*, Science, **357**, no. 6355, 995–1001, 2017.
8. H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, *Probing many-body dynamics on a 51-atom quantum simulator*, Nature, **551**, 579, 2017.
9. J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, *Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator*, Nature, **551**, 601, 2017.
10. E. A. Martinez, C. A. Muschik, P. Schindler, D. Nigg, A. Erhard, M. Heyl, P. Hauke, M. Dalmonte, T. Monz, P. Zoller, and R. Blatt, *Real-time dynamics of lattice gauge theories with a few-qubit quantum computer*, Nature, **534**, 516, 2016.
11. S. R. White, *Density matrix formulation for quantum renormalization groups*, Phys. Rev. Lett., **69**, 2863–2866, Nov 1992.
12. U. Schollwöck, *The density-matrix renormalization group in the age of matrix product states*, Annals of Physics, **326**, no. 1, 96–192, 2011, January 2011 Special Issue.
13. R. Orús, *A practical introduction to tensor networks: Matrix product states and projected entangled pair states*, Annals of Physics, **349**, 117–158, 2014.
14. W. Metzner and D. Vollhardt, *Correlated Lattice Fermions in $d = \infty$ Dimensions*, Phys. Rev. Lett., **62**, 324–327, Jan 1989.
15. A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions*, Rev. Mod. Phys., **68**, 13–125, Jan 1996.
16. D. Vollhardt, *Dynamical mean-field theory for correlated electrons*, Annalen der Physik, **524**, no. 1, 1–19, 2012.
17. J. K. Freericks, V. M. Turkowski, and V. Zlatić, *Nonequilibrium Dynamical Mean-Field Theory*, Phys. Rev. Lett., **97**, 266408, Dec 2006.
18. H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, *Nonequilibrium dynamical mean-field theory and its applications*, Rev. Mod. Phys., **86**, 779–837, Jun 2014.
19. G. Carleo and M. Troyer, *Solving the quantum many-body problem with artificial neural networks*, Science, **355**, no. 6325, 602–606, 2017.

20. M. Schmitt and M. Heyl, *Quantum many-body dynamics in two dimensions with artificial neural networks*, accepted in Phys. Rev. Lett., 2019, arXiv:1912.08828, 2019.
21. H. Burau and M. Heyl, *Unitary Long-Time Evolution with Quantum Renormalization Groups and Artificial Neural Networks*, Phys. Rev. Lett., **127**, no. 5, 050601, July 2021.
22. G. Cybenko, *Approximation by superpositions of a sigmoidal function*, Mathematics of Control, Signals, and Systems, **2**, no. 4, 303–314, Dec 1989.
23. K. Hornik, *Approximation Capabilities of Multilayer Feedforward Networks*, Neural Netw., **4**, no. 2, 251–257, Mar. 1991.
24. N. Le Roux and Y. Bengio, *Representational power of restricted boltzmann machines and deep belief networks*, Neural Comput., **20**, no. 6, 1631–49, 2008.
25. K. Choo, T. Neupert, and G. Carleo, *Two-dimensional frustrated J_1-J_2 model studied with neural network quantum states*, Phys. Rev. B, **100**, no. 12, 125124, Sept. 2019.
26. M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, *Recurrent neural network wave functions*, Phys. Rev. Research, **2**, 023358, Jun 2020.
27. S. Pilati and P. Pieri, *Simulating disordered quantum Ising chains via dense and sparse restricted Boltzmann machines*, Phys. Rev. E, **101**, 063308, Jun 2020.
28. L. Yang, W. Hu, and L. Li, *Scalable variational Monte Carlo with graph neural ansatz*, 2020, arXiv:2011.12453.
29. M. Schmitt and M. Reh, *jVMC: Versatile and performant variational Monte Carlo leveraging automated differentiation and GPU acceleration*, 2021, arXiv:2108.03409.
30. J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, "JAX: composable transformations of Python+NumPy programs", `http://github.com/google/jax`, 2018.
31. M. Schmitt, M. M. Rams, J. Dziarmaga, M. Heyl, and W. H. Zurek, *Quantum phase transition dynamics in the two-dimensional transverse-field Ising model*, arXiv e-prints, June 2021, arXiv:2106.09046.

Cross-Correlators of Conserved Charges from Lattice QCD

Szabolcs Borsányi[1], Zoltán Fodor[1,2,3,4], Jana N. Günther[1], Paolo Parotto[3], and Attila Pásztor[2]

[1] Theoretical Particle Physics, University of Wuppertal, 42119, Wuppertal Germany
E-mail: {borsanyi, fodor, jguenther}@uni-wuppertal.de

[2] Theoretical Physics, Eötvös University, H-1111, Budapest, Hungary
E-mail: apasztor@bodri.elte.hu

[3] Pennsylvania State University, Department of Physics, University Park, PA 16802, USA
E-mail: paolo.parotto@gmail.com

[4] Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany

The grand canonical description of quark matter produced in a heavy ion collision experiment at RHIC (Brookhaven, USA) and LHC (CERN), is investigated with computer simulations of the underlying quantum field theory, Quantum Chromodynamics (QCD). The production of hadrons in a collision event is determined by the thermodynamic conditions (temperature, chemical potential) of the system, and is remarkably well described by equilibrium thermodynamics. Fluctuations and correlations in hadron production are calculated with lattice QCD methods with an unprecedented level of detail. We study which observables are best suited for a comparison between theory and experiment, and we exploit the simulation results to investigate repulsive and attractive interactions between hadrons. Matching our results to experimental results from e.g., the Beam Energy Scan program at RHIC, one can map the final stage of the system's evolution on the QCD phase diagram.

1 Introduction

The purpose of this project is to provide a quantitative description of Quark Gluon Plasma (QGP) thermodynamics, at the stage when it freezes back into ordinary matter, like protons and neutrons and other hadrons. The Quark Gluon Plasma was formed in the Early Universe (10^{-12} s to 10^{-6} s after the Big Bang) and is routinely produced in heavy ion collision experiments at the Large Hadron Collider (LHC) at CERN and at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Lab (BNL), as well as the future FAIR facility at GSI (Germany).

At a temperature of about $2 \cdot 10^{12}$ K matter undergoes a transition and a liquid-like phase emerges, the QGP. Below the transition temperature, matter is best described as a gas of subatomic particles, the so-called hadrons. From experimental evidence, it was learned that the QGP is a fluid with extremely low viscosity (relative to its entropy)[1]. When created in such experiments, it expands until it disintegrates into a wide range of particles, which are then detected and identified with a detector-specific efficiency. The production of the emitted hadrons can be described quantitatively by means of the temperature and chemical potential of the system, which are customarily called freeze-out parameters[2]. The resulting freeze-out temperature is expected to be close to the QCD transition temperature[3].

The QCD transition between the confined and QGP phases has been intensely studied by theorists with several approaches, such as Dyson-Schwinger equations[4], functional

renormalisation group methods[5], and, most systematically, using lattice QCD simulations. Much is known about the thermodynamics of QCD at zero net baryon density, such as the temperature of the crossover transition[6–9] and the equation of state[10–13].

The confined phase of QCD is often assumed to be well described by the ideal hadron resonance gas (HRG) model[14–18]. The HRG model is based on the assumption that a gas of interacting hadrons can be described as a gas of non-interacting hadrons and resonances. The inclusion of the resonances as free particles is an approximate way of taking into account resonant interactions between the stable hadrons[16,17]. The model describes bulk thermodynamic observables - like the pressure or the energy density - obtained from first principle lattice calculations extremely well at zero chemical potential[19–21,11], below the transition temperature.

Baryon number (B), electric charge (Q) and strangeness (S) are conserved in QCD, and, in a grand canonical description, a chemical potential is associated with each. The parameters μ_B, μ_S and μ_Q allow to reflect the net expectation value of baryon number, charge and strangeness present in the experimental setup, respectively. It is relatively straightforward to include the chemical potentials in hadron resonance gas based computations, which can predict hadron yields in collision experiments. After matching these predictions with experimental results, a freeze-out curve can be drawn in the QCD phase diagram[22].

The inclusion of the chemical potentials in lattice QCD is substantially more complicated, if not impossible, due to the infamous sign problem. Although simulations with finite chemical potentials are being produced at high computational expense[23], large scale simulations are possible only at zero or imaginary values of the chemical potentials. Even with this restriction, the thermodynamics of QCD could be explored in detail at least for $\mu \lesssim T$. For example,the QCD transition line was calculated by several groups with high precision[24–27]. The location of the hypothetical critical end-point is still unknown.

The study of fluctuations of conserved charges (B, Q and S) yields a set of complementary observables that can be studied both theoretically and experimentally. One considers the event-by-event fluctuations of identified particles, produced in a sub-system at the time of the freeze-out. In recent years, the STAR Collaboration has published results for the fluctuations of net-proton[28], net-charge[29], net-kaon[30] and net-Λ[31]. Contrary to hadron yields, fluctuations of B, Q and S can be also deduced from a grand canonical ensemble on the theory side. On the lattice, these quantities can be calculated as second derivatives of the pressure with respect to the respective chemical potentials[32,33]. Mixed derivatives are defined analogously, and are related to cross-correlators between hadron species[34].

The existence of such well defined observables raised the hope for direct comparisons between theory and experiment. However, the experimental figures will not exactly match the theoretical expectations without further corrections. First, the assumption of full thermal equilibrium is likely not granted, and a local equilibrium may or may not be sufficient to match the grand canonical predictions. Another issue is that the system volume itself fluctuates on an event-by-event basis[35], distorting the measured fluctuations. Furthermore, canonical effects need to be taken into account[36], as the different charges are conserved globally, and a grand canonical average in a sub-system would be exact only in the thermodynamic limit. Kinematic cuts of the measured hadrons also modify the results[37], and are difficult to reproduce on the lattice. Finally, since not all particles are identified, the net B, Q and S quantum numbers need to be estimated from a small subset of measurable particle species.

The study of cross-correlators opens new channels to match theory with experiment. We studied[34] various combinations of decay products and presented predictions from lattice QCD simulations. The extrapolation to finite baryon density, where the STAR experiment operates, was undertaken[38] using a two-dimensional fugacity expansion. In this contribution we summarise the findings presented in these papers.

2 QCD as a Grand Canonical Ensemble

In a canonical ensemble, the conserved charges are external parameters. In a grand canonical ensemble, the chemical potentials μ_B, μ_Q and μ_S are introduced to allow fluctuations of the conserved charges B,Q and S. They enter the QCD Lagrangian as linear combinations in the individual quark chemical potentials μ_u, μ_d, μ_s, for up, down and strange quarks, respectively:

$$\begin{aligned} \mu_u &= \frac{1}{3}\mu_B + \frac{2}{3}\mu_Q \,, \\ \mu_d &= \frac{1}{3}\mu_B - \frac{1}{3}\mu_Q \,, \\ \mu_s &= \frac{1}{3}\mu_B - \frac{1}{3}\mu_Q - \mu_S \,. \end{aligned} \tag{1}$$

Fluctuations of the conserved charges can be expressed as derivatives of the grand canonical thermodynamic potential (also called susceptibilities), which simplifies to the pressure in infinite volume:

$$\chi^{BQS}_{ijk} = \frac{\partial^{i+j+k} p/T^4}{(\partial\hat{\mu}_B)^i (\partial\hat{\mu}_Q)^j (\partial\hat{\mu}_S)^k} \,, \tag{2}$$

where we introduced the $\hat{\mu} = \mu/T$ notation.

For example, the expectation value of Q in a given volume V is given by

$$\langle \delta Q \rangle = VT^3 \chi^Q_1 \,, \tag{3}$$

and its variance reads:

$$\left\langle \delta Q^2 \right\rangle = VT^3 \chi^Q_2 \,. \tag{4}$$

In a fixed volume, the unknown factors V, T cancel in fluctuation ratios such as

$$\begin{aligned} S\sigma &= \chi_3/\chi_2 \,, & \kappa\sigma^2 &= \chi_4/\chi_2 \,, \\ M/\sigma^2 &= \chi_1/\chi_2 \,, & S\sigma^3/M &= \chi_3/\chi_1 \,, \end{aligned} \tag{5}$$

which relate the susceptibilities of Eq. 2 to the moments (mean M, variance σ^2, skewness S, kurtosis κ) of net-particle distributions. Cross-correlations are defined analogously, e.g. for baryon-strangeness we have

$$\langle \delta B \delta S \rangle = VT^3 \chi^{BS}_{11} \tag{6}$$

Lattice QCD can compute these quantities in terms of quark derivatives (see Eq. 1), e.g.

$$\chi_{11}^{BQ} = \frac{1}{9}\left[\chi_2^u - \chi_2^s - \chi_{11}^{us} + \chi_{11}^{ud}\right] , \tag{7}$$

$$\chi_{11}^{BS} = -\frac{1}{3}\left[\chi_2^s + 2\chi_{11}^{us}\right] , \tag{8}$$

$$\chi_{11}^{QS} = \frac{1}{3}\left[\chi_2^s - \chi_{11}^{us}\right] . \tag{9}$$

In the actual simulations, the up and down quarks are degenerate, which introduce simplifications through the equations $\chi_{11}^{us} = \chi_{11}^{ds}$ and $\chi_2^u = \chi_2^d$.

The picture is more complicated, when an extrapolation to finite chemical potential is required, which is typically the case when theoretical expectations for STAR data are calculated.

One popular method to extend the reach of lattice QCD simulations is the Taylor expansion. In the presence of μ_B and μ_S chemical potentials, such an expansion is written, for example for the χ_{11}^{BQ} coefficient, as:

$$\begin{aligned}\chi_{11}^{BQ}(T,\hat{\mu}_B,\hat{\mu}_S) = \chi_{11}^{BQ}(T,0,0) \\ +\frac{1}{2}\hat{\mu}_B^2\chi_{31}^{BQ}(T,0,0) + \hat{\mu}_S\hat{\mu}_B\chi_{211}^{BQS}(T,0,0) + \frac{1}{2}\hat{\mu}_S^2\chi_{112}^{BQS}(T,0,0) \\ +\mathcal{O}(\hat{\mu}^4) .\end{aligned} \tag{10}$$

Notice the absence of the odd terms in the above expression, which is a consequence of C symmetry. The chemical potentials are further constrained by the strangeness neutrality condition $\left\langle \chi_1^S \right\rangle \equiv 0$, which reflects the zero net number of strange quarks in the colliding nuclei. In terms of susceptibilities one has:

$$\chi_1^S(T,\hat{\mu}_B,\hat{\mu}_S) = \hat{\mu}_B\chi_{11}^{BS}(T,0,0) + \hat{\mu}_S\chi_2^S(T,0,0) + \mathcal{O}(\hat{\mu}^4) = 0 , \tag{11}$$

which in turn fixes the relationship between μ_S and μ_B.

An alternative expansion scheme can be defined through the fugacity expansion. The QCD pressure is expressed as a Laurent series in $e^{\hat{\mu}}$. Due to C symmetry, the series has identical coefficients for the $e^{\hat{\mu}}$ and $e^{-\hat{\mu}}$, thus reducing to a series to an expansion in hyperbolic cosines:

$$P(T,\hat{\mu}_B,\hat{\mu}_S) = \sum_{j,k} P_{jk}^{BS}(T)\cosh(j\hat{\mu}_B - k\hat{\mu}_S) . \tag{12}$$

The coefficients P_{jk}^{BS} are also called fugacity expansion or sector coefficients, alluding to the fact that they get contributions from the Hilbert subspace corresponding to the fixed values of the conserved charges $B = j$ and $S = k$. In the ideal HRG model, the expansion coefficients $P_{00}^{BS}, P_{01}^{BS}, P_{10}^{BS}, P_{11}^{BS}, P_{12}^{BS}, P_{13}^{BS}$ all get sizeable contributions from known hadrons and hadron resonances. In the Boltzmann approximation to the ideal HRG, coefficients like P_{20}^{BS} are zero, while in the full HRG they are non-zero, but very small in magnitude and essentially negligible.

At purely imaginary chemical potentials $\mu_q = i\mu_q^I$, where the sign problem is absent and lattice simulations can be performed, Eq. 12 reduces to a Fourier expansion of the

form:

$$P(T,\hat{\mu}_B^I,\hat{\mu}_S^I) = \sum_{j,k} P_{jk}^{BS}(T)\cos(j\hat{\mu}_B^I - k\hat{\mu}_S^I)\ . \tag{13}$$

3 Lattice Results on Cross-Correlations

We used the 4stout improved staggered action[33] to simulate diagonal and off-diagonal fluctuation coefficients. For the continuum extrapolation we performed simulations on $32^3 \times 8$, $40^3 \times 10$, $48^3 \times 12$, $64^3 \times 16$ as well as on $80^3 \times 20$ lattices. Thus, our study covers the QCD transition region in a broad range of lattice spacings from $a = 0.17$ fm to $a = 0.06$ fm. We focus on the temperature range relevant for the chemical freeze-out $T = 140 - 165$ MeV. The physical scales are set using the pion decay constant $f_\pi = 130.41$ MeV. The continuum extrapolated results must not depend on the details of the scale setting, and we check this by also using an alternative scale setting[41] (w_0) and including any resulting deviations in the systematic error of our results.

For the simulations we use the standard Hybrid Monte Carlo algorithm with multi-scale integrator in the force-gradient scheme. We compute the fluctuations by taking lattice measurements on every tenth configuration. Before each such measurement we project out 256 eigenvectors to accelerate the conjugate gradient solver. These solvers make up for about 90% of the run-time for the measurement. The measurements are configured to use the same amount of resources as used to generate the lattice configurations.

We show our results on the cross-correlation χ_{11}^{BS} at zero chemical potential in Fig. 1. Analogous plots for the other combinations can be found in Ref. 34.

The cross-correlators were extrapolated to real chemical potentials to match STAR's experimental conditions. As discussed in Sec. 2, this can be done in different ways, e.g. using the Taylor or fugacity/sector expansions.

For both expansions we performed simulations not only at $\mu = 0$, but also at imaginary values of the chemical potentials. Higher order μ-derivatives for the higher order Taylor

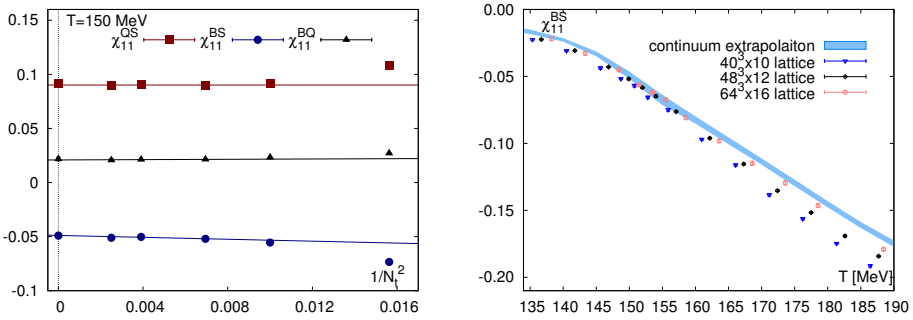


Figure 1. Left: Example for the continuum extrapolation. We show the three cross correlators on the lattices (from right to left): $32^3 \times 8$, $40^3 \times 10$, $48^3 \times 12$, $64^3 \times 16$ and $80^3 \times 20$. The data points correspond to the w_0-based scale setting[41] at $T = 150$ MeV. The error bars in the continuum limit are obtained from the combination of the scale setting, the interpolation, the selection of the continuum extrapolation fit range, and whether a linear or 1/linear function is fitted. Right: Continuum extrapolated result for $\chi_{11}^{BS}(T)$ as a function of the temperature.

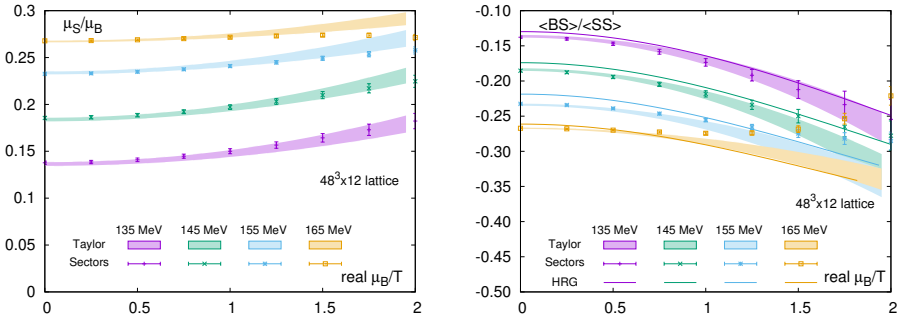


Figure 2. Comparison of two approaches to the finite density extrapolation of two observables. The Taylor result is truncated such that only the leading $\sim \hat{\mu}_B^2$ contribution is considered. In the sector method, contributions up to $|B| = 2$ are included. (The data were generated from a $48^3 \times 12$ lattice, the plots show the intermediate result before continuum extrapolation.) We also show the hadron resonance gas model prediction.

coefficients were obtained from the observed functional behaviour of e.g. χ_1^B, χ_2^S and χ_{11}^{BS} as a function of μ_B. A comparison of the two extrapolations on one of our lattices is shown in Fig. 2.

4 Fugacity Expansion in Two Dimensions

In a subsequent, large scale campaign we collected more data to constrain the fugacity coefficients. For three of our lattice spacings ($N_\tau = 8, 10$ and 12) we simulated a two dimensional mesh of parameters with the imaginary chemical potentials μ_B^I and μ_S^I. The chemical potentials took the values $(\mu_B^I, \mu_S^I) = \frac{\pi}{8}(i, j)$, with $i = 0, 1, \ldots, 15$ and $j = 0, 1, \ldots, 9$, giving a total of $9 \times 16 = 144$ simulation points per temperature. These points are shown in purple in Fig. 3. We studied four different temperatures $T = 145\ \text{MeV}, 150\ \text{MeV}, 155\ \text{MeV}$ and $160\ \text{MeV}$.

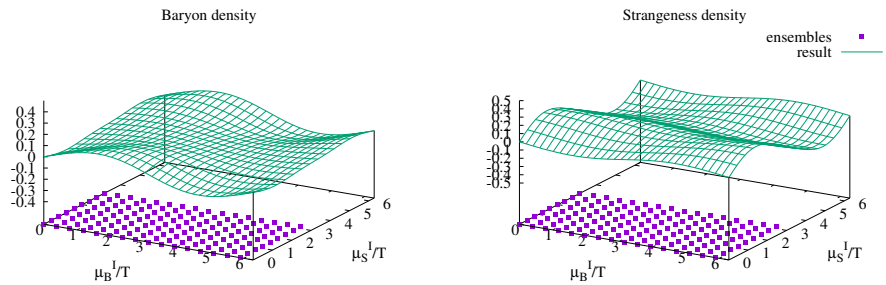


Figure 3. Imaginary baryon (left) and strangeness (right) densities on a 2D mesh of imaginary chemical potentials in μ_S and μ_B. The shown fitted surface includes to the leading eight coefficients in the 2D fugacity expansion. The purple points show the values of $\hat{\mu}_B, \hat{\mu}_S$ where simulations were carried out.

For each temperature T and chemical potential pair $\hat{\mu}_B, \hat{\mu}_S$ in the mesh, low-order susceptibilities are calculated:

$$\begin{aligned}
\mathrm{Im}\,\chi_1^B &= \textstyle\sum_{j,k} j P_{jk}^{BS}(T) \sin(j\hat{\mu}_B^I - k\hat{\mu}_S^I) && (14)\\
\mathrm{Im}\,\chi_1^S &= \textstyle\sum_{j,k} (-k) P_{jk}^{BS}(T) \sin(j\hat{\mu}_B^I - k\hat{\mu}_S^I) && (15)\\
\chi_2^B &= \textstyle\sum_{j,k} j^2 P_{jk}^{BS}(T) \cos(j\hat{\mu}_B^I - k\hat{\mu}_S^I) && (16)\\
\chi_{11}^{BS} &= \textstyle\sum_{j,k} (-jk) P_{jk}^{BS}(T) \cos(j\hat{\mu}_B^I - k\hat{\mu}_S^I) && (17)\\
\chi_2^S &= \textstyle\sum_{j,k} k^2 P_{jk}^{BS}(T) \cos(j\hat{\mu}_B^I - k\hat{\mu}_S^I)\ . && (18)
\end{aligned}$$

A combined fit over all points yields the coefficients P_{ij}^{BS} at each temperature. We truncated the 2D fugacity expansion defined in Eq. 13 to include the terms.

$$\begin{aligned}
&P_{01}^{BS}, \quad P_{10}^{BS}, P_{11}^{BS}, \quad P_{12}^{BS},\\
&P_{13}^{BS}, \quad P_{1,-1}^{BS}, P_{20}^{BS}, \quad P_{21}^{BS},\\
&P_{22}^{BS}, \quad P_{23}^{BS}, P_{02}^{BS}, \quad P_{03}^{BS}.
\end{aligned} \tag{19}$$

The first five of these correspond to sectors that are already present in the ideal HRG in the Boltzmann approximation. However, they also contain contributions from interactions, like non-resonant pion-nucleon interactions contributing to P_{10}^{BS}, or K-Λ interactions contributing to P_{12}^{BS}. The $P_{1,-1}^{BS}$ is an exotic (pentaquark) sector with valence quark content $uudd\bar{s}$, which can also receive contributions from e.g., $p+K^0$ scattering. The sectors P_{2i}^{BS} receive contributions from baryon-baryon scattering: P_{20}^{BS} from $N-N$, P_{21}^{BS} from $N-\Lambda$, P_{22}^{BS} from $N-\Xi$ or $\Lambda-\Lambda$, and P_{23}^{BS} from $N-\Omega$ or $\Lambda-\Xi$. In each case, Σ can replace Λ. The coefficients P_{02}^{BS} and P_{03}^{BS} get contributions from two- and three-kaon scattering, respectively. The inclusion of the P_{03}^{BS} sector with the omission of the P_{30}^{BS} sector is motivated by the lower mass threshold of 3 kaon scattering as compared to 3 baryon scattering. In addition to these, we also performed an analysis in which four sectors with $B = 3$ were included:

$$P_{30}^{BS}, \mathrm{P}_{31}^{\mathrm{BS}}, \mathrm{P}_{32}^{\mathrm{BS}}, \mathrm{P}_{33}^{\mathrm{BS}}\ . \tag{20}$$

These sectors contain contributions from three-baryon scattering, with various strangeness contents. Our data is not yet sufficiently precise to reliably estimate the weight of sectors with $B = 3$, and their inclusion does not improve the χ^2 of the fits. Whether we include these, or not, the results for the $B = 2$ sectors remain consistent.

We perform a continuum extrapolation on each P_{ij}^{BS} coefficient assuming a smooth temperature dependence via the ansatz

$$f(T, N_\tau) = \left(a_0 + a_1 T + a_2 T^2\right) + \left(b_0 + b_1 T + b_2 T^2\right) \frac{1}{N_\tau^2}\,. \tag{21}$$

The results before and after continuum extrapolation are shown in the left panel of Fig. 4. The thermodynamic potential is then defined through the continuum extrapolated coefficients in the truncated form of Eq. 12. We impose the strangeness neutrality condition of Eq. 11 to obtain $\mu_S(T, \mu_B)$ and calculate the fluctuations as functions of T and μ_B only. The results are shown in the right panel of Fig. 4.

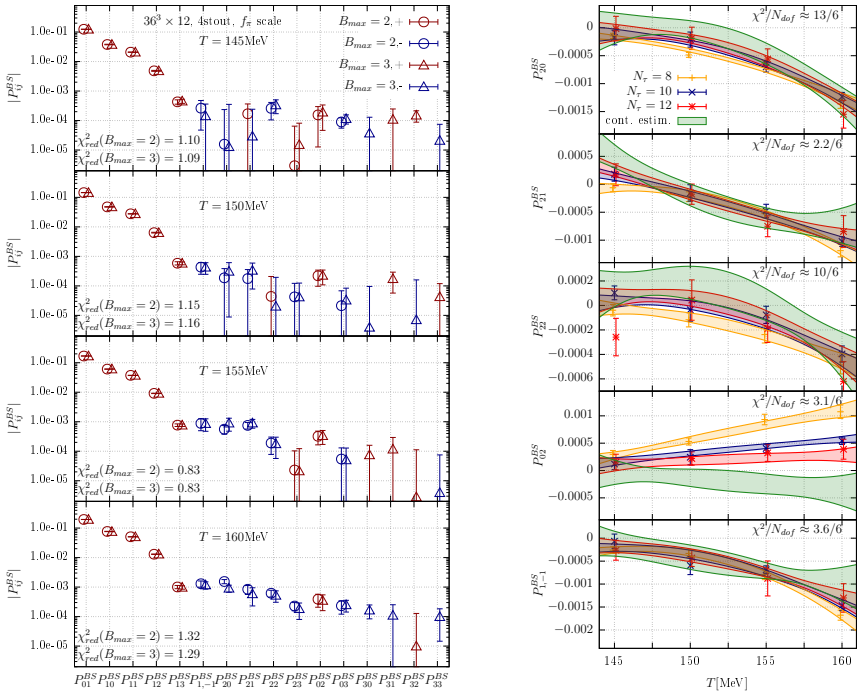


Figure 4. Left: Examples for the combined continuum extrapolation of the extracted P^{BS}_{ij} coefficients. Right: Continuum estimates of the fluctuation ratios χ^B_1/χ^B_2, χ^B_3/χ^B_1, χ^B_4/χ^B_2, χ^{BS}_{11}/χ^S_2 from a fugacity expansion, shown as a function of the dimensionless chemical potential μ_B/T for fixed temperatures, as well as on the crossover line $T_c(\mu_B)$.

5 Discussion

We will now discuss how these findings are experimentally relevant. In the introduction we already mentioned factors that complicate the direct comparison of the experimental findings with the grand canonical averages and fluctuations.

We performed an HRG-model-based study on the second order correlators, both diagonal and non-diagonal[34], showing how they relate to fluctuations of hadronic species which can be measured in heavy ion collision experiments. What percentage of these correlators is accounted for by particles that can actually be detected in the experiment varies quite considerably from observable to observable.

In order to facilitate the comparison to experimental measurements, we focused on

ratios of fluctuations, whose behaviour can be reproduced through commonly measured hadronic observables, which serve as proxies.

We summarise the findings for a ratio containing each of the three possible cross-correlators of baryon (B), electric charge (Q) and strangeness (S). The BQ correlator in equilibrium is dominated by proton fluctuations, with the other contributions – most notably the proton-pion correlation and hyperons self correlations – almost perfectly cancelling each other. Nonetheless, the information loss caused by isospin randomisation[39,40] prevents from constructing successful proxies for ratios including χ_{11}^{BQ}.

Luckily, neither the isospin randomisation, nor the introduction of cuts on the kinematics had a significant effect on either χ_{11}^{BS} or χ_{11}^{QS}. Because of this, we were able to construct proxies for the ratios χ_{11}^{BS}/χ_2^S and χ_{11}^{QS}/χ_2^S that are within 10% of the grand canonical prediction. These two ratios are not independent, since in the isospin symmetric case they are related by the Gell–Mann-Nishijima formula. It is striking that only two measured quantities, namely the variances of net-kaon and net-Lambda distributions, were sufficient to build the proxies for χ_{11}^{BS}/χ_2^S and χ_{11}^{QS}/χ_2^S. We showed that the inclusion of multi-strange hyperons does not improve the quality of the proxy. Moreover, although these are cross correlators of conserved charges, particle species cross correlators do not contribute significantly. In fact, none of the particle cross correlators contributes noticeably to any of the charge fluctuations or cross correlators, with the exception of the proton-pion one, which is however largely affected by isospin randomisation.

Thus, we have a ratio at hand that is available both from lattice simulations and for experimental measurement. The ratio χ_{11}^{BS}/χ_2^S behaves as a strangeness-related thermometer for chemical freeze-out. We provided continuum extrapolated results at zero and finite chemical potential for this quantity.

The sector method also provides a mean for the extrapolation to finite chemical potential, though it tends to diverge at smaller value of μ_B/T. On the other hand, it provides precious insight into the content of different sectors of the Hilbert space, which in turn point to next-to-leading effects in an hadronic gas. These can help constrain the role of interactions, sector-by-sector, and also point at the role of exotic states. All of this can, of course, be done at a temperature-by-temperature basis.

Acknowledgements

This project was partly funded by the DFG grant SFB/TR55 and also supported by the Hungarian National Research, Development and Innovation Office, NKFIH grants KKP126769. The project leading to this publication has received funding from Excellence Initiative of Aix-Marseille University - A*MIDEX, a French "Investissements d'Avenir" programme, AMX-18-ACE-005. The project also received support from the BMBF grant 05P18PXFCA. Parts of this work were supported by the National Science Foundation under grant no. PHY1654219 and by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, within the framework of the Beam Energy Scan Theory (BEST) Topical Collaboration. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725. A.P. is supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the UNKP-20-5 New National Excellence Program of the Ministry of Innovation and Technology. RB acknowledges support from the

U.S. Department of Energy Grant No. DE-FG02-07ER41521. 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 and JURECA Booster at Jülich Supercomputing Centre (JSC). Parts of the computations were performed on the QPACE3 system, funded by the DFG.

References

1. C. Gale, S. Jeon, and B. Schenke, *Hydrodynamic Modeling of Heavy-Ion Collisions*, Int. J. Mod. Phys., **A28**, 1340011, 2013.
2. J. Cleymans and K. Redlich, *Unified description of freezeout parameters in relativistic heavy ion collisions*, Phys. Rev. Lett., **81**, 5284–5286, 1998.
3. P. Braun-Munzinger, J. Stachel, and C. Wetterich, *Chemical freezeout and the QCD phase transition temperature*, Phys. Lett., **B596**, 61–69, 2004.
4. C. S. Fischer, *QCD at finite temperature and chemical potential from DysonSchwinger equations*, Prog. Part. Nucl. Phys., **105**, 1–60, 2019.
5. J. M. Pawlowski, *The QCD phase diagram: Results and challenges*, AIP Conf. Proc., **1343**, 75–80, 2011.
6. Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabo, *The Order of the quantum chromodynamics transition predicted by the standard model of particle physics*, Nature, **443**, 675–678, 2006.
7. Y. Aoki, Z. Fodor, S. D. Katz, and K. K. Szabo, *The QCD transition temperature: Results with physical masses in the continuum limit*, Phys. Lett., **B643**, 46–54, 2006.
8. S. Borsanyi et al., *Is there still any T_c mystery in lattice QCD? Results with physical masses in the continuum limit III*, JHEP, **1009**, 073, 2010.
9. A. Bazavov, T. Bhattacharya, M. Cheng, C. DeTar, H. T. Ding, et al., *The chiral and deconfinement aspects of the QCD transition*, Phys. Rev., **D85**, 054503, 2012.
10. S. Borsanyi, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, et al., *Full result for the QCD equation of state with 2+1 flavors*, Phys. Lett., **B730**, 99–104, 2014.
11. A. Bazavov et al., *Equation of state in (2+1)-flavor QCD*, Phys. Rev., **D90**, no. 9, 094503, 2014.
12. S. Borsanyi et al., *Calculation of the axion mass based on high-temperature lattice quantum chromodynamics*, Nature, **539**, no. 7627, 69–71, 2016.
13. A. Bazavov et al., *Update on the 2+1+1 flavor QCD equation of state with HISQ*, PoS, **LATTICE2013**, 154, 2014.
14. R. Hagedorn, *Statistical thermodynamics of strong interactions at high-energies*, Nuovo Cim. Suppl., **3**, 147–186, 1965.
15. R. Dashen, S.-K. Ma, and H. J. Bernstein, *S Matrix formulation of statistical mechanics*, Phys. Rev., **187**, 345–370, 1969.
16. R. F. Dashen and R. Rajaraman, *Narrow Resonances in Statistical Mechanics*, Phys. Rev. D, **10**, 694, 1974.
17. R. F. Dashen and R. Rajaraman, *Effective Elementarity of Resonances and Bound States in Statistical Mechanics*, Phys. Rev. D, **10**, 708, 1974.
18. R. Venugopalan and M. Prakash, *Thermal properties of interacting hadrons*, Nucl. Phys., **A546**, 718–760, 1992.

19. F. Karsch, K. Redlich, and A. Tawfik, *Hadron resonance mass spectrum and lattice QCD thermodynamics*, Eur. Phys. J., **C29**, 549–556, 2003.
20. P. Huovinen and P. Petreczky, *QCD Equation of State and Hadron Resonance Gas*, Nucl. Phys., **A837**, 26–53, 2010.
21. S. Borsanyi, G. Endrodi, Z. Fodor, A. Jakovac, S. D. Katz, et al., *The QCD equation of state with dynamical quarks*, JHEP, **1011**, 077, 2010.
22. A. Andronic, P. Braun-Munzinger, K. Redlich, and J. Stachel, *Decoding the phase structure of QCD via particle production at high energy*, Nature, **561**, no. 7723, 321–330, 2018.
23. S. Borsanyi, Z. Fodor, M. Giordano, S. D. Katz, D. Nogradi, A. Pasztor, and C. H. Wong, *Lattice simulations of the QCD chiral transition at real baryon density*, Phys. Rev. D, **105**, no. 5, L051506, 2022.
24. 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., **D92**, no. 5, 054503, 2015.
25. R. Bellwied, S. Borsanyi, Z. Fodor, J. Günther, S. D. Katz, C. Ratti, and K. K. Szabo, *The QCD phase diagram from analytic continuation*, Phys. Lett., **B751**, 559–564, 2015.
26. A. Bazavov et al., *Chiral crossover in QCD at zero and non-zero chemical potentials*, Physics Letters B, **795**, 15–21, Aug 2019.
27. S. Borsanyi, Z. Fodor, J. N. Guenther, R. Kara, S. D. Katz, P. Parotto, A. Pasztor, C. Ratti, and K. K. Szabo, *The QCD crossover at finite chemical potential from lattice simulations*, Phys. Rev. Lett., **125**, 052001, 2020.
28. L. Adamczyk et al., *Energy Dependence of Moments of Net-proton Multiplicity Distributions at RHIC*, Phys. Rev. Lett., **112**, 032302, 2014.
29. L. Adamczyk et al., *Beam energy dependence of moments of the net-charge multiplicity distributions in Au+Au collisions at RHIC*, Phys. Rev. Lett., **113**, 092301, 2014.
30. L. Adamczyk et al., *Collision Energy Dependence of Moments of Net-Kaon Multiplicity Distributions at RHIC*, Phys. Lett., **B785**, 551–560, 2018.
31. T. Nonaka, *Recent Results and Methods on Higher Order and Off-diagonal Cumulants of Identified Net-particle Multiplicity Distributions in Au+Au Collisions at STAR*, Nucl. Phys. A, **982**, 863–866, 2019.
32. S. Borsanyi, Z. Fodor, S. D. Katz, S. Krieg, C. Ratti, et al., *Fluctuations of conserved charges at finite temperature from lattice QCD*, JHEP, **1201**, 138, 2012.
33. 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., **D92**, no. 11, 114505, 2015.
34. R. Bellwied, S. Borsanyi, Z. Fodor, J. N. Guenther, J. Noronha-Hostler, P. Parotto, A. Pasztor, C. Ratti, and J. M. Stafford, *Off-diagonal correlators of conserved charges from lattice QCD and experiment*, Physical Review D, **101**, no. 3, Feb 2020.
35. V. Skokov, B. Friman, and K. Redlich, *Volume Fluctuations and Higher Order Cumulants of the Net Baryon Number*, Phys. Rev., **C88**, 034911, 2013.
36. A. Bzdak, V. Koch, and V. Skokov, *Baryon number conservation and the cumulants of the net proton distribution*, Phys. Rev., **C87**, 014901, 2013.

37. P. Alba, W. Alberico, R. Bellwied, M. Bluhm, V. Mantovani Sarti, et al., *Freeze-out conditions from net-proton and net-charge fluctuations at RHIC*, Phys. Lett., **B738**, 305–310, 2014.
38. R. Bellwied, S. Borsanyi, Z. Fodor, J. N. Guenther, S. D. Katz, P. Parotto, A. Pasztor, D. Pesznyak, C. Ratti, and K. K. Szabo, *Corrections to the hadron resonance gas from lattice QCD and their effect on fluctuation-ratios at finite density*, Phys. Rev. D, **104**, no. 9, 094508, 2021.
39. M. Kitazawa and M. Asakawa, *Revealing baryon number fluctuations from proton number fluctuations in relativistic heavy ion collisions*, Phys. Rev. C, **85**, 021901, 2012, doi:10.1103/PhysRevC.85.021901.
40. M. Kitazawa and M. Asakawa, *Relation between baryon number fluctuations and experimentally observed proton number fluctuations in relativistic heavy ion collisions*, Phys. Rev. C, **86**, 024904, 2012, [erratum: Phys. Rev. C, **86**, 069902, 2012], doi:10.1103/PhysRevC.86.024904.
41. S. Borsanyi, S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert and C. McNeile, *et al.*, *High-precision scale setting in lattice QCD*, JHEP, **09**, 010, 2012.

Molecular Dynamics Study of the Glass Transition of Highly Entangled Polymer Melts

Hsiao-Ping Hsu, Torsten Stuehn, Kostas C. Daoulas, and Kurt Kremer

Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany
E-mail: {hsu, stuehn, daoulas, kremer}@mpip-mainz.mpg.de

Understanding the generic properties of bulk polymer melts and polymer films containing highly entangled polymer melts as a function of temperature is important due to their potential applications in real life from both theoretical and technological points of view. Based on the standard bead-spring model, we have developed a coarse-grained weakly semiflexible polymer model for studying polymer systems under cooling. A short-range attractive potential between non-bonded monomers is added such that the pressure of systems is tuned to zero. To keep the local conformational properties of polymer chains, additionally, the commonly used bond-bending potential controlling the chain stiffness is replaced by a new bond-bending potential. The model captures the major features of glass-forming polymers. The viscosity dramatically increases in a non-Arrhenius way close to the glass transition temperature T_g and the Kuhn length and the internal distances along the chains only very weakly depend on the temperature, just as for typical experimental systems. T_g remains almost the same for the chain length longer than $\mathcal{O}(10)$ entanglements. We also find no significant difference in T_g between disentangled and highly entangled polymer melts of long chains, despite rather different chain conformations. We also have implemented an efficient algorithm for equilibrating large polymer films of different thicknesses as well. Applying the new variant of the bead-spring model enables us to study free-standing films or films with one supporting substrate. Furthermore, we can extend the study of the glass transition to polymer films. We see that T_g for both confined and free-standing films of thickness compatible with linear dimensions of bulk melt remains at the bulk value.

1 Introduction

When an amorphous polymer in the liquid more rubbery state is cooled towards the glass transition temperature T_g, the viscosity dramatically increases in a non-Arrhenius way[1–3]. Beyond that, T_g itself is important for processing and applications of commodity amorphous polymer materials[4]. Furthermore, confinement and surface interaction also have significant impact on the dynamic and structural properties of polymers[5–7]. For instance, long chain mobility in confined polymer melts is affected by chain entanglements and T_g can either increase or decrease with decreasing the thickness of films[8]. Experimentally, ways to determine T_g of polymers are by observing the change in the heat capacity of polymers using differential scanning calorimetry (DSC)[9], or by measuring the thermal expansion coefficient using thermo mechanical analysis (TMA)[10]. However, despite decades of research, the nature of the glass transition is still not fully understood[11–16]. This holds for both simple, i.e. low molecular weight, glass formers as well as polymeric systems. It becomes even more complex, when confinement and specific surface interactions come into play.

Computer simulations play an important role in investigating the structure and molecular motion (viscosity) of polymeric systems under a variety of different well-defined conditions. For studying glassy polymers, both atomistic and coarse-grained models are widely

used in the literature[13,14]. In order to study generic properties of large and highly entangled polymer melts in bulk, in confinement and with free surfaces as a function of temperature within accessible computing times, we adopt and modify a highly efficient coarse-grained model[17]. This new model is set to allow for a fast switch between models, for which a wealth of data already exists from our previous simulations. In the next sections, we describe this model and its application for studying the glass transition of several polymer melt systems.

2 Models

In the standard bead-spring model, the bonds linking monomers are described by the FENE potential $U_{\mathrm{FENE}}(r) = -kR_0^2/2[1-(r/R_0)^2]$ with the force constant $k = 30\epsilon/\sigma^2$ and the maximum value of bond length, $R_0 = 1.5\sigma$. In addition excluded volume interaction between monomers is taken care of by a purely repulsive Lennard-Jones (LJ) potential, the Weeks-Chandler-Andersen (WCA) potential[17] $U_{\mathrm{WCA}}(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6 + 1/4]$ as shown in Fig. 1a where the potential and the force are zero at the cut-off $r_{\mathrm{cut}} = 2^{1/6}\sigma$, ϵ is the energy strength of the pairwise interaction, and σ is the size of monomers. While the latter is applied to all pairs of monomer, it results in combination with the FENE potential in a strong bond potential. With a monomer mass of $m = 1$, these Lennard-Jones units also provide a natural time definition via $\tau = \sigma(m/\epsilon)^{1/2}$. Additionally, for controlling the bending elasticity, i.e. chain stiffness, the standard bond-bending potential (Fig. 1b) is given by[18,19] $U^{(\mathrm{old})}_{\mathrm{BEND}}(\theta) = k_\theta(1-\cos\theta)$ with the bond angle θ between two successive bonds. Choosing $k_\theta = 1.5\epsilon$ and the standard bulk melt density ρ_0 of $0.85\sigma^{-3}$ results in an entanglement length of only $N_e = 28$ monomers[20]. We focus on $k_\theta = 1.5\epsilon$ since a huge body of data already exists[21–25].

For confinement, we consider polymer melts confined between two planar, structureless repulsive walls[26]. The walls located at $z = 0$ and $z = L_z$ are described by the 10-4 Lennard-Jones planar wall potential[5,27] $U_{\mathrm{wall}}(z) = 4\pi\varepsilon_w[(\sigma/z)^{10}/5 - (\sigma/z)^4/2 + 3/10]$

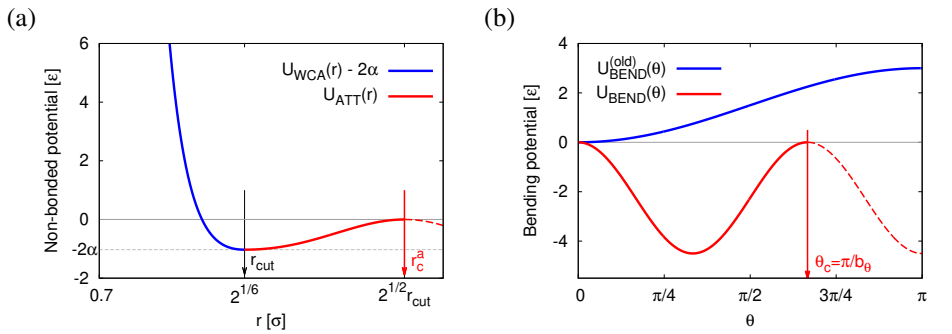


Figure 1. (a) Short-range repulsive and attractive potentials $U_{\mathrm{WCA}}(r)$ and $U_{\mathrm{ATT}}(r)$ with $\alpha = 0.5145\epsilon$ (Eq. 1) plotted as a function of distance r between non-bonded monomers. $U_{\mathrm{WCA}}(r) - 2\alpha = U_{\mathrm{ATT}}(r)$ at $r = r_{\mathrm{cut}}$, and $r_c^a = \sqrt{2}r_{\mathrm{cut}}$. (b) Standard and new bond-bending potentials $U^{(\mathrm{old})}_{\mathrm{BEND}}(\theta)$ with $k_\theta = 1.5\epsilon$ and $U_{\mathrm{BEND}}(\theta)$ with $a_\theta = 4.5\epsilon$ and $b_\theta = 1.5$ (Eq. 2), plotted as a function of bond angle θ between two successive bonds. The cut-off values are pointed by arrows.

where $\varepsilon_w = 0.0032\epsilon$ is the interaction strength between monomers and the walls, and $U_{\rm wall}(z) = 0$ at $z = \sigma$ and $z = L_z - \sigma$.

The excluded volume interaction between monomers is taken care of by $U_{\rm WCA}(r)$, at $T = 1.0\epsilon/k_B$ this results in a rather high pressure ($P \approx 5.0\epsilon/\sigma^3$). This prevents studies of systems with free surfaces, where the experimentally applied atmospheric pressure to a good approximation can be taken as $P \approx 0.0\epsilon/\sigma^3$. We hence introduce a new short-range attractive potential[28] $U_{\rm ATT}(r)$ (Figure 1a) in a way that at the reference/starting temperature $T = 1.0\epsilon/k_B$, the conformations are the same for both non-bonded interaction potentials, while the system pressure is reduced from $P = 5.0\epsilon/\sigma^3$ to $P = 0.0\epsilon/\sigma^3$. The potential is given by

$$U_{\rm ATT}(r) = \begin{cases} \alpha[\cos(\pi(r/r_{\rm cut})^2) - 1] & , r_{\rm cut} \le r < r_c^a \\ 0 & , \text{otherwise.} \end{cases} \quad (1)$$

$\alpha = 0.5145\epsilon$ is the strength parameter, $r_{\rm cut}$ and $r_c^a = \sqrt{2}r_{\rm cut}$ are lower and upper cut-off distances such that the force at $r = r_{\rm cut}$ and $r = r_c^a$ vanishes. To model free-standing films in vacuum, this interaction leads to a stable film with zero pressure in the direction perpendicular to the walls, while the surface tension leads to lateral negative pressure contribution.

To study the glass transition, we have to overcome another problem of the standard semiflexible chain model. As the stiffness of chains is controlled by the standard bond-bending potential $U_{\rm BEND}^{\rm (old)}(\theta)$, chains are getting stretched out with decreasing temperature[27]. In experiments, however, chain conformations only very weakly depend on temperature[29]. To overcome this shortage, we replace the bond-bending potential by

$$U_{\rm BEND}(\theta) = -a_\theta \sin^2(b_\theta\theta)\,, \quad 0 < \theta < \theta_c = \pi/b_\theta \quad (2)$$

where the fitting parameters are set to $a_\theta = 4.5\epsilon$ and $b_\theta = 1.5$ (Fig. 1b) to perfectly match results of the standard model at $T = 1\epsilon/k_B$.

The ESPResSo++package[30,31] is used to perform the standard molecular dynamics (MD) simulations with Langevin thermostat in the NVT ensemble while in the NPT ensemble, Hoover barostat is used together with Langevin thermostat.

3 Viscosity of Bulk Polymer Melts Subjected to Cooling

To test the new coarse-grained model[28], we first perform MD simulations of a polymer melt consisting of $n_c = 2000$ chains of $N = 50$ monomers in the NVT ensemble. Fig. 2a shows that switching between two variants of bead-spring model, chain conformations characterised by the mean square internal distance $\langle R^2(s)\rangle$ remain the same at $T = 1.0\epsilon/k_B$ although the pressure drops to $P = 0.0\epsilon/\sigma^3$ with the potential $U_{\rm ATT}(r)$. We then simulate the system in the NPT ensemble at constant pressure $P = 0.0\epsilon/\sigma^3$ and constant temperature T and apply stepwise cooling[32]. The temperature is reduced in steps of $\Delta = 0.025\epsilon/k_B$ from $T = 1.0\epsilon/k_B$ to $0.1\epsilon/k_B$ with a relaxation time between each step of $\Delta t = 30000\tau \approx 13\tau_e$, $\tau_e \approx 2266\tau_e$ being the entanglement time[20], that subchains of the order of $N_e \approx 28$ can relax easily at higher temperatures close to $T = 1.0\epsilon/k_B$. This results in a cooling rate of $\Gamma = \Delta T/\Delta t = 8.3 \times 10^{-7}\epsilon/(k_B\tau)$, allowing for a full relaxation of this small weakly entangled polymer melt up to the region very close to T_g. This cooling rate serves as the standard cooling rate through out the whole paper.

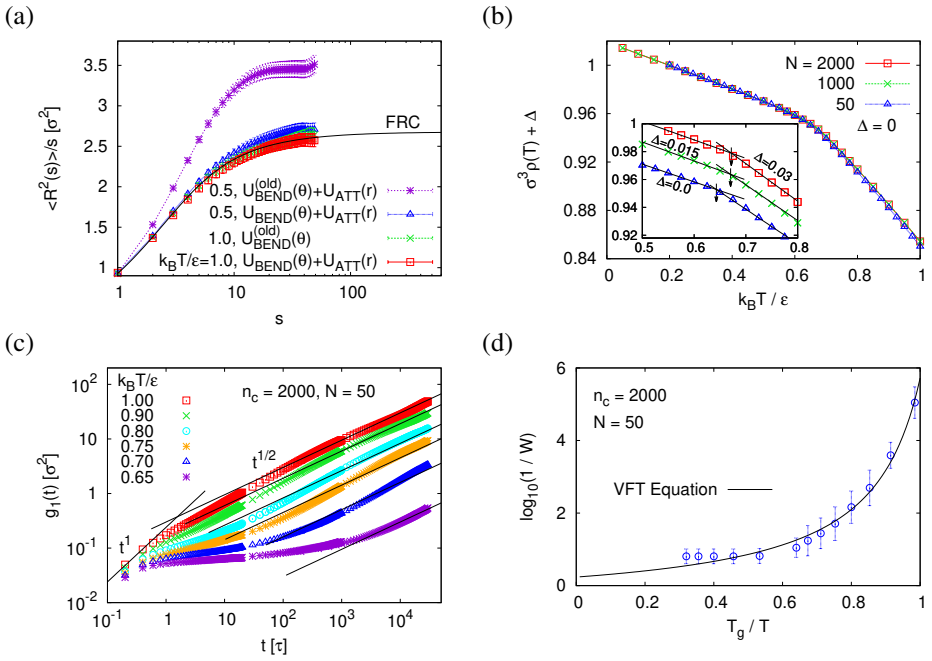


Figure 2. (a) Rescaled mean square internal distance, $\langle R^2(s)\rangle/s$, plotted versus the chemical distance s between two monomers along the same chains for bulk polymer melt of $n_c = 2000$ chains of $N = 50$ monomers described by different variants of the bead-spring model at the temperature $T = 1.0\epsilon/k_B$, and $T = 0.5\epsilon/k_B$ subjected to cooling . (b) Density of bulk polymer melts, $\rho(T)$, plotted versus temperature T for chains containing $N = 50$, 1000, and 2000 monomers in the melts. Only data for $0.5 \leq k_BT/\epsilon \leq 0.8$ are shown in the inset. Data are shifted by a constant Δ for a better illustration. Estimates of T_g are marked by arrows. (c) Time evolution of mean square displacement of inner monomers, $g_1(t)$, at various chosen temperatures T. The predicted scaling laws by the Rouse model are shown by straight lines. (d) Common logarithm of the inverse of Rouse rate W estimated from (c), $\log_{10}(1/W)$, plotted as a function of T_g/T with $T_g \approx 0.64\epsilon/k_B$. The VFT equation $\log_{10}(1/W) = A + B/(T - T_0)$ with $A = 0.24$, $B = 0.45\epsilon/k_B$, and $T_0 = 0.56\epsilon/k_B$ is shown by a solid curve[28].

For $U_{\rm BEND}(\theta)$ chain conformations only very weakly vary locally with temperature, just as in experiment, while this is not the case for $U^{\rm (old)}_{\rm BEND}(\theta)$. At $T = 0.5\epsilon/k_B$, those chains are strongly stretched out as shown in Fig. 2a. $T_g \approx 0.64\epsilon/k_B$ is determined by the change in density $\rho(T)$, i.e., from the intersection of linear extrapolation of $\rho(T)$ between the liquid branch ($\rho_{\rm liquid}(T) = a_{\rm liquid} + \alpha_{\rm liquid}T$) and glass branch ($\rho_{\rm glass}(T) = a_{\rm glass} + \alpha_{\rm glass}T$), $\alpha_{\rm liquid}$ and $\alpha_{\rm glass}$ being thermal expansion coefficients (see Fig. 2b). For $T > T_g$, additional NVT runs are also performed to investigate the monomer mobility through their mean square displacement (MSD), $g_1(t)$. According to the Rouse model, $g_1(t) = \sigma^2(Wt)^{1/2}$ is expressed in terms of the Rouse rate $W = 12k_BT/(\pi\zeta\sigma^2)$. The monomeric friction coefficient $\zeta(\propto D^{-1} \propto \eta)$ is related to the self-diffusion coefficient $D = k_BT/(N\zeta)$ and the viscosity $\eta \propto 1/W$ using the Stokes-Einstein relation. $g_1(t)$ of inner 12 monomers are shown in Fig. 2c. The Rouse rate W at each T is determined by the best fit of a straight line with slope $1/2$ going through our data on log-log scales. We observe the non-Arrhenius rapid increase of the melt viscosity near T_g of a somewhat fragile polymeric glass former in the Vogel-Fulcher-Tammann (VFT) plot[33] in Fig. 2d.

4 Entanglement Effect on the Glass Transition of Polymer Melts

Chain entanglements are known to dominate viscoelastic properties of long chain polymer melts. Over the years there has been a lot of interest to describe the transition from short chain Rouse to long chain reptation behaviour. Similar to that entanglements are also expected to play an important role for the precise location of T_g. To look at that, one can approach this topic from two sides. One can either study long entangled and non-entangled chains or one can study equilibrium melts of increasing chain length. For the first approach we study disentangled and entangled polymer melts containing $n_c = 1000$ chains of $N = 1000 \approx 36N_e$ monomers (Fig. 3a). To produce a disentangled melt, we start with isolated good solvent chains, which are contracted on a time scale much faster than the chain relaxation time, so that no deep knotting can occur. These collapsed chains are squeezed into a box and relaxed for a short time in order to obtain a melt of single-chain nanoparticles with an overall homogeneous monomer density $0.85\sigma^{-3}$, see Refs. 34, 35 for further details. For comparison, we perform NPT simulations of these two polymer systems, the entangled and the non-entangled one, respectively, with new bead-spring model described in Sec. 2 choosing the standard cooling rate. As shown in Fig. 3b, there is no detectable shift in T_g determined from the volume change as a function of T for both disentangled and entangled polymer melt systems, where $T_g \approx 0.67\epsilon/k_B$. The radial non-bonded monomer-monomer distribution function $g(r)$ shown in Fig. 3c also in-

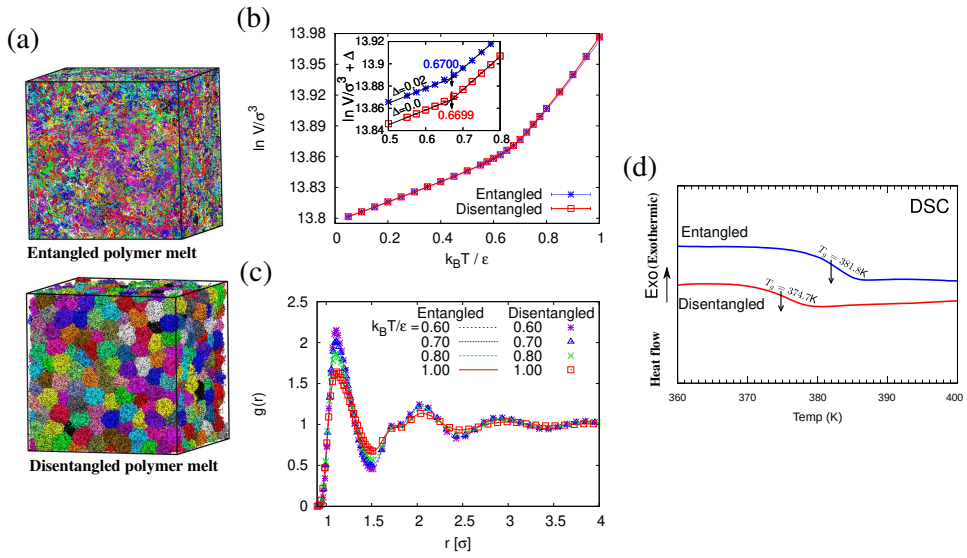


Figure 3. (a) Snapshots of entangled and disentangled polymer melt systems of $n_c = 1000$ chains containing $N = 1000$ monomers, (b) Logarithm of volume of the systems, $\ln V$, plotted versus temperature T. Only data for $0.5 \leq k_BT/\epsilon \leq 0.8$ are shown in the inset. Data are shifted by a constant Δ for a better illustration. Estimates of T_g are marked by arrows. (c) Radial distribution function of non-bonded monomer-monomer pairs, $g(r)$, plotted as a function of r. (d) Comparison of the glass transition temperature T_g of entangled and disentangled polystyrene melt systems through differential scanning calorimetry (DSC) experiment[34].

dicates that the local melt structures are identical, even with rather different global chain conformations. Our results suggest that topological constraints due to chain entanglements do not play a significant role. Motivated by our simulations we have also performed experiments[34]. T_g of the melt of disentangled polystyrene (PS) is reduced by about 2% comparing to T_g of bulk entangled PS measured by DSC shown in Fig. 3d. However, further experiments and simulations on viscoelastic properties are needed in order to see, whether nanoparticle interface effects determine this shift or not.

Furthermore, according to the theoretical prediction by Flory and Fox[36], $T_g(N) = T_{g,\infty} - K/N$, where K is a polymer-specific constant, one should expect that T_g increases with increasing chain size[37,38]. For example, T_g of PS increases with chain length (molecular weight) and then slowly levels off at around the entanglement length N_e. However, the results on the non-entangled nanoparticles rather suggest that it is more an effect of density of free ends rather than an entanglement effect. Still, beyond N_e, T_g does not change much with increasing molecular weight and approaches a constant value of about $T_g = 374$K for high molecular weight PS[39]. Following the exact same cooling protocol, our results presented in Fig. 2b show that $T_g(50)(\approx 0.64\epsilon/k_B) < T_g(1000) \approx T_g(2000)(\approx 0.67\epsilon/k_B)$ by cooling down polymer melts of $N = 50$, 1000, and 2000, i.e., $N \approx 2N_e$, $36N_e$, and $72N_e$. This is in perfect agreement with the observation from experiments.

5 Confined and Free-Standing Films under Cooling

Another matter of discussion since many years[8] is the dependence of the glass transition for thin supported or free standing films. To study this problem we first had to develop an efficient method to equilibrate highly entangled films[26]. As for bulk melts films are first equilibrated in their CG representations where polymer chains are treated as strings of soft spheres (blobs). Then we apply a (one-step) backmapping to reinsert the microscopic details of the underlying bead-spring model[17], see Fig. 4. Finally, short MD simulations about $6\tau_e$ independent of N take care of equilibration of the local bead packing[26]. For the comparison to our reference systems[20] and to test the methodology[26], we here focus on polymer melts of $n_c = 1000$ chains of $N = 2000$ monomers at the bulk melt density $\rho_0 = 0.85\sigma^{-3}$. The two walls are at a distance of $L_z = 133.7\sigma \approx 4.4R_g$ compatible with the spatial extension of bulk reference chains ($R_e/\sqrt{6} = R_g \approx 30.15\sigma$). As shown in Fig. 5, $\langle R^2(s, N_b)\rangle$

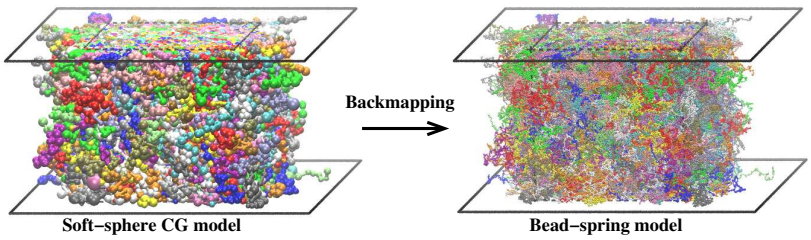


Figure 4. Snapshots of a typical configuration of confined polymer films containing $n_c = 1000$ bead-spring chains of $N = 2000 = N_b N_{\rm CG}$ monomers (right) is obtained from $n_c = 1000$ soft-sphere CG chains (left) of $N_{\rm CG} = 80$ spheres via the backmapping procedure, where $N_b = 25$ monomers[26].

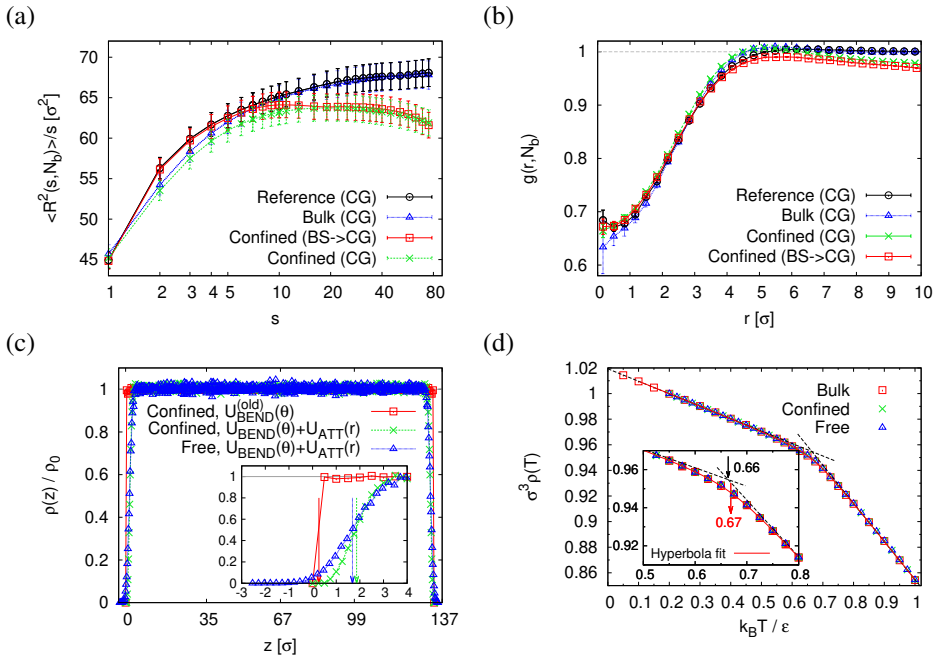


Figure 5. (a) Rescaled mean square internal distance, $\langle R^2(s, N_b = 25)\rangle/s$, plotted as a function s, and (b) Pair distribution of soft-sphere pairs, $g(r, N_b = 25)$, plotted as a function of r for bulk melt and confined film based on the CG model and based on the bead-spring model in a CG representation, as indicated. (c) Monomer density profile along the direction perpendicular to the walls, rescaled by the bulk melt density, $\rho(z)/\rho_0$, plotted as a function of z at $T = 1.0\epsilon/k_B$ for confined and free-standing films based on two variants of bead-spring model, as indicated. For comparison, the centres of all films in the z-direction are matched. In the inset, only data for $z < 4\sigma$ are shown. The Gibbs dividing surfaces at $z = z_G^{(\mathrm{lower})} = 0.27\sigma$, 1.68σ, and 1.86σ from left to right. (d) Density $\rho(T)$ plotted as a function T for bulk melt, and confined and free-standing films. The hyperbola fit is given by the curve. The two tangent lines at high and low temperature are shown by dashed lines.

and $g(r, N_b)$ for the equilibrated confined CG melt, and the equilibrated confined melt in a CG representation are consistent with each other while there characteristic deviations due to the confinement. Tuning the interaction strength $\varepsilon_w (= 0.0032\epsilon)$ between the walls and monomers, the monomer density of confined polymer melts in equilibrium is kept at ρ_0 close to the wall.

For free-standing films and for cooling, we employ the new variant of bead-spring model[28] described in Sec. 2 starting from the fully equilibrated confined polymer film. We first perform MD simulations in the NVT ensemble at $T = 1.0\epsilon/k_B$ to stabilise the system, and then further relax the confined polymer film for about $13\tau_e$ in the NPT ensemble to ensure that the pressure $P = 0.0\epsilon/\sigma^3$. The thickness of the confined film, determined according to the concept of Gibbs dividing surface[40], finally settled at $h \approx 130\sigma$ from 133σ, and the lateral dimensions of film increase slightly from 133σ to 134σ. Under these circumstances, a free-standing film is prepared after removing two walls by turning off the wall potential while keeping its lateral dimensions. Further relaxation for another $13\tau_e$ does not lead to changes, the thickness remains at $h = L_z = 130\sigma$ within small fluctuations. For a

first, preliminary comparison of the glass transition, bulk melt, confined and free-standing films are cooled following the standard cooling protocol. Results for the monomer density, $\rho = N/V_{\text{eff}}$ with the effective volume $V_{\text{eff}} = L_x L_y h$ as a function of temperature T, for these three polymer systems are shown in Fig. 5d. $\rho(T)$ for confined and free-standing films of $h > R_g$ are indistinguishable from the bulk melt. In comparison with the previous bi-linear fit ($T_g \approx 0.67\epsilon/k_B$ in Fig. 2b), here we adopt another approach to estimate T_g by performing a hyperbolic fit[41] $\rho(T) = c - a(T - T_0) - \frac{b}{2}\left(T - T_0 + \sqrt{(T - T_0)^2 + 4e^f}\right)$ where c, T_0, a, b, and f are fitting parameters. T_g is either defined by $T_g = T_0 \approx 0.67\epsilon/k_B$ or the intersection point of the two tangents drawn at the high and the low temperature, $T_g \approx 0.66\epsilon/k_B$. However, uncertainty quantification in the determination of T_g has become an important issue recently[41,42].

6 Conclusion

Following the same implemented strategies, one can examine the effect of confinement and chain topology on T_g, covering the whole range from thick films ($h > R_g$) to very thin films ($h < R_g$) at zero pressure. Furthermore, it also gives the possibility of studying entanglement effects on the rheological properties and the morphology of films subject to deformation.

Acknowledgements

We thank M. K. Singh for preparing the disentangled polymer melt system and performing experiments. This work was supported by the European Research Council under the European Union's Seventh Framework Programme (Grant No. FP7/2007-2013)/ERC Grant Agreement No. 340906-MOLPROCOMP. The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

References

1. R. H. Boyd, R. H. Gee, J. Han, and Y. Jin, *Conformational dynamics in bulk polyethylene: A molecular dynamics simulation study*, J. Chem. Phys., **101**, 788–797, 1994.
2. C. A. Angell, *Formation of glasses from liquids and biopolymers*, Science, **267**, 1924–1935, 1995.
3. L. Berthier and G. Biroli, *Theoretical perspective on the glass transition and amorphous materials*, Rev. Mod. Phys., **83**, 587–645, 2011.
4. J. H. Anantrao, J. C. Motichand, and B. S. Narhari, *A review on: glass transition temperature*, Int. J. Adv. Res., **5**, 671–681, 2017.
5. T. Aoyagi, J. I. Takimoto, and M. Doi, *Molecular dynamics study of polymer melt confined between walls*, J. Chem. Phys., **115**, 552–559, 2001.
6. M. Erber, M. Tress, E. U. Mapesa, Anatoli Serghei, K.-J. Eichhorn, B. Voit, and F. Kremer, *Glassy Dynamics and Glass Transition in Thin Polymer Layers of PMMA Deposited on Different Substrates*, Macromolecules, **43**, 7729–7733, 2010.

7. C. Batistakis, A. V. Lyulin, and M. A. J. Michels, *Slowing Down versus Acceleration in the Dynamics of Confined Polymer Films*, Macromolecules, **45**, 7282–7292, 2012.
8. B. D. Vogt, *Mechanical and Viscoelastic Properties of Confined Amorphous Polymers*, J. Polym. Sci. B: Polym. Phys., **56**, 9–30, 2018.
9. V. B. F. Mathot, *Calorimetry and thermal analysis of polymers*, Hanser, Munich, 1994.
10. R. Bird, C. Curtis, and R. Armstrong, *Dynamics of Polymer Fluids, 2nd ed.*, Wiley, New York, 1987.
11. C. A. Angell, *Perspective on the glass transition*, J. Phys. Chem. Solids, **49**, 863–871, 1988.
12. C. A. Angell, *Relaxation in liquids, polymers and plastic crystals - strong/fragile patterns and problems*, J. Non-Cryst. Solids, **131-133, Part 1**, 13–31, 1991.
13. K. Binder and W. Kob, *Glassy Materials and Disordered Solids*, World Scientific, Singapore, 2005.
14. J.-L. Barrat, J. Baschnagel, and A. Lyulin, *Molecular dynamics simulations of glassy polymers*, Soft Matter, **6**, 3430–3446, 2010.
15. F. H. Stillinger and P. G. Debenedetti, *Glass transition thermodynamics and kinetics*, Annu. Rev. Condens. Matter Phys., **4**, 263–285, 2013.
16. M. D. Ediger and J. A. Forrest, *Dynamics near free surfaces and the glass transition in thin polymer films: A view to the future*, Macromolecules, **47**, 471–478, 2014.
17. K. Kremer and G. S. Grest, *Dynamics of entangled linear polymer melts: a molecular-dynamics simulation*, J. Chem. Phys., **92**, 5057, 1990.
18. R. Faller, A. Kolb, and F. Müller-Plathe, *Local chain ordering in amorphous polymer melts: Influence of chain stiffness*, Phys. Chem. Chem. Phys., **1**, 2071, 1999.
19. R. Everaers, S. K. Sukumaran, G. S. Grest, C. Svaneborg, A. Sivasubramanian, and K. Kremer, *Rheology and microscopic topology of entangled polymeric liquids*, Science, **303**, 823, 2004.
20. H.-P. Hsu and K. Kremer, *Static and dynamic properties of large polymer melts in equilibrium*, J. Chem. Phys., **144**, 154907, 2016.
21. G. Zhang, L. A. Moreira, T. Stuehn, K. Ch. Daoulas, and K. Kremer, *Equilibration of high molecular weight polymer melts: a hierarchical strategy*, ACS Macro Lett., **3**, 198, 2014.
22. L. A. Moreira, G. Zhang, F. Müller, T. Stuehn, and K. Kremer, *Direct equilibration and characterization of polymer melts for computer simulations*, Macromol. Theor. Simul., **24**, 419, 2015.
23. C. Svaneborg and R. Everaers, *Characteristic Time and Length Scales in Melts of Kremer-Grest Bead-Spring Polymers with Wormlike Bending Stiffness*, Macromolecules, **53**, 1917–1941, 2020.
24. H.-P. Hsu and K. Kremer, *Primitive Path Analysis and Stress Distribution in Highly Strained Macromolecules*, ACS Macro Lett., **7**, 107–111, 2018.
25. H.-P. Hsu and K. Kremer, *Chain retraction in highly entangled stretched polymer melts*, Phys. Rev. Lett., **121**, 167801, 2018.
26. H.-P. Hsu and K. Kremer, *Efficient equilibration of confined and free-standing films of highly entangled polymer melts*, J. Chem. Phys., **153**, 144902, 2020; **156**, 019901, 2022.
27. G. S. Grest, *Grafted polymer brushes in polymeric matrices*, J. Chem. Phys., **105**, 5532, 1996.

28. H.-P. Hsu and K. Kremer, *A coarse-grained polymer model for studying the glass transition*, J. Chem. Phys., **150**, 091101, 2019; **150**, 159902, 2019.
29. L. J. Fetters, D. J. Lohse, and R. H. Colby, "Chain dimensions and entanglement spacings", in: Physical Properties of Polymers Handbook, 2nd, J. E. Mark, (Ed.), chapter 25, 447–454, Springer, New York, 2007.
30. J. D. Halverson, T. Brandes, O. Lenz, A. Arnold, S. Bevc, V. Starchenko, K. Kremer, T. Stuehn, and D. Reith, *ESPResSo++: A modern multiscale simulation package for soft matter systems*, Comput. Phys. Commun., **184**, 1129–1149, 2013.
31. H. V. Guzman, N. Tretyakov, H. Kobayashi, A. C. Fogarty, K. Kreis, J. Krajniak, C. Junghans, K. Kremer, and T. Stuehn, *ESPResSo++ 2.0: Advanced methods for multiscale molecular simulations*, Comput. Phys. Commun., **238**, 66–76, 2019.
32. J. Buchholz, W. Paul, F. Varnik, and K. Binder, *Cooling rate dependence of the glass transition temperature of polymer melts: Molecular dynamics study*, J. Chem. Phys., **117**, 7364–7372, 2002.
33. G. Strobl, *The Physics of Polymers, 3nd ed.*, Springer, 2007.
34. M. K. Singh, M. Hu, Y. Cang, H.-P. Hsu, H. Therien-Aubin, K. Koynov, G. Fytas, K. Landfester, and K. Kremer, *Glass transition of disentangled of entangled polyemr melts: single-chain-nanoparticles approach*, Macromolecules, **53**, 7312–7321, 2020.
35. T. Vettorel, G. Besold, and K. Kremer, *Fluctuating Soft-Sphere Approach to Coarse-Graining of Polymer Models*, Soft Matter, **6**, 2282, 2010.
36. T. G. Fox and P. J. Flory, *Second-Order Transition Temperatures and Related Properties of Polystyrene. I. Influence of Molecular Weight*, J. Appl. Phys., **21**, 581–591, 1950.
37. J. Hintermeyer, A. Herrmann, R. Kahlau, C. Goiceanu, and E. A. Rössler, *Molecular Weight Dependence of Glassy Dynamics in Linear Polymers Revisited*, Macromolecules, **41**, 9335–9344, 2008.
38. D. Baker, M. Reynolds, R. Masurel, P. Olmsted, and J. Mattsson, *Chain-length, flexibility and the glass transition of polymers*, arXiv:1911.13278, 2019.
39. P. G. Santangelo and C. M. Roland, *Molecular weight dependece of fragility in polystyrene*, Macromolecules, **31**, 4581–4585, 1998.
40. J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids: with applications to soft matter*, Elsevier Science, 2013.
41. P. N. Patrone, A. Dienstfred, A. R. Browning, S. Tucker, and S. Christensen, *Uncertainty quantification in molecular dynamics studies of the glass transition temperature*, Polymer, **87**, 246–259, 2016.
42. D. Mckechnie, J. Cree, D. Wadkin-Snaith, and K. Johnston, *Glass transition temperature of a polymer thin film: statistical and fitting uncertainties*, Polymer, **195**, 122433, 2020.

Astrophysics

Astrophysics – The Early Universe

Peter L. Biermann

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

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

The existence of a super-massive black hole in the centre of our Galaxy was finally confirmed via imaging, after an earlier image of the black hole at the centre of the giant elliptical galaxy M87. These images are referred to as a "shadow" of the black hole, showing basically a ring of emission. To perform such observations requires a huge library of ray tracing of photons from plasma emission models near to a possibly rotating black hole, with unknown orientation of the spin-axis. The resulting images need to be fitted to interferometric observations at very short wavelengths. Such explorations require first an intercontinental array of mm-telescopes, in the analysis super computer facilities, to do the interferometry and then a library of plasma models near the black hole. The super-massive black hole in the centre of our Galaxy is close to the minimum common mass in the mass distribution of such black holes. How do such black holes form and grow is a central problem in cosmology. UV emission from the activity of first super-massive black holes, stellar mass black holes and initially the massive stars reionises the intergalactic medium, the ubiquitous gas. This epoch needs to be modelled and matched to the data. This provides further constraints on our understanding of the early universe.

Imaging Super-Massive Black Holes

Imaging black holes is the new frontier of General Relativity, with now two nicely detected, ranging in mass from a few $10^6\,M_\odot$ to almost $10^{10}\,M_\odot$. Since many mm-telescope pairs are required, around the globe, this effort is truely international.

Gravitational Wave Emission $->$ Secondary BH Mergers

The statistics of stellar mass black hole mergers, their masses and spins based on the detection of gravitational waves is now so extensive that we learnt that there are multiple generations of successive mergers, with high masses observed. This analysis requires libraries of possible mergers, and a fit to data. Do these mergers provide a path to yet higher masses? Do these observations give a constraint of the spin evolution of black holes?

Reionisation in the Early Universe

The gas of the very early universe is ionised, recombines near redshift 1100, and is then neutral. Around redshift about 6 to about 20 it is reinoised again, probably mostly by the UV radiation of massive stars, but also from the activity of the first black holes. These black holes range in mass from a few $M_\odot$ to many hundreds of $M_\odot$ via multiple mergers in dense star clusters. In the core of a massive star cluster the first super-massive black holes might form, also emitting lots of UV radiation from their immediate proximity. The best test for these concepts is our local environment, the local group of galaxies, and its immediate neighbours: This modelling is fully constrained by the analysis of the microwave

background fluctuations, measured by, e.g., the Planck satellite. Galaxies, from ultra faint dwarf galaxies to more massive galaxies like ours and Andromeda, dominate the local galaxy neighbourhood. With some trial and error as well as initial constraints it is possible to simulate our local group of galaxies. This then serves as a test for the frequency of dwarf galaxies, the orbital planes of dwarf galaxies around the bigger galaxies, as well as the formation of the first super-massive black holes, but foremost the re-ionisation of the intergalactic gas. In a current project "The Reionisation of the Local Universe in the HESTIA Suite" (L. Conaboy, I.T. Iliev, N.I. Libeskind) the emphasis is on the first reionisation by the radiation of massive stars. Today most massive stars are born in binary star systems, often even in triples and quadruples; so also many combinations of a neutron star or a black hole with a massive star can be formed, and can then persist for some limited time. All these stars and their remnants are part of stellar clusters. This enhances the contribution of UV emission from activity around compact objects, as emphasised by Mirabel *et al.* This is especially true in the early stages, when the general heavy element abundance is still low, since then more black holes are formed relative to star formation. It can be expected that this line of investigation will provide strong constraints on the physical concepts for the early universe, such as the nature of dark matter and dark energy.

Intermediate Mass Black Hole Formation in Stellar Clusters

One of the persistent issues for our understanding of the early universe is the formation of the first super-massive black holes. Their mass distribution today suggests a low mass cutoff around a few million Solar masses, just as the mass of our local super-massive black hole in our Galaxy. This is also the mass at which they can start growing by accretion to get to the highest masses observed (work by Caramete, Munyaneza, Wang and many others), close to $10^{10}\,M_{\odot}$, and get there within the age of the universe. As the correlations observed between central supermassive black holes and central compact star clusters with their host galaxies appear to be similar, it is plausible and has been suspected since the 1960s that star clusters provide the source for the first seeds of super-massive black holes. An important step in such an endeavor is being done in the project "Growth of Seed Black Holes in Galactic Nuclei" (R. Spurzem, F. Rizzuto, M. Arca Sedda, A. Kamlah, P. Berczik, Q. Shu, A. Tanikawa, T. Naab). They show that multiple mergers allow to match the observed mass distribution of stellar mass black holes, based on the gravitational wave data of the LIGO/VIRGO/KAGRA collaborations. Their mass spectrum runs to more than $100\,M_{\odot}$ as observed in a number of cases, and so clearly and easily explains intermediate mass black holes. Whether a run-away in the steps towards such black holes then yields the high mass of black holes such as in our Galactic Centre has been suspected for a long time (work by Spitzer and Sanders), but remains a goal at present for the next stage of this work. Another aspect that is getting looked at now is the spin evolution of these black holes; the stellar mass black holes observed typically in a merger have a low spin, but the super-massive black holes in radio galaxies are believed to have extremal high spin. Stellar evolution simulations (work by Chieffi & Limongi and others) and observations suggest that these black holes may start at a high spin, but lose it rapidly due to the action of magnetic fields. Such magnetic fields could also accelerate particles such as protons and nuclei, to the highest energies observed, near to ZeV energies. This line of work is most exciting.

Outlook

Simulations of the early universe have to obey so many observational constraints, now black hole shadows as well as gravitational waves, and ever improving data of the local universe out to the redshift of recombination, that we can hope to achieve a physical understanding of the 95 % of the universe we do not really comprehend yet, as are baryonic intergalactic gas, ultra-high energy cosmic ray particles, dark matter and dark energy.

The Reionisation of the Local Universe in the HESTIA Suite

Luke Conaboy[1], **Ilian T. Iliev**[1], **and Noam I. Libeskind**[2]

[1] Astronomy Centre, Department of Physics & Astronomy, Pevensey II Building, University of Sussex, Falmer, Brighton BN1 9QH, United Kingdom
E-mail: {l.conaboy, i.t.iliev}@sussex.ac.uk

[2] Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
E-mail: nlibeskind@aip.de

The area around our Milky Way is the best-observed patch of the universe, making it the ideal testbed for galaxy formation models, particularly for processes where the environment of a galaxy is thought to play an important role. We present our promising work resimulating the original HESTIA suite using the RAMSES AMR code, with the goal of self-consistently including and studying the impact of ionising radiation from the first generation of stars during the Epoch of Reionisation.

1 Introduction

The neighbourhood in the immediate cosmological vicinity of the Milky Way (MW) is known as the "near field". Its main galaxies are the Milky Way-Andromeda (M31) pair, accompanied by an entourage of smaller galaxies including the Large and Small Magellanic Clouds, the Triangulum galaxy (M33) and a flock of low luminosity dwarf galaxies. Beyond $\sim 3\,\mathrm{Mpc}$, there are a few similarly sized large galaxies (e.g. Centaurus A, M83 and M81). At around $15\,\mathrm{Mpc}$, the Virgo galaxy cluster dominates the scene. A dark matter filament detected by the coherent motion of galaxies stretches from us to Virgo. The entire structure described above is confined to the super-galactic plane, which is itself abutted by the Local Void. On even larger scales, the Virgo cluster and the entire local volume forms part of the local super cluster, Laniakea, which is itself embedded in the "cosmic web". Our near field is the best-observed region of the universe, and consequently of prime interest to study, understand and simulate. Constrained cosmological simulations are ones where the initial conditions have been manipulated such that they are destined to produce predetermined large-scale cosmographic landscape, in our case our own local environment. In order to achieve this, the local cosmography must itself first be quantified. This is done via the measurement of the local peculiar velocity field obtained by comparing "standard candle" distance measures with redshifts. Initial conditions for cosmological simulations can then be constructed.

Such simulations are particularly useful when examining features of the local universe which we believe are unique to it, features which although may not make it "special" in a Copernican sense, are still inherent only to our patch of the universe – a bit like explaining the ridges on a cosmological fingerprint.

A number (one may argue all) of features of the Local Group (LG) of galaxies demand such constrained simulations which accurately model the formation of the local environment. Perhaps the most pressing is to attempt to model or at least understand the impact

of the Epoch of Reionisation (EoR) on the star formation history of the LG and on dwarf galaxies. Was the LG reionised predominantly from the Virgo cluster which, presumably, had UV producing stars forming in its progenitors eons before the LG? Or did the LG reionise "inside out", the reionising radiation being produced within its own confines? These are some of the important questions we will attempt to answer with this project.

2 Methods

2.1 Constrained Simulations

The technique of constrained realisations described here was first suggested by Hoffman & Ribak (1991)[1] and has been used before to successfully reproduce LGs analogues in the correct environment[2]. Simply put, given a data set and a prior model (e.g. ΛCDM with Gaussian fluctuations) which specifies a covariance matrix of the field that is to be reconstructed, the Wiener Filter (WF) seeks a linear minimum variance solution. On linear scales ($\sim 3\,\mathrm{Mpc}$) the velocity and density fields are related by a simple convolution. Thus, reconstructing the density field from the observed velocity field requires a deconvolution in the presence of noise[3]. Obscured or incomplete regions can be handled, although the reconstructed WF density field tends to zero (the mean) at the survey's limits, where the number of constraints goes to zero. The observed (linear) velocity field (reconstructed with a WF) is translated into a linear Zeldovich displacement field which can be reversed in time. This "reversed Zeldovich approximation"[4] combined with the WF and constrained realisation technique has already proved its ability to produce ICs that are fit to task[5]. Once we have determined the large-scale structure through the constrained realisation technique, unconstrained structure or structure below the scale of the reconstruction is added in a trial-and-error fashion, resulting in the HESTIA suite of simulations[6].

2.2 Initial Conditions

The initial conditions for our simulations are based on the ones from the HESTIA suite, but appropriately converted for use in RAMSES zoomed setup, while preserving their constrained nature. The technical details of this are as follows.

Our simulations are performed in a $100\ h^{-1}\ \mathrm{Mpc}$ box at a base resolution of 256^3 dark matter particles and cells (though the number of cells increases as the simulation progresses). We include the small scale information necessary for the formation of the LG by identifying the region surrounding the Milky Way (MW) and M31 at the present day ($z = 0$). We then tracked the particles within this region back to the start time of the simulation ($z = 99$). The size of this LG region at $z = 0$ depends on the resolution of the simulation (i.e. the number of particles and cells used), and varies from a sphere of radius $5\ h^{-1}\ \mathrm{Mpc}$ centred on the LG for the lowest resolution, to the union of two spheres of radius $2.5\ h^{-1}\ \mathrm{Mpc}$, one centred on the MW and the other on M31, for the highest. We then calculated a convex hull containing these particles and all the low resolution particles and cells within this hull are "refined" – replaced with many more high resolution particles. The large scale structure outside the zoom region is simulated with low resolution (higher-mass) particles, in order to keep the computational costs down, while still embedding the LG in the correct large-scale cosmological environment. This technique allows us to simulate the

LG and its surroundings with an effective resolution (i.e. the total number of particles or cells if the whole box were simulated at this resolution) of up to 16384^3, where the mass of the refined dark matter particles is 262144 times smaller than the low resolution particles. Simulating an entire box at this resolution is computationally infeasible, but through the zoom technique we are able to (fairly cheaply) increase the resolution only in our region of interest. We use a modified version of MUSIC[7] to generate masks, which dictate the extent of the refinement region, in the RAMSES-readable GRAFIC format. The density and velocity fields are produced by the GINNUNGAGAP code [a].

2.3 Simulation Methods

We use the adaptive mesh refinement (AMR) code RAMSES[8] to follow the evolution of dark matter, gas and stars within the cosmological context. Ionising radiation from stars is modelled using the publicly available radiative transfer (RT) extension RAMSES-RT[9], which solves the radiation hydrodynamics equations on the same AMR grid as the gas. Comparing resolution between different simulation methods is tricky, so we do not attempt to exactly match the softening length $\epsilon = 220$ pc of the original runs, instead opting to make our maximum cell size a similar order or magnitude, such that $\Delta x = 570$ pc. In the simulations containing gas and radiation, as well as dark matter, we allow star formation in any gas cell where i) the density is above some threshold n_*, ii) the local overdensity is 200 times the critical density of the universe $\rho_{\rm cr}$ and iii) the temperature of the gas is below 2×10^4 K. We test three different values of $n_* = 0.1$, 1 and 10 $\rm cm^{-3}$, in units of the number density of hydrogen atoms. Feedback from supernovae is included through the kinetic feedback model[10].

2.4 Structure Finding

Structure formation in the universe is hierarchical – small initial fluctuations in the distribution of dark matter collapse under gravity to form clumps called haloes. Over time, these haloes grow both by accreting matter and by merging with one another. We use the AHF code[11] to build catalogues of the dark matter haloes in the simulation. Merger trees, which follow the formation and evolution of haloes in time, are constructed using CONSISTENT-TREES[12]. We use the conversion tools supplied with CONSISTENT-TREES to convert our AHF outputs into the CONSISTENT-TREES format.

2.5 Cosmographic Criteria

To identify LG candidates from our $z = 0$ we apply the cosmographic criteria listed in the original HESTIA paper, which are a set of constraints on the local large-scale structure, as well as the properties of the LG itself. With regards to the large-scale structure, there must be a Virgo cluster candidate (with mass $M > 2 \times 10^{14}$ $\rm M_\odot$) within 7.5 Mpc of its expected formation location and there must be no other cluster more massive than the Virgo candidate within 20 Mpc. For the LG itself, it must form within 5 Mpc of its expected location and be within 3.5 Mpc of the true distance to Virgo. The masses of the MW and

[a] https://github.com/spilipenko/ginnungagap

M31 must be between 8×10^{11} $\mathrm{M}_{\odot}$ and 3×10^{12} $\mathrm{M}_{\odot}$, with the MW (the smaller mass halo) no less than half the size of M31. The haloes must be between 0.5 and 1.2 Mpc apart, with no other halo more massive than the MW within 2 Mpc of the midpoint. Finally, the haloes must be infalling.

3 Results

Different simulation methods produce different results, each method having their own strengths and weaknesses[13]. The original HESTIA simulations were carried out using the moving mesh code AREPO, whereas in this study we use the adaptive mesh refinement code RAMSES – so the first steps were to demonstrate convergence between the two codes. To begin with, we ran a series of dark matter-only simulations allowing us to fine-tune the refinement mask in the initial conditions to yield optimum results. Fig. 1 shows the LG produced by these simulations, where the large halo in the centre-left of each panel is the MW, and the halo in the centre-right is M31. The left panel of Fig. 1 shows the result at an effective resolution of 4096^3 in the zoom region, while the right panel shows the result of

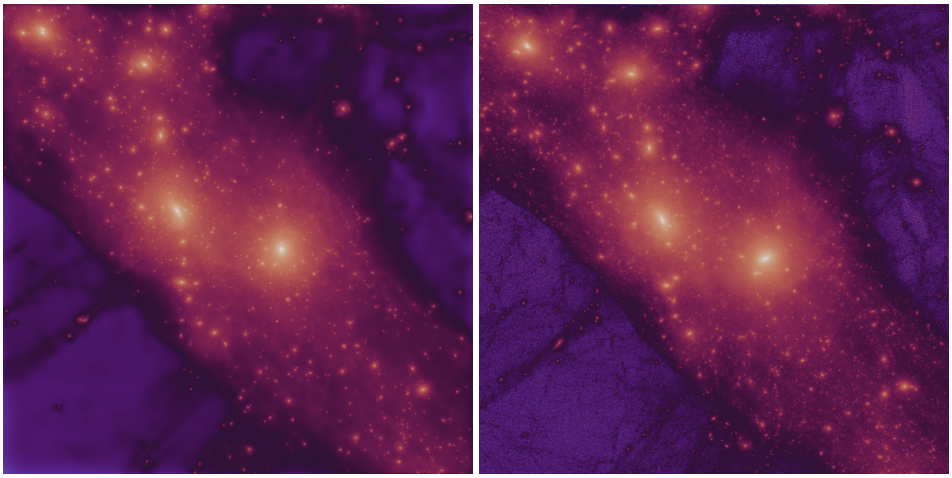


Figure 1. The result of going from 4096^3 (left) to 16384^3 (right) effective resolution on the simulated LG (each centre). With the higher resolution, more structure is able to develop on the smallest scales and on the larger scales more particles means we can resolve the internal dynamics in exquisite detail.

Effective resolution	Δd_{sep} (%)	ΔM_{M31} (%)	ΔM_{MW} (%)
4096^3	-22.20	-2.61	-3.26
16384^3*	-21.48	0.06	-2.41

Table 1. The DMO runs used to verify our ICs and compare to the original HESTIA runs. We calculate the fractional difference between our new runs and the HESTIA runs as $\Delta = (x_{\mathrm{new}} - x_{\mathrm{HESTIA}})/x_{\mathrm{HESTIA}}$. d_{sep} is the separation between MW and M31. M_{M31} is the mass of M31. M_{MW} is the mass of the MW. *The 16384^3 run does not have a counterpart in the original HESTIA suite, so we compare to the nearest available resolution, 8192^3.

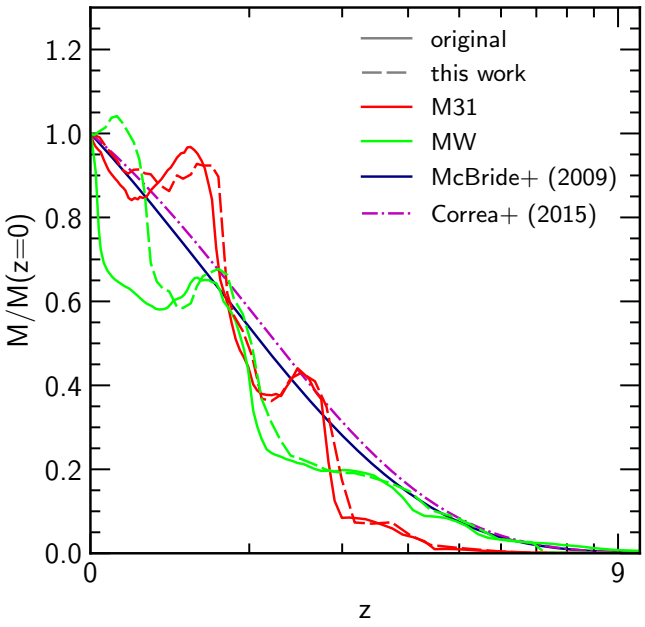


Figure 2. The mass M of M31 (red) and MW (green) as a function of z, normalised to the final mass $M(z = 0)$ for the new (long-dashed) and the original HESTIA (solid) simulations. Also shown are the analytic models of McBride et al. (2009)[14] (blue solid) and Correa et al. (2015)[15] (purple dot-dashed).

increasing this to an effective resolution of 16384^3. At the higher resolution, more structure can develop on the smallest scales (i.e. lowest masses) which does not exist in the lower resolution runs. We are also able to resolve the dynamics of the LG with increased accuracy, as can be seen in Tab. 1 where we show the fractional difference in halo properties between this work and the original HESTIA runs, with increasing resolution giving a better match. Once we have identified the LG in our simulations at $z = 0$ in our simulations, we must trace these haloes back in time to find where they were during the EoR. To do this, we compute merger trees, as described in Sec. 2.4. Merger trees allow us to trace back halo properties, such as mass and position, through cosmic time. Fig. 2 shows the growth of the LG in this work (in the 16384^3 run) and in the original HESTIA suite (in the 8192^3 run), compared to some analytic predictions. We find excellent agreement between this work and the original work, allowing for some expected deviation due to differing simulation codes and resolution as well as slight differences in the size, shape and location of the initial refinement mask. Next, we run a set of mid-resolution test simulations, with an effective resolution of 4096^3, including gas and radiation-hydrodynamics. In these test simulations, we vary the density threshold for star formation $n_\star$, as described in Sec. 2.3. Stars can only form inside the high-resolution region surrounding the LG, so we look at this region when computing quantities related to ionisation. Fig. 3 shows the average fraction of ionised hydrogen $x_{\rm HII}$ in a 2 h^{-1} Mpc sphere surrounding the midpoint of the LG, determined from the high-resolution merger trees. We show the three different star formation thresholds and the averaged calculated using a mass and volume weighting. Reionisation

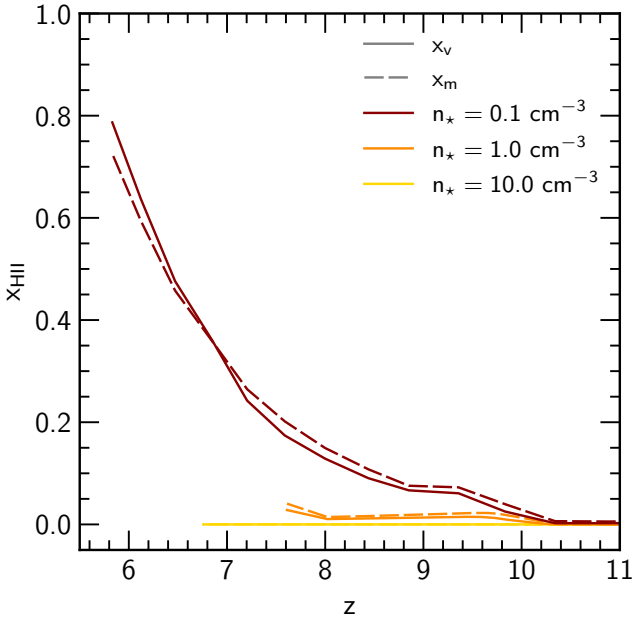


Figure 3. Hydrogen ionised fraction $x_{\rm HII}$ in a series of mid-resolution test simulations, with an effective resolution of 4096^3. The ionised fraction is averaged over a 2 h^{-1} Mpc sphere, centred on the midpoint of the LG and weighted by the computational cell volume (x_v, solid) or the mass within each computational cell (x_m, long-dashed). The position of the midpoint is determined at each timestep from the high-resolution 16384^3 merger trees. We show three different star formation thresholds $n_* = 0.1$ (red), 1 (orange) and 10 (yellow) cm^{-3}, as detailed in Sec. 2.3.

is expected to be complete by $z \sim 5$ (though for the volume size we consider here there can be a lot of scatter around this[16]), so only one of the thresholds, $n_\star = 0.1$ cm^{-3}, produces a realistic looking reionisation history.

4 Conclusion

We have demonstrated excellent agreement between our RAMSES AMR resimulations and the original constrained realisation suite HESTIA, both in the final properties of the LG and in their temporal evolution. Confident that we can accurately reproduce the expected structure at $z = 0$, we were able to produce precursor radiation-hydrodynamics tests in order to calibrate star formation parameters. From here, we are now able to produce the highest resolution radiation-hydrodynamics simulations of the LG, affording us the opportunity to explore the effect of reionisation on our local neighbourhood in exquisite detail. These extremely high-resolution simulations will afford an unprecedented view of the LG, allowing us to unravel the effect that the formation of the earliest stars had on, for example, the star formation history and dwarf galaxies of the LG.

Acknowledgements

The authors would like to make clear the integral contribution of Sergey Pilipenko to this work, through producing the initial conditions with the GINNUNGAGAP code. This work made use of the following packages: NUMPY[17], MATPLOTLIB[18], YT[19], YTREE[20] and PYSPHVIEWER[21]. LC is supported by an STFC studentship. 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[22] at Jülich Supercomputing Centre (JSC).

References

1. Y. Hoffman and E. Ribak, *Constrained realizations of Gaussian fields - A simple algorithm*, ApJL, **380**, L5–L8, Oct. 1991.
2. N. I. Libeskind, G. Yepes, A. Knebe, S. Gottlöber, Y. Hoffman, and S. R. A&A, *Constrained simulations of the Local Group: on the radial distribution of substructures*, MNRAS, **401**, 1889–1897, January 2010.
3. S. Zaroubi, Y. Hoffman, and A. Dekel, *Wiener Reconstruction of Large-Scale Structure from Peculiar Velocities*, ApJ, **520**, no. 2, 413–425, Aug. 1999.
4. T. Doumler, Y. Hoffman, H. Courtois, and S. Gottlöber, *Reconstructing cosmological initial conditions from galaxy peculiar velocities - I. Reverse Zeldovich Approximation*, MNRAS, **430**, no. 2, 888–901, Apr. 2013.
5. J. G. Sorce, H. M. Courtois, S. Gottlöber, Y. Hoffman, and R. B. Tully, *Simulations of the Local Universe constrained by observational peculiar velocities*, MNRAS, **437**, no. 4, 3586–3595, Feb. 2014.
6. N. I. Libeskind, E. Carlesi, R. J. J. Grand, A. Khalatyan, A. Knebe, R. Pakmor, S. Pilipenko, M. S. Pawlowski, M. Sparre, E. Tempel, P. Wang, H. M. Courtois, S. Gottlöber, Y. Hoffman, I. Minchev, C. Pfrommer, J. G. Sorce, V. Springel, M. Steinmetz, R. B. Tully, M. Vogelsberger, and G. Yepes, *The HESTIA project: simulations of the Local Group*, MNRAS, **498**, no. 2, 2968–2983, Oct. 2020.
7. O. Hahn and T. Abel, *Multi-Scale Initial Conditions for Cosmological Simulations*, MNRAS, **415**, no. 3, 2101, Aug. 2011.
8. R. Teyssier, *Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES*, A&A, **385**, 337–364, Apr. 2002.
9. J. Rosdahl, J. Blaizot, D. Aubert, T. Stranex, and R. Teyssier, *RAMSES-RT: Radiation Hydrodynamics in the Cosmological Context*, MNRAS, **436**, 2188–2231, Dec. 2013.
10. Y. Dubois and R. Teyssier, *On the Onset of Galactic Winds in Quiescent Star Forming Galaxies*, A&A, **477**, 79–94, Jan. 2008.
11. S. R. Knollmann and A. Knebe, *AHF: Amiga's Halo Finder*, ApJS, **182**, 608–624, June 2009.
12. P. S. Behroozi, R. H. Wechsler, H.-Y. Wu, M. T. Busha, A. A. Klypin, and J. R. Primack, *Gravitationally Consistent Halo Catalogs and Merger Trees for Precision Cosmology*, ApJ, **763**, no. 1, 18, Jan. 2013.
13. O. Agertz, B. Moore, J. Stadel, D. Potter, F. Miniati, J. Read, L. Mayer, A. Gawryszczak, A. Kravtsov, Å. Nordlund, F. Pearce, V. Quilis, D. Rudd,

V. Springel, J. Stone, E. Tasker, R. Teyssier, J. Wadsley, and R. Walder, *Fundamental differences between SPH and grid methods*, MNRAS, **380**, no. 3, 963–978, Sept. 2007.
14. J. McBride, O. Fakhouri, and C.-P. Ma, *Mass accretion rates and histories of dark matter haloes*, MNRAS, **398**, no. 4, 1858–1868, Oct. 2009.
15. C. A. Correa, J. S. B. Wyithe, J. Schaye, and A. R. Duffy, *The accretion history of dark matter haloes - I. The physical origin of the universal function*, MNRAS, **450**, no. 2, 1514–1520, June 2015.
16. I. T. Iliev, G. Mellema, U. L. Pen, H. Merz, P. R. Shapiro, and M. A. Alvarez, *Simulating cosmic reionization at large scales - I. The geometry of reionization*, MNRAS, **369**, no. 4, 1625–1638, July 2006.
17. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, *Array programming with NumPy*, Nature, **585**, no. 7825, 357–362, Sept. 2020.
18. J. D. Hunter, *Matplotlib: A 2D graphics environment*, Computing in Science & Engineering, **9**, no. 3, 90–95, 2007.
19. M. J. Turk, B. D. Smith, J. S. Oishi, S. Skory, S. W. Skillman, T. Abel, and M. L. Norman, *yt: A Multi-code Analysis Toolkit for Astrophysical Simulation Data*, ApJS, **192**, 9, Jan. 2011.
20. B. D. Smith and M. Lang, *ytree: A Python package for analyzing merger trees*, JOSS, **4**, no. 44, 1881, dec 2019.
21. A. Benitez-Llambay, *py-sphviewer: Py-SPHViewer v1.0.0*, July 2015.
22. Jülich Supercomputing Centre, *JUWELS Cluster and Booster: Exascale Pathfinder with Modular Supercomputing Architecture at Juelich Supercomputing Centre*, Journal of large-scale research facilities, **7**, no. A138, 2021.

Growth of Seed Black Holes in Galactic Nuclei

Rainer Spurzem[1,2], Francesco Rizzuto[3], Manuel Arca Sedda[4], Albrecht Kamlah[5,1], Peter Berczik[1,6,7], Qi Shu (舒琦)[8,2], Ataru Tanikawa (谷川 衝)[9], and Thorsten Naab[10]

[1] Astronomisches Rechen-Institut, Zentrum für Astronomie, University of Heidelberg, Mönchhofstrasse 12-14, 69120 Heidelberg, Germany
E-mail: spurzem@ari.uni-heidelberg.de

[2] Kavli Institute for Astronomy and Astrophysics, Peking University, Yiheyuan Lu 5, Haidian Qu, 100871, Beijing, China

[3] Department of Physics, University of Helsinki, Gustaf Hallstromin katu 2, 00014 Helsinki, Finland
E-mail: francesco.rizzuto@helsinki.fi

[4] Dep. of Physics and Astronomy "Galileo Galilei", University of Padova, Via Marzolo 8, 35131 Padova, Italy
E-mail: m.arcasedda@gmail.com

[5] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
E-mail: albrechtk@hotmail.de

[6] Main Astronomical Observatory, National Academy of Sciences of Ukraine, 27 Akademika Zabolotnoho St., 03143 Kyiv, Ukraine
E-mail: berczik@mao.kiev.ua

[7] Konkoly Observatory, Research Centre for Astronomy and Earth Sciences, Eötvös Loránd Research Network (ELKH), Konkoly Thege Miklós út 15-17, 1121 Budapest, Hungary

[8] Department of Astronomy, School of Physics, Peking University, Yiheyuan Lu 5, Haidian Qu, 100871, Beijing, China
E-mail: shuqi@pku.edu.cn

[9] Department of Earth Science and Astronomy, College of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
E-mail: tanikawa@g.ecc.u-tokyo.ac.jp

[10] Max-Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
E-mail: naab@MPA-Garching.MPG.DE

The evolution of dense star clusters is followed by direct high-accuracy N-body simulation. The problem is to first order a gravitational N-body problem, but stars evolve due to astrophysics and the more massive ones form black holes or neutron stars as compact remnants at the end of their life. After including updates of stellar evolution of massive stars and for the relativistic treatment of black hole binaries we find the growth of intermediate mass black holes and we show that in star clusters binary black hole mergers in the so-called pair creation supernova (PSN) gap occur easily. Such black hole mergers have been recently observed by the LIGO-Virgo-KAGRA (LVK) collaboration, a network of ground based gravitational wave detectors.

1 Introduction

It is one of the grand challenges of theoretical astrophysics to understand the dynamics of dense star clusters, both in their form as galactic globular clusters orbiting in the Milky Way halo as well as nuclear star clusters, surrounding the central supermassive black holes (SMBH) in our Galaxy and in other galaxies. High precision dynamical simulations of star clusters use direct orbit integration under the influence of (in principle) all other stars allowing precise modelling of diffusive transport processes of mass, energy, and angular momentum in the star cluster. This is important to understand relaxation processes in the system. The physical and astrophysical challenge is not only the gravitational million-body problem, but also the presence of a large number of very tight binaries and stellar evolution with black holes, neutron stars and white dwarfs forming. This becomes a strong multi-scale problem, with orbital time scales of days coupling to overall crossing times of the cluster of million years and ages of billion years.

The evolution of dense star clusters is not only governed by the aging of their stellar populations and simple Newtonian dynamics. The stellar densities become so high that stars can interact and collide, stellar evolution and binary stars change the dynamical evolution, black holes can accumulate in their centres and merge with relativistic effects becoming important. Recent high-resolution imaging has revealed even more complex structural properties with respect to stellar populations, binary fractions and compact objects as well as – the still controversial – existence of intermediate mass black holes in clusters of intermediate mass. Dense star clusters therefore are the ideal laboratory for the concomitant study of stellar evolution and Newtonian as well as relativistic dynamics. Last but not least black holes forming and evolving in dense star clusters are one of the prominent sources of gravitational waves (GW) across all frequency windows.

2 Current Astrophysical Updates

2.1 Stellar Evolution of Massive Stars

Before the first LIGO-Virgo-KAGRA (LVK) gravitational wave (GW) detection, many theoretical models of stellar evolution predicted stellar black holes (BHs) masses to be lower than $30\,\mathrm{M}_\odot$. These models remained unchallenged for several years because all stellar BHs observed at the time had masses $\lesssim 20\,\mathrm{M}_\odot$ (Ziółkowski 2008; Özel *et al.* 2010). Surprisingly, the first LVK detection, GW150914, revealed components more massive than $30\,\mathrm{M}_\odot$ (Abbott *et al.* 2016). Such masses had been predicted by stellar evolution models at low metallicity introducing a dependence between stellar winds mass loss and metallicity (see Woosley *et al.* 2002; Vink *et al.* 2001, and references therein).

This highlights the importance of up-to-date stellar evolution models for the correct interpretation and prediction of GW events. An accurate theory for the evolution of massive stars is particularly important to predict the mass distribution of stellar BHs at their formation. For this, precise models of stellar winds and a correct description of the last stages of the stellar evolution before the collapse are required. At the onset of stellar collapse, stars with sufficiently large helium cores undergo a phase of electron-positron pair production that in turn leads to one or more violent explosions. Depending on the initial mass of the core, the star can experience pulsation pair-instability supernovae (PPSN) getting partially

destroyed or it can experience the more violent pair-instability supernovae (PSN) and is destroyed completely[13,49,48]. Due to (P)PSN, isolated massive stars are not supposed to collapse into BHs in the mass range of approximately $50-130\,\mathrm{M}_{\odot}$. This gap in the stellar BH mass distribution is known as the (P)PSN mass gap. The mass limits of this gap are affected by various uncertainties and therefore they depend on the details of the stellar evolution adopted. In this study, the assumed mass gap is $45-195\,\mathrm{M}_{\odot}$.

Most of current updates of stellar evolution have been published for NBODY7[6] and for NBODY6++GPU[27]. The major issues are summarised here:

(i) New stellar wind models following Ref. 9, which in turn follow the wind mass-loss rates given by Ref. 44. With these models, BHs masses that originate from single stars depend strongly on the metallicity. For instance, a $100\mathrm{M}_{\odot}$ isolated main-sequence star would leave a $15\,\mathrm{M}_{\odot}$ BH at so- lar metallicity ($Z = 0.02$). At very low metallicity ($Z = 0.0002$), however, it can form a BH of about $60\mathrm{M}_{\odot}$, in the absence of pair-instability models.

(ii) Pair-instability supernova and pulsation pair-instability supernova models (according to Ref. 8 incorporated in the remnant formation and supernovae models as described in Ref. 14. Stars with helium core with masses $> 40\mathrm{M}_{\odot}$ undergo a violent phase of mass loss. For helium cores in the range between $60{-}135/,\mathrm{M}_{\odot}$ the star is completely destroyed.

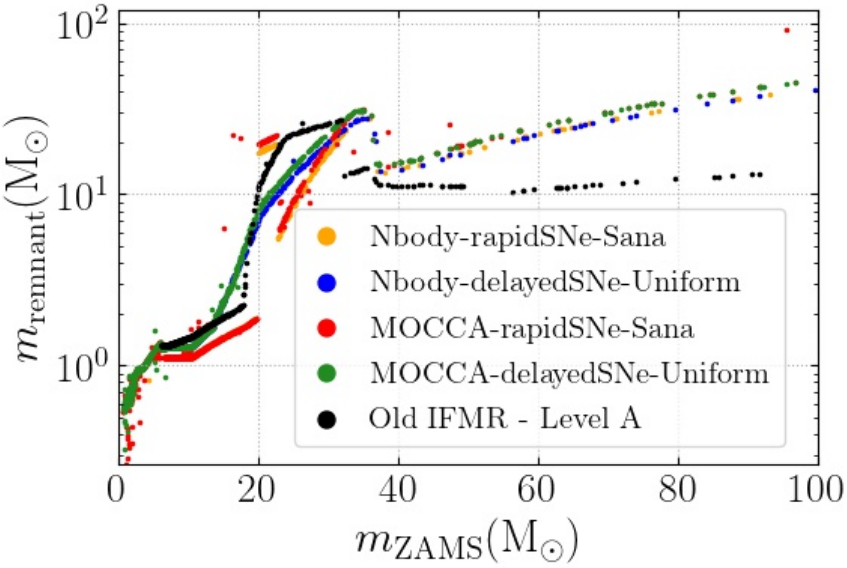


Figure 1. Initial-Final mass relation (IFMR) for the escaping compact objects of the MOCCA and Nbody6++GPU (Nbody) simulations. The keys refer to the Nbody-delayedSNe-Uniform, Nbody-rapidSNe-Sana, MOCCA-delayedSNe-Uniform, and MOCCA-rapidSNe-Sana simulations, respectively. They differ by the use of the code (Nbody or MOCCA), and the remnant mass prescription for the stellar mass black holes up to approximately the pair-instability mass gap (see Fryer *et al.* 2012). The black points show BH masses from another N-body simulation with Level A parameters[10] (Plot taken from Ref. 27).

(iii) New prescription of BHs and NSs natal kick velocities that explicitly depend on the fallback fraction[6].

(iv) A model for electron capture supernovae (ECSN) that produces neutron stars with low-velocity kicks that are therefore likely retained in medium-size star clusters[36,16].

Fig. 1 shows an example of how important the new updates of stellar evolution are. It shows the initial-final mass relation (IFMR), which tells what is the mass of a stellar evolution remnant as a function of the initial mass of a star. On the y-axis all objects more massive than a few solar masses are black holes; the figure compares the old IFMR with two variants of the new one (delayed and rapid supernovae explosions), for our direct N-body simulations as well as for an approximate Monte Carlo model (MOCCA)[18,17].

2.2 Preparation for PopIII Stars

Fitting formulae have been worked out for evolution tracks of massive stars with $8 \lesssim M \lesssim 160\,\mathrm{M}_\odot$ under extreme metal poor (EMP) environments for $\log(Z/\mathrm{Z}_\odot) = -2$, -4, -5, -6, and -8, where $\mathrm{M}_\odot$ and $\mathrm{Z}_\odot$ are the solar mass and metallicity, respectively[43]. They are based on reference stellar models, newly obtained by simulating the time evolutions of EMP stars. The fitting formulae take into account stars ending as blue supergiant (BSG) stars, and stars skipping the Hertzsprung gap (HG) phases and blue loops, which are characteristic of massive EMP stars. Here stars may remain BSG stars when they finish their core Helium burning (CHeB) phase. The fitting formulae are in good agreement with the stellar evolution models, and are now used for our NBODY6++GPU code within the SSE/BSE packages; they are also used for other codes such as PeTar and NBODY7 (see discussion in Sec. 4). The algorithms should be useful to generate theoretical predictions for black holes and black hole mergers under EMP environments. The work with NBODY6++GPU is currently in progress on the JUWELS Booster system (Kamlah, Tanikawa, *et al.*, in preparation).

2.3 Relativistic Dynamics of Compact Objects

The gravitational energy loss and resulting merger of compact objects is computed following the orbit-averaged approach[35], allowing for the final coalescence if the orbit shrinking time due to gravitational wave emission becomes shorter than an orbital time. We have added the following further updates, which affect the formation and evolution of black holes[40]:

(i) For collisions between a compact remnant and a main sequence star or red giant a free parameter f_c is introduced, which describes the mass loss from the system in the process. The previous NBODY6 versions used only $f_c = 1$, i.e. no mass loss in the process.

(ii) Simultaneous treatment of classical tidal interactions (Roche lobe overflow) and Post-Newtonian orbit-averaged orbit shrinking due to gravitational wave emission has been made possible. Both are treated technically in a similar way, and can now be switched on together.

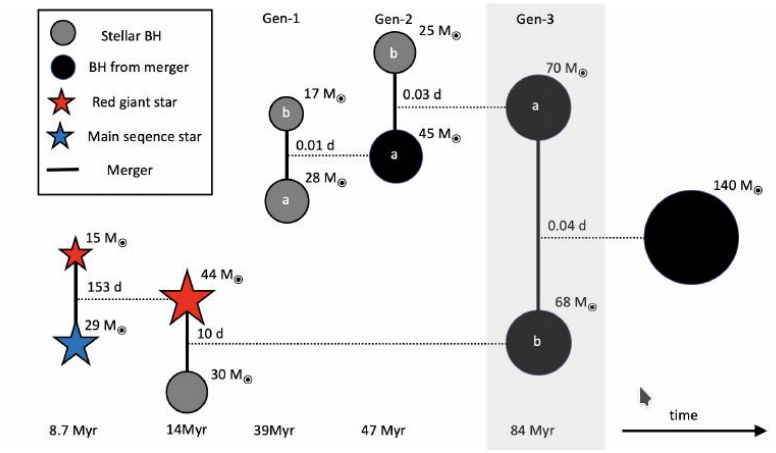


Figure 2. Visualisation of the formation path towards a "mass gap" merger (grey region, third generation) of two black holes with mass $70M_\odot$ and $68M_\odot$ developed in one of our N-body simulations. The more massive BH (third generation) grew by two preceding mergers involving black holes (first and second generation). The lower mass BH (third generation) was created in a stellar merger of a red giant with a main sequence star followed by the collision with a stellar mass BH. The masses of the components (in solar masses) and the orbital periods (in days) are indicated at the respective times of the merger (black horizontal lines) after the start of the simulation (Figure from Ref. 3).

(iii) Strongly bound binaries of two compact objects, which are subject to Post-Newtonian relativistic energy loss are prevented from unperturbed two-body integration, and defined as a new type of binary in the code.

Fig. 2 shows an example, and illustrates, how an intermediate mass black hole is formed through several steps, involving collisions of two massive stars, a collision between a massive star and a black hole and relativistic mergers of black hole binaries. The procedure and parameters described above will affect rates and timescales of the black hole formation process.

2.4 Relativistic Recoils at Coalescence and Spins

Another important physical process that has not yet been used for the published papers so far is relativistic recoil for compact object coalescences. Its absence might artificially enhance the probability of forming massive black holes[5,4]. We have implemented now the relativistic recoils following Ref. 32, and currently tests are ongoing on the JUWELS Booster system (Arca Sedda *et al.*, in prep.).

Regarding spins (have been neglected so far, and are important to determine proper values of recoil at coalescence) we have now included in our simulation code a new spin treatment[7], which provides three alternative models for spin evolution of stars, with variable efficiency of angular momentum transport, as well as further options to choose different spin configurations for single and binary black holes (Arca Sedda *et al.*, in prep.). The

one with highest efficiency by Ref. 15 delivers at the end point of massive star evolution black holes with low spins, consistent with current LVK observations.

Last, but not least the ultimate goal is to use a full Post-Newtonian dynamics inside the regularised binary motion. In such approach relativistic periastron precession and energy loss due to gravitational radiation can be followed at each point of the orbit using a generalised quasi-classical equation of motion[28]. Recently also the inclusion of spin dynamics and spin-spin and spin-orbit interactions are included into the Post-Newtonian approach in our codes[41]. Our cited papers describe the method to integrate this into our codes; any reader interested in the relativistic theory for Post-Newtonian dynamics please refer to papers cited therein.

3 Results

3.1 Intermediate Mass Black Hole Formation

Young dense massive star clusters are a promising environment for the formation of intermediate mass black holes (IMBHs) through collisions (of massive stars) and coalescences (of smaller mass black holes). We have published a set of 80 simulations carried out with NBODY6++GPU using 10 different initial conditions, and shown that an IMBH can form in some cases with – so far – up to $350\,\mathrm{M}_\odot$. We simulated compact star clusters with $1.1 \cdot 10^5$ particles ($\sim 7 \cdot 10^4\,\mathrm{M}_\odot$, core density of $\sim 10^5\,\mathrm{M}_\odot\mathrm{pc}^{-3}$) with a resolved stellar population with 10% initial (primordial) hard binaries, and find that very massive stars with masses up to $\sim 400\,\mathrm{M}_\odot$ grow rapidly by binary exchange and three-body scattering events with main sequences stars in hard binaries. From them IMBHs with masses up to $350\,\mathrm{M}_\odot$ form on timescales of order 15 Myr; the final mass depends critically on an unknown parameter describing how much mass is accreted, if a black hole collides with a main sequence star (see description above of f_c parameter).

It is the first time that such IMBH formation process has been shown in a direct N-body simulation of that quality and particle resolution. This process was qualitatively predicted from Monte Carlo MOCCA simulations[18,17]. After formation, the IMBHs can experience occasional mergers with stellar mass black holes in intermediate mass-ratio inspiral events on a 100 Myr timescale. For more details please compare Ref. 40.

3.2 Black Hole Merger in the Forbidden Zone

The LIGO-Virgo-KAGRA Collaboration (LVC) discovered GW190521, a gravitational wave (GW) source associated with the merger between two black holes (BHs) with masses of 66 and $>85\,\mathrm{M}_\odot$. GW190521 represents the first BH binary merger with a primary mass falling in the PSN mass gap (see explanation in Sec. 2.1) and leaving behind a $\sim 150\,\mathrm{M}_\odot$ remnant. So far, the LVC has reported the discovery of four further mergers having a total mass $> 100\,\mathrm{M}_\odot$, i.e., in the intermediate-mass black hole (IMBH) mass range. In our simulations we discover the development of a GW190521-like system as the result of a third-generation merger, and furthermore four IMBH-BH mergers with total mass $(300 - 350)\,\mathrm{M}_\odot$. We show that these IMBH-BH mergers are low-frequency GW sources detectable with LISA and Decihertz Interferometer Gravitational wave Observatory (DECIGO) out to redshift $z = 0.01 - 0.1$ and $z > 100$, and we discuss how their detection

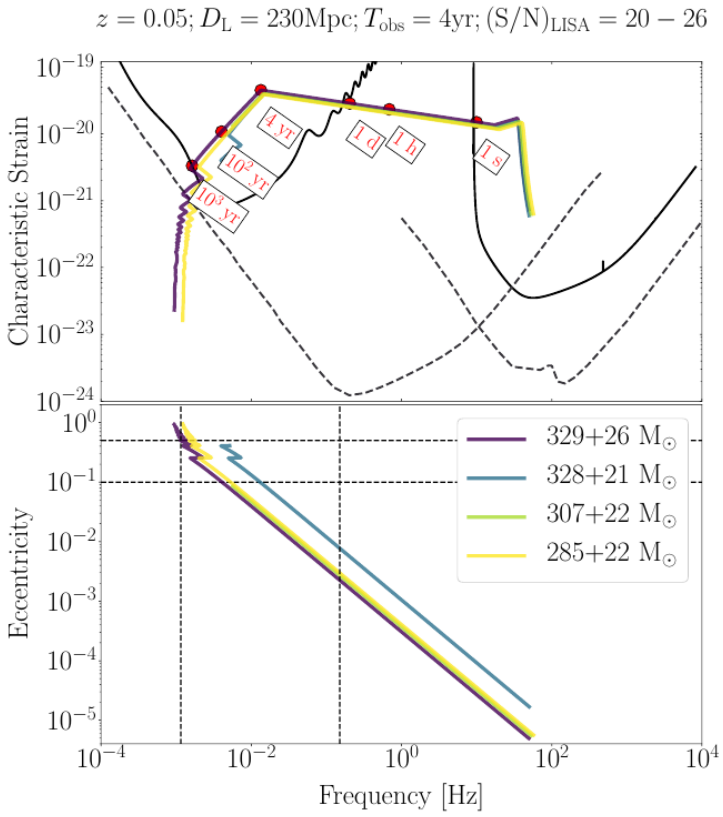


Figure 3. Top panel: GW strain evolution as a function of the frequency for the mergers beyond the mass-gap (coloured lines). Simulated tracks are overlapped to the sensitivity curves of LISA and LIGO-Virgo-KAGRA (LVK, solid black lines), and DECIGO and Einstein Telescope (dashed black lines). The white boxes indicate the time to merger for the heaviest merger. Bottom panel: eccentricity evolution as a function of the frequency. The vertical lines identify the moment in which the IMBH-BH mergers enter and exit the LISA sensitivity window, whereas horizontal lines identify the eccentricity values $e = 0.1, 0.5$. All the mergers are assumed to happen at a redshift $z = 0.05$, corresponding to a luminosity distance $D_L = 230$ Mpc. The figure title report the typical signal-to-noise ratio (S/N) $\sim 20-26$, assuming for LISA a $T_{\rm obs} = 4$ yr long mission. Figure from Ref. 3.

could help unravelling IMBH natal spins. For the GW190521 test case, we show that the third-generation merger remnant has a spin and effective spin parameter that matches the 90% credible interval measured for GW190521 better than a simpler double merger and comparable to a single merger. Due to GW recoil kicks, we show that retaining the products of these mergers require birth sites with escape velocities $\simeq 50-100$ km/s, values typically attained in galactic nuclei and massive clusters with steep density profiles. So, this is an explanation why observed black hole mergers do occur in the PSN mass gap – they are the result of several generation mergers in star clusters[3,39].

We have presented more direct N-body simulations, carried out with NBODY6++GPU, of young and compact low-metallicity (Z = 0.0002) star clusters[39]. Early on, after tens of

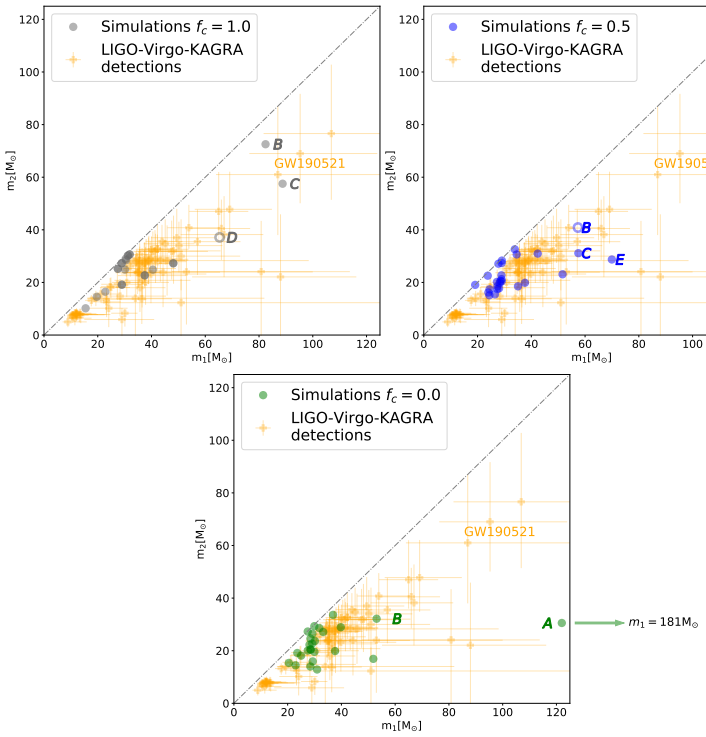


Figure 4. The panels show the primary (m_1) and secondary (m_2) masses of all BH mergers in the simulations for an accretion fraction of $f_c = 1.0$ (left, grey circles), $f_c = 0.5$, (centre, blue circles) and $f_c = 0.0$, (bottom, green circles). The currently available LIGO-Virgo-KAGRA gravitational wave detections including error bars are indicated in orange. BH merger events that might be excluded due to gravitational recoil kicks are indicated with open circles. In general, the simulated events cover a similar parameter space as all currently available observations. The $f_c = 1.0$, simulations provide two possible formation paths for GW190521. Path C is a second-generation event and has a low probability due to a first-generation BH merger. Path B is more likely as the event itself is a first-generation BH merger. One of the $f_c = 0.0$, realisations generated an intermediate-mass ratio inspiral of two black holes with 31 and 181 $M_\odot$), respectively, as shown in the bottom panel (Figure from Ref. 39).

Myrs, every simulated cluster hosts several black hole merger events which nearly cover the complete mass range of primary and secondary black hole masses for current LIGO-Virgo-KAGRA gravitational wave detections. The importance of gravitational recoil is estimated statistically during post-processing analysis. We presented possible formation paths of massive black holes above the assumed lower PSN mass-gap limit ($45\,M_\odot$) into the intermediate mass black hole (IMBH) regime ($> 100\,M_\odot$) which include collisions of stars, black holes and the direct collapse of stellar merger remnants with low core masses. Fig. 3 shows how the black hole binaries found in our simulations would show up in the detection sensitivity diagrams of current ground and future space based gravitational wave detectors. Finally, Fig. 4 compares in a statistical way current LIGO-Virgo-KAGRA detections of binary black hole mergers with a collection of such mergers from our simulations, using in three different panels a different mass loss factor f_c for star-black hole collisions.

While we need more observational and simulation data to improve the statistical quality it can be seen already that such comparisons would allow to constrain the physics of black hole star collisions. In a similar way in the future modelling the spin of black holes could and comparing with observations could tell us something about initial and final spins of black holes (which is quite difficult to observe directly via gravitational waves, at least currently).

4 Initial Models and Codes

4.1 Initial Models

The simulations take into account full stellar evolution as well as the formation and evolution of binary stars. We set the number of primordial binaries typically to be 5 or 10 % of the total number of systems (these are persistent hard binaries with binding energy higher than the r.m.s. random kinetic energy of a star; a much larger number of soft binaries present at the time of formation of the star cluster is being disrupted early on).

4.2 NBODY6++GPU

NBODY6++GPU[a] is a high-precision direct N-body simulation code based on the earlier N-body codes NBODY1-6[1] and NBODY6++[42]. It uses for time integration Taylor series up to 4^{th} order; due to the Hermite scheme it can be based on two time points only. This together with the hierarchically blocked variable time step scheme allows an efficient parallelisation of the code for massively parallel supercomputers (since NBODY6++); gravitational forces between particles are offloaded to graphics processing units (GPUs), used for high-performance general purpose computing (NBODY6++GPU[47]). The parallelisation is achieved via MPI and OpenMP on the top level, distributing work within a group of particles due for time integration, and efficient parallel use of GPU cores at the base level (every MPI process using a GPU), for computing the gravitational forces between particles. The GPU implementation in NBODY6++GPU provides a significant performance improvement, especially for the long-range (regular) gravitational forces (see Ref. 33, 47). Benchmarks and profiling are published in detail in Ref. 46, 20.

Recipes to simulate single, binary, and multiple objects stellar evolution are based on the SSE and BSE programs by Hurley[25,24,23] (see also code references 21, 22). It includes rapid tidal circularisation for binaries with small pericenters and tidal captures[30]. The integrator fully resolves orbits and dynamical evolution of binaries, even during phases of mass loss or when one of the two stars undergoes a supernova explosion. The binary orbit is adjusted to the corresponding loss of mass, energy and angular momentum with appropriate time stepping; in case of a supernova explosion it is always ensuring that the remnant and its companion leave the explosion with the corrected orbital positions and velocities. Further recent improvements are metallicity dependent winds, delayed or rapid white dwarf kicks[14], electron-capture and pair instability supernova events. The reader interested in more details please refer to our papers[6,27], which summarise the code updates and give all necessary further references.

[a] Link to repositories: tarballs/git service at
`https://zenodo.org/record/6511341/files/Nbody6%2B%2BGPU-Jan2022.tgz`
`https://github.com/kaiwu-astro/Nbody6PPGPU-beijing`

4.3 New Developments

A completely new code, called PeTaR[45] has been introduced. This code is written in a coherent, modular way. It is supposed to contain all the physics of NBODY6++GPU, but since it is a new code a lot of testing and comparison is still required. Algorithmically it has two advantages over NBODY6++GPU currently, one is the efficient parallelisation of hard binaries, and the other is the replacement of distant regular gravitational forces from particles by a TREE based scheme, which makes it possible to take into account also small perturbations on binaries from distant objects without too much computational cost. It is possible to upgrade NBODY6++GPU accordingly; we expect that both codes will co-exist for foreseeable time. Nevertheless we have intensively tested PeTaR, see Fig. 5 and may use it also in the next computing period in this project.

A novel hierarchical $4^{\rm th}$ fourth-order forward symplectic integrator and its numerical implementation has been shown in a new GPU-accelerated direct-summation N-body code named FROST[37]. The new integrator also uses an innovative MSTAR chain scheme[38] instead of the classical and algorithmic chains by Ref. 31. The integrator claims to be especially suitable for simulations with a large dynamical range due to its hierarchical nature, and for direct-summation N-body simulations beyond $N = 10^6$ particles on systems with several hundred and more GPUs. In that respect it is very similar to the classic $4^{\rm th}$ order Hermite codes such as φGRAPE[19,11] or HiGPU[12]. These codes, lacking the Ahmad-Cohen neighbour scheme, can easily used hundreds of GPUs, because they compute full long-range gravitational forces even for the smallest time steps. So, they gain parallelism by introducing order N unnecessary computations, which can be avoided by using NBODY6++GPU. In its AC neighbour scheme full force calculations (done on the GPUs) are only done in order ten times larger time intervals than the smallest steps. For the smallest steps only order 50-200 neighbour particle forces are required for force calculation, and this is done efficiently by using OpenMP on the multi-core host CPU[47,46]. Therefore, NBODY6++GPU often appears to be not to scale well to large GPU numbers, because it already obtains physical results comparable to φGRAPE or HiGPU (with many GPUs) by using only few GPUs and OpenMP with CPU cores efficiently.

It should be noted, however, that the new code PeTar uses a TREE scheme for distant

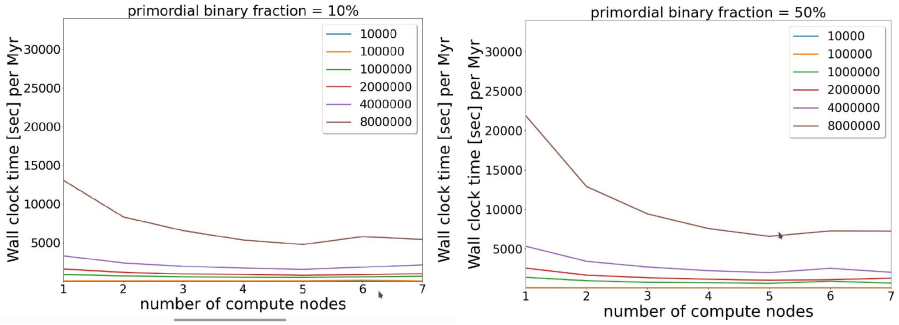


Figure 5. Strong scaling of the novel PeTar code, showing wall clock times obtained on the JUWELS Booster as a function of number of compute nodes.

forces, which in principle can be competitive with the AC scheme. A full quantitative profiling analysis of these issues is still missing. However, Fig. 5 shows first benchmarks done by our team using the PeTar code, which are very promising and open up the path to 10^7 particles in direct N-body simulation, especially with large binary fractions (up to 50%, a region in which NBODY6++GPU currently lags behind).

Acknowledgements

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 (2021) at Jülich Supercomputing Centre (JSC). MAS is a Marie Curie Skłodowska Fellow and Alexander von Humboldt Fellow. AK is a fellow of the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg (IMPRS-HD). PB and RS acknowledge the support of the Volkswagen Foundation under the Trilateral Partnerships 626 grant No. 97778 and the Strategic Priority Research Program (Pilot B) Multi-wavelength gravitational wave universe of Chinese Academy of Sciences (No. XDB23040100). PB acknowledges support from Volkswagen Foundation under the special stipend No. 9B870 (2022), from a President's International Fellowship for Visiting Scientists of Chinese Academy of Sciences, from the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grants No. AP08856184 and AP08856149), from the National Academy of Sciences of Ukraine under the Main Astronomical Observatory GPU computing cluster project No. 13.2021.MM, and by the special program of the National Research Foundation of the Ukraine 'Leading and Young Scientists Research Support' - "Astrophysical Relativistic Galactic Objects (ARGO): life cycle of active nucleus", No. 2020.02/0346. We thank Sambaran Banerjee, Mirek Giersz, Jarrod Hurley, Arek Hypki, Nadine Neumayer, Long Wang, Kai Wu, Roberto Capuzzo-Dolcetta, Andreas Just, M.B.N. (Thijs) Kouwenhoven, Xiaoying Pang for helpful discussions, collaboration, and hospitality during visits. MAS acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101025436 (project GRACE-BH, PI Manuel Arca Sedda).

References

1. S. J. Aarseth, *From NBODY1 to NBODY6: The Growth of an Industry*, PASP **111**, 1333–1346, 1999.
2. B. P. Abbott, et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, PhRvL **116**, 061102-, 2016.
3. M. Arca-Sedda, F. P. Rizzuto, T. Naab, J. Ostriker, M. Giersz, R. Spurzem, *Breaching the Limit: Formation of GW190521-like and IMBH Mergers in Young Massive Clusters*, ApJ **920**, 128-, 2021.
4. S. Banerjee, *Merger rate density of stellar-mass binary black holes from young massive clusters, open clusters, and isolated binaries: Comparisons with LIGO-Virgo-KAGRA results*, PhRvD **105**, 023004-, 2022.

5. S. Banerjee, *Stellar-mass black holes in young massive and open stellar clusters - IV. Updated stellar-evolutionary and black hole spin models and comparisons with the LIGO-Virgo O1/O2 merger-event data*, MNRAS **500**, 3002–3026, 2021.
6. S. Banerjee, K. Belczynski, C. L. Fryer, P. Berczik, J. R. Hurley, R. Spurzem, L. Wang, *BSE versus StarTrack: Implementations of new wind, remnant-formation, and natal-kick schemes in NBODY7 and their astrophysical consequences*, A&A **639**, A41-, 2020.
7. K. Belczynski, J. Klencki, C. E. Fields, et al., *Evolutionary roads leading to low effective spins, high black hole masses, and O1/O2 rates for LIGO/Virgo binary black holes*, A&A **636**, A104-, 2020.
8. K. Belczynski, A. Heger, W. Gladysz, A. J. Ruiter, S. Woosley, G. Wiktorowicz, H.-Y. Chen, T. Bulik, R. O'Shaughnessy, D. E. Holz, C. L. Fryer, E. Berti, *The effect of pair-instability mass loss on black-hole mergers*, A&A **594**, A97-, 2016.
9. K. Belczynski, T. Bulik, C. L. Fryer, A. Ruiter, F. Valsecchi, J. S. Vink, J. R. Hurley, *On the Maximum Mass of Stellar Black Holes*, ApJ **714**, 1217–1226, 2010.
10. K. Belczynski, V. Kalogera, T. Bulik, *A Comprehensive Study of Binary Compact Objects as Gravitational Wave Sources: Evolutionary Channels, Rates, and Physical Properties*, ApJ **572**, 407–431, 2002.
11. P. Berczik, R. Spurzem, L. Wang, S. Zhong, S. Huang, *Up to 700k GPU cores, Kepler, and the Exascale future for simulations of star clusters around black holes*, hpc..conf, 52–59, 2013.
12. R. Capuzzo-Dolcetta, M. Spera, D. Punzo, *A fully parallel, high precision, N-body code running on hybrid computing platforms*, JCoPh **236**, 580–593, 2013.
13. W. A. Fowler, F. Hoyle, *Neutrino Processes and Pair Formation in Massive Stars and Supernovae.*, ApJS **9**, 201-, 1964.
14. C. L. Fryer, K. Belczynski, G. Wiktorowicz, M. Dominik, V. Kalogera, D. E. Holz, *Compact Remnant Mass Function: Dependence on the Explosion Mechanism and Metallicity*, ApJ **749**, 91-, 2012.
15. J. Fuller, L. Ma, *Most Black Holes Are Born Very Slowly Rotating*, ApJL **881**, L1-, 2019.
16. A. Gessner, H.-T. Janka, *Hydrodynamical Neutron-star Kicks in Electron-capture Supernovae and Implications for the CRAB Supernova*, ApJ **865**, 61-, 2018.
17. M. Giersz, A. Askar, L. Wang, A. Hypki, A. Leveque, R. Spurzem, *MOCCA survey data base- I. Dissolution of tidally filling star clusters harbouring black hole subsystems*, MNRAS **487**, 2412–2423, 2019.
18. M. Giersz, N. Leigh, A. Hypki, N. Lützgendorf, A. Askar, *MOCCA code for star cluster simulations - IV. A new scenario for intermediate mass black hole formation in globular clusters*, MNRAS **454**, 3150–3165, 2015.
19. S. Harfst, A. Gualandris, D. Merritt, R. Spurzem, S. Portegies Zwart, P. Berczik, *Performance analysis of direct N-body algorithms on special-purpose supercomputers*, NewA **12**, 357–377, 2007.
20. S.-Y. Huang, R. Spurzem, P. Berczik, *Performance analysis of parallel gravitational N-body codes on large GPU clusters*, Research in Astronomy and Astrophysics **16**, 11-, 2016.
21. J. R. Hurley, O. R. Pols, C. A. Tout, *SSE: Single Star Evolution*, ascl.soft, ascl:1303.015-, 2013.

22. J. R. Hurley, C. A. Tout, O. R. Pols, *BSE: Binary Star Evolution*, ascl.soft, ascl:1303.014-, 2013.
23. J. R. Hurley, O. R. Pols, S. J. Aarseth, C. A. Tout, *A complete N-body model of the old open cluster M67*, MNRAS **363**, 293–314, 2005.
24. J. R. Hurley, C. A. Tout, O. R. Pols, *Evolution of binary stars and the effect of tides on binary populations*, MNRAS **329**, 897–928, 2002.
25. J. R. Hurley, O. R. Pols, C. A. Tout, *Comprehensive analytic formulae for stellar evolution as a function of mass and metallicity*, MNRAS **315**, 543–569, 2000.
26. Jülich Supercomputing Centre, *JUWELS Cluster and Booster: Exascale Pathfinder with Modular Supercomputing Architecture at Juelich Supercomputing Centre*, Journal of large-scale research facilities **7**, A138, 2021.
27. A. W. H. Kamlah, A. Leveque, R. Spurzem, M. Arca Sedda, A. Askar, S. Banerjee, P. Berczik, M. Giersz, J. Hurley, D. Belloni, L. Kühmichel, L. Wang, *Preparing the next gravitational million-body simulations: evolution of single and binary stars in NBODY6++GPU, MOCCA, and MCLUSTER*, MNRAS **511**, 4060–4089, 2022.
28. G. Kupi, P. Amaro-Seoane, R. Spurzem, *Dynamics of compact object clusters: a post-Newtonian study*, MNRAS **371**, L45–L49, 2006.
29. P. Kustaanheimo, E. L. Stiefel, *Perturbation theory of Kepler motion based on spinor regularization*, J. Reine Angew. Math. **218**, 204, 1965.
30. R. A. Mardling, S. J. Aarseth, *Tidal interactions in star cluster simulations*, MNRAS **321**, 398–420, 2001.
31. S. Mikkola, S. J. Aarseth, *An efficient integration method for binaries in N-body simulations*, NewA **3**, 309–320, 1998.
32. J. Morawski, M. Giersz, A. Askar, K. Belczynski, *MOCCA-SURVEY Database I: Assessing GW kick retention fractions for BH-BH mergers in globular clusters*, MNRAS **481**, 2168–2179, 2018.
33. K. Nitadori, S. J. Aarseth, *Accelerating NBODY6 with graphics processing units*, Monthly Notices of the Royal Astronomical Society **424**, 545–552, 2012.
34. F. Özel, D. Psaltis, R. Narayan, J. E. McClintock, *The Black Hole Mass Distribution in the Galaxy*, ApJ **725**, 1918–1927, 2010.
35. P. C. Peters, J. Mathews, *Gravitational Radiation from Point Masses in a Keplerian Orbit*, PhRv **131**, 435–440, 1963.
36. P. Podsiadlowski, N. Langer, A. J. T. Poelarends, S. Rappaport, A. Heger, E. Pfahl, *The Effects of Binary Evolution on the Dynamics of Core Collapse and Neutron Star Kicks*, ApJ **612**, 1044–1051, 2004.
37. A. Rantala, T. Naab, V. Springel, *frost: a momentum-conserving CUDA implementation of a hierarchical fourth-order forward symplectic integrator*, MNRAS **502**, 5546–5562, 2021.
38. A. Rantala, P. Pihajoki, M. Mannerkoski, P. H. Johansson, T. Naab, *MSTAR - a fast parallelized algorithmically regularized integrator with minimum spanning tree coordinates*, MNRAS **492**, 4131–4148, 2020.
39. F. P. Rizzuto, T. Naab, R. Spurzem, M. Arca-Sedda, M. Giersz, J. P. Ostriker, S. Banerjee, *Black hole mergers in compact star clusters and massive black hole formation beyond the mass gap*, MNRAS **512**, 884–898, 2022.
40. F. P. Rizzuto, T. Naab, R. Spurzem, M. Giersz, J. P. Ostriker, N. C. Stone, L. Wang, P. Berczik, M. Rampp, *Intermediate mass black hole formation in compact young massive star clusters*, MNRAS **501**, 5257–5273, 2021.

41. M. Sobolenko, P. Berczik, R. Spurzem, *Merging timescale for the supermassive black hole binary in interacting galaxy NGC 6240*, A&A **652**, A134-, 2021.
42. R. Spurzem, *Direct N-body Simulations*, JCoAM **109**, 407–432, 1999.
43. A. Tanikawa, T. Yoshida, T. Kinugawa, K. Takahashi, H. Umeda, *Fitting formulae for evolution tracks of massive stars under extreme metal-poor environments for population synthesis calculations and star cluster simulations*, MNRAS **495**, 4170–4191, 2020.
44. J. S. Vink, A. de Koter, H. J. G. L. M. Lamers, *Mass-loss predictions for O and B stars as a function of metallicity*, A&A **369**, 574–588, 2001.
45. L. Wang, M. Iwasawa, K. Nitadori, J. Makino, *PETAR: a high-performance N-body code for modelling massive collisional stellar systems*, MNRAS **497**, 536–555, 2020.
46. L. Wang, R. Spurzem, S. Aarseth, M. Giersz, A. Askar, P. Berczik, T. Naab, R. Schadow, M. B. N. Kouwenhoven, *The DRAGON simulations: globular cluster evolution with a million stars*, MNRAS **458**, 1450–1465, 2016.
47. L. Wang, R. Spurzem, S. Aarseth, K. Nitadori, P. Berczik, M. B. N. Kouwenhoven, T. Naab, *NBODY6++GPU: ready for the gravitational million-body problem*, MNRAS **450**, 4070–4080, 2015.
48. S. E. Woosley, *Pulsational Pair-instability Supernovae*, ApJ **836**, 244-, 2017.
49. S. E. Woosley, S. Blinnikov, A. Heger,*Pulsational pair instability as an explanation for the most luminous supernovae*, Natur **450**, 390–392, 2007.
50. S. E. Woosley, A. Heger, T. A. Weaver, *The evolution and explosion of massive stars*, RvMP **74**, 1015–1071, 2002.
51. J. Ziółkowski, *Masses of Black Holes in the Universe*, Chinese Journal of Astronomy and Astrophysics Supplement **8**, 273–280, 2008.

Theoretical Chemistry

Theoretical Chemistry:

From Quantum Monte Carlo Based *ab initio* Molecular Dynamics to QM/MM Simulations of Fluoride Transport

Christine Peter

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

In the proceedings of the NIC Symposium 2022 the two contributions in the chemistry section nicely demonstrate the breadth spanned by electronic structure methods in the investigation of chemical systems.

In *Two-Dimensional Hydrogen Structures at Ultra-High Pressures*, Francesco Calcavecchia, Thomas Kühne, and Markus Holzmann present a new method that combines Quantum Monte Carlo simulations with *ab initio* Molecular Dynamics. The Variational Monte Carlo (VMC) method is used to calculate the electronic Born-Oppenheimer energy surface. The variational principle is used to optimise a trial wave function given the (classical) positions of the nuclei. In a second optimisation layer, the positions of the nuclei are varied to find the zero temperature ground state structure. The method yields coupled electron-ion dynamics that is then used for Quantum Monte Carlo based *ab initio* molecular dynamics. Thus the algorithm follows the principle of Car-Parrinello molecular dynamics (CPMD) but relies on VMC for the electronic energy calculation instead of density functional theory (DFT). VMC is able to treat quantum correlation effects while being computationally efficient and very well scalable – a method perfectly suited for applications on high performance computers. The chapter then shows first tests of the method for a two-dimensional layer of hydrogen atoms. The structures that form at ultra high pressures are determined and the conductivity is investigated. Here, another advantage of the approach is demonstrated since the conductivity can be derived from the localisation properties of the ground state wave function which can be obtained from the VMC wave function. The authors were able to make a prediction for the metallisation transition of hydrogen at high pressures, one of the big open questions of high pressure physics. In summary, the newly developed algorithm opens up new avenues for simulations in a similar vein as the Car-Parrinello approach, with the potential of higher accuracy due to the replacement of DFT by the more accurate Quantum Monte Carlo approach for describing the electronic structure.

In *Molecular Origin of the Unusual Proton/Fluoride Stoichiometry of CLC-Type Fluoride Transporters*, Maria Gabriella Chiariello, Mercedes Alfonso-Prieto, and Emiliano Ippoliti use a combination of classical molecular dynamics (MD) simulations and mixed quantum/classical (QM/MM) metadynamics simulations to study the transport of fluoride ions across cellular membranes with the help membrane proteins of the so-called CLC family. As F^-/H^+ exchangers these proteins allow for an efficient transport of F^- coupled to a H^+ inward movement in a precise 1:1 stoichiometry. The authors study the mechanism of

proton release and the role that the protein conformation and two glutamate residues play which form a triad with the F^- ion that is stabilised by an intricate network of hydrogen bonds. The authors have used classical MD simulations together with well-tempered metadynamics (WT-MTD) simulations to study the protein conformations, more precisely two distinct conformational states involving a critical gating glutamate sidechain. In order to correctly describe the fluoride ion, in particular its formation of partially covalent hydrogen bonds and its capturing of protons from the protein residues, a treatment on the quantum level was required. To this end the authors very successfully used QM/MM simulations and QM/MM-based metadynamics. Combining these methods, the partially covalent binding of fluoride and the subsequent proton release as HF and the dependence of this mechanism on the presence of the two glutamate residues could be elucidated.

These two articles in the chemistry section give a very good impression of the highly efficient methods and highly parallel codes that are nowadays used at high performance computing facilities to study chemical problems – from the high pressure behaviour of atomic solids to the membrane transport mechanisms of protons and ions.

Two-Dimensional Hydrogen Structures at Ultra-High Pressures

Francesco Calcavecchia[1], Thomas D. Kühne[2,3], and Markus Holzmann[1]

[1] LPMMC, UMR 5493 of CNRS, Université Grenoble Alpes, 38042 Grenoble, France

[2] Dynamics of Condensed Matter and Center for Sustainable Systems Design, Chair of Theoretical Chemistry, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany

[3] Paderborn Center for Parallel Computing and Institute for Lightweight Design, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany
E-mail: tkuehne@cp2k.org

We introduce a novel method that combines the accuracy of Quantum Monte Carlo simulations with *ab initio* Molecular Dynamics, in the spirit of the Car-Parrinello approach. This method is then used for investigating the structure of a two-dimensional layer of hydrogen atoms, where the protons are treated as classical and the electrons as quantum particles, at $T = 0$ K and high densities. We find that metallisation is to be expected at $r_s \approx 1.1$, changing from a graphene molecular lattice structure to an atomic phase.

1 Introduction

Hydrogen is considered to be the holy grail of high pressure physics. The origin of this saying is the prediction made by Wigner and Huntington in 1935[1], suggesting the possibility that above the density corresponding to $r_s = 1.63$, hydrogen could turn into a metallic solid with an atomic bcc structure. Here, the density ρ is parametrised by $r_s = a/a_B$, where a is the mean inter-particle distance and a_B is the Bohr radius. Wigner and Huntington estimated the necessary pressure to attain metallisation to be in the order of 25 GPa. In 1967, based on the BCS theory, Ashcroft argued that such a phase would be superconductive at room temperature[2].

These predictions aroused the interest of the scientific community and challenged high-pressure physics to produce metallic hydrogen in the laboratory. After the introduction of diamond anvil cells[3], it became apparent that 25 GPa is insufficient for the metallisation of hydrogen. Nowadays experimentalists are able to compress hydrogen by applying pressures of more than 300 GPa[4,5], revealing a surprisingly rich variety of different phases[6–16].

Experimental advances in high pressure (hydrogen) physics has also triggered the development of new computational methods, often based on accurate quantum Monte Carlo calculations[17–26]. These allow quantitative comparisons and provide accurate theoretical predictions for pressures out of experimental reach. However, despite great experimental and theoretical efforts, determining the stable structures of solid hydrogen at high pressure still remains challenging. Candidate structures for phase III and IV[27,16,28,29] have layer structures. An idealised system of a strictly two-dimensional layer of hydrogen may not only provide a minimal toy model to study possible in-plane structural phases at ultra high pressures, but may also open a new pathway to metallic or superconducting hydrogen under

strong confinement. Here, we present a first quantum Monte Carlo study of 2-D hydrogen layers of classical protons at zero temperature.

For the investigation of the structure of 2-D hydrogen, we introduced a novel algorithm based on Variational Monte Carlo (VMC)[30]. VMC is able to treat quantum correlation effects while remaining computationally affordable, and can take full advantage of modern high-performance computers, since it is embarrassingly parallel. Moreover, as we will see, the use of a variational principle on a chosen Ansatz, allows us to investigate the localisation of electrons, providing additional information on the properties of hydrogen.

2 Method

Our algorithm can be schematised in two layers:

1. At the core of our simulations we employed the VMC method to calculate the electronic Born-Oppenheimer energy surface. For the sake of simplicity, only classical protons are considered, enabling us to treat the Coulomb potential generated by N protons as a static external potential. Therefore, given the protonic positions $Q \equiv (\mathbf{q}_1, \dots \mathbf{q}_N)$ and a trial wave function $\Psi_{\Theta,Q}$ with n variational parameters $\Theta \equiv (\theta_1, \dots \theta_n)$, the VMC energy reads as

$$E_{\text{VMC}}(\Theta, Q) = \frac{\langle \Psi_{\Theta,Q} \mid H \mid \Psi_{\Theta,Q} \rangle}{\langle \Psi_{\Theta,Q} \mid \Psi_{\Theta,Q} \rangle}, \tag{1}$$

2. On top of VMC, we used an optimisation algorithm for finding the Θ and Q which minimise E_{VMC} in order to find the zero temperature ground state structures at different densities.

In the following, the latter optimisation method is illustrated. For that purpose, we begin by introducing a Lagrangian for our system

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^{N} m \dot{\mathbf{q}}_i^2 + \mu \langle \dot{\Psi} \mid \dot{\Psi} \rangle - \langle \Psi \mid H - \Lambda \mid \Psi \rangle, \tag{2}$$

where m is the protonic mass, μ is a fictitious mass, and Λ is a Lagrange multiplier which ensures the normalisation condition $\langle \Psi \mid \Psi \rangle = 1$ of the trial wave function $|\Psi\rangle$.

The corresponding Euler-Lagrange equation for the protons leads to

$$\begin{aligned} m \ddot{\mathbf{q}}_i &= \mathbf{F}_i \\ \mathbf{F}_i &= -\frac{\partial}{\partial \mathbf{q}_i} \left(\langle \Psi \mid H \mid \Psi \rangle - \Lambda \langle \Psi \mid \Psi \rangle + \mu \langle \dot{\Psi} \mid \dot{\Psi} \rangle \right), \end{aligned} \tag{3}$$

where $\mathbf{F}_i$ is the force acting on the protons. Such an equation is related to Born-Oppenheimer molecular dynamics (BOMD). In fact, the latter is recovered by setting $\mu = 0$ and assuming that Ψ is an exact eigenstate, or very close to it, so that the forces can be evaluated using the Hellmann-Feynman theorem

$$\mathbf{F}_i^{BO} = -\left\langle \Psi \middle| \frac{\partial H}{\partial \mathbf{q}_i} \middle| \Psi \right\rangle. \tag{4}$$

The dynamics of the electronic wave function for fixed protons is entirely contained in the time-dependence of the variational parameters

$$\dot{\Psi}(\Theta) = \sum_{i=1}^{n} \dot{\theta}_i \mathcal{O}_i \Psi(\Theta), \tag{5}$$

where $\mathcal{O}_i \equiv \partial \log \Psi / \partial \theta_i - \langle \partial \log \Psi / \partial \theta_i \rangle$, and $\langle \cdot \rangle \equiv \langle \Psi \mid \cdot \mid \Psi \rangle$. The resulting Euler-Lagrange equation for the Θ variables then gives

$$\mu \sum_{i=1}^{n} \ddot{\theta}_i \langle \mathcal{O}_i \mathcal{O}_j \rangle = -\langle \mathcal{O}_j (H - \Lambda) \rangle, \tag{6}$$

or

$$\mu \ddot{\theta}_i = \mathcal{F}_i, \tag{7}$$

where $\mathcal{F}_i$ acts as a generalised force on the electronic parameters

$$\mathcal{F}_i = -\sum_j \langle \mathcal{O}_i \mathcal{O}_j \rangle^{-1} \langle \mathcal{O}_j (H - \Lambda) \rangle. \tag{8}$$

Note that Eq. 3 and 6 represents a coupled electron-ion dynamics that can be utilised to facilitate quantum Monte Carlo based *ab initio* molecular dynamics in the spirit of Car-Parrinello molecular dynamic (CPMD)[31,32], which keeps the electronic degrees of freedom very close to the instantaneous ground state. However, the noise in the nuclear forces computed in this way needs to be compensated by means of a modified Langevin equation[33,32,20].

Our approach therefore generalises CPMD methods using VMC for the electronic energy calculation instead of DFT. As we have already pointed out, BOMD based on VMC forces is recovered by setting $\mu = 0$, developed in Refs. 34–36. All these molecular dynamics methods in principle give also access to the real time dynamics of the (classical) protons in contrast to Coupled-Electron-Ion Monte Carlo method (CEIMC)[37,38], which samples the Boltzmann distribution of protons in the Born-Oppenheimer approximation using the Monte Carlo method. For structural relaxation as addressed later, our gradient based method should yield a higher efficiency compared to simulating annealing within CEIMC. Further, the use of the penalty method drastically reduces the acceptance rate and thus the efficiency of CEIMC for large number of electrons[39].

In this work, we have focused on this set of equations for minimising the variational energy and finding the optimal protonic structure. We have neglected the dynamic, and simply used the forces to reach the minimum, similar to the Stochastic Reconfiguration method[40] as described in[41].

The accuracy of VMC depends on the quality of the underlying trial wave function[42,43]. Here, we have considered basic Slater-Jastrow (SJ) trial wave functions[41],

$$\begin{aligned} \Psi(R) &= \det \phi_n(\mathbf{r}_i) e^{-\sum_{i,j} \mathrm{u}_{\mathrm{YUK}}(r_{ij})}, \text{ where} \\ \mathrm{u}_{\mathrm{YUK}}(r) &\equiv A \, \frac{1 - e^{-Fr}}{r}. \end{aligned} \tag{9}$$

This is to say that a Yukawa form for the electron-electron and electron-proton pair correlation accounts for symmetric correlations (A and F are variational parameters, $R = (\mathbf{r}_1, \ldots \mathbf{r}_N)$ the electronic coordinates) is employed here. For the Slater determinant, we have used four different kinds of orbitals $\phi_n(\mathbf{r})$, i.e.

1. **plane-waves:** $\phi^{\text{pw}}(\mathbf{r}, \mathbf{k}) \equiv \exp(-i\mathbf{k} \cdot \mathbf{r})$, where $\mathbf{k}$ labels the Fermi k-vectors;
2. **atomic:** $\phi^{\text{atm}}(\mathbf{r}, \mathbf{q}) \equiv \exp(-C_{\text{atm}}||\mathbf{q} - \mathbf{r}||)$, where C_{atm} is a variational parameter;
3. **bi-atomic:** $\phi^{\text{bat}}(\mathbf{r}, \mathbf{q}_1, \mathbf{q}_2) \equiv \phi^{\text{atm}}(\mathbf{r}, \mathbf{q}_1) + \phi^{\text{atm}}(\mathbf{r}, \mathbf{q}_2)$, where $\mathbf{q}_1$ and $\mathbf{q}_2$ belong to the same H_2 molecule;
4. **DFT:** orbitals resulting from a DFT calculation employing the PBE exchange-correlation energy functional and the bare Coulomb pseudo-potential.

The only variational parameters Θ involved here are A and F, from the Yukawa Jastrow, and C_{atm} for the atomic and bi-atomic orbitals. In the present study, we use very simple trial wave functions which nevertheless should capture the essence of the involved phenomena. If the plane-waves wave function has a lower variational energy than the atomic or bi-atomic one, we have a clear sign of delocalisation of electrons, characteristic for a metallic phase. This allows us to roughly localise the insulator-to-metal transition.

3 Results

We applied this novel algorithm as described above for investigating the high pressure structures of a 2D layer of hydrogen, where classical, point-like protons are strictly confined to a plane but electrons can move in all three dimensions.

The Hamiltonian of the system writes

$$H = -\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 + \sum_{i<j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,j}\frac{e^2}{|\mathbf{r}_i - \mathbf{q}_j|},$$

where m_e and e represent the electronic mass and charge, respectively. We used a modified version of the Ewald summation described in the additional material for accounting the interactions with the simulation box replicas, as required for simulating an infinite system.

We considered four starting configurations of the protons, forming either a square, triangular, or (atomic or molecular) graphene-like lattice, as illustrated in Fig. 1. Our periodic systems contained $N = 128$ hydrogen atoms for the triangular lattice and $N = 144$ for all of the other structures. Twist averaged boundary conditions were applied to reduce finite size effects[44,45], which are known to be particularly important in the metallic phase. The used code is open source and freely available[46].

For each of these structures, we first computed the variational energy, optimising only the wave function parameters. The resulting energies corresponding to densities in between $1 \leq r_s = a/a_B \leq 3.5$, where $a = (\pi\rho)^{-1/2}$ is the mean inter-particle distance in 2D, are reported in Fig. 2 and Tab. 1. Out of the investigated structures, the molecular graphene-like lattice structure with bi-atomic orbitals turned out to be the most stable one at low density/pressure whereas the triangular atomic lattice with DFT orbitals becomes favourable at high densities, i.e. $r_s \lesssim 1.1$. At this level, energies are rather widespread depending on both the considered lattice structure and on the underlying orbitals used in the Slater determinant.

Based on the generalised forces given above, we have continued to optimise the protonic structure for plane-wave, atomic, and bi-atomic orbitals, starting from the aforementioned crystal structures investigated here. The relaxation using the DFT orbitals was not

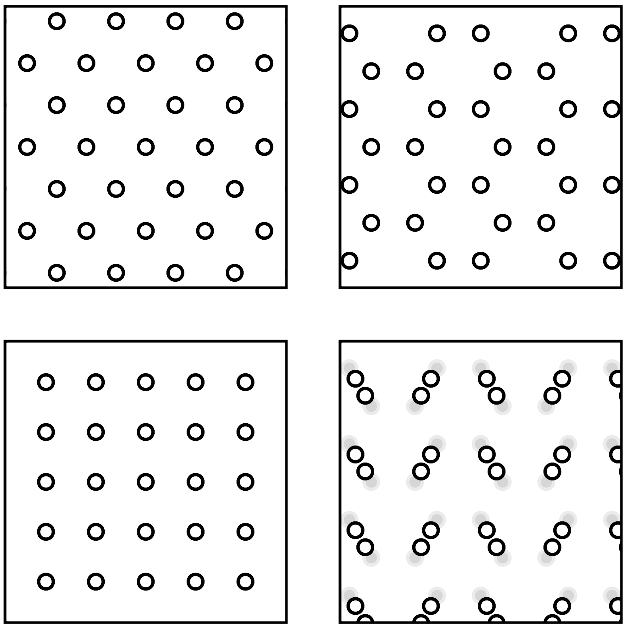


Figure 1. 2D protonic configurations considered in this work. Starting from the bottom-left panel, in clockwise order: Squared, triangular, atomic graphene-like, molecular graphene-like. In the molecular graphene-like structure, light grey markers are used to illustrate the connection to the atomic structure.

r_s		1.0	1.25	1.5	1.75	2.0	2.5	3.0	3.5
square	DFT	-0.8574(13)	-0.965(1)	-0.9910(7)	-0.9734(5)	-0.9459(5)	-0.8888(5)	-0.8521(4)	-0.8278(6)
	pw	-0.8943(3)	-0.9731(4)	-0.9709(3)	-0.9403(4)	-0.9058(2)	-0.8384(4)	-0.7886(5)	-0.7528(9)
	rel.	-0.8938(5)	-0.9725(3)	-0.9698(2)	-0.9417(4)	-0.9220(2)	-0.8754(3)	-0.8597(4)	-0.8585(6)
	atomic	-0.4894(4)	-0.7610(3)	-0.8740(2)	-0.9244(1)	-0.94873(8)	-0.96719(3)	-0.97291(2)	-0.97537(2)
	rel.	-0.7474(5)	-0.9137(3)	-0.9791(3)	-0.9905(2)	-0.9967(2)	-0.9961(2)	-0.98824(7)	-0.98723(5)
	bi-atomic	-0.5429(5)	-0.7876(4)	-0.8783(3)	-0.9092(1)	-0.9162(2)	-0.9085(3)	-0.9034(3)	-0.9093(4)
	rel.	-0.7153(6)	-0.9584(2)	-1.0203(3)	-1.0536(3)	-1.0674(3)	-1.0806(3)	-1.0776(3)	-1.0428(4)
triangle	DFT	-0.9209(12)	-0.9984(5)	-1.0103(4)	-0.9897(6)	-0.9576(4)	-0.8956(4)	-0.8565(5)	-0.8239(6)
	pw	-0.9154(3)	-0.9850(5)	-0.9605(9)	-0.8959(9)	-0.778(1)	-0.5513(9)	-0.4939(5)	-0.3997(4)
	rel.	-0.9154(3)	-0.9865(3)	-0.9744(4)	-0.9482(4)	-0.9347(3)	-0.8911(4)	-0.8677(2)	-0.8548(4)
	atomic	-0.7403(4)	-0.9062(3)	-0.9601(2)	-0.9747(2)	-0.97684(8)	-0.97423(5)	-0.97327(2)	-0.97390(1)
	rel.	-0.7799(5)	-0.9230(5)	-0.9777(3)	-0.9881(3)	-0.9957(2)	-0.98994(8)	-0.99120(7)	-0.99188(5)
	bi-atomic	-0.7091(3)	-0.8762(2)	-0.9270(2)	-0.9360(1)	-0.9300(1)	-0.9103(2)	-0.9028(3)	-0.9097(3)
	rel.	-0.7664(4)	-1.0205(3)	-1.0422(4)	-1.0793(2)	-1.0803(2)	-1.0901(3)	-1.0667(5)	-1.0631(2)
graphene-a	DFT	-0.7679(44)	-0.898(3)	-0.980(2)	-0.974(1)	-0.9566(8)	-0.9009(9)	-0.8658(5)	-0.8146(9)
	pw	-0.8784(4)	-0.9662(4)	-0.9700(4)	-0.9458(3)	-0.9148(3)	-0.8527(7)	-0.8051(7)	-0.713(1)
	rel.	-0.9082(3)	-0.9789(3)	-0.9745(3)	-0.9526(4)	-0.94390(3)	-0.8942(3)	-0.8659(5)	-0.8723(7)
	atomic	-0.8622(3)	-0.9817(2)	-1.0092(2)	-1.0077(2)	-0.9979(1)	-0.98176(7)	-0.97570(5)	-0.97473(2)
	rel.	-0.8609(4)	-0.9816(3)	-1.0090(3)	-1.0075(2)	-1.0037(1)	-0.9925(1)	-0.9933	-0.99476(9)
	bi-atomic	-0.8164(5)	-0.9575(2)	-0.9923(2)	-0.9903(1)	-0.9757(1)	-0.9416(1)	-0.9199(2)	-0.9143(3)
	rel.	-0.8344(5)	-0.9943(3)	-1.0791(2)	-1.0865(2)	-1.1137(1)	-1.1065(3)	-1.1127(2)	-1.0852(3)
graphene-m	DFT	-0.7598(30)	-0.919(3)	-1.003(1)	-1.021(1)	-1.043(1)	-1.0640(8)	-1.0834(6)	-1.0924(8)
	pw	-0.8720(4)	-0.9270(4)	-0.9055(6)	-0.8761(5)	-0.8517(6)	-0.8517(7)	-0.873(1)	-0.900(1)
	rel.	-0.9024(4)	-0.9822(3)	-0.9750(3)	-0.9532(3)	-0.9247(4)	-0.8939(4)	-0.8746(6)	-0.9072(5)
	atomic	-0.8581(3)	-0.9673(4)	-0.9980(4)	-1.019(4)	-1.0286(3)	-1.0419(6)	-1.0556(4)	-1.0724(2)
	rel.	-0.8573(3)	-0.9665(3)	-1.0033(3)	-1.0185(6)	-1.0333(2)	-1.0707(3)	-1.0547(4)	-1.0720(3)
	bi-atomic	-0.8558(3)	-1.0437(2)	-1.1051(2)	-1.1255(2)	-1.1327(1)	-1.1370(2)	-1.1395(1)	-1.1417(1)
	rel.	-0.8553(4)	-1.0433(1)	-1.1048(2)	-1.1255(1)	-1.1322(2)	-1.1369(1)	-1.1392(1)	-1.1417(1)

Table 1. Energies per atom in units of Ry for the different crystal structures: square, triangle, atomic graphene (graphene-a) and molecular graphene (graphene-m) and underlying orbitals in the Slater determinant of the VMC wave function: plane wave (pw), DFT, atomic, and bi-atomic orbitals. In each line below the orbitals types pw, atomic, and bi-atomic, we give the energies of the structural relaxation (rel.) based on the corresponding orbital. The green (orange) fields mark the lowest energies before (after) the geometry optimisation.

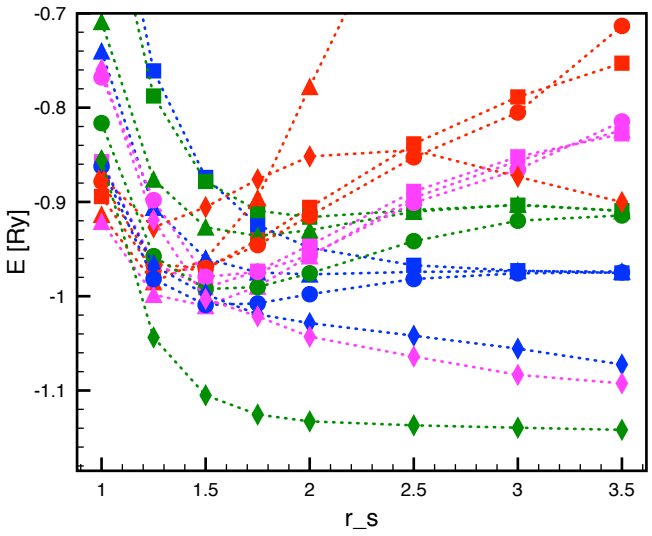


Figure 2. Variational energies obtained by optimising the variational parameters of the wave function. We used a symbol code for labelling different proton configurations (square=squared, triangle=triangular, circle=atomic graphene-like, rhombus=molecular graphene-like) and a colour code for the employed trial wave function (red=plane waves, blue=atomic, green=bi-atomic, magenta=DFT).

considered in this work. The proton relaxation process allows us to obtain new candidate structures within the constraint of fixed number of particles and fixed geometry of the simulation box. Our results after relaxing the positions of the protons are shown in Fig. 3 and Tab. 1. The protonic structure optimisation results in a lower energy for most of the configurations and trial wave functions. Energies after relaxation become less sensitive to the initial structure and group together depending mainly on the choice of the underlying Slater orbitals.

The molecular graphene-like structure is the favoured low density/pressure phase, and it is best described by the bi-atomic orbitals. Energies in this molecular phase are unaffected by the relaxation within our statistical uncertainties. The same holds for the triangular structure at high density/pressure, which is best described by plane wave orbitals, a clear sign of the metallic nature of such phase.

Besides the structure of hydrogen, its conductivity is certainly the most interesting property. Here, instead of attempting a direct calculation of the conductivity[47], we simply deduce metallic or insulating behaviour from the localisation properties of the ground state wave function. Within VMC, the localisation of the reduced single particle density matrix directly reflects the character of the orbitals inside the Slater determinant[48,38]. Since the molecular graphene structure is described by localised bi-atomic orbitals, and the triangular structure with extended DFT/pw orbitals, the metallisation transition occurs together with the structural phase transition around $r_s \approx 1.1$ within our description.

Finally, let us provide a rough estimation for the pressure necessary to reach metallisation in 2D. From the energy differences with respect to density, the two-dimensional pressure can be estimated to be around $54\ \mathrm{N/m} \simeq 1.0 \cdot 10^3 a_0\,\mathrm{GPa}$. The use of DFT orbitals

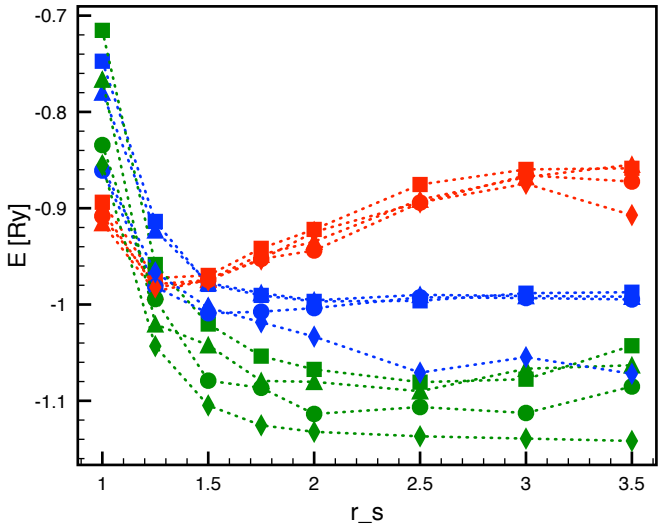


Figure 3. Variational energies obtained by optimising the variational parameters of the wave function and the proton positions. We used a symbol code for labelling different proton configurations (square=squared, triangle=triangular, circle=atomic graphene-like, rhombus=molecular graphene-like) and a colour code for the employed trial wave function (red=plane waves, blue=atomic, green=bi-atomic).

in the structural relaxation is likely to further lower the energy in the metallic phase. Since our current calculation slightly favours the molecular phase, the pressure should present an upper bound for metallisation in strictly 2D hydrogen.

4 Conclusion

In conclusion, we have introduced a novel general algorithm that can be used for simulations in the spirit of the Car-Parrinello approach, with the major difference of replacing density functional theory by more accurate quantum Monte Carlo methods for describing the electronic structure. Whereas this scheme might open up a new generation of *ab initio* simulations of higher accuracy, we have confined ourselves to the case of $T = 0$ K, where the dynamics of the classical protons simply leads to a structure relaxation. We have shown that it is possible to use this method for geometry optimisations, and we have applied it to investigate high pressure 2-D hydrogen structures where the protons are confined within a plane while the electrons are free to move in 3-D. Our simulations indicate metallisation at $r_s \approx 1.1$, which corresponds to a pressure of $1.0 \cdot 10^3$ a_0 GPa, together with a structural transition from a molecular lattice to an atomic phase. However, it is clear, that these numbers correspond to an idealised situation where zero point motion as well as any displacement of the protons in the strongly confined z direction have been neglected.

Appendix

A Ewald Summation in Quasi-2D Layers

In our study, we have considered a $2D$ layer with periodic boundary conditions on the x and y axis. However, the electrons were allowed to move in $3D$ space. Such a peculiarity requires some corrections in the Ewald summation.

In particular, when summing over all $\mathbf{k}$-vectors in the long-range term, one should consider the limit $L_z \to \infty$ and therefore $\Delta k_z \to dk_z$. The sum over all k_z needs to be substituted by an integral:

$$\sum_{k_z} \frac{e^{-\frac{k_z^2}{4\alpha}}}{k_x^2+k_y^2+k_z^2} e^{ik_z r_{ij\,z}} \to \int_{-\infty}^{\infty} dk_z \, \frac{e^{-\frac{k_z^2}{4\alpha}}}{k_x^2+k_y^2+k_z^2} e^{ik_z r_{ij\,z}} \,. \tag{10}$$

This integral can actually be computed analytically. By using the substitution

$$\frac{1}{k_x^2+k_y^2+k_z^2} = -\int_0^{\infty} dt \, e^{-\left(k_x^2+k_y^2+k_z^2\right)t} \,, \tag{11}$$

one can compute the integral in Eq. 10, obtaining

$$\begin{aligned} \frac{\pi e^{\frac{1}{4\alpha}k_x^2+k_y^2-\sqrt{k_x^2+k_y^2}|r_{ij\,z}|}}{2\sqrt{k_x^2+k_y^2}} \Bigg[2 + e^{2\frac{1}{4\alpha}\sqrt{k_x^2+k_y^2}+|r_{ij\,z}|} \mathrm{erfc}\left(\frac{2\frac{1}{4\alpha}\sqrt{k_x^2+k_y^2}+|r_{ij\,z}|}{2\sqrt{\frac{1}{4\alpha}}} \right) \\ - \mathrm{erfc}\left(\frac{-2\frac{1}{4\alpha}\sqrt{k_x^2+k_y^2}+|r_{ij\,z}|}{2\sqrt{\frac{1}{4\alpha}}} \right) \Bigg] \end{aligned} \tag{12}$$

The long-range term in the Ewald summation then reads as

$$\sum_{i,j} \sum_{k_x,k_y} -\frac{4\pi}{\mathcal{V}} e^{-\frac{k_x^2+k_y^2}{4\alpha}} e^{i\left(k_x r_{ij\,x}+k_y r_{ij\,y}\right)} \mathcal{E}(k_x,k_y,\alpha,r_{ij\,z}) \tag{13}$$

where $\mathcal{V}$ is the volume of the system, $\mathcal{E}$ is the expression in Eq. 12, and the sum over k_x, k_y does not include the case $k_x = k_y = 0$.

B Code and Algorithm Reliability Check

Whereas the VMC part was used before[41] and is therefore known to provide reliable results, the structure optimisation algorithm was new.

In order to check that both our novel algorithm and code work as expected, we have applied it to a very simple case: The H_2 molecule. The results of our test simulations are presented in Fig. 4 and demonstrated the reliability of our calculations.

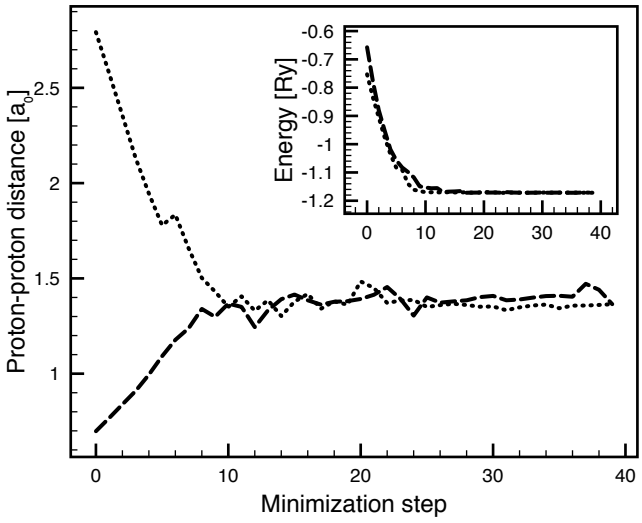


Figure 4. Structure optimisation algorithm applied to the simple H_2 molecule. We used the bi-atomic wave function, and two different starting distances for the protons. The inset shows the corresponding evolution of the variational energy.

C Finite Size Effects

Finite-size effects are known to play a crucial role in solid state physics, and in particular in conductive materials. In our study, we have accounted for them by means of TABC when using plane-wave orbitals, but not when using the atomic and bi-atomic ones. This approximation is justified as long as the simulation box is big enough for containing such localised wave functions. However, at very high densities, it is legitimate to wonder if such an approximation is valid or not.

We verified the validity of our approximation by explicitly computing the kinetic energy for the atomic orbital in a finite and infinite simulation box. In particular, we have considered the atomic graphene-like structure at $r_s = 1$, where $L_x = 27.99$ Bohr, $L_y = 16.16$ Bohr, and $C_{\text{atm}} = 0.421$. We found that our energies per electron are reliable up to $\sim 10^{-3}$ Ry $\simeq 10^{-2}$ eV. Whereas this degree of reliability could in principle be significative for making a structure favourable over another, the discrepancies we have found are such that this possibility can be disregarded. Furthermore, we have computed the energy reliability for the very worst scenario, when $r_s = 1$, and for lower densities this value is expected to be smaller.

Acknowledgements

F. C. and M. H. thank the NanoScience Foundation for support and acknowledge discussions with David Ceperley. 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), as well as computing time granted on the supercomputer Mogon at Johannes Gutenberg University Mainz (hpc.uni-mainz.de). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 716142).

References

1. E. Wigner and H. B. Huntington, The Journal of Chemical Physics **3**, 764, 1935.
2. N. W. Ashcroft, Physical Review Letters **21**, 1748, 1968.
3. A. Jayaraman, Rev. Mod. Phys. **55**, 65, 1983.
4. M. I. Eremets and I. A. Troyan, Nat. Mater. **10**, 2011.
5. R. T. Howie, C. L. Guillaume, T. Scheler, A. F. Goncharov, and E. Gregoryanz, Phys. Rev. Lett. **108**, 125501, 2012.
6. I. F. Silvera and S. Deemyad, Low Temperature Physics **35**, 318, 2009.
7. W. J. Nellis, High Pressure Research **33**, 369, 2013.
8. J. M. McMahon, M. A. Morales, C. Pierleoni, and D. M. Ceperley, Rev. Mod. Phys. **84**, 1607, 2012.
9. M. D. Knudson, M. P. Desjarlais, A. Becker, R. W. Lemke, K. R. Cochrane, M. E. Savage, D. E. Bliss, T. R. Mattsson, and R. Redmer, Science **348**, 1455, 2015.
10. M. Zaghoo, A. Salamat, and I. F. Silvera, Phys. Rev. B **93**, 155128, 2016.
11. P. Dalladay-Simpson, R. Howie, and E. Gregoryanz, Nature **529**, 63, 2016.
12. R. P. Dias and I. F. Silvera, Science **355**, 715, 2017.
13. A. F. Goncharov and V. V. Struzhkin, Science **357**, 2017.
14. X.-D. Liu, P. Dalladay-Simpson, R. T. Howie, B. Li, and E. Gregoryanz, Science **357**, 2017.
15. M. I. Eremets, I. A. Troyan, and A. P. Drozdov, ArXiv, 2016, arXiv:1601.04479.
16. B. Monserrat, R. J. Needs, E. Gregoryanz, and C. J. Pickard, Phys. Rev. B **94**, 134101, 2016.
17. D. M. Ceperley and B. J. Alder, Science **231**, 555, 1986.
18. D. M. Ceperley and B. J. Alder, Phys. Rev. B **36**, 2092, 1987.
19. C. Pierleoni, D. M. Ceperley, and M. Holzmann, Phys. Rev. Lett. **93**, 146402, 2004a.
20. C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501, 2008a.
21. C. J. Pickard and R. J. Needs, physica status solidi (b) **246**, 536, 2009.
22. S. Azadi, W. M. C. Foulkes, and T. D. Kühne, New J. Phys. **15**, 113005, 2013.
23. N. D. Drummond, B. Monserrat, J. H. Lloyd-Williams, P. L. Ríos, C. J. Pickard, and R. J. Needs, Nature Communications **6**, 7794, 2015.
24. F. Calcavecchia and T. D. Kühne, Zeitschrift für Naturforschung A **73**, 845, 2018.
25. S. Azadi, G. H. Booth, and T. D. Kühne, J. Chem. Phys. **153**, 204107, 2020.
26. J. Kessler, F. Calcavecchia, and T. D. Kühne, Adv. Theory Simul. **4**, 2000269, 2021.
27. C. J. Pickard and R. J. Needs, Nature Physics **3**, 473, 2007.
28. C. J. Pickard, M. Martinez-Canales, and R. J. Needs, Phys. Rev. B **85**, 214114, 2012a.
29. C. J. Pickard, M. Martinez-Canales, and R. J. Needs, Phys. Rev. B **86**, 059902, 2012b.
30. W. L. McMillan, Phys. Rev. **138**, A442, 1965.
31. R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471, 1985.

32. T. D. Kühne, M. Krack, F. R. Mohamed, and M. Parrinello, Phys. Rev. Lett. **98**, 066401, 2007.
33. F. R. Krajewski and M. Parrinello, Phys. Rev. B **73**, 041105, 2006.
34. J. C. Grossman and L. Mitas, Phys. Rev. Lett. **94**, 056403, 2005.
35. C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501, 2008b.
36. G. Mazzola, S. Yunoki, and S. Sorella, Nature Communications, 2014.
37. C. Pierleoni, D. M. Ceperley, and M. Holzmann, Phys. Rev. Lett. **93**, 146402, 2004b.
38. C. Pierleoni, M. A. Morales, G. Rillo, M. Holzmann, and D. M. Ceperley, Proc. Nat. Acad. Sci. USA **113**, 4953, 2016.
39. D. M. Ceperley and M. Dewing, J. Chem. Phys. **110**, 9812, 1999.
40. S. Sorella, Phys. Rev. B **71**, 241103, 2005.
41. F. Calcavecchia and T. D. Kühne, EPL (Europhysics Letters) **110**, 20011, 2015.
42. C. Pierleoni, K. T. Delaney, M. A. Morales, D. M. Ceperley, and M. Holzmann, Computer Physics Communications **179**, 89, 2008.
43. M. Taddei, M. Ruggeri, S. Moroni, and M. Holzmann, Phys. Rev. B **91**, 115106, 2015.
44. C. Lin, F. H. Zong, and D. M. Ceperley, Phys. Rev. E **64**, 016702, 2001.
45. M. Holzmann, R. C. Clay, M. A. Morales, N. M. Tubman, D. M. Ceperley, and C. Pierleoni, Phys. Rev. B **94**, 035126, 2016.
46. F. Calcavecchia, `http://github.com/francesco086/HswfQMC`
47. F. Lin, M. A. Morales, K. T. Delaney, C. Pierleoni, R. M. Martin, and D. M. Ceperley, Phys. Rev. Lett. **103**, 256401, 2009.
48. G. Mazzola and S. Sorella, Phys. Rev. Lett. **114**, 105701, 2015.

Molecular Origin of the Unusual Proton/Fluoride Stoichiometry of CLC-Type Fluoride Transporters

Maria Gabriella Chiariello[1,2,4], **Mercedes Alfonso-Prieto**[1,3], **and Emiliano Ippoliti**[1,2]

[1] Institute for Advanced Simulation (IAS-5) and Institute of Neuroscience and Medicine (INM-9), Computational Biomedicine, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {m.alfonso-prieto, e.ippoliti}@fz-juelich.de

[2] JARA-HPC, Forschungszentrum Jülich, 54245 Jülich, Germany

[3] Cécile and Oskar Vogt Institute for Brain Research, University Hospital Düsseldorf, Medical Faculty, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

[4] University of Groningen, Groningen Biomolecular Sciences and Biotechnology Institute, Nijenborgh 4, 9747 AG Groningen, The Netherlands
E-mail: m.g.chiariello@rug.nl

Fluoride (F^-) is present in natural water sources at concentrations that are toxic to unicellular organisms. Thus, prokaryotes have evolved export transporters for this anion to ensure their survival. One such group of fluoride transporters are the F^-/H^+ exchangers belonging to the CLC family (CLC^Fs). The X-ray structure of the *Enterococcus casseliflavus* CLC^F (CLC^F-eca) contains a glutamate (E318) in the central anion-binding site, which is absent in the conventional CLCs and has been suggested as one of the determinants enabling fluoride transport. Here we use classical and subnanosecond QM/MM metadynamics simulations to investigate the molecular basis of the 1:1 proton:fluoride stoichiometry in CLC^F-eca. The proton release mechanism relies on the formation of the E318-F-E118 triad, stabilised by a complex H-bond network in which both glutamates interact with F^-. The interplay between the two glutamates enables protonation of fluoride and proton release as HF into the intracellular solution. Our results demonstrate how glutamate insertion into the central anion-binding site of CLC^F-eca permits F^- release coupled to H^+ inward movement, resulting in effective and selective F^- transport.

1 Introduction

Fluoride ions exert toxic effects at concentrations close to environmental levels by inhibiting enzymes necessary for energy production and nucleic acid synthesis[1–6]. Unicellular organisms accumulate F^- intracellularly via HF permeability of cell membranes, especially in acidic environments. Hence, bacteria, protozoa or yeast utilise specialised F^- transporters to reduce intracellular fluoride below toxic concentrations. One group of F^- transporters, the so-called CLC^F[7,8], belongs to the large CLC family of anion channels and transporters[9–11]. CLC^F transporters exploit the inward directed proton gradient in acidic environment to clear the bacterial cytoplasm from F^-. CLC^F transporters are closely related to the prototypic CLC-type Cl^-/H^+ exchanger from *Escherichia coli*, CLC-ec1[12]. However, CLC-ec1 and CLC^F differ in halide selectivity and proton-coupled stoichiometry: CLC-ec1 is a chloride transporter that is strongly inhibited by F^-[13], whereas CLC^F effectively exchanges H^+ and F^- with a 1:1 stoichiometry[7,12,14]. A comparison of the mechanisms underlying halide selectivity and proton-coupling in CLC^F and CLC-ec1 promises insight into how nature exploits the unique chemical features of fluoride to con-

vert a blocker into a selectively transported substrate. CLC^F transporters share basic transport mechanisms with Cl^-/H^+ exchangers, such as CLC-ec1, which have been extensively studied in recent years[16–23]. In CLC-ec1, Cl^-/H^+ exchange critically depends on two conserved glutamate residues: the gating glutamate (E148)[24] and the internal glutamate (E203). During transport, E148 changes conformation from up to down[25–28], shuttling protons from one side of the membrane to the other, whereas E203 is necessary for proton release into the internal medium. Proton movement drives Cl^- transport in the opposite direction, resulting in a 2 Cl^-:1 H^+ stoichiometry[29]. However, in the presence of fluoride, a strong hydrogen bond (H-bond) is formed between F^- and the gating E148, thereby locking CLC-ec1 in a conformation that does not allow for proton transport[13]. In contrast, CLC^F proteins allow the exchange of fluoride and proton with 1:1 stoichiometry[7,12,15]. CLC^Fs, such as the one of *Enterococcus casseliflavus* (CLC^F-eca), maintain the aforementioned gating glutamate (E118 in CLC^F-eca), but lack the intracellular transport glutamate present in CLC-ec1[12]. Moreover, CLC^Fs exhibit a so-called coordinating glutamate close to the central anion-binding site (E318 in CLC^F-eca, see Fig. 1), absent in CLC-ec1.

Here, we use molecular dynamics (MD) simulations to study how CLC^F transporters are able to exchange fluorides and protons with a net 1:1 stoichiometry. Our simulations show that the first step of the CLC^F-eca transport cycle consists in the up to down movement of the protonated gating glutamate, followed by the formation of a strong H-bonded network in which both the gating and coordinating glutamates interact with F^-. From this E118-F-E318 triad, protons and fluoride anions are co-released as HF into the intracellular solution.

2 Methods

Classical MD simulations have been widely applied to study ion channels and transporters[30]. However, the force field parametrisation of anions, such as fluoride, poses some challenges, especially to describe their polarisability, as recently shown in Ref. 31. Moreover, the partially covalent H-bonds formed by F^-[48,50,49], as well as the ability of this weak base to capture protons from protein residues with similar pKa values (such as glutamate)[32,33], can only be described at the quantum level. Therefore, in this work we used a multiscale approach, combining both classical and quantum mechanics/molecular mechanics (QM/MM) simulations. Furthermore, well-tempered metadynamics (WT-MTD)[34] was used as enhanced sampling and free energy technique.

The crystal structure of the wild-type CLC-type F^-/H^+ antiporter from *Enterococcus casseliflavus* (CLC^F-eca, PDB ID: 6D0J, resolution 3.00 Å) was used to model the initial state of the transport cycle. After an initial equilibration, a 300 ns-long classical MD simulation was performed using the CHARMM-CMAP[37], CHARMM36[38], and TIP3P[39] force fields for protein, lipids and water, respectively, whereas the fluoride parameters were taken from reference[40]. The GROMACS-2018.3 package[41] was used for the classical MD simulations and the classical WT-MTD simulations were performed by introducing in GROMACS the corresponding routines of the PLUMED 2.5.3 library[43]. QM/MM simulations were carried out by using a recently developed framework for multiscale modelling (MiMiC)[35,36]. MiMiC couples two popular and established codes for *ab initio* and classical MD simulations: CPMD 4.3[42] and GROMACS 2018[41], respectively. The QM

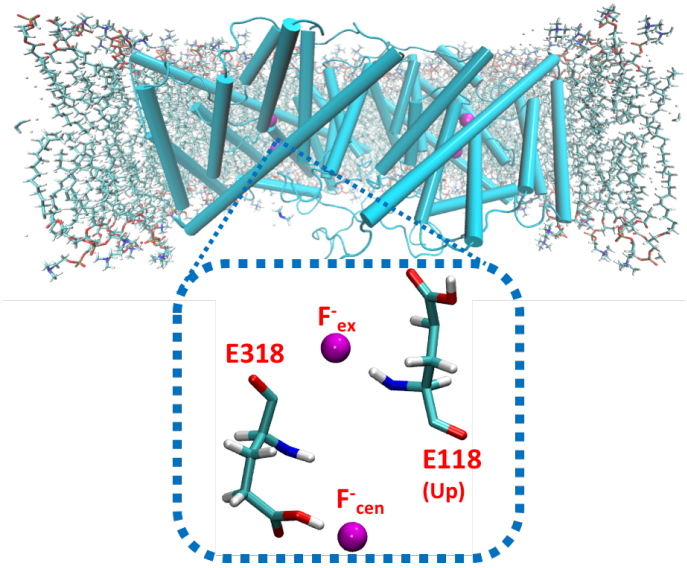


Figure 1. Top panel: *E. casseliflavus* F^-/H^+ antiporter (CLC^F-eca) embedded in a lipid bilayer in our simulation setup. Water and counterions are not shown for clarity. Bottom panel: Anion-binding sites in the wild-type protein. Two fluoride anions bound to the external and central sites (F^-_{ex} and F^-_{cen}, respectively) are included in our simulations, as observed in the corresponding crystal structure (PDB code 6D0J) . F^-_{cen} is coordinated by E318 and the gating glutamate E118 is in the up conformation. Reprinted (adapted) with permission from J. Phys. Chem. Lett. 2021, 12, 18, 4415–4420. Copyright 2021 American Chemical Society.

part was described by density functional theory, using the BLYP[44,45] functional. Core electrons were described using norm conserving pseudopotentials of the Martins-Troullier type[46], while valence electrons were treated explicitly using a plane wave basis set cutoff of 90 Ry. Monovalent carbon pseudopotentials were used to saturate the dangling bonds at the boundaries of the QM and MM regions. The fully Hamiltonian electrostatic coupling scheme developed by Laio *et al.*[47] and implemented in MiMiC was used to describe electrostatic interactions between the QM and MM subsystems. For the QM/MM-based WT-MTD simulations, MiMiC was interfaced with a suitably adapted version of the PLUMED 2.5.3 library.

We would like to note that the implementations of all the codes used in this work (GROMACS, CPMD and MiMiC) are highly parallel and thus very suited to study large membrane proteins systems such as the one investigated here, by taking full advantage of the current state-of-the-art high performance computing platforms (such as those available at the Jülich Supercomputing Centre).

3 Results

Here, we present the results on the H^+ transport from the extracellular to the intracellular solution, the first step of the CLC^F-eca transport cycle, as proposed in Ref. 12. The equilibrated structure of the wild-type CLC^F-eca is used as starting conformation, in which the

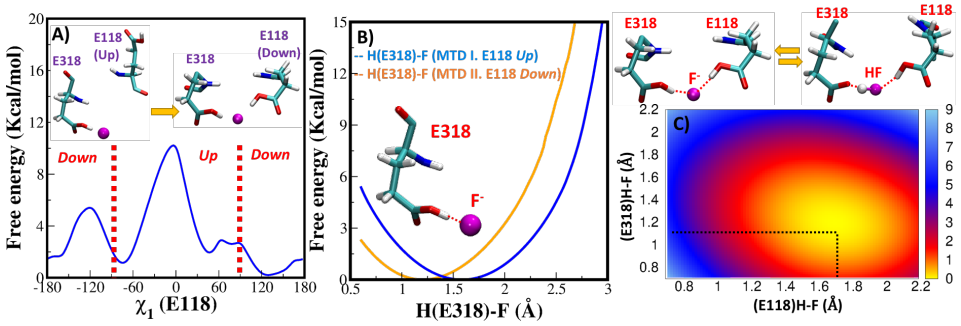


Figure 2. A) Classical MTD simulation of conformational change in the gating glutamate. Up/down conformations are defined by the position of the E118 side chain above/below the plane of its backbone. Red dashed lines indicate the transition between the two conformations. B) QM/MM MTD of proton transfer between fluoride and E318 with the gating glutamate E118 in either the up (MTD.I, blue) or down (MTD.II, orange) conformation. C) QM/MM MTD.II of double proton transfer in the E118-F-E318 triad. The two possible H-bond network configurations are shown on the right, with H-bonds indicated as red dashed lines. All the free energy values are indicated in kcal/mol. Reprinted (adapted) with permission from J. Phys. Chem. Lett. 2021, 12, 18, 4415–4420. Copyright 2021 American Chemical Society.

gating glutamate E118 in the up conformation is exposed to the extracellular solution and thus free to exchange protons. The coordinating glutamate E318 is instead located in the central anion-binding site, in close contact with fluoride (F^-_{cen} in Fig. 1). This suggests that E318 must be H-bonded to F^- and thus protonated in this conformation. Thus, in our model both E118 and E318 are considered as protonated.

After 300 ns of classical MD simulation, we first explored the conformational movement of the gating glutamate via classical WT-MTD using the E118 N-Cα-Cβ-Cγ (χ_1) dihedral angle as collective variable. The free energy profile as a function of χ_1 (Fig. 2A) shows that, once E118 captures a proton from the extracellular solution, its side chain undergoes a spontaneous up to down transition. In the down conformation (Fig. 2), both glutamate protons interact with F^-_{cen} to form an E318-F-E118 triad.

To improve the description of the H-bonds formed by F^- and take into account the possibility of proton transfer events, we next computed free energies profiles through QM/MM WT-MTD simulations. The collective variables used were the H(E318)-F^- distance alone (E118 in the up conformation) or in combination with the H(E118)-F^- distance (E118 in the down conformation). When the E118 is in up conformation, F^- interacts only with the coordinating glutamate E318, forming a strong H-bond, as shown by the single minimum at a H(E318)-F^- distance of 1.5 Å (Fig. 2).

In contrast, when E118 adopts the down position, fluoride is coordinated by both glutamates, forming an E118-F-E318 triad. Fig. 2C shows a single free energy minimum corresponding to an HF molecule placed between protonated E118 and deprotonated E318 (i.e. H(E118)-F = 1.7 Å and H(E318)-F = 1.1 Å). Projection onto the H(E318)-F distance of the two QM/MM free energy profiles for either of the E118 positions (Fig. 2B) makes evident that the up to down movement of E118 is accompanied by a shortening of H(E318)-F distance (from 1.5 Å to 1.1 Å). This indicates a change from F^- that is H-bonded to protonated E318 (up position) to HF that is H-bonded to deprotonated E318 (down position).

To complete the first step of the transport cycle of CLC^F proposed in reference[12], we next calculated the QM/MM WT-MTD free energy surface (Fig. 3A) as a function of the H(E118)-F and H(E118)-O(E318) distances. These two collective variables describe the release of HF from the gating glutamate and the approach of the two glutamates, respectively. The simulation starts with the E118-F-E318 triad (minimum (i) in Fig. 3A), corresponding to protonated E118 H-bonded to HF (H(E118)-F = 1.5 Å). Then, HF dissociates from the gating glutamate (H(E118)-F distance = 3.5 Å), resulting in minimum (ii). Subsequently, a direct H-bond is formed between the protonated E118 and the deprotonated E318 (H(E118)-O(E318) = 1.5 Å), while HF remains coordinated to E318, as shown in minimum (iii). The three free energy minima in Fig. 3A differ by only 2 kcal/mol or less and are separated by low barriers (∼5 kcal/mol), indicating that their interconversion is possible at 300 K. Moreover, the E118-E318 dyad formed in minimum (iii) exhibits a H-bond distance between the two glutamates similar to that observed in the V319G CLC^F-eca X-ray structure[12].

Finally, we investigated the detachment of HF from E318 through QM/MM WT-MTD, using as CV the distance between the HF hydrogen atom and one of the carboxyl oxygen atoms of E318 (O1). The corresponding free energy profile is shown in Fig. 3B. The first two minima (O(E318)-HF = 1.5 and 2.6 Å, respectively) are isoenergetic and correspond to HF H-bonded to either O1 or O2 of the E318 carboxyl group. Such exchange of H-bond partner has a barrier lower than 2 kcal/mol, supporting its feasibility at 300 K. The third minimum corresponds to HF solvated by water molecules in the intracellular side of the channel, while the H-bonded dyad between E318 and E118 in the down conformation is maintained. The HF release from the coordinating glutamate described here requires overcoming a barrier of ∼4 kcal/mol. However, the free energy surface shown in Fig. 3B was obtained while using a harmonic restraint on the HF distance, which precludes HF dissociation into H^+ and F^-. Thus, the computed barrier should be considered a lower bound limit of the energy required for HF release from the glutamate dyad into the intracellular solution. Nonetheless, HF in the intracellular vestibule of CLC^F-eca is solvated exclusively by water molecules. Such an environment, lacking strong H-bond donors, shall disfavour HF dissociation with long-range proton diffusion[48,49], as modelled with the HF distance restraint used in our simulation.

4 Discussion and Concluding Remarks

F^- transport by CLC^F-eca is an outstanding example of proteins evolutionary adapt to the chemical features of substrates to become selective transporters. Our calculations suggest that protonation of the gating E118 in CLC^F-eca leads to a barrierless transition from the up to down conformation and the subsequent formation of a E318-F-E118 triad stabilised by a complex H-bond network (Fig. 4). This triad in turn favours protonation of F^-_{cen}, resulting in an HF molecule H-bonded to both protonated E118 and deprotonated E318. Next, E118 approaches E318, expelling HF from the central binding site into the intracellular solution, and leaving the two glutamates in contact to form a H-bonded dyad. In contrast, CLC-ec1, which lacks the coordinating glutamate, can form only a single glutamate/fluoride interaction in the central binding site, which causes fluoride blockage via proton sharing between the gating glutamate and fluoride[13]. Instead, the two glutamates

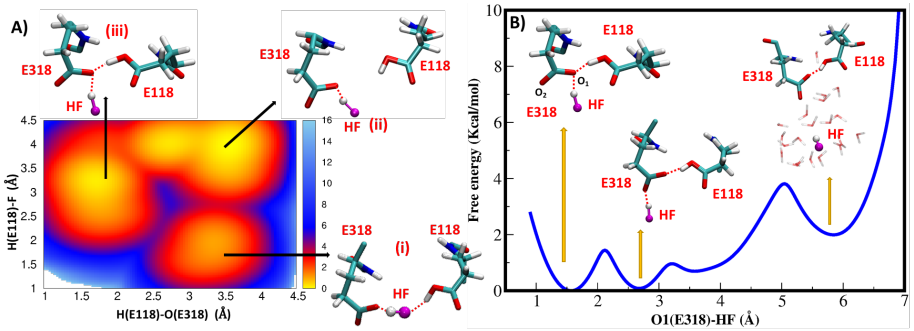


Figure 3. A) Conversion of the E318-F-E118 triad into the intermediate conformation with E318 and E118 in direct contact (MTD.III). B) HF release into the intracellular solution (MTD.IV). O1 and O2 indicate the two carboxyl oxygen atoms of the coordinating E318. H-bonds are shown as red dashed lines and free energy profiles are in kcal/mol. Reprinted (adapted) with permission from J. Phys. Chem. Lett. 2021, 12, 18, 4415–4420. Copyright 2021 American Chemical Society.

present in CLC^F-eca prevent fluoride blockage by forming the E118-F-E318 triad, which permits release of HF from the central binding site.

Based on the crystal structures of CLC^F-eca, Last and coworkers proposed the simultaneous release of protons and fluoride ions into the intracellular solution to explain the reduced stoichiometry of H^+/F^- exchangers compared with H^+/Cl^- antiporters[7,12]. Our calculations complement the static picture provided by the experimental structures (A and F in Fig. 4) by providing the detailed molecular mechanism of this co-release (B–E in Fig. 4): downward rotation of the protonated E118, together with the CLC^F-eca-specific E318 in the central binding site, enables fluoride protonation and subsequent co-release of proton and fluoride as HF into the intracellular solution.

At physiological voltages, CLC^{F}s select against Cl^-, the main physiological halide, through weaker binding of this anion[15]. Our simulations show that F^- forms either partially or fully covalent bonds with the protons in the E118-F-E318 system (Fig. 4). In contrast, Cl^- can only form (weaker) ionic H-bonds with the two glutamates[50]. Moreover, F^- can accept a proton from E318 because their pKa values are close (3.2 and 4.2, respectively). Instead, Cl^- will remain unprotonated, because it is a much weaker base (pKa $\sim$ -7). Combined mutagenesis and functional studies have shown that neutralising the gating glutamate (E118Q and E118A) suppresses proton and fluoride transport, resulting in Cl^- fluxes that are inhibited by F^-[17,51]. In these E118 mutants, anion transport is not linked to the counter transport of H^+ as in WT CLC^F-eca. Our calculations suggest that the F^- inhibition observed in these mutants, which possess only one glutamate (E318), will be due to the presence of a single fluoride-glutamate interaction, as in the case of CLC-ec1[13,50].

In contrast, mutations of the coordinating E318 have been shown experimentally to still permit coupled H^+/F^- transport. However, the E318A and E318Q mutations changed the transport stoichiometry from 1:1 to 2:1[12]. Based on the simulations presented here, we surmise that HF formation and release is probably hampered in the E318 mutants, because of the removal of the coordinating glutamate in the central binding site. Nonetheless, these mutants maintain the gating E118, which may still release protons, thus permitting F^--H^+ exchange. In fact, the ability of E118 to release H^+ is supported by the proton-coupled

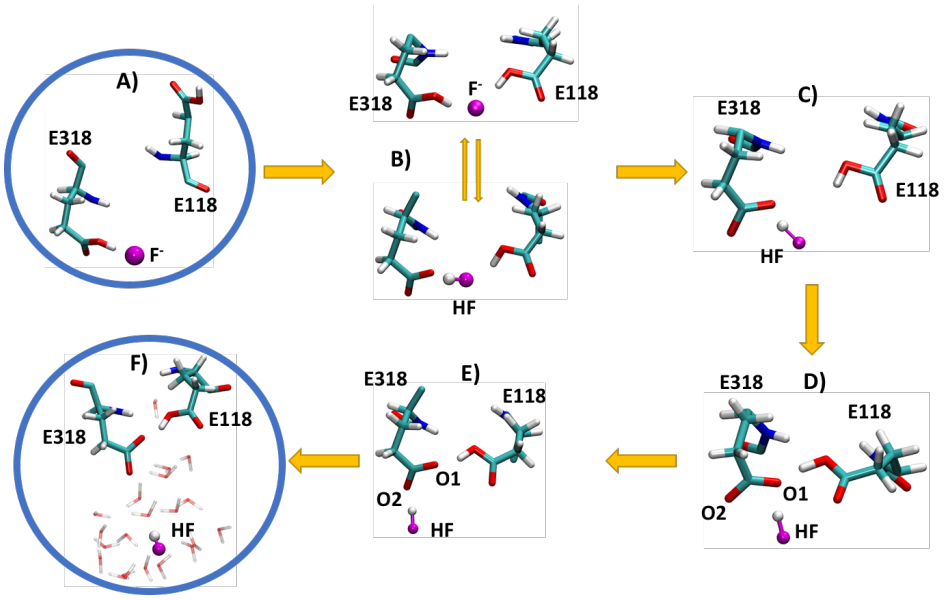


Figure 4. Structures involved in the first step of the transport cycle of CLC^F-eca. A) Conformation corresponding to the wild-type CLC^F-eca crystal structure; E118 is in the up conformation and exposed to the extracellular solution, where it can easily capture a proton. B) Upon protonation, the gating E118 undergoes a spontaneous up to down transition, leading to the formation of the E318-F-E118 triad. F^- accepts a proton from E318, resulting in a strong H-bond network involving deprotonated E318, HF and protonated E118. C) HF dissociates from the gating E118. D) E118 approaches E318 and forms a H-bond glutamate dyad. HF is still H-bonded to one of the E318 carboxyl oxygen atoms. E) HF swings between the two E318 carboxyl oxygen atoms. F) HF dissociates from the glutamate dyad into the intracellular part of the CLC^F-eca vestibule, where it is surrounded by water molecules. E118 in the down conformation has replaced fluoride in the central binding site, as proposed based on the V319G CLC^F-eca X-ray structure. Reprinted (adapted) with permission from J. Phys. Chem. Lett. 2021, 12, 18, 4415–4420. Copyright 2021 American Chemical Society.

Cl^- transport observed in WT CLC^F-eca, especially at high negative voltages[15].

In summary, we have investigated the molecular basis of the first step of the CLC^F-eca transport cycle, providing insights into the partially covalent binding of fluoride at the central anion site and proton release as HF to the internal medium. Based on the simulations presented here, we can make the following prediction to be tested in future experimental studies: since two glutamate residues appear to be crucial for fluoride transport at 1:1 F^-/H^+ stoichiometry, then inserting a second glutamate into the anion coordination sphere of a CLC-type Cl^-/H^+ exchanger should be the first step to convert the chloride transporter into a fluoride transporter.

Acknowledgements

M. A. P. is funded by Deutsche Forschungsgemeinschaft via the Research Unit FOR2518 “Functional Dynamics of Ion Channels and Transporters – DynIon”, project P6. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V.

(www.gauss-centre.eu) for providing computing time on the GCS Supercomputer JUWELS[52] at the Jülich Supercomputing Centre.

References

1. E. A. Orabi, J. D. Faraldo-Gómez, *New Molecular-Mechanics Model for Simulations of Hydrogen Fluoride in Chemistry and Biology*, J. Chem. Theory Comput. **16**, 5105–5126, 2020.
2. F. A. Cotton, G. Wilkinson, C. A. Murillo, M. Bochmann, R. Grimes, *Advanced Inorganic Chemistry*, Wiley New York, 1988.
3. H. Zuo, L. Chen, M. Kong, L. Qiu, P. Lü, P. Wu, Y. Yang, K. Chen, *Toxic Effects of Fluoride on Organisms*, Life Sci. **198**, 18–24, 2018.
4. E. Adamek, K. Pawlowska-Goral, K. Bober, *In Vitro and in Vivo Effects of Fluoride Ions on Enzyme Activity*, Ann. Acad. Med. Stetin. **51**, 69–85, 2005.
5. J. Qin, G. Chai, J. M. Brewer, L. L. Lovelace, L. Lebioda, *Fluoride Inhibition of Enolase: Crystal Structure and Thermodynamics*, Biochemistry **43**, 793–800, 2006.
6. M. Goličnik, L. F. Olguin, G. Feng, N. J. Baxter, J. P. Waltho, N. H. Williams, F. Hollfelder, *Kinetic Analysis of β-Phosphoglucomutase and Its Inhibition by Magnesium Fluoride*, J. Am. Chem. Soc. **131**, 1575–1588, 2009.
7. R. B. Stockbridge, H. Lim, R. Otten, C. Williams, T. Shane, Z. Weinberg, C. Miller, *Fluoride Resistance and Transport by Riboswitch-Controlled CLC Antiporters*, Proc. Natl. Acad. Sci **109**, 15289–15294, 2012.
8. B. C. McIlwain, M. T. Ruprecht, R. B. Stockbridge, *Membrane Exporters of Fluoride Ion*, Ann. Rev. Biochem. **90**, 559-579, 2021.
9. T. J. Jentsch, V. Stein, F. Weinreich, A. A. Zdebik, *Molecular Structure and Physiological Function of Chloride Channels*, Physiol. Rev. **82**, 503–568, 2002.
10. G. Stölting, M. Fischer, C. Fahlke, *CLC Channel Function and Dysfunction in Health and Disease*, Front. Physiol. **5**, 378, 2014.
11. D. R. Poroca, R. M. Pelis, V. M. Chappe, *ClC Channels and Transporters: Structure, Physiological Functions, and Implications in Human Chloride Channelopathies*, Front. Pharmacol. **8**, 151, 2017.
12. N. B. Last, R. B. Stockbridge, A. E. Wilson, T. Shane, L. Kolmakova-Partensky, A. Koide, S. Koide, C. Miller, *A CLC-type F^-/H^+ Antiporter in Ion-Swapped Conformations*, Nat. Struct. Mol. Biol. **25**, 601–606, 2018.
13. M. G. Chiariello, V. Bolnykh, E. Ippoliti, S. Meloni, J. M. H. Olsen, T. Beck, U. Rothlisberger, C. Fahlke, P. Carloni, *Molecular Basis of CLC Antiporter Inhibition by Fluoride*, J. Am. Chem. Soc. **142**, 7254–7258, 2020.
14. M. G. Chiariello, M. Alfonso-Prieto, E. Ippoliti, C. Fahlke, P. Carloni, *Mechanisms Underlying Proton Release in CLC-type F^-/H^+ Antiporters*, J. Phys. Chem. Lett. **12**, 4415–4420, 2021.
15. A. E. Brammer, R. B. Stockbridge, C. Miller, *F^-/Cl^- Selectivity in CLC^F-type F^-/H^+ Antiporters*, J. Gen. Physiol **144**, 129–136, 2014.
16. R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait, R. MacKinnon, *X-ray Structure of a ClC Chloride Channel at 3.0 Å Reveals the Molecular Basis of Anion Selectivity*, Nature **415**, 287–294, 2002.

17. A. Accardi, C. Miller, *Secondary active transport mediated by a prokaryotic homologue of ClC Cl^- channels*, Nature **427**, 803–807, 2004.
18. A. Accardi, M. Walden, W. Nguitragool, H. Jayaram, C. Williams, C. Miller, *Separate Ion Pathways in a Cl^-/H^+ Exchanger*, J. Gen. Physiol. **126**, 563–570, 2005.
19. R. Dutzler, *The ClC family of chloride channels and transporters*, Curr. Op. Struct. Biol. **16**, 439–446, 2006.
20. M. Walden, A. Accardi, F. Wu, C. Xu, C. Williams, C. Miller, *Uncoupling and Turnover in a Cl^-/H^+ Exchange Transporter*, J. Gen. Physiol **129**, 317–329, 2007.
21. D. Basilio, K. Noack, A. Picollo, A. Accardi, *Conformational Changes Required for H^+/Cl^- Exchange Mediated by a CLC Transporter*, Nat. Struct. Mol. Biol **21**, 456–463, 2014.
22. W. Han, R. C. Cheng, M. C. Maduke, E. Tajkhorshid, *Water Access Points and Hydration Pathways in CLC H^+/Cl^- Transporters*, Proc. Natl. Acad. Sci. **111**, 1819–1824, 2014.
23. S. Lobet, R. Dutzler, *Ion-Binding Properties of the ClC Chloride Selectivity Filter*, EMBO J. **25**, 24–33, 2006.
24. R. Dutzler, E. B. Campbell, R. MacKinnon, *Gating the Selectivity Filter in ClC Chloride Channels*, Science **300**, 108–112, 2003.
25. H. B. Mayes, S. Lee, A. D. White, G. A. Voth, J. M. J. Swanson, *Multiscale Kinetic Modeling Reveals an Ensemble of Cl^-/H^+ Exchange Pathways in ClC-ec1 Antiporter*, J. Am. Chem. Soc. **140**, 1793–1804, 2018.
26. S. Lee, H. B. Mayes, J. M. J. Swanson, G. A. Voth, *Origin of Coupled Chloride and Proton Transport in a Cl-/H+ Antiporter*, J. Am. Chem. Soc **138**, 14923–14930, 2016.
27. S. Lee, J. M. J. Swanson, G. A. Voth, *Multiscale Simulations Reveal Key Aspects of the Proton Transport Mechanism in the ClC-ec1 Antiporter*, Biophys. J. **110**, 1334–1345, 2016.
28. M. Vien, D. Basilio, L. Leisle, A. Accardi, *Probing the Conformation of a Conserved Glutamic Acid Within the Cl^- Pathway of a CLC H^+/Cl^- Exchanger*, J. Gen. Physiol. **149**, 523–529, 2017.
29. A. Picollo, Y. Xu, N. Johner, S. Bernèche, A. Accardi, *Synergistic Substrate Binding Determines the Stoichiometry of Transport of a Prokaryotic H^+/Cl^- Exchanger*, Nat. Struct. Mol. Biol. **19**, 525–531, 2012.
30. V. Carnevale, L. Delemotte, R. J. Howard, *Molecular Dynamics Simulations of Ion Channels*, Trends Biochem. Sci. **46**, 621–622, 2021.
31. Z. Yue, Z. Wang, G. A. Voth, *Ion permeation, selectivity, and electronic polarization in fluoride channels*, Biophys. J. **121**, 1336–1347, 2022.
32. D. L. Bostick, C. L. Brooks III, *Statistical determinants of selective ionic complexation: ions in solvent, transport proteins, and other "hosts"*, Biophys. J. **96**, 4470–4492, 2009.
33. M. Cametti, K. Rissanen, *Highlights on contemporary recognition and sensing of fluoride anion in solution and in the solid state*, Chem. Soc. Rev. **42**, 2016–2038, 2013.
34. A. Barducci, M. Bonomi, M. Parrinello, *Metadynamics*, Wiley Interdiscip. Rev. Comput. Mol. Sci. **5**, 826–843, 2011.

35. J. M. H. Olsen, V. Bolnykh, S. Meloni, E. Ippoliti, M. P. Bircher, P. Carloni, U. Rothlisberger, *MiMiC: A Novel Framework for Multiscale Modeling in Computational Chemistry*, J. Chem. Theory Comput. **15**, 3810–3823, 2019.
36. V. Bolnykh, J. M. H. Olsen, S. Meloni, M. P. Bircher, E. Ippoliti, P. Carloni, U. Rothlisberger, *Extreme scalability of DFT-based QM/MM MD simulations using MiMiC*, J. Chem. Theory Comput. **15**, 5601-5613, 2019.
37. A. D. Mackerell Jr, M. Feig, C. L. Brooks III, *Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations*, J. Comput. Chem. **11**, 1400–1415, 2004.
38. J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, R. W. Pastor, *Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types*, J. Phys. Chem. B **114**, 7830–7843, 2010.
39. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *Comparison of simple potential functions for simulating liquid water*, J. Phys. Chem. B **79**, 926–935, 1983.
40. H. M. Senn, D. O'Hagan, W. Thiel, *Insight into Enzymatic C–F Bond Formation from QM and QM/MM Calculations*, J. Am. Chem. Soc. **127**, 13643–13655, 2005.
41. M. J. Abraham, T. Murtola, S. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers*, SoftwareX **1-2**, 19–25, 2015.
42. CPMD program, `http://www.cpmd.org`, Copyright IBM Corp 1990-2019, Copyright MPI für Festkörperforschung Stuttgart 1997-2001.
43. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, *PLUMED 2: New feathers for an old bird*, Comput. Phys. Commun. **185**, 604–613, 2014.
44. A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A **38**, 3098–3100, 1988.
45. C. Lee, W. Yang, R. G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B **37**, 785–789, 1988.
46. N. Troullier, J. L. Martins, *Efficient pseudopotentials for plane-wave calculations*, Phys. Rev. B **43**, 1993-2006, 1991.
47. A. Laio, J. VandeVondele, U. Rothlisberger, *A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations*, J. Chem. Phys. **116**, 6941–6947, 2002.
48. A. M. Elena, S. Meloni, G. Ciccotti, *Equilibrium and Rate Constants, and Reaction Mechanism of the HF Dissociation in the HF·$(H_2O)_7$ Cluster by Ab Initio Rare Event Simulations*, J. Phys. Chem. A **117**, 13039–13050, 2013.
49. A. J. Sillanpää, C. Simon, M. L. Klein, K. Laasonen, *Structural and Spectral Properties of Aqueous Hydrogen Fluoride Studied Using Ab Initio Molecular Dynamics*, J. Phys. Chem. B **106**, 11315–11322, 2002.
50. B. Dereka, Q. Yu, N. H. C. Lewis, W. B. Carpenter, J. M. Bowman, A. Tokmakoff, *Crossover from Hydrogen to Chemical Bonding*, Science **160**, 160–164, 2021.
51. M. Walden, A. Accardi, F. Wu, C. Xu, C. Williams, C. Miller, *Uncoupling and turnover in a Cl^-/H^+ exchange transporter*, J.Gen.Physiol **129**, 317–329, 2007.
52. Jülich Supercomputing Centre, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, Journal of large-scale research facilities **5**, A135, 2019.

Elementary Particle Physics

Elementary Particle Physics

Stefan Schaefer

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

Most of the visible matter is made of protons and neutrons forming the atomic nuclei. The theory describing these two constituents has been known for five decades: quantum chromodynamics (QCD). It describes protons and neutrons as composed of quarks held together by gluons.

What makes studying this theory interesting after so many years? The answers the contributions in this section give is quite simple: There is a lot we do not know about this important part of the Standard Model of Particle Physics and computational techniques to compute quantities of interest directly from the theory have become available only recently.

The ultimate aim of the contributions in this volume is to make theory more predictive and by doing so enhance the possibilities of the internationally ongoing experimental programmes which test the Standard Model of particle physics. In collider experiments like at the LHC at CERN in Geneva, the properties of the known elementary particles and their interactions are studied with the hope to find unknown particles or effects. These searches so far confirmed the particle spectrum and interactions as described by the five decades old Standard Model and culminated with the discovery of the Higgs boson a decade ago. This is a spectacular success, however, there are good reasons to believe that there is physics beyond the Standard Model motivating these experimental efforts.

In many of these searches, what we know about quarks and gluons is a limiting factor. This is because in most experiments protons (or other hadrons) are used as a probe in scattering experiments at high energy and for the interpretation their structure needs to be known. In principle, many of the required observables could be extracted from experimental data, but this limits the accuracy and the discovery potential.

Here is where efforts like the ones presented in this section enter. By performing *ab initio* computations in the relevant part of the Standard Model, information relating to how quarks and gluons form hadrons like protons, neutrons or pions can be obtained. As explained in the first contribution by Alexandrou *et al.*, we do not have a clear picture of this. Unlike the electron and the nucleus forming an atom held together by photons, we do not observe free quarks and gluons. They are strongly coupled and aspects of the structure of their bound states have been revealed only by experiments like the ones at the LHC or the former DESY collider HERA in Hamburg. The aim is to supersede these with direct computations and the following contributions make it clear that significant progress has been made into this direction.

The first contribution by Alexandrou *et al.* highlights the challenges faced when trying to reveal the structure how pions are made of quarks and gluons. Conceptual work together with the development of algorithms and efficient use of computational resources have led to the results that constitute an important breakthrough. It could be shown that the independently computed contributions from quarks and gluons to the pion do add up correctly

to the expected total. This has not been the case in the past and demonstrates the progress that could be achieved in recent years.

A highlight of the article by Bali for the RQCD collaboration is the computation of so-called nucleon σ-terms which parametrise the scattering of baryons, like the proton, with scalar particles. These are important for the interpretation of dark matter searches where these hypothetical particles would interact with the protons inside the atomic nucleus. There are many such experiments currently on-going and the presented computation of the σ-terms is an example of a quantity where QCD computations have become the most accurate source.

The final contribution by Jansen for the ETM collaboration gives a flavour of the efforts needed to obtain the results and check their reliability. All collaborations employ a Markov Chain Monte Carlo method and here ergodicity is a mayor concern. Once this is verified, results of several important quantities can be presented: the masses of the quarks (for which these *ab initio* computations have become the most precise source) as well as several observables related to the structure of protons and their behaviour in scattering experiments. Also here, algorithmic progress and a clear exposition of the challenges have paved a way to substantial advances and precise results.

All contributions together give a good overview of the level of accuracy which can currently be achieved and which is such that it can indeed make an important impact on the many large scale experimental efforts around the world which try to get hold of what comes beyond the Standard Model of Particle Physics.

The Quark and Gluon Momentum Fractions in the Pion

Constantia Alexandrou[1,2], Simone Bacchio[2], Georg Bergner[3], Jacob Finkenrath[2], Andrew Gasbarro[4], Kyriakos Hadjiyiannakou[1,2], Karl Jansen[5], Bartosz Kostrzewa[6], Konstantin Ottnad[7], Marcus Petschlies[8], Ferenc Pittler[1], Fernanda Steffens[8], Carsten Urbach[8], and Urs Wenger[4,9]

[1] Department of Physics, University of Cyprus, 4CV6+V63, Aglantzia, Nicosia, Cyprus

[2] Computation-based Science and Technology Research Center, The Cyprus Institute, 20 Konstantinou Kavafi Street, 2121, Aglantzia, Nicosia, Cyprus

[3] Institute of Theoretical Physics, Friedrich-Schiller-University Jena, Fröbelsteig 1, 07743 Jena, Germany

[4] Albert Einstein Center, Institute for Theoretical Physics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland

[5] NIC, DESY, Platanenallee 6, 15738 Zeuthen, Germany

[6] High-Performance-Computing & Analytics Lab, Rheinische Friedrich-Wilhelms-Universität Bonn, Friedrich-Hirzebruch-Allee 8, 53115 Bonn, Germany

[7] PRISMA$^+$ Cluster of Excellence and Institut für Kernphysik, Johannes-Gutenberg-Universität Mainz, Johann-Joachim-Becher-Weg 45, 55128 Mainz, Germany

[8] Helmholtz-Institut für Strahlen- und Kernphysik, Bethe Center for Theoretical Physics, Rheinische Friedrich-Wilhelms-Universität Bonn, Nussallee 14-16, 53115 Bonn, Germany
E-mail: urbach@hiskp.uni-bonn.de

[9] Department of Theoretical Physics, CERN, 1211 Geneva, Switzerland

1 Introduction

The standard model (SM) of particle physics combines three of the four known fundamental forces in nature, and is extremely successful: up to now there are no deviations between the theoretical predictions of the SM and its experimental verification which are beyond reasonable doubt.

One of the three forces combined in the SM is the so-called strong force, which is governed by a theory called quantum chromodynamics (QCD). QCD describes the interaction of quarks and gluons, where quarks represent the fundamental fermionic matter and gluons the force carriers. This is very similar to quantum electrodynamics (QED), where electrons represent the fundamental matter particles with photons the force carriers.

However, the similarity between QCD and QED ends quickly, because in QCD the gluons can interact with themselves due to the $\mathrm{SU}(3)$ gauge structure, something photons are not allowed to do. This very fact has a plethora of consequences. Most importantly, QCD is a theory that looks radically different depending on the energy scale that is involved: at low energies (long distances) QCD is confined and strongly coupled, at large energies (short distances) QCD is asymptotically free. The low energy regime in QCD is

of particularly high interest, because all the matter surrounding us, i.e. mostly protons and neutrons bound in atomic nuclei, is composed out of quark bounds states of QCD.

Returning to the similarities between QED and QCD, one immediate question is how the structure of the QCD bound states is comparable to the QED bound states. In QED an atom is classically imagined as a bound state of a positively charged core and negatively charged electrons where the electrons surround the core on approximately circular trajectories. However, there is no such simple classical analogous picture for the structure of a proton or neutron. The closest picture would be that of quarks confined in a bag, where the internal walls of the bag play the role of a confining potential. Reality, however, is much more complex, and information about the sub-structure of protons and neutrons, or more generally of hadrons, which are all states generated in QCD, is encoded in so-called parton distribution functions (PDFs), where parton is a generic name for both quarks and gluons. Thus, the knowledge of PDFs will boost our understanding of how QCD generates, for instance, the observed mass and spin of hadrons. Hadron structure is also an important topic for phenomenology in general, as PDFs are important ingredients for reliable analyses of experimental data, for instance those produced at the Large Hadron Collider (LHC) at CERN in Switzerland.

Beyond protons and neutrons there are other hadrons generated by QCD. Of particular importance are the pions, which is a bound state of light quark-antiquark pairs. Their particular importance stems from the fact that they are the Goldstone bosons related to spontaneously broken chiral symmetry in QCD, which makes them the lightest states in QCD. At the same time, they are crucial to understand the nuclear force keeping protons and neutrons bound in atomic nuclei. Given their fundamental importance towards our comprehension of the visible world around us, it is pressing to answer questions like *What are the quark and gluon energy contributions to the pion mass?* and *Is the pion full or empty of gluons as viewed at large squared momentum transfer Q^2?*, as recently emphasised in the Electron Ion Collider (EIC) yellow report[1].

2 Pion PDFs from Lattice QCD

PDFs are fundamental quantities in QCD, as argued above. Because they are defined at low energy, where the QCD coupling constant is large, they are inaccessible to perturbation theory. This non-perturbative nature leaves two possibilities to extract information about PDFs: use experimental data or appropriate *ab initio* theoretical methods. Lattice QCD, a discretised version of QCD, represents the most successful non-perturbative tool to investigate QCD numerically, with systematically improvable uncertainties.

However, since lattice QCD is defined in Euclidean space-time, it was not known for a long time how to compute the PDFs directly, as these quantities are light-cone dominated. In the last few years, new techniques have been developed, within which encouraging results have been obtained[5]. Yet, a precise computation of PDFs using lattice QCD is still out of reach. The computation of the so-called Mellin moments of PDFs, on the other hand, can be made with higher precision. The first of these moments, which is simply the integral over the momentum fraction of all quarks making up a hadron, also has a nice interpretation: it represents the average momentum fraction carried by the different constituents, i.e. quarks and gluons, in the hadron. Per definition, the sum of these different fractions must equal to one. And there is decisive progress in the computation of moments

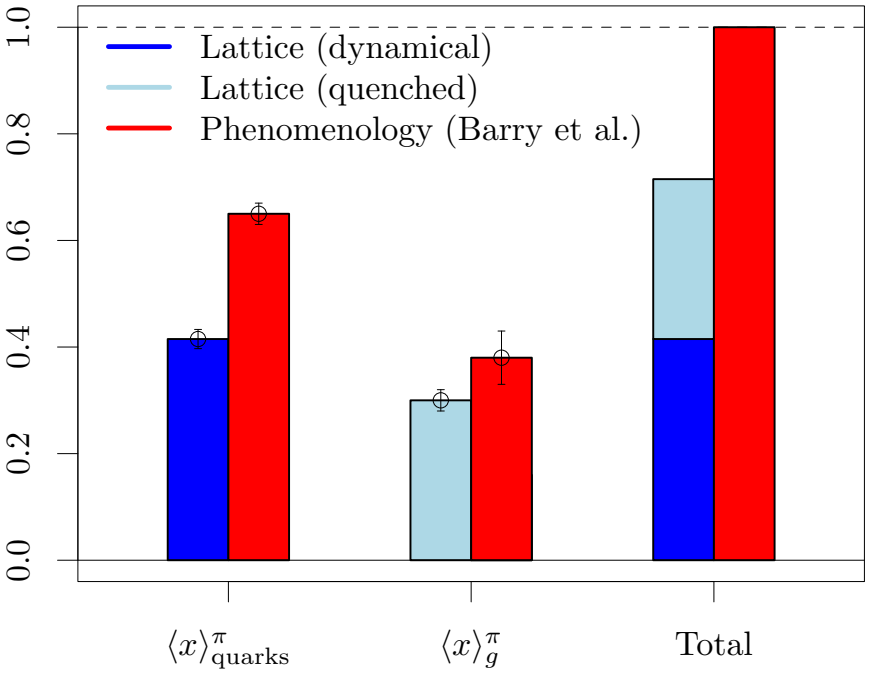


Figure 1. Lattice QCD status of contributions to the average momentum of the pion comparing phenomenology[2], lattice QCD[3] and quenched lattice QCD[4] before the computations made possible by the NIC computing resources. Results are at 2 GeV in the $\overline{\text{MS}}$ renormalisation scheme.

of PDFs from lattice QCD over the recent years. But, despite the fundamental importance of the pion, there was little known on the structure of the pion from lattice QCD.

The status of lattice calculations for the pion before the calculation enabled by the NIC computer resources is summarised in Fig. 1. From the lattice side there were calculations of the quark momentum fraction $\langle x \rangle^{\pi}_{\text{quarks}}$ available, the most recent one[3] shown in the figure: this calculation, even though performed in lattice QCD with dynamical up/down, strange and charm quark flavours, did not take the so called fermionically disconnected contributions into account. For the gluon contribution $\langle x \rangle^{\pi}_{g}$ there was only one computation in the so called quenched approximation available[4]. In the figure we compare to a phenomenological determination by Barry and co-authors[2]. While the different phenomenological contributions sum to one per construction, this is not true for the lattice contributions.

It is clear that important contributions are missing: first of all, every quark flavour is expected to give a contribution. The largest contribution is expected to come from the two light quarks, the up and down quarks. Following the mass hierarchy, the next largest contribution is expected from the strange and then the charm quark. The contribution from the bottom quark is expected to be negligible and the top quark is too heavy.

Moreover, in the process of computing the momentum fractions there are different contributions to be taken into account. These are visualised diagrammatically in Fig. 2. In these three diagrams time runs from left to right and, thus a pion π is created and anni-

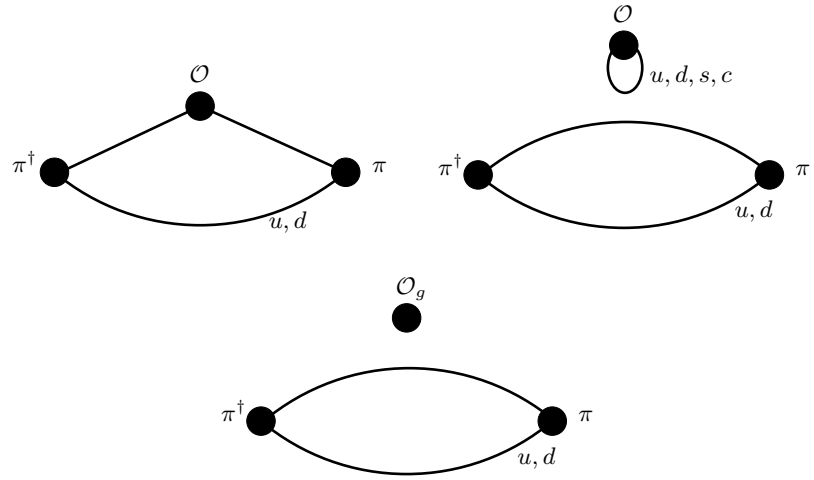


Figure 2. Diagrammatic representation of the so-called contractions to be computed for $\langle x \rangle$. The upper left diagram corresponds to the fermionically connected contribution, the upper right to the fermionically disconnected contribution and the lower one to the gluon contribution to the momentum fraction.

hilated after some time span. At some intermediate time an operator $\mathcal{O}$ is inserted. The upper left diagram corresponds to the fermionically connected contribution. The quark lines are only the light up and down type quarks. The upper right diagram is the fermionically disconnected contribution. Here the light and strange and charm quarks contribute to the quark line connected to the operator. And lastly, the lower diagram corresponds to the gluon contribution. Oehm and co-authors[3] have only considered the fermionically connected contribution, while the gluon contribution had only been computed in the quenched approximation.

Finally, the computation of the momentum fractions is complicated by the fact that the operators making up the pion field must be renormalised. Renormalisation is needed to remove infinities in quantum field theories. On top of it, in the process of renormalisation there occurs mixing between all operators with the same quantum numbers of that used in the pion field, when symmetries are not sufficient to eliminate such mixing.

3 Results

The result of this calculation is published in Ref. 6, where more details of the calculation can be found.

We have used one simulation ensemble generated by Extended Twisted Mass Collaboration (ETMC)[7], for which the bare quark mass parameters were tuned such as to reproduce a number of physical values of hadron masses and decay constants. Up to artefacts from the discretisation and the finite volume – both needed for the simulation to be practical – this ensemble should reproduce experimental results, if QCD is the correct theory. Most importantly, the pion mass is tuned to its physical value.

On this ensemble we have computed all contractions shown in Fig. 2. The renormalisation and mixing coefficients have been computed in another project[8]. We then compute

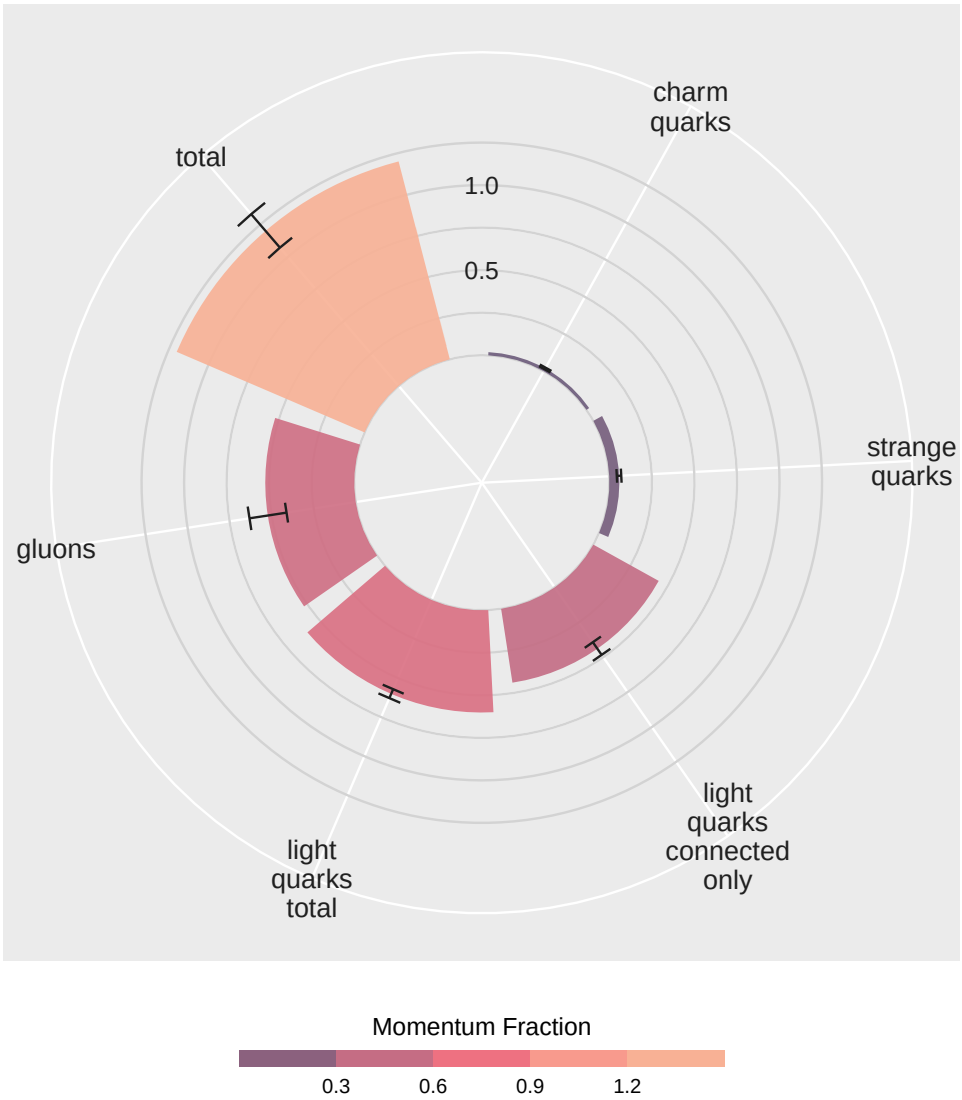


Figure 3. The final result for the momentum fractions of quarks and gluons in the pion. With total we denote the sum of light, strange, charm and gluon fractions. The light quark connected only contribution is part of light quarks total.

the renormalised momentum fractions from

$$\begin{pmatrix} \sum_f \langle x \rangle_f^{\mathrm{R}} \\ \langle x \rangle_g^{\mathrm{R}} \end{pmatrix} = \begin{pmatrix} Z_{qq}^s & Z_{qg} \\ Z_{gq} & Z_{gg} \end{pmatrix} \begin{pmatrix} \sum_f \langle x \rangle_f \\ \langle x \rangle_g \end{pmatrix} \tag{1}$$

with Z_{qq}^s the quark-singlet renormalisation constant, Z_{qg} and Z_{gq} the mixing coefficients and Z_{gg} the gluon renormalisation constant. The sum over f denotes the sum over the different quark flavours, i.e. up, down, strange and charm.

The result is shown in Fig. 3, where we display the separate contributions from light, strange and charm quarks, as well as the contribution from the gluons. In addition we show the connected only contribution, which exists only for the light quarks and which is a part of the total contribution by the light quarks.

Since the distribution between quark flavours and gluons is depending on the renormalisation scale, which is chosen to be 2 GeV in the $\overline{\mathrm{MS}}$-scheme in Fig. 3, the most interesting question is for the sum of all contributions, denoted as total in Fig. 3: it should equal one independent of the scale and the renormalisation scheme. And this is the case for our calculation within the still sizable statistical uncertainties.

	this work	Ref. 9	Ref. 10	Ref. 11
$\langle x \rangle_l^{\mathrm{R}}$	$0.601(28)(_{-21})$	–	–	–
$\langle x \rangle_s^{\mathrm{R}}$	$0.059(13)(_{-10})$	–	–	–
$\langle x \rangle_c^{\mathrm{R}}$	$0.019(05)(_{-10})$	–	–	–
$\langle x \rangle_g^{\mathrm{R}}$	$0.52(11)(^{+02})$	–	0.42(4)	0.25(13)
$\sum_f \langle x \rangle_f^{\mathrm{R}}$	$0.68(05)(_{-03})$	0.220(207)	0.58(9)	0.75(18)

Table 1. Compilation of results for the momentum fraction in the pion and comparison to literature. All values are at 2 GeV in the $\overline{\mathrm{MS}}$ scheme.

The results are compiled again in Tab. 1. In this table we also compare to results available in the literature.

4 Outlook

Even though this was the first time all contributions to the momentum fraction in the pion could be computed, thanks to the computing resources available to us in Jülich, there is more to do: first of all, lattice artefacts need to be studied. As mentioned above, this calculation is based on a single ensemble at a single value of the lattice spacing. The limit of vanishing lattice spacing needs to be investigated. Second, the mixing coefficients need to be computed non-perturbatively just like the diagonal coefficients. Third, statistical uncertainties could be reduced. With all this combined we hope to be able to finally check the sum rule more stringently.

This calculation is currently being repeated for other meson states. For instance kaons, which are very similar to pions but with one light quark replaced by a strange quark will help us to understand the role of the different quark flavours in mesons better, as well as the process of mass generation from strong forces.

Acknowledgements

This project was carried out as part of ETMC. We thank all members of ETMC for the most enjoyable collaboration. We thank JSC and NIC for making the computing time available to us. Parts of the results were created within the EA program of JUWELS Booster and we are thankful for the help and support of the JUWELS Booster Project Team (JSC,

Atos, ParTec, NVIDIA). We also thank for computing time granted on Piz Daint at Centro Svizzero di Calcolo Scientifico (CSCS) via projects s849, s702 and s954 supported by Swiss National Science Foundation SNSF grant No.200021_175761. This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and the NSFC through the funds provided to the Sino-German Collaborative Research Center CRC 110 "Symmetries and the Emergence of Structure in QCD" (DFG Project-ID 196253076 - TRR 110, NSFC Grant No. 12070131001) as well as the STFC Consolidated Grant ST/T000988/1. The open source software packages tmLQCD[12–14], Lemon[15], QUDA[16–18] and R[19] and hadron[20] have been used.

References

1. R. Abdul Khalek et al., *Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report*, 3 2021.
2. P. C. Barry, N. Sato, W. Melnitchouk, and C.-R. Ji, *First Monte Carlo Global QCD Analysis of Pion Parton Distributions*, Phys. Rev. Lett., **121**, no. 15, 152001, 2018.
3. M. Oehm, C. Alexandrou, M. Constantinou, K. Jansen, G. Koutsou, B. Kostrzewa, F. Steffens, C. Urbach, and S. Zafeiropoulos, *$\langle x \rangle$ and $\langle x^2 \rangle$ of the pion PDF from lattice QCD with $N_f = 2 + 1 + 1$ dynamical quark flavors*, Phys. Rev. D, **99**, no. 1, 014508, 2019.
4. H. B. Meyer and J. W. Negele, *Gluon contributions to the pion mass and light cone momentum fraction*, Phys. Rev. D, **77**, 037501, 2008.
5. X. Ji, Y.-S. Liu, Y. Liu, J.-H. Zhang, and Y. Zhao, *Large-momentum effective theory*, Rev. Mod. Phys., **93**, no. 3, 035005, 2021.
6. C. Alexandrou et al., *Quark and Gluon Momentum Fractions in the Pion from Nf=2+1+1 Lattice QCD*, Phys. Rev. Lett., **127**, no. 25, 252001, 2021.
7. C. Alexandrou et al., *Simulating twisted mass fermions at physical light, strange and charm quark masses*, Phys. Rev. D, **98**, no. 5, 054518, 2018.
8. C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, H. Panagopoulos, and G. Spanoudes, *Complete flavor decomposition of the spin and momentum fraction of the proton using lattice QCD simulations at physical pion mass*, Phys. Rev. D, **101**, no. 9, 094513, 2020.
9. M. Löffler, D. Jenkins, R. Rödl, A. Schäfer, L. Walter, P. Wein, S. Weishäupl, and T. Wurm, *Mellin moments of spin dependent and independent PDFs of the ρ and π*, 8 2021.
10. P. C. Barry, C.-R. Ji, N. Sato, and W. Melnitchouk, *Global QCD analysis of pion parton distributions with threshold resummation*, 8 2021.
11. I. Novikov et al., *Parton Distribution Functions of the Charged Pion Within The xFitter Framework*, Phys. Rev. D, **102**, no. 1, 014040, 2020.
12. K. Jansen and C. Urbach, *tmLQCD: A Program suite to simulate Wilson Twisted mass Lattice QCD*, Comput.Phys.Commun., **180**, 2717–2738, 2009.
13. A. Abdel-Rehim, F. Burger, A. Deuzeman, K. Jansen, B. Kostrzewa, L. Scorzato, and C. Urbach, *Recent developments in the tmLQCD software suite*, PoS, **LATTICE2013**, 414, 2014.
14. A. Deuzeman, K. Jansen, B. Kostrzewa, and C. Urbach, *Experiences with OpenMP in tmLQCD*, PoS, **LATTICE2013**, 416, 2013.

15. A. Deuzeman, S. Reker, and C. Urbach, *Lemon: an MPI parallel I/O library for data encapsulation using LIME*, Comput. Phys. Commun., **183**, 1321–1335, 2012.
16. M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi, *Solving Lattice QCD systems of equations using mixed precision solvers on GPUs*, Comput. Phys. Commun., **181**, 1517–1528, 2010.
17. R. Babich, M. A. Clark, B. Joo, G. Shi, R. C. Brower, and S. Gottlieb, *Scaling Lattice QCD beyond 100 GPUs*, in: SC11 International Conference for High Performance Computing, Networking, Storage and Analysis Seattle, Washington, November 12-18, 2011.
18. M. A. Clark, B. Joó, A. Strelchenko, M. Cheng, A. Gambhir, and R. Brower, *Accelerating Lattice QCD Multigrid on GPUs Using Fine-Grained Parallelization*, 2016.
19. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2019.
20. B. Kostrzewa, J. Ostmeyer, M. Ueding, and C. Urbach, "hadron: package to extract hadronic quantities", `https://github.com/HISKP-LQCD/hadron`, 2020, R package version 3.0.1.

Baryons, their Flavour Structure and Low Energy Constants

Gunnar Bali
for the RQCD Collaboration

Institut für Theoretische Physik, Universität Regensburg, 93040 Regensburg, Germany
E-mail: gunnar.bali@ur.de

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. To relate experimental cross sections and decay rates to an underlying fundamental theory, i.e. the Standard Model of particle physics or theories beyond the Standard Model, that describes interactions on the quark and gluon level rather than with the nucleons, a precise and sound derivation of the properties of nucleons from those of their quark and gluon constituents is required. The part of the Standard Model that needs to be solved to provide this missing link is quantum chromodynamics (QCD). This requires simulations on four-dimensional spacetime lattices (Lattice QCD) with a lattice spacing a and a subsequent continuum limit extrapolation. We include so-called hyperons that contain strange quarks, in addition to the up and down quarks of the nucleons, at many different quark mass combinations, in our study. This enables us to test the validity of quark flavour symmetry relations and to predict low energy constants of SU(3) chiral perturbation theory (ChPT). Here we report on results of the mass spectrum and on SU(3) ChPT low energy constants.

The simulations are carried out in QCD with two mass-degenerate light quarks and a strange quark, 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 full order a improvement of lattice spacing effects. The main computational task in the analysis that was mostly carried out on the Xeon-Phi Booster module of JURECA was the multiple solution of large sparse linear systems with a dimension of up to $(2 \cdot 10^9)^2$ complex variables. This is achieved by an adaptive algebraic multigrid algorithm.

1 Introduction

With the discovery of the Higgs boson at the LHC at CERN in 2012 the existence of the last particle predicted by the Standard Model was confirmed. The Standard Model, which is the accepted theory of elementary particles and their interactions, successfully passed an impressive number of precision tests. 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. Therefore, interactions between this matter and Standard Model matter must be very weak. Once scattering between dark matter and baryonic matter is detected, matrix elements like those computed in this project will be required to relate experimental observations to interactions between dark matter particles and quarks.

The Standard Model unifies the electromagnetic and weak interactions and contains the theory of strong interactions (QCD) that binds 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 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. 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. Clearly, deriving the structure and properties of nucleons in terms of those 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, lattice spacings and statistical precision required to make an impact in this field. Here we highlight some of our recent results.

2 The Gauge Ensembles Employed and the Role of ChPT

Lattice QCD is well suited for computing properties like hadron masses, decay constants and so-called three-point functions, that reveal further aspects of hadron structure. However, it is clear that regarding more complex experimental set-ups, our *ab initio* approach has its practical limitations. Based on the assumption of chiral symmetry breaking and the global symmetries of massless QCD, an effective field theory can be constructed, ChPT, that provides a systematic expansion in terms of pseudoscalar meson masses and momenta, of key hadronic observables. The interplay between ChPT and Lattice QCD simulations is particularly exciting: while in nature the quark masses are fixed, in lattice simulations these (and other simulation parameters) can be varied and the precision and the range of validity of truncated ChPT expansions explored systematically. Moreover, some of the otherwise unknown low energy constants (LECs) of this effective field theory can be constrained or determined from lattice data, which complements phenomenological fits to experimental data that are limited to the physical quark mass point. *Vice versa*, ChPT augments Lattice QCD simulations, providing parametrisations of the dependence of the results on the light quark masses and on the simulation volume.

While the light pseudoscalar masses, decay constants, the chiral condensate and related mesonic quantities have been well explored in Lattice QCD simulations and confronted with SU(2) ChPT predictions – see, e.g., the recent Flavour Lattice Averaging Group (FLAG) review[1] – this is less so regarding baryonic observables or SU(3) ChPT. On the one hand, the lattice data are less precise for baryons, in particular towards small values of the quark masses. On the other hand, the number of independent LECs is larger and also the convergence properties of ChPT are less well explored in the baryonic sector. Including hyperons, i.e., the Λ, the Σ (containing two light and one strange quark) and the Ξ (one light and two strange quarks), into the ChPT analysis, in addition to the nucleon N (three light quarks), provides a wealth of additional information, whereas the

number of baryonic LECs of flavour SU(3) ChPT increases only moderately relative to SU(2) ChPT. This makes SU(3) ChPT a particularly popular choice in the description of processes that involve baryons. One concern regarding phenomenological applications, however, is the convergence of SU(3) ChPT at the physical point itself, where neither the mass $M_{\eta_8} \approx (\frac{4}{3}M_K^2 - \frac{1}{3}M_\pi^2)^{1/2} \approx 565$ MeV of the heaviest pseudoscalar meson nor the average meson mass $\overline{M} = (\frac{2}{3}M_K^2 + \frac{1}{3}M_\pi^2)^{1/2} \approx 411$ MeV are particularly small in comparison to the chiral symmetry breaking scale $\Lambda_\chi := 4\pi F_0 \approx 900$ MeV[2]. While this may limit the practical applicability of SU(3) ChPT regarding some observables, the corresponding LECs are well-defined and can in principle be obtained from Lattice QCD.

Within most lattice simulations the mass of the light quark $m_\ell = m_u = m_d$ is varied while that of the strange quark m_s is kept approximately fixed near its physical value. In a few cases, instead $\mathrm{tr}\, M == 2m_\ell + m_s$ is kept constant[3–5]. The former setting is ideal regarding SU(2) ChPT while neither choice is sufficient to determine SU(3) LECs, unless other quark mass combinations are added; in particular, one may want to reduce the trace of the quark mass matrix $\mathrm{tr}\, M$ below its physical value. This can be achieved via a partially quenched strategy or, ideally, including additional sea quark mass combinations[6,7,2].

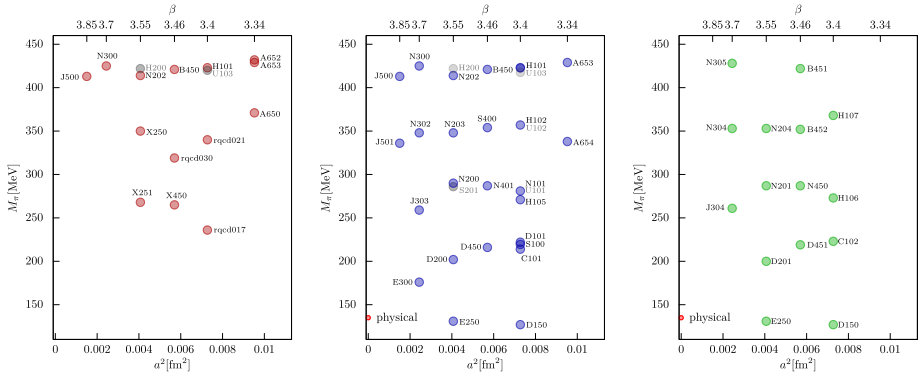


Figure 1. Overview of the analysed ensembles: three different quark mass trajectories (left: $m_s = m_\ell$, centre: $\mathrm{Tr}\, M = \text{const}$, right: $\widehat{m}_s \approx \text{const}$) have been analysed at six (four for $\widehat{m}_s \approx \text{const}$) different lattice spacings. In the case $m_s = m_\ell$ six additional ensembles with $M_\pi > 450$ MeV exist (A651, rqcd019, rqcd029, B250, N303 and N500), which do not enter our present analysis. We also omit the grayed out ensembles with $LM_\pi < 2.3$ fm. Additional ensembles are in production, e.g., on JUWELS.

We determine our observables on ensembles generated by the Coordinated Lattice Simulations (CLS) effort, employing $N_f = 2+1$ flavours of non-perturbatively improved Wilson fermions. The hadron masses are interpolated and extrapolated within the quark mass plane, utilising three trajectories: (a) along the "symmetric" line $m_s = m_\ell$ where strange and light quark masses are equal, (b) along a nearly-physical fixed value of the trace of the quark mass matrix and (c) varying the light quark mass while keeping the renormalised strange quark mass $\widehat{m}_s$ almost constant near its physical value. The latter two trajectories intersect close to the physical quark mass point. We analyse a large number of ensembles along these three trajectories, most of them with open boundary conditions in time[8] to circumvent critical slowing down towards the continuum limit, due to the freezing of the

topological charge[9]. An overview of the ensembles analysed so far is shown in Fig. 1. The lattice spacing is varied from $a \approx 0.098\,\text{fm}$ down to $a \approx 0.039\,\text{fm}$, the pion mass from $M_\pi \approx 429$ MeV down to $M_\pi \approx 127\,\text{fm}$, where the leading discretisation correction is quadratic in the lattice spacing and the pion mass that is realised in nature ("physical") is $M_\pi \approx 135$ MeV. The lattice dimensions $N_t \cdot N_s^3$ increase with decreasing a and M_π since $L = N_s a \propto 1/M_\pi$. For instance, the ensembles E250 and E300 contain $192 \cdot 96^3$ lattice sites while the A65n ensembles correspond to only $24 \cdot 48^3$ sites. At each point four SU(3) valued gauge links are attached, while the quark spinors have 12 complex components at each site.

In the isospin symmetric limit, to order p^3 in covariant baryon ChPT (BChPT), all four octet baryon masses and the finite volume effects only depend on the leading order LEC m_0, the next-to-leading order (NLO) LECs b_0, b_F and b_D and, at order p^3, the LECs D and F that correspond to the axial baryon charges in the chiral limit, in units of the pion decay constant F_0[10]. With these six parameters, it is possible to fit all our baryon masses, where we use the extended-on-mass-shell (EOMS)[11] regularisation of BChPT.

3 Selected Results

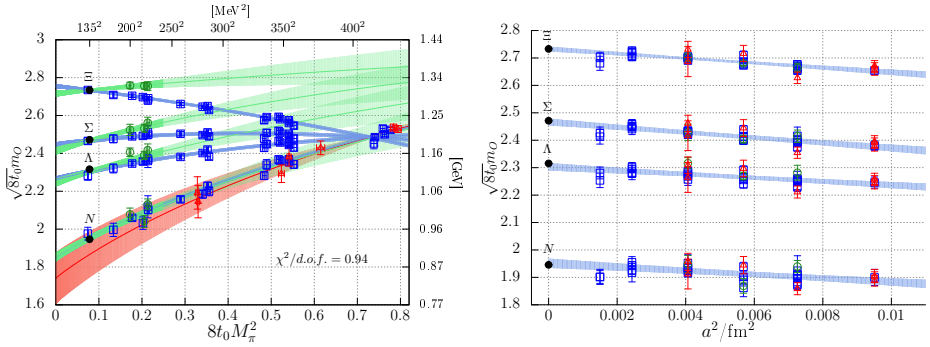


Figure 2. Left: the EOMS BChPT fit with the best χ^2/N_{DF}, imposing cuts on the average pseudoscalar mass $\overline{M} = [(2M_K^2 + M_\pi^2)/3]^{1/2} < 440$ MeV and on the lattice size $LM_\pi > 4$. The data are shifted to infinite volume and the continuum limit according to the fit. The experimental values correspond to $\sqrt{8t_0} = 0.4093\,\text{fm}$, which is obtained using the Ξ baryon mass as an input. The green and blue error bands ($\widehat{m}_s \approx$ const and $\text{tr}\, M =$ const) intersect at the physical point while the red band ($m_s = m_\ell$) goes to the chiral limit. Right: the continuum limit extrapolation for the same fit, with the data shifted to the physical quark mass point and infinite volume.

We carry out a variety of fits, using polynomial expansions, the EOMS BChPT parametrisation and heavy baryon ChPT (HBChPT). The best results are obtained in the EOMS BChPT approach with six parameters describing the volume and quark mass dependence of all the four baryons in the continuum limit and another six parameters describing the quadratic lattice spacing effects. We show the resulting 12 parameter fit to 92 data points, projected onto the quark mass plane in the infinite volume continuum limit, in Fig. 2 (left) and singling out the lattice spacing dependence at the physical point (right).

Not only is the fit quality very good but also the resulting LECs F and D are in agreement with those obtained from the axial decay constants of the octet baryons[2]. It appears that for average pseudoscalar masses smaller than 450 MeV, within our present statistical accuracy, the mass data are consistently described by order p^3 SU(3) EOMS BChPT. This, in addition to the predicted LECs, is a very valuable result since this includes the physical point $\overline{M} = 411$ MeV. For the future we plan a simultaneous analysis of the spectrum and nucleon structure observables within SU(3) BChPT, extending to our pilot study[2] to points away from the symmetric $m_s = m_\ell$ line and incorporating additional observables, e.g., the σ-terms[12].

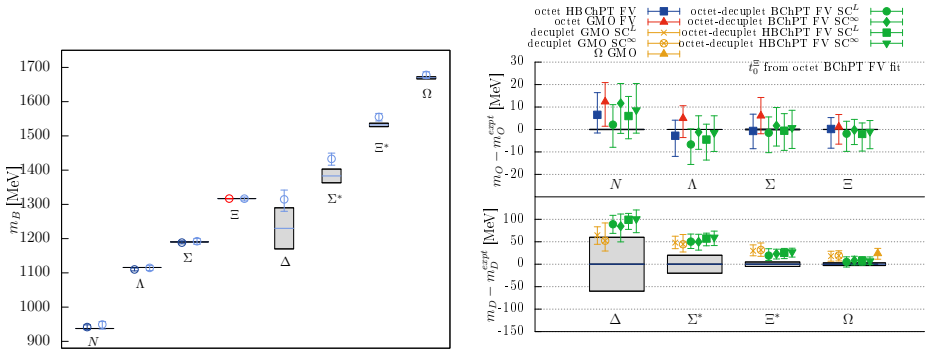


Figure 3. Left: the baryon spectrum, obtained from the SU(3) EOMS BChPT parametrisation, including systematic uncertainties. Dark symbols are obtained from a fit to the octet baryons only, where the Ξ baryon mass (red) was used as an input to set the scale. The light symbols correspond to a joint fit to octet and decuplet baryons. The boxes are the experimental values, corrected for isospin breaking effects and including (for the Δ, Σ^* and Ξ^*) the experimental decay widths. Right: differences between our predictions and the experimental values, using a multitude of different fit forms.

We also analysed the decuplet baryons Δ (three light quarks), Σ^* (two light quarks and one strange quark), Ξ^* (one light quark and two strange quarks) and Ω (three strange quarks) jointly with the octet baryons in the framework of SU(3) BChPT, also including pion-baryon loops. In the right panel of Fig. 3 we compare our results from the octet baryon fits (dark symbols) and from the combined octet and decuplet baryon fits (light symbols) to the known baryon masses (boxes). The Ξ baryon mass (red) was used as an input to set the scale. In nature all decuplet baryons other than the Ω undergo strong decay into a pion and an octet baryon. The respective decay widths are added to the experimental uncertainty. Our predictions for these resonances must be taken with a grain of salt since we ignore the decay. Ideally, one would carry out a finite volume study with additional baryon-pion interpolators to determine the scattering parameters. The results shown are weighted averages of BChPT fits with different cuts regarding the volume, the lattice spacings included and the maximum average quark mass. The errors that include systematics are therefore somewhat bigger than those of the individual fit discussed above. For the nucleon mass we obtain a precision of about 0.8%, including all systematics, which was unprecedented so far. Isospin breaking effects account for a difference of 0.2% between the proton and the neutron mass and should be included if we were to increase our precision further. In

the right panel of Fig. 3 we compare results on the spectrum obtained, using different fit ansätze: due to the good coverage of the quark mass plane, including the physical point, and the large number of different lattice spacings, the results are data driven and are quite independent of the parametrisation used for the extrapolation. As mentioned above, we also computed the corresponding ChPT LECs.

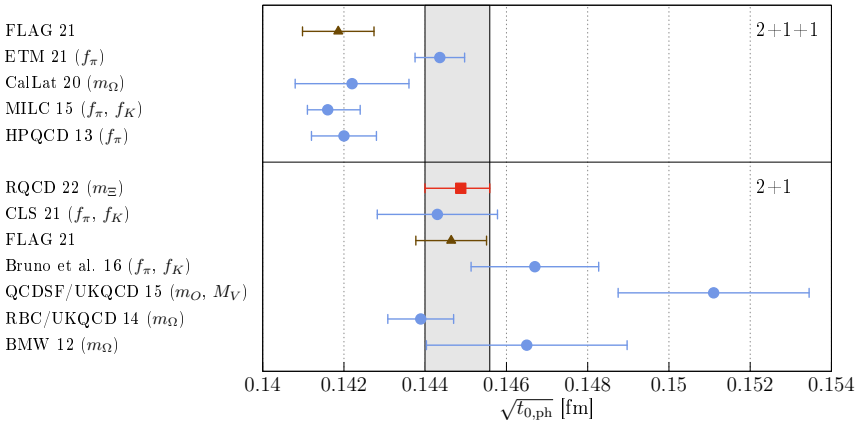


Figure 4. Comparison of our determination of $\sqrt{t_{0,\mathrm{ph}}}$ from the Ξ baryon mass (red square and shaded grey region) to other determinations of this quantity with $N_f = 2+1$ flavours (CLS 21[13], Bruno et al. 16[5], QCDSF/UKQCD 15[14], RBC/UKQCD 14[15] and BMW 12[16]) and $N_f = 2+1+1$ flavours (ETM 21[17], CalLat 20[18], MILC 15[19] and HPQCD 13[20]). The experimental quantities used as an input to set the scale are indicated in brackets. Also shown as green triangles are the averages of the FLAG Review 2021[1].

The quark masses and the scale are not determined by QCD but by the electroweak sector of the Standard Model. We use the experimental pion and kaon masses as input to determine the physical point in the quark mass plane and the Ξ baryon mass to set the scale. Once this is carried out, everything else can be predicted. In lattice simulations the scale setting is particularly important since its uncertainty limits the precision that can be achieved for any dimensionful quantity. A popular choice to relate different lattice spacings with each other is the scale t_0[21] that can be determined easily and very accurately. Using the experimental value $m_\Xi = 1317$ MeV and determining the product $\sqrt{t_0} m_\Xi$ at the physical point gives the value $\sqrt{8t_0} = 0.4098^{(20)}_{(25)}$ fm. In Fig. 4, we compare this result with literature values: in contrast to some of the other values shown, we carefully studied and included all systematic sources of error, in particular regarding the continuum limit extrapolation (see the right panel of Fig. 2). Nevertheless, in terms of the total error, our determination is among the most precise ones.

Having a parametrisation of the baryon masses in terms of pseudoscalar masses and relating these pseudoscalar masses to quark masses allows us to determine the so-called σ-terms for all the octet and decuplet baryons. The σ-terms are defined as the scalar singlet matrix elements

$$\sigma_q^B = \langle B|\bar{q}q|B\rangle = m_q \frac{\partial m_B}{\partial m_q}, \quad \sigma_{\pi B} = \sigma_u^B + \sigma_d^B = m_\ell \frac{\partial m_B}{\partial m_\ell}, \tag{1}$$

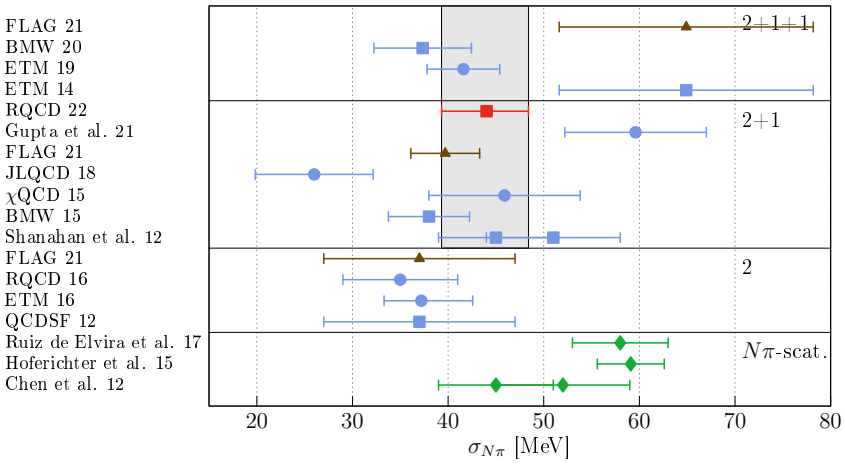


Figure 5. Comparison of our determination of $\sigma_{\pi N}$ (red square and shaded grey region) to recent determinations of this quantity from pion-nucleon scattering data (Chen *et al.* 12[22], Hoferichter *et al.* 15[23], Ruiz de Elvira *et al.* 17[24]), lattice simulations with $N_f = 2$ flavours (QCDSF 12[25], ETM 16[26], RQCD 16[27]), $N_f = 2 + 1$ flavours (Shanahan *et al.* 12[28], BMW 15[29], χQCD 15[30], JLQCD 18[31], Gupta *et al.* 21[32]) and $N_f = 2 + 1 + 1$ flavours (ETM 14[33], ETM 19[34], BMW 20[35]). Also shown as green triangles are the averages of the FLAG Review 2021[1]. Circles correspond to direct determinations, squares to fits to the quark mass dependence of the nucleon mass.

where $m_\ell = m_u = m_d$, $q \in \{u, d, s\}$, $B \in \{N, \Lambda, \Sigma, \Xi\}$. These are proportional to the coupling of the baryon either directly to a scalar dark matter particle or indirectly through exchange of the Higgs. On a theoretical level, these are the direct contributions of the quark masses (generated via the Higgs mechanism) to the nucleon mass, the remainder, which will not vanish as $m_q \to 0$, being due to kinetic and gluon contributions as well as from the broken scale invariance (trace anomaly). In Fig. 5 we compare our determination of this important quantity with literature values.

4 Conclusions

We have computed baryonic observables on a large set of gauge ensembles, allowing us to take data driven continuum, infinite volume and physical quark mass limits. Here we restricted the discussion to the spectrum, scale setting and σ-terms. Simultaneous analyses within the framework of SU(3) BChPT with baryon structure observables are ongoing[7,12,36,37,2] and additional mesonic quantities are also being investigated[38].

Acknowledgements

We acknowledge support through BMBF grant agreement no. 05P18WRFP1, the EU's Horizon 2020 research and innovation programme under grant agreement nos. 813942 (ITN EuroPLEx) and 824093 (STRONG 2020). This work was also supported by DFG

grant NFDI 39/1 (PUNCH4NFDI). We thank all CLS and RQCD colleagues for discussions and the joint production of the gauge ensembles.

We gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on JUWELS[39] and on the Booster partition of JURECA[40] at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the BMBF and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF). Additional computations were carried out on the QPACE 3 Xeon Phi cluster of SFB/TRR-55 hosted at JSC and the Regensburg Athene 2 cluster. The authors also thank the JSC for their support and for providing services and computing time on the HDF Cloud cluster[41] at JSC, funded via the Helmholtz Data Federation (HDF) programme.

Most of the ensembles were generated using OPENQCD[42]. For the computation of hadronic two- and three-point functions we used a modified version of the CHROMA[43] software package along with the LIBHADRONANALYSIS library and our algebraic multigrid solver[44] implementation[45,46].

References

1. Y. Aoki et al., "FLAG Review 2021", 2111.09849, 11 2021.
2. G. S. Bali, S. Collins, W. Söldner, and S. Weishäupl, *Leading order mesonic and baryonic SU(3) low energy constants from Nf=3 lattice QCD*, Phys. Rev. D, **105**, no. 5, 054516, 2022.
3. W. Bietenholz et al., *Tuning the strange quark mass in lattice simulations*, Phys. Lett. B, **690**, 436, 2010.
4. M. Bruno et al., *Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions*, JHEP, **02**, 043, 2015.
5. M. Bruno, T. Korzec, and S. Schaefer, *Setting the scale for the CLS $2 + 1$ flavor ensembles*, Phys. Rev. D, **95**, no. 7, 074504, 2017.
6. G. S. Bali, E. E. Scholz, J. Simeth, and W. Söldner, *Lattice simulations with $N_f = 2 + 1$ improved Wilson fermions at a fixed strange quark mass*, Phys. Rev. D, **94**, no. 7, 074501, 2016.
7. G. S. Bali, S. Collins, P. Korcyl, R. Rödl, S. Weishäupl, and Thomas Wurm, *Hyperon couplings from $N_f = 2 + 1$ lattice QCD*, PoS, **LATTICE2019**, 099, 2019.
8. M. Lüscher and S. Schaefer, *Lattice QCD without topology barriers*, JHEP, **07**, 036, 2011.
9. S. Schaefer, R. Sommer, and F. Virotta, *Critical slowing down and error analysis in Lattice QCD simulations*, Nucl. Phys. B, **845**, 93, 2011.
10. P. J. Ellis and K. Torikoshi, *Baryon masses in chiral perturbation theory with infrared regularization*, Phys. Rev. C, **61**, 015205, 2000.
11. J. Gegelia and G. S. Japaridze, *Matching heavy particle approach to relativistic theory*, Phys. Rev. D, **60**, 114038, 1999.
12. P. L. Jones Petrak, G. Bali, S. Collins, J. Heitger, D. Jenkins, S. Weishäupl, and T. Wurm, *Towards the determination of sigma terms for the baryon octet on $N_{\mathrm{f}} = 2 + 1$ CLS ensembles*, in: 38th International Symposium on Lattice Field Theory, 12 2021.

13. B. Straßberger et al., *Scale Setting for CLS 2+1 Simulations*, in: 38th International Symposium on Lattice Field Theory, 12 2021.
14. V. G. Bornyakov et al., *Wilson flow and scale setting from lattice QCD*, 8 2015, arXiv:1508.05916.
15. T. Blum et al., *Domain wall QCD with physical quark masses*, Phys. Rev. D, **93**, no. 7, 074505, 2016.
16. S. Borsányi et al., *High-precision scale setting in Lattice QCD*, JHEP, **09**, 010, 2012.
17. C. Alexandrou et al., *Ratio of kaon and pion leptonic decay constants with Nf=2+1+1 Wilson-clover twisted-mass fermions*, Phys. Rev. D, **104**, no. 7, 074520, 2021.
18. N. Miller et al., *Scale setting the Möbius domain wall fermion on gradient-flowed HISQ action using the omega baryon mass and the gradient-flow scales t_0 and w_0*, Phys. Rev. D, **103**, no. 5, 054511, 2021.
19. A. Bazavov et al., *Gradient flow and scale setting on MILC HISQ ensembles*, Phys. Rev. D, **93**, no. 9, 094510, 2016.
20. R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile, *Vus from pi and K decay constants in full lattice QCD with physical u, d, s and c quarks*, Phys. Rev. D, **88**, 074504, 2013.
21. M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, JHEP, **08**, 071, 2010, [Erratum: JHEP **03** (2014) 092].
22. Y.-H. Chen, D.-L. Yao, and H. Q. Zheng, *Analyses of pion-nucleon elastic scattering amplitudes up to $O(p^4)$ in extended-on-mass-shell subtraction scheme*, Phys. Rev. D, **87**, 054019, 2013.
23. 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.
24. 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**, no. 2, 024001, 2018.
25. G. S. Bali et al., *Nucleon mass and sigma term from lattice QCD with two light fermion flavors*, Nucl. Phys. B, **866**, 1, 2013.
26. A. Abdel-Rehim, C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, Ch. Kallidonis, G. Koutsou, and A. Vaquero Aviles-Casco, *Direct Evaluation of the Quark Content of Nucleons from Lattice QCD at the Physical Point*, Phys. Rev. Lett., **116**, no. 25, 252001, 2016.
27. G. S. Bali, S. Collins, D. Richtmann, A. Schäfer, W. Söldner, and A. Sternbeck, *Direct determinations of the nucleon and pion σ terms at nearly physical quark masses*, Phys. Rev. D, **93**, no. 9, 094504, 2016.
28. P. E. Shanahan, A. W. Thomas, and R. D. Young, *Sigma terms from an SU(3) chiral extrapolation*, Phys. Rev. D, **87**, 074503, 2013.
29. S. Dürr et al., *Lattice computation of the nucleon scalar quark contents at the physical point*, Phys. Rev. Lett., **116**, no. 17, 172001, 2016.
30. Y.-B. Yang, A. Alexandru, T. Draper, J. Liang, and K.-F. Liu, *πN and strangeness sigma terms at the physical point with chiral fermions*, Phys. Rev. D, **94**, no. 5, 054503, 2016.
31. N. Yamanaka, S. Hashimoto, T. Kaneko, and H. Ohki, *Nucleon charges with dynamical overlap fermions*, Phys. Rev. D, **98**, no. 5, 054516, 2018.

32. R. Gupta, S. Park, M. Hoferichter, E. Mereghetti, B. Yoon, and T. Bhattacharya, *Pion–Nucleon Sigma Term from Lattice QCD*, Phys. Rev. Lett., **127**, no. 24, 242002, 2021.
33. C. Alexandrou, V. Drach, K. Jansen, C. Kallidonis, and G. Koutsou, *Baryon spectrum with $N_f = 2 + 1 + 1$ twisted mass fermions*, Phys. Rev. D, **90**, no. 7, 074501, 2014.
34. C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, and A. Vaquero Aviles-Casco, *Nucleon axial, tensor, and scalar charges and σ-terms in lattice QCD*, Phys. Rev. D, **102**, no. 5, 054517, 2020.
35. S. Borsanyi, Z. Fodor, C. Hoelbling, L. Lellouch, K. K. Szabo, C. Torrero, and L. Varnhorst, *Ab-initio calculation of the proton and the neutron's scalar couplings for new physics searches*, 7 2020, arXiv:2007.03319.
36. G. S. Bali, L. Barca, S. Collins, M. Gruber, M. Löffler, A. Schäfer, W. Söldner, P. Wein, S. Weishäupl, and T. Wurm, *Nucleon axial structure from lattice QCD*, JHEP, **05**, 126, 2020.
37. D. Jenkins, S. Collins, and G. Bali, *Nucleon Charges and Sigma Terms from $N_f = 2 + 1$ QCD*, in: 38th International Symposium on Lattice Field Theory, 11 2021.
38. G. S. Bali, V. Braun, S. Collins, A. Schäfer, and J. Simeth, *Masses and decay constants of the η and η' mesons from lattice QCD*, JHEP, **08**, 137, 2021.
39. Jülich Supercomputing Centre, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, J. of large-scale research facilities, **5**, A135, 2019.
40. Jülich Supercomputing Centre, *JURECA: Modular supercomputer at Jülich Supercomputing Centre*, J. of large-scale research facilities, **4**, A132, 2018.
41. Jülich Supercomputing Centre, *HDF Cloud – Helmholtz Data Federation Cloud Resources at the Jülich Supercomputing Centre*, J. of large-scale research facilities, **5**, A137, 2019.
42. M. Lüscher and S. Schaefer, *Lattice QCD with open boundary conditions and twisted-mass reweighting*, Comput. Phys. Commun., **184**, 519, 2013.
43. R. G. Edwards and B. Joó, *The Chroma software system for Lattice QCD*, Nucl. Phys. B Proc. Suppl., **140**, 832, 2005.
44. 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, 2014.
45. S. Heybrock, M. Rottmann, P. Georg, and T. Wettig, *Adaptive algebraic multigrid on SIMD architectures*, PoS, **LATTICE2015**, 036, 2016.
46. P. Georg, D. Richtmann, and T. Wettig, *DD-αAMG on QPACE 3*, in: 35th International Symposium on Lattice Field Theory (Lattice 2017) Granada, Spain, June 18–24, 2017.

Towards Reaching Physical Conditions: Lattice QCD Results for Hadron Structure and Quark Masses

Karl Jansen
for the Extended Twisted Mass Collaboration

Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, 15738 Zeuthen, Germany
E-mail: karl.jansen@desy.de

In this project we have used three gauge ensembles with two mass degenerate light, a strange, and a charm quark (N_f=2+1+1) all tuned close to their physical values. Having three values of the lattice spacing available allowed us for the first time to obtain results in the continuum limit using only physical point ensembles. Results include the flavour decomposition of the nucleon axial form factors, and the calculation of the light, strange, and charm quark masses.

1 Introduction

The substantial progress in developing highly improved algorithms in combination with new and faster generations of supercomputers has allowed us to study lattice QCD in essential physical conditions. In particular, in this contribution we will report about results on nucleon structure quantities obtained on a set of N_f=2+1+1 twisted mass fermion ensembles at the physical point with three values of the lattice spacing a and two lattice volumes V for the coarsest lattice spacing. Here, N_f=2+1+1 refers to a situation with two light, mass degenerate up and down quarks as well as a strange and a charm quark with all masses of these quarks being at or very close to their physical values.

After tuning simulation parameters for the generation of twisted mass fermion N_f=2+1+1 ensembles we have now gauge field configurations at three values of the lattice spacing, a=0.080, 0.069, and 0.057 fm. The volume of these ensembles in physical units is $L^3 \simeq$ (5.1-5.5 fm)^3, while a fourth ensemble is available at a larger volume of $L^3 \simeq$ (7.7 fm)^3 and a=0.080 fm. We note that these are the first ensembles using Wilson-type improved fermions to be generated at the physical point spanning three values of a.

The parameters of our simulations are summarised in Tab. 1, where we also include the total number of Molecular Dynamics Updates (MDUs) carried out. A detailed discussion of the techniques employed in the simulations as applied to the generation of the gauge field configurations is documented in Ref. 1, while details on the simulations for all ensembles are provided in Ref. 2.

2 Ensemble Generation and Physical Observables

The twisted mass action that is being employed is especially suited for nucleon structure calculations, since it provides for automatic improvement of discretisation effects at maximal twist which is achieved by requiring the Partially Conserved Axial Current mass ($m_{\rm PCAC}$) to be zero within a margin of error of 10 %. It is, thus, important to monitor

name	MDUs	$m_\pi L$	a [fm]	m_N/m_π	m_π [GeV]	m_N [GeV]	L [fm]
cB64	3200	3.62	0.0798(3)	6.75(3)	0.1398(5)	0.945(5)	5.11(2)
cB96	2500	5.43	0.0798(3)	6.77(5)	0.1397(5)	0.946(7)	7.66(3)
cC80	3200	3.78	0.0686(2)	6.92(2)	0.1359(4)	0.939(4)	5.49(2)
cD96	2500	3.90	0.0569(3)	6.70(3)	0.1410(7)	0.944(6)	5.46(3)

Table 1. Physical point N_f=2+1+1 ensembles used in this contribution. The columns, from left to right, give the abbreviated ensemble name, the number of molecular dynamics units (MDUs) available, the pion mass m_π times the linear extent of the box, L, $m_\pi L$, which is a measure of volume effects, the lattice spacing in fm as determined from the nucleon mass[3], the ratio of the nucleon to pion mass, the nucleon mass and pion mass in physical units, and the lattice box length in fm.

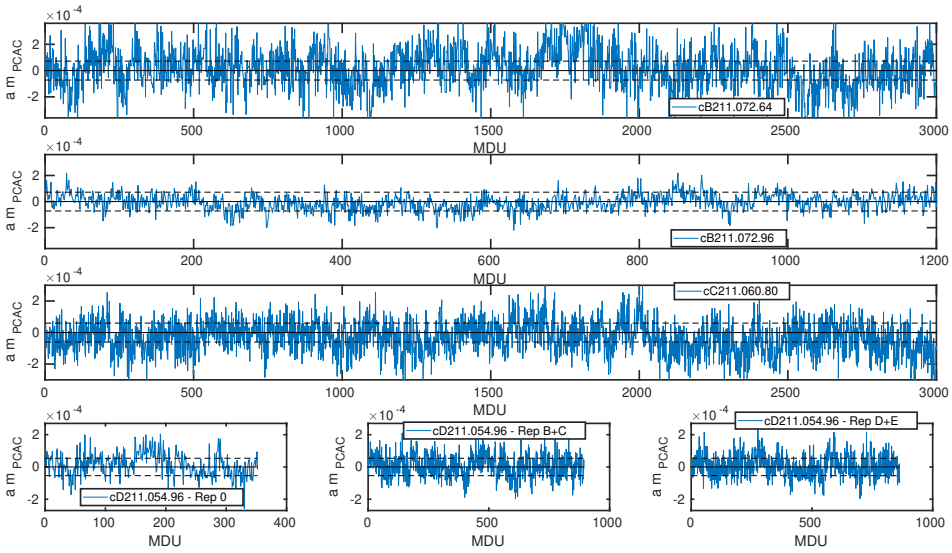


Figure 1. m_{PCAC} for the cB64, cB96, and cC80 ensembles (top, second, and third rows respectively), as well as for the cD96 ensemble (last row) using five replica chains. See Tab. 1 for information on the gauge field ensembles used.

m_{PCAC} during simulations and to check, whether this criterion is met. We show the Monte Carlo time evolution of the PCAC mass for the ensembles employed here in Fig. 1.

On the ensembles listed in Tab. 1 several observables were measured in order to monitor the progress of the Markov Chain. One key observable is the topological charge Q, which is important to monitor possible long autocorrelation times in the Monte Carlo simulation. In particular, at small lattice spacings around $a \simeq 0.05$ fm, it is important to ensure that the topological charge is indeed changing frequently, and not frozen for a long sequence of trajectories into a single sector, which would lead to long autocorrelation times in observables, a phenomenon known as "critical slowing down".

In Fig. 2, we compare the topological charge measured for the cD96 ensemble to

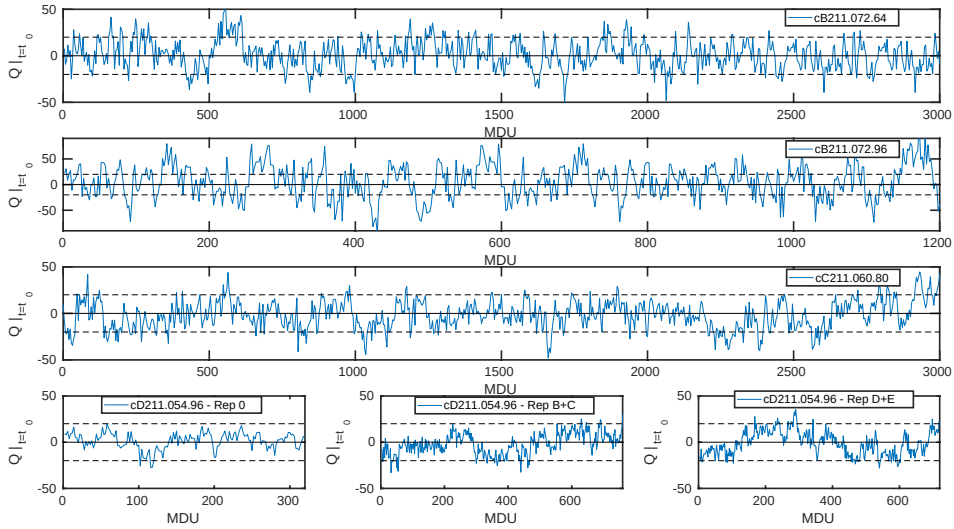


Figure 2. Topological charge Q measured for the cB64, cB96, and cC80 ensembles (top, second, and third rows respectively), as well as for the cD96 ensemble (last row) simulated during the past year using five replica chains.

the one measured using the other ensembles that have larger values of a. For the cD96 ensemble, which is shown in the last row, we have split the simulation into five replica runs that can run in parallel, thus improving throughput. We see that the fluctuations in the topological charge are similar for the three values of a, an indication that tunnelling is still satisfactory for the cD96 ensemble, thus demonstrating that our simulations are safe.

We consider the availability of ensembles with three values of a at the physical point to be milestone for our Extended Twisted Mass (ETM) collaboration in which this project is embedded. For the study of nucleon structure this set of ensembles allows for an analysis of cut-off and volume effects, directly at the physical pion mass. In Fig. 3, we show results for the nucleon mass obtained on ensembles that include those with pion mass larger than the physical value. We use continuum SU(2) chiral perturbation theory to next-to-leading order given by

$$m_N(m_\pi) = m_N^0 - 4c_1 m_\pi^2 - \frac{3g_A^2}{16\pi f_\pi^2} m_\pi^3 \tag{1}$$

to fit the data, where the four values of the lattice spacing a are determined as fit parameters, as explained in Ref. 3. The extracted values of a are listed in Tab. 1, where we also give the ratio of m_N/m_π confirming that the ensembles are indeed at the physical point.

2.1 Calculation of the Light, Strange, and Charm Quark Masses

In addition to determining nucleon observables, we used the ensembles at three lattice spacings for a calculation of the up and down, the strange, and the charm quark masses. Quark masses are essential inputs of the Standard Model (SM) and play a crucial role in the

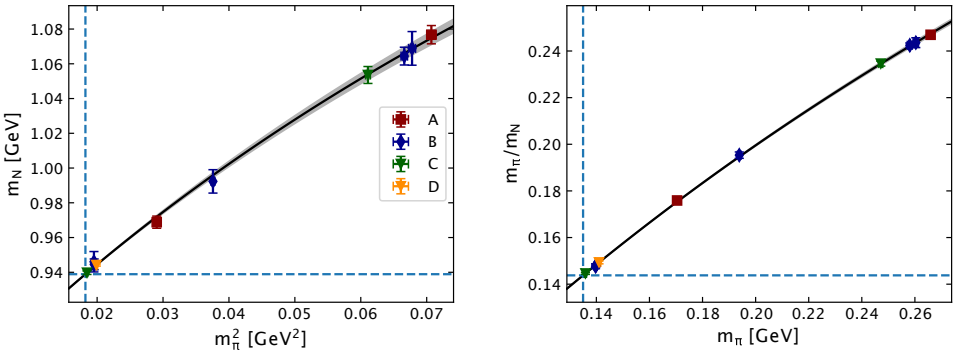


Figure 3. The nucleon mass (left) and ratio of pion to nucleon mass (right) as a function of the pion mass squared (left) and pion mass (right), respectively. We collect values using ensembles with lattice spacing a=0.093 fm denoted as A-ensembles that do not include an ensemble at the physical point (red squares), 0.080 fm denoted as B (blue diamonds), 0.067 fm (green triangles) denoted as C, and 0.057 fm denoted as D (yellow triangles). The grey bands are fits using SU(2) chiral perturbation theory to one-loop[4]. The dashed lines show the physical point.

	m_{ud} [MeV]	m_s [MeV]	m_c [MeV]	m_s/m_{ud}	m_c/m_s
Meson	$3.689(80)(66)$	$101.0(1.9)(1.4)$	$1039(15)(8)$	$27.30(24)(14)$	$11.43(9)(10)$
Baryon	$3.608(58)(^{+32}_{-19})$	$94.9(2.4)(^{+4.1}_{-1.0})$	$1030(21)(^{+22}_{-5})$	$26.30(61)(^{+1.17}_{-0.33})$	$12.04(31)(^{+58}_{-15})$
Average	$3.636(66)(^{+60}_{-57})$	$98.7(2.4)(^{+4.0}_{-3.2})$	$1036(17)(^{+15}_{-8})$	$27.17(32)(^{+56}_{-38})$	$11.48(12)(^{+25}_{-19})$

Table 2. The renormalised quark masses determined in the meson sector (first row) and baryon sector (second row) in the $\overline{\text{MS}}$ scheme. In the third row we give the average over the values obtained in the meson and baryon sectors. The light quark mass, m_{ud} (second column), and the strange quark mass, m_s (third column), are given at 2 GeV, while the charm quark mass, m_c (fourth column), is given at 3 GeV. In the fifth and sixth columns we give the ratios m_s/m_{ud} and m_c/m_s, respectively. Details regarding the error analysis and renormalisation are provided in Ref. 3.

phenomenological description of the plethora of complex phenomena governed by strong nuclear forces. In our calculation[3], ensembles with three values of the lattice spacing and spanning pion masses in the range from about 135 to 350 MeV were analysed, enabling us to perform a combined chiral and continuum extrapolation. Systematic effects were studied using two sets of observables to set the lattice spacing, one from the meson sector and one from the baryon sector. In the former case we used the pion mass, the pion decay constant, and the kaon and D-meson masses. In the latter case, instead, the masses of the pion and nucleon are employed to set the scale and fix the average up/down quark mass, while the Ω and the Λ_c masses determine the strange and charm quark masses, respectively. We tabulate the results obtained in Tab. 2 and compare to our previous results and the ones given in the FLAG report in Fig. 4.

More details of the analysis are reported in Ref. 3, where we also provide discussions regarding the error analysis and the computation of the renormalisation factor needed for extracting the renormalised quark masses.

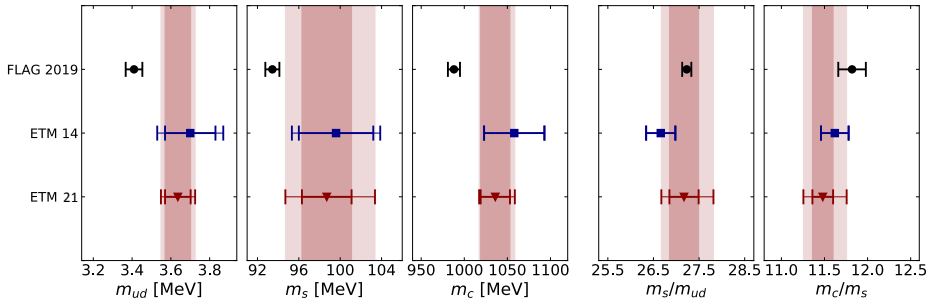


Figure 4. Comparison of our results for the quark masses, averaged between the values determined in the meson and baryon sectors (red triangles) with the values obtained using twisted mass fermions in Ref. 5 (blue squares) and the N_f=2+1+1 averages given in the 2019 FLAG report[6] (black circles). The inner error bar is the statistical error only, and the outer is the total error, obtained by summing in quadrature the statistical and the systematic errors.

2.2 Nucleon Charges and Moments of PDFs in Lattice QCD

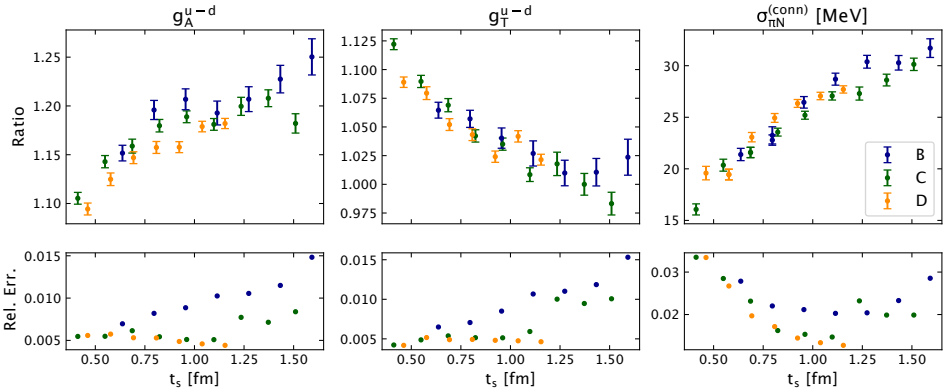


Figure 5. Nucleon isovector axial (g_A^{u-d}) (left) and tensor (g_T^{u-d}) (centre) charge and connected contribution to the light σ-term ($\sigma_{\pi N}$, right) measured using the cB64 (blue), cC80 (green), and cD96 (yellow) ensembles for multiple values of t_s.

Nucleon charges and moments of Parton Distribution Functions (PDFs) are obtained by computing the nucleon matrix elements ($\mathcal{M}$), that are extracted by taking an appropriate ratio of three- to two-point functions:

$$R(t_s, t_{\rm ins}) = \frac{G_{\rm 3pt}(t_s, t_{\rm ins})}{G_{\rm 2pt}(t_s)} \xrightarrow[\Delta E t_{\rm ins} \gg 1]{\Delta E(t_s - t_{\rm ins}) \gg 1} \mathcal{M} \tag{2}$$

where $G_{\rm 3pt}(t_s, t_{\rm ins})$ is the three-point correlation function at sink time t_s and current insertion time $t_{\rm ins}$, $G_{\rm 2pt}(t_s)$ is the two-point function, and ΔE is the energy gap between the

ground and first excited state. To extract the ground state we use a multi-state fit approach and several values of t_s increasing the statistics as we increase t_s to keep the statistical error approximately constant. This is essential to reliably assess excited states[7,8]. Results for the axial, tensor, and scalar charges, as well as the nucleon σ-terms, using the `cB211.072.64` ensemble have been published in Refs. 9, 10.

2.3 Nucleon Form Factors

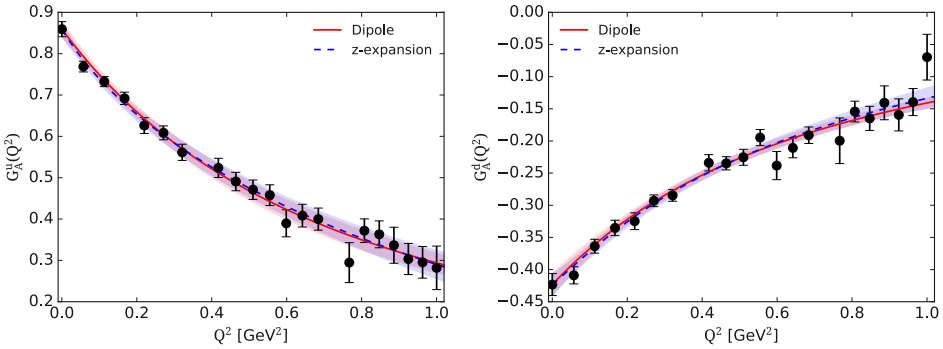


Figure 6. The up (left) and down (right) axial form factor obtained using the `cB64` ensemble. We use a dipole form (solid red curve) as well as the so called z-expansion (dashed blue curve) to fit the data, as explained in Ref. 14.

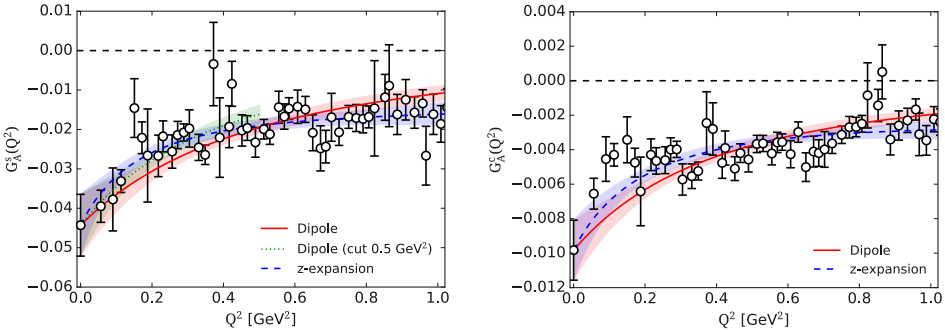


Figure 7. The strange (left) and charm (right) axial form factor obtained using the `cB64` ensemble. The rest of the notation is as in Fig. 6.

Results obtained using the `cB211.072.64` ensemble include the electromagnetic and axial form factors of the nucleon[11–14], the first such computation using an N_f=2+1+1

ensemble at the physical point. For the analysis GCS resources were used for the calculation of the disconnected contributions to the axial form factors and, therefore, obtain the individual quark flavour contributions to these quantities. The disconnected quark loops required are difficult to compute due to large statistical fluctuations, but due to the GPU resources such as JUWELS-Booster it became possible to determine them in a quantitatively significant way. Representative results of this analysis are shown in Fig. 6 for the light-quark contributions to the axial form factor and in Fig. 7 for the strange- and charm-quark contributions. Additional results computed include the SU(3) singlet and octet flavour combinations, as well as the corresponding induced pseudoscalar form factor, as reported in detail in Ref. 14.

3 Conclusion

In this contribution we have reported on a number of observables which were computed on ensembles at three values of the lattice spacing and in a situation where the light up and down, the strange and the charm quarks assume their physical values. To reach such a situation needed substantial algorithmic and methodological developments, which we could achieve in the last years. Still, without the employment of ever increasing super computer power made available through new architectures it would not have been possible to reach such a goal.

On the generated ensembles a number of physical quantities could be computed, comprising the nucleon and pion mass, renormalised quark masses, nucleon charges and moments of parton distribution functions as well as nucleon form factors. Having these quantities now available in physical conditions and at three values of the lattice spacing is a major breakthrough which eventually allows to make a direct contact to experiments or phenomenological analyses.

Acknowledgements

This work is embedded in the broad physics program of the Extended Twisted Mass Collaboration (ETMC). The results reported upon here would not have been possible without the great collaborative effort of ETMC which also includes the simulations of the here used gauge field ensembles. We gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for providing computing time.

References

1. C. Alexandrou et al., *Simulating twisted mass fermions at physical light, strange and charm quark masses*, Phys. Rev. D, **98**, no. 5, 054518, 2018.
2. J. Finkenrath et al., *Twisted mass gauge ensembles at physical values of the light, strange and charm quark masses*, in: 38th International Symposium on Lattice Field Theory, 1 2022.
3. C. Alexandrou et al., *Quark masses using twisted-mass fermion gauge ensembles*, Phys. Rev. D, **104**, no. 7, 074515, 2021.

4. J. Gasser, M. E. Sainio, and A. Švarc, *Nucleons with chiral loops*, Nuclear Physics B, **307**, no. 4, 779–853, 1988.
5. N. Carrasco et al., *Up, down, strange and charm quark masses with N_f = 2+1+1 twisted mass lattice QCD*, Nucl. Phys. B, **887**, 19–68, 2014.
6. S. Aoki et al., *FLAG Review 2019: Flavour Lattice Averaging Group (FLAG)*, Eur. Phys. J. C, **80**, no. 2, 113, 2020.
7. C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, C. Kallidonis, G. Koutsou, and A. Vaquero Aviles-Casco, *Nucleon axial form factors using N_f = 2 twisted mass fermions with a physical value of the pion mass*, Phys. Rev. D, **96**, no. 5, 054507, 2017.
8. G. von Hippel, T. D. Rae, E. Shintani, and H. Wittig, *Nucleon matrix elements from lattice QCD with all-mode-averaging and a domain-decomposed solver: an exploratory study*, Nucl. Phys. B, **914**, 138–159, 2017.
9. C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, and A. Vaquero Aviles-Casco, *Nucleon axial, tensor, and scalar charges and σ-terms in lattice QCD*, Phys. Rev. D, **102**, no. 5, 054517, 2020.
10. C. Alexandrou, *Recent progress on the study of nucleon structure from lattice QCD and future perspectives*, SciPost Phys. Proc., **3**, 015, 2020.
11. C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, and G. Koutsou, *Nucleon strange electromagnetic form factors*, Phys. Rev. D, **101**, no. 3, 031501, 2020.
12. C. Alexandrou et al., *Nucleon axial and pseudoscalar form factors from lattice QCD at the physical point*, 11 2020.
13. C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, and A. Vaquero, *Nucleon form factors from N_f=2+1+1 twisted mass QCD at the physical point*, in: 38th International Symposium on Lattice Field Theory, 12 2021.
14. C. Alexandrou, S. Bacchio, M. Constantinou, K. Hadjiyiannakou, K. Jansen, and G. Koutsou, *Quark flavor decomposition of the nucleon axial form factors*, Phys. Rev. D, **104**, 074503, 2021.

Materials Science

Materials Science

Gustav Bihlmayer

Peter Grünberg Institut and Institute for Advanced Simulation (PGI-1/IAS-1),
Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: g.bihlmayer@fz-juelich.de

Being at the interface between theoretical condensed matter physics and more applied fields of science and engineering, computational materials science tries to bridge the gap between fundamental physical principles and the understanding of technologically relevant phenomena. This chapter provides three such examples: improving the performance of catalysts, the quality of thin-film growth, and the minimisation of friction between surfaces by numerical simulations. The computational investigation of such topics faces the challenge to explore either a large configuration space (e.g. of all possible reaction pathways in a catalytic process on non-uniform surfaces) or to simulate the dynamics of a huge ensemble of particles under different external conditions (e.g. the friction of materials separated by a lubricant). Despite the growth of computational resources, as provided by the John von Neumann Institute for Computing, the solution of quantum mechanical equations is still out of reach for these problems. However, due to the intrinsically quantum mechanical nature of these problems, the application of semi-empirical principles is not a promising strategy. New concepts must be applied in materials science to bridge the scales.

In the first contribution by Xu *et al.*, the authors explore transition metal oxide surfaces and their catalytic properties for the oxygen evolution reaction as well as their interaction with more complex organic molecules. The authors apply data-efficient machine-learning models to predict properties such as adsorption enthalpies of different oxygen-containing intermediates of the reaction. The so-called SISSO method is employed to create scaling relations between certain "features" computed by density functional theory (DFT), and the adsorption enthalpy of the reaction intermediates. The method finds a non-linear relation between the features, in this case the adsorption enthalpy of oxygen, the electronegativity of the surface atoms, or some parameters of its d-bands, and this target. The results can be tested against direct DFT calculations of the targeted adsorption enthalpies, and the authors conclude that the root-mean-square errors are less than the inherent DFT errors. Moreover, such a machine-learning approach is not only computationally efficient, it also allows insight into the parameters that are relevant for the catalytic reaction itself. While the parameters here are all scalar quantities, the procedure must be extended for more complex molecules, where the orientation with respect to the surface also matters. In a second example, the authors show a graph-based machine learning model that can be applied to these cases.

Another example of a complex process, relevant for application, is the growth of thin films from precursor molecules that impinge on a heated substrate surface in a method called chemical vapour deposition (CVD). Often used to produce graphene, this method can also be applied to fabricate hexagonal boron nitride (hBN) monolayers that are proposed as dielectrics for graphene devices. However, a three-dimensional growth of hBN is

often observed, and Dabrowski *et al.* used DFT calculations to simulate the growth process on germanium surfaces. Again, the number of possible processes that lead from the supplied borazine molecules to the hBN layers is huge, and many processes in the reaction chamber must be taken into account to explain the observed growth. In the second example given in their work about the structural and compositional changes in the surface of the wide-bandgap semiconductor $ZnGa_2O_4$ the true structure could be identified only in close collaboration with experiment. Under realistic conditions many processes happen far from equilibrium, where a naive ground-state search cannot be applied.

This is certainly also true in tribology where – despite their importance in everyday life – even the most fundamental mechanisms that determine friction are still a matter of debate. In the work of Reichenbach *et al.*, the dry friction between two diamond surfaces, terminated by hydrogen or fluorine, is considered where the choice of the termination crucially influences the behaviour. It was not clear whether this is a steric effect or caused by the different electronegativity of the terminating atoms or the chemical bonding, etc. Simulations with more than 100.000 atoms could clarify the situation, but such system sizes are beyond the realm of DFT, and classical molecular dynamics (MD) calculations are necessary. Nevertheless, extensive DFT calculations allow the creation of reliable force fields that can be employed in a specific setting. A similar problem is encountered in the study of surfaces separated by a lubricant, where MD simulations of 75.000 atoms must be performed for different sliding conditions over tens of nanoseconds. In this second case study, the authors show how ordering phenomena in the lubricant influence the evolution of friction in such a system.

Common to all the examples from materials science discussed here are (i) their basis in density functional theory, sometimes complemented by techniques to improve the description of certain material classes like oxides and (ii) the combination with other methods to get closer to the targeted reality. This can be machine learning models to sample a large configuration space, scale bridging models to allow simulating systems that are large enough to reflect the actual processes, or working hand-in-hand with experimental methods. In all cases, the first step requires setting up a database of DFT results, something that profits from automated workflows developed in the last years and the extended computational resources at hand. In combination with the tools mentioned in point (ii), materials science becomes closer to real processes that are relevant for our society and environment.

Data-Efficient and Physics-Inspired Machine-Learning Models to Treat Complex Catalyst Materials and Reaction Networks

Wenbin Xu[1,2], Karsten Reuter[2], and Mie Andersen[3,4]

[1] Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany

[2] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

[3] Department of Physics and Astronomy - Center for Interstellar Catalysis, Aarhus University, Aarhus C, DK-8000 Denmark

[4] Aarhus Institute of Advanced Studies, Aarhus University, Aarhus C, DK-8000 Denmark
E-mail: mie@phys.au.dk

Recently developed machine learning (ML) methods hold great promise for simultaneously reducing the computational cost and increasing the accuracy in catalysis modelling, allowing us to capture more complexity, make our models more realistic and perhaps even obtain new physical insights. Here we focus on two aspects relevant for the computational screening of heterogeneous catalyst materials. One is the computational screening of transition metal (TM) oxides, which is challenging due to their more localised and intricate electronic structure as compared to TM catalysts. The other is the prediction of adsorption motifs and associated adsorption enthalpies for complex adsorbates that may exhibit mono-, bi- or higher-dentate adsorption motifs at TMs. The ML methods employed range from the compressed sensing SISSO method, which seeks descriptors in the form of analytical functions, to Gaussian Process Regression (GPR) using a customised version of the recently developed Wasserstein Weisfeiler-Lehman graph kernel.

1 Introduction

Over the last decades computational screening has evolved into an invaluable supplement to experimental efforts for the identification of new and improved catalyst materials[1,2]. Descriptors are crucial ingredients of the screening process, as they provide a means for predicting the reactivity of candidate materials at low computational cost and thus allow to cover a much larger number of materials than would be possible using costly *ab initio* methods such as density-functional theory (DFT). For TMs, with their comparably simple electronic structure, even very simple descriptors such as the energetic position of the d-band[3] or related scaling relations[4] have been highly successful. The latter connect the adsorption enthalpies of few (1-2) key intermediates with the remaining intermediates, e.g. for the oxygen evolution reaction (OER) the OH* and OOH* adsorption enthalpies have been shown to scale linearly with the O* adsorption enthalpy[5]. Furthermore, the existence of scaling relations has been shown to lead to fundamental limitations to the theoretical OER overpotentials achievable with TM oxide catalysts[6]. However, scaling relations do not always work well for complex materials. It has been shown that they can be completely broken at both TM oxides[7] and at interfaces between oxides and metals[8]. In general, the identification of simple, descriptive models for TM oxides is challenging due to their more localised and intricate electronic structure.

Another outstanding challenge in catalyst screening is the prediction of adsorption motifs and associated adsorption enthalpies for complex adsorbates that may exhibit mono-, bi- or higher-dentate adsorption motifs[9]. For example, already the identification of the energetics and adsorption motifs of adsorbates involved in ethanol synthesis at a relatively simple monometallic catalyst such as Rh(111) is a formidable task[10]. The investigation of more complex classes of materials such as TM alloys is generally out of reach due to the combinatorial explosion of possible active sites and adsorption motifs. Nevertheless, it is highly important for enabling computational screening of novel catalyst materials for a wide range of reactions that are critical to our society and the green transition toward renewable energy sources such as Fischer-Tropsch, methanol, or higher oxygenate synthesis.

Here we present two recent projects that were made possible owing to generous compute time allowances at the JUWELS supercomputer at Jülich Supercomputing Centre (JSC): the development of descriptors and refined scaling relations for doped TM oxides[11] using the compressed sensing method SISSO (sure independence screening and sparsifying operator)[12] and the prediction of adsorption motifs and associated adsorption enthalpies for complex adsorbates that may exhibit mono-, bi- or higher-dentate adsorption motifs at TMs using a customised Wasserstein Weisfeiler-Lehman (WWL) graph kernel[13] and GPR[14].

2 Methodology

2.1 DFT Calculations and Databases

All DFT calculations were carried out with the Quantum ESPRESSO code[15] employing the Bayesian error estimation functional with van der Waals correlation (BEEF-vdW)[16] and ultrasoft pseudopotentials. For the TM oxides we focused on O*, OH* and OOH* as they are established key OER intermediates. Their adsorption enthalpies were calculated at the coordinatively unsaturated (cus) sites of all five symmetry-inequivalent low-index facets of rutile IrO_2 and RuO_2 (see Fig. 1), when doping them with a wide range of TMs (Ti, Mn, Fe, Co, Ni, Cu, Zn, Mo, Ru, Ag, W, Ir).

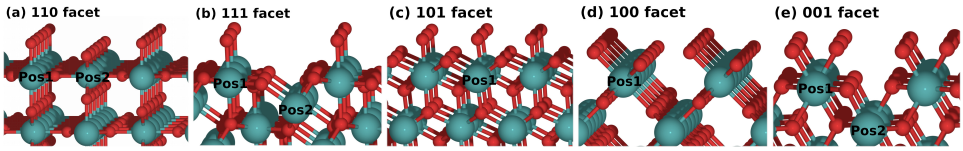


Figure 1. (a-e) Perspective side views of oxygen-rich terminations of low-index facets of rutile-structured oxides. The TM atoms (Ru or Ir) are shown in green and O atoms are shown in red. The considered substitutional doping positions are indicated. Reprinted with permission from Ref. 11. Copyright 2020 American Chemical Society.

The database of complex adsorbates included 41 different small to large adsorbates involved in ethanol synthesis on fcc(111) and fcc(211) facets of Cu, Rh, Pd, and Co. Examples of complex adsorbates are CHCO, CCHOH, CH_2CH_2O and CH_3CH_2OH. This database was constructed using an automated workflow where initial geometries of the surface/adsorbate systems were generated using the CatKit software[17].

2.2 SISSO Method

SISSO[12] is a compressed sensing-based method for identifying the key features necessary to describe and predict a property of interest (here adsorption enthalpies). SISSO makes the ansatz that the property of interest $\mathbf{P}$ can be expressed as

$$\mathbf{P} = \sum_{i}^{M} c_i \mathbf{f}_i \tag{1}$$

where c_i are free parameters (coefficients) and $\mathbf{f}_i$ are a small number of selected features, i.e., the sparse solution, which provides an analytic solution enabling the extraction of chemical insights.

The primary features are the inputs from which the predictive model is built. In Fig. 2 we show examples of different types of primary features used for TMs and TM oxides. Features are constructed as non-linear functions of the primary features in an iterative manner by applying functional/algebraic operators (e.g. addition, multiplication, exponential, logarithm, square root) to the primary features. The number of iterations performed,

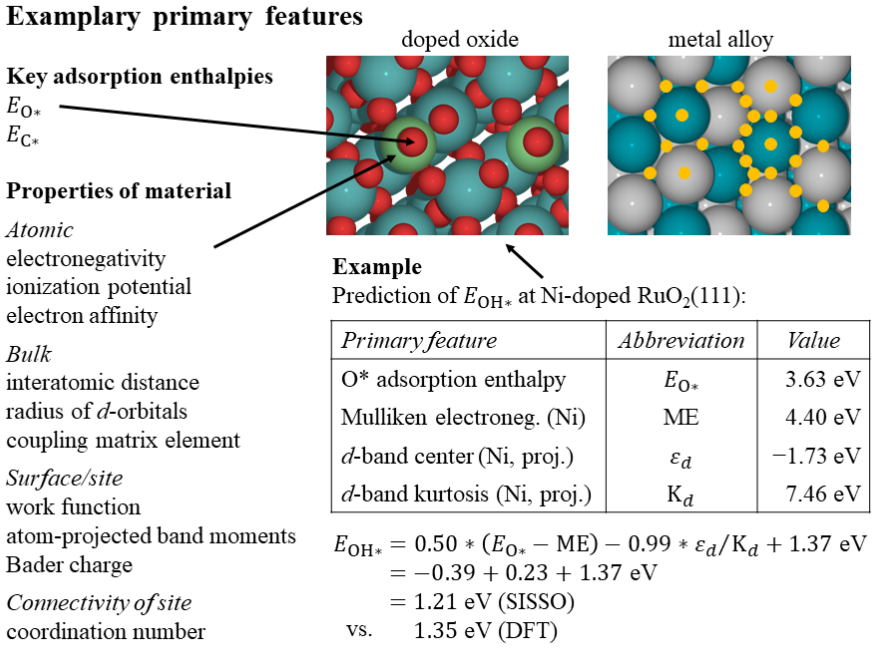


Figure 2. Examples of primary features considered in our works and example prediction using one of the simpler identified models specific for TM oxide surfaces. Symmetry-inequivalent adsorption sites (top, bridge, three-fold and four-fold coordinated sites) on the stepped TM alloy are marked in yellow. In case the adsorption site is made up of several atoms (all but top sites), the primary features are averaged over these atoms. Reprinted with permission from Ref. 18. Copyright 2021 American Chemical Society.

N (also termed the rung, Φ_N), determines the complexity and total number of the features. The size of the feature space grows rapidly in each iteration and typically reaches around 10^{12} in Φ_3 when the size of the initial (Φ_0) feature space (number of primary features employed) is about 20. Dimensionality and rung are hyperparameters of the SISSO method, which can e.g. be optimised in a grid search using cross-validation. Feature selection is typically carried out using the ℓ_0 constraint in a smaller feature subspace, the size of which is controlled by the SIS parameter.

2.3 Graph-Based Machine Learning

For complex adsorbates we rely on a graph representation of the surface/adsorbate system in order to take into account how the adsorbates bind to the surface, possibly in complex mono-, bi- or higher-dentate adsorption motifs. The developed WWL-GPR model[14], which is based on a customised version of the WWL graph kernel[13] and GPR, is shown schematically in Fig. 3. We use similar primary features as described above for the SISSO models. However, owing to the graph representation, we can now use a separate feature vector for every atom in the system (i.e. by using the primary features as node attributes), instead of using only a single feature vector containing global properties of the system (e.g. work function) or averaged properties of the system (e.g. average coordination number of the surface atoms to which the adsorbate coordinates). For comparison, we show also results using vector-based input to a GPR model together with the standard radial basis function kernel (RBF-GPR model).

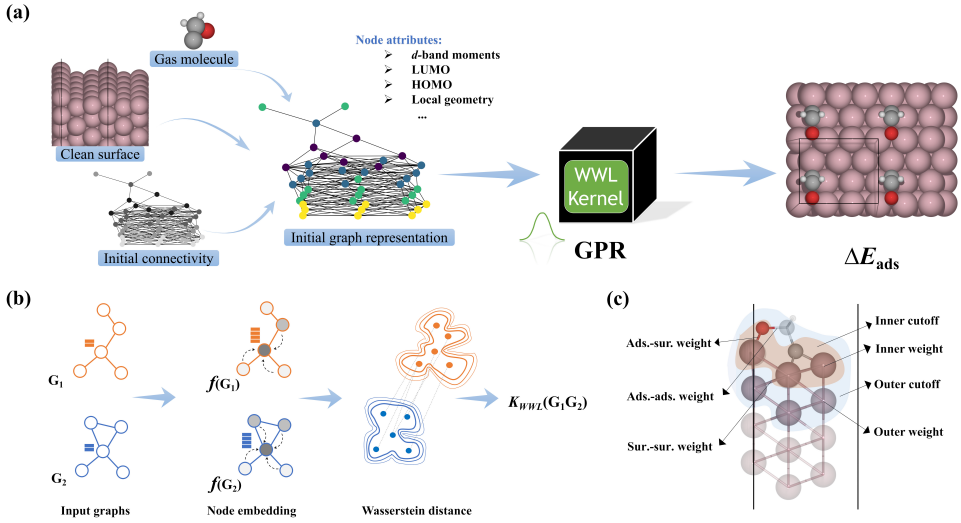


Figure 3. Schematic of the developed WWL-GPR model. (a) The adsorption enthalpy corresponding to the relaxed structure of the surface/adsorbate system is predicted from a graph representation of the initial structure. We use node attributes that are computed from the isolated molecule in the gas phase or from the adsorbate-free surface. Predictions are performed via the WWL graph kernel and GPR. (b) Illustration of the customised WWL graph kernel. (c) Hyperparameters that were introduced to customise the WWL kernel, which are motivated by general considerations in surface adsorption. Adapted from Ref. 14, Springer Nature Ltd.

3 Results

3.1 SISSO-Refined Scaling Relations for TM Oxides

Based on the established TM oxide DFT database we can construct scaling relations between the OER intermediates O*, OH* and OOH* (see our original work in Ref. 11) However, as discussed above, scaling relations are not that accurate for oxides. Indeed, we find that even the training mean absolute error for our constructed scaling relations is 0.22 eV, which is significantly higher than for corresponding oxygenate scaling relations at TMs (about 0.10 eV[19]). In this section we show how SISSO can be employed to unravel the underlying reason for this different performance of scaling relations in the two materials classes. Specifically, we elucidate the missing ingredients (primary features) in the standard scaling description by identifying SISSO-refined scaling relations. In practice, this is done by explicitly including the O* adsorption enthalpy itself among the primary features and training the SISSO descriptors based on the database of OH* and OOH* adsorption enthalpies alone. We carry out five-fold cross validation, that is, the TM oxide database is shuffled and partitioned into five equal-sized subsamples stratified by adsorbates. The training is then carried out based on four of the subsamples while retaining the fifth subsample for validation. This is repeated five times until all data points have been used once for validation.

The predictive performance of the descriptors is shown in Fig. 4 and the composition of the simpler 1D and 2D descriptors (as identified by ultimately training on the entire *OH and *OOH database) is given in Tab. 1. Note that the simplest 1D Φ_0 descriptor corresponds to the standard OH* vs. O* and OOH* vs. O* scaling relations. As seen

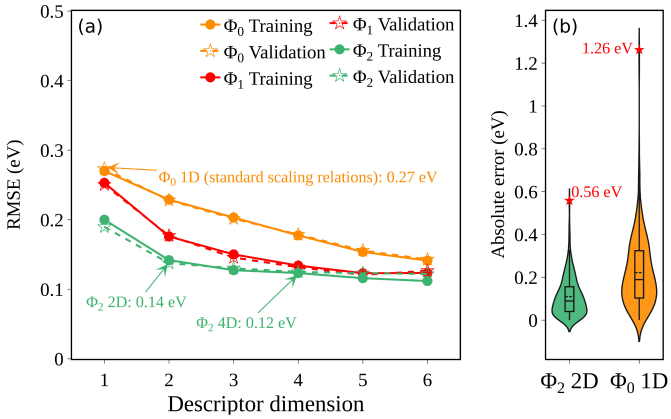


Figure 4. SISSO-refined scaling relations identified by including the O* adsorption enthalpy among the primary features and training and validating only on the OH* and OOH* adsorption enthalpies. (a) Training and validation RMSE from 5-fold cross validation (averaged over the 5 folds) for all tested SISSO hyperparameters. (b) Violin plot of the distribution of the absolute validation errors (combined from the 5 folds) for the 5D Φ_3 descriptor. The internal solid line marks the median, the internal dashed line marks the mean, the black box marks the 75% and 25% percentiles, the whiskers mark the 95% and 5% percentiles, the red star marks the maximum absolute error, and the blue region is the density plot. Reprinted with permission from Ref. 11. Copyright 2020 American Chemical Society.

Descriptor	RMSE	First term	Second term
1D Φ_0	0.27	$(E_{\mathrm{O}*})$	
2D Φ_0	0.23	$(E_{\mathrm{O}*})$	(Q_3)
1D Φ_1	0.25	$(E_{\mathrm{O}*}/\mathrm{IP})$	
2D Φ_1	0.18	$(E_{\mathrm{O}*} - \mathrm{ME})$	$(\varepsilon_d/\mathrm{K}_d)$
1D Φ_2	0.20	$(E_{\mathrm{O}*} - \mathrm{ME}) - (\varepsilon_d/\mathrm{K}_d)$	
2D Φ_2	0.14	$(E_{\mathrm{O}*} - \mathrm{ME}) \cdot (Q_1 + Q_4)$	$(\mathrm{EA} - \varepsilon_d) \cdot (\mathrm{K}_d/d_1)$

Table 1. RMSE (in eV) of various descriptors in predicting the OH* and OOH* adsorption enthalpies in the database. The 1D Φ_0 descriptor corresponds to the standard scaling relation ($E_{\mathrm{OH}*} = C_1 * E_{\mathrm{O}*} + C_0$ and analogous for $E_{\mathrm{OOH}*}$). The SISSO-refined scaling relations have the form $E_{\mathrm{OH}*} = C_1 * \mathrm{first\,term} + C_2 * \mathrm{second\,term} + C_0$. Reprinted (adapted) with permission from Ref. 11. Copyright 2020 American Chemical Society.

in Fig. 4, already the simplest refinements offer substantially improved prediction errors over the standard scaling relations with the validation root-mean-squared error, RMSE (maximum absolute error) reaching 0.14 eV (0.56 eV) for the 2D Φ_2 descriptor. The 4D Φ_2 descriptor even lowers the RMSE to 0.12 eV.

Considering the composition of the simplest 2D Φ_0 and $1-$2D Φ_1 descriptors (Tab. 1) we identify the ionisation potential (IP) or the Mulliken electronegativity (ME) of the cus atom to which the adsorbates coordinate as the most important corrections to the O* adsorption enthalpy. We interpret their presence in the correction terms as an indication that it is this physical information that is missing in the standard scaling description and propose that this is the principal reason why scaling relations do not work very well on oxides. In fact, as we will show below, it is precisely the inclusion of these SISSO-identified charge correction terms that is crucial to correctly capture promising dopants for OER catalysis.

3.2 Screening for OER Catalysts

A particularly interesting application of the above-demonstrated possibilities for low-cost and accurate prediction of adsorption enthalpies is the tailoring of doped TM oxide catalysts for the OER through high-throughput screening. From previous scaling relation-based approaches it has been established that the existence of a universal scaling relation between the critical OER intermediates OH* and OOH* leads to a constraint on the lowest possible theoretical overpotential[6]. In Fig. 5(a) we plot the here calculated DFT OOH* vs. OH* adsorption enthalpies and indicate both this universal scaling relation and the ideal scaling relation that would allow for desirable catalyst operation already at the equilibrium potential for OER of $U = 1.23$ V vs. the reversible hydrogen electrode (more precisely, this is a necessary but not sufficient condition as detailed below)[6]. Among the here studied dopants some clear trends emerge, i.e., independently of the studied facet the dopants Ru, Ir (and thus also the undoped facets), W, Mo and Ti tend to lie close to the universal scaling line, whereas the dopants Ni, Co, and Ag tend to lie closer to the ideal scaling line. The remaining dopants Cu, Fe, Mn and Zn are intermediate cases.

Within the classic volcano picture of OER established by Nørskov and coworkers[20,5,6], the ideal catalyst should not only follow the ideal scaling line, but also possess an ideal absolute O* (or OH*) adsorption enthalpy to satisfy the Sabatier principle. To get a quantitative estimate of the theoretical overpotential of all materials in our database, we

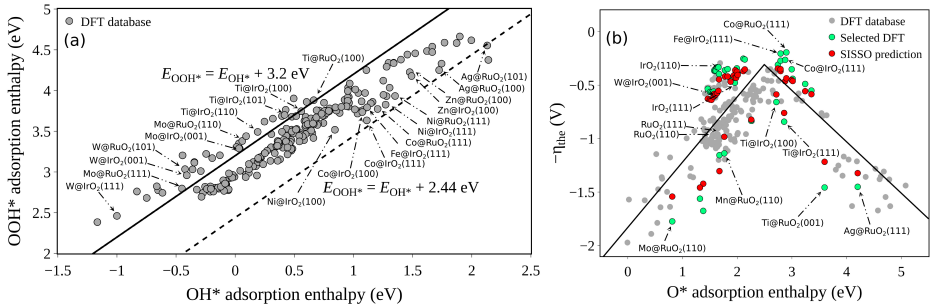


Figure 5. (a) OOH* vs. OH* adsorption enthalpies from our DFT database. The solid (dashed) lines indicate the universal (ideal) scaling relations[6]. (b) Negative theoretical overpotential η_{the} of all materials in our database as a function of the O* adsorption enthalpy. The black curve is the volcano predicted from the standard OH* vs. O* and OOH* vs. O* scaling relations. Selected DFT data points (in green) and corresponding SISSO predictions using refined scaling relations (2D Φ_2, in red) are also indicated. Reprinted with permission from Ref. 11. Copyright 2020 American Chemical Society.

carry out the corresponding thermodynamic analysis using the computational hydrogen electrode (CHE) approach[20] for the peroxide OER reaction mechanism consisting of four proton-coupled electron transfer steps[5,6]. In this analysis the theoretical overpotential is equated with the potential required to make all four reaction steps exergonic. The results are presented in Fig. 5(b). It is seen that the scatter of the points around the scaling lines in Fig. 5(a) translates into an uncertainty of about ± 0.5 V in the theoretical overpotential as compared to the known scaling-derived volcano curve with its top at the ideal $E_{\mathrm{O}*} = 2.46$ eV. Some of the DFT data points that exhibit particularly large deviations from the volcano are highlighted in green in the figure. Intriguingly, already the quite simple 2D Φ_2 SISSO-refined scaling relation (cf. Tab. 1) instead quite reliably captures those data points that are outliers to the volcano curve. Compared to the actual DFT-computed theoretical overpotential the remaining uncertainty of the SISSO-derived overpotential is on average only around the targeted 0.2 V, as visually reflected in Fig. 5(b) by the closeness of the green (DFT) and red (SISSO) data points.

The SISSO-refined scaling relations allow to differentiate between materials which, in spite of very similar O* adsorption enthalpies, may in reality show up to $\sim$1 V difference in the theoretical overpotential. As apparent from Fig.5(b) it is in particular Co and Fe that show a favourable deviation from the scaling line and lead to catalysts that exhibit lowest theoretical overpotentials. This deviation is correctly captured within SISSO, i.e. a SISSO-based computational screening study would have correctly identified these doped systems. In contrast, as their O* adsorption enthalpy does not match the "ideal" value, these systems are located to the right or left of the volcano maximum and therefore would likely have been dismissed as sub-optimal in a traditional scaling relation-based screening. As discussed above, we attribute this superior SISSO performance to its ability to appropriately capture the local charge transfer properties of these dopants. Our independent theoretical results are furthermore impressively backed up by experiments, where recent studies have in particular highlighted the favourable electrocatalytic OER activities for Co-doping of RuO_2 and IrO_2[21], Fe/Mn-doping of IrO_2[22], and Ni- and Co-doping of IrO_2[23].

We note that Mn and Ni are also present in many of the here predicted good OER catalysts, even though not exhibiting lowest theoretical overpotentials.

3.3 Complex Adsorbates at TM Surfaces

Next, we turn to a second project concerning complex adsorbates and TM catalysts. We here evaluate and compare the performances of the graph-based WWL-GPR model (i.e. employing a graph representation of the surface/adsorbate system with features as node attributes) and a GPR model with a single feature vector as input (the RBF-GPR model). Fig. 6 shows the parity plot resulting from five-fold cross validation of DFT-calculated against ML-predicted adsorption enthalpies for the combined validation set from the five folds as well as violin plots of the absolute error distributions. The results clearly show that we need the graph representation for describing and predicting adsorption enthalpies of complex adsorbates. The accuracy obtained with the WWL-GPR model is very promising for screening applications with a RMSE of 0.18 eV and a maximum absolute error of 0.91 eV.

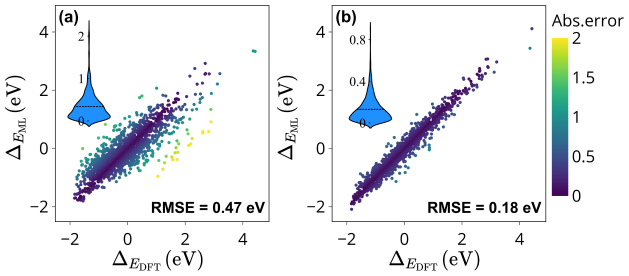


Figure 6. Parity plot of DFT-calculated vs. ML-predicted adsorption enthalpies from the combined validation set in 5-fold cross validation using (a) RBF-GPR and (b) WWL-GPR for the complex adsorbates database. The violin plots in the insets illustrate the absolute error distributions (in eV). The internal dashed line marks the mean. Adapted from Ref. 14, Springer Nature Ltd.

4 Conclusions

In this work we constructed DFT databases of adsorption enthalpies of small to large adsorbates at various types of materials relevant for heterogeneous catalysis, in particular doped TM oxides and TMs. Since DFT calculations are in general too computationally expensive for materials screening or design studies, we used the databases to train descriptors or predictive models using compressed sensing and ML techniques. In one project we used the compressed sensing method SISSO to identify descriptors for the prediction of the adsorption enthalpies of the OER intermediates O*, OH* and OOH* at the cus sites of various facets of doped IrO_2 and RuO_2 electrocatalysts[11]. The descriptors were constructed as algebraic expressions of electronic and geometric primary features. By including the

O* adsorption enthalpy into the training data, we identified so-called SISSO-refined scaling relations. Their compositions help us to identify primary features related to the local charge transfer to be the primary correction (and thereby missing ingredient) to standard scaling relations. For the showcased screening of dopants these corrections turn out as crucial to reliably identify Co and Fe in agreement with recent experimental works.

In a second project we developed the graph-based WWL-GPR model for predicting adsorption enthalpies and adsorption motifs of complex adsorbates at TMs. This is a data-efficient and physics-inspired ML approach based on a customised Wasserstein Weisfeiler-Lehman graph kernel and Gaussian Process Regression. The task that is solved is to directly predict the relaxed adsorption enthalpies corresponding to a range of plausible initial guesses of the adsorption motif based on graph representation. Thereby, for a given surface/adsorbate combination of interest, both the most stable and all meta-stable adsorption motifs as well as their associated adsorption enthalpies can be predicted. Rather than relying on graph representation alone, as is primarily being done in more data-greedy approaches implemented with deep neural networks, our WWL-GPR model is augmented with physically inspired node attributes representing local geometric or electronic information about each atom in the structure as well as surface adsorption motivated hyperparameters that highlight the role of the atoms near the active site and the surface/adsorbate interface. The model achieves an in-domain prediction of adsorption enthalpies with a RMSE of 0.18 eV, i.e. it is en par with the expected intrinsic uncertainty in DFT.

Acknowledgements

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). W.X. is grateful for support through the China Scholarship Council (CSC). M.A. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754513, the Aarhus University Research Foundation, the Danish National Research Foundation through the Center of Excellence "InterCat" (Grant agreement no.: DNRF150) and VILLUM FONDEN (grant no. 37381).

References

1. K. Tran and Z. W. Ulissi, *Active learning across intermetallics to guide discovery of electrocatalysts for CO2 reduction and H2 evolution*, Nat. Catal., **1**, no. 9, 696–703, 2018.
2. M. Andersen, S. V. Levchenko, M. Scheffler, and K. Reuter, *Beyond Scaling Relations for the Description of Catalytic Materials*, ACS Catal., **9**, no. 4, 2752–2759, 2019.
3. B. Hammer and J. K. Nørskov, *Electronic factors determining the reactivity of metal surfaces*, Surf. Sci., **343**, no. 3, 211–220, 1995.
4. F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skulason, T. Bligaard, and J. Kehlet Nørskov, *Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces*, Phys. Rev. Lett., **99**, no. 1, 016105, 2007.

5. J. Rossmeisl, Z.-W. Qu, H. Zhu, G.-J. Kroes, and J. K. Nørskov, *Electrolysis of water on oxide surfaces*, J. Electroanal. Chem., **607**, no. 1-2, 83–89, 2007.
6. I. C. Man, H.-Y. Su, F. Calle-Vallejo, H. A. Hansen, J. I. Martinez, N. G. Inoglu, J. Kitchin, T. F. Jaramillo, J. K. Nørskov, and J. Rossmeisl, *Universality in oxygen evolution electrocatalysis on oxide surfaces*, ChemCatChem, **3**, no. 7, 1159–1165, 2011.
7. A. Vojvodic, F. Calle-Vallejo, W. Guo, S. Wang, A. Toftelund, F. Studt, J. I. Martínez, J. Shen, I. C. Man, J. Rossmeisl, T. Bligaard, J. K. Nørskov, and F. Abild-Pedersen, *On the behavior of Brønsted-Evans-Polanyi relations for transition metal oxides*, J. Chem. Phys., **134**, no. 24, 244509, 2011.
8. P. Mehta, J. Greeley, W. N. Delgass, and W. F. Schneider, *Adsorption Energy Correlations at the Metal-Support Boundary*, ACS Catal., **7**, no. 7, 4707–4715, 2017.
9. Z. Wang, Y. Li, J. Boes, Y. Wang, and E. Sargent, *CO_2 Electrocatalyst Design Using Graph Theory*, 2020, Preprint at doi:10.21203/rs.3.rs-66715/v1.
10. T. Gu, B. Wang, S. Chen, and B. Yang, *Automated Generation and Analysis of the Complex Catalytic Reaction Network of Ethanol Synthesis from Syngas on Rh(111)*, ACS Catal., **10**, no. 11, 6346–6355, 2020.
11. W. Xu, M. Andersen, and K. Reuter, *Data-driven descriptor engineering and refined scaling relations for predicting transition metal oxide reactivity*, ACS Catalysis, **11**, no. 2, 734–742, 2020.
12. R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, and L. M. Ghiringhelli, *SISSO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates*, Phys. Rev. Mater., **2**, no. 8, 083802, 2018.
13. M. Togninalli, E. Ghisu, F. Llinares-López, B. Rieck, and K. Borgwardt, *Wasserstein Weisfeiler-Lehman Graph Kernels*, 2019, Preprint at arXiv:1906.01277.
14. W. Xu, K. Reuter, and M. Andersen, *Predicting binding motifs of complex adsorbates using machine learning with a physics-inspired graph representation*, Nat. Comput. Sci., no. 2, 443–450, 2022.
15. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials*, J. Phys.: Condens. Matter, **21**, no. 39, 395502, Sep 2009.
16. J. Wellendorff, K. T Lundgaard, A. Møgelhøj, V. Petzold, D. D. Landis, J. K. Nørskov, T. Bligaard, and K. W. Jacobsen, *Density functionals for surface science: Exchange-correlation model development with Bayesian error estimation*, Phys. Rev. B, **85**, no. 23, 235149, 2012.
17. J. R. Boes, O. Mamun, K. Winther, and T. Bligaard, *Graph Theory Approach to High-Throughput Surface Adsorption Structure Generation*, J. Phys. Chem. A, **123**, no. 11, 2281–2285, 2019.
18. M. Andersen and K. Reuter, *Adsorption Enthalpies for Catalysis Modeling through Machine-Learned Descriptors*, Acc. Chem. Res., **54**, no. 12, 2741–2749, 2021.

19. F. Calle-Vallejo, D. Loffreda, M. T. M. Koper, and P. Sautet, *Introducing Structural Sensitivity into Adsorption-Energy Scaling Relations by Means of Coordination Numbers*, Nat. Chem., **7**, 403–410, 2015.
20. J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. R. K. J. Lindqvist, J. R. Kitchin, T. Bligaard, and H. Jonsson, *Origin of the overpotential for oxygen reduction at a fuel-cell cathode*, J. Phys. Chem. B, **108**, no. 46, 17886–17892, 2004.
21. R. G. Gonzalez-Huerta, G. Ramos-Sánchez, and P. B. Balbuena, *Oxygen evolution in Co-doped RuO2 and IrO2: Experimental and theoretical insights to diminish electrolysis overpotential*, J. Power Sources, **268**, 69–76, 2014.
22. Y. Wu, M. Tariq, W. Q. Zaman, W. Sun, Z. Zhou, and J. Yang, *Bimetallic Doped RuO2 with Manganese and Iron as Electrocatalysts for Favorable Oxygen Evolution Reaction Performance*, ACS Omega, **5**, no. 13, 7342–7347, 2020.
23. T. Reier, Z. Pawolek, S. Cherevko, M. Bruns, T. Jones, D. Teschner, S. Selve, A. Bergmann, H. N. Nong, R. Schlögl, K. J. J. Mayrhofer, and P. Strasser, *Molecular Insight in Structure and Activity of Highly Efficient, Low-Ir Ir-Ni Oxide Catalysts for Electrochemical Water Splitting (OER)*, J. Am. Chem. Soc., **137**, no. 40, 13031–13040, 2015.

Ab Initio Simulations for Microelectronics: $ZnGa_2O_4$ and Hexagonal BN

Jarek Dabrowski, Felix Reichmann, Max Franck, and Mindaugas Lukosius

IHP – Leibniz-Institut für innovative Mikroelektronik,
Im Technologiepark 25, 15236 Frankfurt (Oder), Germany
E-mail: {dabrowski, reichmann, franck, lukosius}@ihp-microelectronics.com

The computational project IHPms21 provides assistance to the IHP Materials Department in the pursuit for new materials that might be integrated into the existing microelectronic and optoelectronic technologies. We report on the results of *ab initio* calculations pertinent to two distinct subjects: (1) to the surface preparation of $ZnGa_2O_4$ (ZGO), which is an ultra-wide gap transparent semiconductor with properties attractive for high-power electronic devices and deep-UV optoelectronic devices, and (2) to the growth of hexagonal BN (hBN) films, which is a layered 2D material structurally akin to the famous graphene and considered as promising dielectric to be used in graphene devices. With the example of ZGO we illustrate the interpretation of photoemission spectra in the terms of surface structure and composition. With the example of hBN we illustrate the seeking for physicochemical processes, which determine the growth of 2D insulators by industrially viable methods.

1 Introduction

The results reported here have been obtained within the project IHPms21, which uses *ab initio* density functional calculations to interpret and to guide the experiments conducted by the IHP[1]. The IHP is a Leibniz institute for basic and applied research and development in the fields of silicon-based microelectronics and technology for wireless and broadband communication. This includes work on new materials and on methods to integrate them into the existing technologies. The *ab initio* theory employed by IHPms21 is the plane wave pseudopotential DFT as implemented in Quantum Espresso[2–4], an actively developed open-source package that since decades has been available to the users of the JSC cluster.

The numerical approach is outlined in Sec. 2. The motivation for and the goal of the calculations are given in Sec. 3. Secs. 4 and 5 describe the results for ZGO (completed study)[5] and for hBN (ongoing research)[6], respectively. ZGO ($ZnGa_2O_4$) is a wide-gap transparent semiconductor with properties attractive for high-power electronic and deep-UV optoelectronic devices, and hBN (hexagonal BN) is a layered 2D material structurally akin to graphene and considered as a promising dielectric for graphene devices.

2 Approach

Atomic structures in question are optimised by the Quantum Espresso (QE) plane wave pseudopotential density functional theory (DFT) package[2], publicly available on JUWELS. Its pw utility constructs and diagonalises the Kohn-Sham Hamiltonian (usually, by the Davidson method)[7]. It solves the Kohn-Sham equations self-consistently and embeds them in a loop, in which the atomic coordinates relax following the Hellmann-Feynman forces. Numerous semilocal and hybrid functionals for exchange-correlation energy are available,

including the libxc library[8]. We mostly use GGA type[9] of exchange and correlation energy (PBE form)[10]. Van der Waals interactions needed for proper description of 2D materials (like hBN) are taken in the non-local meta-GGA form (rVV10)[11]. For selected cases we adapt hybrid (i.e., with a portion of HF exchange) functionals in the HSE form[12]; this is computationally expensive and therefore used here mostly in the case of doubt, as means to establish the validity limit of less heavy and less accurate GGA calculations.

DFT is designed for efficient computation of atomic structures. The electronic structure it renders corresponds to a fictitious system of non-interacting electrons. It resembles the real electronic structure and in many cases can be used as an informative approximation, but in order to derive the physical quantities which, as photoemission, are closely related to the electronic structure, one must often augment it with other means. One way is to employ the above-mentioned hybrid functionals. A far less expensive, though also less precise, less general and more arbitrary approach is to use the so-called Hubbard U correction[13]. It works well for materials such as ZGO and for this reason it was adapted in this work.

Photoemission spectra were simulated from the projection of the crystal states onto the states of isolated atoms. We used HF atomic photoionisation cross-sections tabulated as a function of the photon energy[14]. The inelastic mean free paths of electrons were approximated by the TPP2M formula[15] corrected at low energies E by a $1/E^2$ term[16].

3 Motivation

The path from an idea to its technological realisation may be long, particularly when the confronted roadblocks depend on details evading experimental access. What cannot be seen, must then be computed. Viewed from this perspective, *ab initio* theory becomes just another characterisation tool, with its pros and cons. Being applicable only to smallest dimensions and shortest times, it resembles the experimental tools allowing one to visualise individual atoms, like the scanning tunnelling microscopy (STM) does. And as in the case of STM, when properly applied it can provide understanding helpful in the "debugging" of the concepts and also in the optimisation of the technological process.

The calculations undertaken by IHPms21 and exemplified by the cases reported here are intended to interpret and guide experimental work and to evaluate new ideas by educated guess. In particular, their aim is to infer structure and chemistry of surfaces and to identify the structures, including defective arrangements of atoms, which are responsible for the encountered behaviour of the systems of interest. The main observations, which prompted us to examine ZGO and hBN by *ab initio* means, were:

- X-ray photoemission spectroscopy (XPS) and UV photoemission spectroscopy (UPS) indicate unintended stoichiometry changes of the ZGO surface, apparently induced by the surface preparation process. What is happening on the surface, and does this indicate problems, or new chances, or is just a side effect?

- During the hBN growth by chemical vapour deposition (CVD) from borazine (BZ, $B_3N_3H_6$), the substrate becomes covered by apparently amorphous islands and it is difficult to obtain an ordered, layered film of intended thickness. Why is it so, and what can be changed in the process in order to improve the outcome?

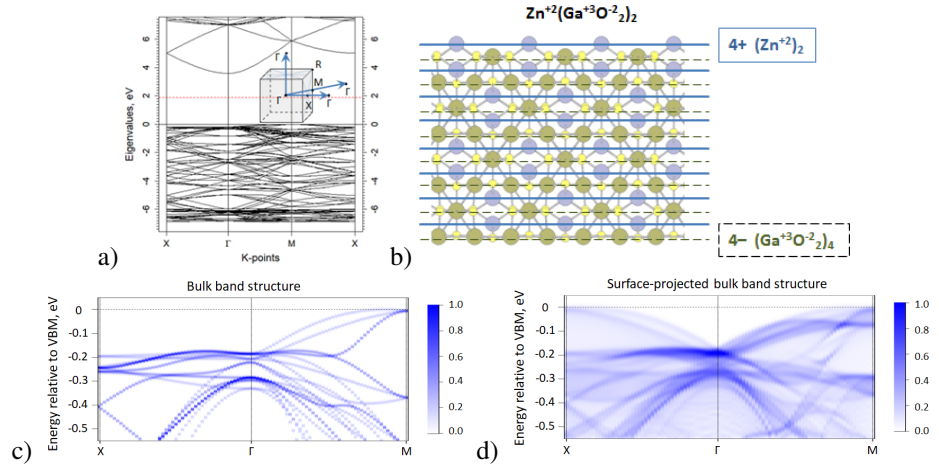


Figure 1. $ZnGa_2O_4$ bulk: a) Band structure. b) Atomic structure: alternating (001) Zn and GaO_2 layers; Zn is blue, Ga is green, O is yellow. c) Top of the valence band. d) Top of the surface-projected valence band, which is the bulk band structure integrated along the Γ-X line, see the Brillouin zone inset in the panel a) of this Figure.

4 The (001) Surface of $ZnGa_2O_4$ Crystals

$ZnGa_2O_4$ is a cubic semiconductor with light electrons and heavy holes (Fig. 1a). Its bulk is composed of alternating (001)-oriented, one atom thick layers of two types: (a) positively charged metal layers, Zn, and (b) negatively charged Ge oxide layers, GaO_2 (Fig. 1b). The gap is indirect (Fig. 1c). Because the surface removes the translational symmetry along the surface normal, the band structure measured by techniques which, like angle resolved ultraviolet photoemission spectroscopy (ARUPS) applied by us, feel the presence of the surface, corresponds to the surface projection of the band structure, with states visible on both sides of the Brillouin zone centre at Γ (Fig. 1d). The band structures measured in our experiments by ARUPS are compatible with this simulation.

We have observed significant nominal loss of Zn from the surface of ZGO crystals as the result of treatments needed to prepare the samples for examination: the intensity of the photoelectrons ejected from Zn by Al Kα photons decreased with respect to those emitted from Ga and from O atoms, indicating stoichiometry changes and not just creation of isolated defects. On the other hand, the prepared surfaces retain high degree of crystallinity. As direct imaging methods (like STM) provided us with no clue regarding the chemical and structural status of the surface region, we addressed the issue by our *ab initio* tools.

Because the surface preparation is complex, involving removal of atoms by sputtering with Ar^+ ions, the thermodynamical status is unclear and it is not practicable to draw conclusions solely on the basis of computed total energies and entropies. By the selection of the structures for close examination we have thus took a semi-empirical approach, excluding the cases, for which (a) well-ordered arrangements with periodicity higher than the observed 1×1 are favourable, (b) there are empty states at more than about 50 meV below the conduction band minimum (CBM), and/or (c) strong photoemission (PE) from states within the band-gap and/or near the valence band maximum (VBM) is expected.

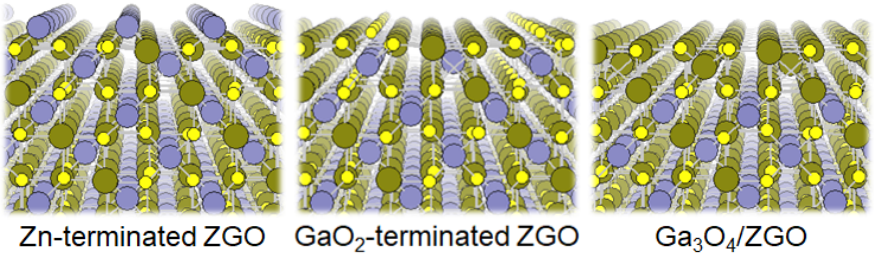


Figure 2. $ZnGa_2O_4(001)$: left: Zn-terminated, middle: GaO_2-terminated, right: Ga_2O_3-terminated. Colours as in Fig. 1b.

Because of the (001) bilayer structuring of the bulk (Fig. 1b), one may have either Zn-terminated or GaO_2-terminated ZGO(001) surface. The simplest explanation of the Zn loss would be to assume that as-loaded samples are Zn-terminated (Fig. 2a) and fully prepared samples are GaO_2-terminated (Fig. 2b). The simulated PE changes are however too small, in agreement with a simple estimate: given the electron escape depth and the measured PE intensity changes, one expects that the loss amounts up to two Zn monolayers. This could be in principle be accounted for by assumption that the as-loaded samples are covered by two Zn monolayers (more precisely, by two ZnO monolayers, because Zn oxidises readily), but the corresponding simulated PE spectra do not fit to those measured.

After having examined numerous structures and stoichiometries and having verified the reliability of the simulations by checking various numerical conditions and approximations (type of pseudopotentials, Hubbard U energies, hybrid functionals) we arrived at the conclusion that the experimental and theoretical data are at the best agreement for the oxide-terminated surface, in which the subsurface Zn atoms are substituted by Ga to form an ultrathin β-Ga_3O_4 film (Fig. 2c). Fig. 3 shows the PE spectra simulated for this case. We also concluded that pseudopotential calculations introduce intrinsic uncertainties to the electron distributions, which cannot be alleviated by Hubbard U or even by HSE. On the other hand, the peak shapes are well assessed, when the position dependence of Hubbard U is accounted for: U must be computed self-consistently, for each surface atom separately.

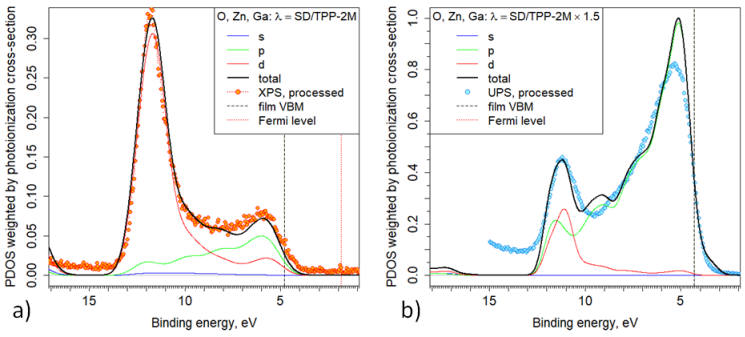


Figure 3. Measured (points) and simulated (lines) PE spectra: a) XPS: Al Kα, 1486.7 eV. b) UPS: He II, 40.81 eV.

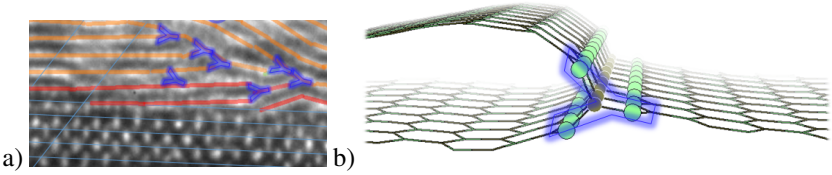


Figure 4. a) HRTEM image of hBN at Ge(001). Ge atoms are visible as big white spots, the blue lines indicate the Ge crystallographic planes. The red lines mark the hBN layers nucleated on Ge, the orange lines mark the other hBN layers, the blue symbols mark the layer connection points. b) Hypothetical structure of a layer interconnection at an antiphase boundary in the bottom layer. Boron atoms are brown, nitrogen atoms are green.

5 Growth of hBN Films from Gaseous Borazine

CVD growth of hBN is achieved by exposing a hot (around 900 °C) Ge(001) sample to gaseous $B_3N_3H_6$ (borazine, BZ). For this purpose, a carrier gas (Ar) is pumped through liquid BZ. The resulting Ar+BZ mixture is ultimately delivered by a stainless steel pipe, which ends about 5 cm away from the sample kept in a vacuum chamber.

We observed that with this approach there is a pronounced tendency for 3D growth: there are high hBN islands (aspect ratio close to 1) surrounded by flat hBN areas. Depending on the technical details, one obtains films, which in high-resolution transmission electron microscopy (HRTEM) images appear to be amorphous with nanocrystalline hBN inclusions, or films built predominantly of nanocrystalline hBN. In the latter case one can see that the deepest layers nucleate at Ge(001) surface steps (Fig. 4a, red lines) and that some layers are interconnected (Fig. 4a, blue symbols). The observed nucleation at Ge(001) steps is in accordance with the computed energy differences and, given the outcome of the calculations, the growth on Ge can be understood as promoted by catalytic activity of the Ge substrate, which is able to decompose BZ down to single atoms. The occurrence of interlayer connections can be readily associated with the natural presence of antiphase boundaries in the nanocrystalline film: we find that such boundaries are the only place at which a new layer may nucleate on an intrinsic layer and remain connected to it. Other plausible nucleation sites on hBN are boron substitutionals (nitrogen vacancies filled with boron), but the layers nucleated there are expected to eventually become detached.

In contrast to the growth directly on Ge, the relatively high hBN growth rate achieved by CVD from BZ on hBN is however not easy to explain. Namely, as soon as the whole Ge surface becomes covered by hBN, the catalytic activity of the substrate is expected to go down practically to zero. Indeed, calculations show that no BZ decomposition can take place on perfect, intrinsic hBN. It follows that from the second hBN layer on, the only BZ molecules contributing to the hBN growth should be those arriving at the layer *edge* undamaged, either directly from the gas or from a narrow collection zone along the edge[17]. The computed desorption energy of BZ (0.4 eV) is however too small and the computed energy barriers for attachment of undamaged BZ at the hBN edge (2.0 eV) are too high to allow for noticeable growth of this type, at least at the temperatures and pressures used in the experiment: the frequency of collisions between the molecules and the step edges and the probability that the impinging or diffusing molecule goes then over the energy barrier to make a bond with the edge atom are then much too small. We estimate that pressures in the atmospheric range would be needed for such a process to be efficient, but the actual pres-

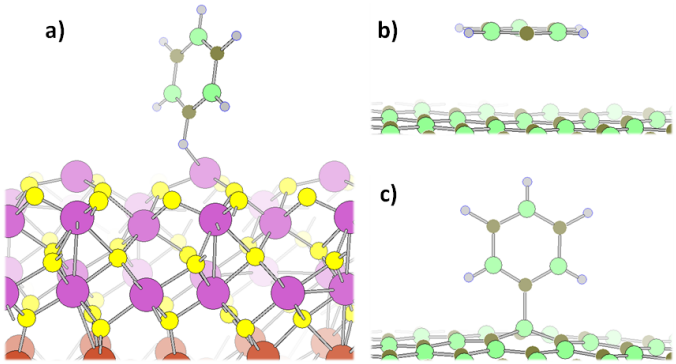


Figure 5. a) $B_3N_3H_6$ adsorbed on a Cr atom of stainless steel. Cr atoms are violet, Fe atoms are red, O atoms are yellow. Dehydrogenation occurs by proton jump from between B and Cr onto one of the surface O atoms. b) $B_3N_3H_5$ is physisorbed on intrinsic hBN with the adsorption energy 0.7 eV. c) $B_3N_3H_5$ is chemisorbed on p-type hBN, with the DFT-CGA adsorption energy of 3.2 eV.

sures in the reaction chamber are at least six *orders of magnitude* below the atmospheric pressure. Even though the DFT-GGA barrier heights are not precise, the required corrections to the barrier heights (from 2.0 eV down to at most 1.6 eV) are significant. Moreover, the critical reaction step is here the reordering of H: the moment when two H atoms – one from the hBN edge and the other one from the incoming molecule – leave the B and N atoms to combine into H_2. The barriers computed by us for this process are comparable to those obtained in similar BN systems by other authors using more precise methods[18].

In order to solve this puzzle, we considered a number of alternative reaction schemes. First, the readings from the mass spectrometer suggest that the gas pumped out of the chamber contains considerable amount of BH_3 of unclear origin, possibly already dissolved in liquid BZ, and also some NH_3. A hBN film can in principle be assembled also from such a BZ + BH_3 + NH_3 mixture. However, although the barriers for BH_3 attachment appear be low enough, the barriers for NH_3 attachments, while lower than those for BZ attachments, may be still too high. Second, the mass spectrometer reveals the presence of water; however, according to the calculations, even if water would take part in the reactions on the edge, it tends to stabilise H there (in the form of OH), instead of destabilising it. Third, we considered the possibility that some of the BZ molecules lose a H atom before reaching the sample. We verified that such $B_3N_3H_5$ molecules would indeed facilitate the growth by abstracting single H atoms from the edge. Some BZ molecules would dehydrogenate to $B_3N_3H_5$ if they come in contact with hot stainless steel (Fig. 5a). Given the computed barrier heights, we estimate that dehydrogenation requires the temperature of the steel to exceed 500 °C. It is unclear if the gas injector heats up to temperatures needed to achieve sufficient partial pressure of $B_3N_3H_5$. A more realistic source of $B_3N_3H_5$ radicals could be the heater or the hot sample holder. Furthermore, the conditions for BZ to lose H are similar for those for H_2 dissociate, so that the reaction H(g) + H(ads) $\Leftrightarrow$ H_2(g) may be responsible for partial dehydrogenation of the edge. Finally, we found that those hBN surface regions, where multiple positive charge carriers (holes) are present in the valence band, are able to dehydrogenate BZ to $B_3N_3H_5$ (Fig. 5b-c) and that this $B_3N_3H_5$ may re-

main on the surface even for about a second at 900 °C and diffuse on it far enough to assure that sufficiently high amount of the radicals take place in the reactions at the step. Such doping with holes may be associated with the formation of defects during growth, nevertheless the doping levels (in the range of several $10^{13}\,\mathrm{cm}^{-2}$) needed for this mechanism to be efficient may be too high to be realistic. The hypotheses involving $B_3N_3H_5$ and/or H radicals formed on hot parts in the chamber appear to be the most promising. We are now exploring them, attempting to find more theoretical support and seeking for ways of experimental verification.

6 Concluding Remarks

We have presented two examples from our research on new materials for microelectronics and optoelectronics, in which we used *ab initio* calculations treated as characterisation tools in parallel to experimental characterisation techniques. For the case of zinc gallate ($ZnGa_2O_4$ZGO), a wide gap semiconductor of potential interest to high power microelectronics and deep UV optoelectronics, we have provided interpretation of photoemission spectra in order to strengthen the evidence for the formation of ultrathin β-Ga_3O_3 film on top of ZGO in the course of a standard surface preparation procedure. For the case of hexagonal boron nitride (hBN), we have analysed various film growth mechanisms in order to explain the observed growth rate and morphology of the film.

Acknowledgements

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)[19]. JD, MF and ML gratefully acknowledge the support from the EU FLAG-ERA project 2DHetero[6].

References

1. https://www.ihp-microelectronics.com/en/start.html
2. P. Giannozzi et al., *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials*, J. Phys. Cond. Mat. **21**, 395502, 2009.
3. P. Giannozzi et al., *Advanced capabilities for materials modelling with Quantum ESPRESSO*, J. Phys. Cond. Mat. **29**, 465901, 2017.
4. *Quantum ESPRESSO*, http://www.quantum-espresso.org/whatcanqedo.php
5. F. Reichmann, J. Dabrowski, A. P. Becker, W. M. Klesse, K. Irmscher, R. Schewski, Z. Galazka, and M. Mulazzi, *Experimental and Theoretical Investigation of the Surface Electronic Structure of $ZnGa_2O_4$(100) Single-Crystals*, Phys. Stat. Sol. B **259**, 2100452, 2021.
6. *2DHetero: hBN/Graphene 2D Heterostructures: from scalable growth to integration*, https://www.flagera.eu/wp-content/uploads/2020/06/FLAG-ERA_JTC2019_GRA-BR_2DHetero.pdf

7. E. R. Davidson, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices*, J. Comp. Phys. **17**, 87, 1975.
8. *Libxc - a library of exchange-correlation functionals for density-functional theory*, `https://www.tddft.org/programs/libxc/`
9. S. Grimme, *Semiempirical GGA-type density functional constructed with a long-range dispersion correction*, J. Comp. Chem. **27**, 1787, 2006.
10. J. P. Perdew, K. Burke, and M. Ernzerhof *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **78**, 1396, 1997.
11. R. Sabatini, T. Gorni, and S. de Gironcoli, *Nonlocal van der Waals density functional made simple and efficient*, Phys. Rev. B **87**, 041108(R), 2013.
12. J. Heyd, G. E. Scuseria, and M. Ernzerhof, *Hybrid functionals based on a screened Coulomb potential*, J. Chem. Phys **118**, 8207, 2003.
13. M. Cococcioni and S. de Gironcoli, *Linear response approach to the calculation of the effective interaction parameters in the LDA+U method*, Phys. Rev. B **71**, 035105, 2005.
14. J. J. Yeh and I. Lindau, *Atomic subshell photoionization cross sections and asymmetry parameters: $1 \leq Z \leq 103$*, Atomic Data and Nuclear Data Tables **32**, 1, 1985.
15. S. Tanuma, C. J. Powell, and D. R. Penn, *Calculations of electron inelastic mean free paths. V. Data for 14 organic compounds over the 50–2000 eV range*, Surface and Interface Analysis **21**, 165, 1994.
16. A. Jablonski, S. Tanuma, and C. J. Powell, *Modified predictive formula for the electron stopping power*, J. Appl. Phys. **103**, 063708, 2008.
17. J. Dabrowski, G. Lippert, T. Schroeder, and G. Lupina, *Role of defects in the process of graphene growth on hexagonal boron nitride from atomic carbon*, Appl. Phys. Lett. **105**, 191610, 2014.
18. V. Kumar, B. Roy, and P. Sharma, *Kinetics of borazine formation from ammonia borane dehydrocoupling reaction through Ab initio analysis*, Journal of Hydrogen Energy **44**, 22022, 2019.
19. D. Krause, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, Journal of large-scale research facilities **5**, A171, 2019, doi:10.17815/jlsrf-5-171.

Atomic-Scale Insights into Friction from Molecular Dynamics Simulations: Fluorinated Carbon Coatings and PTFE

Thomas Reichenbach[1], **Leonhard Mayrhofer**[1], **Gianpietro Moras**[1], **and Michael Moseler**[1,2]

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

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

We use atomistic simulations to investigate the dynamic behaviour of atoms at tribological interfaces and gain insights into the mechanisms that govern friction but are virtually inaccessible experimentally. The main goal of our research within this project is to identify structure-property relationships that link the chemical structure of surfaces and lubricants to friction. Achieving this goal is crucial to complementing experimental analyses in an effort to minimise frictional losses and replace harmful substances often used to reduce friction and wear. To be able to perform accurate large-scale molecular dynamics simulations, we adopt a multiscale approach that hierarchically couples quantum-mechanical and non-reactive force field methods. The force fields are tailored to the specific tribological system under investigation and optimised using information gathered at the quantum-mechanical level. Here, we show two case studies that focus on the tribological properties of fluorinated carbon materials. This class of materials has remarkable tribological properties but poses a health and environmental hazard that makes their replacement a priority. In the first study, we investigate how the chemical properties of C–F bonds influence the dry friction properties of F-passivated diamond and diamond-like carbon coatings, while in the second study we analyse shear-induced transformations in polytetrafluoroethylene and their role in solid lubrication of highly-loaded bearings.

1 Introduction

Friction is a ubiquitous phenomenon and its reduction is of crucial importance in the energetic challenges our society is facing. In order to make progress in this direction and achieve sustainable tribological solutions that reduce emissions and the use of harmful substances, it is necessary to identify the physical and chemical processes that underlie friction and take place at sliding interfaces that are hardly accessible experimentally. Atomistic simulations are proving to be a valuable tool to provide the missing information as they enable a direct investigation of the atomic-scale processes in tribological interfaces. Quantum-mechanical simulation methods such as density-functional theory (DFT) guarantee an accurate description of the complex chemistry of tribological interfaces. However, their use is limited by their high computational cost. Indeed, following the dynamic behaviour of a tribological system using DFT-based molecular dynamics (MD) can only be afforded for model systems composed by up to 1000 atoms and for no longer than tens of picoseconds, even with computational resources from top-tier high-performance-computing facilities such as the Jülich Supercomputing Centre. More efficient atomistic methods based on reactive interatomic potentials would be a valuable alternative to DFT

and allow the simulation of thousands of atoms for hundreds of nanoseconds. Unfortunately, even in their latest developments[1], these methods cannot yet replace DFT in the modelling of most tribological systems, mainly owing to the limited number of chemical elements they can accurately handle simultaneously.

In this project, we adopt a hierarchical multiscale approach that combines DFT and non-reactive force fields (FF)[2] and allows us to perform accurate large-scale MD simulations to gain insights into the structure-property relationship that link the atomic-scale chemical structure of tribological interfaces to friction. DFT simulations on small model systems are used to characterise the tribological interface in terms of the chemical and mechanical properties that are relevant to their frictional properties. The information gathered in this way serves as a "learning" set to develop non-reactive FFs that are specifically tailored for the tribological system under investigation and are used to perform dynamical simulations in which the frictional shear stress can be directly "measured" and related to the physical and chemical properties of the atomistic system. The simplicity and efficiency of the FFs, and their ability to describe many chemical elements, come at a price: they cannot describe chemical reactions (i.e. bond breaking and formation). Therefore, in this approach we rely on the assumption that the tribological system is in a steady-state in which no chemical reactions that can influence friction occur within the simulation time (about 10 ns). This is a reasonable assumption for systems like those we investigate in this project, in which low steady-state friction coefficients are achieved thanks to a chemically stable passivation of the tribological interfaces.

In the next two sections, we present two case studies in which this multiscale approach is applied to fluorinated carbon materials. More specifically, in the first study[3] we investigate the correlation between the properties of C–F bonds and friction in dry interfaces between F-terminated diamond and diamond-like carbon (DLC). In the second case study[4], we focus on the tribologically induced structural transformations that control adhesion and friction in PTFE (polytetrafluoroethylene) used as a solid lubricant in highly-loaded rolling-element bearings. There are two main reasons for the choice of this class of fluorinated carbon materials. On the one hand, they are of considerable technological importance owing to their exceptional friction properties, whose origins are, however, still poorly understood. On the other hand, understanding how friction of CF-based materials works is a first step towards their replacement, which is already being enforced in some fields because of their negative impact on health and environment.

2 Dry Friction of H- and F-Terminated Carbon Coatings: The Importance of Steric Effects

The outstanding friction properties of diamond and DLC coatings in boundary lubrication conditions (i.e. when the lubricant cannot prevent the direct contact between surface asperities) are ascribed to a combination of surface smoothness and a stable passivation of surface dangling bonds[5,6]. The latter is crucial as it prevents the formation of strong covalent C–C bonds between the surfaces in relative motion, which would lead to extremely high friction forces and tribologically induced plastic processes[7]. However, there are many ways to chemically passivate a carbon dangling bond, and there is no general consensus about the structure-property relationships between surface chemistry and friction. In this context, it is particularly instructive to look at an apparently simple and technologically

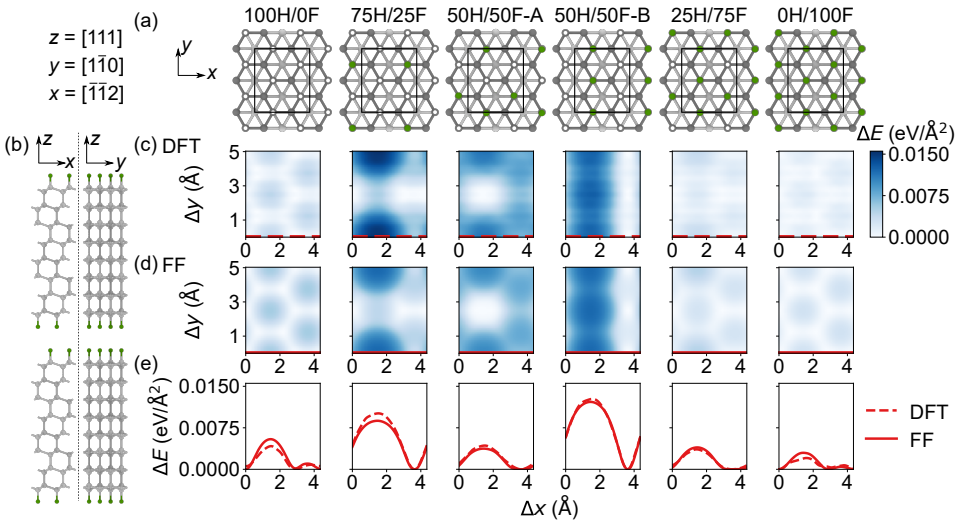


Figure 1. Tailoring a FF for tribological applications using contact potential energy surfaces (CPES) as benchmark data. (a) Model systems of H/F-terminated diamond (111) surfaces. Grey, white, and green spheres represent carbon, hydrogen, and fluorine atoms, respectively. By pairing two such surfaces (b) and evaluating the adhesive energy for different relative lateral positions (Δx, Δy), we obtain the CPES for DFT (c) and for the optimised FF (d). The horizontal red lines in (c) and (d) are cross sections through the CPES, whose energy profiles are shown in (e). Adapted with permission from Ref. 3. Copyright 2020 American Chemical Society.

relevant example: the substitution of hydrogen (the simplest monoatomic termination of a C dangling bond) with fluorine (another monoatomic termination). The effect of such substitution on the friction properties of carbon surfaces is still not clear and previous studies have come to different conclusions regarding which properties of the C–H and C–F surface bonds can be responsible for the different friction properties of these surfaces, namely the different atomic masses and atomic radii of H and F and the different polarities of the C–H and C–F bonds.

To shed light on this problem, we use the multiscale approach mentioned in the Introduction and start by developing a non-reactive FF[2] that can accurately describe the properties of the material system that are relevant for friction[3]. We use the results of accurate van-der-Waals-corrected DFT reference simulations to parameterise a FF that is able to describe diamond's elastic constants, the structure of H- and F-terminated diamond (111) surfaces as well as the electrostatic field induced by the surface terminations. For this purpose, we consider different mixtures of H- and F-terminations: a pure H termination (100H/0F), a pure F termination (0H/100F), and four mixed cases (e.g. 25%H and 75%F which is denoted by 25H/75F). All model systems are shown in Fig. 1a. Fitting the aforementioned quantities allows us to optimise the FF parameters that describe the bonding and Coulomb interactions[2]. The remaining parameters control the Lennard-Jones interactions (i.e. the short-range Pauli repulsion and the long-range dispersion interactions) and are optimised by fitting another quantity that is crucial to friction: the so-called contact potential energy surface (CPES). This quantity describes the variation of the adhesive en-

ergy between two surfaces (Fig. 1b) as a function of their relative position on the xy-plane. Fig. 1c shows the CPES obtained by DFT simulations for the self-mated systems in Fig. 1a (each self-mated interface model system contains 144 atoms and the DFT minimum energy is found for 340 relative positions of the two surfaces). After the optimisation of the Lennard-Jones parameters, the FF can accurately reproduce the DFT CPES, as can be observed by comparing Fig. 1c (DFT) and Fig. 1d (FF). Just by analysing the CPES, we can obtain a first qualitative description (as a normal load and elastic deformations are not considered yet) of the energy barriers the different systems have to overcome upon relative motion. More precisely, any line on the xy-plane defines a sliding path and the corresponding cross section of the CPES gives the energy profile that the system experiences during sliding along this specific path. Fig. 1e shows an example for a sliding path along the x-axis of Fig. 1c (DFT) and Fig. 1d (FF), which corresponds to sliding along the crystallographic $[\overline{1}\overline{1}2]$ direction. Based on these energy profiles, under the assumption that the sum of energy barriers determines friction during sliding, we expect the following qualitative ranking for the shear stresses in ascending order: 0H/100F $<$ 50H/50F-A $\approx$ 25H/75F $<$ 100H/0F $<$ 75H/25F $<$ 50H/50F-B.

The parameterisation of a bespoke, computationally efficient FF allows us to go beyond these semi-quantitative estimates as it enables large-scale MD simulations in which a normal load and elastic deformations can be taken into account and the shear stress during dynamic sliding can be directly calculated. To achieve this goal, we scale the self-mated systems from Fig. 1a up to 114560 atoms (see Fig. 2a). We initially thermalise the systems at $T = 300$ K and impose a constant normal load $P_n = 5$ GPa. After equilibration, we turn off the thermostat to avoid spurious effects on the shear stress calculation (more details in Ref. 3) and drive the system at a constant sliding velocity of $v = 10$ m/s along the $[\overline{1}\overline{1}2]$ direction, i.e. the direction for which we have previously analysed the energy profiles for the small model systems (Fig. 1e). In Fig. 2b we plot the dissipated energy per unit area as a function of the sliding distance. The slope of the curves corresponds to the average shear stress $\langle\tau\rangle_{\mathrm{MD}}$. We find that the ordering of the shear stresses roughly follows the qualitative ordering we estimated by analysing the CPES corrugations of Fig. 1e. A more precise and sophisticated estimate of the energy barriers the tribological systems have to overcome during sliding can be obtained by means of nudged elastic band simulations that are performed directly on the large model systems used for the MD simulations[3]. This analysis results in a very close correlation between the sum of the energy barriers and $\langle\tau\rangle_{\mathrm{MD}}$, as shown in Fig. 2c, and provides us with two important pieces of information. First, for tribological purposes, the choice to fit the FF to the CPES is sound, as the CPES fully determines the friction behaviour of tribological systems with the same stiffness. Second, differences in shear stress between the differently terminated systems are not caused by the different masses of H and F since the CPES corrugation (i.e. the energy barriers) are independent of the masses.

To identify which properties of the C–H and C–F terminations determine the CPES (and thus friction), we take advantage of the fact that single parameters of the FF (e.g. the parameters that control the atomic charge, the atomic radius, the bond length and stiffness) can be arbitrarily varied, leaving the other parameters unaltered. This allows us to perform a systematic study to identify the parameters (i.e. the C–H/C–F physical properties) that are relevant to the frictional behaviour of our systems. We here show the results of this study for the 100H/0F and 0H/100F systems. The horizontal lines in Fig. 2d represent

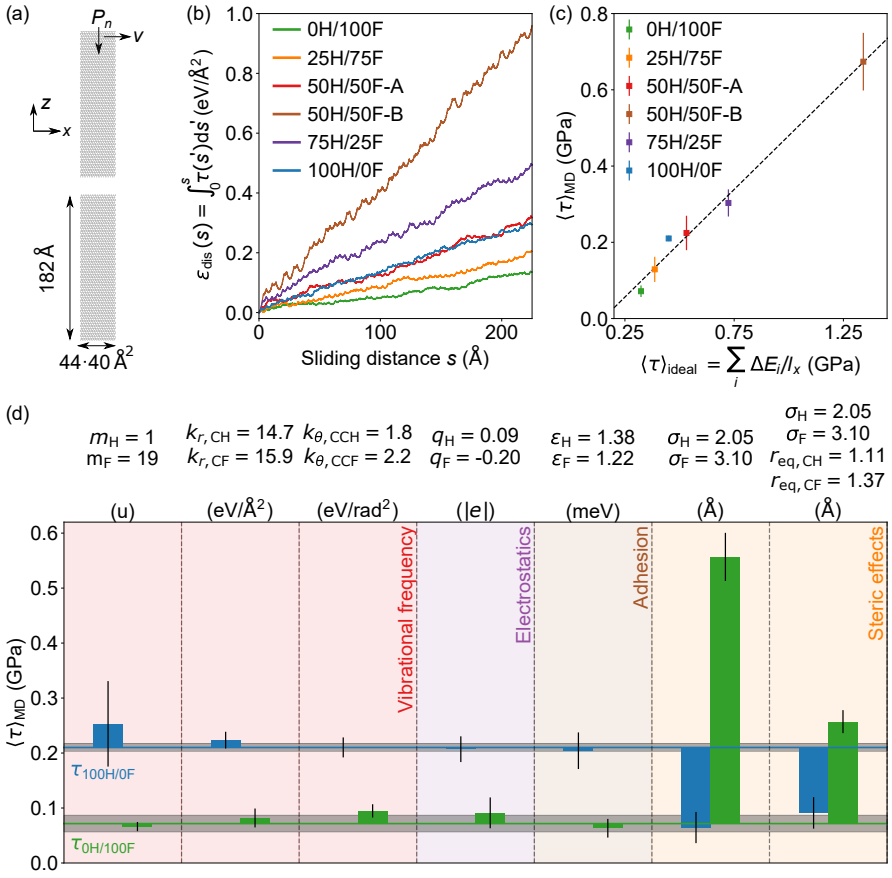


Figure 2. Large-scale MD sliding simulations of self-mated H/F-terminated diamond C(111) surfaces. (a) Simulation setup, where a layer of atoms at the bottom of the lower slab is kept fixed, while a rigid layer of atoms at the top of the upper slab that is subject to a normal load P_n is driven with constant velocity v. Periodic boundary conditions are applied in the interface plane. (b) Dissipated energy as a function of the sliding distance for the system in (a) with different terminations. The slope of each curve corresponds to the respective average shear stress $\langle\tau\rangle_{\mathrm{MD}}$. (c) $\langle\tau\rangle_{\mathrm{MD}}$ as a function of the normalised sum of the CPES barriers determined by nudged elastic band calculations. (d) Variation of $\langle\tau\rangle_{\mathrm{MD}}$ upon parameter exchange for the 100H/0F and 0H/100F systems. The horizontal blue and green line correspond to the reference $\langle\tau\rangle_{\mathrm{MD}}$ for 100H/0F and 0H/100F from (b). Blue bars show the variation of $\langle\tau\rangle_{\mathrm{MD}}$, when the F parameter above the column is assigned to H for the 100H/0F case (e.g. the fist blue bar shows the variation of $\langle\tau\rangle_{\mathrm{MD}}$ for 100H/0F with respect the original value, when the mass of H is increased to the mass of F, while keeping all other parameters fixed). Green bars show the variation of $\langle\tau\rangle_{\mathrm{MD}}$, when the H parameter above the column is assigned to F for the 0H/100F case. The considered parameters are the masses of the terminations m, the force constants associated with C–H/C–F bonds k_r, the force constants for the C–C–H/C–C–F angles k_θ, the H/F charges q, the depth of the H–H and F–F Lennard-Jones potential wells ε, the atomic sizes σ, and a combination of σ and the C–H/C–F bond lengths r_{eq}. Adapted with permission from Ref. 3. Copyright 2020 American Chemical Society.

the reference $\langle\tau\rangle_{\mathrm{MD}}$ values from Fig. 2b without any parameter modification. For each relevant FF parameter, we perform a H $\leftrightarrow$ F parameter exchange: for the 100H/0F system

we assign the value of the F parameter to the H terminations while for the 0H/100F systems we perform the opposite exchange. We then measure the resulting $\langle\tau\rangle_{\mathrm{MD}}$ in a sliding MD simulation. The resulting variations of$\langle\tau\rangle_{\mathrm{MD}}$ with respect to the original $\langle\tau\rangle_{\mathrm{MD}}$ are shown as blue and green bars for the two systems, respectively.

As expected, an exchange of the masses does not yield any significant variation of $\langle\tau\rangle_{\mathrm{MD}}$. The same holds for an exchange of the force constants for bond stretching and angle bending. Interestingly, at odds with most explanations proposed in the literature, also the charge differences between H and F terminations do not account for the measured frictional differences, as also suggested by our previous study on the polar hydrophobicity of fluorinated carbon compounds[8]. Moving now to the exchange of Lennard-Jones parameters, we can observe that also the ε parameters that control the energy of the H–H and F–F Lennard-Jones interactions only cause negligible $\langle\tau\rangle_{\mathrm{MD}}$ variations. Interestingly, however, an exchange of the σ parameters that can be used as a measure of the atomic radii results in very different shear stresses compared to the original values. Furthermore, exchanging both the σ parameters and the C–H/C–F bond lengths r_{eq} results in an inversion of the shear stresses, i.e. upon this parameter exchange the H-terminated system behaves like the original F-terminated one and *vice versa*. These results show that the frictional differences between the 100H/0F and the 0H/100F systems are almost exclusively determined by steric effects. The size of the terminating atoms together with the length and surface density of terminating bonds determine the geometric and energetic smoothness of the interface. In other words, "small" H atoms cannot prevent the "interlocking" of C–H terminal groups sliding against each other and the interlocking depth is determined by the C–H bond length. Conversely, interlocking is suppressed by larger F atoms independently of the C–F bond length.

Further investigations[3] reveal that these steric effects (i.e. the geometric corrugation of the interface) are more general and determine the shear stress also in the case of mixed H/F surface terminations (Fig. 1a), incommensurate diamond contacts, and even sliding interfaces between tetrahedral amorphous carbon (ta-C) surfaces. Moreover, we find that for atomically flat, incommensurate diamond contacts, a substitution of about 40% of the H terminations with F decreases friction significantly. However, for ta-C/ta-C contacts another geometric factor, namely the atomic-scale roughness of the surfaces, plays an additional role. In this case, the substitution of H with F is not necessarily beneficial as the large F terminations can also act as sterical obstacles if their packing is not dense enough.

3 PTFE as Solid Lubricant in Highly-Loaded Rolling-Element Bearings

In the second study, we continue our investigation of the tribological properties of CF compounds and focus on steel surfaces that are lubricated by PTFE. On the one hand, this allows us to extend the findings of the previous study, which only considered C–F bonds as chemical passivations on carbon surfaces, to another class of materials, i.e. fluoropolymers. On the other hand, this second study has a direct connection with a specific technological application since PTFE is a popular solid lubricant for plain bearings and rolling-element bearings (REBs) under low loading conditions. However, in highly-loaded REBs a long-lasting lubrication with PTFE is not possible because PTFE is easily squeezed out of the tribological contact. In order to exploit the excellent tribological properties of PTFE in

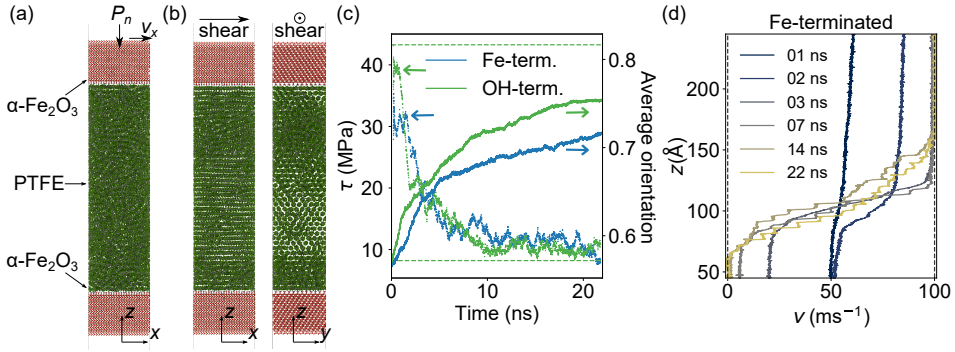


Figure 3. Large-scale MD simulations of PTFE-lubricated α-Fe_2O_3 contacts. (a) Simulation setup consisting of an amorphous PTFE layer that is inserted between two α-Fe_2O_3 slabs. A constant sliding velocity v_x is applied to the outermost atoms of the upper slab, while the outermost atoms of the lower slab are kept fixed. Periodic boundary conditions are applied in the interface plane and the simulation is performed at $T = 300$ K and a constant normal load $P_n = 100$ MPa. White, red, brown, grey, and green spheres represent H, O, Fe, C, and F atoms, respectively. (b) Snapshots of the system in (a) after sliding for 22 ns. (c) Shear stress τ (dotted lines, left scale) and average alignment of the PTFE molecules with the sliding direction (solid lines, right scale) as a function of the sliding time for self-mated Fe- (blue) and OH-terminated (green) surfaces. (d) Velocity profile of PTFE in the Fe-terminated case for different times. Adapted from Ref. 4 under a Creative Commons Attribution 4.0 International License[11].

highly-loaded REBs, a different lubrication concept is required. A promising strategy is to constantly resupply PTFE using an external reservoir. To further develop and optimise this alternative lubrication strategy, a detailed understanding of the PTFE lubrication mechanisms and of the transfer of PTFE from the reservoir to the steel bearing surfaces is essential. For this purpose, we combine large-scale MD simulations with experiments that are performed in the research group of Prof. Georg Jacobs at the RWTH Aachen[4].

The lack of MD studies on the friction and transfer mechanisms of PTFE-lubricated steel contacts can probably be traced back to the lack of FFs that can suitably describe PTFE in combination with α-Fe_2O_3, which is normally used as atomistic model of oxidated steel surfaces[9]. To tackle this issue, we use an approach similar to the one adopted in the first study and develop a non-reactive FF for Fe- and OH-terminated α-Fe_2O_3 surfaces and their interaction with PTFE molecules that are modelled using a non-reactive FF by Watkins and Jorgensen[10]. The new FF's parameters are determined by fitting the elastic constants of α-Fe_2O_3, the bonding properties of the Fe- and OH-terminated α-Fe_2O_3 surfaces and their interactions with perfluorinated linear carbon molecules as obtained by DFT simulations.

In a first step, we use this newly developed, DFT-informed FF to perform large-scale MD sliding simulations with about 75000 atoms (Fig. 3a) to understand the friction mechanisms of PTFE-lubricated α-Fe_2O_3 contacts. Even though PTFE is typically semi-crystalline, we start with an amorphous PTFE phase that is modelled with 400 $C_{40}F_{82}$ molecules to explore the shear-induced structural transformations. The PTFE molecules are placed between two α-Fe_2O_3 surfaces for which we consider both Fe and OH surface terminations. Upon sliding with $v = 100$ m/s under a normal load $P_n = 100$ MPa, we observe a pronounced rearrangement of PTFE's supramolecular structure with molecules

that increasingly align along the sliding direction as sliding proceeds (Fig. 3b) leading to a pseudo-crystalline region. To quantify the molecular alignment, we measure the average absolute projection of the normalised C–C bond vectors of the PTFE molecules on the sliding direction. For both types of terminations, we find a qualitatively similar monotonic increase of the alignment with increasing sliding distance (Fig. 3c). Importantly, the shear stress is closely related to the rearrangement of the molecules: with increasing alignment the shear stress decreases (Fig. 3c). This correlation can be explained by looking at the velocity profile within PTFE, as exemplary shown for the Fe-terminated case at different simulations times in Fig. 3d (the OH-terminated case delivers analogous results). Initially the shear deformation is localised at the PTFE/α-Fe_2O_3 interfaces. However, with increasing alignment of the molecules, the shear deformation localises within the pseudo-crystalline phase and the wall slip disappears. This causes the shear stress to decrease significantly. The velocity profile is approximately linear within this region, i.e. the local shear rate decreases as the thickness of pseudo-crystalline region increases. This is likely the reason for the further decrease of the shear stress.

In a second step, we investigate the atomic-scale mechanisms involved in the transfer of PTFE lubricant from an external reservoir to the steel tribological contact. Tribological experiments of a rolling contact between a sphere and a rotating disc[4] show that the resupply of PTFE in the highly-loaded rolling contact can be achieved by adding a sliding contact between an external PTFE reservoir and the rotating steel disc. In this sliding contact, PTFE is transferred from the reservoir onto the steel surface, and is then transported into the rolling contact by the rotating disc surface, thus enabling an effective relubrication. Our MD simulations confirm that when a α-Fe_2O_3 surface is brought into sliding contact with a PTFE-coated α-Fe_2O_3 surface, a transfer of PTFE to the uncoated surface occurs. We find that the transfer film formation is enabled by the strong adhesion of PTFE to α-Fe_2O_3, which exceeds the molecular cohesion within the PTFE film. Remarkably, the physisorption of PTFE on α-Fe_2O_3 is sufficiently strong to drive the transfer film formation, which does not necessarily require the chemisorption of PTFE on the iron oxide surface.

4 Concluding Remarks

The two case studies that we have chosen to represent this project provide a general example of an information-passing multiscale simulation approach that combines DFT and non-reactive MD, enabling large-scale MD simulations of friction processes. Using this approach, we were able to link the frictional behaviour of a tribological model system to its atomic-scale chemical features. The accuracy of the FF is ensured by a parameter optimisation that is specifically tailored for the friction properties of the investigated systems. Its simplicity and flexibility allow for systematic parametric studies that can help generate design rules for optimising the chemical structure of low-friction systems or for replacing undesired chemical elements. The approach can be extended to a variety of tribological systems that present steady-state low-friction states with negligible chemical transformations after the running-in phase. Moreover, the developed FFs can be reused for similar systems or extended with relatively little effort, for example to include new lubricant molecules. Yet, the number of DFT simulations that are required to develop a FF for the specific system under study, the size of the MD model systems and the number of

parametric studies make our approach computationally intensive, and the work presented here would not have been possible without the computational resources of the Jülich Supercomputing Centre. However, we believe that this approach allows a more effective use of computing resources when compared with fully *ab initio* MD simulations, with which it would not be possible to achieve the space and time scales required by these studies.

Acknowledgements

We gratefully acknowledge funding by the BMWi (Bundesministerium für Wirtschaft und Energie) within the project Poseidon II (03ET1477D), and by the DFG (Deutsche Forschungsgemeinschaft) within the SPP 2074 (project numbers MO 879/22-1 and MO 879/22-2). We are grateful to the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for making these studies possible by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS[12] at Jülich Supercomputing Centre (JSC). Classical MD simulations were carried out with LAMMPS[13], while GPAW[14] and VASP[15] were used to perform the DFT simulations. Visualisation of atomic structures was carried out with OVITO[16].

References

1. A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons*, Physical Review Letters **104**, 136403, 2010.
2. W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, *Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids*, Journal of the American Chemical Society **118**, 45, 11225–11236, 1996.
3. T. Reichenbach, L. Mayrhofer, T. Kuwahara, M. Moseler, and G. Moras, *Steric Effects Control Dry Friction of H- and F-Terminated Carbon Surfaces*, ACS Applied Materials and Interfaces **12**, 7, 8805–8816, 2020.
4. S. von Goeldel, T. Reichenbach, F. König, L. Mayrhofer, G. Moras, G. Jacobs, and M. Moseler, *A Combined Experimental and Atomistic Investigation of PTFE Double Transfer Film Formation and Lubrication in Rolling Point Contacts*, Tribology Letters **69**, 136, 2021.
5. T. Kuwahara, G. Moras, and M. Moseler, *Friction Regimes of Water-Lubricated Diamond (111): Role of Interfacial Ether Groups and Tribo-Induced Aromatic Surface Reconstructions*, Physical Review Letters **119**, 096101, 2017.
6. M. Z. Baykara, M. R. Vazirisereshk, and A. Martini, *Emerging Superlubricity: A Review of the State of the Art and Perspectives on Future Research*, Applied Physics Reviews **5**, 041102, 2018.
7. 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*, Physical Review Materials **2**, 083601, 2018.
8. L. Mayrhofer, G. Moras, N. Mulakaluri, S. Rajagopalan, P. A. Stevens, and M. Moseler, *Fluorine-Terminated Diamond Surfaces as Dense Dipole Lattices: The Electrostatic Origin of Polar Hydrophobicity*, Journal of the American Chemical Society **138**, 12, 4018–4028, 2016.

9. C. A. Latorre, J. P. Ewen, C. Gattinoni, and D. Dini, *Simulating Surfactant-Iron Oxide Interfaces: From Density Functional Theory to Molecular Dynamics*, Journal of Physical Chemistry B **123**, 31, 6870–6881, 2019.
10. E. K. Watkins and W. L. Jorgensen, *Perfluoroalkanes: Conformational Analysis and Liquid-State Properties from Ab Initio and Monte Carlo Calculations*, Journal of Physical Chemistry A **105**, 16, 4118–4125, 2001.
11. `http://creativecommons.org/licenses/by/4.0/`
12. Jülich Supercomputing Centre, *JUWELS Cluster and Booster: Exascale Pathfinder with Modular Supercomputing Architecture at Juelich Supercomputing Centre*, Journal of Large-Scale Research Facilities **7**, A183, 2021.
13. A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, *LAMMPS – a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales*, Computer Physics Communications **271**, 108171, 2022.
14. J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, *Real-Space Grid Implementation of the Projector Augmented Wave Method*, Physical Review B **71**, 035109, 2005.
15. G. Kresse and J. Furthmüller, *Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set*, Physical Review B **54**, 11169, 1996.
16. A. Stukowski, *Modelling and Simulation in Materials Science and Engineering Visualization and Analysis of Atomistic Simulation Data with OVITO – the Open Visualization Tool*, Modelling and Simulation in Materials Science and Engineering **18**, 015012, 2009.

Condensed Matter Theory

Condensed Matter Theory

Frithjof B. Anders

Condensed Matter Theory, Department of Physics, TU Dortmund University,
44221 Dortmund, Germany
E-mail: frithjof.anders@tu-dortmund.de

Strong electron-electron interactions dominate not only the elementary quasi-particle excitations and the band gaps but also influence the exciton dynamics in semiconductors and collective modes such as magnons in magnetically ordered material. Furthermore, they are responsible for the occurrence of magnetic phase transitions emerging from correlated insulator states.

The three rather different examples of state-of-the-art condensed matter theory approaches demonstrate the variety of different problems that can be addressed by the use of top-class supercomputers. Driven by the interplay between the substantial increase in the supercomputing power made available through the John von Neumann Institute for Computing (NIC) and the technological advances in the algorithms, remarkable progress of our understanding of such complex correlation driven systems is reported.

One of the traditional work horses in material science is the local-density approximation (LDA) derived from the the density functional theory. While this approach is highly successful in uncorrelated materials, it has its limitations in semiconductors and in strongly correlated materials. Two of the three contributions to this chapter on condensed matter theory employ the LDA to construct a material adapted effective single-particle starting point and include the necessary many-body corrections to correctly capture the essentials physics of the materials with predictive power.

Rohlfing presents the concept for combining the LDA with a many-body perturbation theory in the so-called GW approximation that includes the electron-electron interaction in leading order and also accounts for screening effects of the Coulomb interaction: Quasiparticle excitations as well as the properties of excitons become accessible. Rohlfing applies this approach to transition-metal dichalcogenides such as MoS_2. Since the discovery of a direct band gap in monolayers of transition-metal dichalcogenides these semiconductor material class have become highly topical: Spin-orbit coupling and the change of the optical properties in layered structures makes these semiconductor materials very attractive for technological applications. In a monolayer of MoS_2, a significantly enhanced band gap is found as well the direct band gap in contrary to the LDA predictions. The exciton energies are modified in a three-dimensional embedding as shown by Rohlfing's approach.

Assad *et al.* target emerging collective phenomena in more generic model Hamiltonians. The aim of the approach is to identify the relevant fundamental interactions in a class of materials, to include these interactions into an effective quantum field theory and to accurately solve the thermodynamic properties of the largest system modern supercomputer can handle without further approximations. In the last decade, Assad *et al.* have developed and employed an auxiliary field Monte Carlo (AFQMC) method for fermions which is greatly benefiting from the supercomputing resources made available through the

John von Neumann Institute for Computing (NIC). The group is in the process of porting their ALF library ("Algorithms for Lattice Fermions") to make use of the growing GPU boosters in a modular supercomputer architecture. The ALF code is made available for everybody as open source via a gitlab server. The Assad group applied their approach to the heavy fermion state which emerges from an interplay between itinerant charge carriers and localised electrons providing a finite magnetic moment. A big challenge is to account for the competition between a magnetically ordered phase and the screening of the local moments by forming a heavy fermi liquid coining the name of a large class of Ce and U based materials. Results for the single-particle gap, the temperature dependent composite fermion spectral function as well as phase boundary between a Kondo insulator and antiferromagnetic phase are presented.

The article by Zhang *et al.*, entitled "Massively Parallel Simulations of Correlated Materials", contains an excellent short tutorial to the so-called dynamical mean-field theory (DMFT) where a periodic lattice problem is mapped onto an effective correlated single site or local cluster problem that is embedded in a self-consistency loop. This effective single-site problem is addressed by the continuous-time quantum Monte Carlo (CTQMC) method where contributions to the partition function and expectation values are evaluated to all orders by a Metropolis algorithm. Due to the massively parallel approach on modern supercomputers correlated materials with multiple electronically active local orbitals can be simulated at low temperatures which is essential to understand the correlation driven physics of transition metal compounds. When using the band structure obtained by a local-density approximation as input into the DMFT, the magnetically order phase of a Mott insulator becomes accessible, and the effective spin-spin exchange interactions can be calculated. A comparison between the experimental data and the DMFT(CTQMC) result for the spin wave dispersion of Sr_2IrO_4 is presented as an example for this rather powerful approach.

Optical Excitations of Two-Dimensional Semiconductors

Michael Rohlfing

University of Münster, Institute of Solid-State Theory, 48149 Münster, Germany
E-mail: Michael.Rohlfing@uni-muenster.de

Two-dimensional semiconductors have particular optical properties. On the one hand, electrons, holes, and excitons are vertically confined to the two-dimensional layer. On the other hand, the screened Coulomb interaction among the particles is significantly enhanced as compared to conventional three-dimensional semiconductors. In combination this leads to strongly bound excitons. However, three-dimensional environment as well as hybridisation between adjacent layers yield characteristic coupling mechanisms, and finally leads to a large variety of characteristic excitonic states. Here we will discuss both intra-layer excitons and their modification by the dielectric environment, as well as, their modification by the admixture of inter-layer configurations that do not exist in a single monolayer.

1 Introduction

Two-dimensional semiconductors are attracting much attention due to their distinct electronic and optical properties[1–7]. Materials like transition-metal dichalcogenides (TMDC, like MoS_2) or hexagonal boron nitride (h-BN) are composed of two-dimensional layers that are stacked on top of each other (bound by van der Waals forces) to form extended bulk crystals. Exfoliation can be used to isolate a single monolayer, a bilayer system, etc., which enables the unique possibility that a monolayer and thicker samples (including extended bulk samples) can be systematically compared to each other. One of the most striking feature is the existence of excitons with very strong optical signature in these materials, in particular in strictly two-dimensional monolayers.

In a monolayer, electrons, holes, and excitons are vertically confined. In addition, the screened Coulomb interaction among the particles is significantly enhanced as compared to conventional three-dimensional semiconductors. Both effects lead to strongly bound excitons that dominate the optical spectra in a much more distinct manner than commonly observed in traditional semiconductor materials. However, when going from a monolayer to thicker samples, hybridisation between the layers causes characteristic modifications of the properties of the two-dimensional excitons.

A systematic description of excitons can be achieved by many-body perturbation theory (MBPT), in particular the GW method followed by the solution of the Bethe-Salpeter equation (BSE), which is a state-of-the-art method to reliably describe excited electronic states[8,9]. Its numerical cost, however, is significant and requires large-scale computational power[10,11].

In our work we are in particular interested in the precise determination of *relative* effects as resulting from differences between related systems, e.g., system-specific states of nanostructures (like monolayer states as compared to bulk states), and modifications due to changes in the non-local dielectric polarisability. In many of our calculations we employ a specific LDA+GdW version of MBPT[12], which is more efficient than a conventional GW calculation while describing the same physics in a complete way, including inter-electronic correlation effects and electron-hole interaction relevant for the excitons. This

LDA+GdW approach is particularly suited to discuss small systematic differences due to environmental issues. We employ these techniques to investigate in particular the red-shift of excitons due to environmental polarisability, as well as, the emergence of exciplex states due to the admixture of charge-transfer configurations if a monolayer is brought in close contact to further layers, like in a bilayer or bulk system.

2 Basic Theory and Computational Aspects

The general framework of many-body perturbation theory (MBPT) has been discussed in numerous publications (see, e.g., Ref. 8 and references therein). In short, MBPT describes single-particle and two-particle excitations as effective individual particles on the background of the interacting many-electron system of the material, in terms of Green functions and their equation of motion[13–15]. Single-particle excitations refer to the addition of removal of a single electron, the energetics of which is reflected in the band structure or energy-level diagram of a system (occupied valence states refer to the removal of an electron, empty conduction states to the addition of an electron), and are described by the single-particle Green function. The net effects of electronic interaction and of exchange effects (from Pauli's principle) are expressed in form of the electronic self-energy operator, which is commonly evaluated on the level of the so-called GW approximation[8,13]. Two-particle excitations occur in form of correlated electron-hole pair states, commonly known as excitons, and are described by the two-particle Green function and its equation of motion (the Bethe-Salpeter equation, BSE)[8,9]. Here we only mention some aspects relevant for the application of the established theory to the case of two-dimensional semiconductor systems[16].

It has been well established that excitons in a semiconducting material are described by the Bethe-Salpeter equation (BSE), applied to a representation of the exciton in terms of a linear combination of independent interband transitions. Within the so-called Tamm-Dancoff approximation, an exciton can be expanded as

$$|S\rangle = \int d^3k \sum_{v,c} A_{v,c}^{(S)}(\mathbf{k}) \, |(v\mathbf{k}) \rightarrow (c, \mathbf{k}+\mathbf{Q})\rangle \quad , \tag{1}$$

In here, v/c denotes occupied and empty bands, $\mathbf{Q}$ is the total momentum of the exciton (often close to zero) and the $\mathbf{k}$-space integration covers the first Brillouin zone. For numerical feasibility we have to replace the continuous integration by a finite summation over $\mathbf{k}$ points:

$$|S\rangle = \sum_{\mathbf{k}_i} \sum_{v,c} A_{v,c}^{(S)}(\mathbf{k}_i) \, |(v\mathbf{k}_i) \rightarrow (c, \mathbf{k}_i+\mathbf{Q})\rangle \quad . \tag{2}$$

Each $\mathbf{k}_i$ represents a volume V_i in reciprocal space (usually all V_i are of equal size and shape). The expansion coefficient $A(\mathbf{k}_i)$ is supposed to represent the average of the original ones, $A(\mathbf{k}_i) = 1/V_i \int_{V_i} A(\mathbf{k}) d^3k$. Finite sampling makes only sense if $A(\mathbf{k})$ varies only weakly within V_i. This set of $\mathbf{k}$ points defines all further requirements of the algorithm.

After defining the excitation in Eq. 1, its equation of motion is given by the BSE in the following way (omitting band indices for brevity sake):

$$\Delta E(\mathbf{k}) A(\mathbf{k}) - \int W(\mathbf{k}-\mathbf{k}') A(\mathbf{k}') d^3\mathbf{k}' = \Omega A(\mathbf{k}) \tag{3}$$

with $\Delta E(\mathbf{k})$ denoting band-energy differences, $W(\mathbf{q})$ denoting the screened Coulomb interaction and Ω the excitation energy. When using Eq. 2, the BSE turns into

$$\Delta E(\mathbf{k}_i)A(\mathbf{k}_i) - \sum_j \tilde{W}(\mathbf{k}_i - \mathbf{k}_j)A(\mathbf{k}_j) = \Omega A(\mathbf{k}_i) \quad . \tag{4}$$

In here,

$$\tilde{W}(\mathbf{k}_i - \mathbf{k}_j) := 1/V_j \int_{V_j} W(\mathbf{k}_i - \mathbf{k}')d^3\mathbf{k}' \tag{5}$$

is the integral of $W(\mathbf{q})$ over a (little) volume V_j around ($\mathbf{k}_i$-$\mathbf{k}_j$). For small reciprocal-space distance $\mathbf{k}_i$-$\mathbf{k}_j$, we employ an analytically known model for $W(\mathbf{k}_i - \mathbf{k}')$ and carry out the integration of Eq. 5 numerically[16]. For large reciprocal-space distance, $\tilde{W}(\mathbf{k}_i - \mathbf{k}_j) \approx W(\mathbf{k}_i - \mathbf{k}_j)$. Note that the treatment of W in the electron-hole interaction for the BSE must be equivalent to the treatment of W in the underlying GW (or LDA+GdW) band-structure calculation, especially concerning anisotropic behaviour of $W(\mathbf{q})$ at small momentum and the finite set of $\mathbf{k}$ points used in Eq. 2. These two issues, i.e. details of the band-structure calculation and of the BSE, must exactly correspond to each other to reach the numerical stability we need for the discussion in the next sections[16].

The underlying reason is the one-to-one correspondence (within MBPT) between the GW self-energy operator Σ^{GW} and the direct part of the corresponding electron-hole interaction kernel derived from Σ^{GW}:

$$\begin{aligned} \Sigma(1,2) &= iG(1,2)W(1^+,2) \\ \Longrightarrow K^d(13,24) &= \partial\Sigma(1,2)/\partial G(4,3) = iW(1^+,2)\delta(1,4)\delta(2,3) \end{aligned}$$

where we have made the usual approximation that $\partial W(1^+,2)/\partial G(4,3) \approx 0$. Apparently, for consistency between GW and BSE, the GW part of the MBPT should employ the identical screened Coulomb interaction as the BSE. This implies using exactly the same $\mathbf{q}$-point grid for the internal summation leading to the self energy, as well as, employing exactly the same modified interaction $\tilde{W}(\mathbf{q})$.

3 Excitons in a Two-Dimensional Monolayer

The starting point of the excitons is a semiconducting band structure. As illustration, Fig. 1 shows our calculated band structure of a monolayer of MoS_2, resulting from many-body perturbation theory[10,11,16–20]. As can be seen in Fig. 1, the band structure is direct, with minimum gap at the K point of the two-dimensional Brillouin zone of the material. An important aspect concerns the fundamental band gap of 2.7 eV, which is significantly larger than commonly assumed (e.g., within density-functional theory), and is also significantly larger than found in a three-dimensional MoS_2 crystal. It is one of the most important observations of careful MBPT analysis that a single monolayer (or, similarly, a fictitious three-dimensional supercell arrangement with large vertical distance between the monolayers) has a drastically larger band gap than a three-dimensional structure. In the latter, the much stronger dielectric polarisability density (resulting from the much larger material density) causes a significant narrowing of the gap ("band-gap renormalisation") compared to a system with a lot of vacuum. In the case of MoS_2 the difference amounts to about

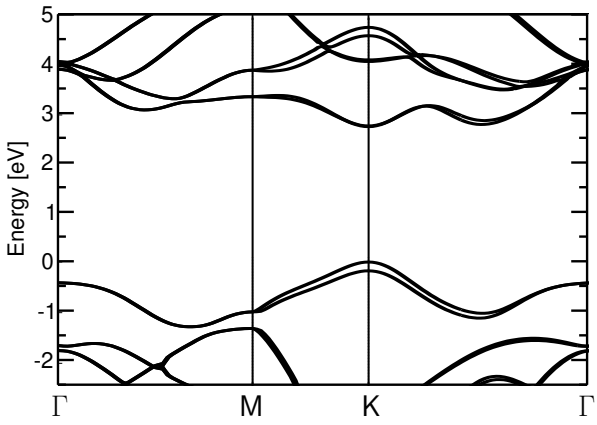


Figure 1. Band structure of a single monolayer of MoS_2, as obtained from many-body perturbation theory (LDA+GdW).

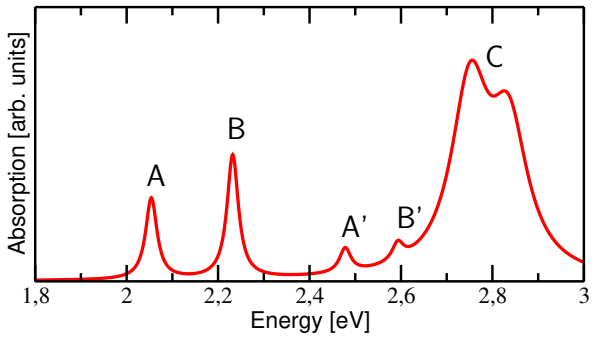


Figure 2. Optical absorption spectrum of of a single monolayer of MoS_2, as obtained from the Bethe-Salpeter equation based on the band-structure data as shown in Fig. 1. A and B denote the lowest-energy excitons stemming from the highest and second-highest valence band of the band structure of Fig. 1. A' and B' denote the next state from the Rydberg series belonging to A and B, respectively. The broad structure labelled C results from other transitions throughout the band structure.

0.6 eV. Another important aspect is the splitting (of about 0.2 eV) between the two uppermost valence bands at the K point, resulting from spin-orbit interaction. Such splitting also occurs in the lowest conduction band, but to much smaller extent.

Based on the band structure, Fig. 2 shows the monolayer's optical excitation spectrum (here: absorption of light shone on the sample from the vertical direction). The spectrum is dominated by two low-energy excitons (A and B) with binding energies of about 0.7 eV, formed between the highest valence and lowest conduction band (exciton A) and between the second-highest valence and second-lowest conduction band (exciton B). The spin-orbit splitting of the valence band leads to the corresponding energy difference between A and B. At higher energy, two excitons A' and B' are observed, which are related to A and B

in the same way as higher atomic levels are related to the ground state of an atom. For example, a hydrogen atom has an $1s$ ground state, followed by a $2s$ excited state at higher energy, with a wave function which has a node to be orthogonal to $1s$. Likewise, in the present system, A' also results from the highest valence state (just as A), but has a wave function $|S\rangle$ orthogonal to A, with a corresponding nodal structure, which forces it to have higher excitation energy. Such a sequence of excitons is expected for all systems with excitonic spectra, and has in fact been observed for the two-dimensional semiconductors. Likewise, B' is related to B, with nodal wave function $|S\rangle$, both coming from the second-highest valence band. In addition, at and above the band gap, a large number of further excitations (all over the Brillouin zone) yield a broad feature labelled C in the spectrum.

4 Modification of Excitons due to Three-Dimensional Embedding

4.1 Redshift from Environmental Polarisability

We have observed in many systems that additional polarisability causes a redshift of excitons[21,22]. As mentioned in Sec. 3, a very drastic effect is on the fundamental band gap of the monolayer. Typical results are that a monolayer, when brought into the environmental polarisability of further layers nearby (for instance, in a threedimensional crystal) is reduced by up to 1 eV (depending on the specifics of the system). The effect is particularly pronounced if the polarisability of the material in question is high (i.e., metals or semiconductors), and is less pronounced if the polarisability is lower (e.g., insulators or molecular systems).

The main reason for the effect is that the band structure corresponds to the physics of one additional particle (electron or hole, depending on whether an occupied valence or empty conduction band is considered). Such a particle induces a charge response in the other layers, with which it interacts (traditionally known as image-charge or image-potential effects for charges in front of a surface, often a metal). This effect raises the energy of a hole state (i.e., valence band), and lowers the energy of an electron state (i.e., conduction band). In total, the fundamental gap between them is narrowed. As noted in Sec. 3, in the case of MoS_2 the renormalisation amounts to 0.6 eV.

The situation is fundamentally different for an intra-layer exciton. Here the excited state is composed of a hole plus an electron, and is thus charge neutral and does not induce an image charge in nearby layers. Nonetheless, it can have a dynamical dipole (in particular in case of an optically allowed state), which will then induce image dipoles in the other layers. Again, there is Coulomb interaction between these dipoles (weaker than between charges, but not negligible), and as a net effect the excitation energy of the exciton is reduced if there is polarisable material nearby. This effect is commonly known as bathochromatism in chemistry, when the excitation of a molecule in solution is compared to that in the gas phase. In our present situation of layered materials we observe redshifts in similar magnitude, i.e. up to about 0.1 eV.

A characteristic case is shown in Fig. 3 which shows the excitation energy of the lowest exciton in hexagonal boron nitride (h-BN). Without the polarisability effect of nearby layers, we find the excitation energy as 5.99...6.01 eV (see open circles). These calculations were performed in a unit cell of finite height c (between 6 and 7 Å). In principle, this energy should be independent of c. In practice, we observe the slight variation as shown

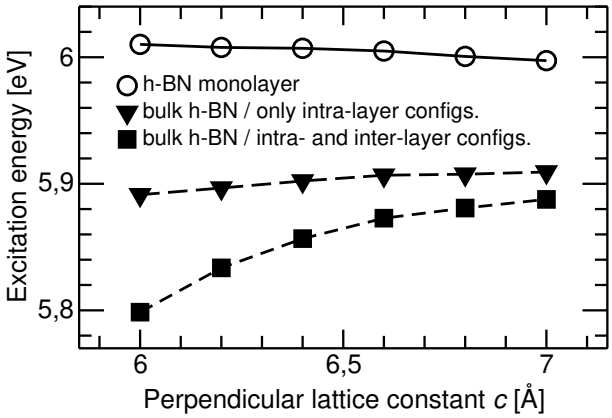


Figure 3. Energy of the lowest exciton in hexagonal boron nitride (h-BN). The upper data (open circles) refer to a single monolayer (showing a little variation with respect to the inter-layer-distance, 6 to 7 Å, for numerical rather than physical reason). The middle and lower data refer to bulk h-BN with vertical lattice constant between 6 and 7 Å, i.e. inter-layer distance between 3 and 3.5 Å. The filled triangles restrict the exciton to intra-layer configurations, while the filled squares also allow for the admixture of inter-layer (or charge-transfer) configurations.

in Fig. 3. We consider this not as a weakness, but rather as an asset of our algorithm: different from many other codes, we here observe a stability (with respect to c) of better than 0.02 eV, which is the basis for discussing the effects as outlined above. Without this stability, numerical noise and systematic falsification would simply prevent a systematic discussion and quantitative assessment of the effects in question.

The main effect of this section is reflected in the difference between the open circles and filled triangles in Fig. 3. Apparently, the environmental polarisability of the neighbouring layers reduces the excitation energy of the intra-layer exciton by about 0.1 eV. Furthermore, the effect depends on the distance between the layers, which is given by $c/2$: at a distance of 3 Å (3.5 Å), the redshift amounts to -0.12 eV (-0.09 eV), showing a slope of 0.06 eV/Å.

Such effects are not just academic theory, but can be expected to occur when the perpendicular lattice constant of a layered material is changed. A prototypical example would be the compression of such a system by external pressure.

4.2 Inter-Layer Band-Structure Hybridisation

It should be noted that the direct nature of the band structure of MoS_2 and other TMDC systems is lost if the layer is coupled to a second layer with perfect registry, as commonly found in naturally grown crystals[1]. Here orbital hybridisation across the van der Waals gap occurs throughout the Brillouin zone, except for the K point. Consequently, orbitals from neighbouring layers will form even and odd linear combinations, and the bands (that are initially degenerate in the various layers) split in energy, by up to 1 eV. This occurs in bilayer, trilayer, and thicker samples up to the threedimensionally extended crystal. Since the K point is excluded from this behaviour for symmetry reasons, the valence-band maximum and conduction-band minimum are no longer located at the K point, and the materials turns into indirect semiconductors. The property of being a direct semiconductor, with the

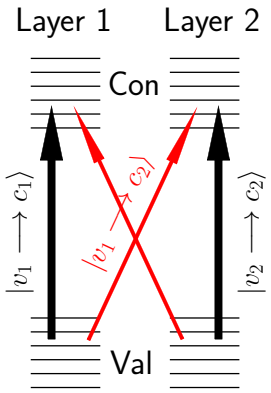


Figure 4. Schematic diagram of electron-hole excitations in a two-layer system. The black vertical arrows indicate intra-layer configurations in which the hole (from the valence bands) and the electron (from the conduction bands) reside in the same layer. Such intra-layer configurations exist in each layer. In contrast, the red slanted arrows indicate inter-layer configurations in which the hole and the electron reside in different layers. Since the charge is moved from one layer to the other, such configurations are also denoted as "charge-transfer configurations". Quantum-mechanical coupling of all these configurations leads to hybridisation among them and to the admixture of charge-transfer configurations to the original intra-layer excitations that would be obtained for well separated monolayers. Such superpositions are also known as "exciplex" states, following the observation of "excimer" states in molecular physics from the same mechanism.

resulting advantages for optoelectronic behaviour and application, is thus mostly restricted to the monolayers. The modifications we have observed in our calculations due to this effect are reported in Refs. 23–25.

4.3 Exciplex States / Charge-Transfer Admixture

Independent of the above-mentioned single-particle like hybridisation effects mentioned in Sec. 4.2, there is a second hybridisation effect on the two-particle level.

The basic physics of such effects is shown in Fig. 4, for the simplest case of two identical objects at close distance. In our present context, these would be two monolayers (of the same material), indicated as "Layer 1" and "Layer 2", but they could equally well be two molecules or two noble-gas atoms[26–29]. In fact, the mechanism to be discussed here was first observed in molecular "excimer" complexes. In short, each of the two objects hosts the same kind of occupied valence and empty conduction states, from which the hole and electron originate when an excited state is formed. An excitation in only one of the layers, e.g. the black arrow in "Layer 1", would constitute an intra-layer excitation as discussed in Fig. 2 or as shown by the triangles in Fig. 3. Note that these intra-layer excitations occur equally well in the other layer, and are energetically degenerate. Alternatively, the excited electron might also hop from one layer to the other, such that the hole resides in, e.g., layer 1 and the electron in layer 2. This would be an inter-layer or charge-transfer excitation, as indicated by the slanted arrows in Fig. 4. Again, for a system made from two identical objects both directions of this charge transfer would cause the same excitation energy.

If the two systems are spatially well separated, the intra-layer excitation energies are

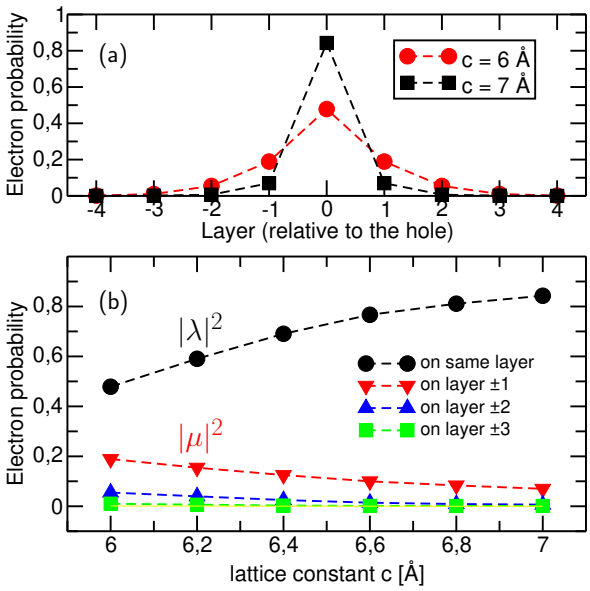


Figure 5. Probability distribution of the excited electron relative to the hole, for the lowest-energy exciton in bulk h-BN. (a) Shown as a function of distance from the layer on which the hole is located. (b) Same data as in (a), but shown as a function of perpendicular lattice constant c.

significantly lower than the inter-layer excitation energies, because the former observe much stronger electron-hole attraction than the latter (the attraction is inversely proportional to the distance and thus vanishes for large separation). Consequently, intra-layer and inter-layer excitations would be decoupled. At closer distance, on the other hand, the inter-layer excitations become lowered in energy and are admixed to the intra-layer excitations. For two layers, the resulting states can be denoted as

$$|S\rangle = \lambda \underbrace{\frac{1}{\sqrt{2}}\left(|1 \to 1\rangle + |2 \to 2\rangle\right)}_{\text{Intra-layer}} + \mu \underbrace{\frac{1}{\sqrt{2}}\left(|1 \to 2\rangle + |2 \to 1\rangle\right)}_{\text{Inter-layer}} \tag{6}$$

with λ and μ strongly depending on the distance between the two layers.

The balance between intra-layer and inter-layer contributions for layered materials depends sensitively on the inter-layer distance (i.e., on the vertical lattice constant, c). As an example, we discuss the intra-layer and inter-layer contributions for the lowest vertical exciton in bulk h-BN. Fig. 5(a) shows relative to the hole the amplitude of the excited electron on the various layers, for two values of the unit-cell height c (i.e. inter-layer distance $c/2$, since h-BN has two layers per unit cell). At small perpendicular lattice constant ($c = 6$ Å) the electron has an amplitude of $|\lambda|^2 \sim 0.5$ on the layer of the hole, accompanied by significant amplitude on the neighbouring layers, i.e. it is rather delocalised. At $c = 7$ Å, on the other hand, the electron is much more strongly localised on the same layer as the hole, with an amplitude of $|\lambda|^2 \sim 0.85$.

The distance-dependence of these effects results from the overlap of and hopping be-

System	Intra-layer contribution	Inter-layer contributions
MoS_2	0.89	0.11
$MoSe_2$	0.92	0.08
$MoTe_2$	0.90	0.10
WS_2	0.87	0.13
WSe_2	0.88	0.12
h-BN	0.73	0.27

Table 1. Calculated intra-layer and inter-layer contributions to the lowest singlet excitons in bulk structures of two-dimensional semiconductors. All values are obtained at the experimental equilibrium distance of vertical stacking. Similar to h-BN, the intra-layer contribution becomes smaller at compression and larger at increased vertical stacking.

tween the orbitals of adjacent layers. Both the overlap of the orbitals and thus the charge-transfer admixture depend exponentially on the inter-layer distance, $c/2$. This is demonstrated in Fig. 5(b), which displays the same data as in panel (a) as a function of c. The behaviour is analogous to the behaviour one would find in a two-layer system (cf. Eq. 6), but for a system of more layers. The black dots correspond to the intra-layer amplitude $|\lambda|^2$, while the red triangles (hole on one layer, electron on one of the neighbouring layers) refer to the inter-layer amplitude $|\lambda|^2$ of Eq. 6. For increasing distance, the quantum-mechanical coupling between intra- and inter-layer configurations decreases, and the inter-layer configurations rise in energy (due to the distance-dependent decrease of electron-hole attraction). Consequently, intra- and inter-layer configurations become disentangled, and the excitons regain their pure intra-layer nature as observed in a single monolayer.

At the natural vertical lattice constant of c = 6.6 Å (inter-layer spacing of 3.3 Å), the intra-layer contribution to the exciton amounts to 0.73. Such behaviour is also observed in the semiconducting TMDC materials, with the values for some materials listed in Tab. 1[17]. In all cases there is significant charge-transfer admixture, and pure intra-layer excitons do not exist in any of the systems.

5 Concluding Remarks

The purpose of this paper was to present a theory to describe excitonic states in two-dimensional semiconductors, and to investigate their behaviour, in particular comparing an isolated monolayer with thicker samples like a bulk crystal. Our method of choice is many-body perturbation theory in its standard form, i.e. combining the GW band-structure framework with the Bethe-Salpeter equation. Equally well, one can use a more efficient LDA+GdW version which describes all the same physical mechanisms at much reduced computational effort. We have found that excitons of an isolated monolayer are significantly modified when the monolayer is contacted with additional material, e.g. if a monolayer is embedded in a crystalline environment. Here two important mechanisms occur: on the one hand, the additional polarisability of the surrounding material red-shifts all excitons by up to 0.1 eV, depending on the specifics of the material. On the other hand, the admixture of charge-transfer configurations (in which the excited electron and

hole reside in different layers) further lowers the excitation energy, leading to a mixture of intra- and inter-layer excitons which can equally well be called an exciplex state.

Acknowledgements

The author thanks T. Deilmann, M. Drüppel, M.-C. Heissenbüttel, P. Krüger, P. Marauhn, and J. Noky for many fruitful discussions, and gratefully acknowledges 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[30] at Jülich Supercomputing Centre (JSC).

References

1. K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, *Atomically Thin MoS2: A New Direct-Gap Semiconductor*, Phys. Rev. Lett. 105, 136805, 2010.
2. A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C.-Y. Chim, G. Galli, and F. Wang, *Emerging photoluminescence in monolayer MoS2*, Nano. Lett. 10, 1271, 2010.
3. A. Ramasubramaniam, *Large excitonic effects in monolayers of molybdenum and tungsten dichalcogenides*, Phys. Rev. B 86, 115409, 2012.
4. H.-P. Komsa and A. V. Krasheninnikov, *Effects of confinement and environment on the electronic structure and exciton binding energy of MoS2 from first principles*, Phys. Rev. B 86, 241201, 2012.
5. J. Feng, X. Qian, C.-W. Huang, and J. Li, *Strain-engineered artificial atom as a broad-spectrum solar energy funnel*, Nat. Photonics 6, 866, 2012.
6. I. Ben Amara, E. Ben Salem, and S. Jaziri, *Optoelectronic response and excitonic properties of monolayer MoS2*, J. Appl. Phys. 120, 051707, 2016.
7. H. Zeng, J. Dai, W. Yao, D. Xiao, and X. Cui, *Valley polarization in MoS2 monolayers by optical pumping*, Nat. Nanotechnol. 7, 490, 2012.
8. G. Onida, L. Reining, and A. Rubio, *Electronic excitations: Density-functional versus many-body Green's-function approaches*, Rev. Mod. Phys. 74, 601, 2002.
9. M. Rohlfing and S. G. Louie, *Electron-hole excitations and optical spectra from first principles*, Phys. Rev. B 62, 4927, 2000.
10. F. Hüser, T. Olsen, and K. S. Thygesen, *How dielectric screening in two-dimensional crystals affects the convergence of excited-state calculations: Monolayer MoS2*, Phys. Rev. B 88, 245309, 2013.
11. D. Y. Qiu, F. H. da Jornada, and S. G. Louie, *Screening and many-body effects in two-dimensional crystals: Monolayer MoS2*, Phys. Rev. B 93, 235435, 2016.
12. M. Rohlfing, *Electronic excitations from a perturbative LDA+GdW approach*, Phys. Rev. B 82, 205127, 2010.
13. L. Hedin, *New method for calculating the one-particle green's function with application to the electron-gas problem*, Phys. Rev. 139, A796, 1965.
14. M. S. Hybertsen and S. G. Louie, *Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies*, Phys. Rev. B 34, 5390, 1986.
15. G. Strinati, *Application of the Green's functions method to the study of the optical properties of semiconductors*, Rivista del Nuovo Cimento 11, 1, 1988.

16. M. Drüppel, T. Deilmann, J. Noky, P. Marauhn, P. Krüger, and M. Rohlfing, *Electronic excitations in transition metal dichalcogenide monolayers from an LDA+GdW approach*, Phys. Rev. B 98, 155433, 2018.
17. M.-C. Heissenbüttel, P. Marauhn, T. Deilmann, P. Krüger, and M. Rohlfing, *Nature of the excited states of layered systems and molecular excimers: Exciplex states and their dependence on structure*, Phys. Rev. B. Phys. Rev. B 99, 035425, 2019.
18. H. Shi, H. Pan, Y.-W. Zhang, and B. I. Yakobson, *Quasiparticle band structures and optical properties of strained monolayer MoS2 and WS2*, Phys. Rev. B 87, 155304, 2013.
19. D. Y. Qiu, H. Felipe, and S. G. Louie, *Optical Spectrum of MoS2: Many-Body Effects and Diversity of Exciton States*, Phys. Rev. Lett. 111, 216805, 2013.
20. A. Molina-Sanchez, D. Sangalli, K. Hummer, A. Marini, and L. Wirtz, *Effect of spin-orbit interaction on the optical spectra of single-layer, double-layer, and bulk MoS2*, Phys. Rev. B 88, 045412, 2013.
21. M. Rohlfing, *Redshift of Excitons in Carbon Nanotubes Caused by the Environment Polarizability*, Phys. Rev. Lett. 108, 087402, 2012.
22. M. Drüppel, T. Deilmann, P. Krüger, and M. Rohlfing, *Diversity of trion states and substrate effects in the optical properties of an MoS_2 monolayer*, Nat. Comm. 8, 2117, 2017.
23. A. Arora et al., *Valley-contrasting optics of interlayer excitons in Mo- and W-based bulk transition metal dichalcogenides*, Nanoscale 10, 15571, 2018.
24. Y. Niu et al., *Thickness-Dependent Differential Reflectance Spectra of Monolayer and Few-Layer MoS_2, $MoSe_2$, WS_2 and WSe_2*, Nanomaterials 8, 725, 2018.
25. R. Wallauer, P. Marauhn, J. Reimann, S. Zoerb, F. Kraus, J. Güdde, M. Rohlfing, and U. Höfer, *Momentum-resolved observation of ultrafast interlayer charge transfer between the topmost layers of MoS_2*, Phys. Rev. B 102, 125417, 2020.
26. J. B. Birks, Rep. Prog. Phys. 38, 903, 1975.
27. Y. Wang, O. Haze, J. P. Dinnocenzo, S. Farid, R. Farid, I. R. Gould, *Bonded Exciplexes. A New Concept in Photochemical Reactions*, J. Org. Chemistry 72, 6970, 2007, doi:10.1021/jo071157d.
28. J. Liang, Q. L. Nguyen, and S. Matsika, *Exciplexes and conical intersections lead to fluorescence quenching in π-stacked dimers of 2-aminopurine with natural purine nucleobases*, Photochemical and Photobiological Sciences 12, 1387, 2013, doi:10.1039/c3pp25449f.
29. J. Kalinowski, M. Cocchi, D. Virgili, V. Fattori, and J. A. G. Williams, *Mixing of Excimer and Exciplex Emission: A New Way to Improve White Light Emitting Organic Electrophosphorescent Diodes*, Adv. Mater. 19, 4000, 2007.
30. Jülich Supercomputing Centre, *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.

Numerical Simulations of Strongly Correlated Electron Systems

Fakher Assaad, Bimla Danu, Florian Goth, Anika Götz, Zihong Liu, Francesco Parisen Toldin, Marcin Raczkowski, Gabriel Rein, Toshihiro Sato, Jonas Schwab, Jefferson Stafusa E. Portela, and Maksim Ulybyshev

Institut für Theoretische Physik, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany
E-mail: {assaad, bimla.danu, florian.goth, anika.goetz, zihong.liu, francesco.parisentoldin, marcin.raczkowski, gabriel.rein, toshihiro.sato, jonas.schwab, jstafusa, maksim.ulybyshev}@physik.uni-wuerzburg.de

Our research is devoted to correlation induced phenomena in the solid state. Owing to the quantum to classical mapping that, in some cases, maps d-dimensional quantum systems to d+1-dimensional classical ones, this includes collective and especially critical phenomena in classical systems [a]. Correlations lead to collective phenomena with seemingly unlimited richness, which, within a reductionist approach to physics, includes life. Here we will be much more modest and concentrate on model systems in thermodynamic equilibrium. However, we will carry out no other approximation other than the system size. Our topics broadly fall into two categories. On one hand we aim at modelling materials, and comparing model results directly to experiments. On the other hand we strive at understanding fundamental phenomena and novel emergent phases and phase transitions. We use various quantum Monte Carlo (QMC) methods, and in particular are maintaining and constantly further developing a general implementation of the auxiliary field QMC algorithm: the Algorithms for Lattice Fermions library (ALF). In this article, we will briefly give an account of various research subjects carried out within our grant proposal and then consider the special case of magnetic impurities in metallic environments.

1 Introduction and Motivation

Our overarching aim is to understand emergent collective phenomena in thermodynamic equilibrium in classical and/or in quantum systems. In fact as physicists we follow the ambitious reductionist point of view that simple elementary rules will suffice to describe the seemingly unlimited richness of nature. The elementary rules are described in terms of model Hamiltonians, that dictate time evolution and thermodynamic properties. Our goal is to make no further approximation and to solve this many body problem on as large systems as possible. For the class of stochastic Monte Carlo methods we use, the computational time, on a single core, scales as a method dependent power-law of the volume. To achieve large systems, we rely entirely on supercomputing resources in Germany, and in particular at the John von Neumann Institute for Computing (NIC). One may ask why the system size is so important. Emergent and critical phenomena is characterised by a hierarchy of energy scales. The inverse lattice size acts as an energy cutoff such that we are blind to low energy phenomena if the lattice size is too small. Optimally, we would like the inverse lattice size to be the lowest scale in the problem, but since it is *a priori* unknown, we have to systematically enhance the system size to see if novel phenomena emerges. Hence, on

[a] As a canonical example the Ising model in a transverse field on a d-dimensional hyper-cubic lattice maps onto the (anisotropic) Ising model maps on a d+1-dimensional hyper-cubic lattice.

general grounds the size of the system acts as a figure of merit. This quest of bigger and bigger system sizes is similar to the quest of high energies in particle physics. In particle physics, higher energy scales allow to discover new particles.

Our workhorse is the so-called auxiliary field quantum Monte Carlo algorithm[1–3] , that we will briefly touch upon in the next section. We also use the related so-called hybrid Monte Carlo approach[4,5] that allows in special cases to reach unprecedentedly large lattices[6]. In Sec. 2 we will provide a very succinct overview of the implementation of our open source general auxiliary field QMC code[7,8]. Over the last years we have invested considerable time and effort to promote the **A**lgorithms for **L**attice **F**ermions (ALF) library. As we will describe below, it is unique in its generality. At this stage in the development of the ALF library, major challenges include ease of use, implementation of novel ideas, GPU support, and quality control. This is a daunting task. It can only be tackled if funding agencies deem it important enough to dedicate a line of funding for these types of endeavours that do not fully qualify as fundamental research but that are fundamental for efficient research. In particular, aspects of the FAIR principles of data stemming from computer simulations hinge on sustainable and documented open source program packages.

Let us now concentrate on some of the physics questions we have been able to address with the allocated computer resources. The physics of a Fermi liquid is captured by ^{3}He. The low energy physics of this liquid, is described by a collection of fermions subject to quantum statistics and experiencing no further interactions. Clearly ^{3}He is a strongly interacting fermion system. However one can argue that the combination of Fermi statistics and the short ranged nature of the Lenard-Jones potential suffice to show the stability of the Fermi liquid state[9]. It turns out that the Fermi liquid breaks down in graphene. Graphene is a realisation of interacting Dirac fermions. Since in free standing graphene the density of states vanishes at the Fermi level, the Coulomb repulsion is not screened and the Fermi liquid state breaks down. In fact the Coulomb repulsion at the Fermi liquid fixed point is marginally relevant and one expects the Fermi velocity to diverge logarithmically[10]. This expectation is supported by experiments[11]. It is remarkable to realise that, by using state of the art hybrid Monte Carlo GPU based simulations of realistic quantum models of this material[6], we are now in a position to reach system sizes that allow for direct comparison with experiments. Hence the possibility of realising a long standing goal: a quantitative comparison between correlation driven phenomena and model calculations.

It is fascinating to see that the physics of magnetic impurities embedded in a metallic environment maps onto U(1) gauge theories, akin to quantum electro-dynamics[12]. This mapping is at the basis of our understanding of fractionalised phases in these models. Here, novel elementary particles with different quantum numbers than those of the electron (spin-1/2 and charge e) emerge. It also provides a very detailed understanding and classification of the possible phases of matter that can occur in these system. Interestingly, all these phases of matter can be realised in a simple, experimental relevant system: a spin-1/2 magnetic chain on a metallic surface[13–16]. Note that even if we can foresee novel emergent phenomena at the theoretical level it is the computer that will tell us if these phases are realised in model calculations and ultimately in materials. In this overview we will concentrate on the Kondo lattice model (KLM) on the square lattice and, show that the large-N approximation of the Kondo insulating phase is adiabatically connected to the physical N=2 one[17]. We will also show that the fundamental quantity to understand the so called heavy fermion state is a composite particle carrying the quantum numbers of the electron[18].

2 Computational Methods

Auxiliary field Monte Carlo (AFQMC) methods for fermions[1,2,8] are based on the following representation of the partition function of a given Hamiltonian $\hat{\mathcal{H}}$

$$Z = \mathrm{Tr}\left[e^{-\beta\hat{\mathcal{H}}}\right] = \int D\Phi(\boldsymbol{i},\tau)e^{-S(\Phi(\boldsymbol{i},\tau))}. \tag{1}$$

Here the integration runs over a space ($\boldsymbol{i}$) and imaginary time (τ) dependent scalar field $\Phi(\boldsymbol{i},\tau)$ and S is the action of a single particle Hamiltonian subject to this field. The crucial question is if the action is real. In this case, $e^{-S(\Phi(\boldsymbol{i},\tau))}$ can be interpreted as a weight function and the high dimensional integration can be carried out stochastically with Monte Carlo importance sampling. Thereby the exponential problem (the dimension of the Hilbert space grows exponentially with the system size V) is reduced to a polynomial one and the AFQMC approach scales as $V^3\beta$ where β is the inverse temperature. If S is not real, then we are confronted to a so-called negative sign of phase problem. Clearly the partition function is positive, but in the presence of a negative sign problem it results from the cancellation of positive and negative contributions. As the system size grows this cancellation is exponential, such that the noise dominates. The sign problem is formulation dependent and notion of how severe the problem is reflects in the average sign:

$$\langle \mathrm{sign} \rangle = \frac{\int D\Phi(\boldsymbol{i},\tau)\; e^{-S(\Phi(\boldsymbol{i},\tau))}}{\int D\Phi(\boldsymbol{i},\tau)\; |e^{-S(\Phi(\boldsymbol{i},\tau))}|}. \tag{2}$$

In the past years, we have witnessed sustained progress in defining the class of models that are free of the negative sign problem. Using symmetry considerations[19–21] one can design models that capture the desired physics and that are so-called sign free, $\langle \mathrm{sign} \rangle = 1$. For instance, in Ref. 21, it is shown that the negative sign problem is absent if one can find two anti-unitary operators that mutually anti-commute and that commute with the aforementioned single particle Hamiltonian for each field configuration. This set of very powerful tools has allowed us to define a class of frustrated Kondo lattice models[22] as well as a variety of models showing de-confined quantum critical points and de-confined phases[23–25]. We note however that even in the presence of a negative sign problem, very useful information at *high* temperatures can be extracted as shown in Ref. 26.

The generality of the ALF code[8] lies in the definition of the class of Hamiltonians $\hat{\mathcal{H}}$ that can be simulated at minimal programming costs. The master Hamiltonian can be written as:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_T + \hat{\mathcal{H}}_V + \hat{\mathcal{H}}_I + \hat{\mathcal{H}}_{0,I}\,, \text{ where} \tag{3}$$

$$\hat{\mathcal{H}}_T = \sum_{k=1}^{M_T}\sum_{\sigma=1}^{N_{\mathrm{col}}}\sum_{s=1}^{N_{\mathrm{fl}}}\sum_{x,y}^{N_{\mathrm{dim}}} \hat{c}^{\dagger}_{x\sigma s} T^{(ks)}_{xy} \hat{c}_{y\sigma s} \equiv \sum_{k=1}^{M_T} \hat{T}^{(k)}\,, \tag{4}$$

$$\hat{\mathcal{H}}_V = \sum_{k=1}^{M_V} U_k \left\{ \sum_{\sigma=1}^{N_{\mathrm{col}}}\sum_{s=1}^{N_{\mathrm{fl}}} \left[\left(\sum_{x,y}^{N_{\mathrm{dim}}} \hat{c}^{\dagger}_{x\sigma s} V^{(ks)}_{xy} \hat{c}_{y\sigma s} \right) + \alpha_{ks} \right] \right\}^2 \equiv \sum_{k=1}^{M_V} U_k \left(\hat{V}^{(k)} \right)^2, \tag{5}$$

$$\hat{\mathcal{H}}_I = \sum_{k=1}^{M_I} \hat{Z}_k \left(\sum_{\sigma=1}^{N_{\mathrm{col}}}\sum_{s=1}^{N_{\mathrm{fl}}}\sum_{x,y}^{N_{\mathrm{dim}}} \hat{c}^{\dagger}_{x\sigma s} I^{(ks)}_{xy} \hat{c}_{y\sigma s} \right) \equiv \sum_{k=1}^{M_I} \hat{Z}_k \hat{I}^{(k)}\,. \tag{6}$$

Here $\hat{c}^{\dagger}_{x\sigma s}$ is a fermion operator that has a generalised site index x, a colour index σ and a flavour index s. The hermitian matrices, T, V and I are independent on the colour index and do not mix flavour sectors. As a consequence, our Hamiltonian enjoys a global SU(N_{col}) symmetry. The first term corresponds to the kinetic energy. Correlation effects, $\hat{\mathcal{H}}_V$, have to be written in terms of sums of perfect squares. $\hat{\mathcal{H}}_I$ allows the coupling of a scalar field with user defined dynamics, $\hat{\mathcal{H}}_{0,I}$ to a fermion bilinear. These terms allow for the simulations of electron-phonon interactions[27] and lattice gauge theories[24,25,28].

Given the above Hamiltonian, it is in principle possible to write a Python code that generates automatically the Fortran 2003 Hamiltonian module required to use the ALF-library. However, with very modest Fortran knowledge, adapting the Hamiltonian module template provided in the source code, is straightforward. The lion share of the computations are carried out by MKL routines. Finally the code runs in a hybrid MPI/OpenMP environment.

3 Magnetic Impurities in Metals: The Heavy Fermion State

The key ingredient in the physics of heavy fermions is the interplay between two types of charge carriers: one being delocalised and described in terms of a large band-width Fermi liquid, the other being on the verge of localisation and possessing a local moment. There are many experimental realisations of this unique situation. In particular, recent experiments on bilayer ^{3}He[29] have been interpreted in terms of heavy fermion physics. Here the second layer of ^{3}He starts forming before the first solidifies. Being far from the solidification concentration, the second layer corresponds to the Fermi liquid[9]. In contrast, the first layer is on the verge of solidification, exhibits a large effective mass and local moments[30]. Heavy fermion physics is historically present in Ce-based compounds. Here, a strong Coulomb repulsion on the $4f$ shell leads to single occupancy and to the formation of local magnetic moments[31].

Heavy fermion systems are characterised by a hierarchy of scales. The Kondo scale marks the screening of local magnetic moments. This screening is a many-body effect which entangles, via the formation of a spin singlet, the spins of the conduction electrons and local moments. Below the coherence temperature, which is believed to track the Kondo scale, the paramagnetic heavy fermion liquid emerges. Even in the Kondo limit, where charge fluctuations of the impurity spins are completely suppressed, this paramagnetic state is characterised by a large Fermi surface with Luttinger volume including both the magnetic moments and conduction electrons. The coherence temperature of this metallic state is small or, equivalently, the effective mass large. In Ref. 32 we have provided evidence that the periodic ensemble of already 25 magnetic impurities exhibits the electronic band structure characteristic of a coherent Bloch state of heavy quasiparticles in the infinite Kondo lattice. Thus our calculations rationalise the interpretation of the spectroscopic characteristics of the artificially engineered superlattice of iron(II) phthalocyanine molecules on a Au(111) surface in terms of Kondo lattice behaviour[33].

Kondo screening competes with the Ruderman-Kittel-Kasuya-Yosida (RKKY) interaction, which indirectly couples the local moments via the magnetic polarisation of the conduction electrons. The competition between Kondo screening – favouring paramagnetic ground states – and the RKKY interaction – favouring magnetically ordered states – is at the heart of quantum phase transitions[34], the detailed understanding of which is still

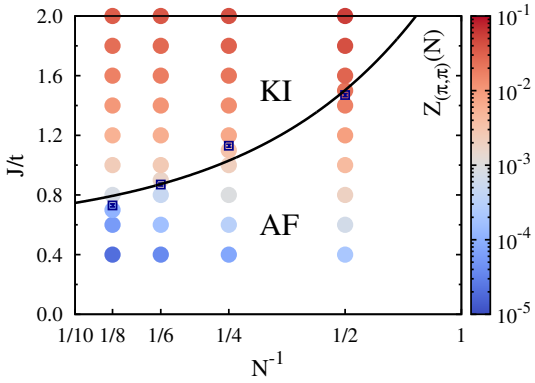


Figure 1. Ground-state phase diagram of the SU(N) KLM as a function of the inverse of the number of fermionic flavours N and Kondo coupling J with antiferromagnetic (AF) and Kondo insulating (KI) phases at half-filling; empty squares indicate magnetic order-disorder transition points extracted from the behaviour of the staggered moment in the thermodynamic limit; solid line is a fit to a finite-N functional form of the critical coupling $J_c \propto 1/\log N$ obtained by comparing the Kondo temperature T_K to the magnetic energy scale T_{RKKY}. Colour-coded circles correspond to the quasiparticle (QP) residue $Z_{\boldsymbol{k}}$ of the doped hole at momentum $\boldsymbol{k} = (\pi, \pi)$. This figure is reproduced from Ref. 17.

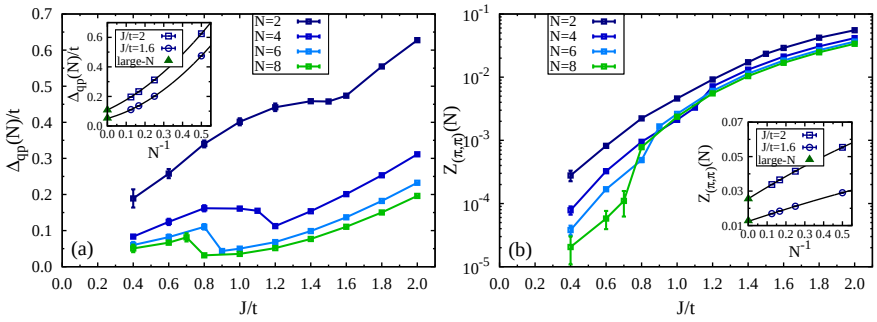


Figure 2. (a) Single-particle gap Δ_{qp} at momentum $\boldsymbol{k} = (\pi, \pi)$ and (b) the corresponding QP residue $Z_{(\pi,\pi)}$ after extrapolation the QMC data to the thermodynamic limit. Insets show second-order polynomial fits to the QMC data in order to extract Δ_{qp} and $Z_{(\pi,\pi)}$ in the $N \to \infty$ limit; the extrapolated values $\Delta_{qp}^{N\to\infty} = 0.109(3)$ and $Z_{(\pi,\pi)}^{N\to\infty} = 0.0256(3)$ at $J/t = 2$ and $\Delta_{qp}^{N\to\infty} = 0.051(2)$ and $Z_{(\pi,\pi)}^{N\to\infty} = 0.0125(2)$ at $J/t = 1.6$ match well those obtained using the large-N approximation (triangles). This figure is reproduced from Ref. 17.

under debate. Here, two radically different scenarios have been put forward to describe this quantum phase transition. In the *standard* Hertz-Millis picture[35,36], the quasi-particles of the heavy fermion liquid remain intact across the transition and undergo a spin density wave transition. Since many experimental observations such as the almost wave vector independent spin susceptibility in $CeCu_{6-x}Au_x$[37], or the *jump* in the low temperature Hall coefficient in $YbRh_2Si_2$[38] are not accounted for by this theory alternative scenarios have been put forward[39,40]. In those scenarios, the quantum critical point is linked to the very breakdown of the quasi-particle of the heavy fermion state, and a topological reorganisation of the Fermi surface across the transition is expected[38].

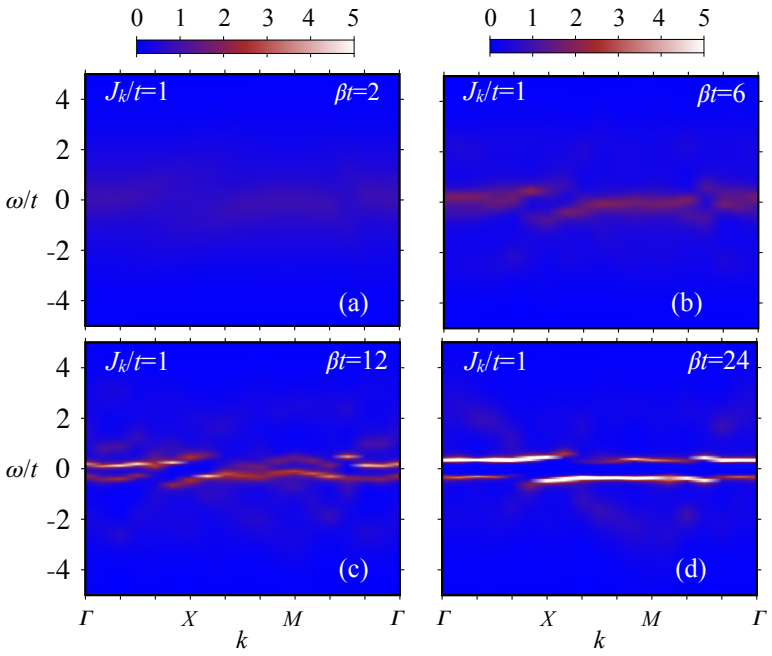


Figure 3. Temperature dependence of the composite fermion spectral function on a 12×12 lattice. Reproduced from Ref. 18.

The canonical model to study the physics of heavy fermions is the Kondo lattice model, in which charge fluctuations on the localised orbitals are suppressed. It reads,

$$\hat{H}_{KLM} = \sum_{i,j} T_{i,j} \hat{\boldsymbol{c}}_i^\dagger \hat{\boldsymbol{c}}_j + \frac{J_k}{2} \sum_i \hat{\boldsymbol{c}}_i^\dagger \boldsymbol{\sigma} \hat{\boldsymbol{c}}_i \cdot \hat{\boldsymbol{S}}_i, \tag{7}$$

where $\hat{\boldsymbol{c}}_i^\dagger = \left(\hat{c}_{i,\uparrow}^\dagger, \hat{c}_{i,\downarrow}^\dagger\right)$ is a Wannier-state spinor, $\boldsymbol{\sigma}$ the vector of Pauli matrices, and $T_{i,j}$ the hopping matrix between delocalised orbitals. It is convenient to adopt an Abrikosov representation of the spin operator, $\hat{\boldsymbol{S}}_i = \frac{1}{2} \hat{\boldsymbol{f}}_i^\dagger \boldsymbol{\sigma} \hat{\boldsymbol{f}}_i$ with $\hat{\boldsymbol{f}}_i^\dagger = \left(\hat{f}_{i,\uparrow}^\dagger, \hat{f}_{i,\downarrow}^\dagger\right)$ and constraint $\hat{\boldsymbol{f}}_i^\dagger \hat{\boldsymbol{f}}_i = 1$.

Unravelling the phases of this model is a daunting task. One limiting case where analytical progress can be made is the large-N limit. Here the SU(2) symmetry of the model is generalised to SU(N). [b] In the limit $N \to \infty$ the saddle point approximation becomes exact. In Ref. 17 we have carried out extensive calculations of the SU(N) KLM model for even values of N where the negative sign problem, for particle-hole symmetric conduction bands, can be avoided. The phase diagram is plotted in Fig. 1. In this limit the RKKY interaction turns out to be a 1/N effect and for each value of N we observe a magnetically ordered state. At N=2 the magnetic transition is continuous and it becomes first order at N=4 and beyond. In the paramagnetic phase, we see in the insets of Figs. 2(a)(b) that the

[b] Here we consider the representation of SU(N) described by a single column Young tableau with N/2 rows.

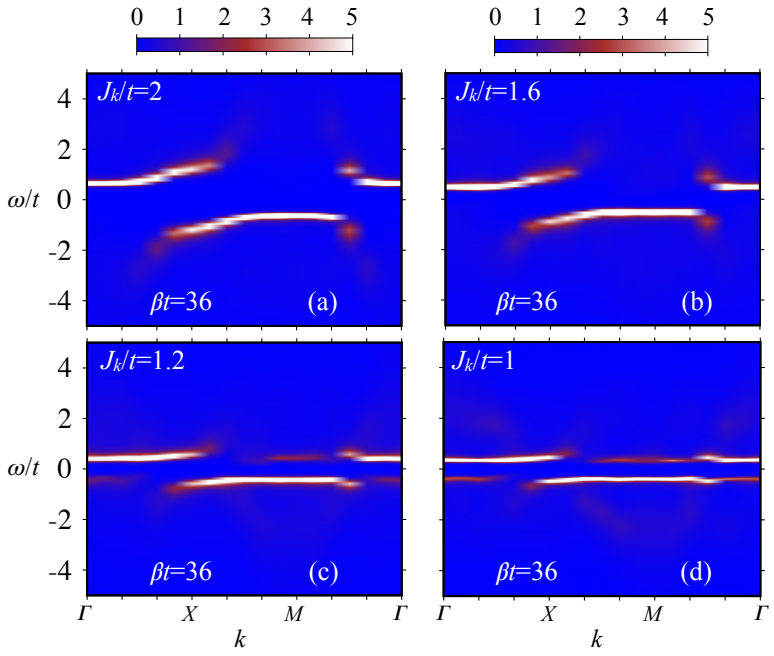


Figure 4. J_k dependence of the composite fermion spectral function 12×12 lattice. Reproduced from Ref. 18.

single particle gap as well as the quasiparticle residue smoothly extrapolate to the large-N results. This is an important result since it shows numerically that the large-N limit is a valid starting point for our understanding of the heavy fermion phase.

From the field theory, that in the large-N limit maps onto the aforementioned saddle point, we can understand the nature of the heavy fermion state. The constraint $\hat{\boldsymbol{f}}_{\boldsymbol{i}}^{\dagger}\hat{\boldsymbol{f}}_{\boldsymbol{i}} = 1$ endows the f-fermion with a U(1) local gauge charge. The key quantity in the theory is a bosonic mode $b_i(\tau)$ that carries U(1) gauge charge and physical electric charge and that couples to the f-fermion. The combined quantity $\Psi_{i,\sigma} \propto b_i(\tau) f_{i,\sigma}(\tau)$ has the quantum numbers of an electron, spin-1/2 and charge e. The heavy fermion state corresponds to the emergence of this composite fermion as an elementary excitation of the system. If the temporal fluctuations of the phase of the bosonic mode can be neglected then one can show that the field theory of the KLM involves interacting composite fermions that hybridise with the conduction electrons. Since the composite fermions carry charge e and spin-1/2 they contribute to the Luttinger volume alongside the conduction electrons. This freezing of the gauge fluctuations, or Higgs phenomena, corresponds to our present understanding of the heavy fermion phase in the KLM. In Ref. 18 we have computed the spectral function of the composite fermion operator $\Psi_{i,\sigma}$. Fig. 3 shows this quantity as a function of temperature. One observes the emergence of a well-define pole corresponding to the composite fermion. Fig. 4 shows the low temperature composite fermion spectral function across the magnetic transition at $J_k/t \simeq 1.4$. As apparent, the pole remains intact thereby showing that magnetism and Kondo screening co-exist.

To summarise, using computer simulations we have considerably progressed in our

fundamental understanding of the physics of the Kondo lattice model. In fact, we are now able to characterise precisely the heavy fermion quasiparticle in terms of the composite fermion operator. The emergence of this particle stems from a Higgs transition in which gauge fluctuations freeze out.

4 Conclusions

The above provides a glimpse of the research that we are carrying out. A key point for progressing is the ability to quickly and efficiently play with models so as to assess if they show interesting or surprising collective phenomena. Our ALF code [c] combined with access to supercomputer resources certainly allows us to do this. Ultimately, the combination of field theoretical approaches and computer simulations that guide our intuition will lead to sustained progress.

Acknowledgements

The authors gratefully acknowledge the John von Neumann Institute for Computing (NIC) for providing computing time on the GCS Supercomputer JUWELS[41] at the Jülich Supercomputing Centre (JSC).

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, Oct 1981.
2. S. White, D. Scalapino, R. Sugar, E. Loh, J. Gubernatis, and R. Scalettar, *Numerical study of the two-dimensional Hubbard model*, Phys. Rev. B, **40**, 506–516, Jul 1989.
3. F. F. Assaad and H. G. Evertz, "World-line and determinantal quantum monte carlo methods for spins, phonons and electrons", in: Computational Many-Particle Physics, H. Fehske, R. Schneider, and A. Weiße, (Eds.), vol. 739 of *Lecture Notes in Physics*, 277–356. Springer, Berlin Heidelberg, 2008.
4. S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Hybrid Monte Carlo*, Phys. Lett., **B195**, 216–222, 1987.
5. S. Beyl, F. Goth, and F. F. Assaad, *Revisiting the hybrid quantum Monte Carlo method for Hubbard and electron-phonon models*, Phys. Rev. B, **97**, 085144, Feb 2018.
6. M. Ulybyshev, S. Zafeiropoulos, C. Winterowd, and F. Assaad, *Bridging the gap between numerics and experiment in free standing graphene*, 2021, arXiv:2104.09655.
7. 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.
8. ALF Collaboration, F. F. Assaad, M. Bercx, F. Goth, A. Götz, J. S. Hofmann, E. Huffman, Z. Liu, F. Parisen Toldin, J. S. E. Portela, and J. Schwab, *The ALF (Algorithms for Lattice Fermions) project release 2.0. Documentation for the auxiliary-field quantum Monte Carlo code*, 2020, arXiv:2012.11914.

[c] `https://alf.physik.uni-wuerzburg.de`

9. D. Vollhardt, *Normal $He3$: an almost localized Fermi liquid*, Rev. Mod. Phys., **56**, no. 1, 99, Jan 1984.
10. 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**, no. 6402, 570–574, 2018.
11. D. C. Elias, R. V. Gorbachev, A. S. Mayorov, S. V. Morozov, A. A. Zhukov, P. Blake, L. A. Ponomarenko, I. V. Grigorieva, K. S. Novoselov, F. Guinea, and A. K. Geim, *Dirac cones reshaped by interaction effects in suspended graphene*, Nature Phys., **7**, no. 9, 701–704, Sep 2011.
12. S. Saremi and P. A. Lee, *Quantum critical point in the Kondo-Heisenberg model on the honeycomb lattice*, Phys. Rev. B, **75**, 165110, Apr 2007.
13. R. Toskovic, R. van den Berg, A. Spinelli, I. S. Eliens, B. van den Toorn, B. Bryant, J. S. Caux, and A. F. Otte, *Atomic spin-chain realization of a model for quantum criticality*, Nature Physics, **12**, 656–660, Apr 2016.
14. 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, Oct 2019.
15. B. Danu, M. Vojta, F. F. Assaad, and T. Grover, *Kondo Breakdown in a Spin-1/2 Chain of Adatoms on a Dirac Semimetal*, Phys. Rev. Lett., **125**, 206602, Nov 2020.
16. B. Danu, M. Vojta, T. Grover, and F. F. Assaad, *Spin chain on a metallic surface: Dissipation-induced order vs. Kondo entanglement*, 2022, arXive:2204.00029.
17. M. Raczkowski and F. F. Assaad, *Phase diagram and dynamics of the $\mathrm{SU}(N)$ symmetric Kondo lattice model*, Phys. Rev. Research, **2**, 013276, Mar 2020.
18. B. Danu, Z. Liu, F. F. Assaad, and M. Raczkowski, *Zooming in on heavy fermions in Kondo lattice models*, Phys. Rev. B, **104**, 155128, Oct 2021.
19. 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, Apr 2005.
20. 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, Jun 2016.
21. 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, Dec 2016.
22. T. Sato, F. F. Assaad, and T. Grover, *Quantum Monte Carlo Simulation of Frustrated Kondo Lattice Models*, Phys. Rev. Lett., **120**, 107201, Mar 2018.
23. 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**, no. 1, 2658, 2019.
24. F. F. Assaad and T. Grover, *Simple Fermionic Model of Deconfined Phases and Phase Transitions*, Phys. Rev. X, **6**, 041049, Dec 2016.
25. 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, 2018.
26. T. Sato and F. F. Assaad, *Quantum Monte Carlo simulation of generalized Kitaev models*, Phys. Rev. B, **104**, L081106, Aug 2021.

27. A. Götz, S. Beyl, M. Hohenadler, and F. F. Assaad, *Valence-bond solid to antiferromagnet transition in the two-dimensional Su-Schrieffer-Heeger model by Langevin dynamics*, Phys. Rev. B, **105**, 085151, Feb 2022.
28. S. Gazit, F. F. Assaad, and S. Sachdev, *Fermi Surface Reconstruction without Symmetry Breaking*, Phys. Rev. X, **10**, 041057, Dec 2020.
29. M. Neumann, J. Nyéki, B. Cowan, and J. Saunders, *Bilayer 3He: A Simple Two-Dimensional Heavy-Fermion System with Quantum Criticality*, Science, **317**, 1356, 2007.
30. A. Casey, H. Patel, J. Nyéki, B. P. Cowan, and J. Saunders, *Evidence for a Mott-Hubbard Transition in a Two-Dimensional ^{3}He Fluid Monolayer*, Phys. Rev. Lett., **90**, 115301, 2003.
31. Z. Fisk, H. R. Ott, T. M. Rice, and J. L. Smith, *Heavy-Electron metals*, Nature, **320**, no. 13, 124, 1986.
32. M. Raczkowski and F. F. Assaad, *Emergent Coherent Lattice Behavior in Kondo Nanosystems*, Phys. Rev. Lett., **122**, 097203, Mar 2019.
33. N. Tsukahara, S. Shiraki, S. Itou, N. Ohta, N. Takagi, and M. Kawai, *Evolution of Kondo Resonance from a Single Impurity Molecule to the Two-Dimensional Lattice*, Phys. Rev. Lett., **106**, 187201, May 2011.
34. S. Doniach, *The Kondo lattice and weak antiferromagnetism*, Physica B, **91**, 231, 1977.
35. J. A. Hertz, *Quantum critical phenomena*, Phys. Rev. B, **14**, no. 3, 1165–1184, Aug 1976.
36. A. J. Millis, *Effect of a nonzero temperature on quantum critical points in itinerant fermion systems*, Phys. Rev. B, **48**, no. 10, 7183–7196, Sep 1993.
37. A. Schröder, G. Aeppli, R. Coldea, M. Adams, O. Stockert, H. v. Löhneysen, E. Bucher, R. Ramazashvili, and P. Coleman, *Onset of antiferromagnetism in heavy-fermion metals*, Nature, **407**, 351, 2000.
38. S. Paschen, T. Lühmann, S. Wirth, P. Gegenwart, O. Trovarelli, C. Geibel, F. Steglich, P. Coleman, and Q. Si, *Hall-effect evolution across a heavy-fermion quantum critical point*, Nature, **432**, 881–885, Dec 2004.
39. Q. Si, S. Rabello, K. Ingersent, and J. L. Smith, *Locally critical quantum phase transitions in strongly correlated metals.*, Nature, **413**, 804, 2001.
40. T. Senthil, M. Vojta, and S. Sachdev, *Weak magnetism and non-Fermi liquids near heavy-fermion critical points*, Phys. Rev. B, **69**, 035111, Jan 2004.
41. Jülich Supercomputing Centre, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, Journal of large-scale research facilities, **5**, no. A135, 2019.

Massively Parallel Simulations of Correlated Materials

Xue-Jing Zhang, Neda Rihaisamani, Elaheh Adibi, Erik Koch, and Eva Pavarini

Institute for Advanced Simulation, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {x.zhang, n.rihaisamani, e.adibi, e.koch, e.pavarini}@fz-juelich.de

The realistic description of strongly-correlated systems is one of the great challenges in condensed-matter physics. In the last decade, impressive progress has been achieved by combining new algorithms with the power of massively-parallel supercomputers. Recent developments made it possible to systematically explore not only spectral functions but also static and dynamical responses. Furthermore, new categories of systems, such as those with important spin-orbit interaction effects, including topologically non-trivial materials, became accessible. Here we briefly describe the numerical approach behind this success.

1 Introduction

Electronic-structure theory is the basis of modern technologies such as electronics and computing. Electronic properties of materials are determined by quantum mechanics. Thus, in principle, by solving the Schrödinger equation, we should be able to predict the properties of real materials, or even design new ones with superior qualities. In practice, unfortunately, solving this equation is not easy at all. The essential complication comes from the inherent quantum many-body nature of the problem. As a result, a brute-force solution is impossible, except in the simplest cases. As an illustration let us consider a single atom of iron. Having 26 electrons, its wave function is a function of 26 times 3 coordinates. Neglecting spin, already an extremely crude representation of this function at merely 10 values of each variable would thus require storage of 10^{78} numbers. Even after reducing this number by exploiting symmetries, there is simply not enough matter available in our galaxy for building the required memory.

Given this example, electronic-structure theory seems a hopeless enterprise. Nevertheless, it is a thriving discipline. This is largely due to density-functional theory (DFT). In practice, thanks to the Kohn-Sham (KS) picture, this approach drastically simplifies the many-body problem by assuming that the electrons retain their individuality and experience the other electrons via a static mean field. In this picture electrons occupy states that extend over the whole crystal, forming the band structure of the material.

For many important classes of materials a density-functional description fails even qualitatively, however. Striking effects like the breakdown of the Fermi-liquid picture at the Mott metal-insulator transition, heavy Fermion behaviour, exotic one-dimensional Luttinger phases, or high-temperature superconductivity cannot be addressed by such a simple approach. All these materials are called *strongly-correlated* systems. This name originates from the fact that, because of the strong repulsion between the electrons, the latter lose their individuality, and the single-particle picture breaks down. Furthermore, because of the strength of the interaction, non-perturbative many-body techniques have to be used, so that very powerful computers are essential for reliable calculations. And still, calculations are restricted to quite small model systems. This means that the full Hamiltonian of a crystal has to be approximated by a small lattice Hamiltonian, which describes only (few of)

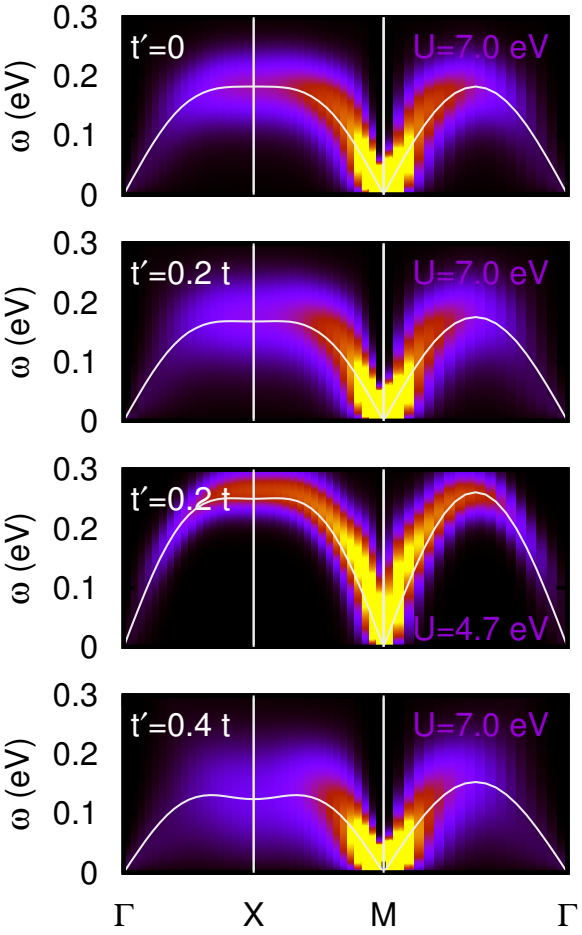


Figure 1. Dynamical response function yielding bosonic Goldstone modes for a model simulating high-temperature superconducting cuprates (reproduced from Ref. 14).

the strongly-correlated electrons. All other electrons have to be included in the calculation in an average way[1].

The modern approach to solving the many-body problem is dynamical mean-field theory (DMFT)[2–7]. It reduces the lattice Hamiltonian to a correlated impurity embedded in a self-consistent dynamical medium, which mimics the other lattice sites. This approximation simplifies the problem significantly. Still, a complicated quantum-impurity problem remains, which has to be solved. One of the most flexible methods to this aim is quantum Monte Carlo (QMC)[8–12]. Thanks to DMFT, it was possible, for the first time, to understand the physics of the Mott transition. In Mott insulators the electronic band structure loses its meaning. Instead, physics becomes more local and it is more appropriate to think about the electrons as occupying atomic-like orbitals.

In this strongly-correlated regime we find a number of fascinating ordering phenom-

ena. Most well known is anti-ferromagnetism, where spins on neighbouring lattice sites point in opposite directions. When there are many correlated orbitals, a similar ordered phase can exist: occupied orbitals on neighbouring sites point in different directions. This directionality can give rise to highly anisotropic transport properties. Coupling of spin- and orbital degrees of freedom can make transport properties strongly dependent on magnetic field. Such a mechanism is believed to be the basis of the colossal magneto-resistance effect (CMR). Like the giant magneto-resistance (GMR), this effect, once fully understood, holds the promise of, e.g., another vast increase in hard-disk capacity.

The realistic description of most interesting systems requires that both spin and orbital degrees of freedom are accounted for. This enlarges the size of the quantum impurity problem; the computational time scales with a power (typically larger than one) of the number of the degrees of freedom; which power exactly depending on the specific algorithm and problem. A further doubling of size comes from the spin-orbit interaction, which is key for the description of many families of compounds, in particular topologically non-trivial systems. Finally, complicated spatial patterns cannot be described by a single-site approach such as DMFT, which assumes that all lattice sites are equivalent. In order to add the required spatial degrees of freedom the single impurity of DMFT has to be replaced by a cluster of sites. This approach is accordingly called cluster DMFT (CDMFT). Unfortunately, treating a large cluster instead of a single site increases the already high computational cost of a calculation even further.

With an efficient parallelisation of the QMC solver and of the DMFT self-consistency loop, we can however exploit the spectacular increase in performance offered by massively parallel machines of the class of JUWELS. Thanks to advances in algorithms and supercomputers, it is now possible to reach experimental temperatures, where, before massively-parallel supercomputers, calculations of low temperature physics were limited to temperatures of about 1000 K. In addition, for lack of computer time, uncontrolled approximations had to be introduced in the model Hamiltonians. Now it is instead possible to check these approximations by explicit calculations. In short, calculations are becoming significantly more reliable and consequently gain predictive power. In most recent years, it became possible to calculate systematically statical and dynamical response functions, obtain bosonic excitation spectra (an example is shown in Fig. 1), and study spin-orbit correlated materials[13–20], included topologically non-trivial systems.

2 Method

2.1 DMFT

The Hubbard Hamiltonian is the simplest model for the description of the Mott metal-insulator transition. In the tight-binding approximation it becomes

$$\hat{H} = \varepsilon_d \sum_{\sigma i} \hat{n}_{i\sigma} - t \sum_{\sigma \langle ii' \rangle} c^\dagger_{i\sigma} c_{i'\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \tag{1}$$

where $c^\dagger_{i\sigma}$ ($c_{i\sigma}$) creates (destroys) an electron at site i and with spin σ, and $\langle ii' \rangle$ is a sum over first neighbours. Here the competing energy scales are the hopping integral, t, and the on-site Coulomb repulsion, U. The on-site energy, ε_d, instead, only modifies the chemical

potential and thus does not affect physical properties. For $U{=}0$, at half-filling, the Hamiltonian in Eq. 1 describes a metallic band. For $t{=}0$ it describes an insulating collection of disconnected atoms. Somewhere in between, at a critical value of t/U, a metal to insulator transition must occur. Here we will discuss the DMFT solution of Eq. 1 and the picture of the metal-insulator transition emerging from it. The first step consists in mapping Eq. 1 into an effective quantum-impurity model, such as the Anderson Hamiltonian

$$\hat{H}^A = \underbrace{\sum_{\mathbf{k}\sigma} \varepsilon^s_{\mathbf{k}} \hat{n}_{\mathbf{k}\sigma}}_{\hat{H}_{\rm bath}} + \underbrace{\sum_{\mathbf{k}\sigma} \left(V^s_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{d\sigma} + \text{h.c.} \right)}_{\hat{H}_{\rm hyb}} + \underbrace{\varepsilon_d \sum_\sigma \hat{n}_{d\sigma} + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow}}_{\hat{H}_{\rm imp}} . \tag{2}$$

In this model the on-site Coulomb repulsion U appears only in the impurity Hamiltonian, $\hat{H}_{\rm imp}$, while the terms $\hat{H}_{\rm bath}$ and $\hat{H}_{\rm hyb}$, describe, respectively, the bath and the bath-impurity hybridisation. In the next step, the quantum-impurity model is solved. This requires non-perturbative numerical methods, such as exact diagonalisation, the numerical renormalisation group, density-matrix renormalisation group or QMC. Here we describe the DMFT self-consistency loop for a QMC quantum-impurity solver. Solving the quantum-impurity model yields the impurity Green function $G^\sigma_{d,d}(i\nu_n)$. From the impurity Dyson equation we can calculate the impurity self-energy

$$\Sigma^\sigma_A(i\nu_n) = \left(G^{0\sigma}_{d,d}(i\nu_n)\right)^{-1} - \left(G^\sigma_{d,d}(i\nu_n)\right)^{-1}. \tag{3}$$

Next, we adopt the local self-energy approximation, i.e., we assume that the self-energy of the Hubbard model equals the impurity self-energy. The local Dyson equation is used once more, this time to calculate the bath Green function $\mathcal{G}^\sigma(i\nu_n)$, which in turn defines a new quantum-impurity model. This procedure is repeated until self-consistency is reached, i.e., the number of electrons is correct and the self-energy does not change anymore (within a given numerical accuracy). It is important to underline that self-consistency is key to the success of DMFT in describing the metal-to-insulator transition. The self-consistency loop is shown schematically in Fig. 2 for the general case.

Although the extension of DMFT to Hubbard models with many orbitals might appear straightforward, in practice it is not. The bottleneck is the solution of the generalised multi-orbital quantum-impurity problem via QMC. Despite being flexible, QMC-based approaches have limitations. These can be classified in two types. First, with increasing the number of degrees of freedom, calculations become very quickly computationally too expensive – how quickly depends on the specific QMC algorithm used and the actual implementation. Thus, going beyond a rather small number of orbitals and reaching the zero-temperature limit is unfeasible in practice. The second type of limitation is more severe. Increasing the number of degrees of freedom leads, eventually, to the infamous sign problem; when this happens, QMC calculations cannot be performed at all. In order to deal with limitations of the first type, it is crucial to restrict QMC calculations to the essential degrees of freedom; furthermore, we should exploit symmetries, develop fast algorithms and use the power of massively parallel supercomputers to reduce the actual computational time. For the second type of problems not a lot can be done; nevertheless, it has been shown that a severe sign problem might appear earlier with some basis choices than with others[12]. Although eventually we cannot escape it, this suggests that the model set-up can be used as a tool to expand the moderate sign-problem zone.

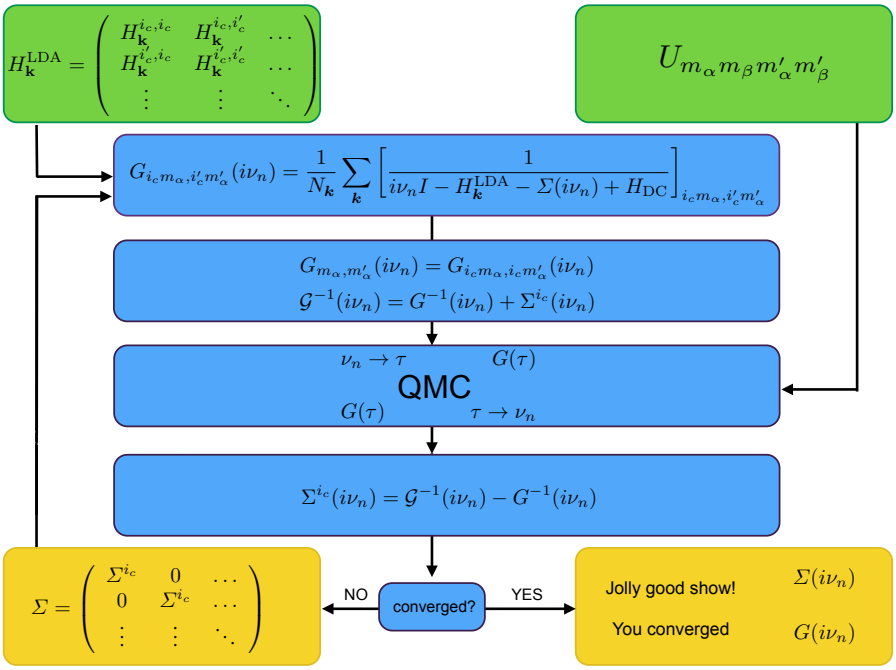


Figure 2. DMFT self-consistency loop. The one-electron Hamiltonian is built in the basis of Bloch states obtained from localised Wannier functions, e.g., in the local-density approximation (LDA); this gives $H^{\mathrm{LDA}}_{\mathbf{k}}$. The set $\{i_c\}$ labels the equivalent correlated sites inside the unit cell. The local Green-function matrix is at first calculated using an initial guess for the self-energy matrix. The bath Green-function matrix is then obtained via the Dyson equation and used to construct an effective quantum-impurity model. The latter is solved via a quantum-impurity solver, here quantum Montecarlo (QMC). This yields the impurity Green-function matrix. Through the Dyson equation the self-energy is then obtained. The procedure is repeated until self-consistency is reached (reproduced from Ref. 7).

2.2 Quantum-Impurity Solvers: Continuous-Time Quantum Monte Carlo

Here we explain the principles of the continuous-time QMC method[8–12], perhaps the most successful among QMC-based quantum-impurity solvers. In particular, we illustrate the hybridisation expansion version of the approach. To make things simpler, we consider the case of a Hubbard model in Eq. 1 with only two sites, $i = 1, 2$, and, correspondingly, a quantum-impurity model of form Eq. 2 also made of two sites, one representing the impurity, labelled with d, and one representing the bath, labelled with s. The first step consists in splitting the quantum-impurity Hamiltonian into bath ($\hat{H}_{\mathrm{bath}}$), hybridisation ($\hat{H}_{\mathrm{hyb}}$) and local ($\hat{H}_{\mathrm{loc}}$) terms

$$\hat{H}^{\mathrm{A}} = \underbrace{\varepsilon_s \sum_\sigma \hat{n}_{s\sigma}}_{\hat{H}_{\mathrm{bath}}} \underbrace{-t \sum_\sigma \left(c^\dagger_{d\sigma} c_{s\sigma} + c^\dagger_{s\sigma} c_{d\sigma} \right)}_{\hat{H}_{\mathrm{hyb}}} + \underbrace{\varepsilon_d \sum_\sigma \hat{n}_{d\sigma} + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow}}_{\hat{H}_{\mathrm{loc}}} . \qquad (4)$$

Next, we write the partition function Z as a perturbation series in the hybridisation. To this end, we define $\hat{H}_0 = \hat{H}_{\rm bath} + \hat{H}_{\rm loc}$ and rewrite the partition function as

$$Z = \mathrm{Tr}\left(e^{-\beta(\hat{H}_0-\mu\hat{N})}\hat{V}(\beta)\right), \tag{5}$$

where the operator $\hat{V}(\beta)$ is given by

$$\hat{V}(\beta) = e^{\beta(\hat{H}_0-\mu\hat{N})}\, e^{-\beta(\hat{H}_0+\hat{H}_{\rm hyb}-\mu\hat{N})} = \sum_m \underbrace{\int_0^\beta d\tau_1 \cdots \int_{\tau_{m-1}}^\beta d\tau_m}_{\int d\tau^m} \underbrace{(-1)^m \prod_{l=m}^1 \hat{H}_{\rm hyb}(\tau_l)}_{\hat{O}^m(\tau)}, \tag{6}$$

and

$$\hat{H}_{\rm hyb}(\tau_l) = e^{\tau_l(\hat{H}_0-\mu\hat{N})}\, \hat{H}_{\rm hyb}\, e^{-\tau_l(\hat{H}_0-\mu\hat{N})}. \tag{7}$$

In this expansion, the only terms that contribute to the trace are even order ones ($m = 2k$) and they are products of impurity (d) and bath (s) creator-annihilator pairs. Thus

$$\int d\tau^{2k} \longrightarrow \int d\tau^k \int d\bar{\tau}^k \quad \text{and} \qquad \hat{O}^{2k}(\tau) \longrightarrow \sum_{\sigma,\bar{\sigma}} \hat{O}^{2k}_{\sigma,\bar{\sigma}}(\tau,\bar{\tau}), \tag{8}$$

where

$$\hat{O}^{2k}_{\sigma,\bar{\sigma}}(\tau,\bar{\tau}) = (t)^{2k} \prod_{i=1}^k \left(c^\dagger_{d\bar{\sigma}_i}(\bar{\tau}_i) c_{s\bar{\sigma}_i}(\bar{\tau}_i) c^\dagger_{s\sigma_i}(\tau_i) c_{d\sigma_i}(\tau_i)\right). \tag{9}$$

The vector $\sigma = (\sigma_1, \sigma_2, ..., \sigma_k)$ gives the spins $\{\sigma_i\}$ associated with the k impurity annihilators at imaginary times $\{\tau_i\}$, while $\bar{\sigma} = (\bar{\sigma}_1, \bar{\sigma}_2, ..., \bar{\sigma}_k)$ gives the spins $\{\bar{\sigma}_i\}$ associated with the k impurity creators at imaginary times $\{\bar{\tau}_i\}$. It follows that the local and bath traces can be decoupled and the partition function can be rewritten as

$$\frac{Z}{Z_{\rm bath}} = \sum_k \int d\tau^k \int d\bar{\tau}^k \sum_{\sigma,\bar{\sigma}} d^k_{\bar{\sigma},\sigma}(\tau,\bar{\tau})\, \mathbf{t}^{\mathbf{k}}_{\sigma,\bar{\sigma}}(\tau,\bar{\tau}), \tag{10}$$

$$d^k_{\bar{\sigma},\sigma}(\tau,\bar{\tau}) = \frac{t^{2k}}{Z_{\rm bath}}\, \mathrm{Tr}_{\rm bath}\left(e^{-\beta(\hat{H}_{\rm bath}-\mu\hat{N}_s)} \mathcal{T} \prod_{i=k}^1 c^\dagger_{s\sigma_i}(\tau_i) c_{s\bar{\sigma}_i}(\bar{\tau}_i)\right), \tag{11}$$

$$t^k_{\sigma,\bar{\sigma}}(\tau,\bar{\tau}) = \mathrm{Tr}_{\rm loc}\left(e^{-\beta(\hat{H}_{\rm loc}-\mu\hat{N}_d)} \mathcal{T} \prod_{i=k}^1 c_{d\sigma_i}(\tau_i) c^\dagger_{d\bar{\sigma}_i}(\bar{\tau}_i)\right), \tag{12}$$

where $Z_{\rm bath} = 1 + 2e^{-\beta(\varepsilon_s-\mu)} + e^{-2\beta(\varepsilon_s-\mu)}$. The trace involving only bath operators is simple to calculate, since $\hat{H}_{\rm bath}$ describes an independent-electron problem for which Wick's theorem holds. It is given by the determinant

$$d^k_{\bar{\sigma},\sigma}(\tau,\bar{\tau}) = \det\left(\mathcal{F}^k_{\bar{\sigma},\sigma}(\tau,\bar{\tau})\right), \tag{13}$$

of the $k\times k$ non-interacting hybridisation-function matrix, with elements

$$\left(\mathcal{F}^k_{\bar{\sigma},\sigma}(\tau,\bar{\tau})\right)_{i',i} = \mathcal{F}^0_{\bar{\sigma}_{i'},\sigma_i}(\bar{\tau}_{i'}-\tau_i), \tag{14}$$

where

$$\mathcal{F}^0_{\bar{\sigma},\sigma}(\tau) = \delta_{\bar{\sigma},\sigma} \frac{t^2}{1+e^{-\beta(\varepsilon_s-\mu)}} \times \begin{cases} -e^{-\tau(\varepsilon_s-\mu)} & \tau > 0, \\ +e^{-(\beta+\tau)(\varepsilon_s-\mu)} & \tau < 0. \end{cases} \tag{15}$$

The calculation of the local trace is in general more complicated. In the case discussed here, the Hamiltonian does not flip spins. Thus only terms with an equal number of creation and annihilation operators *per spin* contribute to the local trace, and we can express the partition function in expansion orders per spin, k_σ. This yields[12]

$$\frac{Z}{Z_{\rm bath}} = \left(\prod_\sigma \sum_{k_\sigma=0}^{\infty} \int d\tau_\sigma^{k_\sigma} \int d\bar{\tau}_\sigma^{k_\sigma} \right) d^k_{\bar{\sigma},\sigma}(\tau,\bar{\tau}) t^k_{\sigma,\bar{\sigma}}(\tau,\bar{\tau}) \tag{16}$$

where the vectors $\sigma = (\sigma_\uparrow, \sigma_\downarrow)$ and $\bar{\sigma} = (\bar{\sigma}_\uparrow, \bar{\sigma}_\downarrow)$ have $(k_\uparrow, k_\downarrow)$ components, and for each k_σ component $\sigma_i = \bar{\sigma}_i = \sigma$. Thus

$$t^k_{\sigma,\bar{\sigma}}(\tau,\bar{\tau}) = {\rm Tr}_{\rm loc} \left(e^{-\beta(\hat{H}_{\rm loc}-\mu\hat{N}_d)} \mathcal{T} \prod\nolimits_\sigma \prod\nolimits_{i=k_\sigma}^{1} c_{d\sigma}(\tau_{\sigma_i}) c^\dagger_{d\sigma}(\bar{\tau}_{\bar{\sigma}_i}) \right). \tag{17}$$

The latter can be calculated analytically. To do this, first we parametrise all configurations for a given spin via a timeline $[0,\beta)$ plus a number of creator/annihilator pairs which define segments on the timeline. At zeroth order two possible configurations exist per spin, an empty timeline, which corresponds to the vacuum state $|0\rangle$, and a full timeline, which corresponds to the state $c^\dagger_{d\sigma}|0\rangle$. A given configuration yields, at order $k = k_\uparrow + k_\downarrow$

$$t^k_{\sigma,\bar{\sigma}}(\tau,\bar{\tau}) = \left(\prod_\sigma s_\sigma^{k_\sigma} \right) e^{-\sum_{\sigma\sigma'}((\varepsilon_d-\mu)\delta_{\sigma\sigma'} + \frac{U}{2}(1-\delta_{\sigma,\sigma'}))l_{\sigma,\sigma'}}, \tag{18}$$

where $l_{\sigma,\sigma'}$ is the length of the overlap of the τ segments for spins σ and σ', respectively, while $s_\sigma = {\rm sign}(\tau_{\sigma_1} - \bar{\tau}_{\sigma_1})$ is the fermionic sign. Possible configurations at order $k = 0, 1, 2$ are shown in Fig. 3. Taking all k values into account, the partition function can be expressed as the sum over all configurations $\{c\}$, i.e., in short

$$Z = \sum_c w_c = \sum_c |w_c| \ {\rm sign}\ w_c. \tag{19}$$

In a compact form, we can write $w_c = d\tau_c \ d_c \ t_c$ where $d\tau_c = \prod_\sigma \prod_i^{k_\sigma} d\tau_{\sigma_i} d\bar{\tau}_{\bar{\sigma}_i}$, and d_c and t_c are the bath and local traces for the configuration c. This expression of the partition function shows that we can interpret $|w_c|$ as the sampling weight of configuration c. A generic observable $\hat{O}$ can then be obtained as the Monte Carlo average on a finite number of configurations N_c

$$\langle \hat{O} \rangle = \frac{\sum_c \langle \hat{O} \rangle_c |w_c| \ {\rm sign}\ w_c}{\sum_c |w_c| \ {\rm sign}\ w_c} \approx \frac{\frac{1}{N_c} \sum_c^{N_c} \langle \hat{O} \rangle_c \ {\rm sign}\ w_c}{\frac{1}{N_c} \sum_c {\rm sign}\ w_c}. \tag{20}$$

The term $\frac{1}{N_c} \sum_c {\rm sign}\ w_c$ in the denominator is the average fermionic sign. When this is small, much longer runs are required to obtain data of the same quality; eventually the computational time can become so long that the calculation is unfeasible – in these cases we have a sign problem. In practice, the QMC simulation starts from a random

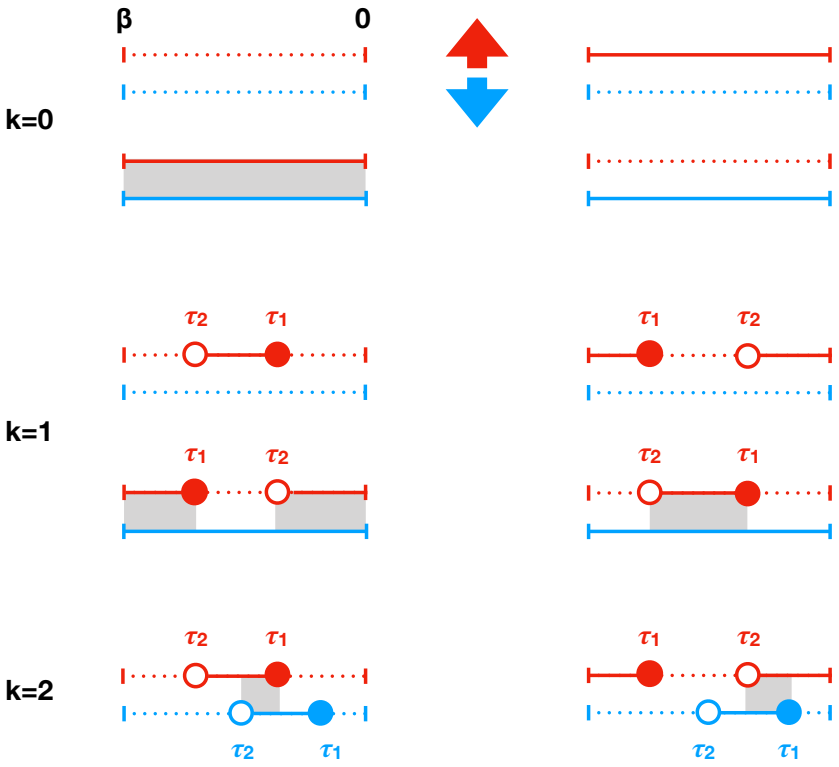


Figure 3. Representative configurations contributing to the local trace at zeroth, first and second order (reproduced from Ref. 7). The timelines for spin up are red and those for spin down are blue. The filled circles correspond to the insertion of a creator (time τ_1), and the empty circles to the insertion of an annihilator (time τ_2). Dotted lines represent the vacuum state for a given spin, full lines the occupied state. The grey boxes indicate the regions in which $l_{\uparrow,\downarrow} \neq 0$.

configuration c. Next we propose an update $c \to c'$. Within the Metropolis algorithm, the acceptance ratio is

$$R_{c\to c'} = \min\left(1, \frac{p_{c'\to c}}{p_{c\to c'}} \frac{|w_{c'}|}{|w_c|}\right), \tag{21}$$

where $p_{c\to c'}$ is the proposal probability for the update $c \to c'$. In the approach described here, known as *segment solver*, the basic updates are addition and removal of segments, antisegments (segments winding over the borders of the timeline, see Fig. 3), or complete lines. As example, let us consider the insertion of a segment for spin σ. A segment is made by a creator and an annihilator. The creator is inserted at time $\tau_{\rm in}$; the move is rejected if $\tau_{\rm in}$ is in a region where a segment exists. If created, the segment can have at most length $l_{\rm max}$, given by the distance between $\tau_{\rm in}$ and the time at which the next creator is, hence

$$p_{c\to c'} = \frac{d\bar{\tau}}{\beta} \frac{d\tau}{l_{\rm max}} . \tag{22}$$

The proposal probability of the reverse move (removing a segment) is instead given by the inverse of the number of existing segments

$$p_{c' \to c} = \frac{1}{k_\sigma + 1} \,. \tag{23}$$

The acceptance ratio for the insertion of a segment becomes then

$$R_{c \to c'} = \min\left(1, \frac{\beta l_{max}}{k_\sigma + 1} \left| \frac{d_{c'}}{d_c} \frac{t_{c'}}{t_c} \right| \right). \tag{24}$$

In multi-orbital and multi-site systems the updates are more complex due to the coupling between flavours, but the general principle holds. Dynamical responses require sampling three independent imaginary times, sizably increasing the computational cost.

3 Concluding Remarks

We have described the key steps behind numerical simulation of correlated electrons in materials. In recent years the ability of calculating systematically static and dynamical response functions (see Fig. 1) and to study spin-orbit materials (see Fig. 4), included topologically non-trivial matter, has opened new horizons in understanding correlated systems. This is not only key for basic physics but also for practical applications. Thanks to their high sensitivity to external parameters such as pressure, temperature or magnetic/electric fields, correlated systems hold the promise to play a key role in future electronics and information technology. Theoretically the challenge is to solve systematically ever larger models, in order to eventually reach predictive power. Experience so far tells us that ultimately this can only be achieved with the help of large-scale simulations on novel generations of massively parallel supercomputers.

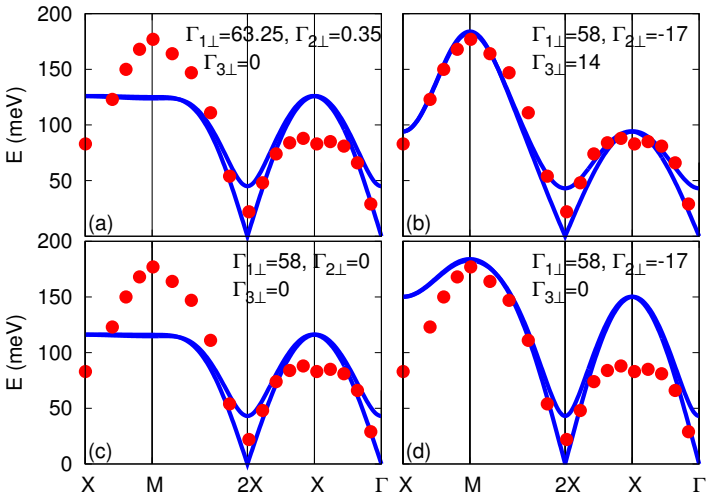


Figure 4. Spin waves in the iridates (reproduced from Ref. 13).

Acknowledgements

Initial code development and benchmark computations were performed with a grant of computer time provided by the VSR/JARA of 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 through the John von Neumann Institute for Computing (NIC) on the GCS.

References

1. E. Pavarini, Riv. Nuovo Cimento, Springer-Nature **44**, 597, 2021.
2. W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324, 1989.
3. A. Georges, G. Kotliar, W. Krauth, M. J. Rozenberg, Rev. Mod. Phys. **68**, 12, 1996.
4. M. A. Korotin, A. O. Anokhin, V. I. Anisimov, A. I. Poteryaev, and G. Kotliar, J. Phys. Cond. Mat. **9**, 7539, 1998.
5. A. I. Lichtenstein and M. I. Katsnelson, Phys. Rev. B **57**, 6884, 1998.
6. E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein (eds.), *DMFT at 25: Infinite Dimensions*, Modeling and Simulation, Vol. 4, Verlag des Forschungszentrum Jülich, 2014, ISBN 978-3-89336-953-9.
7. E. Pavarini, E. Koch, A. Lichtenstein, and D. Vollhardt (eds.), *DMFT: From Infinite Dimensions to Real Materials*, Modeling and Simulation, Vol. 8, Verlag des Forschungszentrum Jülich, 2018, ISBN 978-3-95806-313-6.
8. A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Phys. Rev. B **72**, 035122, 2005.
9. P. Werner and A. J. Millis, Phys. Rev. B **74**, 155107,2006.
10. E. Gull, P. Werner, A. J. Millis, and M. Troyer, Phys. Rev. B **76**, 235123, 2007.
11. E. Gorelov, M. Karolak, T. O. Wehling, F. Lechermann, A. I. Lichtenstein, and E. Pavarini, Phys. Rev. Lett. **104**, 22641, 2010.
12. A. Flesch, E. Gorelov, E. Koch, E. Pavarini, Phys. Rev. B **87**, 195141, 2013.
13. G. Zhang and E. Pavarini, Phys. Rev. B **104**, 125116, 2021.
14. J. Musshoff, A. Kiani, and E. Pavarini, Phys. Rev. B **103**, 075136, 2021.
15. X-J. Zhang, E. Koch and E. Pavarini, Phys. Rev. B **102**, 035113, 2020.
16. G. Zhang and E. Pavarini, Phys. Rev. B **101**, 205128, 2020.
17. G. Zhang and E. Pavarini, Phys. Rev. B **99,** 125102, 2019.
18. J. Musshoff, G. Zhang, E. Koch, E. Pavarini, Phys. Rev. B **100**, 045116, 2019.
19. E. Pavarini and E. Koch (eds.), *Simulating Correlations on Computers*, Modeling and Simulation, Vol. 11, Verlag des Forschungszentrum Jülich, 2021, ISBN 978-3-95806-529-1.
20. E. Pavarini and E. Koch (eds.), *Topology, Entanglement, and Strong Correlations*, Modeling and Simulation, Vol. 10, Verlag des Forschungszentrum Jülich, 2020, ISBN 978-3-95806-466-9.

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 is a generic term for molecular systems where typically intra-molecular or inter-molecular degrees of freedom strongly fluctuate so that intra- and inter-molecular entropies contribute in the same way to the overall free energy of the system as energy. This decisive relevance of entropy also defines the thermal energy k_BT as the relevant energy scale, orders of magnitude smaller than typical electronic excitations. This relevant free energy scale can be connected to small characteristic energy densities, making systems soft, i.e. yielding to very small forces. Biological tissue is of course the most prominent example of soft matter. As such most soft matter typically consists of large molecules, e.g. polymers, which can be of both biological or synthetic origin. In general one is dealing with an enormously large class of materials/systems covering such diverse fields as biomaterials, electronic materials, medical applications, food or even simple commodities, illustrating that these materials are of universal importance. And, their importance is even growing in view of the current societal and technological needs. From a theoretical point of view the delicate interplay between small local changes and potentially resulting huge macroscopic response and the interplay of universal and materials specific features pose special challenges and opportunities. Not surprising for many years computer simulations of soft matter have played a central role in research and also have benefited significantly from support by NIC and its predecessor organisation HLRZ.

This year's NIC symposium volume contains two contributions concerning co-polymer systems at interfaces on the one hand and about transport in nanowires and confined colloidal systems on the other. The latter also illustrates the role soft matter can play as classical model systems for atomic analogs.

The first contribution *Controlling the Orientation of Mesophases by Processing: Insights from Computer Simulation* by Niklas Blagojevic and Marcus Müller deals with the importance of sample history for structure formation in block-copolymer systems. Different species of long polymer chains usually are not miscible, since very small interaction preferences lead to macroscopic phase separation. This complication for blending materials is turned into a design concept in the case of block copolymers. There, a polymer chain contains connected blocks of different chemical structure, in the present case two such blocks. The macroscopic phase segregation only can happen on a microscopic level and scientifically and technologically important microstructured systems appear. Structures range from lamellar systems to complicated plumbers nightmare morphologies depending on segregation strength and relative length of the blocks. During structure formation systems can undergo several transitions between metastable states, which makes equilibration difficult. It also means that the outcome of an experiment often delicately depends on details of the sample history. This complication, however, in turn also can be used to design materials with tailored properties. The chapter by Blagojevich and Müller addresses

exactly such questions by employing highly optimised software for a soft core polymer model, which is well suited to take advantage of acceleration through Graphical Processor Units (GPUs). They look at a surface active systems upon solvent evaporation as well as at consequences of mechanical deformation. This allows them not only to follow structure formation but also to calculate effective free energy barriers between metastable states.

The second contribution *Transport in Nanowires and Colloidal Systems and Crystal Growth* by Konstantin Beck *et al.* presents a number of simulations studies elucidating similarities and differences between a one dimensional arrangement of colloidal particles and nanowires. Their studies on two seemingly very different classes of systems are presented. Nanowires, i.e. systems where junctions are formed by very few or even single atoms are very interesting nanotechnology objects. Besides electronic also classical phonon contributions play an important role for heat transport along these wires. Here the focus is on the phononic part of heat transport. For that methodologies and problems are actually often quite similar to characteristic questions for soft matter subject to confinement. Beck *et al.* demonstrate heat transport details as a function of the length of monoatomic nanowires and showed that phonon phonon scattering becomes less important for longer nanowires. In contrast transport of and in colloidal systems has to take into account that one is dealing with colloidal suspensions. For driven transport in a linear channel of a mixture of colloids of two different sizes this can lead to drift induced lane formation, and thus a nonequilibrium phase segregation, as shown by the simulations of the Nielaba group. While these studies focused on spherical colloids, anisotropy can become very important for dynamics as well as structure formation upon density variation. Here, symmetric dumbell colloids are studied. As shown here the dumbell structure couples to the different diffusion mechanisms, namely parallel and perpendicular to the axis as well as rotational. Furthermore in collaboration with an experimental colleague (Wittemann, Konstanz) structure formation in 2-d layers has been investigated. Beyond that self propelled colloidal rods as a minimal model systems for active matter as well as colloidal crystal growth has been studied.

Though very different, these two contributions reveal the power and the need of high level computer simulations for soft matter problems.

Controlling the Orientation of Mesophases by Processing: Insights from Computer Simulation

Niklas Blagojevic and Marcus Müller

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

Using a soft, coarse-grained, particle-based model in conjunction with the GPU-accelerated MC program SOMA, we study the kinetics of structure formation and ordering in a thin film of cylinder-forming diblock copolymers and provide direct microscopic insights into the complex, nonequilibrium changes of domain topology by large-scale computer simulations. Strategies for controlling the orientation of the incipient, microphase-separated structure after solvent evaporation and the kinetics of grain coarsening in the late stages are investigated and rationalised *via* phenomenological concepts.

1 Introduction

Polymers are formed by covalently bonding monomeric repeat units into a long, flexible macromolecule. These macromolecules adopt random-walk-like configurations in a dense melt. Small structural differences of monomer species often result in phase separation. In the following, we consider a copolymer that is comprised of two blocks of different monomer species, A and B. Since the two blocks are covalently connected into a single macromolecule, they cannot macroscopically phase separate but, instead, form a spatially modulated morphology. In equilibrium, periodically alternating lamellae of A and B form in case of equal fractions, $f = 1/2$, of A and B species or cylinders of the minority species that arrange on a hexagonal lattice in case of larger asymmetry of composition – more complex phases[1] additionally occur but will not be considered in the following.

The free-energy landscape of this soft-matter system, however, is rugged, *i. e.* there are many local minima that represent metastable states such as *e. g.* alternate periodic structures, multi-grain morphologies, or defects, such as *e. g.* dislocations or disclinations[2,3]. These metastable basins on the free-energy landscape are separated by free-energy barriers that scale like $\sqrt{\bar{N}}k_{\mathrm{B}}T$, where $k_{\mathrm{B}}T$ and $\bar{N}$ denote the thermal energy scale and the invariant degree of polymerisation, respectively. Thus, for long polymers, these metastable states are long-lived and are relevant for application properties. In the following, we consider two competing structures that only differ in their orientation with respect to external boundaries or fields.

Which (meta)stable structure forms is chiefly dictated by the processing of the sample[4]. Either, the experimentally observed structure is formed in the course of structure formation or it is dictated by subsequent application of external fields. We illustrate both processing strategies by studying solvent evaporation from a thin film and mechanical deformation.

The contribution is arranged as follows: In Sec. 2 we introduce a highly coarse-grained model and the Monte-Carlo (MC) simulation techniques that are well suitable for the JUWELS Booster. In the subsequent Sec. 3, we discuss the two processing strategies – solvent evaporation to control the early stages and mechanical deformation to induce sub-

sequent grain-boundary motion and coarsening. The article closes with a brief summary and outlook.

2 Models and Techniques

We consider the formation and reorientation of a spatially periodic morphology. The periodicity of the morphology is set by the molecular size, quantified by the mean-squared end-to-end distance, R_e^2, of the random-walk-like, single-chain configurations. Due to the fractal shape of the molecules, one long, flexible molecule interacts with many neighbours, $\sqrt{\bar{N}} \gg 1$. Thus, studying the collective dynamics poses a serious computational challenge because of the large number of molecules and the long time scales. One way to mitigate the computational expense of studying changes of the block copolymer morphology consist in using a highly coarse-grained, top-down model that only accounts for the relevant interactions – chain connectivity, repulsion between unlike segment species, A and B, and a limited compressibility of the dense polymer liquid[5].

Since one highly coarse-grained segment represents a large number of monomeric repeat units, the length distribution of such a subchain is Gaussian, and we use a discretised Edwards Hamiltonian to describe the universal aspects of the macromolecular architecture. The centres of mass of highly coarse-grained segments can overlap, even if their constituent monomeric repeat units exhibit harsh excluded-volume interactions. Therefore we use soft, pairwise, nonbonded interactions to describe the repulsion between unlike segment species (*i. e.* thermodynamic incompatibility) and the limited compressibility of the dense polymer fluid. The Flory-Huggins parameter, χN, quantifies the repulsion between blocks, whereas the parameter, κN, is related to the inverse, isothermal compressibility of the liquid. In the limit of fine discretisation, N, of the chain contour, the bonded interactions are much stronger than the computationally expensive, nonbonded interactions.

An experimentally relevant, large invariant degree of polymerisation, $\bar{N}$, can be achieved by increasing the number, $N \sim \bar{N}$, of segments per chain at constant segment number density or by increasing the segment number density, $\rho_0 \sim \sqrt{\bar{N}}$, at fixed chain-contour discretisation. The latter strategy is computationally much more efficient and results in a highly coarse-grained model with soft, nonbonded interactions and overlapping segments. The large segment density can be coped with by (i) evaluating the nonbonded interactions of a segment with the large number of its neighbours *via* a collocation grid rather than a neighbour list and (ii) by employing the single-chain-in-mean-field algorithm[6] that replaces the weakly changing but computationally expensive nonbonded interactions by external fields for a short time period (*i. e.* 1 MCS). The soft, coarse-grained polymer model is efficiently implemented in the MC program SOft coarse grained Monte-carlo Acceleration (SOMA)[7] that parallelises over multiple Graphics Processing Unit (GPU)s.

3 Results

3.1 Solvent Evaporation

In the following, we discuss the structure formation after evaporation of a volatile solvent from a diblock polymer film[8]. The solvent-swollen polymer film in contact with the gas

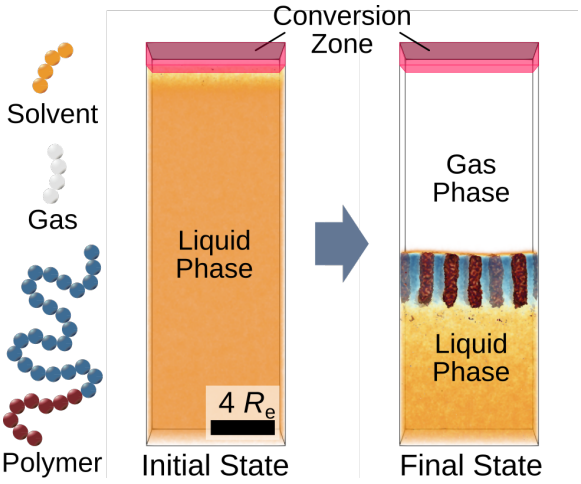
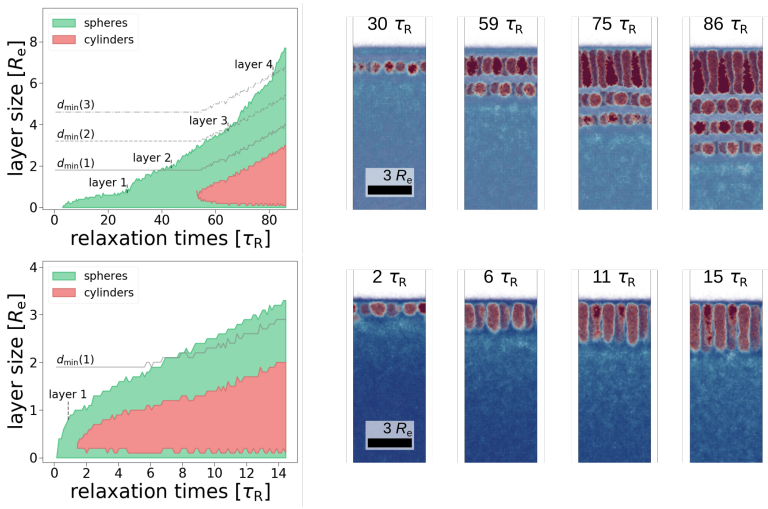


Figure 1. Left: Molecular architectures of solvent, gas, and diblock copolymer, $f_A = 0.25$, are depicted. Schematic illustration of the simulation box in the initial (middle image) and final stage (right panel) of solvent evaporation. In the course of the simulation, solvent particles that reach the conversion zone on top of the simulation box or a distance, d_{vap}, above the film surface are converted to gas particles. Adapted with permission from Ref. 8. Copyright 2022 American Chemical Society.

phase is sketched in Fig. 1. It is comprised of cylinder-forming diblock copolymers, solvent and gas molecules. Copolymer and gas are strongly incompatible and segregate into a solvent-swollen polymer film and a gas domain that are separated by a narrow surface. The volatile solvent, which has been initially dissolved in the polymer film, diffuses across the narrow film surface and is converted into gas molecules. This alchemical conversion occurs either at a specified distance, d_{vap}, above the film surface or at the top of the simulation cell. In the course of solvent evaporation, the film thickness shrinks and the film surface moves downward with velocity, v_i. The ratio between this retraction velocity, chain extension, R_{e}, and the self-diffusion coefficient, D, of the polymer defines the Péclet number, Pe$\equiv v_i R_{\mathrm{e}}/D$, which adopts a value of order 10^{-4} in the simulation, in agreement with experimental estimates.

As the film thickness decreases, the polymer concentration rises inside the film, the effective incompatibility between the A and B blocks of the copolymer increases, and microphase separation commences. A perpendicular orientation of the cylindrical morphology is sought for membrane applications[9], where the minority component of the diblock copolymer forms channels that are narrowly distributed in diameter and can be tailored by varying the molecular weight of the diblock copolymer. The goal of our project consists in understanding how the orientation of the self-assembled cylindrical morphology can be controlled by the thermodynamic and kinetic processing conditions, such as the selectivity of the gas phase for one of the components of the block copolymer or the evaporation rate.

In the course of solvent evaporation, a density gradient inside the film builds up – the density of the polymer at the surface is increased ("polymer skin") whereas the density of the volatile solvent is lower at the surface than at the bottom of the film. Thus, the self-assembly of the copolymers starts at the film surface. Once the polymer concentration, ϕ_P,



system	f_A	$\chi_{AS}N_0$	d_{vap} [R_e]	$\rho_{P,ODT}$ [ρ/ρ_0]	$\rho_{P,OOT}$ [ρ/ρ_0]
top	7/32	14	14	0.70	0.82
bottom	9/32	20	8	0.57	0.70

Figure 2. Left: Time evolution of the positions and thicknesses of the layers, in which micelles form (green, $\phi_P > \phi_{ODT}$) and in which micelles transform into cylinders (red, $\phi_P > \phi_{OOT}$) for two systems. Observed layer formations are marked, the integer indicating the total number of layers at the respective time. Right: Snapshots of the minority-block density at different times, as indicated in the key in units of the diffusion time, τ_R. The table below summarises the difference between the two systems investigated in the top and bottom panels. Adapted with permission from Ref. 8. Copyright 2022 American Chemical Society.

crosses the first threshold, ϕ_{ODT}, spherical micelles form in the disordered solution. Only when the polymer concentration increases further to ϕ_{OOT} do these micelles transform into cylindrical domains that correspond to the equilibrium structure of the "dry" film, *i. e.* in the absence of solvent. This transformation from micellar order to cylindrical order may occur either (i) *via* incorporating copolymers that were dissolved in the solvent-rich matrix between the micelles or (ii) by fusing of neighbouring micelles. Our approach is to rationalise the interplay between processing conditions and orientation of the self-assembled microphase *via* the laterally averaged composition profiles inside the polymer film during the early stages.

Two illustrative examples are presented in Fig. 2. The graph on the left-hand side presents the time evolution of the layer positions, in which the thresholds, ρ_{ODT} for micelle formation and ρ_{OOT} for the transformation to cylinders are passed, by green and red shaded areas, respectively. Images on the left depict the morphologies at various times, in units of the time, τ_R, that a copolymer needs to diffuse its end-to-end distance, R_e. In the first system, the polymer density increases above the first threshold, and a layer of spherical micelles forms. Subsequently, the thickness of the surface layer, in which micelles are thermodynamically stable, increases but the threshold for transforming micelles into cylinders is nowhere reached. Thus, a second layer of micelles forms. Eventually, at $t \approx 53\tau_R$,

ϕ_{OOT}, is passed at the surface, and the spherical micelles closest to the surface elongate into cylinders. These cylinders grow perpendicular to the surface and push the subjacent micelle layers downward. As time progresses, both layers become thicker but the layer, in which micelles self-assemble from the disordered phase, increases faster than the inner layer, where micelles can transform into thermodynamically stable cylinders. Thus, the number of micelle layers in front of the elongating perpendicular cylinders increases.

If we decrease the asymmetry of the diblock copolymer, $f_A = 7/32 \rightarrow 9/32$, both thresholds, ϕ_{ODT} and ϕ_{OOT}, decrease and the formation of cylindrical domains is thermodynamically favoured. Moreover, a decrease of the distance, d_{vap}, between the film surface and the position, above which solvent is converted into gas, increases the evaporation rate. In this case, both layers thicken at comparably high rate. Thus, the micelles are transformed into perpendicularly oriented cylinders before an additional layer of micelles can from in front of the cylinder heads. Note that the high rate also lowers the probability of lateral fusion of micelles into cylinders parallel to the film surface because such a thermally activated process has to overcome a free-energy barrier that increases like $\sqrt{N}k_{\mathrm{B}}T$ and gives rise to an induction time.

Our simulation study indicates that the thermodynamic and kinetic parameters of this processing window are (i) the surface selectivity and (ii) the evaporation rate. Gas or solvent should be slightly favourable to the matrix-forming, majority block of the copolymer in order to compensate the entropic enrichment of the minority block at the film surface and suppress a parallel layering of the composition. A high evaporation rate, in turn, gives rise to faster propagation of the ordering away from the surface, leaving insufficient time for micelles to laterally fuse into parallel cylinders. Moreover, a high evaporation rate results in a steep gradient of the polymer density inside the polymer film so that the thickness of the layer with $\phi_{\mathrm{OOT}} < \phi_P < \phi_{\mathrm{ODT}}$ in front of the cylinder heads remains small and no additional layer of micelles can form.

3.2 Grain-Boundary Motion in Response to Deformation

In equilibrium, cylinders of the minority component arrange on an hexagonal lattice with the preferred cylinder spacing. In a macroscopically extended film, however, defects of the morphology may arise in the course of structure formation, *i. e.* due to the stochasticity of the self-assembly, some grains, in which the cylindrical domains are not aligned perpendicular to the film surface, are expected to form. Grains with different orientations are separated by grain boundaries, and the motion of these boundaries dictates the kinetics of structure formation (coarsening) in the late stages. In the following, we consider the motion of a boundary between two domains that are oriented at an angle of 90 °.

As the final solvent leaves the microphase-separated film, the film thickness decreases, whereas the lateral dimensions remain unaltered. Such a uniaxial deformation affects grains with different orientations in different ways: If a grain is perpendicularly oriented to the film surface, the uniaxial shrinkage by a factor, $1 - \sigma_z$, along the cylinder axis will not alter the cylinder spacing. For a grain with a parallel orientation to the film surface, in turn, the mechanical deformation will distort the hexagonal lattice of cylinders and, therefore, increase the free-energy density of this grain, $\Delta f = f_{\parallel} - f_{\perp} > 0$. Thus, the parallel oriented grain will shrink and the perpendicularly oriented domain of cylinders will grow by the motion of the grain boundary. In order to focus on grain coarsening *via* the motion of

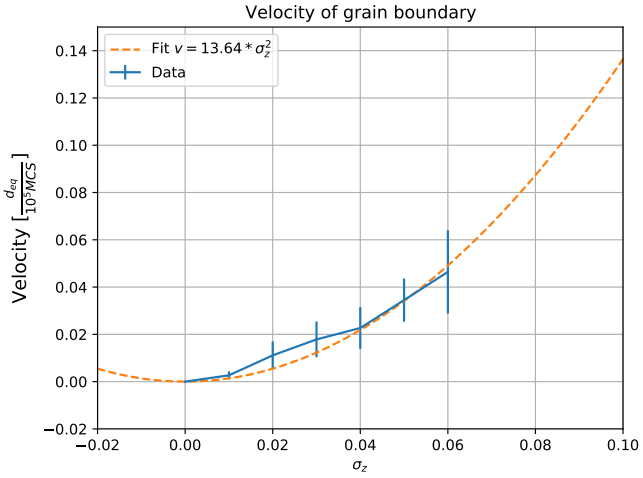


Figure 3. Measured velocity of a grain boundary as a function of the compression factor, σ_z, away from the equilibrium spacing, $\sigma_z = 0$ for systems with $\sqrt{N} = 65$. The quadratic fit illustrates the phenomenological behaviour motivated by Eq. 1. $\chi N = 23$ and $f = \frac{1}{4}$. Adapted from Ref. 10.

grain boundaries, we consider systems with periodic boundary conditions in the following, thereby avoiding additional effects due to the narrow film surface.

Phenomenologically, one can motivate a kinetic equation for the directed grain-boundary motion by a Rayleighian that is the sum of the rate of free-energy change, $\dot{F}$, and the dissipation, $T\Sigma$.

$$\mathcal{R} = \dot{F} + \frac{1}{2}T\Sigma = -\Delta f A v + \frac{1}{2}A\zeta v^2 \tag{1}$$

where v denotes the velocity of the grain boundary, and ζ is a phenomenological friction constant of the grain-boundary motion. Minimisation of the Rayleighian with respect to the velocity, v, yields $v = \Delta f/\zeta$, where the free-energy difference is given by $\Delta f \sim \sigma_z^2$. This overdamped motion of the grain boundary is indeed compatible with previous arguments[11] and the simulation data, presented in Fig. 3.

The microscopic processes involved in the motion of the grain boundary are phenomenologically described by the friction coefficient, ζ. The particle-based MC simulations provide direct insights into the sequence of changes of the domain topology inside the grain boundary. For small σ_z, two classes of microscopic processes can be identified: (i) the fusion of two neighbouring heads of parallel cylinders to a perpendicular cylinder portion, and (ii) the rupture of connections between perpendicular and parallel cylinders. These two processes in an ideal, planar grain boundary are illustrated on the left-hand side of Fig. 4.

Whereas the topology changes only locally, the structure transformation is accompanied by a long-range distortion of the domain morphology. This effect is clearly visible in the density difference between the metastable states before and after the fusion of two cylinder heads (cf. Fig. 4 right) and stems from the long-range strain fields that surround a

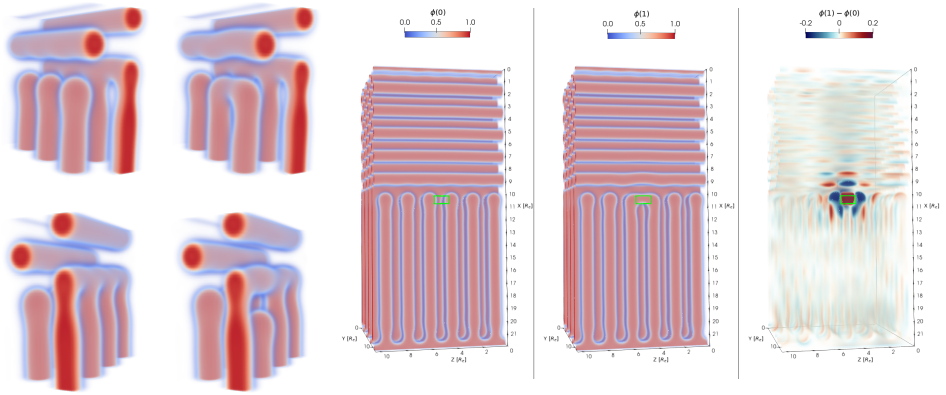


Figure 4. Left: Close-up of the fusion of two neighbouring heads of parallel cylinders to a perpendicular cylinder portion (top) and the rupture of connections between perpendicular and parallel cylinders bottom. The image indicates the normalised density of the minority component, A. Right: Illustration of the (small) system size used for calculating the MFEP of the fusion process. Both, the reference density of the metastable starting state (two unconnected neighbouring cylinder heads) and that of the final state (fused cylinder heads) are depicted, as well as density change of this transformation. The green box indicates the reduced region, employed to compute the reaction coordinate, *i. e.* arc length of the MFEP. (Note that the perpendicular z-direction is horizontal.) Adapted from Ref. 10.

defect in a spatially modulated phase. Due to these long-range strain fields, defects interact with each other, and the free-energy difference between the two metastable states, as well as the free-energy barrier that has to be overcome, depend on the environment of the two cylinder heads. Computationally, the dependence of the free-energy profile of the elementary microscopic processes on the environment necessitates the use of large systems.

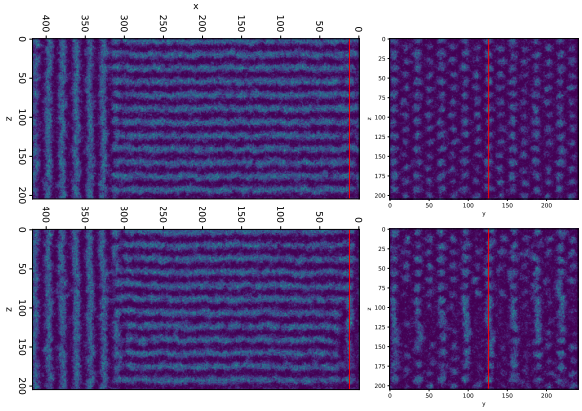


Figure 5. Cross-sectional snapshots of the minority-block density at different times, $t = 10^5$ MCS (initial grain boundary, top) and $t = 5 \cdot 10^6$ MCS (bottom). The left panel depicts a xz-cross-section perpendicular to the grain boundary at $y = 126R_\mathrm{e}$, indicated by the red line in right panel. The right panel presents the $yz-$plane parallel to the grain boundary at position, $x = 30R_\mathrm{e}$. $\sigma_z = 0.02$ and $\sqrt{\mathcal{N}} = 65$. Adapted from Ref. 10.

By virtue of the long-range interaction between defects, the grain-boundary motion involves a sequence of topological changes, correlated in space and time. Such a correlation is illustrate in Fig. 5 that presents simulation snapshot after a small deformation, $\sigma_z = 0.02$. The system size is $41.8 \times 24.1 \times 20.9 R_e^3$ and the low value of the invariant degree of polymerisation, $\sqrt{\mathcal{N}} = 65$, reduces all free-energy barriers and shifts the time scale of thermally activated processes into the window accessible to direct, kinetic MC simulations. The left panel depicts an ideal planar grain boundary and the cross-section parallel to the grain boundary shows alternating lines of parallel cylinder heads and junctions between parallel and perpendicular cylinders. The right panel presents the same cross-section at a later time. One observes that some neighbouring cylinder heads have fused and some cylinder junctions have ruptured. One can appreciate that the fusions of cylinder heads are not isolated and uncorrelated but that there is a higher probability to fuse a cylinder head with a neighbouring perpendicular cylinder portion, resulting in the extension of a perpendicular cylinder (oriented along the z-direction). Likewise, one notes that ruptures of connections between parallel and perpendicular cylinders preferentially occur next to fused perpendicular cylinders.

To quantify this observation, we have calculated the Minimum Free-Energy Path (MFEP) of the fusion and rupture processes for selected environments. The calculations employ a smaller system, see Fig. 4 right, but an increased invariant degree of polymerisation, $\sqrt{\mathcal{N}} = 1300$. The latter increases the free-energy changes in the course of structure transformation. Examples of MFEPs are presented in Fig. 6, where we study ruptures of junctions between parallel and perpendicular cylinders. In the first case, denoted "r", we consider the isolated rupture of a junction in an otherwise ideal, planar grain boundary. The second case, "fr", refers to the rupture in the neighbourhood of a fused pair of parallel cylinders, and in the third case, "2fr", we study the rupture if the rupture site is sandwiched between two perpendicular cylinder portions. The left panels of Fig. 6 presents the density distributions of the metastable starting and ending states, as well as the concomitant density difference. The latter panels indicate that the neighbouring cylinder heads of perpendicular cylinder portions move inwards. This response of the neighbouring structure elements distinguishes the free-energy profiles of rupture processes in the three different environments.

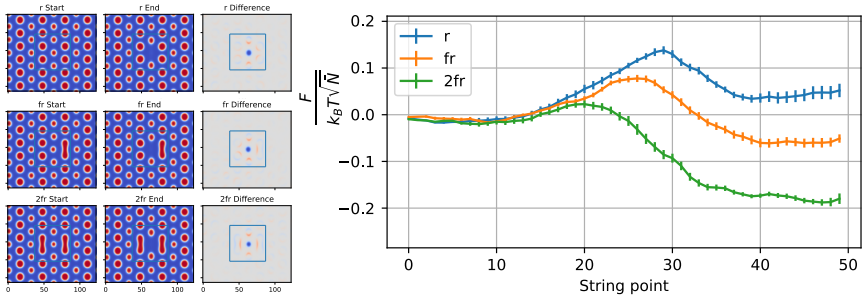


Figure 6. Left: Two-dimensional density plot in a plane parallel to the grain boundary for the three rupture processes considered. The metastable stating and ending states, as well as the density difference are shown. Right: MFEPs of the three rupture processes obtained by the improved string method[12] in conjunction with field-theoretic umbrella sampling[5]. Adapted from Ref. 10.

The free-energy profiles along the MFEPs that have been obtained by the improved string method[12] in conjunction with field-theoretic umbrella sampling[5] are presented in the right panel. Indeed we observe that the rupture of the junction between parallel and perpendicular cylinders becomes more favourable in the vicinity of neighbouring perpendicular cylinder portions. Whereas the isolated rupture in an ideal, planar grain boundary increases the free energy, the rupture in the vicinity of one perpendicular cylinder portion reduces the free energy, and the free-energy reduction increases further if the cylinder heads on both sides are fused into portions of perpendicular cylinders. Importantly, the free-energy barriers and therefore the inverse rate (waiting time), with which these transformations occur in kinetic simulations, decrease by the presence of neighbouring perpendicular cylinders. This analysis rationalises the shape of the cluster of fused cylinder heads and ruptured junctions in Fig. 5.

This simulation study also demonstrates that the motion of a grain boundary cannot be simply reduced to two elementary microscopic processes – fusion of neighbouring cylinder heads and rupture of junction between parallel and perpendicular cylinders. Instead, fusion and rupture are classes of processes that are comprised of different processes, whose free-energy differences and barriers subtly depend on the environment *via* the transformation of long-range strain fields associated with the local change of the domain topology.

Even with the resources of JUWELS Booster it is not feasible to calculate the free-energy profiles of all rupture and fusion processes for a rough (non-planar) grain boundary. Therefore, we assume that the environment will add a correction to the free-energy profile that is approximately linearly in the reaction coordinate, *i. e.* the arc length of the MFEP. For selected cases we have validated this assumption. The environment-dependent correction for other cases can be obtained by consistency requirements, asserting that (i) sequences of processes that connect the identical starting and ending states must yield the same free-energy difference and (ii) that the sum of free-energy differences of a sequence of processes that translates the grain boundary at $\sigma_z = 0$ by one lattice plane must add up to zero. This allows us to enlarge the list of free-energy changes and barriers of fusion and rupture processes and study the grain-boundary motion by kinetic MC simulation. Indeed, this approximate but computationally efficient scheme qualitatively reproduces the anisotropic shape of the cluster of rupture and fusion sites, revealed by the particle-based simulations presented in the right, bottom panel of Fig. 5 and allows us to estimate the phenomenological mobility, $1/\zeta$, of the grain boundary.

4 Summary

We have presented particle-based simulations that provide direct insights into the ordering kinetics of cylinder-forming diblock copolymers in response to the evaporation of a solvent from a thin polymer film. This process is the first step of the formation of isoporous membranes where the minority component, A, forms cylindrical channels of specific and narrowly distributed diameters that are oriented perpendicular to the film surface. The simulations provided a simple mechanistic rational for the early stages of structure formation and identified two characteristics – surface selectivity and evaporation rate – that control the orientation of the incipient mesophase. These two thermodynamic and kinetic characteristics can be tailored in experiments in multiple ways, *e. g. via* the preference of the gas phase, solvent selectivity, molecular architecture, incompatibility between gas and solvent

or plasticising effect of the solvent.

The later stage of ordering is dictated by the growth and shrinkage of grains, comprised of differently oriented cylinders. As the residual solvent evaporates after the microphase-separated structure has formed, the uniaxial compression along the normal of the film surface distorts the hexagonal structure of grains with parallel orientation of cylinders to the film surface but leaves the spatial arrangement of cylinders in perpendicular grains unaltered. Therefore, we expect that perpendicular grains grow at the expense of parallel grains *via* the motion of grain boundaries. Both, simulations and phenomenological considerations predict that the grain-boundary velocity is linearly proportional to the free-energy-density difference between the grains; the constant of proportionality depends on the microscopic processes that change the domain topology in the grain boundary. Our simulations have identified two classes of processes – fusion of cylinder heads and rupture of connections between parallel and perpendicular cylinders. The elementary microscopic processes, however, sensitively depend on the environment because of the change of the long-range strain fields associated with defects in these spatially modulated phases (qualitatively similar to what is observed in liquid crystalline phases of small, anisotropic molecules). This results in a sequence of elementary processes in the grain boundary, correlated in space and time. We also note that the free-energy profiles depend on molecular architecture and solvent content[13], indicating additional experimental control strategies.

We hope that the direct microscopic insights into the complex ordering kinetics and the simple, phenomenological concepts, employed to rationalise the simulations, will provide some useful guidance to experimentalists in judiciously choosing process parameters in the high-dimensional space of thermodynamic and kinetic characteristics that can be controlled in experiment[9].

Acknowledgements

It is great pleasure to thank Volker Abetz, Oliver Dreyer, Gregor Ibbeken, Louis Pigard, Maryam Radjabian, Ludwig Schneider, and Qiyun Tang for fruitful and enjoyable collaboration. Financial support has been provided by the DFG under grants Mu 1674/15-2 and 1674/16-2. 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 Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

References

1. L. Leibler, *Theory of Microphase Separation in Block Co-Polymers*, Macromolecules, **13**, 1602–1617, 1980.
2. W. H. Li and M. Müller, *Directed self-assembly of block copolymers by chemical or topographical guiding patterns: optimizing molecular architecture, thin-film properties, and kinetics*, Prog. Polym. Sci., **54-55**, 47–75, 2016.
3. Y. Ren and M. Müller, *Kinetics of pattern formation in symmetric diblock copolymer melts*, J. Chem. Phys., **148**, 204908, 2018.
4. M. Müller, *Process-directed self-assembly of copolymers: Results of and challenges for simulation studies*, Prog. Polym. Sci., **101**, 101198, 2020.

5. M. Müller, *Studying amphiphilic self-assembly with soft coarse-grained models*, J. Stat. Phys., **145**, 967, 2011.
6. K. C. Daoulas and M. Müller, *Single Chain in Mean Field simulations: Quasi-instantaneous field approximation and quantitative comparison with Monte Carlo simulations*, J. Chem. Phys., **125**, 184904, 2006.
7. L. Schneider and M. Müller, *Multi-architecture Monte-Carlo (MC) simulation of soft coarse-grained polymeric materials: SOft coarse grained Monte-Carlo Acceleration (SOMA)*, Comput. Phys. Commun., **235**, 463–476, 2019.
8. O. Dreyer, G. Ibbeken, L. Schneider, N. Blagojevic, M. Radjabian, V. Abetz, and M. Müller, *Simulation of Solvent Evaporation from a Diblock Copolymer Film: Orientation of the Cylindrical Mesophase*, Macromolecules, submitted, 2022.
9. M. Müller and V. Abetz, *Nonequilibrium Processes in Polymer Membrane Formation: Theory and Experiment*, Chem. Rev., **121**, 14189–14231, 2021.
10. N. Blagojevic, Ph.D. thesis Göttingen, in preparation.
11. J. L. Goveas and S. T. Milner, *Dynamics of the Lamellar-Cylindrical Transition in Weakly Segregated Diblock Copolymer Melts*, Macromolecules, **30**, 2605–2612, 1997.
12. W. E, W. Ren, and E. Vanden-Eijnden, *Simplified and improved string method for computing the minimum energy paths in barrier-crossing events*, J. Chem. Phys., **126**, 164103, 2007.
13. D. Duque, K. Katsov, and M. Schick, *Theory of T Junctions and Symmetric Tilt Grain Boundaries in Pure and Mixed Polymer Systems*, J. Chem. Phys., **117**, 10315–10320, 2002.

Transport in Nanowires and Colloidal Systems and Crystal Growth

Konstantin Beck, Bastian Gast, Kay Hofmann, Jacob Holder, Marc Isele, Anton Lüders, Fabian Müller, Markus Ring, Ralf Schmid, Philipp Stengele, Ellen Zander, and Peter Nielaba

Physics Department, University of Konstanz, 78457 Konstanz, Germany
E-mail: peter.nielaba@uni-konstanz.de

Transport in nanowires and colloidal systems as well as crystal growth are interesting research areas, and by the support of NIC we have explored many interesting details by computer simulations. These simulations range from Brownian Dynamics (BD) over Molecular Dynamics (MD), to Monte Carlo (MC) simulations and rare event techniques, and to electronic and phonon transport studies in metallic wires. Here we report on recent results.

1 Introduction

In the last years, our research on transport and structure formation in nanowires and colloidal systems as well as on crystal growth has been supported by NIC in two projects (CHMZ07 and TNCSCG) and has been done in the framework of the recently ended SFB 1214 with two projects (A4 and B4), the former SFBs 513, 767, TR6, and two DFG-projects on layer/lane formation in colloidal mixtures and in nanowire electromigration. Several important results have been achieved for structural and transport properties of nanowire systems[1–12], for structural and dynamical properties of colloidal systems[14–26], and for nucleation and crystal growth[27]. Here we report on recent results obtained in the projects CHMZ07 and TNCSCG. Details will be outlined in Sec. 2.

2 Scientific Work Accomplished and Results Obtained

2.1 Nanowires

In the last years, we have investigated the structural and transport properties of nanowires by a combination of Molecular Dynamics-, Non-Equilibrium Greens Function-, Tight-Binding-, and Density Functional- methods, using the Landauer-Büttiker formalism[1–12]. Electronic conductance histograms have been computed for wires with Au particles[1], with Ag, Pt, Ni particles[2], and in Fe, Co, Ni systems[3]. The thermopower of Ag, Au, and Pt nanocontacts has been explored[4] as well as the features of a current-driven single-atom memory[5]. At room temperature, shot noise variation has been studied in Au atomic break junctions[6] as well as the quantum thermopower of metallic contacts[7]. The orbital origin of the electrical conduction in ferromagnetic atomic-size contacts has been investigated[8], and the quantised thermal transport in single-atom junctions[9]. The thermal conductance has been analysed further[10], as well as the electronic and phonic transport[11] and electromigration phenomena[12].

A) Stability, Electromigration, and Switching in Atomic Sized Contacts

The potential to construct electronic devices with functional building blocks of atomic size is a major driving force of nanotechnology. In Ref. 12 we investigated the stability of gold, copper, aluminium and lead, see Fig. 1, in direct comparison with experimental results. We studied as well the contact evolution under electromigration. Extending our previous work on the effect of current-induced forces on the onset of switching in metallic contacts[12], we also approached the whole switching process. For this we used the unstable modes found to connect locally stable atomic configurations (states of the switch)[13] and investigated the process in between with reaction path analysis. For this we repeatedly looked for new stable configurations after moving the contacts along the runaway modes. This resulted in a sampling of the phase space of stable contact configurations together with conductance information in these states as well as a runaway voltage and direction of switching.

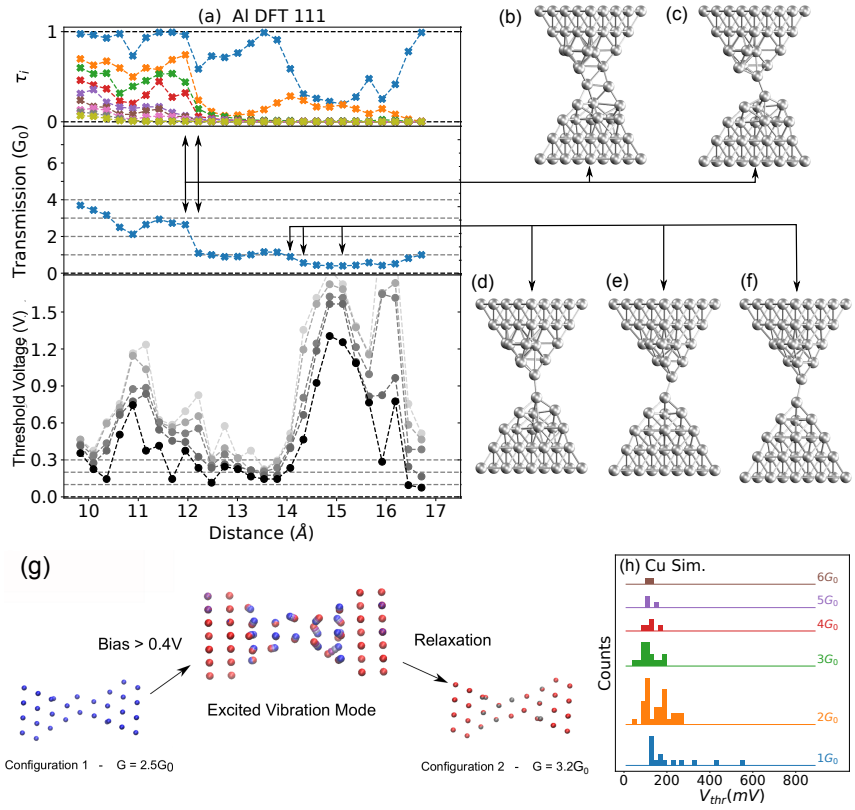


Figure 1. Simulation results for the investigated systems. Shown in (a-f) are the transport and stability threshold voltage results for an aluminium push contact. Subfigure (g) shows a simulated electromigration of a lead contact exhibiting a switch in conductance. The last figure (h) shows statistics for the threshold voltages of simulated copper contacts sorted into bins of different conductances, which has been compared with experimental data[12].

B) Phononic Contribution to the Thermal Conductance in Nanostructures

Understanding and efficiently controlling heat transport through nanoscale devices is a major driving force of nanotechnology. Besides via electrons (being the major force in terms of heat transport in metallic junctions), heat can also be transported via lattice vibrations, i.e. the phonons, and (to some extent) radiative processes, i.e. photons. We therefore extended our methods to include the phononic contribution to the thermal conductance in atomic metallic junctions.

With a non-equilibrium molecular dynamics (NEMD) approach we were able to calculate the phononic thermal conductance of atomic-sized metallic junctions under stretching alongside the electronic contribution computed via a tight-binding formalism to probe the Wiedemann-Franz law on the atomistic scale with a high statistical significance[11].

We investigated the role of inelastic effects in the process of phononic thermal transport. We therefore calculated the thermal conductance in the phase-coherent regime using the Landauer-Büttiker formalism[33]. Since both the NEMD and the phase-coherent approach are using a classical Molecular Dynamics simulation we can directly compare those two results. We did this comparison in a systematic study of atomistic gold chains varying their lengths. In Fig. 2 the phononic thermal conductance as a function of temperature is depicted for exemplarily chosen junction lengths. Also we show the absolute difference between the coherent and incoherent calculation methods as a function of the chain length in the bottom panel of Fig. 2.

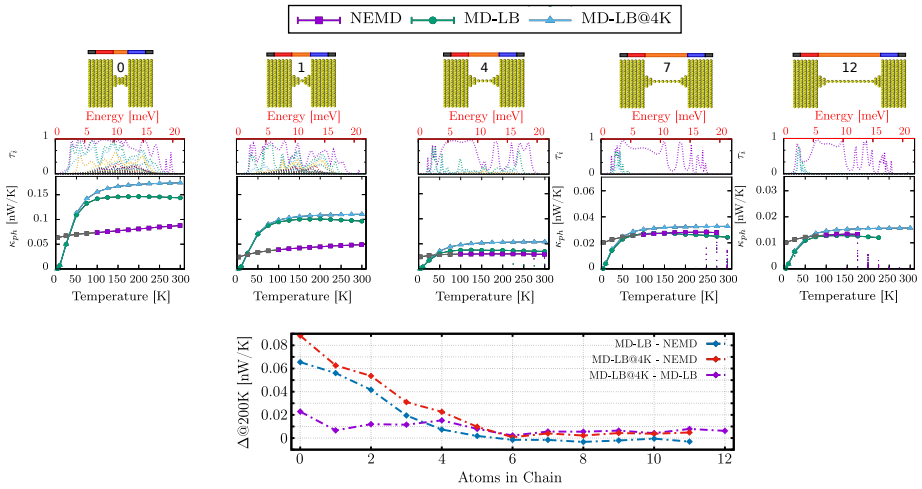


Figure 2. Phononic thermal conductance κ_{ph} as a function of temperature. Compared are the results of the NEMD (purple squares), our novel coherent method MD-LB (green circles) and a low-temperature approximation MD-LB@4K (blue triangles). The results are presented for 5 different junction lengths exemplarily. In the bottom panel the absolute difference between the three different transport descriptions are shown as a function of junction length at 200 K[33].

Our results could show that incoherent effects like phonon-phonon scattering even appear at the atomistic scale of metallic junctions[33]. Only for longer chains the transport can be seen as fully phase coherent.

2.2 Colloidal Systems

In the last years, we have investigated the structural and dynamical properties of colloidal systems by a combination of Monte Carlo-, Brownian Dynamics-, and Molecular Dynamics- simulations[14–26]. We studied the layering reduction in colloidal systems through microchannels in two dimensional[14,18] and three dimensional systems[19], phase transitions in two dimensional colloids in external periodic fields[15–17], a non-monotonic crossover from single-file to regular diffusion in micro-channels[20], and transport and diffusion properties of colloids in confinement and external fields[21–23], and with counterflow[24]. Recently, we investigated the single-file diffusion in a multilayered channel[25] and the dynamic ordering of driven spherocylinders in a nonequilibrium suspension of small colloidal spheres[26].

Dispersions of colloidal particles are the foundation of a significant percentage of materials and phenomena we witness constantly in our everyday life. Naturally, colloids are actively researched in fields such as medicine, biology, technology and physics. From the theoretical point of view, conventional Brownian Dynamics simulations are a simple yet good approach to analyse simple colloid-based systems by approximating the corresponding Langevin-equations numerically. In our work, we employ this algorithm to analyse the ordering phenomena and the formation of superstructures of colloidal model systems in confined geometries.

In our many-body studies, the interaction of the particles are realised via simple pair potentials like the Yukawa potential and the Weeks-Chandler-Andersen (WCA) potential[45]. To model the confinement in the particular systems, we use two distinct approaches: firstly, a hard-core repulsion given by the numerical treatment of the Smoluchowski equation for a single sphere in front of a wall[46] or secondly, a soft-core repulsion based on the WCA potential.

Originally, we focused our studies on spherical particles[14–25], but in recent years we broadened our research to non-spherical particles like rod- and dumbbell-shaped particles[26,28,32,49]. The following paragraphs will give a short overview of our current projects in the research area of colloids.

A) Particles in a Linear Channel

Binary mixtures of spherical colloids with two different particle sizes were simulated in a linear channel on a slope[38]. The gravitational force leads to a difference in velocity of the particle types which in turn can result in the formation of lanes. With a larger tilting angle, the forces on the colloids increase, leading to a higher lane formation parameter Φ. When plotting the Φ-value over the driving force (Fig. 3(a)) , a value of 0.5 corresponds to an extrapolated critical force F_{crit} which can describe the lane formation tendency of a whole system. In Fig. 3(b) we can see that for low particle densities, F_{crit} is stable over the length of the periodic channel. For larger densities however, a clear increase can be seen the longer the channel gets. (Additionally, for very long channels and high densities we can see a jamming phase forming at lower forces.)

In Fig. 4(a) we can see that the channel width has almost no influence on the lane formation properties of the system. Only in the very slim channels we can identify an influence which we assume to be occurring due of wall effects. In Fig. 4(a) we also analysed the effect of the diameter ratio a of the two particle types. For a larger a the difference in force gets larger as well, meaning lane formation occurs at lower F_{small}, the force on

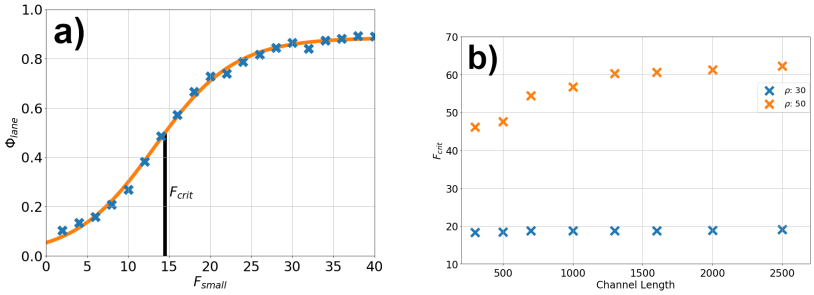


Figure 3. a) Order parameter Φ_{lane} over driving force F of the small colloids. The blue crosses represent simulations at the given driving force and the orange line is a sigmoidal fit over the data points. b) Critical forces F_{crit} over the channel length for different densities.

the smaller particles. If we analyse this more closely by looking at the difference in forces acting on the different particles, see Fig. 4(b), we can clearly see that there is barely any influence of the diameter ratio on lane formation tendencies. There is one exception, however, being the smallest particle size difference with a diameter ratio of $a = 1.4$. Here, a clear gap appears between those points and those of the larger particles.

In other studies, we investigated lane and band formation of oppositely driven colloidal particles in two-dimensional ring geometries[39].

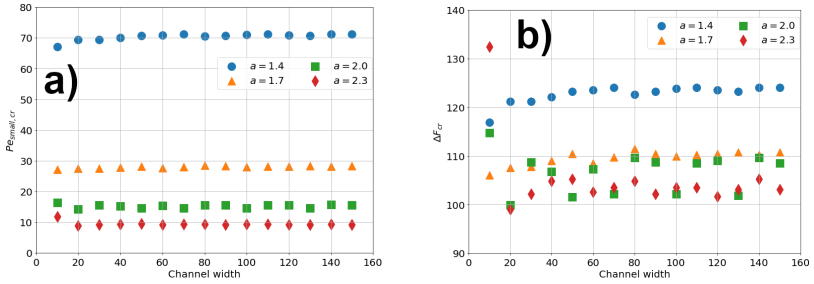


Figure 4. a) Critical force of the small particles $F_{small,crit}$ over channel width for different diameter ratios a. b) The same data but not the critical forces are plotted but rather the difference in force between the different colloid types ΔF_{crit}.

B) Diffusion of Anisotropic Colloids

Using bead-shell models and the HYDRO++ program[37] which is based on the methods described in Ref. 38, it is possible to calculate the microscopic single-particle diffusion coefficients of complex anisotropic particles[32]. Fig. 5 (left) depicts an example of a bead-shell model for symmetric dumbbells used in Ref. 48.

In the calculations based on the bead-shell models, the surface of the colloids is represented by small spherical friction elements, which are called beads. These elements interact with the surrounding liquid via Stokes friction and with each other via hydrodynamic interactions. The hydrodynamic interactions are approximated up to the third order

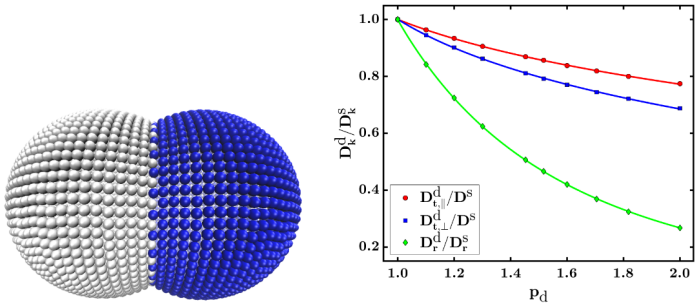


Figure 5. Left: Bead-shell model of a symmetric dumbbell. The different colours are used to highlight the transition between the two segments of the dumbbell[48]. Right: Resulting diffusion coefficients of symmetric dumbbells. The points are the data calculated with HYDRO++, the lines are the corresponding fits[48]. For further details see Ref. 32.

and no-slip boundary conditions are utilised. With this, we calculate the diffusion coefficients for different numbers and sizes of beads. Afterwards, we are able to extrapolate to the limit of infinite beads with zero diameter using the values based on the different shells. The calculations are repeated for different size parameters (for example the length of rod-like particles) so that we are able to calculate empirical formulas for the diffusion coefficients.

The points in Fig. 5 (right) depict the results of the calculations based on the bead-shell models and the HYDRO++ program for symmetrical dumbbells with different aspect ratios p_d. The coefficients $D^{\mathrm{d}}_{\mathrm{t},\parallel}$ and $D^{\mathrm{d}}_{\mathrm{t},\perp}$ model the diffusion parallel and perpendicular to the orientation axis of the particle, and the constant $D^{\mathrm{d}}_{\mathrm{r}}$ is the rotational diffusion coefficient. The quantities with the index "s" are their corresponding counterparts for spheres. Using a suitable fit, the calculated diffusion coefficients can be used to determine empirical formulas. The result of the fitting is depicted by the lines in Fig. 5 (right). A detailed discussion of the resulting empiric formulas for dumbbells is given in Ref. 32.

Using the determined microscopic diffusion coefficients, it is possible to perform Brownian dynamics simulations of multiple particles. In these simulations, it is possible to determine dynamic properties like self-diffusion or long-time diffusion coefficients shaped by particle interactions. For example, we recently studied these properties in monolayers of rod-sphere mixtures in Ref. 32.

C) Structure Formation of Dumbbell-Shaped Colloids in Monolayers

In cooperation with the Wittemann group[43] at the University of Konstanz, we study the structure formation and the order of dumbbell-shaped colloids in monolayers generated under geometrical confinement[28]. In the experiments performed by the Wittemann group, repulsive dumbbells with a width of approximately $170\,\mathrm{nm}$ and an aspect ratio of 1.55 are compacted using the Langmuir-Blodgett method[40,47]. Using this technique, monolayers with a high area fraction at the air/water interface are obtained. The structure and order in these layers of colloids can be compared to two-dimensional Brownian dynamics simulations, afterwards.

In the corresponding simulations performed by us, the dumbbells are realised as rigid

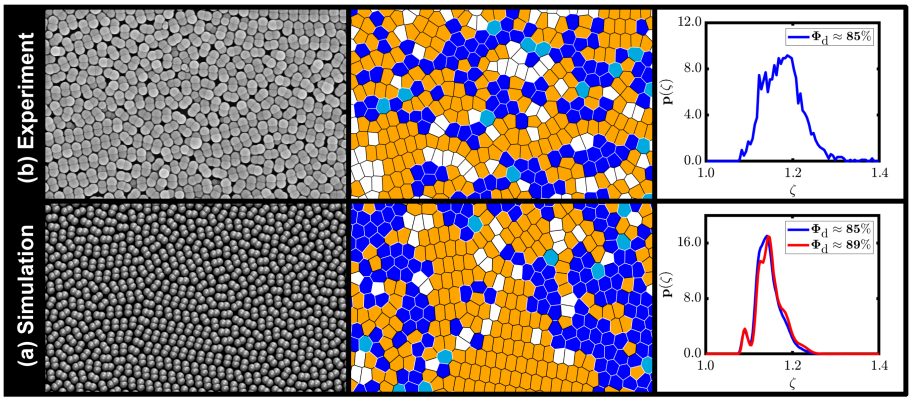


Figure 6. Comparison of experimental and numerical results regarding high density dumbbell monolayers. The upper image shows experimental data of the Research Group of Alexander Wittemann, University Konstanz are depicted. The lower image shows the results of our Brownian dynamics simulations. Corresponding Voronoi tessellations are depicted next to the particular images. The orange coloured Voronoi cells indicate especially ordered dumbbells. The curves on the right depict early results for the probability density of the Voronoi shape factor $p(\zeta)$ for a given final area fraction ϕ of the monolayers. The experiments and the simulations show similar qualitative progressions.

particles based on two overlapping spheres which can interact via a modified WCA potential[28]. To model the Langmuir-Blodgett setup, soft wall boundary conditions with a variable position are utilised. By reducing the distance between the walls, high density dumbbell monolayers are created which can then be compared to the experimental layers.

Using Voronoi tessellations and quantities like the Voronoi shape factor[42] (which quantifies the nonsphericity of the particular Voronoi cells), we analyse the superstructure of the monolayers under variation of the parameters of the Langmuir-Blodgett setup given by the target surface pressure and the barrier speed. In the same way, the final area fraction and the barrier speed is adjusted in the simulations. We find that higher target surface pressures and lower barrier speeds lead to more ordered monolayers. The experiments and the simulations are in good agreement, see Ref. 28 for a detailed discussion. An example of early results for the shape factor probability density $p(\zeta)$ are depicted in Fig. 6.

D) Colloidal Rods with a Visual Perception

Active matter systems like self-propelled particles[35] are an important and often discussed topic in modern science. Using self-propelled particles as a minimum model, it is possible to understand the fundamental dynamics of simple living organisms[34] like bacteria. To understand more complex systems of living organisms, further effects and influences have to be taken into account: In systems of living organisms, environmental stimuli like chemical or optic signals modify the formations of groups and particle clusters[41]. In our studies (see Ref. 49) we introduce a simple model system of two-dimensional colloidal spherocylinders which become self-propelled under visual stimuli triggered by their neighbouring particles. With this minimum model of a colloid system with additional cognitive ability, we strive to further the understanding of the formation and the dynamics of swarms or other collective behaviour.

The particles in our simulations interact via a Weeks-Chandler-Andersen potential. In our model system, every particle is linked to a predefined field of view. A specific particle moves according to active Brownian motion if there are enough other close particles inside the corresponding field of view but diffuses normally otherwise. Fig. 7 (left side) shows a schematic depiction of a particle with its cone of sight. In our most recent studies, we use field of views with an infinite cone of sight radius $r_v^{\max} \to \infty$ and a visual perception function (based on Ref. 41) quantifying the visual stimulus.

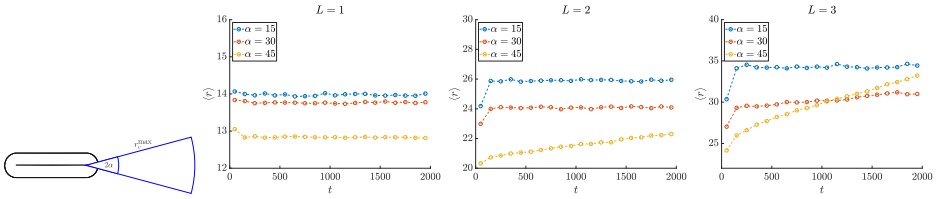


Figure 7. Left: Sketch of a spherocylinder with its corresponding sight cone. Right: Average distance $\langle r \rangle$ from the centre of mass for different field of view angles α and particle lengths L. Only for specific parameters, $\langle r(t) \rangle$ is constant and thus stable swarms are formed.

In our simulations, we place, for example, a small number of spherocylinders with an infinite cone of sight characterised by an opening angle α in the centre of a large simulation box. The length of this box is chosen large enough that the colloids are not able to reach the boundaries of the simulation box during the simulations. By calculating the average distance from the centre of mass $\langle r(t) \rangle$, we analyse the stability of formed swarms for different particle lengths L.

Fig. 7 (right side) depicts some resulting $\langle r(t) \rangle$ curves for different L and α. It can be seen that only for specific parameter combinations, the average distance to the centre of mass is constant and thus only for these parameters, stable swarms are formed. This indicates that the length of the particle is crucial for the maximum field of view angle α for which swarms are possible. A detailed discussion of this phenomena is given in Ref. 49.

2.3 Nucleation Phenomena and Crystal Growth

How do anisotropic particles grow? Where in the process of nucleation and growth does anisotropy appear? Which factors steer the formation of certain shapes and morphologies? In recent years exciting experimental studies have shed new light on the processes and structures that occur during the early stages of mineralisation, and the traditional view that had been dominated by classical nucleation theory and classical growth models was significantly expanded. New pathways and intermediates have been investigated[27,30], for example (liquid-like, amorphous) prenucleation clusters and metastable amorphous precursors that exist up to macroscopic length scales, requiring solid-solid phase transformations during mineralisation. In addition, interesting effects on the structure formation of nanocube mesocrystals have been investigated by MD, in particular the effect of corner rounding[31].

In Ref. 27 we employed the well-tempered parallel-bias metadynamics algorithm to study the stability of nanoparticles in a lattice gas for crystallisation from solution. The model allows us to give a description for the transition from amorphous to crystalline

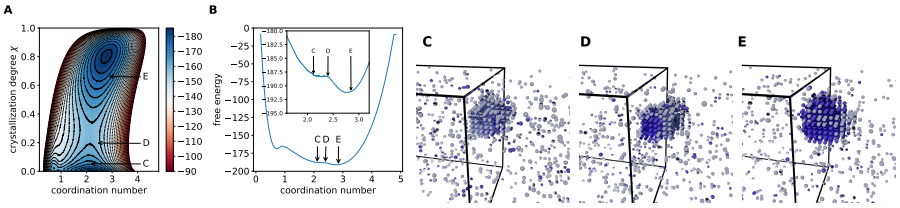


Figure 8. Free energy surface of two-step nucleation (A,B), with example configurations (C-E). The solute particles form an amorphous (grey) cluster, which crystallises (blue) from the inside. [From Ref. 30 under Creative Commons Attribution license]

nanoparticles by introducing parameters directly related to the surface tensions of the two phases and also the differences of the entropy per particle in each phase. By examining the parameter space, we find a critical cluster size of crystalline stability, whose temperature and size dependencies follow the Gibbs-Thomson equation. An additional melting point depression due to cluster surface fluctuations was observed, leading to a non-classical nucleation barrier of cluster growth. In Ref. 30 we analysed the impact of confinement on two-step nucleation. The depletion effect prevents further growth of the amorphous state. By crystallisation of the amorphous cluster additional binding energy get produced, which then allows further growth. This effect should be also applicable to slow transport. In Fig. 8 the free energy surface and example configurations are displayed.

3 Concluding Remarks

We have described a variety of computer simulation results for transport in nanowires and colloidal system as well on crystal growth. The computer time granted by NIC was essential to achieve these results.

Acknowledgements

We thank NIC for computer time granted in the projects CHMZ07 and TNCSCG as well as the DFG for support (SFB1214, NI 259/14-1, project 299102402). We further thank the groups of Profs. Wittemann, Coelfen, Scheer, Leiderer (all from U. Konstanz), Cuevas (U. Madrid), Pauly (U. Augsburg), and Erbe (HZDR Dresden) for the excellent cooperations.

References

1. M. Dreher, F. Pauly, J. Heurich, J. C. Cuevas, E. Scheer, P. Nielaba, Phys. Rev. B **72**, 075435, 2005.
2. F. Pauly, M. Dreher, J.K. Viljas, M. Häfner, J. C. Cuevas, P. Nielaba, Phys. Rev. B **74**, 235106, 2006.
3. M. Häfner, J. K. Viljas, D. Frustaglia, F. Pauly, M. Dreher, P. Nielaba, J. C. Cuevas, Phys. Rev. B **77**,104409, 2008.

4. F. Pauly, J. K. Viljas, M. Bürkle, M. Dreher, P. Nielaba, J. C. Cuevas, Phys. Rev. B **84**, 195420, 2011.
5. C. Schirm, M. Matt, F. Pauly, J. C. Cuevas, P. Nielaba, E. Scheer, Nature Nanotechnology **8**, 645, 2013.
6. R. Chen, M. Matt, F. Pauly, P. Nielaba, J. C. Cuevas, D. Natelson, J. Phys. - Cond. Mat. **26**, 474204, 2014.
7. C. Evangeli, M. Matt, L. Rincon-Garcia, F. Pauly, P. Nielaba, G. Rubio-Bollinger, J. C. Cuevas, N. Agrait, Nano Lett. **15**, 1006, 2015.
8. R. Vardimon, M. Matt, P. Nielaba, J. C. Cuevas, O. Tal, Phys. Rev. B **93**, 085439, 2016.
9. L. J. Cui, W. Jeong, S. Hur, M. Matt, J. C. Klöckner, F. Pauly, P. Nielaba, J. C. Cuevas, E. Meyhofer, P. Reddy, Science **355**, 1192, 2017.
10. J. C. Klöckner, M. Matt, P. Nielaba, F. Pauly, J. C. Cuevas, Phys. Rev. B **96**, 205405, 2017.
11. D. O. Möhrle, F. Müller, M. Matt, P. Nielaba, F. Pauly, Phys. Rev. B **100**, 125433, 2019.
12. M. Ring, D. Weber, P. Haiber, F. Pauly, P. Nielaba, E. Scheer, Nano Lett. **20**, 5773, 2020.
13. M. Ring, E. Scheer, F. Pauly, P. Nielaba, *Simulating switching through phonon-pumping*, in preparation.
14. M. Köppl, P. Henseler, A. Erbe, P. Nielaba, P. Leiderer, Phys. Rev. Lett. **97**, 208302, 2006.
15. F. Bürzle, P. Nielaba, Phys. Rev. E **76**, 051112, 2007.
16. K. Franzrahe, P. Nielaba, Phys. Rev. E **76**, 061503, 2007.
17. K. Franzrahe, P. Nielaba, Phys. Rev. E **79**, 051505, 2009.
18. P. Henseler, A. Erbe, M. Köppl, P. Leiderer, P. Nielaba, Phys. Rev. E **81**, 041402, 2010.
19. N. Schwierz, P. Nielaba, Phys. Rev. E **82**, 031401, 2010.
20. U. Siems, C. Kreuter, A. Erbe, N. Schwierz, S. Sengupta, P. Leiderer, P. Nielaba, Sci. Rep. **2**, 1015, 2012.
21. C. Kreuter, U. Siems, P. Henseler, P. Nielaba, P. Leiderer, A. Erbe, J. Phys. - Cond. Mat. **24**, 464120, 2012.
22. C. Kreuter, U. Siems, P. Nielaba, P. Leiderer, A. Erbe, Eur. Phys. J. **222**, 2923, 2013.
23. U. Siems, P. Nielaba, Phys. Rev. E **91**, 022313, 2015.
24. B. Heinze, U. Siems, P. Nielaba, Phys. Rev. E **92**, 012323, 2015.
25. U. Siems, P. Nielaba, Phys. Rev. E **98**, 032127, 2018.
26. A. Lüders, U. Siems, P. Nielaba, Phys. Rev. E **99**, 022601, 2019.
27. R. Schmid, P. Nielaba, J. Chem. Phys. **150**, 054504, 2019.
28. R. Stuckert, A. Lüders, A. Wittemann and P. Nielaba, Soft Matter **17**, 6519, 2021.
29. Z. Chen, R. Schmid, X. Wang, M. Fu, E. Scheer, M. Huang, P. Nielaba, H. Cölfen, *Universal wet-chemical green synthesis of two-dimensional nanomaterials*, preprint.
30. J. Holder, R. Schmid, P. Nielaba, J. Chem. Phys. **156**, 124504, 2022.
31. K. Beck, Master thesis, University of Konstanz, 2022.
32. A. Lüders, E. Zander, P. Nielaba, J. Chem. Phys. **155**, 104113, 2021.
33. F. Müller, J. C. Cuevas, F. Pauly and P. Nielaba, *Heat conductance in Au single-atom contacts: Coherent versus incoherent transport*, draft preprint.

34. L. Angelani, R. Di Leonardo, and G. Ruocco, Phys. Rev. Lett. **102**, 048104, 2009.
35. C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, Rev. Mod. Phys. **88**, 045006, 2016.
36. M. X. Fernandes and J. G. de la Torre, Biophysical Journal **83**, 3039, 2002.
37. J. Garcia de la Torre and S. Harding, Biophysics Reviews **5**, 195, 2013.
38. K. Hofmann, Bachelor thesis, University of Konstanz, 2020.
39. T. Vater, M. Isele, U. Siems, P. Nielaba, *Lane and band formation of oppositely driven colloidal particles in two-dimensional ring geometries*, preprint.
40. S. A. Hussain, S. Deb, and D. Bhattacharjee, J. Env. Sc. Res. **4**, 25, 2005.
41. F. A.Lavergne, H. Wendehenne, T. Bäuerle, C. Bechinger, Science **364**, 70, 2019.
42. F. Moucka and I. Nezbeda, Physical Review Letters **94**, 040601, 2005.
43. R. Stuckert, C. S. Plüisch, and A. Wittemann, Langmuir **34**, 13339, 2018.
44. B. ten Hagen, F. Kümmel, R. Wittkowski, D. Takagi, H. Löwen, and C. Bechinger, Nature Communications **5**, 4829, 2014.
45. J. D. Weeks, D. Chandler, and H. C. Andersen, J. Chem. Phys. **54**, 5237, 1971.
46. H. Behringer and R. Eichhorn, Phys. Rev. E **83**, 065701(R), 2011.
47. H. Wu, G. Niu, W. Ren, L. Jiang, O. Liang, J. Zhao, Y. Liu, and Y.-H. Xie, Langmuir **36**, 10061, 2020.
48. E. Zander, Master thesis, University of Konstanz, 2020.
49. P. Stengele, A. Lüders, P. Nielaba, *Group formation and collective motion of colloidal rods with an activity triggered by visual perception*, preprint.

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 increasingly recognised as a research area that tackles a number of grand challenges of high socio-economic relevance. Prominent examples include climate change and associated impacts such as sea level rise, biodiversity loss and geo-hazards such as volcanic eruptions, earthquakes and tsunamis. Numerical experimentation with models constitutes one of the main tools – augmenting theory, experiment and more recently data science – to make progress. In this context, the use of computers has a long tradition as expressed, for example, by the fact that weather forecasting has been considered by John von Neumann and his team as one of the first prime applications. Nowadays, modelling is routinely being used by scientists in the field for numerical experimentation that enhances system understanding (establishing causality), for predictions (statements about possible future trajectories of systems), and for data assimilation (advanced monitoring and 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 making progress in science – as illustrated by the two projects in the field of Earth and Environment, described in this issue.

Sophie Stolzenberger and colleagues use to Finite Element Sea Ice-Ocean Model (FESOM) – the first mature global sea ice-ocean model formulated on unstructured meshes – to study the impact of Greenland ice sheet melting on the North Atlantic Ocean, along with its sensitivity to horizontal resolution. In fact, the high-resolution setup, going down to a resolution of 6 km in the region of interest is rather competitive by international standards. The authors clearly demonstrate a resolution dependence of the response to a freshwater forcing of the ocean around Greenland. This finding supports the notion that the response of the climate system to anthropogenic warming sensitively depends on how realistic we simulate certain key processes. In this regard, the results by Stolzenberger *et al.* make the case for using a new class of km-scale climate models that will be run on the first (pre-)exascale HPC systems in the coming years.

In the second study, Sandro Jahn and colleagues from the University of Cologne explore hydrothermal fluids that are key agents in the Earth's crust and upper mantel. More specifically, they consider solute-solvent interactions in supercritical fluids as a function of density and the speciation of rare earth element cations in supercritical solutions, using molecular dynamics (MD) methods on JUWELS (typically 8 nodes and 384 cores). The authors conclude that more classical approaches are well-suited to the problem. However, they also point out certain limitations related, for example, to finite size and time effects. This leads Jahn *et al.* to call for new types of interaction models to be developed, of which

reactive force fields and neural networks are promising candidates, both of which require the availability of sufficient HPC resources.

Finally, I would like to remind the reader that emerging technologies related to upcoming exascale HPC systems provide both great opportunities and major challenges to modellers in Earth and Environment. The computational resources that will become available in the coming years, are bound to provide the opportunity for breakthroughs in our ability to simulate and predict the Earth system – for example by ramping up spatial resolution and thus make more use of the laws of physics; and the same time, making our models and associated workflows exascale-ready provides a major challenge, that will require disruptive changes to our software. Clearly, exciting times lie ahead of us.

The Impact of Greenland Ice Sheet Melting on the North Atlantic Simulated with the FESOM Sea Ice-Ocean Model

Sophie Stolzenberger[1], Roelof Rietbroek[2], Claudia Wekerle[3], Bernd Uebbing[1], and Jürgen Kusche[1]

[1] Institute of Geodesy and Geoinformation, University of Bonn,
Nussallee 17, 53115 Bonn, Germany
E-mail: {stolzenberger, uebbing, kusche}@geod.uni-bonn.de

[2] ITC-WRS, Faculty of Geo-Information Science and Earth Observation, University of Twente,
Hengelosestraat 99, 7514AE Enschede, Netherlands
E-mail: r.rietbroek@utwente.nl

[3] Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research,
Klußmannstr. 3d, 27570 Bremerhaven, Germany
E-mail: claudia.wekerle@awi.de

Increased melting of the Greenland ice sheet has led to increased Greenland freshwater discharge, particularly in recent years. This study examines the extent to which melting effects are visible in the North Atlantic. To do this, we perform ocean model simulations with the coupled sea ice-ocean model FESOM (Finite Element Sea Ice-Ocean Model). FESOM is a global ocean model, which is able to use multi-resolution meshes. We run the simulations at two different grid resolutions to investigate the sensitivity of ocean variables at a higher model resolution. The simulations are compared to a variety of different data sets including *in situ* data in the ocean (Argo floats), satellite measurements and ocean reanalysis. We find that the Greenland freshwater impact is visible up to a depth of 100 m for salinity and temperature. Baffin Bay (west of Greenland) is a sensitive region for Greenland freshwater discharge due to the local currents close to the Greenland coast. We find that simulated melting effects are often small compared to the effect of an increased model resolution.

1 Introduction and Motivation

In recent years, the Greenland freshwater discharge has increased primarily due to increased melting of Greenland's glaciers and ice sheet. This melting is a major contributor to global sea level rise over the past two and a half decades, causing a mean sea level rise of 10.8 ± 0.9 mm since the 1990s (The IMBIE Team, 2020[1]). Since the mid 1990s, accelerated melting rates could be observed, e.g. with the help of satellite measurements. Gravimetric satellite missions such as the Gravity Recovery Climate Experiment (GRACE, launched in 2002) and its successor GRACE-FO (launched in 2018) have become essential for measuring Earth gravity changes. With altimeter satellite missions, such as Topex, it has become possible to observe sea surface heights above the ellipsoid by measuring the distance between the known satellite position and the ocean. To separate mass-induced sea level changes from volume-induced changes, a combination of gravimetric and altimetric satellite observations can be performed. For example, the separation of the mass induced sea level in an inversion framework results in a simultaneous estimation of mass changes and steric changes, i.e. changes due to volume expansion caused by temperature and/or salinity variations (Rietbroek *et al.*, 2016[2]).

Oceanic (and atmospheric) processes influence surface melting and retreat of Greenland outlet glaciers and *vice versa*. Ocean models forced with atmospheric data can be used to perform simulations of changing ocean dynamics. The additional input of Greenland freshwater into the model can be done by so-called hosing experiments, where Greenland freshwater is artificially added to the ocean (e.g. Gerdes *et al.*, 2006[3]). For example, Brunnabend *et al.* (2012)[4] assume a constant Greenland mass loss rate of 200 Gt/year in their ocean model simulations and study the effects of regional sea level change and the circulation. After 48 years, the whole North Atlantic is affected by sea level rise (except for the subtropical gyre) with highest amplitudes in Baffin Bay (west of Greenland).

In this study, we show the effects of Greenland freshwater on the surroundings of Greenland and the North Atlantic. For this, we perform ocean model simulations using the Finite Element Sea Ice-Ocean Model (FESOM) and compare them to a variety of different observational types, such as Argo floats, satellite measurements and ocean reanalysis. Variables of interest are salinity, temperature and steric height changes. In addition, we are interested in the effects of an increased horizontal model resolution. This study and the following sections are essentially based on Stolzenberger et al. (under review)[9]. It was conducted as part of the research programme *Greenland Ice Sheet Ocean Interaction* (GROCE) funded by the German Federal Ministry of Education and Research, where we aim to improve our understanding of the changing dynamics of the Greenland Ice Sheet by using satellite measurements and other observations as well as model simulations in atmosphere, cryosphere and ocean (e.g. Turton *et al.*, 2019[5]; Schaffer *et al.*, 2020[6]; Huhn *et al.*, 2021[7]; Kappelsberger *et al.*, 2021[8]).

2 FESOM Model and Data

2.1 Model Description and Experiment Setup

We use the global multi-resolution sea ice-ocean model FESOM Version 1.4 (FESOM1.4, see Wang *et al.*, 2014[10]). FESOM1.4 consists of a finite element ocean model (e.g. Wang *et al.* 2008[11]) and is coupled to a dynamic-thermodynamic finite element sea ice model (Danilov *et al.*, 2015[12]). In 2D, FESOM uses an unstructured triangular surface mesh. In 3D, 2D elements are used to form prisms, which are cut into tetrahedras. The model configuration provides z-levels with 47 vertical layers with a 10 m resolution in the upper 100 m.

Temperature and salinity climatologies provided by the Polar Science Center[a] are used to initialise FESOM. Atmospheric variables, such as zonal and meridional wind (10 m height), air temperature (10 m height), specific humidity (10 m height), precipitation and downward radiation (shortwave and longwave) provided by the JRA-55 reanalysis (Kobayashi *et al.*, 2015[13]) are used as forcing. We set the spin-up phase of the model to 13 years, and the analysing time period is from 1993 to 2016.

One benefit of FESOM is that it can deal with multi-resolution meshes. This means that a global model domain is used, with local mesh refinement in the area of interest. This has the advantage that a high resolution regional analysis can be performed in the global model, without the need for additional boundary conditions for the subdomain. In our

[a] http://psc.apl.washington.edu/nonwp_projects/PHC/Climatology.html

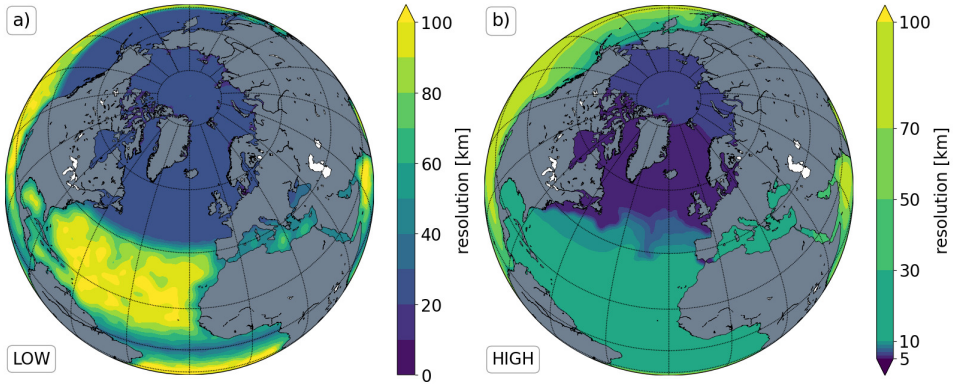


Figure 1. Global mesh resolution (in km) for LOW (a, number of 2D-vertices is 137,186) and HIGH (b, number of 2D-vertices is 919,338); Note the different colour scales. Source: modified from Stolzenberger *et al.* (under review)[9].

study, we perform simulations on two different meshes: one with a low resolution (we call it LOW in the following), i.e. 24 km in the North Atlantic and Arctic region, and one with a high resolution (we call it HIGH in the following), i.e. 6 km in the mentioned regions. The remaining ocean areas have a resolution of 1 degree (see Fig. 1). The HIGH mesh configuration is eddy-permitting in some regions of the northern North Atlantic, e.g. Baffin Bay, but not in the Greenland shelf areas while the LOW mesh configuration cannot resolve or permit eddies.

To investigate the impact of Greenland freshwater on the North Atlantic, we perform two different model experiments: in the "realistic" run we include Greenland freshwater and in the reference run we exclude Greenland freshwater. We use the Greenland freshwater data set by Bamber *et al.* (2018)[15], which is derived from satellite observations and a regional climate model. We distribute this data set horizontally with a radius of 60 km from the Greenland coast. It is included in the top model layer (ocean surface) in FESOM.

Unstructured mesh models have the advantage of being more flexible than traditional structured models (Wang *et al.*, 2014[10]). In general, the model efficiency of typical structured meshes is higher than that of unstructured meshes (time improvement by a factor of 10). However, FESOM is well suited for the Arctic region because of its particular geometry with a lot of land and shelf areas, in contrast to ocean models with structured grids where many grid points cover land. Experiences with FESOM show that it is computationally advantageous to have most of the grid nodes within the refinement region (Wang *et al.*, 2014[10]). For mesh HIGH, approximately 80 % of all surface nodes are located in the refinement area (see Fig. 1).

FESOM1.4 has been tested on different high performance computing systems in the past and it scales well. Parallelisation is based on MPI-only and is done by splitting the grid into small areas. The newer FESOM version (FESOM2) also shows promising results in terms of throughput, is continuously optimised and can keep up with known structured ocean models (Koldunov et al., 2019[14]).

2.2 Observations and Other Data Sets

For salinity and temperature, we compare our modelling results to Argo float measurements. Argo floats are heterogeneously distributed, autonomously powered floats in the ocean that provide vertical profiles of salinity, temperature, and pressure from sea surface to 2000 m depth. We use the profiles from 2005 until 2016 (Argo, 2020[16]). The FESOM simulations are interpolated to each Argo float position and are aggregated to given depth levels.

For the comparison with observed steric height changes, we use a joint combination of gravimetric and altimetric satellite data in an inversion framework (see Rietbroek *et al.*, 2016[2] and Uebbing *et al.*, 2019[17] for method details). Gravimetric data resulting from changes in ocean mass are obtained from the GRACE mission. GRACE data are used from 2002 to 2016. In this setup, altimetric data include the missions of Envisat (2002-2012),

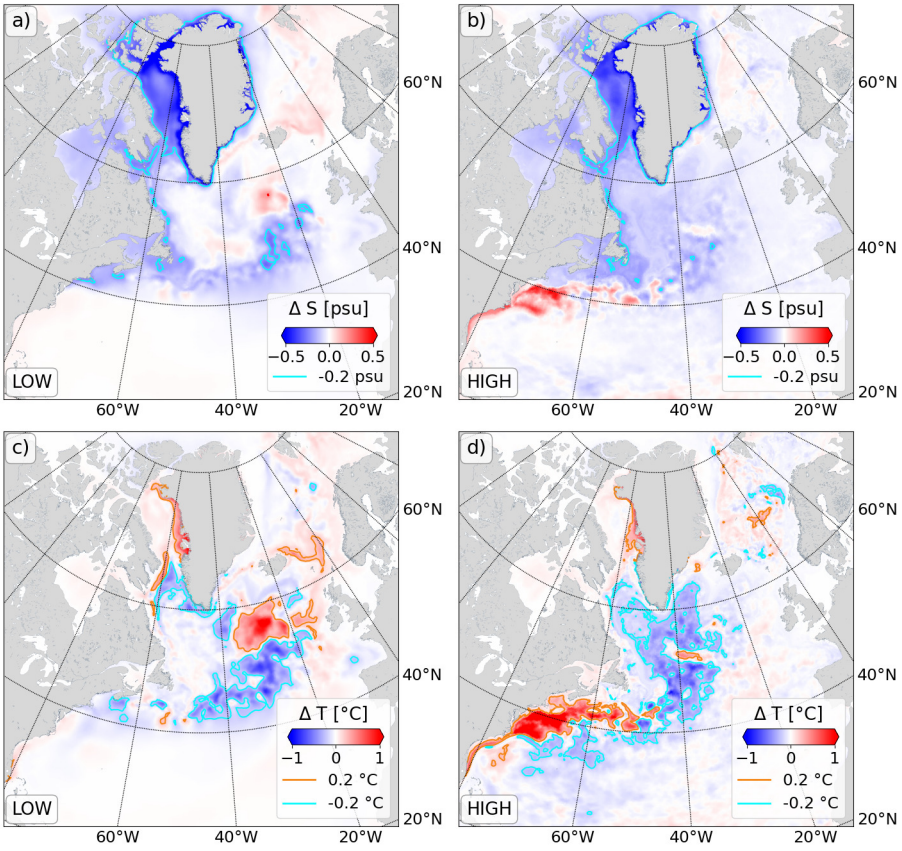


Figure 2. Differences between the "realistic" Greenland freshwater simulations and reference simulations for LOW (a, c) and HIGH (b, d) averaged over the upper 100 m for the time period 1993-2016. The differences are shown for salinity (top) and temperature (bottom). Blue shadings represent cooling and freshening effects due to Greenland freshwater. Source: Stolzenberger *et al.* (under review)[9].

SARAL/AltiKa (2013-2016), Jason1 (2001-2012), and Jason2 (2008-2016) and measure the total contribution to sea level change. With the fingerprint inversion method we can resolve for steric and mass-driven sea level changes at the same time. Furthermore, we derive steric heights from temperature and salinity fields of the Ocean Reanalysis System 5 (ORAS5, Zuo *et al.*, 2019[18]) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). This reanalysis provides an ensemble of five members.

3 Results and Discussion

The effect of Greenland freshwater is mainly visible as large scale salinity and temperature anomalies in Baffin Bay (west of Greenland) and south of Greenland (Fig. 2). We find

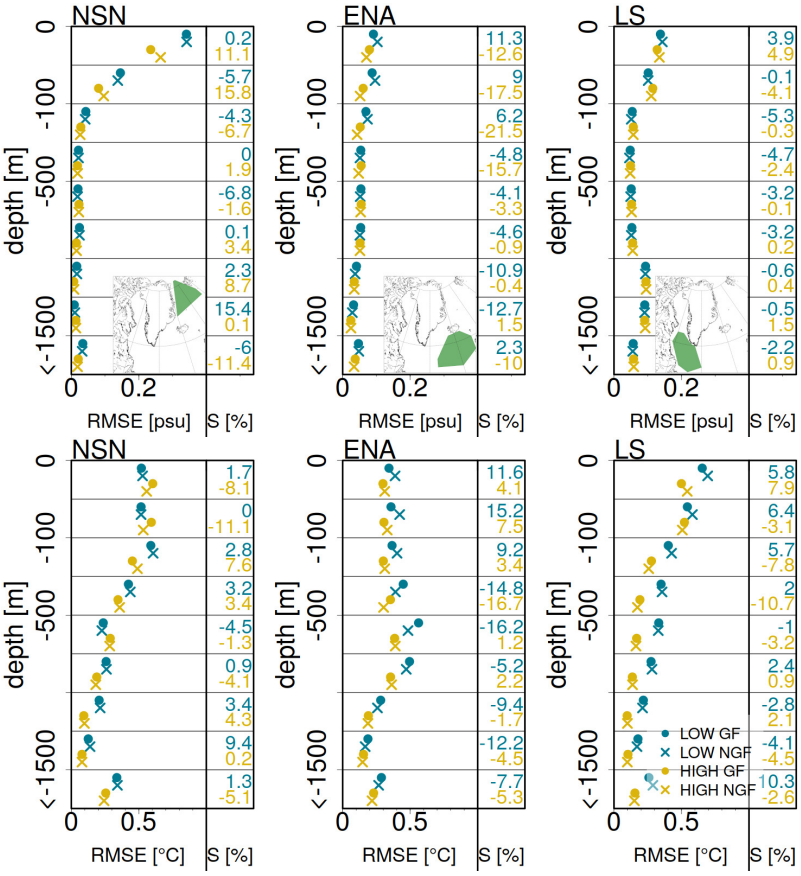


Figure 3. Root mean square error (RMSE) between each experiment (GF is linked to the Greenland freshwater experiment, NGF to the reference run) and Argo float measurements for three areas (green areas in the subplots): Nordic Seas North (NSN), Eastern North Atlantic (ENA) and Labrador Sea (LS). In addition, the relative improvement (regarding the RMSE) of the Greenland freshwater experiment compared to the reference run is shown (S) for salinity (top) and temperature (bottom). Source: modified from Stolzenberger *et al.* (under review)[9].

freshening effects around the Greenland coast and in Baffin Bay of up to 1 psu compared to the reference run where no Greenland freshwater is included. The pattern are similar for LOW and HIGH in the surroundings of Greenland. In terms of temperature, there is cooling south of Greenland and warming west of Greenland due to the freshwater input. The latter is caused by rising sea surface heights leading to changes in the local current system and transport of warmer water northwards along the coast (Castro de la Guardia *et al.*, 2015[19]). For HIGH there are even warming effects at the Canadian / North-American coast, which are not visible for LOW. Although we force the ocean with Greenland water at the ocean surface, the effects are also visible at deeper levels up to 100 m depth.

For comparing the model output to the heterogeneously distributed Argo float observations, we interpolate the simulations to the corresponding positions and average them in vertical layers. We define three regions around Greenland: Northern Nordic Seas, Eastern North Atlantic and Labrador Sea. The higher model resolution has mostly smaller root mean square errors (RMSE) compared to the low model resolution configuration (Fig. 3). This can be clearly seen for salinity at the upper levels at the Nordic Seas North and for temperature in Labrador Sea. This is due to i) an improved mean circulation structure and ii) eddies that are partially resolved or permitted in the higher resolution (Wekerle *et al.*, 2017[20]). Looking at the different experiments, the impact of Greenland freshwater in the offshore regions is small compared to the impact of a higher model resolution. At the upper 100 m in the Nordic Seas, there is an improvement regarding the RMSE for high resolution salinity of more than 11 % when including Greenland freshwater. The Greenland experiment for LOW shows some improvements at the upper layers of the Eastern North Atlantic. One reason we do not see larger freshwater effects may be that the areas defined here, based on bathymetry, are too large and potential smaller signals are removed due to averaging effects.

However, we find areas where the inclusion of Greenland freshwater plays a stronger role compared to model resolution, e.g. for steric height changes in Baffin Bay (Fig. 4). In the early 1990s, the interannual signal of the Greenland freshwater experiments is close

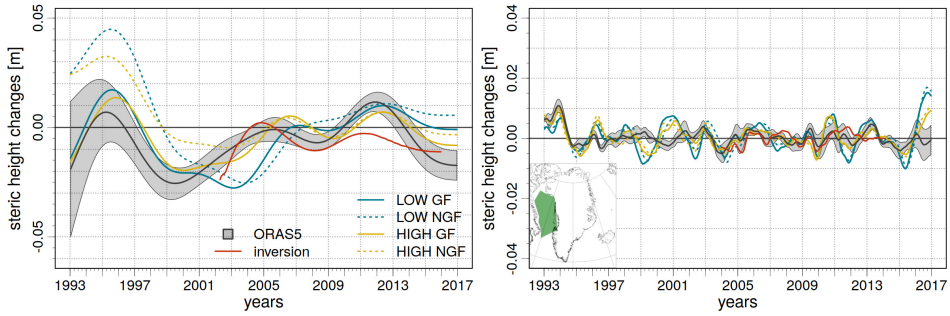


Figure 4. Interannual, five years filtered (left) and month-to-month (right) signal of steric height changes in Baffin Bay, west of Greenland (green area in subplot). GF refers to the Greenland freshwater experiment and NGF to the reference simulations. The gray shading represents the ORAS5-ensemble. Inversion results based on gravimetric and altimetric data are available from 2002 as they depend on GRACE data. For the inversion month-to-month signal, the first and last two years of the time series are not shown due to filter effects at the margins. Source: modified from Stolzenberger *et al.* (under review)[9].

to the ORAS5-ensemble compared to the corresponding reference simulations. We find a benefit of around 37 % for LOW and 32 % for HIGH when including Greenland freshwater with respect to the ORAS5 ensemble spread. From the month-to-month signal (difference between filtered and original time series), we find time periods where the Greenland freshwater experiments are within the ORAS5 ensemble spread (e.g. years 2003-2006). However, the freshwater experiment and the reference run show a similar evolution. The inversion output is comparable to ORAS5. The difficulty for the inversion framework is that Baffin Bay is an area that is seasonally covered with sea ice. This can cause problems for altimeter missions, as estimates can be biased due to missing data during the ice-covered periods.

4 Concluding Remarks and Outlook

4.1 Summary

In this study we investigate the effects of Greenland freshwater on the North Atlantic. For this we use the sea ice-ocean model FESOM. FESOM uses unstructured, triangular meshes that are well suited for the Greenland land-ocean geometry. We perform two different experiments: one where we include a Greenland freshwater data set and one reference run where we exclude Greenland freshwater. We find freshening effects in Baffin Bay, warming effects at the west coast of Greenland and cooling effects south of Greenland. Baffin Bay is also a sensitive region for steric height changes, where we find good agreement between the Greenland freshwater simulations and the ORAS5-ensemble and the satellite data output (at least for several years). These two experiments are performed with one low resolution configuration and one high resolution configuration to see how sensitive ocean variables are with respect to the horizontal mesh resolution. When we compare with Argo float observations, it seems that the model resolution has a larger impact on the results compared to the inclusion of Greenland freshwater. This can be due to uncertainties in the observations, in the simulations, and the spatial aggregations within the predefined areas. However, this can be the basis for further studies optimising spatial fields where freshwater signatures are identified.

4.2 Linking Oceanography to Machine Learning Techniques

The opportunities offered by Machine Learning are currently becoming more and more popular in the geo- and ocean science community. One example for combining oceanography to deep learning methods is the ongoing DFG funded project EDDY [b]. This project looks at the identification, tracking and classification of ocean eddies, i.e. cyclonically or anti-cyclonically rotating water masses generated by ocean currents, from along-track altimetry and not from gridded maps. To do this, deep neural network architectures such as transformers, which have recently become popular in the field of computer vision, are used. Eddy statistics from high-resolution ocean model simulations can provide a significant additional input to these deep learning algorithms.

[b] https://gepris.dfg.de/gepris/projekt/444762031

Acknowledgements

We acknowledge the funding from the Federal Ministry of Education and Research in Germany (BMBF) through the research programme "GROCE" (FKZ 03F0778E) and "GROCE-2" (FKZ 03F0855F) . 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. The IMBIE Team *Mass balance of the Greenland Ice Sheet from 1992 to 2018*, Nature, **579**, 233-239, 2020.
2. R. Rietbroek, S. E. Brunnabend, J. Kusche, J. Schröter, C. Dahle, *Revisiting the contemporary sea-level budget on global and regional scales*, Proceedings of the National Academy of Science, **113(6)**, 1504, 2016.
3. R. Gerdes, W. Hurlin, S. M. Griffies, *Sensitivity of a global ocean model to increased runoff from Greenland*, Ocean Modelling, **12(3-4)**, 416-435, 2006.
4. S. E. Brunnabend, J. Schröter, R. Timmermann, R. Rietbroek, J. Kusche, *Modeled steric and mass-driven sea level change caused by Greenland Ice Sheet melting*, Journal of Geodynamics, **59-60**, 219-225, 2012.
5. J. V. Turton, T. Mölg, D. van As, *Atmospheric processes and climatological characteristics of the 79N glacier (northeast Greenland)*, Monthly Weather Review, **147(4)**, 1375-1394, 2019.
6. J. Schaffer, T. Kanzow, W. J. von Appen, L. von Albedyll, J. E. Arndt, D. H. Roberts, *Bathymetry constrains ocean heat supply to Greenland's largest glacier tongue*, Nature Geoscience, **13**, 227–231, 2020.
7. O. Huhn, M. Rhein, T. Kanzow, J. Schaffer, J. Sültenfuß, *Submarine meltwater from Nioghalvfjerdsbræ (79 North Glacier), Northeast Greenland*, Journal of Geophysical Research: Oceans **126**, e2021JC017224, 2021.
8. M. T. Kappelsberger, U. Strößenreuther, M. Scheinert, M. Horwath, A. Groh, C. Knöfel, S. Lunz, S. A. Khan, *Modeled and observed bedrock displacements in north-east Greenland using refined estimates of present-day ice-mass changes and densified GNSS measurements*, Journal of Geophysical Research: Earth Surface, **126**, e2020JF005860, 2021.
9. S. Stolzenberger, R. Rietbroek, C. Wekerle, B. Uebbing, J. Kusche, *Simulated signatures of Greenland melting in the North Atlantic: a model comparison with Argo floats, satellite observations, and ocean reanalysis*, Journal of Geophysical Research: Oceans, under review.
10. Q. Wang, S. Danilov, D. Sidorenko, R. Timmermann, C. Wekerle, X. Wang, T. Jung, J. Schröter, *The Finite Element Sea Ice-Ocean Model (FESOM) v.1.4: formulation of an ocean general circulation model*, Geoscientific Model Development, **7**, 663-693, 2014.
11. Q. Wang, S. Danilov, J. Schröter, *Finite element ocean circulation model based on triangular prismatic elements, with application in studying the effect of topography representation*, Journal of Geophysical Research: Oceans, **113**, C05015, 2008.

12. S. Danilov, Q. Wang, R. Timmermann, N. Iakovlev, D. Sidorenko, M. Kimmritz, T. Jung, J. Schröter, *Finite-Element Sea Ice Model (FESIM), version 2*, Geoscientific Model Development, **8**, 1747-1761, 2015.
13. S. Kobayashi, Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, K. Takahashi, *The JRA-55 Reanalysis: General Specifications and Basic Characteristics*, Journal of the Meteorological Society of Japan. Ser. II, **93(1)**, 5-48, 2015.
14. N. V. Koldunov, V. Aizinger, N. Rakowsky, P. Scholz, D. Sidorenko, S. Danilov, T. Jung, *Scalability and some optimization of the Finite-volumE Sea ice-Ocean Model, Version 2.0 (FESOM2)*, Geoscientific Model Development, **12**, 3991-4012, 2019.
15. J. Bamber, A. J. Tedstone, M. D. King, I. M. Howat, E. M. Enderlin, M. R. van den Broeke, B. Noel, *Land Ice Freshwater Budget of the Arctic and North Atlantic Oceans: 1. Data, Methods, and Results*, Journal of Geophysical Research: Oceans, **123(3)**, 1827-1837, 2018.
16. Argo 2020, *Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) [Dataset]*, doi:10.17882/42182.
17. B. Uebbing, J. Kusche, R. Rietbroek, F. W. Landerer, *Processing choices affect ocean mass estimates from GRACE*, Journal of Geophysical Research: Oceans, **124(2)**, 1029-1044, 2019.
18. H. Zuo, M. A. Balmaseda, S. Tietsche, K. Mogensen, M. Mayer, *ECMWF operational ensemble reanalysis-analysis system for ocean and sea ice: a description of the system and assessment*, Ocean Science, **15(3)**, 779-808, 2019.
19. L. Castro de la Guardia, X. Hu, P. G. Myers, *Potential positive feedback between Greenland Ice Sheet melt and Baffin Bay heat content on the west Greenland shelf*, Geophysical Research Letters, **42(12)**, 4922-4930, 2015.
20. C. Wekerle, Q. Wang, S. Danilov, V. Schourup-Kristensen, W. J. von Appen, T. Jung, *Atlantic Water in the Nordic Seas: Locally eddy-permitting ocean simulation in a global setup*, Journal of Geophysical Research: Oceans, **122**, 914-940, 2017.

Molecular Modelling of Hydrothermal Fluids

Sandro Jahn, Maximian Schulze, and Rajorshi Chattopadhyay

Institut für Geologie und Mineralogie, Universität zu Köln,
Zülpicher Str. 49b, 50674 Köln, Germany
E-mail: s.jahn@uni-koeln.de

Hydrothermal fluids are key agents of many geological processes occurring in the Earth's crust and upper mantle. They are involved in chemical transport and mineral replacement reactions as well as in transport of heat within the Earth. However, their structural, thermodynamic and physical properties are not easily accessible due to the high temperature and high pressure conditions as well as the chemical complexity. In collaboration with experimental groups we investigate model hydrothermal fluids using molecular modelling techniques, which includes both classical and *ab initio* molecular dynamics simulations. Here, we present results of recent studies of solute-solvent interactions in supercritical fluids as function of fluid density. We discuss how structural changes in the solvent with increasing density affect solute properties such as the Krichevskii parameter and the partial molar volume. In a second case study, we investigate the speciation of rare earth element cations in supercritical aqueous solutions to understand better the transport mechanism of such metals in hydrothermal systems.

1 Introduction

The geochemical evolution of the Earth's crust and mantle is largely determined by the interactions between aqueous fluid phases and the surrounding rocks. Fluids can move through rocks, e.g., through fractures, and alter the local chemical composition of the environment by dissolving existing minerals or precipitating new minerals from the respective fluid phase. In this sense, both the chemical compositions of the fluid and the composition of the surrounding rocks cannot be considered static in space and time. In natural environments, hydrothermal fluids occur at temperatures of up to several hundred degrees Celsius and densities ranging from less than 0.1 g/cm^3 in parts of geothermal systems up to about 1.4 g/cm^3 in subduction zones[1]. Physical and thermodynamic properties of hydrothermal fluids such as pressure-volume-temperature (P-V-T) equations of state, viscosity and other transport properties, and molecular speciation models are important input parameters for large-scale reservoir and reactive process models. The ability to make reliable predictions about the solubility of chemical components is essential not only for typical "hard rock geology" questions such as the formation of ore deposits, but also for more applied problems such as the sustainable exploitation of geothermal energy systems.

Despite their importance, the thermodynamic and physical properties of hydrothermal fluids are not well constrained at many of the relevant conditions, especially in the high temperature range above the critical point of water (374 °C). To date there is no universal model available that can predict the thermodynamic properties of aqueous fluids under all geologically relevant conditions up to high temperatures and pressures. The existing thermodynamic models are only applicable in a restricted range of fluid density, temperature and chemical compositions, and extrapolations bear a large uncertainty. Especially in the high temperature and low pressure region the most popular thermodynamic model for aqueous electrolytes at high P and T by Helgeson, Kirkham and Flowers, HKF[2] is

not valid anymore. But many natural hydrothermal systems exist exactly in this range of conditions (300 to 700 °C, 0.1 to 100 MPa). This includes magmatic-hydrothermal systems, mid-ocean ridge hydrothermal systems and deep roots of high enthalpy geothermal systems[3].

In this context, conditions at and near the critical point of water pose a particular challenge for the development of new thermodynamic models. The reason for this lies in the strongly divergent nature of certain thermodynamic properties of the solutes such as partial molar volumes, whose derivatives at the critical point approach infinity. Therefore, an important task in the development of next-generation models for supercritical aqueous fluids is to find appropriate correlations between physical quantities that allow to avoid the direct treatment of anomalously behaving properties of solutes under critical point conditions. In this project, we use molecular dynamics simulations to develop structure models for hydrothermal fluids, to predict thermodynamic and physical properties that are not available experimentally, and to reveal the underlying mechanisms of structure-property relations that may help to improve the physical models. The project is focused on two types of hydrothermal fluids: (1) (super)hot geothermal fluids in the context of geothermal energy exploitation and (2) magmatic-hydrothermal fluids in the context of ore-forming processes.

2 Methods

Most of the simulations use the molecular dynamics (MD) method, which is a finite difference method to solve Newton's equations of motion numerically. Atomic interactions are either described by classical interatomic potentials or by electronic structure calculations in the framework of density-functional theory (DFT). Depending on the type of potential, the simulations are referred to as classical MD or *ab initio* MD. Both MD methods described in well-known text books, e.g. by Frenkel and Smit[4] and Marx and Hutter[5]. For the *ab initio* and classical MD simulations on JUWELS we employ community codes such as CP2K[6], CPMD[7] and LAMMPS[8]. All of these codes are well suited to be used on HPC architectures. *Ab initio* MD simulations are computationally very demanding, which limits accessible system sizes to a few hundred atoms and simulation times to tens or hundreds of picoseconds. The scaling of the *ab initio* codes is usually quite good up to a few hundred CPU cores. On JUWELS, we typically use 8 nodes and 384 cores for production runs. The use of supercomputers for classical MD simulations allows to study systems with thousands of atoms and to explore the parameter space (density, temperature, chemical composition) to gain a systematic understanding of the investigated system. Especially for aqueous systems, changing conditions lead to continuous changes of the structure and physical properties, which makes use of HPC resources indispensable. Scaling of classical MD simulations using LAMMPS is reasonable up to about 100 cores.

As MD approaches do not provide direct access to Gibbs or Helmholtz energies, the derivation of such thermodynamic properties requires more advanced simulation techniques. Thermodynamic integration is a well known method of statistical mechanics to obtain the change of Helmholtz free energy F of a system in terms of changes in the total energy U[9]. Knowing the free energy of a reference system (0), the free energy of the unknown system (1) is obtained by stepwise integrating system 0 towards system 1:

$$F_1 = F_0 + \int_0^1 \langle \mathrm{d}U_\lambda / \mathrm{d}\lambda \rangle_\lambda \, \mathrm{d}\lambda \tag{1}$$

For determining the stability constants of aqueous species we usually employ the potential of mean force approach, in which short MD simulations are performed with one or more constrained atomic distances. For the bond of interest the distance is increased in steps from the equilibrium distance of a contact ion pair to a distance where the mean force between the two atoms becomes negligible. Integration over the mean force then yields the change in Helmholtz free energy. Another approach to estimate the free energy of chemical reactions is metadynamics[10,11]. Metadynamics allows to overcome these hindrances and enforce chemical reactions by adding a time dependent bias potential function of so-called collective variables to the free energy surface of the reaction. In this way the configurational space can be sampled in an affordable computation time. For MD simulations in conjunction with advanced sampling methods the PLUMED library[12] is used.

3 Structure and Thermodynamic Properties of Superhot Geothermal Fluids

Rigorous thermodynamic models are crucial for understanding the properties of geological fluids as part of planning for exploration, design and operation of geothermal energy facilities. However, these models are currently incomplete and do not give accurate enough results for reliable planning. Existing thermodynamic models were originally developed over 20 years ago and are increasingly unable to cope with modern demands, both in computational practicality and in resolving scientific and technological questions on project planning and geothermal field management. Specifically, the relevant chemical and physical geothermal fluid properties are described by sets of equations of state across the relevant range of temperature, pressure, and composition. Currently, there is no generally applicable equation of state for predicting the thermodynamic properties of superhot aqueous fluids, which are fluids that remain in the single phase during isenthalpic decompression, i.e. they do not unmix into a gas and a liquid phase. Within the EU-funded HORIZON 2020 project GEOPRO, our main goal is to explore suitable ways to establish such an equation of state, with special attention to conditions close to the critical point of water.

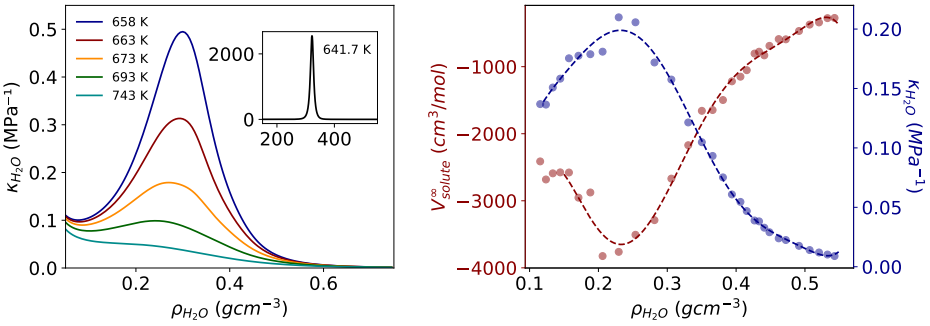


Figure 1. (a) Isothermal compressibility of water κ_{H_2O} from IAPWS95[13] as a function of density at different temperatures close to the critical point, and (b) partial molar volume $\overline{V}^{\infty}_{solute}$ and κ_{H_2O} from classical MD simulations at 673 K.

From the computational chemistry side, we approach this problem with the help of classical MD simulations using the LAMMPS code[8]. Salt solutions at infinite dilution are simulated using the TIP4P/2005 water model[14] in conjunction with the solute parameters from Joung and Cheatham[15]. In a first step, we study the relation between the partial molar volume of a solute at infinite dilution $\overline{V}^{\infty}_{solute}$ and the isothermal compressibility of water κ_{H_2O}, which is given by the following equation:

$$\overline{V}^{\infty}_{solute} = V_{H_2O}(1 + \kappa_{H_2O} \cdot \gamma) \tag{2}$$

where V_{H_2O} is the molar volume of water and γ is called the Krichevskii parameter. Considering the well-behaved nature of γ close to the critical point (see below), it becomes obvious from Eq. 2 that the strongly divergent character of $\overline{V}^{\infty}_{solute}$ must arise from the isothermal compressibility, which is a pure solvent property. In Fig. 1a we show κ_{H_2O} as taken from the IAPWS95 equation of state[13] for various isotherms in the vicinity of the critical point. It can be seen that κ_{H_2O} exhibits a sharp peak at and near the critical point, which gradually disappears with increasing T. To ensure that the interaction models of our choice can reproduce the anomalous behaviour of solvent and solutes, we first calculate κ_{H_2O} from the variance of volume fluctuations in an NPT ensemble (i.e. constant number of particles N, P and T) simulation

$$\kappa_{H_2O} = \frac{\langle V^2 \rangle - \langle V \rangle^2}{\langle V \rangle k_B T} \tag{3}$$

where k_B is the Boltzmann constant. Then, the partial molar volume of NaCl is determined by running a sequence of NPT simulations with different solute concentrations or mole fraction, X. The volume, V, of the simulation box changes with an increasing number of solutes. Accordingly, the partial molar volume can be estimated as the slope of a linear fit through the resulting V vs. X plot. The result for the 673 K isotherm is shown in Fig. 1b. The near critical behaviour of both, solute and solvent is reasonably reproduced by the chosen approach. We now come back to the Krichevskii parameter γ, which is defined as the change of the pressure caused by the insertion of a solute into the solvent at infinite dilution

$$\gamma = \left[\frac{\partial P}{\partial X}\right]_{VT} \tag{4}$$

where the subscripts refer to constant volume and temperature. As already mentioned, γ describes a smooth function even close to the critical point, making it a promising key variable for new thermodynamic models. We compute γ at a given density by running a sequence of MD simulations with a varying number of solutes in the NVT ensemble. The density dependence of γ is shown in Fig. 2a for NaCl at different temperatures, and in Fig. 2b at constant temperature for different solutes. As can be seen, γ always takes on negative values at low densities, which emphasises the attractive nature of the solute–solvent interactions under these conditions. At higher densities, the values of γ are positive and increase sharply as a function of pressure, showing that solute-solvent interactions have a pronounced repulsive character. Furthermore, our results show that both solute type and temperature have less influence on the behaviour of γ at low densities than at higher densities. The sign of γ determines whether the partial molar volume of the solute is negative or positive. Among other things, our ongoing research is focused on developing a

comprehensive, mechanistic picture of these observations that can be used as a molecular-scale guideline for future thermodynamic models for supercritical aqueous fluids.

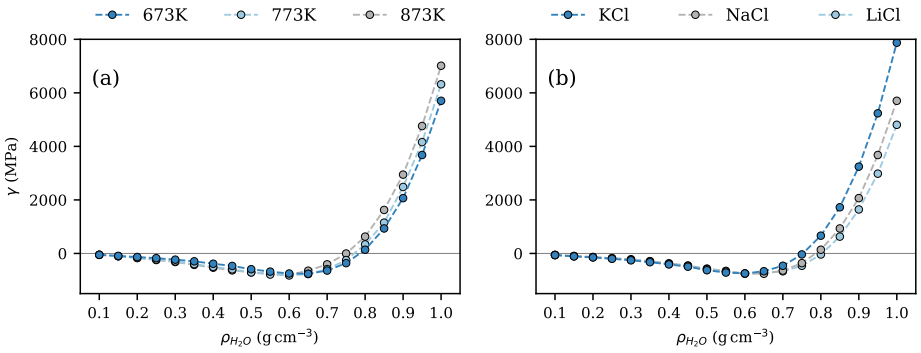


Figure 2. Density dependence of the Krichevskii parameter γ (a) for NaCl as a function of temperature and (b) for different solutes at 673 K.

4 Speciation and Transport of REE in Hydrothermal Systems

Rare earth elements (REE) are highly demanded metals with a wide range of applications, e.g., for new technologies in the energy and environmental sectors[16]. Many products of our daily life such as smart phones or computers critically depend on the availability of REE. However, exploitable resources of REEs are largely concentrated in China, which makes their supply vulnerable and therefore REEs are considered critical metals. Although REEs occur in most minerals of the Earth's crust, their abundance in exploitable concentrations is limited to specific rocks such as carbonatites or pegmatites. Hydrothermal fluids are considered important for the mobilisation and concentration of REEs in natural systems. To understand the formation of deposits with REE concentrations high enough to be economically exploitable the underlying processes of hydrothermal transport and concentration need to be studied.

Fluid inclusions in natural mineralisations[17] provide evidence for the composition and temperature at which initial mineral precipitation started. Usually such hydrothermal fluids are rich in NaCl, KCl and/or sulfate (up to 44 wt% NaCl) but other ligands such as hydroxide, phosphate and carbonate groups may also be important[16]. Metal transport by hydrothermal fluids is largely determined by metal complexation in the respective fluids. So far, the thermodynamic description of REE speciation in high temperature hydrothermal fluids mainly relied on extrapolations from thermodynamic data at ambient conditions[18]. Experimentally, the speciation of REEs in hydrothermal fluids at high temperatures has been studied, e.g. by *in situ* UV-vis and X-ray absorption spectroscopy (XAS) or by solubility measurements[16]. Due to technical challenges only a few of these experiments have been performed at temperatures above 300 °C.

We have been using *ab initio* and classical MD simulations to study the REE complexation in aqueous fluids at hydrothermal conditions. Fig. 3 shows, as an example, structural

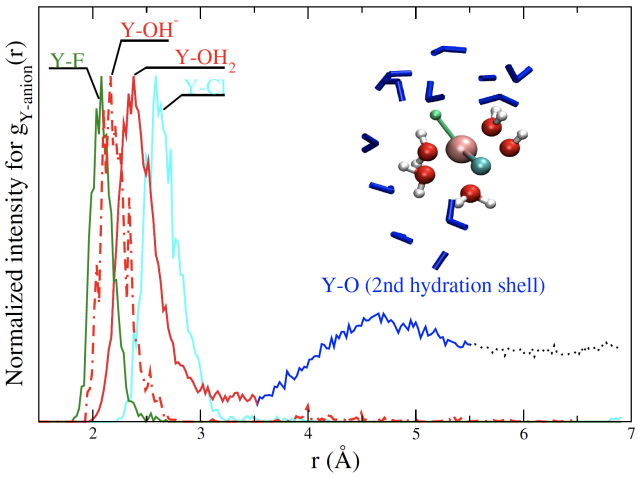


Figure 3. Partial radial distribution functions $g_{Y-anion}(r)$ between Y and three different anions (O, Cl and F) scaled to maximum peak height and snapshot of a [YClF(OH)(H_2O)$_4$] complex from *ab initio* molecular dynamics simulations at 1.3 GPa and 800 °C. The first peak positions reflect the different cation-anion distances. The second peak in $g_{Y-O}(r)$ is indicative of the second hydration shell (modified from Ref. 19).

information of Y speciation in mixed fluoride and chloride supercritical solutions from *ab initio* MD simulations using the CPMD code[19]. Peaks of the renormalised partial radial distribution functions exhibit a systematic shift of the metal ion – ligand distance from Y–F to Y–Cl. A second hydration shell formed by the solvent molecules can also be identified. In order to obtain the stability constants of different complexes, the dissociation constants were derived using the potential of mean force approach of thermodynamic integration. This involves determining the change in free energy during stepwise removal of individual ligands (e.g. Cl or F). In practice, this is done by performing constraint molecular dynamics simulations in which the distance between the cation and the dissociating anion is held constant and the average force between these two ions is determined. Repeating this procedure for different distances between the equilibrium distance of the associated complex and a distance at which the complex is considered dissociated, the potential of mean force is obtained. The free energy of dissociation is given by the integral of this potential. Stability constants describe the reverse process of stepwise association, but they are easily derived from the (standardised) dissociation energies. Species distribution diagrams are constructed based on the computed species stability constants (Fig. 4). In the example presented here[19], the fluoride species become significant and increasingly dominant at fluoride concentrations above a few millimole per kg H_2O.

The *ab initio* MD simulation approach is usually considered a rather accurate and predictive method and has been used recently for investigations of REE complexation by others as well[20,21]. In our current project, we not only extend the simulations of REE complexation from to other relevant hydrothermal conditions and to other ligands and REEs but we also address some of the fundamental problems that remain. One major issue is related to the finite size of the simulation box in *ab initio* MD simulations, which typically

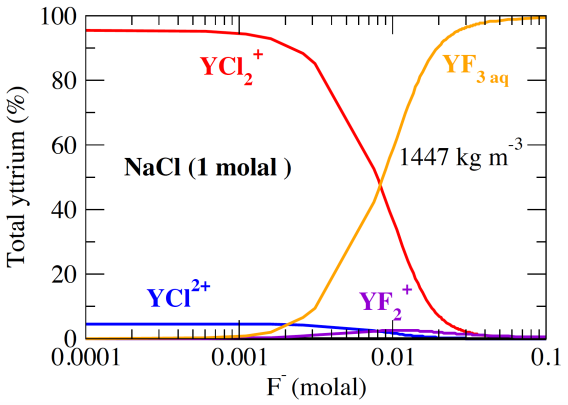


Figure 4. Yttrium speciation model in a fluoride-bearing 1 molal NaCl solution at subduction zone conditions ($P = 4.5$ GPa, $T = 800\,^{\circ}$C) as a function of fluoride concentration (modified from Ref. 19).

has cell lengths in the order of 16 to 18 Å. Thus, the maximum separation between dissociated anions and cations is less than 10 Å due to periodic boundary conditions. Potential of mean force simulations using both classical and *ab initio* approaches show a minimum at the typical cutoff distance of 8 Å and often this is interpreted as convergence of the free energy integration. Preliminary classical MD simulations of much larger simulation boxes with cell lengths of about 40 Å performed with Lennard Jones potential parameters from Migliorati *et al.* (2017)[22] suggest that full convergence is obtained at significantly larger distances. However, the classical force fields are not reactive and their applicability to hydrothermal conditions remain questionable. Therefore new approaches such as machine-learning potentials[23] or reactive force fields, such as ReaxFF[24], appear very attractive but so far they have not been applied in the context of geological fluids. Both methods have a high potential for bridging the methodological gap between the existing DFT-based *ab initio* approach and conventional classical potentials.

5 Concluding Remarks

Ab initio and classical molecular dynamics simulations are well-suited approaches for studying the structural and thermodynamic properties of hydrothermal aqueous fluids. They provide unique insights into structure-property relationships in fluids under extreme conditions, which are used as input for the construction and parameterisation of advanced thermodynamic models for solvents and solutes in hydrothermal systems. Aware of their limitations, finite size and time effects of *ab initio* MD simulations, and the problems with accuracy and transferability when using classical force fields, new types of interaction models need to be developed. Reactive force fields and neural network potentials are first candidates in this respect. They would open new perspectives for the study of complex geological fluids and their interaction with minerals and rocks. HPC resources are essential for this endeavour.

Acknowledgements

We would like to thank former group member Johannes Stefanski for his contributions to this work. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (GCS) 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). Part of this work was funded by EU2020 project GEOPRO, grant agreement ID 851816, and DFG Priority Program SPP2238 Dynamics of Ore Metals Enrichment - DOME, project ID JA1469/13-1.

References

1. C. E. Manning, *The chemistry of subduction-zone fluids*, Earth Planet. Sci. Lett., **223**, 1–16, 2004.
2. H. C. Helgeson, D. H. Kirkham, and G. C. Flowers, *Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures: IV. Calculation of activity coefficients, osmotic coefficients, and apparent molal and standard and relative partial molal properties to 600 °C and 5 kbar*, Am. J. Sci., **281**, 1249–1516, 1981.
3. T. Driesner, *The molecular-scale fundament of geothermal fluid thermodynamics*, Rev. Mineral. Geochem., **76**, 5–33, 2013.
4. D. Frenkel and B. J. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, San Diego, 2002.
5. D. Marx and J. Hutter, *Ab initio molecular dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, 2009.
6. T. D. Kühne, M. Iannuzzi, M. D. Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borstnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter, *CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations*, J. Chem. Phys., **152**, 194103, 2020.
7. J. Hutter and M. Iannuzzi, *CPMD: Car-Parrinello molecular dynamics*, Z. Krist., **220**, 549–551, 2005.
8. A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, *LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales*, Comp. Phys. Comm., **271**, 10817, 2022.
9. M. J. Gillan, D. Alfe, J. Brodholt, L. Vocadlo, and G. D. Price, *First-principles modelling of Earth and planetary materials at high pressures and temperatures*, Rep. Prog. Phys., **69**, 2365–2441, 2006.
10. A. Laio and M. Parrinello, *Escaping free-energy minima*, PNAS, **99**, 12562–12566, 2002.

11. G. Bussi and D. Branduardi, *Free-energy calculations with metadynamics: theory and practice*, Rev. Comput. Chem., **28**, 1–49, 2015.
12. The PLUMED consortium, *Promoting transparency and reproducibility in enhanced molecular simulations*, Nat. Methods, **16**, 670–673, 2019.
13. W. Wagner and A. Pruss, *The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use*, J. Phys. Chem. Ref. Data, **31**, 387–535, 2002.
14. J. L. F. Abascal and C. Vega, *A general purpose model for the condensed phases of water: TIP4P/2005*, J. Chem. Phys., **123**, 234505, 2005.
15. I. S. Joung and T. E. Cheatham III, *Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations*, J. Phys. Chem. B, **112**, 9020–9041, 2008.
16. A. Migdisov, A. E. Williams-Jones, J. Brugger, and F. A. Caporuscio, *Hydrothermal transport, deposition, and fractionation of the REE: Experimental data and thermodynamic calculations*, Chem. Geol., **439**, 13–42, 2016.
17. M. Wilson, *Transferable ionic simulation models from ab initio calculations*, Comp. Mat. Science, **17**, 234–238, 2000.
18. S. A. Wood, *The aqueous geochemistry of the rare-earth elements and yttrium*, Chem. Geol., **88**, 99–125, 1990.
19. J. Stefanski and S. Jahn, *Yttrium speciation in subduction-zone fluids from ab initio molecular dynamics simulations*, Solid Earth, **11**, 767–789, 2020.
20. Q. Guan, Y. Mei, B. Etschmann, D. Testemale, M. Louvel, and J. Brugger, *Yttrium complexation and hydration in chloride-rich hydrothermal fluids: A combined ab initio molecular dynamics and in situ X-ray absorption spectroscopy study*, Geochim. Cosmochim. Acta, **281**, 168–189, 2020.
21. Q. Guan, Y. Mei, B. Etschmann, D. Testemale, E. Bastrakov, and J. Brugger, *Yttrium speciation in sulfate-rich hydrothermal ore-forming fluids*, Geochim. Cosmochim. Acta, **325**, 278–295, 2022.
22. V. Migliorati, A. Serva, F. M. Terenzio, and P. D'Angelo, *Development of Lennard-Jones and Buckingham Potentials for Lanthanoid Ions in Water*, Inorg. Chem., **56**, 6214–6224, 2017.
23. J. Behler, *Four Generations of High-Dimensional Neural Network Potentials*, Chem. Rev., **121**, 10037–10072, 2021.
24. T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, and A. C. T. van Duin, *The ReaxFF reactive force-field: development, applications and future directions*, npj Comput. Mater., **2**, 15011, 2016.

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. The first one concerns the modelling of biological cell populations and the second one semantic segmentation.

In the first contribution "Massively Parallel Likelihood-Free Parameter Inference for Biological Multi-Scale Systems" by E. Alamoudi *et al.* multi-scale multi-cellular models are used to describe organisation and development of cell populations and tissues, e.g. viral infection or tumour development. Typically, such models depend on unknown parameters that need to be estimated on experimental data, in order to make predictions or compare models. The main tool for getting numerical approximations of solutions of the underlying models are the Approximate Bayesian Computation (ABC) method, which is rooted in Bayesian statistics. In all model-based statistical interference, the likelihood function expresses the probability of the observed data, is of central importance and will help to choose particular values of parameters and to choose the model among different models. For simple models the likelihood function can be specified explicitly, for more complex models the calculation of the likelihood function is computationally very expensive. The ABC method avoids the calculation of the likelihood function and still allows to estimate parameters and to select models. Particularly for applications from biology this method has proved itself.

In this paper the authors first briefly summarise the underlying methodology, they describe concrete implementations of different algorithms via software packages developed by the authors and describe the realisation on high performance computers.

This work was supported by the German Federal Ministry of Education and Research (BMBF) and the German Research Foundation (DFG).

In the second contribution "Uncertainty Quantification and Resource-Demanding Computer Vision Applications of Deep Learning" by J. Burghoff *et al.* different current projects that use GPU compute resources on the JUWELS are described. These projects range from the detection of unknown objects in semantic segmentation, domain adaptation with generative adversarial networks (GANs), active learning for object detection to automated network design. The first one means that the algorithm is able to assign each pixel in a given image to an object category, e.g. to a car or a human person. The algorithm is based on "deep learning" and neural networks. In particular the algorithm is able to handle objects from classes that are semantically unknown as these objects are mostly associated with high model prediction uncertainty. For the domain adaptation, they demonstrate the performance of the semi-supervised method they have developed using the example of street scenes. It turns out, that even though they still find some artefacts in the transformed images, an improvement in the overall scene understanding can be seen. In addition, they present an active learning approach based on previous work in uncertainty quantification

for deep object detection and a method to do rapid prototyping to compare to a wide range of baselines for a compute demanding task. Finally, they investigate network architecture optimisation to automatically find network architectures with high accuracies but few parameters where no readily trained networks are available. All methods typically require plenty of GPU resources.

This project is in part funded by the German Federal Ministry for Economic Affairs and Climate Action, and by the state Ministry of Economy, Innovation and Energy of North Rhine-Westphalia (MWIDE) and the European Fund for Regional Development.

Massively Parallel Likelihood-Free Parameter Inference for Biological Multi-Scale Systems

Emad Alamoudi[1], Jörn Starruß[2], Nils Bundgaard[3], Robert Müller[2], Felipe Reck[1], Frederik Graw[3,4], Lutz Brusch[2], Jan Hasenauer[1,5,6], and Yannik Schälte[1,5,6]

[1] Faculty of Mathematics and Natural Sciences, Rheinische Friedrich-Wilhelms-Universität Bonn, 53115 Bonn, Germany
E-mail: {emad.alamoudi, jan.hasenauer, yannik.schaelte}@uni-bonn.de

[2] Center of Information Services and High Performance Computing (ZIH), Technische Universität Dresden, 01062 Dresden, Germany

[3] BioQuant - Center for Quantitative Biology, Heidelberg University, 69120 Heidelberg, Germany

[4] Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany

[5] Institute of Computational Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany

[6] Center for Mathematics, Technische Universität München, 85748 Garching, Germany

Biological processes such as virus spread and tumour growth are highly complex. In order to accurately describe and understand such systems, increasingly complex models are being developed. In particular, multi-scale multi-cellular models are used to describe organisation and development of cell populations and tissues. Typically, such models depend on unknown parameters that need to be calibrated on experimental data, in order to test hypotheses and make predictions. Approximate Bayesian Computation (ABC) is a likelihood-free parameter inference method applicable to a wide range of models, where other methods fail. However, it is computationally demanding, such that computation time is often the limiting factor, especially for expensive multi-scale models. In this article, we demonstrate how the use of high-performance computing (HPC) infrastructure together with efficient parallelisation strategies enables ABC inference on challenging application problems. Further, we demonstrate how HPC enables the development and study of novel algorithms.

1 Introduction

Quantitative mechanistic models are important tools in many research areas to describe, study, and understand real-world systems[1,2]. Commonly, these models depend on unknown parameters. In order to test hypotheses, make predictions, or compare models, these parameters need to be calibrated on experimental data[3]. The Bayesian framework allows doing so by combining prior beliefs about parameters with the likelihood of data given parameters. However, especially for complex mechanistic models, evaluating the likelihood quickly becomes computationally infeasible[4,5]. Approximate Bayesian computation (ABC) is a widely applicable parameter estimation method that, in a nutshell, circumvents likelihood evaluation by simulating data from the model, and accepting corresponding parameters if a distance between simulated and observed data is sufficiently small[6].

A particular model type of interest are multi-scale multi-cellular models, which are used to describe interacting processes on a cellular or tissue level, such as viral infection

or tumour development. Such models provide a realistic description and are often easy to formulate as a computer program, taking an agent-based perspective. However, they can be highly complex, often combining different sub-model types to describe different sub-processes, and can be computationally expensive[4,7,8].

In this article, we first briefly summarise the underlying methodology. Then, we outline concrete implementations of the various algorithms via software packages developed by the authors, and outline their deployment and parallelisation on HPC infrastructure. Then, we give concrete biological applications demonstrating how bringing ABC methods to HPC infrastructure allows to tackle new problems in the life sciences, and further how it allows to develop and test novel algorithms, specifically on outlier-robust adaptive distances, and the use of inverse machine learning models to construct summary statistics in ABC.

2 Sequential Approximate Bayesian Computation

Consider a mechanistic model specified via its likelihood $\pi(y|\theta)$ of observing data $y \in \mathbb{R}^{n_y}$ given model parameters $\theta \in \mathbb{R}^{n_\theta}$. The parameters are assumed unknown and can comprise e.g. reaction rates or initial concentrations. Assuming we have observed experimental data $y_{\text{obs}} \in \mathbb{R}^{n_y}$, the goal of Bayesian inference is to obtain information about the posterior distribution

$$\pi(\theta|y_{\text{obs}}) \propto \pi(y_{\text{obs}}|\theta)\pi(\theta),$$

where $\pi(\theta)$ encodes beliefs prior to observing the data. Unlike many other methods, ABC is still applicable when evaluating the likelihood is computationally infeasible, and only requires that the model is generative, i.e. one can generate data $y \sim \pi(y|\theta)$ given parameters $\theta \sim \pi(\theta)$. This amounts to assuming that the model exists as an executable, potentially stochastic, forward simulation program. Then, at its core, ABC consists of three steps (Fig. 1)[10,5]: First sample parameters $\theta \sim \pi(\theta)$, second simulate data $y \sim \pi(y|\theta)$, and third accept (θ, y) if $d(y, y_{\text{obs}}) \leq \varepsilon$. Here, d is a distance measure comparing simulated and observed data, and $\varepsilon \geq 0$ is an acceptance threshold. This is repeated until sufficiently many, say N, particles have been accepted. Denoting $\pi_{\text{ABC},\varepsilon}(\theta|y_{\text{obs}})$ the distribution of

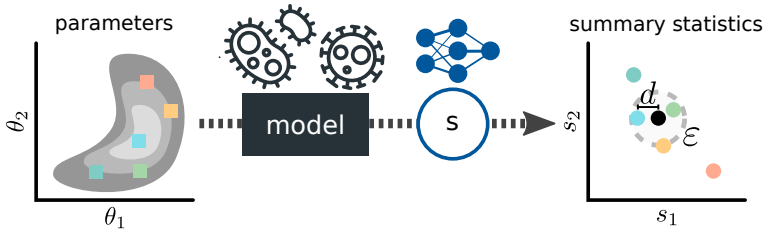


Figure 1. **Basic ABC algorithm.** Parameters $\theta \sim \pi(\theta)$ are sampled from the prior or a proposal distribution, and passed to a black-box model generating potentially stochastic simulated data according to the likelihood $y \sim \pi(y|\theta)$. These are optionally passed through a summary statistics function giving a low-dimensional representation $s(y)$. Summary statistics of simulated and observed data are compared via a distance metric d, and the underlying parameters accepted if the distance is below an acceptance threshold ε. This figure is taken from Schälte *et al.*, 2022[9].

accepted particles, it can be shown under mild conditions that $\pi_{\mathrm{ABC},\varepsilon}(\theta|y_{\mathrm{obs}}) \to \pi(\theta|y_{\mathrm{obs}})$ for $\varepsilon \to 0$ in an appropriate sense[11]. To tackle the curse of dimensionality, the comparison of simulations and data is often in terms of low-dimensional summary statistics $s(y)$ in place of y.

ABC is frequently combined with a sequential Monte Carlo (SMC) scheme[12,13]. In ABC-SMC, a series of particle populations $P_t = \{(\theta_i^t, y_i^t, w_i^t)\}_{i \leq N}$, $t = 1, \ldots, n_t$, are generated, with thresholds $\varepsilon_1 > \ldots > \varepsilon_{n_t}$. In generation t, importance sampling based on generation $t-1$ is used, with Radon-Nikodym derivatives w^t. It successively reduces the acceptance threshold and thus improves the posterior approximation, while maintaining high acceptance rates.

3 Implementation

3.1 Systematic Inference for Multi-Scale Multi-Cellular Models via FitMultiCell

We developed a scalable, modular ABC-SMC framework in the open-source Python toolbox pyABC [a][9]. Beyond the core ABC-SMC routine sketched above, pyABC provides various algorithms to adapt to the problem structure and automatically tune hyperparameters, such as transitions kernels, acceptance thresholds, and population sizes[14,15]. pyABC can be used with models written in arbitrary languages, with tight support of Python, R, and Julia. See Klinger *et al.*, 2018, and Schälte *et al.*, 2022, for details[16,9].

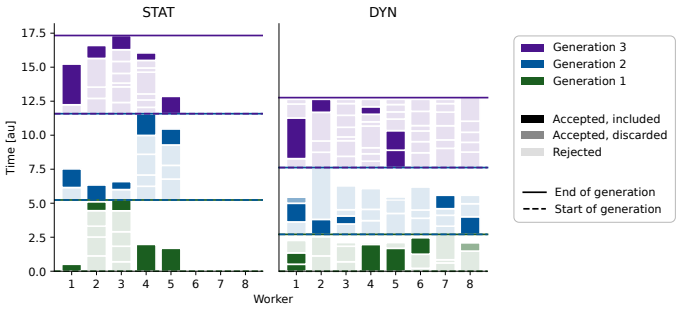


Figure 2. **Static (STAT) and dynamic (DYN) scheduling.** For 3 sequential populations of $N = 5$ particles on $W = 8$ workers. The colour shades from dark to light indicate whether a sample satisfies the acceptance criterion and is included in the population, satisfies the acceptance criterion but is discarded from the population due to late start, or does not satisfy the acceptance criterion and is rejected. For details, see Klinger *et al.*, 2018[16].

The forward model $\pi(y|\theta)$ is to (py)ABC a black-box generating data from parameters. A particular application are multi-scale and multi-cellular models of cell populations and tissues. The open-source C++ toolbox Morpheus [b] is a graphical user interface (GUI) based modelling and simulation environment for such models. It allows to couple various modelling formalisms, including ordinary, partial, and stochastic differential equations,

[a] `https://github.com/icb-dcm/pyabc`
[b] `https://morpheus.gitlab.io`

and cellular Potts models, to describe interactions of cells with each other and intercellular substances. See Starruß *et al.*, 2014, for details[8].

To enable systematic parameter inference for multi-scale multi-cellular models, we connected the tools pyABC and Morpheus in the platform FitMultiCell [c], using a standardised interface of parameter definitions and simulation outputs, as well as a problem formulation based on the PEtab standard[17].

3.2 Parallelisation

Each ABC-SMC population consists of conditionally independent particles, which allows to parallelise the sampling. pyABC currently supports two strategies of parallelisation for both multi-processing and multi-machine distributed execution: Static (STAT) and dynamic (DYN) scheduling (Fig. 2). STAT, which is implemented in many ABC tools, generates only as many sampling tasks and processes as needed, N. However, this does not always use all available cores, especially leaving cores idle while waiting for a few to finish. This is especially a problem on common HPC infrastructure, where resources cannot be easily dynamically re-assigned, leading to inefficient resource usage. In comparison, DYN continues sampling on all available cores until sufficiently many particles have been accepted, afterwards correcting for execution time bias. Thus, it uses all available cores at almost all times, except shortly at the end of generations. DYN has been shown to reduce the overall wall-time substantially over STAT on distributed environments[16]. In all numerical experiments in this work, we used DYN for parallelisation.

3.3 Automated Resource Allocation on HPC Systems

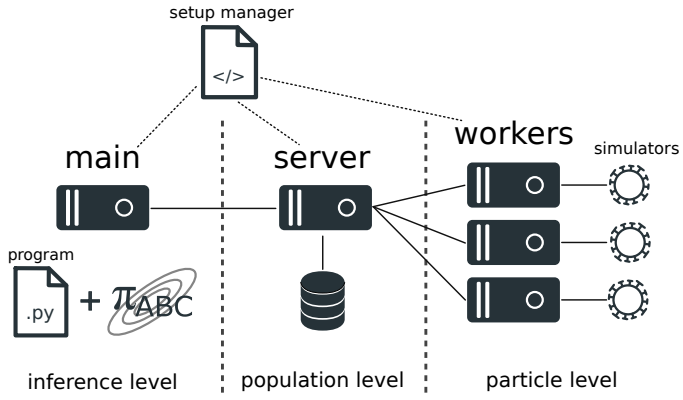


Figure 3. **Distributed ABC-SMC architecture.** The main inference process executes the actual ABC routine via pyABC and an execution program. It transmits tasks to generate particle populations to the central server, from which a number of workers read. The workers then perform the actual model simulations single- or multi-core, e.g. via Morpheus. A central script manages the architecture setup.

[c] e.g. `https://gitlab.com/fitmulticell/fit`

As part of the FitMultiCell effort, we wrote an automatised job management, resource allocation, and execution system for SLURM workload managers, which are in use on many computer clusters. It maintains three types of processes (Fig. 3): a server process that acts as a central broker to distribute work across workers, a number of worker processes that wait for tasks to perform, and an inference process to execute the actual ABC routine. The processes are linked in a hierarchical way uncommon in classical distributed execution applications. The server process runs per default a Redis [d] server that builds a central in-memory database acting as a message broker. Once the server has been launched, its address is dynamically retrieved and passed along with a specified port to all worker processes, enabling their unidirectional communication with the server via Pub/Sub. Simultaneously with starting the workers, the inference process is started, which executes the ABC routine. In each SMC generation, the inference process sends a request to simulate particles to the server, and processes the results once the server has gathered N successful answers from the workers.

Number and resources of worker processes, port, execution time, and target program can be specified via a central shell script [e]. Resources can be shared among the three process types to ensure full exploitation. As (py)ABC is agnostic of model specifics, it is straightforward to replace Morpheus by other simulation programs.

4 Case Studies

4.1 Application Examples

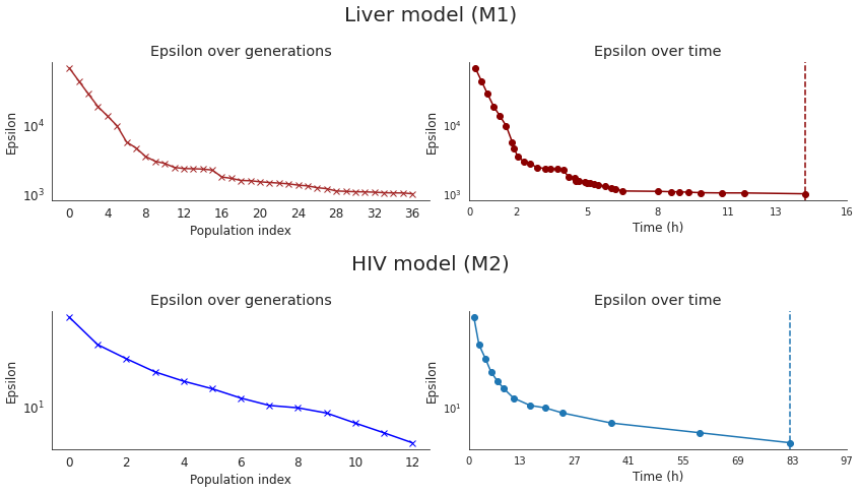


Figure 4. **Acceptance thresholds.** Over generations (left), and over time (right), for models M1+2.

[d] https://redis.io
[e] https://gitlab.com/fitmulticell/Distribution_manager

To assess and validate the FitMultiCell pipeline, we considered two models of different biological processes: Model M1 describes the process of mouse liver regeneration[18]. On different sections of the liver tissue, the mean of yes-associated protein (YAP), which drives the regeneration process, was quantified in 10 zones at different time points. These data were used as observation data for fitting. Model M2 describes the motility of uninfected and HIV-infected lymphocytes within complex tissue environments via a stochastic cellular Potts model (CPM) accounting for biophysical properties of individual cell types to analyse the spread of viruses between cells[19].

The FitMultiCell pipeline was used on model M1 to fit 14 reaction rate parameters. The fit was performed on the JUWELS supercomputer at the Jülich Supercomputing Centre, using up to 84 nodes with 48 cores each, of type 2x Intel Xeon Platinum 8168 CPU, 2.7 GHz. We used $N = 1,000$ particles per generation and $4,032$ workers. The fitting process for this model was run for $n_t = 36$ generations, with a total of $3,803,776$ simulations. The average time for a single forward simulation of this model was about $53.42\,s$, the whole analysis took around $14\,h$ wall-time. Over time, the epsilon acceptance threshold decreased by orders of magnitude and converged gradually (Fig. 4 red). The ABC approximate posterior distribution evolved over the generations, yielding tight posteriors for some of the 14 parameters, while others could not be as precisely inferred (Fig. 5 red, using kernel density estimates).

As an additional application to test the platform, we used model M2, fitting the model's 13 reaction rate parameters. Here, we ran the analysis on the BioQuant Cluster. Each node in the cluster is equipped with IBM xSeries 2 x Quad Intel Xeon E5530 with 20 GB Memory. This demonstrates that the developed platform can be used on different HPC environments. We employed a population size of $N = 256$ using 256 cores for the worker processes. As termination criterion, a minimum discrepancy of $\varepsilon_{\text{min}} = 5$ was used, which

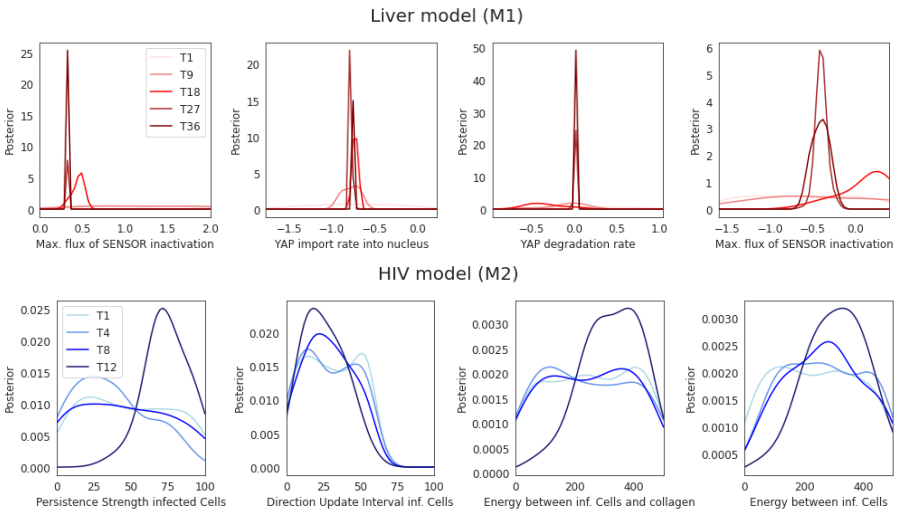


Figure 5. **Parameter fits.** Kernel density estimates in different generations of the iterative fitting process for models M1+2. Shown is only a subset of 4 of the models' 14 resp. 13 parameters each.

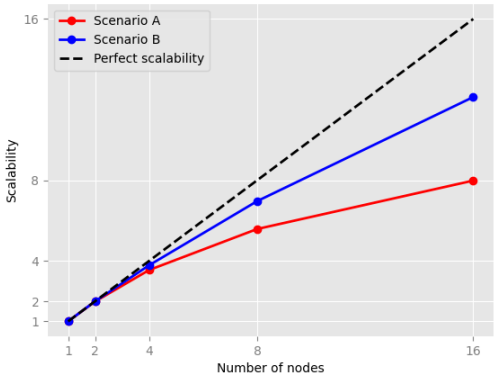


Figure 6. **Scaling behaviour.** Of the FitMultiCell pipeline on JUWELS, for model M3.

was met after $n_t = 12$ generations, resulting in a total of $60,215$ simulations. The simulation of this model is computationally more expensive, taking on average about $1280.25\,s$, the whole analysis required around $83\,h$ wall-time. Similar to before, the epsilon thresholds of different generations, as well as ABC posterior approximations are shown in Fig. 4 and Fig. 5 (blue).

Note that the analyses performed here serve as a first proof-of-concept. They demonstrate that using HPC parallelisation allows us to consider models whose inference would be far too time-consuming otherwise. To assess validity and robustness of obtained posterior approximations, further studies would be necessary.

4.2 Performance Assessment

To assess the speed-up achieved using the massive parallelisation facilitated by the FitMultiCell pipeline, we evaluated parallel scaling on JUWELS standard compute nodes using a model that describes the spread of HCV virus (M3)[20]. We performed the scaling study on two different scenarios of M3, running three sequential populations of size $N = 1,000$ (Scenario A) or $N = 10,000$ (Scenario B) particles. For Scenario A, we observed for 8 nodes (384 cores) a parallel efficiency of 69% and for 16 nodes (768 cores) a parallel efficiency of 49%. For Scenario B, the parallel efficiency was higher, namely 87% and 75% (Fig. 6).

In typical ABC-SMC applications, 15 to 40 sequential populations are run. With each population, usually the acceptance rate decreases, i.e. more simulations have to be performed. This implies an increase of the simulations-to-cores ratio, and in particular the idle time at the end of each population has proportionally less of an impact. This is similar to what can be observed in the scaling study for the larger population size. As we here only ran three sequential populations, we thus expect in an application setting a further improved overall parallel efficiency in later iterations. This indicates that the wall-time of our ABC pipeline scales well, nearly inversely proportionally, with the number of cores.

5 Algorithm Development: Robustness and Informativeness

While the core principle of ABC is simple, its practical performance relies on a number of factors. Two essential problems that we faced are robustness especially to outliers, and assessment of data informativeness.

Outliers, resulting from errors in the data generation process not captured by the model, can severely impact the analysis and lead to biased or uncertain results[22,21]. We showed that popular ABC distance measures can unfortunately be highly sensitive to outliers. This holds especially for a popular distance measure that adjusts in each ABC-SMC generation to the problem structure by weighting different data points or summary statistics based on samples in previous generations, aiming to automatically level the impact of heterogeneous data on the analysis by weights accounting for scale differences[23]. We extended this idea to outlier-corrupted data, by the use of robust norms, and having weights also account for discrepancy of simulations and data in order to identify data points that cannot be explained by the model[21]. This allowed to identify and down-weight outliers, see Schälte *et al.*, 2021, for details[21].

The second problem is that data points are often not homogeneously informative of parameters, e.g. when some only depend on background noise, or different data points are informative of different parts of the model. A popular line of approaches uses inverse regression models, e.g. linear or neural network models, of parameters on data, to extract low-dimensional informative summary statistics[24–26]. We showed first that such approaches benefit on heterogeneous data scales from the combination with scale-normalised distance measures as above[23]. Second, we developed an approach that uses inverse regression models not to define low-dimensional summary statistics, which can possibly lead to a loss of information, but to define additional sensitivity weights quantifying informativeness

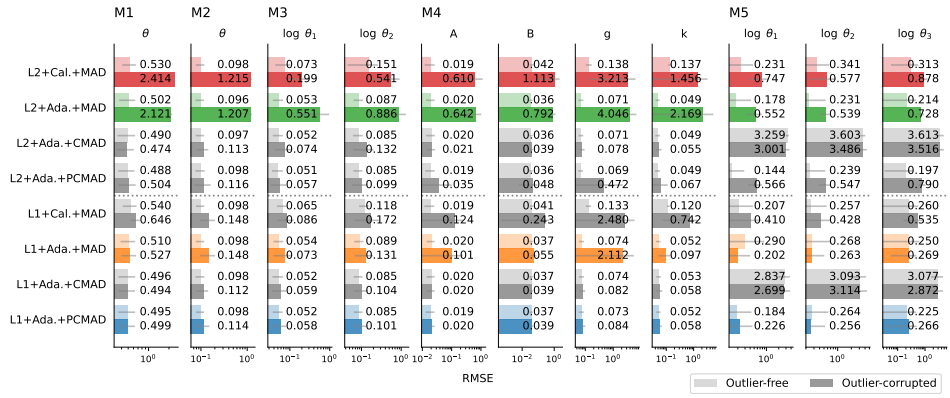


Figure 7. **Distance robustness study.** Mean root mean square errors (RMSE) for the parameters of 5 test models (columns) obtained for 8 distance functions (rows), using L2 or L1 distances, calibrated only in the first ("Cal.") or every ("Ada.") generation, and using MAD, CMAD, and PCMAD for distance weight calculation. Each RMSE is averaged over 20 data sets, grey lines indicate standard deviations. For each distance, the upper, lighter bar is based on outlier-free data, while the lower, darker bar is based on outlier-corrupted data. Distances of interest are coloured, alternative distance combinations are shown in grey for reference. This figure is taken from Schälte *et al.*, 2021, see there for details[21].

of each data point on parameters. Third, we addressed problems of established regression-based approaches with non-identifiability, and presented a solution using regression target augmentation. We observed substantially superior performance of the novel approaches, see Schälte and Hasenauer, 2022, for details[27].

In both studies, we performed extensive method evaluations and comparisons on sets of test and application problems, using up to 20 data replicates (see exemplarily Fig. 7 on robust distances). This was to obtain average and deviation statistics, as results can vary substantially for specific runs and data types. The studies were performed on JUWELS, using up to 384 cores per analysis, and running various analyses in parallel. Given an overall wall-time for all analyses of still several days, a restriction to non-HPC setups would have substantially slowed down method development and robust evaluation.

See Schälte *et al.*, 2021, and Schälte and Hasenauer, 2022, for details on the methods and results[21,27].

6 Conclusion

In this article, we have shown that the use of massive parallelisation enables systematic likelihood-free parameter estimation for biological multi-scale models, and aids the development of novel algorithms. While we have obtained some promising results, the deployment of analysis pipelines on HPC infrastructure could still be improved. For example, existing parallelisation approaches do not make use of all available resources at all times yet, as also dynamic scheduling leaves cores idle at the end of generations, which could e.g. be used to start the next generation already. Further, common HPC infrastructures limit the wall-time of jobs, which can be problematic for longer-running analyses. While pyABC allows to continue previous runs, the amount of manual intervention and information loss could be reduced by automatic checkpointing and job continuation.

We anticipate that approaches as presented here will facilitate the development and analysis of holistic models and gain novel and deeper insights into the organisation of cellular systems.

Acknowledgements

This work was supported by the German Federal Ministry of Education and Research (BMBF) (FitMultiCell/031L0159 and EMUNE/031L0293) and the German Research Foundation (DFG) under Germany's Excellence Strategy (EXC 2047 390873048 and EXC 2151 390685813). 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. N. A. Gershenfeld and N. Gershenfeld, *The nature of mathematical modeling*, Cambridge university press, 1999.

2. H. Kitano, *Systems Biology: A Brief Overview*, Science, **295**, no. 5560, 1662–1664, Mar. 2002.
3. A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, 2005.
4. N. Jagiella, D. Rickert, F. J. Theis, and J. Hasenauer, *Parallelization and High-Performance Computing Enables Automated Statistical Inference of Multi-scale Models*, Cell Systems, **4**, no. 2, 194–206, 02 2017.
5. S. Tavaré, D. J. Balding, R. C. Griffiths, and P. Donnelly, *Inferring coalescence times from DNA sequence data*, Genetics, **145**, no. 2, 505–518, 1997.
6. S. A. Sisson, Y. Fan, and M. Beaumont, *Handbook of approximate Bayesian computation*, Chapman and Hall/CRC, 2018.
7. J. Hasenauer, N. Jagiella, S. Hross, and F. J. Theis, *Data-Driven Modelling of Biological Multi-Scale Processes*, Journal of Coupled Systems and Multiscale Dynamics, **3**, no. 2, 101–121, Sept. 2015.
8. J. Starruß, W. de Back, L. Brusch, and A. Deutsch, *Morpheus: A user-friendly modeling environment for multiscale and multicellular systems biology*, Bioinformatics, **30**, no. 9, 1331–1332, Jan. 2014.
9. Y. Schälte, E. Klinger, E. Alamoudi, and J. Hasenauer, *pyABC: Efficient and robust easy-to-use approximate Bayesian computation*, arXiv preprint arXiv:2203.13043, 2022.
10. J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman, *Population growth of human Y chromosomes: a study of Y chromosome microsatellites.*, Molecular biology and evolution, **16**, no. 12, 1791–1798, 1999.
11. S. Barber, J. Voss, and M. Webster, *The rate of convergence for approximate Bayesian computation*, Electronic Journal of Statistics, **9**, no. 1, 80–105, 2015.
12. P. Del Moral, A. Doucet, and A. Jasra, *Sequential Monte Carlo samplers*, J. R. Stat. Soc. B, **68**, no. 3, 411–436, 2006.
13. S. A. Sisson, Y. Fan, and M. M. Tanaka, *Sequential Monte Carlo without likelihoods*, Proc. Natl. Acad. Sci., **104**, no. 6, 1760–1765, Jan. 2007.
14. E. Klinger and J. Hasenauer, *A scheme for adaptive selection of population sizes in Approximate Bayesian Computation - Sequential Monte Carlo*, in: Computational Methods in Systems Biology. CMSB 2017, J. Feret and H. Koeppl, (Eds.), vol. 10545 of *Lecture Notes in Computer Science*, Springer, Cham. 2017.
15. Y. Schälte and J. Hasenauer, *Efficient exact inference for dynamical systems with noisy measurements using sequential approximate Bayesian computation*, Bioinformatics, **36**, no. Supplement 1, i551–i559, 7 2020.
16. E. Klinger, D. Rickert, and J. Hasenauer, *pyABC: distributed, likelihood-free inference*, Bioinformatics, **34**, no. 20, 3591–3593, 10 2018.
17. L. Schmiester, Y. Schälte, F. T. Bergmann, T. Camba, E. Dudkin, J. Egert, F. Fröhlich, L. Fuhrmann, A. L. Hauber, S. Kemmer, P. Lakrisenko, C. Loos, S. Merkt, W. Müller, D. Pathirana, E. Raimúndez, L. Refisch, M. Rosenblatt, P. L. Stapor, P. Städter, D. Wang, F.-G. Wieland, J. R. Banga, J. Timmer, A. F. Villaverde, S. Sahle, C. Kreutz, J. Hasenauer, and D. Weindl, *PEtab – Interoperable specification of parameter estimation problems in systems biology*, PLOS Computational Biology, **17**, no. 1, 1–10, January 2021.

18. K. Meyer, H. Morales-Navarrete, S. Seifert, M. Wilsch-Braeuninger, U. Dahmen, E. M. Tanaka, L. Brusch, Y. Kalaidzidis, and M. Zerial, *Bile canaliculi remodeling activates YAP via the actin cytoskeleton during liver regeneration*, Molecular systems biology, **16**, no. 2, e8985, 2020.
19. A. Imle, P. Kumberger, N. D. Schnellbächer, J. Fehr, P. Carrillo-Bustamante, J. Ales, P. Schmidt, C. Ritter, W. J. Godinez, B. Müller, et al., *Experimental and computational analyses reveal that environmental restrictions shape HIV-1 spread in 3D cultures*, Nature Communications, **10**, no. 1, 2144, 2019.
20. P. Kumberger, K. Durso-Cain, S. Uprichard, H. Dahari, and F. Graw, *Accounting for Space – Quantification of Cell-To-Cell Transmission Kinetics Using Virus Dynamics Models*, Viruses, **10**, no. 4, 200, Apr 2018.
21. Y. Schälte, E. Alamoudi, and J. Hasenauer, *Robust adaptive distance functions for approximate Bayesian inference on outlier-corrupted data*, bioRxiv, 2021.
22. C. Maier, C. Loos, and J. Hasenauer, *Robust parameter estimation for dynamical systems from outlier-corrupted data*, Bioinformatics, **33**, no. 5, 718–725, Mar. 2017.
23. D. Prangle, *Adapting the ABC Distance Function*, Bayesian Analysis, **12**, no. 1, 289–309, 2017.
24. P. Fearnhead and D. Prangle, *Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation*, J. R. Stat. Soc. B, **74**, no. 3, 419–474, 2012.
25. B. Jiang, T.-Y. Wu, C. Zheng, and W. H. Wong, *Learning summary statistic for approximate Bayesian computation via deep neural network*, Statistica Sinica, 1595–1618, 2017.
26. A. Borowska, D. Giurghita, and D. Husmeier, *Gaussian process enhanced semi-automatic approximate Bayesian computation: parameter inference in a stochastic differential equation system for chemotaxis*, Journal of Computational Physics, **429**, 109999, 2021.
27. Y. Schälte and J. Hasenauer, *Informative and adaptive distances and summary statistics in sequential approximate Bayesian computation*, bioRxiv, 2022.

Uncertainty Quantification and Resource-Demanding Computer Vision Applications of Deep Learning

Julian Burghoff, Robin Chan, Hanno Gottschalk, Annika Mütze, Tobias Riedlinger, Matthias Rottmann, and Marius Schubert

School of Mathematics & Natural Sciences, IMACM, IZMD, University of Wuppertal, 42119 Wuppertal, Germany
E-mail: {burghoff, rchan, hanno.gottschalk, muetze, riedlinger, rottmann, mschubert}@uni-wuppertal.de

Bringing deep neural networks (DNNs) into safety critical applications such as automated driving, medical imaging and finance, requires a thorough treatment of the model's uncertainties. Training deep neural networks is already resource demanding and so is also their uncertainty quantification. In this overview article, we survey methods that we developed to teach DNNs to be uncertain when they encounter new object classes. Additionally, we present training methods to learn from only a few labels with help of uncertainty quantification. Note that this is typically paid with a massive overhead in computation overhead of an order of magnitude and more compared to ordinary network training. Finally, we survey our work on neural architecture search which is also an order of magnitude more resource demanding then ordinary network training.

1 Introduction

Deep learning (DL), convolutional neural networks in particular, in the last 20 years has revolutionised computer vision[1,2]. This has brought new applications into reach that previously have been considered impossible. The application of deep learning technology in safety critical fields – with autonomous driving or assistance of medical diagnosis being prominent examples – however poses new challenges, as deep learning is a statistical machine which errors with a certain probability. Additional information on the reliability of the prediction of AI algorithms should therefore be considered as an integral part of the model's inference. This will either be a piece of valuable information that is provided by the AI-driven assistance to a human decision maker, like in medicine, or to downstream algorithms that, in the case of autonomous driving, choose a more defensive driving policy when facing high uncertainty.

Uncertainty quantification (UQ) for deep neural networks has been under intense study in the past years. The classification of probabilistic uncertainty – epistemic, aleatoric – has been analysed, e.g., using Bayesian neural networks[3] or Monte Carlo dropout[4] as a proxy for Bayesian epistemic uncertainty. For an introduction to UQ, we refer to a survey[5]. UQ is also needed in open worlds problems: As the semantics of AI based classifiers only contains from few to several dozen categories, unknown objects outside this closed semantic world are forcibly classified as an instance of one of the known categories. Control and traceability in such scenarios therefore requires an adequate expression of uncertainty by the network[6–8] or even abstention[9,10].

Advanced deep learning models enable the task of object detection and localisation: on a high resolution image, multiple objects are detected[11,12] or the entire image is segmented into semantic categories providing a pixel-wise classification, see e.g. Ref. 13–15. The

training of such models requires heavy computer resources, in particular general purpose graphic processing units (GPGPU)[14]. Combining such advanced prediction models with an adequate UQ therefore is an important task.

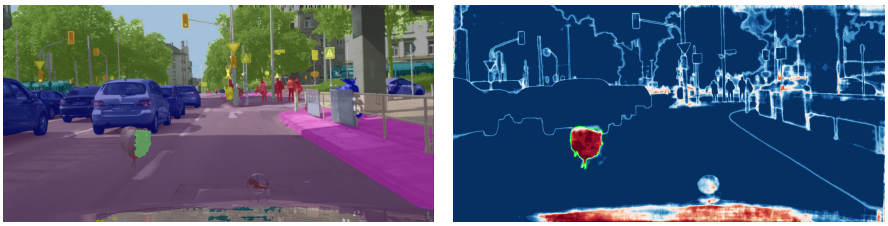
In the past years, UQ for image segmentation and object detection has been initiated in a series of works[16–21] focusing on false positive instances and false negative instances[22,23], respectively, we also refer to a survey[17]. While many works focus on street scene recognition for autonomous driving,[16,24] cover medical image segmentation applications as well.

Further application of UQ is the usage in active[25,26] and semi-supervised training strategies[3], which is eligible to reduce the required ground truth by almost an order of magnitude. Labelling real-world data is time consuming, costly and error-prone[27]. To mitigate these problems several simulation tools and synthetic datasets[28–30] were published in the recent years. Using this data mitigates the label effort but leads to uncertainty of the network as it predicts on a different domain than it was trained on. Reducing via so called domain adaptation methods and analysing the uncertainty is an active research field. It has also been demonstrated that network architectures that were pruned can be more robust[31]. The UQ and the training of deep networks for uncertainty awareness however require additional hardware resources compared with the already resource hungry standard deep learning technology, which requires plenty of GPU and CPU compute resources.

In the remainder of this work, we give an overview of current projects of the authors that use GPU compute resources on the JUWELS[32] supercomputer at the Jülich Supercomputing Centre (JSC). We present insights into results and discuss open questions. More precisely, the remainder of this work is structured as follows: In Sec. 2, we present a method to identify objects from semantically unknown categories. Intuitively, these objects should come with high model prediction uncertainty, which we exploit to detect and localise unknown objects in semantic segmentation. In Sec. 3 we analyse the domain gap as a source of uncertainty of a neural network. In the context of urban street scenes and simulations we present a semi-supervised domain adaptation approach for the semantic segmentation task. In Sec. 4, we describe our active learning endeavours in the context of deep object detection. We compare a method based on previously developed, highly informative uncertainty quantification methods against common baseline approaches. Sec. 5 is about the automated optimisation of neural networks with respect to their architecture. We show how we have extended Google's MorphNet approach to build networks from scratch. This is demonstrated in numerical experiments on the CIFAR10 dataset.

2 Detecting Unknown Objects in Semantic Segmentation

Semantic segmentation is the computer vision task of assigning each pixel in a given image to an object category. For this complex task, one typically employs very deep convolutional neural networks (DNNs). Their training typically requires plenty of GPU resources. Depending on the architecture of the DNN, such a training can take up to one week on a single GPU. DNNs are trained to operate on a closed and pre-defined set of object classes. Therefore, those models are ill-equipped to handle objects from classes that are semantically unknown. These outlined objects are also called *out-of-distribution* (OoD), as they are associated with an extremely low probability of occurrence. We introduce a method to detect such unknown objects in semantic segmentation[33], which is commonly known as the task of *anomaly segmentation*. Our proposed training approach enforces high prediction



(a) Semantic segmentation mask (b) Pixel-wise softmax entropy scores

Figure 1. Illustration of the successful detection of an unknown object by means of the softmax entropy in semantic segmentation.

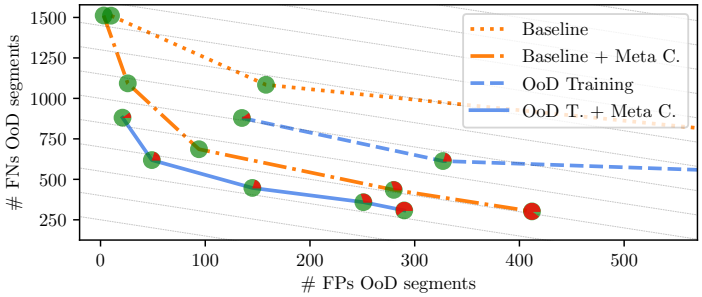


Figure 2. Detection errors of OoD objects. In this plot, the number of false positive and false negative errors are plotted against each other for different combinations of entropy maximisation (denoted as OoD Training) and meta classification. For each combination the detection threshold is additionally varied, i.e., the entropy value at which an object prediction (or segment in the case of semantic segmentation) is identified as OoD. The corresponding pie-chart markers indicate the loss in original semantic segmentation performance, being entirely red if this loss is greater than 1%.

uncertainty on objects from unknown categories, while at the same time retaining segmentation performance on original categories. Our approach outperforms several anomaly segmentation baselines, such as Mahalanobis distance[34] or Monte Carlo dropout[35]. We achieve significant anomaly segmentation performance gains over those established baselines of up to 26 percent points in the relevant metric area under precision recall curve, yielding one of the best performing methods on the public benchmarks Fishyscapes[36] and SegmentMeIfYouCan[37] at the time of writing. For an illustration of our method we refer to Fig. 1. For finding the appropriate training hyperparameters of our anomaly segmentation model but also other parameters such as the size of the OoD dataset and OoD weighting in the multi-criteria loss function, extensive computational experiments were necessary. To this end, we used single-GPU training on single (accelerated) nodes performing four runs simultaneously. In each run a semantic segmentation DNN with more than 60 million weight parameters was trained for multiple epochs.

Furthermore, we introduce meta classification for anomaly segmentation. A meta classifier is a model that flags incorrect predictions. In this work, we employ a logistic regression as meta classifier to detect false indications of unknown objects based on hand-crafted

metrics derived from the probabilistic output of the semantic segmentation DNN. In our experiments, the additional usage of meta classification consistently reduces the number of false OoD indications, i.e. false positives[38]. Compared to the baseline, which is no entropy maximisation and no meta classification, we reduce the number of false positives by 66% while at the same time reducing the number false negatives (overlooked OoD objects) by 50%. Importantly, this significant improvement in the error rates is associated with only a marginal loss in original semantic segmentation performance of at most 1%, see also Fig. 2. Moreover, as logistic regression models are simple linear models, it allows us to track features contributing to the detection of unknown objects. Therefore, besides observing an additional gain in anomaly segmentation performance the combination of entropy maximisation and meta classification particularly contributes to safer DNNs for real-world applications.

3 Downstream Task Informed Domain Adaptation GAN for Semantic Segmentation

As introduced in the previous section, deep neural networks (DNNs) are used for semantic segmentation tasks. For generalising well on unseen scenes, these models need plenty of labelled images but a manual label process is time and cost consuming and usually error-prone. With the help of simulations we have the possibility of generating arbitrarily many labelled data samples to train an expert network on the simulated domain via supervised learning. However, DNNs trained on one domain can perform arbitrarily poor when switching to another domain, i.e., when changing the data generating distribution[39]. With our approach we try to mitigate the so-called domain gap via style transfer and guidance towards the down stream task on the domain where labelled data is rare. Our work is based on the publication of[40] and is extended by the semi-supervised training routine of the GAN and an additional loss term to achieve a task aware generator. Our method can be split into three steps:

- *Semantic Segmentation:* Training a semantic segmentation network on the synthetic domain (expert network);
- *Pre-training:* Transfer the real images into the synthetic domain via Image-to-Image translation;
- *Fine-tuning:* Guide the generator to the downstream task with the help of a couple of ground truth masks and the semantic segmentation loss.

Training neural networks and the style transfer in particular on images with a high resolution is computationally expensive and takes about 45 minutes per epoch on a single NVIDIA V100 GPU each. As the network training and the GAN pre-training are independent, we can train both in parallel on different GPUs. For our experiments, we focus on the task of semantic segmentation for real world street scenes where we have few labels and use a simulation to train our expert segmentation network. Therefore we focus on "real to sim(ulation)" as domain shift. For now we consider random sampling as query strategy for the fine tune process. We train the networks completely from scratch to prevent a bias towards the real world, to evaluate the domain gap accurately. As a consequence

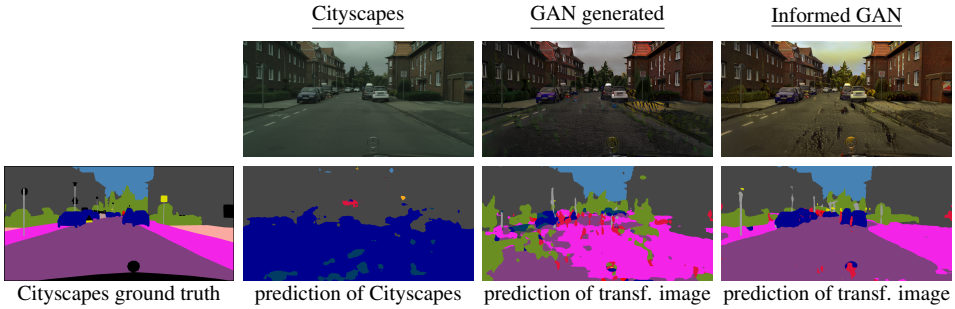
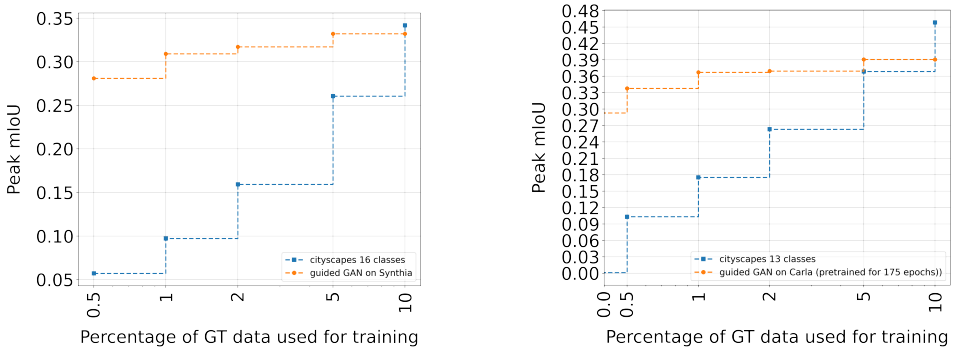


Figure 3. Comparison of prediction results of original Cityscapes image (left), simple style transfer (mid) and our approach (right).



(a) Network trained on Synthia

(b) Network trained on CARLA

Figure 4. Comparison of the performance given different amount of ground truth data. The blue graph represents the supervised Cityscapes training and the orange graph our approach. Please note the different scaling of the y-axis in the two plots.

we accept a reduction of the total accuracy. The Cityscapes dataset[41] serves as real domain in our experiments and we either use Synthia_Rand_Cityscapes[30] or a self generated CARLA dataset[28] as synthetic domain. Our experiments show that even though our overall performance lies between 28–39% mean Intersection over Union (mIoU) we can mitigate the domain gap. We vary the amount of labelled data in our experiments between 0.5% (5 images) to 10% (297 images) of the Cityscapes trainingset. We achieve a gap reduction by 18–30 percent points mIoU depending on the amount of labels used during fine-tuning.

An example image from the validation set of Cityscapes is shown in Fig. 3. For a qualitative comparison we show its transformation with plain CycleGAN[40] and with an informed GAN which was trained with our approach and 148 randomly sampled images with labels during the fine tuning step. Furthermore the corresponding prediction of the network trained on Synthia is illustrated in the second row. One can see, that even though we still find some artefacts in the transformed images, an improvement in the overall scene understanding. In addition we compare our approach to a plain supervised Cityscapes training with the same amount of labels as used for the fine tuned. In Fig. 4 the comparison of the

strongest performance achieved for a given amount of ground truth (GT) is shown. On the x-axis the relative amount of the Cityscapes training data set (2975 images) is plotted. The plot shows that with our approach we can improve segmentation results when only a small fraction of labelled data is available. But it is worth noting, that if the amount of labelled data increases, a supervised training approach should be preferred as the information of the labels can be learned directly.

4 Active Learning Strategies for Deep Object Detection

It is widely known that, as a rule of thumb, for the training of any kind of artificial neural network, an increase of training data leads to increased performance and better generalisation. Obtaining labelled training data, however, is often very costly since in most cases, labels can only be given by a human "oracle" / expert. In applications where labelled training data is scarce (such as in deep object detection), the application of active learning is promising for an effective cost reduction with respect to training labels. The generic active learning algorithm works as follows: after training a network on some (potentially rather small) initial set of labelled data $\mathcal{L}$ to convergence, the network's uncertainty with respect to samples from some unlabelled set of data $\mathcal{U}$ is inferred. It has been shown in[42,43] that in many cases, labels corresponding to uncertain samples can increase the network's performance significantly when adding those samples to the training data.

Our proposed active learning strategy builds upon previous work in uncertainty quantification for deep object detection[20,21]. The MetaDetect[20] and gradient uncertainty[21] approaches have shown state-of-the-art uncertainty quantification in terms of false positive detection of individual predicted instances as well, as localisation uncertainty estimation. Moreover, the proposed methods do not affect the training process of the neural network, therefore, fitting as a post-processing module fitted on a validation dataset $\mathcal{V}$ on top of any object detector making the implementation into an active learning cycle flexible. Fig. 5 shows the experimental setup of this strategy starting with the initially labelled dataset $\mathcal{L}$ on the very left of the schematics.

Applying active learning to tasks that are close to real-world applications, such as object detection, requires large amounts of computational resources. In each single experiment (e.g., 10-15 active learning cycles/query steps), a deep neural network for object detection needs to be trained until convergence several times which comes with a large computational overhead, even when utilising GPUs. Apart from comparing to various

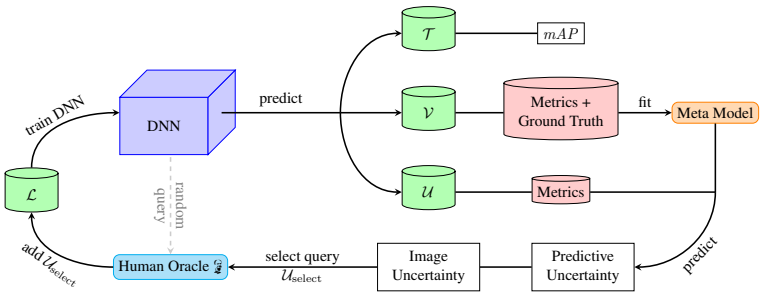


Figure 5. Schematic setup of active leaning cycle for MetaDetect-based query strategies.

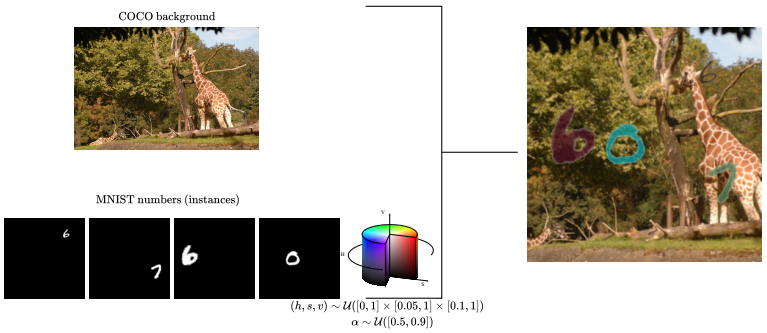


Figure 6. Semi-synthetic dataset generation for MNIST digit and EMNIST letter detection on a complex background.

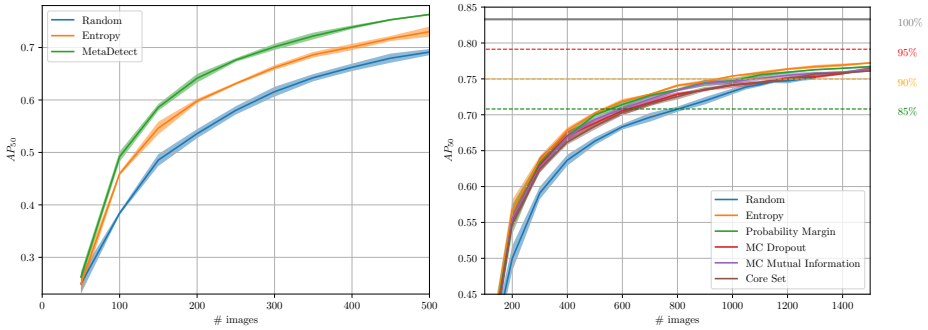


Figure 7. *Left*: MetaDetect-based query strategy on our MNIST-based toy model in comparison with the random and the entropy baseline (y-axis shows detection performance in terms of AP_{50}). *Right*: Comparison of common literature baselines on our EMNIST-based toy model.

baselines, in order to obtain reliable studies, each experiment needs to be repeated under different random seeds for upward of four times. Moreover, contemporary active learning literature for object detection tends to use setups widely differing in the utilised hyperparameters. When developing a query strategy and comparing to different baseline approaches, a rapid prototyping frequency is essential for the reduction of overall compute expense. We, therefore, base our investigation on an efficient, comparable and trackable toy model environment, in addition to standard setups using standard object detection benchmarks and networks. Our toy model framework consists of two semi-synthetic data sets for supervised learning of detecting MNIST digits or, respectively, EMNIST letters on an MS COCO background image, see Fig. 6. Moreover, we use a slim network architecture in RetinaNet with a short ResNet18 backbone for fast training and inference, combined with the data sets cutting the time used for an active learning step (one cycle in Fig. 5) to up to 10% of the time consumed in the standard setting. This leads to an effective time consumption of up to 1 day per experiment as opposed to 5 to 10 days (when using a single GPU). However, even this experimental setup still requires a considerable amount of GPU resources.

Central to our prototyping investigation is the comparison to a wide range of baselines beyond the random image selection and entropy-based query strategies which is computationally expensive despite the use of an efficient framework. Our considerations lie on baselines which are not heavily architecture-dependent like the probability margin of the classifier, Monte-Carlo dropout-based selection strategies or the core set approach which was initially introduced for the classification setting. See the right-hand side of Fig. 7 for a comparison of the implemented baselines. We see that all baselines show significant improvement in comparison with the random strategy which represents uninformed image selection.

5 Automated Design of Neural Network Architectures

Since a neural network usually delivers high performance only in a small, predefined area of application (e.g. image recognition for a given set of classes), research is being conducted into how suitable architectures of neural networks can be found automatically. To this end, we developed an architecture search method for convolutional neural networks, which is based on Google's MorphNet[44]. Our extension for the first time allows to adaptively add and remove layers as well as whole residual blocks[13] alongside with MorphNet, enabling users to automatically obtain neural networks architectures with close to state-of-the-art performance without any manual architecture development. This is particularly useful in exotic applications and/or new datasets where there are no readily trained networks available.

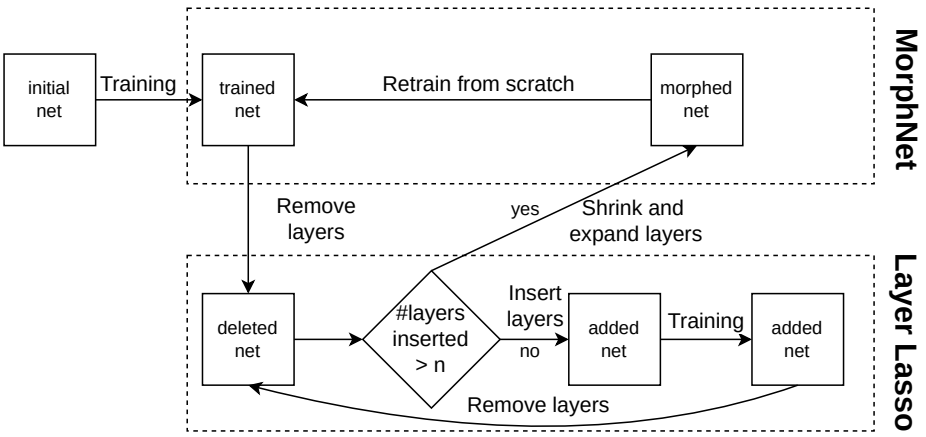


Figure 8. Training pipeline for generating a well-fitted architecture for a specific dataset. Starting with an initial architecture (which can consist of only one convolutional layer or can be a state-of-the-art architecture) we train this net from scratch and then apply the Layer Lasso procedure. This adds and deletes residual blocks of convolutional layers according to the importance of the layer in terms of the sum of the absolute values of the weights. After a chosen number of Layer Lasso iterations n, the MorphNet procedure starts where the number of channels in each layer is shrunk (dependant on the sum of the weights of the filter) and expanded by a linear factor again to take use of the full computation capacity. At the end of the MorphNet routine the net is trained from scratch and the Layer Lasso method starts again with 0 layers inserted.

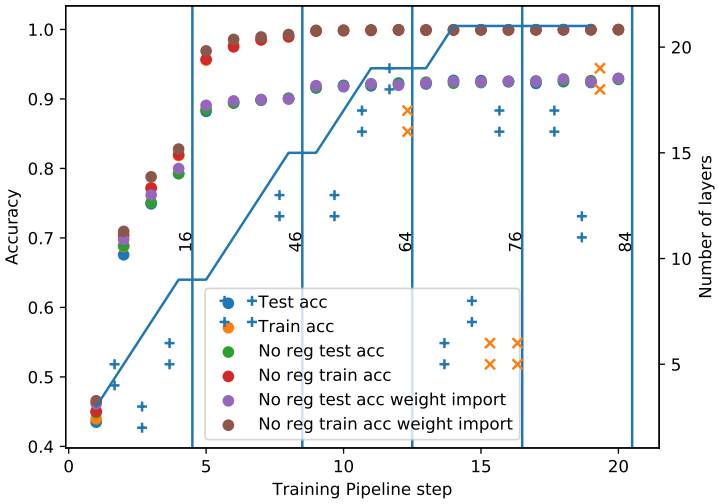


Figure 9. Example run on the CIFAR10 dataset. Each pipeline step stands for a different tested architecture whereas points show the reached accuracy of this architecture in three different ways: Blue/orange points show the accuracy achieved by the network when all regularisation terms (L2 + MorphNet + Layer Lasso) apply. Green and red points show the accuracy after training the same architecture from scratch but only with L2 regularisation and violet/brown points are also trained only with L2 regularisation, but we import the weights we achieved from the run with all regularisation terms. The vertical blue lines indicate that a morphnet steps happens and the number next to the lines explain how many channels the initial convolutional layer, i.e. also all end layers of the different residual blocks, have. A blue plus/orange x show at which position a layer was added/deleted, respectively.

Fig. 8 depicts the architecture search pipeline which is used to automatically search for new architectures of neural networks fitted to a given dataset. Besides the ability of adding and deleting blocks of layers in a similar way Google's MorphNet does, we also implemented a momentum on the layer insertion where we skip the deletion of layers, if the accuracy we achieved with the new architecture is significantly higher than the accuracy of the previous architecture. Another feature is a learning rate on the MorphNet suggestions: We do not use these shrinkage suggestions exactly to get a better cooperation of the two basic structures MorphNet and LayerLasso.

Neural architecture search requires an order of magnitude more compute resources than an ordinary training. In addition, our architecture search maintains a number of hyper parameters such as different regularisation strengths and adjusting the momentum on layer insertion. Therefore, the tuning of these hyper parameters consumes high amounts of compute resources.

A preliminary result on the CIFAR10 dataset is given in Fig. 9. We started with a minimal architecture containing only a single convolutional layer and then automatically added ResNet blocks with our architecture search, achieving a test accuracy of up to 92.97% which is fairly close to comparable hand-crafted ResNet architectures.

6 Conclusion & Outlook

In this overview article, we gave insight into our work on some of the most resource demanding applications in the field of deep learning and the role of uncertainty quantification therein.

For future work building upon our proposed anomaly segmentation method, our resulting model could be integrated in an retrieval-based approach to identify relevant OoD objects. Via clustering methods, the frequency of occurrence of specific OoD objects can then be assessed to determine whether new concepts are required to be learned. Ideally, this is pursued in an entirely unsupervised fashion, in this way enabling DNNs to operate in the open world[45,46]. In the context of the domain adaptation we will investigate if an active learning strategy based on the uncertainty of the network can help sampling the most informative samples in a pool of unlabelled data which are then labelled and used for fine-tuning the generator. With a focus on active learning for object detection, current investigations show that query strategies which perform better than the random baseline in terms of queried images do not, however, consistently outperform the random baseline in terms of queried bounding boxes. Since annotation costs are proportional to the amount of boxes queried, we plan on refining our method and investigating the problem of reducing the amount of labelled instances queried while maintaining performance via an informed selection strategy. With regard to our automated design of neural networks, we will evaluate our methods on real-world data and will focus on optimising state-of-the-art networks for a number of benchmark datasets to further develop the method.

Acknowledgements

Supported by:

Federal Ministry for Economic Affairs and Climate Action

on the basis of a decision by the German Bundestag

EUROPEAN UNION
Investing in our Future
European Regional Development Fund

The research leading to these results is in part funded by the German Federal Ministry for Economic Affairs and Climate Action within the project "Methoden und Maßnahmen zur Absicherung von KI basierten Wahrnehmungsfunktionen für das automatisierte Fahren (KI-Absicherung)", grant no. 19A19005R and "KI Delta Learning - Scalable AI for Automated Driving", grant no. 19A19013Q. Furthermore, the authors gratefully acknowledge financial support by the state Ministry of Economy, Innovation and Energy of Northrhine Westphalia (MWIDE) and the European Fund for Regional Development via the FIS.NRW project BIT-KI, grant no. EFRE-0400216. The authors would like to thank the consortiums for the successful cooperation. 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 John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS[32] at Jülich Supercomputing Centre (JSC).

References

1. Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. E. Hubbard, and L. D. Jackel, *Handwritten digit recognition with a back-propagation network*, in: Advances in neural information processing systems, 396–404, 1990.

2. I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*, MIT press, 2016.
3. K. Doya, S. Ishii, A. Pouget, and R. P. N. Rao, *Bayesian brain: Probabilistic approaches to neural coding*, MIT press, 2007.
4. Y. Gal and Z. Ghahramani, *A theoretically grounded application of dropout in recurrent neural networks*, in: Advances in neural information processing systems, 1019–1027, 2016.
5. E. Hüllermeier and W. Waegeman, *Aleatoric and epistemic uncertainty in machine learning: A tutorial introduction*, 2019, preprint arXiv:1910.09457.
6. A. Meinke and M. Hein, *Towards neural networks that provably know when they don't know*, 2019, preprint arXiv:1909.12180.
7. A.-A. Papadopoulos, M. R. Rajati, N. Shaikh, and J. Wang, *Outlier exposure with confidence control for out-of-distribution detection*, 2019, preprint arXiv:1906.03509.
8. D. Hendrycks and K. Gimpel, *A baseline for detecting misclassified and out-of-distribution examples in neural networks*, 2016, preprint arXiv:1610.02136.
9. V.-L. Nguyen and E. Hüllermeier, *Reliable Multi-label Classification: Prediction with Partial Abstention*, 2019, preprint arXiv:1904.09235.
10. C. Cortes, G. DeSalvo, and M. Mohri, *Boosting with abstention*, in: Advances in Neural Information Processing Systems, 1660–1668, 2016.
11. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *You only look once: Unified, real-time object detection*, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 779–788, 2016.
12. J. Redmon and A. Farhadi, *Yolov3: An incremental improvement*, 2018, preprint arXiv:1804.02767.
13. K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778, 2016.
14. F. Chollet, *Xception: Deep learning with depthwise separable convolutions*, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 1251–1258, 2017.
15. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, *Mobilenetv2: Inverted residuals and linear bottlenecks*, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 4510–4520, 2018.
16. M. Rottmann, P. Colling, T.-P. Hack, R. Chan, F. Hüger, P. Schlicht, and H. Gottschalk, *Prediction Error Meta Classification in Semantic Segmentation: Detection via Aggregated Dispersion Measures of Softmax Probabilities*, in: 2020 International Joint Conference on Neural Networks, IJCNN 2020, Glasgow, United Kingdom, July 19-24, 2020, 1–9, IEEE, 2020.
17. M. Rottmann, K. Maag, R. Chan, H. Fabian, P. Schlicht, and H. Gottschalk, *Detection of False Positive and False Negative Samples in Semantic Segmentation*, CoRR, 2019, arXiv:1912.03673.
18. M. Rottmann and M. Schubert, *Uncertainty Measures and Prediction Quality Rating for the Semantic Segmentation of Nested Multi Resolution Street Scene Images*, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2019, Long Beach, CA, USA, June 16-20, 2019, 1361–1369, Computer Vision Foundation / IEEE, 2019.

19. K. Maag, M. Rottmann, and H. Gottschalk, *Time-Dynamic Estimates of the Reliability of Deep Semantic Segmentation Networks*, in: IEEE International Conference on Tools with Artificial Intelligence (ICTAI), 11 2020.
20. M. Schubert, K. Kahl, and M. Rottmann, "Metadetect: Uncertainty quantification and prediction quality estimates for object detection", 2020, accepted for publication in the proceedings of the International Joint Conference on Neural Networks (IJCNN) 2021.
21. T. Riedlinger, M. Rottmann, M. Schubert, and H. Gottschalk, *Gradient-Based Quantification of Epistemic Uncertainty for Deep Object Detectors*, 2021, preprint arXiv:2107.04517.
22. R. Chan, M. Rottmann, F. Hüger, P. Schlicht, and H. Gottschalk, *Application of Maximum Likelihood Decision Rules for Handling Class Imbalance in Semantic Segmentation*, in: The 30th European Safety and Reliability Conference (ESREL), 2020.
23. R. Chan, M. Rottmann, R. Dardashti, F. Hüger, P. Schlicht, and H. Gottschalk, *The Ethical Dilemma when (not) Setting up Cost-based Decision Rules in Semantic Segmentation*, CoRR, 2019.
24. A. Mehrtash, W. M. Wells III, C. M. Tempany, P. Abolmaesumi, and T. Kapur, *Confidence Calibration and Predictive Uncertainty Estimation for Deep Medical Image Segmentation*, 2019, preprint arXiv:1911.13273.
25. R. Mackowiak, P. Lenz, O. Ghori, F. Diego, O. Lange, and C. Rother, *Cereals-cost-effective region-based active learning for semantic segmentation*, 2018, preprint arXiv:1810.09726.
26. P. Colling, L. Roese-Koerner, H. Gottschalk, and M. Rottmann, *MetaBox+: A New Region based Active Learning Method for Semantic Segmentation using Priority Maps*, in: Proceedings of the 10th International Conference on Pattern Recognition Applications and Methods - Volume 1: ICPRAM,, 51–62, SciTePress. INSTICC, 2021.
27. T.-Y. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, *Microsoft COCO: Common Objects in Context*, CoRR, 2014, arXiv:1405.0312.
28. A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, *CARLA: An Open Urban Driving Simulator*, in: Proceedings of the 1st Annual Conference on Robot Learning, 1–16, 2017.
29. S. R. Richter, V. Vineet, S. Roth, and V. Koltun, *Playing for Data: Ground Truth from Computer Games*, in: European Conference on Computer Vision (ECCV), Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, (Eds.), vol. 9906 of *LNCS*, 102–118, Springer International Publishing. 2016.
30. G. Ros, L. Sellart, J. Materzynska, D. Vázquez, and A. López, *The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes*, 3234–3243, June 2016.
31. Y. Guo, C. Zhang, C. Zhang, and Y. Chen, *Sparse DNNs with Improved Adversarial Robustness*, CoRR, 2018.
32. Jülich Supercomputing Centre, *JUWELS Cluster and Booster: Exascale Pathfinder with Modular Supercomputing Architecture at Juelich Supercomputing Centre*, Journal of large-scale research facilities, **7**, no. A138, 2021.

33. R. Chan, M. Rottmann, and H. Gottschalk, *Entropy Maximization and Meta Classification for Out-of-Distribution Detection in Semantic Segmentation*, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 5128–5137, virtual conference, Oct. 2021.
34. K. Lee, K. Lee, H. Lee, and J. Shin, *A Simple Unified Framework for Detecting Out-of-Distribution Samples and Adversarial Attacks*, in: Advances in Neural Information Processing Systems, vol. 31, 7167–7177, Curran Associates, Inc. 2018.
35. Y. Gal and Z. Ghahramani, *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning*, in: Proceedings of The 33rd International Conference on Machine Learning, vol. 48 of *Proceedings of Machine Learning Research*, 1050–1059, PMLR, New York, New York, USA, 20–22 Jun 2016.
36. H. Blum, P.-E. Sarlin, J. Nieto, et al., *Fishyscapes: A Benchmark for Safe Semantic Segmentation in Autonomous Driving*, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, October 2019.
37. R. Chan, K. Lis, S. Uhlemeyer, H. Blum, S. Honari, R. Siegwart, P. Fua, M. Salzmann, and M. Rottmann, *SegmentMeIfYouCan: A Benchmark for Anomaly Segmentation*, in: Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track, virtual conference, Dec. 2021.
38. R. Chan, M. Rottmann, F. Hüger, P. Schlicht, and H. Gottschalk, *Controlled False Negative Reduction of Minority Classes in Semantic Segmentation*, in: 2020 International Joint Conference on Neural Networks, IJCNN 2020, Glasgow, United Kingdom, July 19-24, 2020, 1–8, IEEE, 2020.
39. G. Csurka, *Domain Adaptation for Visual Applications: A Comprehensive Survey*, Mar. 2017, arXiv:1702.05374 [cs].
40. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks*, in: Computer Vision (ICCV), 2017 IEEE International Conference on, 2017.
41. M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, *The Cityscapes Dataset for Semantic Urban Scene Understanding*, in: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
42. Y. Gal, R. Islam, and Z. Ghahramani, *Deep Bayesian Active Learning with Image Data*, CoRR, 2017, arXiv:1703.02910.
43. B. Settles, *Active Learning Literature Survey*, Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009.
44. A. Gordon, E. Eban, O. Nachum, B. Chen, H. Wu, T.-J. Yang, and E. Choi, *MorphNet: Fast & Simple Resource-Constrained Structure Learning of Deep Networks*, 2017.
45. S. Uhlemeyer, M. Rottmann, and H. Gottschalk, *Towards Unsupervised Open World Semantic Segmentation*, 2022, arXiv:2201.01073.
46. R. Chan, S. Uhlemeyer, M. Rottmann, and H. Gottschalk, *Detecting and Learning the Unknown in Semantic Segmentation*, 2022, preprint arXiv:2202.08700.

Fluid Mechanics

Fluid Mechanics

Christian Stemmer

Lehrstuhl für Aerodynamik und Strömungsmechanik, Technische Universität München,
Boltzmannstraße 15, 85748 Garching, Germany
E-mail: Christian.Stemmer@tum.de

This year's fluid mechanics contributions to the NIC symposium circle around multi-scale and multi-physics simulations. Numerical schemes have been expanding in the past decades to go from high-resolution simulations towards coupled problems incorporating more and more nonlinear effects from other disciplines than the core topics of solving the fundamental fluid mechanics equations as accurate as possible. Interface problems as the interaction between solids and fluids (fluid-structure interaction) or fluids and gaseous media (fluid-bubble interactions) among others have become feasible and the predictive capabilities of these multi-physics, multi-scale problems have addressed topics from material science, added manufacturing or medical topics to name a few.

The first paper in this series is a nice example of a joint effort from the Korea University Guro Medical Center and the Jülich Supercomputing Centre where the GPU capabilities on the JUWELS booster are exploited to investigate the change in acoustic impedance in presence of variable density fluids as encountered with biomedical fluids to control wave diffraction. Acoustic and fluid scales are far from each other and require a smart way to overcome resolution issues. The presented simulations open the possibility for the optimisation of the variable density region to obtain acoustic wave patterns of the required type.

A second paper by researchers from the University of Siegen investigates the oscillatory behaviour and control of turbulent compressible channel flow. The active control concept which requires the input of energy into the fluid systems shall be used to reduce the turbulence intensity and therewith the friction drag at the surface of the channel. The introduction of streamwise modulation of the lateral velocity to change the streaky turbulent structures requires high-resolution simulations executed on the GPUs of the JUWELS Cluster module and the Booster module. Depending on the Mach number in the range from 0.3 to 3, the reduction in the wall-friction coefficient c_f was in the range of 40% for the higher Reynolds number case of $Re_\tau = 400$ which is notable.

The third paper is presented by two groups from the TU Illmenau and the Occidental College in Los Angeles. The Rayleigh-Bénard convection is investigated in its turbulent regime. Large Rayleigh numbers (the ratio of the time scale for diffusive thermal transport to the time scale for convective thermal transport at speed) need to be looked at. Different parameters for the size, the Rayleigh number and the compressibility together with rotation rates of the problem are presented and the potential to investigate the structures emerging with machine-learning algorithms (recurrent neural networks) is investigated. The potential for a reduced-order modelling concept is shown.

A similar task has been taken up by the last article in the fluid mechanics section by a mixed group from the University of Magdeburg and Princeton University where turbulent

channel flows and data-driven decomposition of data containing very-large-scale structures is investigated. The method of Proper Orthogonal Decomposition (POD) which was also used in the previous article as a data-analysis tool shows the very-large-scale structures which contain a substantial amount of the disturbance energy of the turbulent channel flow.

Simulation of Sound Wave Focusing by Variable Density Medium

Seong-Ryong Koh[1], Ilya Zhukov[1], Jong-Hyun Kim[2], and Wolfgang Frings[1]

[1] Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {s.koh, i.zhukov, w.frings}@fz-juelich.de

[2] Korea University Guro Medical Center, Gurodong-ro 148, 08308 Seoul, Korea

High-performance computing with GPU acceleration is performed to determine the acoustic field controlled by a diffraction grating approach. The numerical analysis considers synthesised structures which manipulate the wave propagation and the acoustic intensity in the targeted region. A simple convex lens showed that the wave beam is localised and focused down to a small spot. The acoustic intensity in the area is increased by 15dB compared to the result by a 5-slits configuration. The programming language exploits OpenACC directives to achieve a large amount of performance gain by using GPU on the JUWELS Booster. The computational speed of production run is doubled with GPU acceleration compared to the performance of a pure message-passing interface parallelisation.

1 Introduction

Flow-induced sound occurs in many engineering applications, yet it is one of the important energy and environment problems. Powerful high-frequency waves beyond the limit of human audibility are used in many ultrasonic imaging devices for medicine and industry. In the last decade, acoustofluidics, i.e., the fusion of micro- to nano-scale acoustics and fluid mechanics, is one of the fast-developing research topics in science and technology. The acoustic energy in this field is of ultrasonic nature such that the waves are controlled to interact with bio-tissues and micro-particles in fluidic structures. The challenges of acoustofluidics are closely related to physics, engineering, and biomedical applications. The precise wave control is essential to optimise the fluidic-acoustic interaction in real devices. Therefore, new concepts are essential to progress the noise-control techniques and to manufacture the industrial products in a form of new artificial materials.

The design of new acoustic materials requires abilities to manipulate and control sound waves such that the physical behaviours have been impossible with conventional materials. A review of metamaterials researches[1] introduced the designs and properties of acoustic materials with special parameters e.g., negative refractive index, which decreases the wave phase in the medium. The advances of material technologies require the identification of technical challenges related to the future engineering designs such as wave-control devices. Furthermore, the development of new wave-control structures influences expanding researches on new materials that control vibrations, waves and the motion of solid materials with a highly efficient noise reduction.

In the present study the physical problem is configured to develop wave control devices by using variable density media. The diffraction grating is designed to obtain wave focusing and acoustic energy transformation. As improving the mechanical performance the acoustic meta-material could bring a silent driving to vehicles as conventional sound

insulation materials have been suffering from heavy weight and cost constraints. The combination of transformation acoustics theory and highly anisotropic medium enables precise control over the deformation of sound fields, e.g., to hide or cloak objects from incident acoustic energy.

2 Motivation

Towards the further step to develop real-world applications the physical problem tackles the acoustic waves in a channel flow, e.g., frequently visited in the biomedical fluids[2], and employs synthesised structures to control the sound power in the targeted region.

For large-scale engineering applications, the local-reacting properties of fluid-contacting structures, e.g., metal or porous medium, allow their acoustic behaviour to be characterised by a single complex quantity called acoustic impedance. This quantity is greatly affected by high sound pressure levels and grazing flows. That is, the pressure fluctuations accompany the turbulent flow forces exerted on the surface such that the sound generation occurs as the radiation condition is satisfied at a given surface impedance. The numerical analysis of full scale acoustic fields, which frequently concern high Reynolds number flows, increases computational cost due to the large disparity between the flow and the acoustic scale.

In the acoustofluidics the underpinning physics involves acoustics and fluid mechanics in micro-structures. The flow Reynolds number is usually very small, i.e., the viscous stresses easily dominate the flow motion. Furthermore, the applications are combinations of physics, electronics, medicine, and engineering. The numerical simulations are quite challenging to overcome the coupling problems which drastically increase the computational cost and the complexity of scientific interpretation. The computational analysis needs to resolve the temporal and the spatial scales which range from a micro unit to several meters in conventional mechanical designs. From the computational aspects it should be emphasised that the high-resolution numerical schemes and the high-efficiency high-performance computing (HPC) algorithms are essential to develop the noble wave control devices.

Therefore, this project aims for consolidation of the high-resolution computational fluid dynamics (CFD) workflow by exploiting graphics processing units (GPU) on a HPC system. The simulation of physical problems focuses on materials designed to control the propagation of acoustic waves generated in fluids. The GPU programming will include performance assessment and optimisation using OpenACC and CUDA-aware MPI libraries.

3 Sound Wave Control

A sound wave of wavelength λ is diffracted when the wavefront encounters an object or aperture. The sound intensity behind such obstructions is distributed by the shape of the aperture that the wave passes through. The diffraction characteristics are determined by interference between different portions of the wavefront. The resulting intensity distribution is called a diffraction pattern. When sound penetrates multiple apertures or slits with fixed spacing on a thin screen, the emerging wavefronts constructively interfere to produce a diffraction pattern with intensities peaked in certain directions as shown in Fig. 1. These

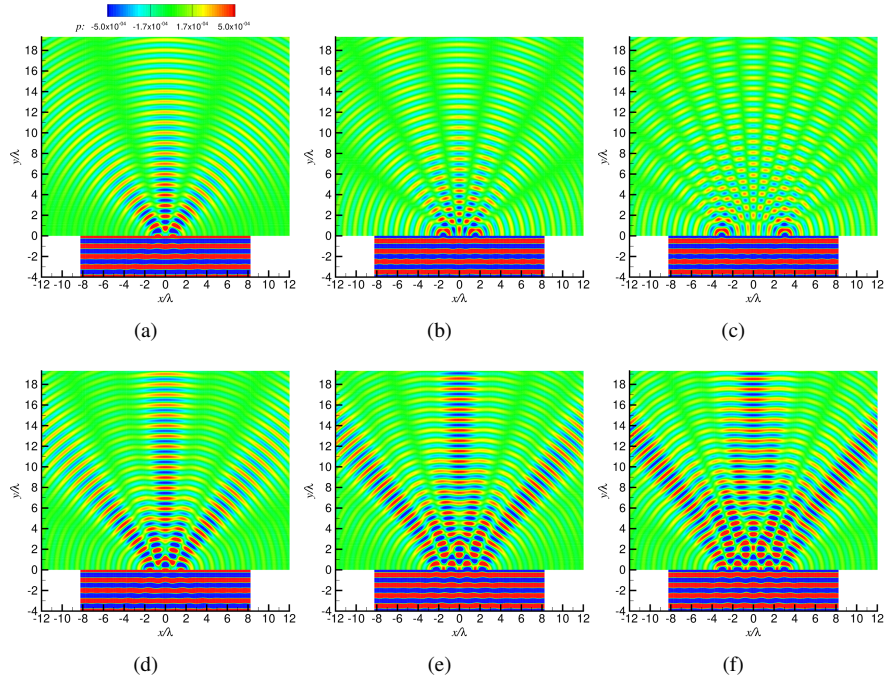


Figure 1. Instantaneous contours of acoustic pressure generated by various slits positioned at the coordinates $y = 0$, (a) 2 slits with $d_G = 1.5\lambda$, (b) 2 slits with $d_G = 3\lambda$, (c) 2 slits with $d_G = 6\lambda$, (d) 3 slits, (e) 4 slits, (f) 5 slits with $d_G = 1.5\lambda$.

directions are strongly dependent on both the slit spacing and wavelength of the incident wave. The direction of wave is determined by the basic grating equation for a wave with an incident angle α, i.e., $m\lambda = d_G(\sin\alpha + \sin\beta_m)$ where d_G is the spacing between slits, m is the order of diffraction, and β is the diffraction angle. Therefore, surfaces with well-defined slit locations are able to direct the acoustics of certain wavelengths into specific directions.

In the 1930s the variable density was reported as a new method of producing acoustic spectra[3]. The method already employed a variable-density sound-film as a diffraction grating for optical waves and created an instantaneous separation of sound into its components. Starting from a pure tone the diffraction angle of which increases with frequency, the diffraction pattern of numerous frequencies was considered in a diffraction grating which could record the range of frequencies over seven octaves. In recent studies an analytical model for a surface wave diffraction[4] was formulated to predict the diffraction pattern generated by the phased diffraction grating. The direction of waves is guided by a unit element of the diffraction grating. The diffraction waves are generated by the entire grating composed of a linear array of unit elements. The width, thickness, and grating length of the unit grating are w, δ, and d. The incident plane wave with an angular frequency $\omega = 2\pi f$ and amplitude A_i is defined by $A_i e^{i\omega t}$. Assuming the transmission coefficient is C, transmission waves at the exit of the unit element can then be expressed as $A_i C e^{i(\omega t - kd + \Delta\phi)}$ with

a phase difference $-kd + \Delta\phi$ by adding superstrate. The additional superstrates manipulate the dispersion relation of the travelling waves on the surface. The guided waves are diffracted at the exit of a unit element. A broadband metagrating design[5] detailed the density of periodical micro-structures which drives the wave diffraction and the Bessel beam implementation. An iterative algorithm was used to control the wave form by recording the phase profiles. The theoretical phase distribution and the simulation results show a good agreement in development of focused beams.

4 Numerical Method

The flow and the acoustic field of variable-density diffraction grating are analysed by solving the conservation equations for compressible fluid flows. The equations describing the sound propagation are the acoustic perturbation equations (APE)[6]. Since a compressible flow problem is tackled in this study the APE-4 system is used, i.e., the formulation of which is derived from the viscous conservation equations. The non-linear terms containing entropy fluctuations occur as additional source terms in the right-hand side of the governing equations. The contribution of viscous dissipation is negligible in the momentum equation of the APE system. To accurately resolve the acoustic wave propagation, a sixth-order finite difference scheme with the summation by parts property[7,8] is used for the spatial discretisation and an alternating 5-6-stage low-dispersion and low-dissipation Runge-Kutta method for the temporal integration[9]. The high-order discretisation with the summation by parts property (SBP)[8] improves the numerical stability of continuous problems via SBP operators. The dispersion relation preserving scheme (DRP)[7] is reformulated with the SBP properties. The sixth order discretisation in space uses the 9-points numerical stencil with a high-order 11-points spatial filter to prevent the spurious wave generation[10].

The flow and the acoustic simulation easily demand numerical meshes of $\mathcal{O}(10^9)$ degree of freedom which is only manageable by HPC systems. Due to the high-order discretisation schemes a typical message-passing-interface (MPI) approach suffers from the serious increase of computational overhead in mesh decomposition. To obtain the large gain of performance acceleration with a minimum number of MPI domains, a GPU programming with OpenACC is adopted to enhance the parallel efficiencies and to accelerate the computational performance in the simulation of the flow and the acoustic field of an acoustofluidic problem.

5 GPGPU Programming

GPUs are not new in the HPC ecosystems and most HPC centres in the world are equipped with GPUs. Usage of GPUs allows researchers and data scientists to improve their computations in orders of magnitude. An initial version of the CFD/CAA solver supports MPI and OpenMP on HPC systems. The behaviour of code subroutines in runtime is evaluated on the JUWELS[11] with GPU accelerators, i.e., JUWELS Booster.

5.1 General Structure

The numerical solver has the typical structure of a scientific application, i.e. initialisation, computational loop and post-processing. A graphical representation of 100 iterations executed on 4 MPI ranks on JUWELS Booster is shown in Fig. 2. The long green area in

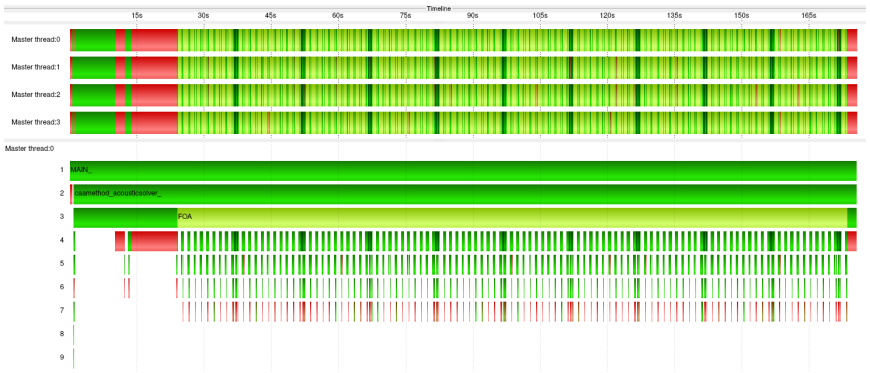


Figure 2. Execution timeline of the CFD/CAA solver, each row represents a single MPI process. A process timeline shows the different levels of function calls in a stacked bar chart at MPI rank 0 by MPI routines (red) and the user functions (green).

the top four bars is an actual compute loop (marked as a focus of analysis (FOA) in the stack diagram at the bottom) and is the most significant part of CFD/CAA solver. FOA consists of a large loop (100 iterations in the depicted testcase), where after each iteration occurs short non-blocking communication within `dom_dom_upgrategl_` routine, and every tenth iteration has a 3d filtering step (short dark green stripes). Short communication based on MPI non-blocking communication allowed to overlap communication with computation in a very efficient way. At the end of the computation loop all results were aggregated with MPI collective operations, i.e, `MPI_Reduce` and `MPI_Allreduce`. As computation is a lion's share of FOA, it is possible and reasonable to parallelise it, e.g., by distributing computations over available CPU cores or GPUs. The CFD/CAA solver can benefit already from OpenMP parallelisation which was implemented in the past.

5.2 Parallelisation Decisions and Strategies

There are several alternatives for GPU parallelisation, i.e., CUDA, OpenACC, HIP[a] and various libraries. Considering all these options we were tempted by OpenACC due to multiple reasons, e.g., OpenACC is relatively easy to implement as it uses compiler directives that are very similar to OpenMP; it gives better portability and readability of the code in comparison to other methods; OpenACC code can be compiled to CPU and GPU versions easily and can co-exist with OpenMP directives. Note that there are several drawbacks of OpenACC, i.e., so far only NVHPC compilers provide full OpenACC support, whereas newer GCC compilers provide only basic support; the low-level tuning for a specific computer architecture can be difficult. To port the CFD/CAA solver to the GPUs the following general rule was used, i.e. avoid costly memory transfers between CPUs and GPUs as much as possible. To accomplish this rule we reused data on GPUs as long as possible and used a special MPI library that allows exchanging memory data from one GPU to another without an intermediate copy to CPU memory (such libraries are called CUDA-aware MPI). CUDA-aware MPI routines can be included into OpenACC applications.

[a]HIP is a C++ Heterogeneous-Compute Interface for Portability:
`https://github.com/ROCm-Developer-Tools/HIP`

5.3 Current Implementation and Results

To integrate OpenACC in the CFD/CAA solver in desired fashion the following approach was used: allocate and initialise all necessary data before FOA and copy results back to the CPU at the end of the computation loop. All communication across GPUs performs via CUDA-aware MPI. A small part of the computation, e.g. 3d filtering step was not ported to GPUs due to the necessity of intensive communication across domains and frequent data transfer from host to device and backwards. This defect will be addressed in a follow-up project.

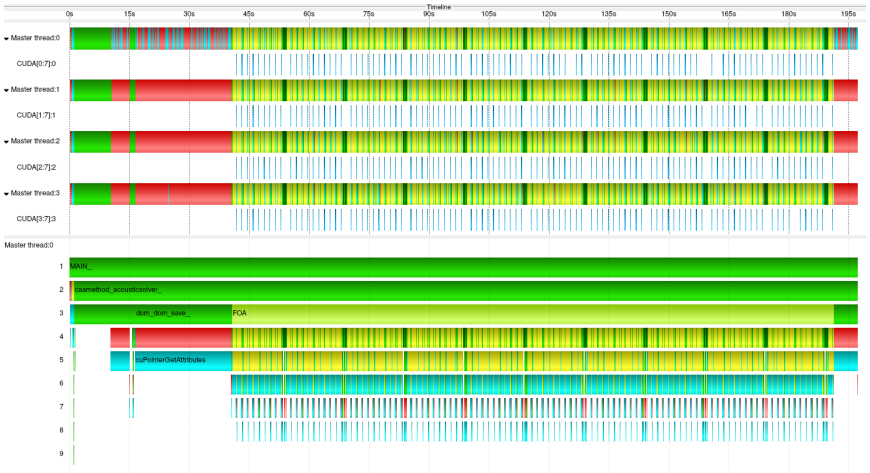


Figure 3. Execution timeline of the CFD/CAA solver (at the top): depicts a sequence of events in time, each row represents a single MPI process or CUDA stream. A process timeline (at the bottom) shows the different levels of function calls in a stacked bar chart for MPI rank 0. MPI routines are depicted in red, user functions are shown in all shades of green, OpenACC routines are yellow, CUDA kernels are dark blue, CUDA API routines are light blue.

A graphical representation of 100 iterations executed on 4 MPI ranks with A100 GPU each on JUWELS Booster is illustrated in Fig. 3. The total runtime with OpenACC implementation in Fig. 3 is longer than in the case in Fig. 2 due to the measurement overhead. The overhead is caused by instrumentation of the code by Score-P measurement infrastructure[b]. The non-instrumented code executions with a relatively heavy computational load on 2 nodes show that the runtime of FOA with a pure MPI approach on JUWELS Booster (96 MPI ranks) is 348 seconds and the one with the OpenACC also on JUWELS Booster (16 MPI ranks with 8 A100 GPUs) is 153 seconds, i.e., 2.3 times faster with GPU acceleration. With production runs where GPUs' load is more realistic, one can get even better speedup. The result shows a good agreement with the measurement in the GPU Hackathon 2021[12]. Note that the number of MPI ranks per GPU is considered an important factor to tune the code performance on the JUWELS Booster.

[b]Score-P measurement infrastructure: `https://www.score-p.org/`

5.4 Future Work

Although we achieved quite good results during the project, there are some aspects we want to improve or investigate, i.e. port as much as possible computations from CPUs to GPUs, investigate possibilities of OpenMP offloading, investigate GPU kernels for possible low-level optimisations.

6 Hardware and Software Configuration

All measurements were performed on the JUWELS Booster[c] which is a part of the high-performance cluster JUWELS. The JUWELS Booster consists of 936 compute nodes equipped with four NVIDIA A100 GPUs for each node. The GPUs are hosted by AMD EPYC Rome CPUs. The compute nodes are connected with HDR-200 InfiniBand in a DragonFly+ topology. The following software was used for the tests on JUWELS Booster: NVHPC/22.3, ParaStationMPI/5.5.0-1, CUDA/11.5 and HDF5/1.12.1. Furthermore, Score-P/7.1 was used for measurements and Vampir/10.0.0[d] for timeline visualisations.

7 Simulation Results

7.1 Simulation Setup

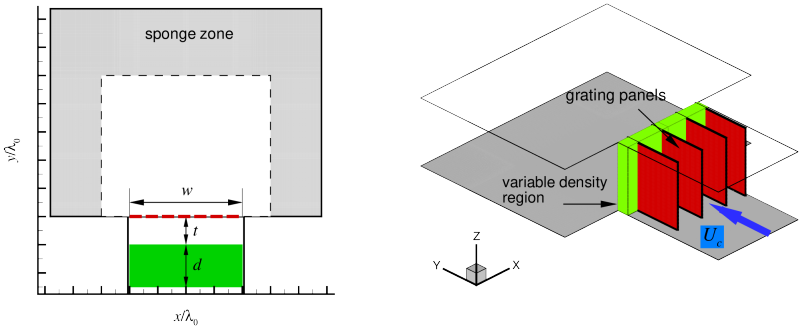


Figure 4. Schematic of the acoustic domain, w denotes the width of diffraction grating, t is the tolerance between the grating exit and the variable-density region, of which the thickness is d, the waves move in the positive y direction and the convection speed of a channel flow is U_c.

The computational study is performed to obtain the flow and the acoustic field of diffraction gratings with variable density distribution. The test cases included slits defined by three different widths d_G. The number of slits varies up to five with which the

[c] JUWELS Booster: `https://apps.fz-juelich.de/jsc/hps/juwels/booster-overview.html`

[d] Vampir trace visualizer: `https://vampir.eu/`

diffraction pattern is obtained in Fig. 1. Fig. 4 shows the schematic of an acoustic domain illustrated with a set of parameters. The grating panels are defined with the thickness d which varies from λ_0 to $4\lambda_0$ where λ_0 is the reference length of incident waves. The diffraction grating consists of eight elements distributed in the x direction. Each element is defined with a variable-density function which enables phase-shifts of propagating waves. The phase-shift difference is related with the distribution function and the distance d. The mixed forms of a quadratic and a linear function are used to define the distribution of variable density inside the grating elements.

7.2 Physical Solution

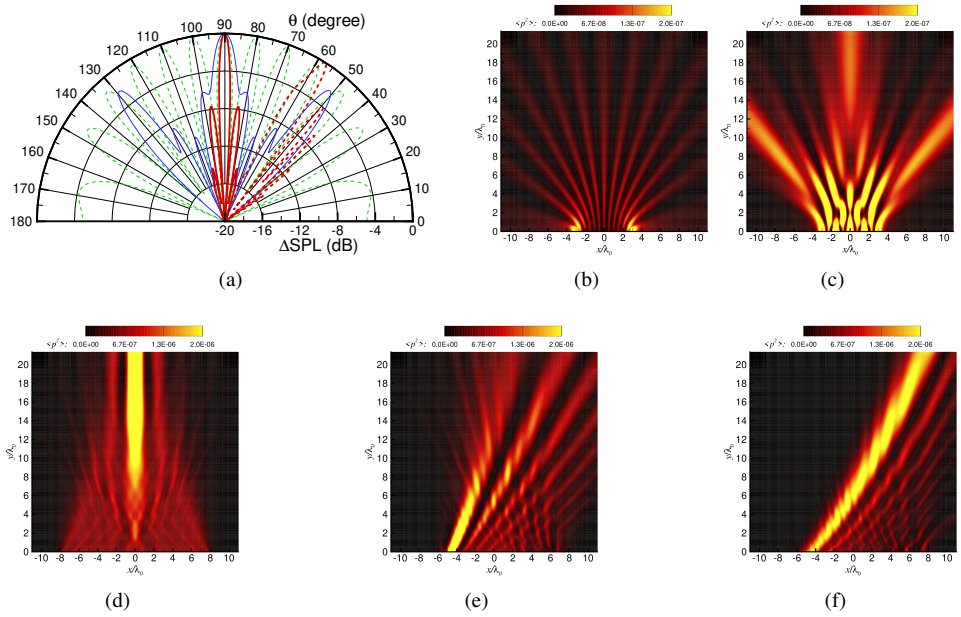


Figure 5. Contours of acoustic intensity obtained by guided waves, (a) directivity with the azimuthal angle θ, the intensity contours determined by (b) 2 slits with $d_G = 6\lambda$, (c) 5 slits with $d_G = 1.5\lambda$, (d) a density lens, (e) grating with a quadratic distribution of density, and (f) grating with a linear distribution of density.

In Fig. 5 the distributions of acoustic intensity are determined in the downstream of the diffraction gratings. The directivity in Fig. 5(a) is calculated with the pressure signals obtained by the virtual microphones at 20λ away from the grating exits. The origin is locates at the xyz-coordinates (0,0,0) for (b), (c), and (d) and (-5,0,0) for (e) and (f). The intensity maximum of each case is used for a normalised sound pressure level ΔSPL, i.e., 18 dB at the propagation angle $\theta = 90°$ for green dashes by (b), 26 dB for blue solid lines by (c), 41 dB for red solid lines by (d), and 38 dB at $\theta = 58°$ for red dashes by (f).

Fig. 6 illustrates the guided circular waves transformed to a focused beam (waves coloured by red and blue) using phased diffraction grating. Colours indicate contours of

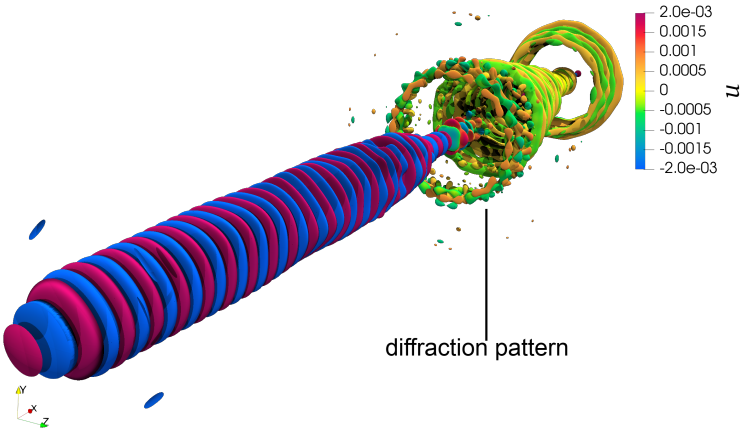


Figure 6. Contours of the acoustic pressure determined by phase diffraction grating. Colours indicate the perturbed velocity component in the x direction.

perturbed velocity component u. The full three-dimensional computational domain consists of ca. 1.77 billion meshes with the minimum cell spacing of $0.02\lambda_0$.

8 Summary and Outlook

A computational aeroacoustics approach is used to analyse the acoustic field generated by a diffraction grating. Simulations showed that the synthesised structures are able to control the wave propagation and the acoustic intensity in the targeted region. A numerical solver exploits the OpenACC directives to achieve a significant performance gain by using GPGPU on the JUWELS Booster. A performance analysis was detailed by using Score-P and the visualisation via Vampir. Production runs on the JUWELS Booster show a 2.3-fold runtime improvement in FOA, which is consistent with results from the Helmholtz GPU Hackathon[12].

The outlook addresses the relationship of the material parameters with the acoustic variables which is essential to define an optimal structure. The correlation of the flow variables with the characteristic length scales of the optimum design will be deduced by quantifying the structure effects on the flow field and the acoustic impedance. This prospective study includes a deep learning technique that is available via the open-source reinforcement learning library *smarties*[13]. The library performance has been proved in studying various unsteady dynamics problems. In simulation of acoustic fields the agents are distributed evenly along a spherical surface perpendicular to the three-dimensional incident waves, with each agent obtaining state information (local acceleration of fluid particles) at a polar angle ϕ^a, computing the reward on the surface and feeding into the policy function to obtain actions in the next step. The details will be presented with providing an optimised phased diffraction grating for targeted acoustics.

Acknowledgements

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[11] at Jülich Supercomputing Centre (JSC).

References

1. S. Cummer, J. Christensen, and A. Alù, *Controlling Sound with Acoustic Metamaterials*, Nature Reviews Materials, **1**, 16001, February 2016.
2. S. R. Koh, J.-H. Kim, and A. Lintermann, *Numerical Analysis of Oscillatory Flows in the Human Brain by a Lattice-Boltzmann Method*, in: Proceedings of 14th WCCM-ECCOMAS Congress 2020, 2021.
3. D. Brown, *Sound-Films as Diffraction Gratings for the Visual Fourier Analysis of Sound-Waves*, Nature, **140**, 1099, 1937.
4. Z. Tian and L. Yu, *Elastic Phased Diffraction Gratings for Manipulation of Ultrasonic Guided Waves in Solids*, Physical Review Applied, **11**, 024052, 2019.
5. Y. Wang, Y. Cheng, and X. Liu, *Modulation of Acoustic Waves by a Broadband Metagrating*, Science Reports, **9**, 7271, 2019.
6. R. Ewert and W. Schröder, *Acoustic Perturbation Equations Based on Flow Decomposition via Source Filtering*, Journal of Computational Physics, **188**, 365–398, 2003.
7. C. K. W. Tam and J. C. Webb, *Dispersion-Relation-Preserving Finite Difference Schemes for Computational Acoustics*, Journal of Computational Physics, **107**, no. 2, 262–281, 1993.
8. S. Johansson, *High Order Finite Difference Operators with the Summation by Parts Property Based on DRP Schemes*, Tech. Rep. 2004-035, Uppsala University, 2004.
9. F. Q. Hu, M. Y. Hussaini, and J. L. Manthey, *Low-Dissipation and Low-Dispersion Runge-Kutta Schemes for Computational Acoustics*, Journal of Computational Physics, **124**, 177–191, 1996.
10. C. Bogey and C. Bailly, *A Family of Low Dispersive and Low Dissipative Explicit Schemes for Flow and Noise Computations*, Journal of Computational Physics, **194**, no. 1, 194–214, 2004.
11. Jülich Supercomputing Centre, *JUWELS Cluster and Booster: Exascale Pathfinder with Modular Supercomputing Architecture at Juelich Supercomputing Centre*, Journal of large-scale research facilities, **7**, no. A138, 2021.
12. OpenACC Organization, "Helmholtz GPU Hackathon 2021", 2021, March 15–24 (Digital Event), Jülich Supercomputing Centre.
13. G. Novati and P. Koumoutsakos, *Remember and Forget for Experience Replay*, in: Proceedings of the 36th International Conference on Machine Learning, 4851–4860, 2019.

Oscillatory Control of Compressible Channel Flow Using Direct Numerical Simulations

Marius Ruby and Holger Foysi

Chair of Fluid Dynamics, University of Siegen, 57072 Siegen, Germany
E-mail: {marius.ruby, holger.foysi}@uni-siegen.de

An oscillatory turbulence control technique has been implemented with the aim of obtaining altered, energetically more favourable flow characteristics. The aim of this active control method, where energy input is required in order to manipulate the flow, is to achieve the goal of a lower turbulence intensity and thus friction drag. The latter effects outweigh the required input power yielding large positive net power reductions throughout the range of investigated parameters. The control method is applied to turbulent compressible channel flow. This well-studied canonical flow case is characterised by two confining horizontal static and rigid plane walls. In the controlled case, by contrast, we prescribe a sinusoidal modulation of spanwise velocity at the walls dependent on the streamwise coordinate, originally introduced by Viotti *et al.*[1]. This way the momentum transfer between the wall region and the outer high speed flow is disturbed and thus the important self-sustaining turbulence mechanisms. The consequence is a lower turbulence intensity accompanied with lower friction drag at the walls. Most applications in the context of turbulence control including the present principle are related to incompressible flow. In this work, however, the specific control method is applied to supersonic flow as well. Particular attention is directed to characteristics purely associated with compressible flow and interactions of the control method with the latter. We make use of direct numerical simulations in order to resolve the relevant features temporally and spatially. In virtue of the arising problem size, the A100 GPUs of Jülich Supercomputing Centre had to be used providing high performance and excellent scaling.

1 Introduction

Making use of the computing time opportunities for research facilities in Germany, offered by Gauss Centre for Supercomputing and John von Neumann Institute for Computing, an oscillatory flow control technique has been applied to turbulent, compressible channel flow. Flow control in general aims for altering the flow in respect of a specific desired benefit, where leaving the flow to its own would waste potential. Depending on the case of application such a desirable benefit could be a delay of the transition from laminar to turbulent flow, the prevention of flow separation at curved boundaries or a reduction of the flow induced noise. Even locally increasing the turbulence intensity is sometimes favoured in order to enhance mixing of multiple phases or to assist in the ability of the flow to follow the shape of blunt body. An example for the latter is the particular shape of a golf ball that allows for significantly longer drives compared to a smooth spherical ball surface. An overview of application fields, classification and advances of flow control is given in Gad-el-Hak[2] for example. In this work we focus on reducing the turbulence level by means of an oscillatory wall control technique. In wall-bounded flows the velocity builds up from zero at the wall to the high speed outer flow within a small vertical distance, resulting in large velocity gradients that contribute to the friction drag. In turbulent flow the gradient is even higher owing to enhanced momentum transfer between the high velocity and low-speed wall region. Despite the extreme complexity of the turbulence problem (an impression of

the multitude of length scales involved is given in Fig. 1), there are certain well-known near wall flow structures and motions that contribute to turbulence production to a great extent[3,4].

An effective way to influence these flow structures is imprinting oscillations of suitable frequency directly at the wall. Momentum transfer between near wall flow and the outer layer is disturbed this way, leading to a lower velocity gradient at the wall and thus lower friction drag. Various methods related to oscillatory control distinguish themselves in the specification of the velocity distribution at the wall (e.g. direction, spatial and temporal dependence). There are spanwise oscillatory wall motions[5], streamwise (longitudinal) travelling waves of spanwise velocity[6] or even spanwise travelling transversal waves through flexible surfaces[7]. The oscillation technique used in this project is a pure space dependent formulation by Viotti *et al.*[1,8], where the spanwise velocity at the wall is described by

$$W = A \sin\left(\frac{2\pi}{\lambda_x} x\right) . \tag{1}$$

Here, A denotes the oscillation amplitude, the streamwise coordinate is referred to as x and λ_x is the wavelength in that streamwise direction.

The flow setup and underlying equations along with employed solution procedures will be outlined in Sec. 2. Parameters used for characterising the flow and arising computational challenges are discussed in Sec. 3. Sec. 4 demonstrates the fundamental operation of the applied control method as well as peculiarities with reference to the application in compressible flow.

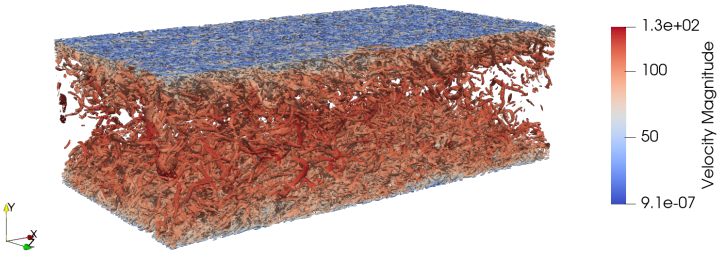


Figure 1. Turbulent channel with dimensionless flow parameters $Ma_b = 0.3$ and $Re_\tau = 800$. Vortical structures are made visible by aid of the Q criterion and coloured with the vorticity magnitude.

2 Flow Setup and Governing Equations

A sketch of the flow setup demonstrating the control method at the same time is given in Fig. 2. The rectangular flow domain with dimensions L_x, L_y and L_z is confined by the top and bottom walls, where the steady spanwise velocity distribution according to Eq. 1 is determined in the controlled case. The coordinate system is set in a way that x and z represent the homogeneous streamwise and spanwise directions, respectively, with periodic boundary conditions. y indicates the wall normal direction ranging from $-h$ to h, h being

the channel half width. Moreover the resulting mean streamwise velocity distribution $u(y)$ is schematically depicted. The underlying problem involves solving the three dimensional

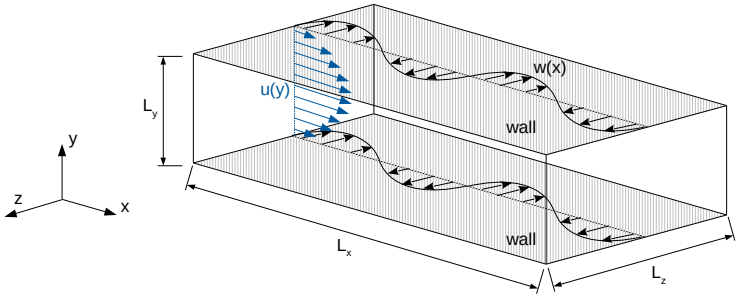


Figure 2. Sketch of the controlled channel flow configuration along with corresponding dimensions, directions and velocity distributions.

compressible Navier-Stokes equations comprising conservation laws for mass, momentum and energy. Using conservative variables ρ (density), ρu_i (momentum) and E (total energy), they read

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho u_j\right) = 0\,, \tag{2}$$

$$\frac{\partial}{\partial t}\left(\rho u_i\right) + \frac{\partial}{\partial x_j}\left(\rho u_i u_j\right) + \frac{\partial p}{\partial x_i} + f\delta_{i1} = \frac{\partial \sigma_{ij}}{\partial x_j}\,, \tag{3}$$

$$\frac{\partial E}{\partial t} + \frac{\partial}{\partial x_j}\left(E + p\right)u_j + f u_1 = \frac{\partial}{\partial x_j}\sigma_{ij}u_i - \frac{\partial}{\partial x_j}q_j\,, \tag{4}$$

with time t, spatial coordinates x_i and σ_{ij} representing the viscous stress tensor. The flow is driven by an external force f ensuring a constant flow rate. The total energy E is given by

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho u_j u_j\,, \tag{5}$$

where $\gamma = c_p/c_v$ is the ratio of specific heats at constant pressure c_p and constant volume c_v. The temperature is obtained with the aid of the ideal gas law as

$$T = \frac{1}{c_v}\frac{1}{\gamma - 1}\frac{p}{\rho}\,. \tag{6}$$

The system of equations is solved numerically for the solution vector composed of density, momentum components and total energy. For that purpose the flux reconstruction approach by Huynh[9], implemented in the flow solver PyFR and developed at Imperial College London is used[10]. Since the numerical scheme features high temporal and spatial

locality, it is appealing for massively parallel applications on modern hardware platforms. The majority of operations within a time-step is matrix multiplication, which can be done by efficient BLAS libraries designed for the appropriate hardware platform. PyFR has been selected as finalist of the Gordon Bell Price in High Performance Computing due to its features to efficiently target multiple platforms. It is written in Python and automatically generates/compiles platform specific C++ code at runtime. The simulations here were performed on the accelerated nodes of JUWELS Cluster and Booster module in Jülich using the CUDA backend of the PyFR framework.

3 Flow Parameters and Computational Challenges

This section introduces several dimensionless quantities that are used to characterise the flow. An important quantity in compressible channel flow is the bulk Mach number, that is defined as the ratio of the bulk flow velocity and the speed of sound at the wall

$$M_b = \frac{U_b}{c_w}\,. \tag{7}$$

Previous work showed, that variable density effects started to become noticeable for Mach numbers greater than 1.5[3,4]. The bulk velocity

$$U_b = \frac{1}{2h\rho_b}\int_{-h}^{h} <\rho u> dy \tag{8}$$

and bulk density

$$\rho_b = \frac{1}{2h}\int_{-h}^{h} <\rho> dy \tag{9}$$

are mean quantities with $<\cdot>$ denoting temporal and spatial averaging over homogeneous directions. Further important quantities are the bulk Reynolds number

$$Re_b = \frac{\rho_b u_b h}{\mu_w}, \tag{10}$$

composed of the latter bulk quantities as well as the channel half-height h and the viscosity at the walls, μ_w and the friction Reynolds number to characterise the near-wall flow,

$$Re_\tau = \rho_w u_\tau h/\mu_w. \tag{11}$$

Here, ρ_w is the density at the wall and a characteristic velocity called the friction velocity in the near-wall region, $u_\tau = (\tau_w/\rho_w)^{1/2}$, derived from the wall shear stress τ_w. In compressible wall-bounded flow, another Reynolds number, defined as

$$Re^*_\tau = Re_\tau \sqrt{\frac{<\rho>}{<\rho_w>}}\frac{\mu}{\mu_w} \tag{12}$$

is used, that supplements the latter by an additional factor $\sqrt{\frac{<\rho>}{<\rho_w>}}\frac{<\mu>}{<\mu_w>}$, accounting for substantial variations in viscosity, density and temperature in the direction of the wall normal coordinate. It is called semi-local Reynolds number and it turned out that overall flow characteristics coincide very well when comparing flows with matching semi-local Reynolds number evaluated at the channel centre $Re^*_{\tau,c}$ [11,3,4,8]. The investigated flow cases

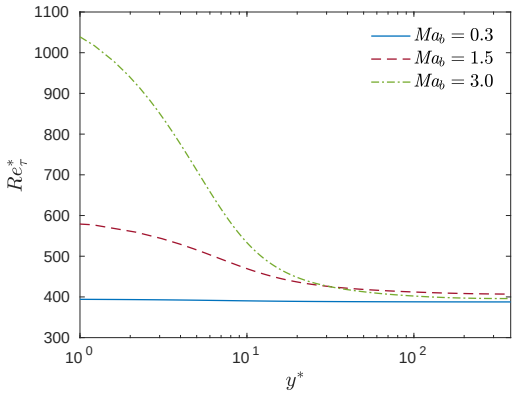


Figure 3. Semi-local Reynolds number Re^*_τ for different Mach numbers Ma_b and with $Re^*_{\tau,c} = 400$ dependent on the semi-local wall coordinate y^*.

of the present project comprise three different Mach numbers $Ma_b = 0.3$, 1.5 and 3.0 each for three different Reynolds numbers $Re^*_{\tau,c} = 150$, 400 and 800 to study the influence of these two nondimensional parameters. The associated numerical grids contain between 4 and 168 million cells and require utilisation of up to 64 nodes on JUWELS Booster equipped with four NVIDIA A100 GPUs each. The computational effort increases with growing Reynolds number, as the spectrum of time- and length scales to be resolved gets wider. The same holds true for increasing the Mach number as Fig. 3 illustrates. First of all, the semi-local Reynolds number for the nearly incompressible case at $Ma_b = 0.3$ shows barely any alteration throughout the wall-normal coordinate. However, at $Ma_b = 3$ it exhibits nearly triple the value at the wall compared to the central region accompanied with associated requirements concerning spatial resolution and time step size. Details of flow conditions and computational domains of the proposed runs are given in Tab. 1.

Flow conditions				**computational domain**		
Ma_b	Re_b	Re_τ	$Re^*_{\tau c}$	$\frac{L_x \times L_y \times L_z}{h}$	$N_x \times N_y \times N_z$	# cells [10^6]
0.3	3000	192	188	$6\pi \times 2 \times 2\pi$	$288 \times 128 \times 152$	3.7
1.5	3000	218	146	$6\pi \times 2 \times 2\pi$	$288 \times 128 \times 152$	5.6
3.0	4880	454	148	$6\pi \times 2 \times 3/2\pi$	$752 \times 224 \times 288$	48.5
0.3	6890	396	388	$3\pi \times 2 \times \pi$	$300 \times 160 \times 200$	9.6
1.5	9450	604	407	$3\pi \times 2 \times \pi$	$380 \times 200 \times 252$	19.2
3.0	14000	1150	398	$3\pi \times 2 \times \pi$	$564 \times 336 \times 375$	71.1
0.3	15400	804	784	$2\pi \times 2 \times \pi$	$400 \times 448 \times 400$	71.7
1.5	20000	1182	807	$2\pi \times 2 \times \pi$	$396 \times 336 \times 396$	52.7
3.0	33035	2522	855	$3/2\pi \times 2 \times \pi/2$	$600 \times 700 \times 400$	168

Table 1. Details of flow conditions and computational domains of the proposed runs. Configurations of approximate matching semi-local Reynolds number at the channel centre $Re^*_{\tau,c}$ are combined.

The scaling behaviour up to 96 GPUs of a flow configuration that is part of the project is shown in Fig. 4. Mach and semi-local Reynolds number were $Ma_b = 3$ and $Re^*_{\tau,c} = 400$, respectively. The order of the solution polynomials was chosen such that the number of cells added up to $112 \cdot 10^6$. Results are shown for both the GPUs of the *JUWELS Cluster module* and the *JURECA DC GPU* partition. They start at the smallest possible number of GPUs in each case, whereby the NVIDIA A100 GPUs at *JURECA DC* module make use of 40 GB HBM2e memory compared to 16 GB HBM memory of JUWELS. The run time on the more powerful A100 GPUs is roughly half of the corresponding one using V100 GPUs. Regarding the accumulated computing time for the project, its division into several parallel runs and the available resources per call, we found the *JUWELS Booster* module most suitable.

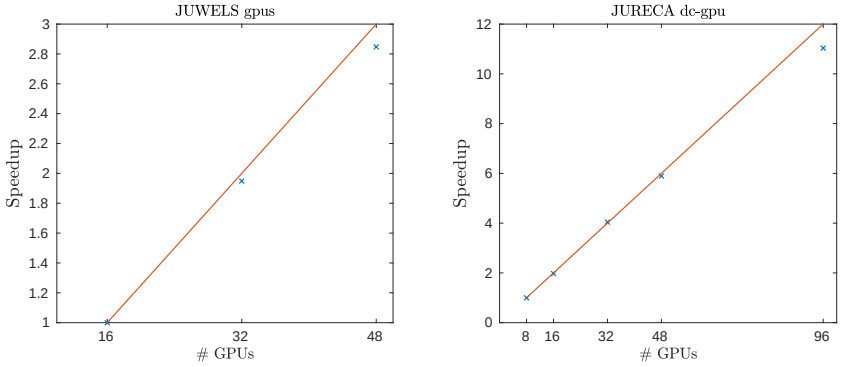


Figure 4. Scaling behaviour using a test case part of the project on the GPU architectures of supercomputers JUWELS and JURECA at Jülich. This data was obtained with a problem size of 112 Mio. cells.

4 Application of Spanwise Wall Oscillations

The control method is applied for the whole set of flow configurations discussed in Sec. 3 and specified in Tab. 1, each with varying control parameter λ_x according to Eq. 1. The general method of operation by the wall oscillation technique is visualised with the aid of coherent structures in close vicinity to the wall (Fig. 5). Instabilities of these vortical structures play an eminent role for the interchange of low / high speed fluid and finally for the generation of wall friction drag. As it can be seen on the right hand side of Fig. 5, the control method is capable of distinctly modifying the streaks in terms of distance, strength and alignment within the range of parameters used in this work. The temporal evolution of the normalised friction coefficient dependent on the applied wavelength λ_x is exemplary shown for flow conditions $Ma_b = 1.5$ and $Re^*_\tau = 400$ in Fig. 6.

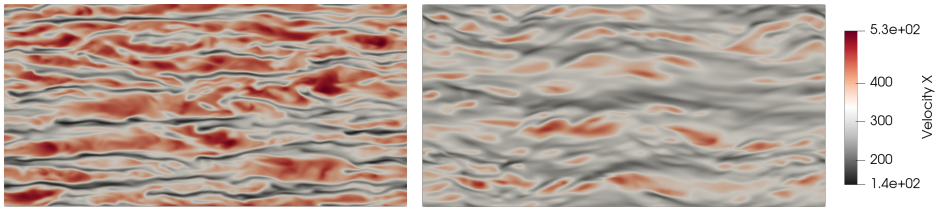


Figure 5. Visualisation of Coherent structures at $Ma_b = 1.5$ and $Re^*_{\tau,c} = 400$ in horizontal wall-parallel planes with wall-normal distance $y^* = 20$. Left: reference flow, right: controlled case.

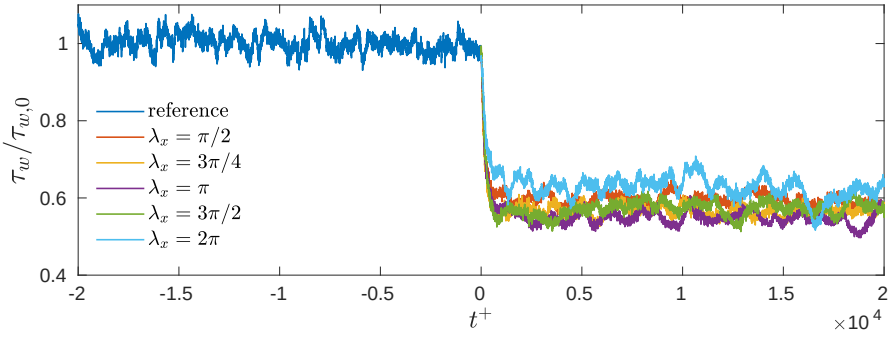


Figure 6. Normalised friction coefficient at $Ma_b = 1.5$ and $Re^*_{\tau,c} = 400$ for different control parameters λ_x. t^+ denotes a dimensionless time, its zero value indicates the beginning of the control period.

4.1 Influence of Compressibility

An overview of the drag reducing property by the applied control method is given in Fig. 7. Values of the skin friction coefficient $2\tau_w/(\rho_b u_b^2)$ normalised by the uncontrolled ones are shown for simulations with varying control parameters λ_x and flow conditions accomplished so far. While the decreasing effectiveness with higher Reynolds numbers is a well known effect of wall based methods like the one that is used in this project, it is more distinct at higher control wavelengths λ_x. The occurrence of a pronounced optimum oscillation wavelength is visible for all studied Mach numbers. However, the optimum wavelength is shifted to higher values for the higher Mach number cases. Furthermore, the optimum is less distinct, allowing for low values of the friction coefficient throughout the whole range of investigated wavelengths. In addition, the achieved maximum drag reduction slightly increases with growing Mach number. In the case of flow parameters $Ma_b = 1.5$ and $Re^*_{\tau,c} = 140$ relaminarisation occurred, apparent in the large drop in the skin friction coefficient from a certain wavelength on. The turbulence self-sustaining cycle is interrupted entirely in that case. It is possible through a favourable interplay of an overall low Reynolds number and increasingly larger vortex structures owing to compressibility effects. Thus it can rather be seen as an anomaly related to the remaining data set. A detailed discussion of the effects of Mach and Reynolds numbers, using a decomposition of the wall shear stress, for example, can be found in Ruby & Foysi (2022)[8].

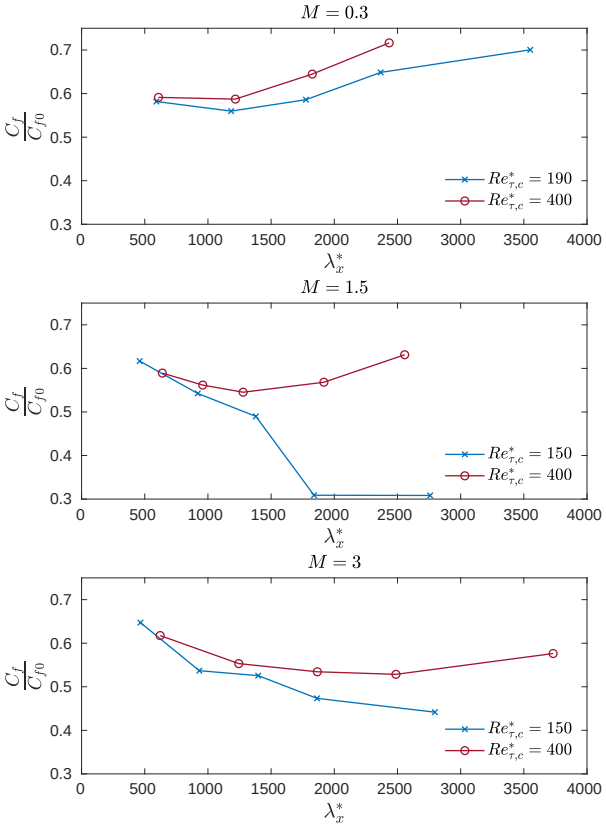


Figure 7. Skin friction coefficient C_f with reference to the respective uncontrolled reference flow with skin friction coefficient C_{f0} in various flow conditions and with different control parameters λ_x. Top: $Ma_b = 0.3$, centre: $Ma_b = 1.5$, bottom: $Ma_b = 3.0$.

5 Concluding Remarks

We have applied a wall-based oscillatory flow control method to supersonic turbulent channel flow. By means of a fairly large database of different flow and control parameters, we want to gain knowledge how the performance behaves under these different circumstances compared to purely incompressible conditions. Due to the requirements concerning resolution and time-step size as well as the large number of different runs, the computing capabilities of John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC) are used for this purpose.

A comparison of the results obtained so far, points to a better performance with increasing influence of compressibility. The method is more indulgent to deviations from the optimum control parameter under these flow conditions. The preference of larger wavelengths is furthermore appealing with regard to an easier implementation of derived experimental methods.

Acknowledgements

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) during GCS Call 1/2021 (osccompchannel21). Additionally we thank our local IT centre ZIMT for providing scientific computing capabilities to members of Uni Siegen.

References

1. C. Viotti, M. Quadrio, and P. Luchini, *Streamwise oscillation of spanwise velocity at the wall of a channel for turbulent drag reduction*, Physics of Fluids, **21**, no. 11, 115109, 2009.
2. M. Gad-el Hak, *Flow Control: Passive, Active, and Reactive Flow Management*, Cambridge University Press, 2000.
3. H. Foysi, S. Sarkar, and R. Friedrich, *Compressibility Effects and Turbulence Scalings in Supersonic Channel Flow*, J. Fluid Mech., **509**, 207–216, 2004.
4. S. Gosh, H. Foysi, and R. Friedrich, *Compressible turbulent channel and pipe flow: similarities and differences*, J. Fluid Mech., **648**, 155–181, 2010.
5. E. Touber and M. A. Leschziner, *Near-wall streak modification by spanwise oscillatory wall motion and drag-reduction mechanisms*, Journal of Fluid Mechanics, **693**, 150–200, 2012.
6. M. Quadrio, P. Ricco, and C. Viotti, *Streamwise-travelling waves of spanwise wall velocity for turbulent drag reduction*, Journal of Fluid Mechanics, **627**, 161–178, 2009.
7. D. Roggenkamp, W. Jessen, W. Li, and W. Schröder, *Experimental investigation of turbulent boundary layers over transversal moving surfaces*, CEAS Aeronautical Journal, **6**, 471–484, 09 2015.
8. M. Ruby and H. Foysi, *Active control of compressible channel flow up to $Ma_b = 3$ using direct numerical simulations with spanwise velocity modulation at the walls*, GAMM Mitteilungen, Special Issue: Direct Numerical Simulations of Turbulent Flows – Part I, **45 (1)**, 2022.
9. H. T. Huynh, *A Flux Reconstruction Approach to High-Order Schemes Including Discontinuous Galerkin Methods*, AIAA Paper AIAA 20074079, 1–42, 01 2007.
10. F. Witherden, B. C. Vermeire, and P. Vincent, *Heterogeneous Computing on Mixed Unstructured Grids with PyFR*, Computers & Fluids, **120**, 09 2014.
11. P. G. Huang, G. N. Coleman, and P. Bradshaw, *Compressible turbulent channel flows: DNS results and modelling*, J. Fluid Mech., **305**, 185–218, 1995.

Analysis of the Large-Scale Order in Turbulent Mesoscale Convection

Philipp P. Vieweg[1], Florian Heyder[1], John Panickacheril John[1], Janet D. Scheel[2], and Jörg Schumacher[1,3]

[1] Institute of Thermodynamics and Fluid Mechanics, Technische Universität Ilmenau, 98684 Ilmenau, Germany
E-mail: {philipp.vieweg, florian.heyder, john.panickacheril-john, joerg.schumacher}@tu-ilmenau.de

[2] Department of Physics, Occidental College, 1600 Campus Road M21, Los Angeles, CA 90041, USA
E-mail: jscheel@oxy.edu

[3] Tandon School of Engineering, New York University, New York, NY 11201, USA

Highly turbulent flows in nature display often regular and gradually evolving patterns at an intermediate range of scales which we term an order at the mesoscale. Here, we summarise our analyses on the basis of comprehensive direct numerical simulations to reveal the physical origins of this order, its relevance for the turbulent transport of momentum and heat, as well as our efforts to reduce the complex dynamics in lower-dimensional models by applying machine learning methods. At the core of these studies is the Rayleigh-Bénard convection model, which we recently extended to the fully compressible convection in order to incorporate non-Boussinesq effects and thus to come closer to real flow behaviour.

1 Introduction

Mesoscale convection is a central link for the turbulent transport of momentum, mass and heat in many natural flows which bridges the small-scale turbulence and the global circulation processes, e.g. in planetary atmospheres or in stellar interiors to mention some examples[1,2]. The mesoscale implies that these flows proceed in domains, where the horizontal extensions L are large in comparison to the characteristic height H of the layer. Thus the aspect ratio $\Gamma = L/H \gg 1$. Mesoscale convection processes in nature are typically multi-physics fluid motions which include for example the rotation of the layer, phase changes or radiative transport. Convection flows in large aspect ratio domains tend to form a large-scale order, partly even a hierarchical order of different regular patterns, despite being highly turbulent. The numerical detection of these *turbulent superstructures* requires direct numerical simulations (DNS) on big grids that can cover extended domains at sufficiently high Rayleigh numbers Ra[3,4]. These superstructures are found to evolve very slowly, which we analysed in the past NIC supercomputing projects in detail.

Typically DNS of turbulent convection are run with constant prescribed temperatures at the top and bottom planes. The situation was found to change drastically when the boundary conditions for the temperature field are changed to that of a constant heat flux[5]. In the latter simulation configuration, one big supergranule cell is formed when the simulations are run long enough for the non-rotating flow case. This supergranule aggregation proceeds until the whole computational domain is filled. The process proceeds over several

thousands of convective time units. It is superposed with smaller convection cells, known as granules, as visible in Figs. 1(a–d). The subsequent analysis in the Fourier space clearly indicates an accumulation of thermal variance or turbulent kinetic energy at the smallest horizontal wavenumbers (which correspond to the largest scales in the system), see panels (e–l) in the middle and bottom rows of Fig. 1.

In this article, we report comprehensive direct numerical simulations to understand this (hierarchical) order and its relevance for the turbulent transport of momentum and heat. We will also discuss the impact of an increase of the complexity of the flow problem, particularly by including rotation and a switch from the idealised Boussinesq to the non-Boussinesq case. Non-Boussinesq effects, such as asymmetric boundary layers at the top and bottom arise for example when the model is generalised to the compressible case. We present furthermore results to reduce the complex dynamics in lower-dimensional models by applying machine learning methods.

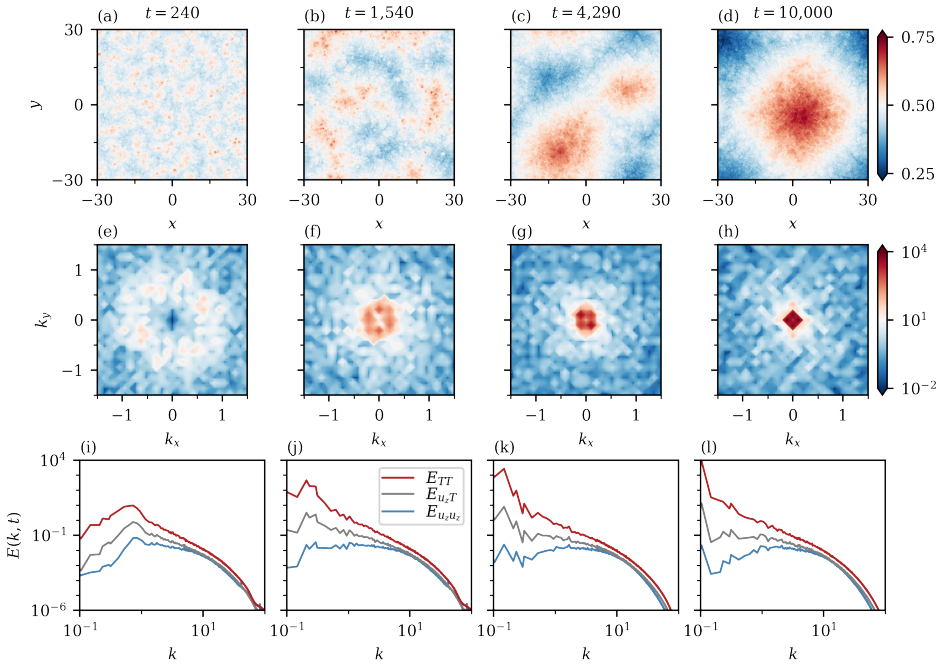


Figure 1. Gradual aggregation of the supergranule in physical and wavevector space for the non-rotating case. The process is shown for the supergranule that forms at a Rayleigh number of $\mathrm{Ra} = 3.93 \times 10^6$ for a time period of 10^4 free-fall times as an example. (a–d) Temperature contour plots close to the top plane. Times for the snapshots are indicated. The colour bar applies for all panels in this row. (e–h) Corresponding Fourier spectra taken from temperature fluctuations at the same horizontal plane with respect to the horizontal wave vector components (k_x, k_y). The wavenumber range on both axes magnifies the smallest value to demonstrate the variance accumulation in the largest possible modes. The colour bar is again the same for all panels in this row. (i–l) Additionally azimuthally averaged Fourier spectra with respect to the horizontal wavenumber $k = (k_x^2 + k_y^2)^{1/2}$ are shown. The Fourier transform of three different quantities is displayed as indicated in the legend in panel (j): temperature variance spectrum, co-spectrum of the turbulent convective heat flux, and the kinetic energy spectrum with respect to the vertical velocity component.

2 Turbulent Convection Model

2.1 Fully Compressible Non-Boussinesq Case

The complete compressible Navier-Stokes equations comprising the conservation equations of mass, momentum, and internal energy densities which are denoted as ρ, ρu_i, and ρe, respectively, are given by[2,6]

$$\frac{\partial \rho}{\partial t} = -\frac{\partial (\rho u_i)}{\partial x_i}\,, \tag{1a}$$

$$\frac{\partial (\rho u_i)}{\partial t} = -\frac{\partial (\rho u_i u_j)}{\partial x_j} - 2\rho\epsilon_{ijk}\Omega_j u_k - \frac{\partial p}{\partial x_i} + \frac{\partial \sigma_{ij}}{\partial x_j} - \rho g \delta_{i3}\,, \tag{1b}$$

$$\frac{\partial (\rho e)}{\partial t} = -\frac{\partial (\rho e u_j)}{\partial x_j} - p\frac{\partial u_i}{\partial x_i} + \frac{\partial}{\partial x_i}\left(k\frac{\partial T}{\partial x_i}\right) + \sigma_{ij} S_{ij}\,, \tag{1c}$$

which are supplemented by an equation of state which connects pressure p, temperature T, and mass density ρ to

$$p = \rho R T\,. \tag{1d}$$

Here, $e = c_v T$ with the specific heat at constant volume c_v. Furthermore, k is the thermal conductivity, g is the acceleration due to gravity, and $\Omega_j = \Omega\delta_{j3}$ is the vector of constant rotation rate. The equation of state is that of an ideal gas with the specific gas constant $R = c_p - c_v$; the specific heat at constant pressure is c_p. The viscous stress tensor and strain rate tensor are defined as

$$\sigma_{ij} = \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\delta_{ij}\frac{\partial u_k}{\partial x_k}\right)\,, \tag{2a}$$

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\,. \tag{2b}$$

In the convection case, periodicity is applied only in two directions, x_1 and x_2. No-slip or free-slip boundary conditions at $x_3 = 0$ and $x_3 = H$ are applied for the velocity field u_i. Constant temperatures or constant heat flux boundary conditions follow for the temperature field. We will sometimes denote $x_1 = x$, $x_2 = y$, and $x_3 = z$.

We use a pencil 3d domain decomposition. This allows us to efficiently implement compact finite difference schemes with spectral like resolution[7,8] for the spatial derivatives appearing at the right hand side of the compressible equations using BLAS library subroutines. The configurations are changed using collective "MPI-ALL TO ALL" calls. We use a 6th-order Padé scheme in the present simulations. Time advancement is done by a third-order Runge-Kutta scheme. The most important dimensionless parameters that characterise the convection flow are the Rayleigh number Ra with the superadiabaticity $\varepsilon = \Delta T/T_0$, the Prandtl number Pr, the Rossby number Ro, the convective Mach number M_c, and the aspect ratio $\Gamma = L/H$,

$$\mathrm{Ra} = \frac{c_p g \rho_0 \varepsilon H^3}{\nu k}, \quad \mathrm{Pr} = \frac{c_p \rho_0 \nu}{k}, \quad \mathrm{Ro} = \frac{U_f}{2\Omega H} \quad \text{and} \quad M_c = \frac{U_f}{c_s}\,, \tag{3}$$

where T_0 and ρ_0 are reference temperature and densities, respectively, which are taken at the bottom plate. The speed of sound is c_s and the free-fall velocity $U_f = (g\varepsilon H)^{1/2}$.

2.2 Incompressible Boussinesq Case

For cases $|u_i| \ll c_s$, the mass density becomes constant, $\rho \simeq \rho_0$ and the velocity field is to a good approximation free of divergences. An exception is the volume forcing term on the right hand side of the Navier-Stokes equations, see Eq. 1b. The Boussinesq approximation assumes in this regime that the mass density is a linear function of the (small) temperature deviation from the reference state, $\rho = \rho_0[1 - \alpha(T - T_0)]]$. This case is also known as the Rayleigh-Bénard convection (RBC) model. Eqs. 1a–1c simplify to

$$\frac{\partial u_i}{\partial x_i} = 0\,, \tag{4a}$$

$$\frac{\partial u_i}{\partial t} = -\frac{\partial\,(u_i u_j)}{\partial x_j} - 2\epsilon_{ijk}\Omega_j u_k - \frac{1}{\rho_0}\frac{\partial p}{\partial x_i} + \nu\frac{\partial^2 u_i}{\partial x_j^2} + g\alpha(T - T_0)\delta_{i3}\,, \tag{4b}$$

$$\frac{\partial T}{\partial t} = -\frac{\partial\,(u_j T)}{\partial x_j} + \frac{k}{\rho_0 c_p}\frac{\partial^2 T}{\partial x_j^2}\,, \tag{4c}$$

Here, T_0 is a reference temperature and α is the thermal expansion coefficient. In the incompressible limit, the pressure looses its physical content as an independent state variable. Its role is to readjust the incompressibility of the velocity field. For the incompressible cases of turbulent convection, we applied a spectral element method[9]. Time advancement is done by a backward difference method BDF2.

3 Effects of Moderate Rotation and Compressibility

In the following, we discuss in brief the effects of moderate rotation and compressibility on the structure formation in mesoscale convection. In the case of moderate rotation, we solve the Eqs. 4a–4c; in the compressible case we apply the model which was outlined in Sec. 2.1.

The results of Fig. 1 showed that a simple Boussinesq model is able to generate a hierarchy of convection structures, namely granules and a big supergranule in the background, once the appropriate thermal boundary conditions are chosen and the integration is done for a very long time. Such a hierarchy is observed in natural flows, e.g. solar convection. We could show that this growth process is connected with an inverse cascade in the Fourier spectral space. For example, in the case of free-slip velocity boundary conditions in combination with constant flux temperature boundary conditions, the subset of vertically homogeneous Fourier modes, which depend on wavevector components k_x and k_y only and $k_z = 0$, forms an inverse cascade that dominates the transport of thermal variance and turbulent kinetic energy from smaller to the largest scales. Furthermore, we found that the mean amount of transferred heat from the bottom to the top remains constant as the supergranule aggregation proceeds.

A limitation of the growth to a supergranule at a scale smaller than the horizontal box length is obtained when the convection layer rotates around the vertical axis with $\Omega_j = \Omega\delta_{j3} \neq 0$. Slow or moderate rotation implies that the dimensionless Rossby number is $\mathrm{Ro} > 1$. The effect of the rotation on the large-scale structure is underlined in Figs. 2(a–c) where the fragmentation of the original supergranule into smaller structures

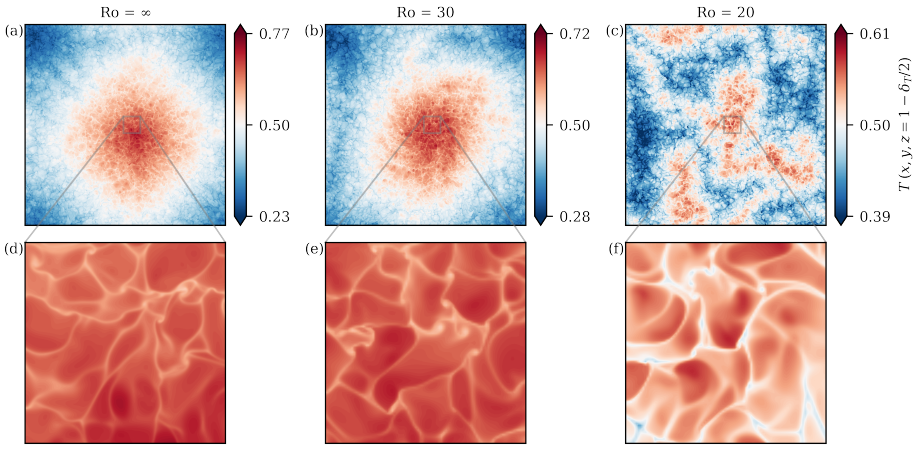


Figure 2. Large-scale structure of the temperature field close to the top surface. Panels (a)–(c) show the whole simulation cross-section of $60H \times 60H$ for different dimensionless Rossby numbers Ro which quantify the strength of rotation. The non-rotating case at $\mathrm{Ro} = \infty$ is found in the left column. Panels (d)–(f) show a magnification of a small region of $5H \times 5H$ to highlight the granular structure that is superposed to the supergranule. All snapshots are taken in the final statistically stationary regime of the dynamical evolution for a case of $\mathrm{Ra} = 3.93 \times 10^6$.

is demonstrated for a sufficiently small dimensionless Rossby number Ro. Recall, the smaller Ro the stronger the rotation. The magnifications in Figs. 2(d–f) highlight once more the superposed granules – the convection cells with an aspect ratio of one – which persist to exist in this change of the large-scale structure as a pattern that covers the whole convection layer. Note also that the resulting convection patterns still have a characteristic length Λ that is larger than the height of the layer. The rotation rate Ω was found to be directly proportional to Λ. The proportionality constant is in addition found to depend on the Rayleigh number Ra. We could also show that the inverse cascades of thermal variance and turbulent kinetic energy are interrupted for a finite moderate rotation.

Deviations from the Boussinesq regime can be immediately observed in the compressible convection case, as seen Fig. 3. In these simulations, we applied the Sutherland model for the dynamic viscosity and the thermal conductivity of compressible gases[10]. They are given by

$$\mu(T) = \mu_0 \frac{T^{3/2}}{C_\mu + T} \quad \text{and} \quad k(T) = k_0 \frac{T^{3/2}}{C_k + T}\,, \tag{5}$$

where μ_0 and C_μ as well as k_0 and C_k are constants. In the present studies, we took $C_\mu = C_k = 0$. A DNS snapshot of the fluid density $\rho(\mathbf{x}, t)$ is shown in panel (a) of the figure. One observes a strong increase of the mean fluid density from bottom to top. In panels (b) and (c), we plot the mean profiles of the temperature $\langle T(z)\rangle_{A,t}$ and the mean convective heat flux $\langle u_z T(z)\rangle_{A,t}$, respectively. Both figures show top-down-asymmetric profiles, which stand for a clear non-Boussinesq behaviour, as present in most natural convection flows. Detailed parametric investigations of the system are planned for the coming project year.

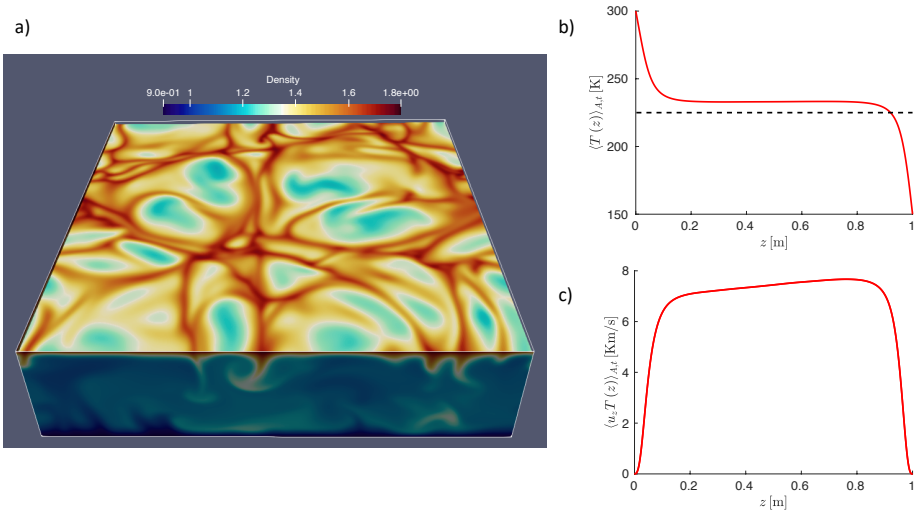


Figure 3. Compressible convection at $\mathrm{Ra} = 10^6$ and $\mathrm{Pr} = 0.72$. (a) Contour plot of a mass density snapshot with a colour bar given in kg/m^3. (b) Mean temperature profile across the convection layer in K. (c) Mean convective heat flux in Km/s. The aspect ratio of the cell is 4, the superadiabaticity is $\varepsilon = \Delta T/T_0 = 0.5$. The mean profiles are obtained by a combined average with respect to the horizontal cross section area A and time t.

4 Reduced Dynamical Models of Convection by Reservoir Computing

In the following, we discuss the application of recurrent neural networks (RNN) to develop reduced order models of turbulent convection. This specific class of machine learning models can process sequential data and is thus able to learn a dynamics[11,12]. As all machine learning models, they are data-driven and do not use the underlying partial differential equations which we have outlined in Sec. 2. Different recurrent neural network architectures are possible. Our focus is on reservoir computing models (RCM). The reservoir is a big, sparse, and randomly initialised adjacency matrix that substitutes the convoluted hidden layers in deep neural networks. The reservoir state dynamics is given by

$$\mathbf{r}(n) = (1-\gamma)\mathbf{r}(n-1) + \gamma \tanh\left[W^{\mathrm{r}}\mathbf{r}(n-1) + W^{\mathrm{in}}\mathbf{x}(n)\right]. \tag{6}$$

where $\mathbf{r}(n) \in \mathbb{R}^{N_{\mathrm{r}}}, \mathbf{x}(n) \in \mathbb{R}^{N_{\mathrm{in}}}$ are the reservoir state and input at time step n respectively. W^{r} and W^{in} are the reservoir and input weight matrices, respectively. The parameter $\gamma \in [0,1]$ is the constant leaking rate. In this formulation, the components of the reservoir state r_i are automatically $r_i \in [-1,1]$. The reservoir output $\hat{\mathbf{y}} \in \mathbb{R}^{N_{\mathrm{in}}}$ is computed by a linear mapping of the extended reservoir state $\tilde{\mathbf{r}}(n) = [\mathbf{x}(n), \mathbf{r}(n)]$ (vertical concatenation of reservoir input and state),

$$\hat{\mathbf{y}}(n) = W_*^{\mathrm{out}}\tilde{\mathbf{r}}(n). \tag{7}$$

The fitted output weights W_*^{out} are chosen as to minimise the mean square cost function

$$C(W^{\mathrm{out}}) = \sum_{n=-T_L}^{-1} \|\mathbf{y}(n) - W^{\mathrm{out}}\tilde{\mathbf{r}}(n)\|_2^2 + \lambda\|W^{\mathrm{out}}\|_F^2, \tag{8}$$

where $\mathbf{y}$ are the target outputs, which are part of the training data. T_L is the number of training time steps, $\|\cdot\|_F$ denotes the Frobenius norm. The last term penalises large values of the rows of the output weight matrix (and thus overfitting) by adjusting the regression parameter λ. The optimal output matrix W_*^{out} is thus directly obtainable by established methods. Fig. 4 summarises the architecture of the RCM in a sketch.

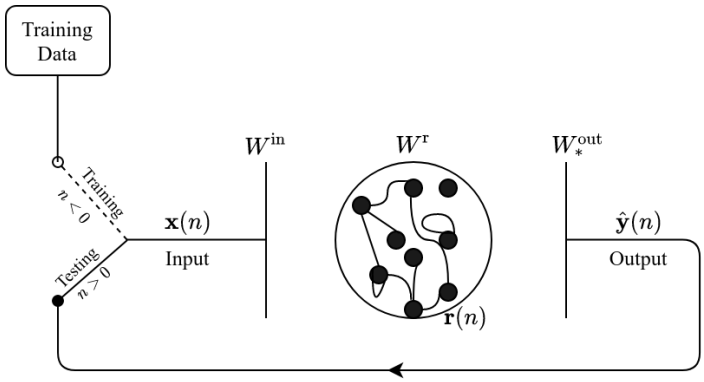


Figure 4. Sketch of the reservoir computing model architecture. In the training phase for times $n < 0$ the network learns the dynamics on training data.. In the testing phase ($n > 0$) it runs in a mode of autonomous prediction, where the last network output is fed back to the input layer to be used as new input for the next time step.

The inexpensive training procedure comes at a cost of finding a suitable set of hyperparameters, i.e. parameters which are not learned and have to be tuned beforehand. Here we restrict ourselves to $\mathrm{h} = \{\gamma, \lambda, N_{\mathrm{r}}, D, \varrho\}$. The last two quantities are the reservoir density D and spectral radius ϱ. They are algebraic properties of the reservoir weight matrix and represent the number of non-zero elements and largest absolute eigenvalue of W^{r}, respectively. Finding a right setting of these hyperparameters is crucial, as they influence the memory capacity of the reservoir. In most cases, we applied a grid search to obtain the best set and left D fixed to reduce the dimensionality of the hyperparameter space. We have applied RCM for different tasks in connection to turbulent convection:

- Dynamical evolution of the large-scale structure in a two-dimensional RBC flow at a Prandtl number $\mathrm{Pr} = 7$ and reproduction of the low-order temperature and convective heat flux statistics[13]. This flow is characterised by a top-down symmetry with respect to the horizontal midplane. The reduction of the original simulation data was done by a Proper Orthogonal Decomposition (POD), i.e., the time dependence of the POD expansion coefficients is learned by the RCM.
- Evolution of the large-scale cloud patterns in a two-dimensional moist Rayleigh-Bénard convection flow at a Prandtl number $\mathrm{Pr} = 0.7$ and reproduction of the liquid

water flux statistics and the cloud cover[14]. In the presence of phase changes from vapour to liquid water and *vice versa*, the top-down symmetry in the convection is broken is broken. This enhances the complexity of the machine learning task. The reduction of the original simulation data was done again by a POD snapshot method.

- Dynamical modelling of a two-dimensional flux-driven convective boundary layer (CBL) at a Prandtl number $\mathrm{Pr} = 0.7$ with highly asymmetric buoyancy flux profiles across the layer. In the focus of this work was the generalisation capability of RCMs[15]. The model was trained with one specific prescribed top plane flux. The test data comprised unseen simulation data at different prescribed fluxes which lead to a different flow behaviour. The RCM reproduced the statistics and flow structure fairly well. We also compared the reduction by means of POD with that of a convolutional autoencoder (CAE). The latter encodes the data in a low-dimensional latent space in which the RCM is trained and subsequently predicts dynamics. In Fig. 5, we display CBL reconstructions by the CAE and predictions by the joint CAE-RCM model.

- Direct data-driven forecast of the local convective heat flux of turbulent RBC by a combined CAE-RNN network. The local convective heat flux is a derived quantity that is composed as the product of the temperature and the vertical velocity component, $u_z T$. In this work, an RCM was compared with a gated recurrent unit, a further implementation of RNNs. The results of both architectures agreed fairly well with the ground truth[16].

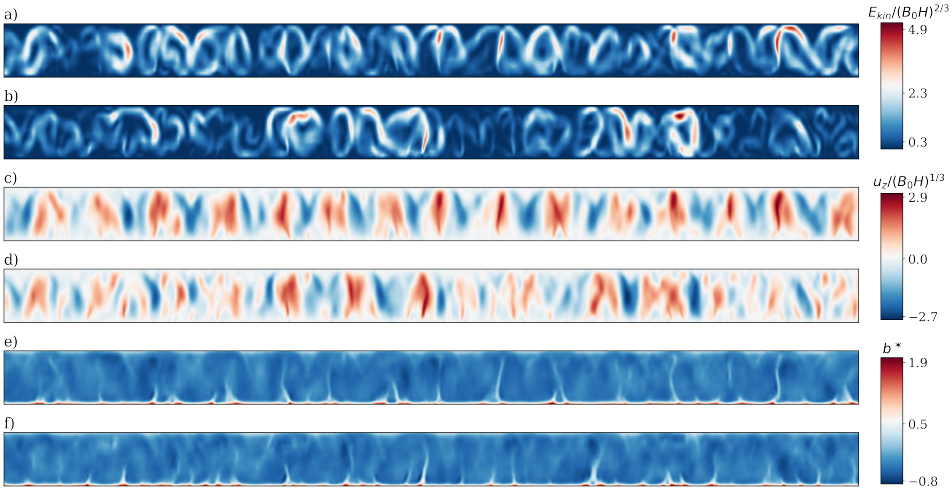


Figure 5. Prediction of the combined RCM-CAE network for unseen test data at a different prescribed flux at the top plane of the convective boundary layer (CBL). Instantaneous snapshots of the turbulent kinetic energy in panels (a,b), the vertical velocity component u_z in panels (c,d) and the normalised buoyancy b^* in panels (e,f) at discrete time step $n = 350$ are shown. CAE reconstructions of the data are shown in panels (a), (c), and (e). The corresponding RCM predictions are displayed in panels (b), (d), and (f). The aspect ratio of the CBL is 24.

As these results demonstrate, recurrent neural networks, specifically reservoir computing models, can serve as data-driven reduced models in turbulent convection flows. They

can reproduce the transport fluxes and their fluctuations across convection layers well. Such models might thus serve as dynamic parametrisations of mesoscale convection processes in global circulation models, such as climate models.

5 Concluding Remarks

Mesoscale convection appears at intermediate scales in global climate models of the Earth or in convection models of the Sun that study the solar activity. Here, we presented some results from our supercomputer simulations to understand these processes in simplified settings better. We also discussed methods to reduce and model the mesoscale turbulence and the related ordered flow patterns. This is done here by recurrent neural networks in combination with data reduction techniques, such as the Proper Orthogonal Decomposition and convolutional autoencoders.

Our numerical simulations show that the specific boundary conditions of the temperature field have a significant impact on the structure formation. Constant flux conditions lead to the formation of a supergranule structure that is superposed by a granule pattern. This aggregation seems to be connected to the behaviour at the onset of convection. Already a simple RBC setup can thus generate a hierarchy of structures, as present in natural and much more complex convection flows. Moderate additional rotation of the layer is able to control the size of the mesoscale convection patterns[17]. The current efforts focus on an extension of the incompressible Boussinesq case to the compressible non-Boussinesq case. In the latter system that top-down symmetry is immediately broken which brings our convection studies closer to the real behaviour in natural flows. The studies of compressible convection have just started and will be continued in the coming years.

Recurrent neural networks in the form of reservoir computing models can serve as reduced dynamical parametrisations of mesoscale convection. The next important step is to extend the application to three-dimensional flows. Furthermore, we plan to implement the multiscale character of turbulence into the reservoir network architecture which would be an alternative route to the recently developed physics-informed neural networks.

Acknowledgements

Our work is supported by the Deutsche Forschungsgemeinschaft (P. P. V. and J. D. S.), the Carl Zeiss Foundation (F. H.), and the Alexander von Humboldt Foundation (J. P. J.). We would also like to thank the John von Neumann Institute for Computing for the continued support of our research in turbulence with supercomputing resources and professional advice.

References

1. R. Markson, *Atmospheric electrical detection of organized convection*, Science **188**, 1171, 1975.
2. J. Schumacher and K. R. Sreenivasan, *Colloquium: Unusual dynamics of convection in the Sun*, Rev. Mod. Phys. **92**, 041001, 2020.

3. M. S. Emran and J. Schumacher, *Large-scale mean patterns in turbulent convection*, J. Fluid Mech. **776**, 96, 2015.
4. A. Pandey, J. D. Scheel, and J. Schumacher, *Turbulent superstructures in Rayleigh-Bénard convection*, Nat. Commun. **9**, 2118 (11 pages), 2018.
5. P. P. Vieweg, J. D. Scheel, and J. Schumacher, *Supergranule aggregation for constant heat flux-driven turbulent convection*, Phys. Rev. Res. **3**, 013231, 2021.
6. J. Verhoeven, T. Wiesehöfer, and S. Stellmach, *Anelastic versus fully compressible turbulent Rayleigh-Bénard convection*, Astrophys. J. **805**, 62, 2015.
7. S. K. Lele, *Compact finite differences with spectral like resolution*, J. Comp. Phys. **103**, 16, 1992.
8. J. Panickacheril John, D. A. Donzis, and K. R. Sreenivasan, *Does dissipative anomaly hold for compressible turbulence?* J. Fluid. Mech. **920**, A20, 2021.
9. `https://nek5000.mcs.anl.gov/`
10. F. M. White, *Viscous Fluid Flow* (McGraw Hill, New York, 2006).
11. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (The MIT Press, Cambridge, MA, 2017).
12. S. Pandey, J. Schumacher, and K. R. Sreenivasan, *A perspective on machine learning in turbulent flows*, J. Turbul. **21**, 567, 2020.
13. S. Pandey and J. Schumacher, *Reservoir computing model of two-dimensional turbulent convection*, Phys. Rev. Fluids **5**, 113506, 2020.
14. F. Heyder and J. Schumacher, *Echo state network for two-dimensional turbulent moist Rayleigh-Bénard convection*, Phys. Rev. E **103**, 053107, 2021.
15. F. Heyder, J. P. Mellado, and J. Schumacher, *Generalizability of reservoir computing for flux-driven two-dimensional convection*, 2022, arXiv:2108.06195v3.
16. S. Pandey, P. Teutsch, P. Mäder, and J. Schumacher, *Direct data-driven forecast of local turbulent heat flux in Rayleigh-Bénard convection*, Phys. Fluids **34**, 045106, 2022.
17. P. P. Vieweg, J. D. Scheel, R. Stepanov, and J. Schumacher, *Inverse cascades of kinetic energy and thermal variance in three-dimensional horizontally extended turbulent convection*, 2022, arXiv:2207.12606v1.

Identification of Very-Large-Scale Motions in Turbulent Channel Flows by Direct Numerical Simulations and Data-Driven Decomposition

Cheng Chi[1], **Alexander J. Smits**[2], **Steve Wolligandt**[3], **Holger Theisel**[3], **and Dominique Thévenin**[1]

[1] Laboratory of Fluid Dynamics and Technical Flows,
University of Magdeburg "Otto von Guericke", 39106, Magdeburg, Germany
E-mail: {cheng.chi, thevenin}@ovgu.de

[2] Department of Mechanical and Aerospace Engineering, Princeton University,
Princeton, NJ 08544, USA
E-mail: asmits@princeton.edu

[3] Faculty of Computer Science, University of Magdeburg "Otto von Guericke",
39106, Magdeburg, Germany
E-mail: {steve, theisel}@isg.cs.ovgu.de

To identify and characterise very-large-scale motions (VLSM) in turbulent channel flows, direct numerical simulations are performed at large Reynolds numbers, Re_τ = 906 and 3216, thanks to the impressive computing power provided by JSC. The turbulence structure is then analysed by multi-scale proper orthogonal decomposition (mPOD). The VLSM structures can be clearly visualised from the most energetic spatial modes of the mPOD analysis at both Reynolds numbers. Besides, the mPOD analysis also identify a new energetic mode with characteristic length covering the whole domain length. This new mode (called eVLSM mode) contains substantial energy and is straight along the streamwise direction, making it an important component of the VLSM in the flow field, with potential impact regarding all transport processes in the flow. The large-scale structures (apart from eVLSM) are inclined to the streamwise direction and appear to be responsible for the typical meandering behaviour or even for the breakup of VLSM.

1 Introduction

Very-large-scale motions (VLSM) are a class of organised structures that have been identified relatively recently in a variety of wall-bounded turbulent flows[1–9]. VLSM are found in the logarithmic and wake regions of wall turbulence, and they are characterised by very long (of the order of 10δ), meandering low- and high-speed regions. Here, δ is the thickness of the boundary layer, the radius of the pipe or the channel half-height. These structures significantly impact near-wall flow properties and turbulent energy distributions, and their influence becomes more pronounced at higher Reynolds number[10–12]. VLSM can contain up to 60% of the cumulative fraction of the Reynolds shear stress[3,4,13] and they can have a large influence on turbulence statistics[14,15].

Proper orthogonal decomposition (POD) analysis can be used to capture the most energetic contributions to the turbulent kinetic energy and to form a low-dimensional representation of the turbulent flow[13]. Therefore, POD is useful to extract the large-scale coherent motions in turbulent flows. Liu *et al.*[16] found that large-scale coherent motions in a turbulent channel flow can be reconstructed using only a few low-order POD modes. Wu *et al.*[17,18] further showed that for a turbulent boundary layer these low-order POD modes

essentially capture the large-scale signature of the hairpin packets. POD was also used by Hellström *et al.*[13] to analyse turbulent pipe flows. They found that the 10 most energetic modes encompassed all the principal characteristics of the VLSM.

While POD ensures that the most energetic contributions are captured in the lower order modes, these modes still contain contributions from all frequencies (i.e., all scales). To refine the analysis, multi-scale POD (mPOD) was proposed by Mendez *et al.*[19,20]. mPOD combines multi-resolution analysis (MRA) and standard POD to obtain the energetic POD modes within a certain range of frequencies, and it was originally designed for flows that demonstrate clear scale separations. Here, we use mPOD for the first time (to the best of the authors understanding) to examine the structure of turbulent channel flows, thereby allowing a more precise visualisation and analysis of VLSM, as marked by the identification of a new mode that is important for the understanding of VLSM.

Throughout the paper, the superscript $+$ refers (as usual) to quantities normalised by the friction velocity u_τ and the viscous wall unit $l^* = \nu/u_\tau$.

2 Numerical Database

Direct numerical simulations (DNS) were performed for fully developed turbulent flows in long channels using the incompressible version of the in-house hybrid finite-difference/spectral DNS solver DINO[21–23]. The code employs a 6th-order centred explicit scheme for spatial derivatives and a 3rd-order explicit Runge-Kutta scheme for temporal integration. The Poisson equation for pressure is solved with spectral accuracy.

We consider two cases, Case A with $Re_\tau = 906$ and Case B with $Re_\tau = 3216$, as summarised in Tab. 1. Here, $Re_\tau = hu_\tau/\nu$, where h is the half-width of the channel and ν is the kinematic viscosity of the fluid ($= 1.557 \times 10^{-5}$ m^2/s). We use u, v, and w to denote streamwise, wall-normal, and spanwise velocity fluctuations, respectively, corresponding to the x-, y-, and z-directions. No-slip conditions are applied to all channel walls. Case A has a domain of $20h \times 2h \times 4h$ with $h = 0.0125$ m. This domain is discretised by a grid measuring 1024×1025×256, which is uniform in the streamwise and spanwise directions and stretched in the wall-normal direction. The final spatial resolution is Δx^+ = 17.56, Δz^+ = 14.05, and $0.35 \leq \Delta y^+ \leq 8.40$. Case B has a domain measuring $20h \times 2h \times 2h$ with $h = 0.05$ m. This domain is discretised by a grid of 4096×2049×512, with a final spatial resolution of Δx^+ = 15.68, Δz^+ = 12.54, and $0.54 \leq \Delta y^+ \leq 11.22$. The resolutions of both cases are considered sufficient to capture the small scales, and the domains are long enough ($20h$) to correctly capture VLSM and LSM.

Case	Re_τ	Re_b	U_b (m/s)	u_τ (m/s)	l^* (μm)	η (μm)
A	906	38,432	23.9	1.12	13.8	27.6
B	3216	153,330	23.9	1.00	15.5	31

Table 1. Summary of the two DNS cases. Here, U_b is the bulk velocity, $Re_b = 2hU_b/\nu$ is the bulk Reynolds number, u_τ is the friction velocity, Re_τ is the friction Reynolds number, $l^* = \nu/u_\tau$ is the viscous length scale, and η is the Kolmogorov length scale near the wall.

Fig. 1 shows a single realisation of the full channel flow for Case B coloured by streamwise velocity on an iso-surface of Q-criterion. The simulations were performed on the

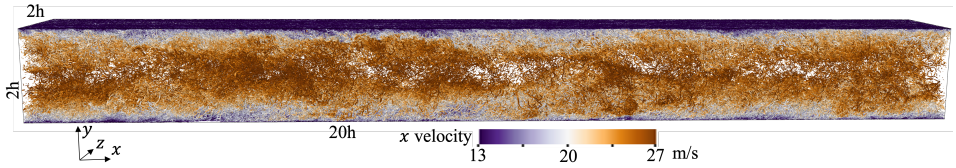


Figure 1. Iso-surface of Q-criterion with $Q = 5 \times 10^6$ s^{-2} for a channel flow at Re_τ = 3216 (Case B) at $t = 3.47t_F$ (where $t_F = 20h/U_b$ is the flow-through time), coloured by streamwise (x) velocity.

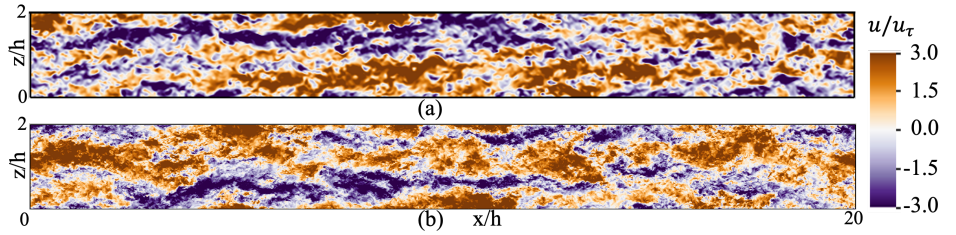


Figure 2. Iso-surfaces of velocity at $y/h = 0.19$ for (a) Re_τ = 906 (Case A); (b) Re_τ = 3216 (Case B). For reference, $U^+ \approx 18.5$ for Case A, and ≈ 20 for Case B.

supercomputer JUWELS at the Jülich Supercomputing Centre (JSC) using 8192 processors. In total, 8.4 million CPU hours were necessary for these two simulations, producing 3.4 TB of raw data for Case A and 45 TB of raw data for Case B.

To give an overview of the large-scale structures in our test cases, Fig. 2 shows the iso-surfaces of streamwise velocity fluctuations u (scaled by the friction velocity u_τ) at a wall distance $y/h = 0.19$ for Case A and Case B. To achieve a fair comparison with Case B, only $2h$ spanwise wide domain (from the full $4h$ wide domain) has been chosen for all the analysis in Case A. The VLSM are evident in Case B by the low speed regions separated by high speed regions with streamwise lengths of the order of $10h$. These low and high speed regions can be also seen in Case A, with a similar spanwise spacing but possibly a shorter streamwise length than in Case B. The maximum velocity difference between low and high-speed regions is typically greater than $\pm 15\%$.

3 Analysis and Discussion

In the following sections, the overline $\overline{q}$ means the average of the quantities over the x-z plane, and the capital letter U indicates the temporally and spatially (over the x-z plane) averaged statistics.

3.1 mPOD Analysis

mPOD were performed on 2D slices of the 3D DNS datasets $\vec{\mathbf{u}}(\mathbf{x}, t_k) = (\mathbf{u}(\mathbf{x}, t_k), \mathbf{v}(\mathbf{x}, t_k), \mathbf{w}(\mathbf{x}, t_k))$. Here, $\mathbf{x}$ indicates all the grid points in the 2D slices, and

t_k indicates different time instants. The analysis is performed in a plane which is parallel to the wall, resulting in a uniform grid for **x** (the grid is only stretched in wall-normal direction). The distribution of t_k is also uniform along the temporal axis, with $t_{k+1} - t_k = \Delta t$, where $\Delta t = 1/f_s$ is the constant time step and f_s is the sampling frequency.

mPOD is computed based on POD by adding spectral constraints to the temporal structures of the classical POD. In classical POD the full spatial-temporal data $\vec{\mathbf{u}}(\mathbf{x}, t_k)$ is decomposed into a linear combination of spatial structures $\vec{\boldsymbol{\phi}}_n(\mathbf{x})$ and temporal structures $\psi_n(t_k)$, where n is the mode number, so that

$$\vec{\mathbf{u}}(\mathbf{x}, t_k) = \sum_{n=1}^{N} \delta_n \vec{\boldsymbol{\phi}}_n(\mathbf{x}) \psi_n(t_k), \tag{1}$$

where δ_n is the corresponding mode amplitude. The rank $N = \min(n_s, n_t)$ is the full rank, with $n_s = n_p \times n_u$, where n_p is the number of grid points, n_u is the number of velocity components, and n_t is the number of snapshots. In mPOD, the temporal structures are split into different scales before the orthogonal decomposition is done, so that the most energetic modes are more representative without scale mixing. Detailed information concerning mPOD is given by Mendez *et al.*[19].

3.2 Scale Separation Cutoffs

A spectral cut-off needs to be defined to separate different scales, and a number of levels have been proposed in the literature. For example, Mathis *et al.*[24] used the cutoff $\lambda_x/\delta = 1$ in the near-wall layer of the turbulent boundary layer, while Guala *et al.*[3] used $\lambda_x/\delta = \pi$ in the outer layer of pipe flow where Bailey *et al.*[7] instead used $\lambda_x/\delta = 2\pi$. λ_x is the streamwise wavelength. Based on the slope of the spectrum curve, the cutoff $\lambda_x/\delta = 5$ was proposed by Önder and Meyers[25] in a turbulent boundary layer, whereas for turbulent channel flow Balakumar *et al.*[4] used $\lambda_x/\delta = \pi$. In the present study, a cutoff at $\lambda_x/h = 1.65$ (corresponding to λ_x^+ = 1495 for Case A and λ_x^+ = 5306 for case B) was adopted to distinguish contributions due to LSM from those due to VLSM. Also a cutoff at $\lambda_x/h = 0.88$ was adopted to distinguish contributions due to LSM from those due to small scale motions (see Tab. 2). The choice of these two cutoffs are obtained by the preliminary study on the pre-multiplied spectra plot (not shown here). Tab. 2 summarises the splitting frequencies for Case A and Case B.

Case	Re_τ	$F_V^+ \times 10^3$	$F_V h/U_b$	λ_x^+	λ_x/h
A	906	6.23	0.264	2970	3.3
		12.46	0.528	1485	1.65
		23.36	0.986	792	0.88
B	3216	1.88	0.252	10613	3.3
		7.07	0.948	2830	0.88

Table 2. Splitting frequencies for Cases A and B.

3.3 Results for Case A

For the lower Reynolds number Case A ($Re_\tau = 906$), 254 snapshots were used for the analysis with a sampling frequency $f_s^+ = f_s\nu/u_\tau^2 = 0.1559$ and $f_s h/U_b = 6.569$. The 2D slice at wall distance $y/h = 0.19$ was selected since it is located in the outer region, nominally near the end of the logarithmic region, although at these Reynolds numbers the logarithmic region is not yet fully formed[26,27]. Here, the sampling wavelength is $\lambda_x^+ = U^+/f_s^+$ = 118.78. At this wall distance, the spectral contributions from different scales are comparable. For mPOD, three splitting frequencies were chosen as $F_V^+ = [6.23, 12.46, 23.36] \times 10^{-3}$, which correspond to $F_V h/U_b = [0.264, 0.528, 0.986]$, $\lambda_x^+ = [2970, 1485, 792]$, and $\lambda_x/h = [3.3, 1.65, 0.88]$. See Tab. 2. The cutoffs at $\lambda_x/h = 0.88$ and 1.65 separate VLSM from LSM, and LSM from small scale motions, but a 3rd cutoff at $\lambda_x/h = 3.3$ is also used based on the full temporal mode (not shown here) that suggest the presence of structures even longer than VLSM.

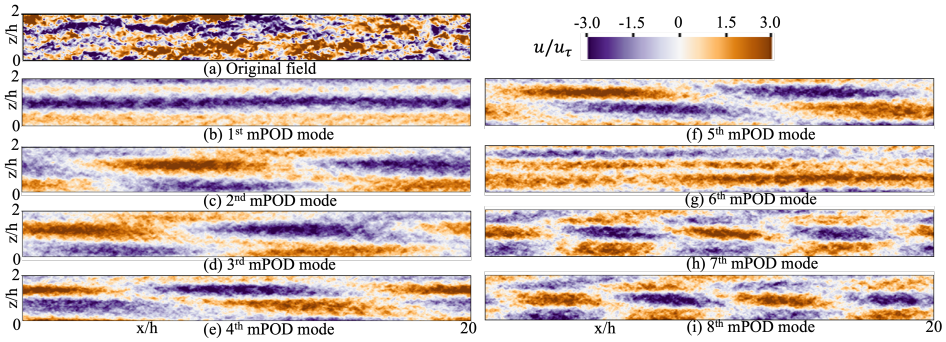


Figure 3. Spatial modes for Case A at $y/h = 0.19$. (a) Original streamwise velocity fluctuation flow field; (b) 1st mPOD mode; (c) 2nd mPOD mode; (d) 3rd mPOD mode, (e) 4th mPOD mode; (f) 5th mPOD mode; (g) 6th mPOD mode; (h) 7th mPOD mode; (i) 8th mPOD mode. Note that the 1st and 6th mPOD have $\lambda_x/h \geq 3.3$; the 2nd, 3rd, 4th and 5th mPOD have $1.65 \leq \lambda_x/h < 3.3$; the 7th and 8th mPOD have $0.88 \leq \lambda_x/h < 1.65$.

The spatial modes of mPOD are given in Fig. 3. It can be seen that some modes come in pairs, as in the 2nd and 3rd Modes, the 4th and 5th Modes, and the 7th and 8th Modes. The spatial arrangement of the pairs is consistent with the evolutionary cycle of large scale motion described by Hellström it et al.[28], or possibly the presence of helical modes. Long and straight coherent structures appear in the 1st mPOD mode with a length that covers the full domain, much longer than that typically associated with VLSM. A similar behaviour is seen in the 6th mode, albeit with a reduced spanwise spacing. The temporal filtering inherent in mPOD suggests that these streamwise invariant spatial modes, not previously identified in channel flows, are an important part of the most energetic motions. In what follows, we will call these streamwise invariant spatial modes "extended VLSM" (eVLSM).

It is important to note that we have labelled the coherent structures in the decomposed modes as VLSM, LSM and small-scale motions using their typical streamwise lengths. These structures, however, are component contributions to the full velocity field. For ex-

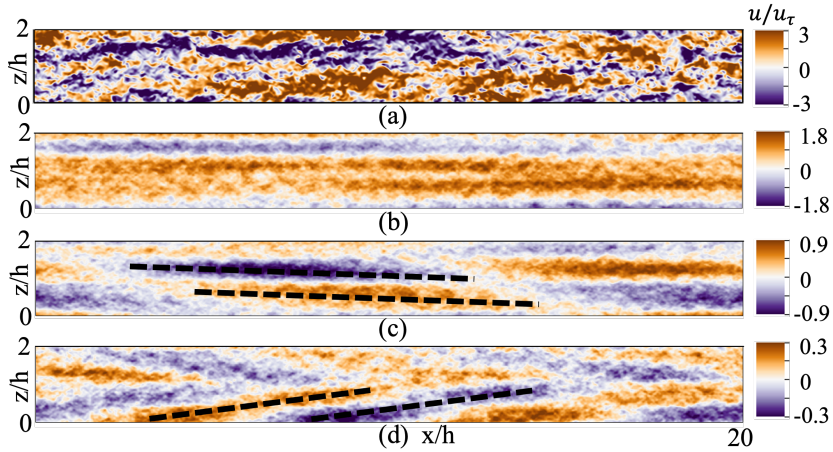


Figure 4. Reconstructed streamwise velocity fluctuation flow field for Case A at $y/h = 0.19$. (a) Original velocity field; (b) reconstructed field using the 4 most energetic mPOD modes with $\hat{f}_s^+ \leq 6.23\times10^{-3}$ ($\lambda_x/h \geq 3.3$); (c) using the 4 most energetic mPOD modes with $6.23 \times 10^{-3} < \hat{f}_s^+ \leq 12.46 \times 10^{-3}$ ($1.65 \leq \lambda_x/h < 3.3$); (d) using the 4 most energetic mPOD modes with $12.46 \times 10^{-3} < \hat{f}_s^+ \leq 23.36 \times 10^{-3}$ ($0.88 \leq \lambda_x/h < 1.65$).

ample, the VLSM and LSM seen in the full velocity field are the sum of the decomposed modes. Therefore, identification of the new mode eVLSM does not necessarily identify a new coherent structure in the original flow field.

As to the characteristic lengths of the other modes, the 7th and 8th mPOD modes have lengths in the range 2-4h, which are typical for LSM, while for the 2nd, 3rd, 4th and 5th mPOD mode they are 8-10h, which are in the range expected for VLSM. This result confirms that the cutoff at $\lambda_x/h = 1.65$ is a good choice for distinguishing between VLSM and LSM.

The reconstructed flow field using the most energetic modes within different scales are shown in Fig. 4. eVLSM (typical length covering the domain length 20h) can be clearly identified from the reconstruction using the 4 most energetic mPOD modes with $\hat{f}_s^+ \leq 6.23\times10^{-3}$ ($\lambda_x/h \geq 3.3$). VLSM with a typical length of 8-10h appear in the reconstruction using the 4 most energetic mPOD mode with $6.23\times10^{-3} < \hat{f}_s^+ \leq 12.46\times10^{-3}$ ($1.65 \leq \lambda_x/h < 3.3$). LSM with typical length of 2-4h can be identified from the reconstruction using the 4 most energetic mPOD mode with $12.46\times10^{-3} < \hat{f}_s^+ \leq 23.36\times10^{-3}$ ($0.88 \leq \lambda_/h < 1.65$). Hence, with the help of mPOD the coherent structures at different scales are clearly separated. Furthermore, it can be observed that the reconstructed structures in Fig. 4(c) are slightly inclined to the streamwise direction, with an angle of approximately 5°, which may be related to the helix angle of 5° seen for the most energetic POD mode in pipe flow[29]. The structures in Fig. 4(d) are inclined with an angle of approximately 10°. The meandering of the VLSM observed in the original flow field is clearly identified with the spanwise tilt of the LSM, as seen in Fig. 4(d).

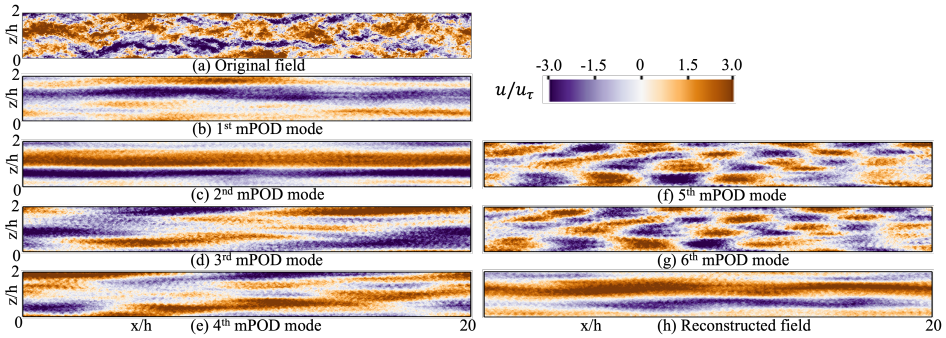


Figure 5. Spatial modes for Case B at $y/h = 0.19$. (a) Original streamwise velocity fluctuation flow field; (b) spatial structures of the 1st mPOD mode; (c) the 2nd mPOD mode; (d) the 3rd mPOD mode; (e) the 4th mPOD mode; (f) the 5th mPOD mode; (g) the 6th mPOD mode (h) reconstructed streamwise velocity fluctuation flow field using the first 4 mPOD modes. Note that the 1st and 2nd mPOD have $\lambda_x/h \geq 3.3$; the 3rd and 4th mPOD have $0.88 \leq \lambda_x/h < 3.3$, and the 5th and 6th mPOD have $\lambda_x/h < 0.88$.

3.4 Results for Case B

For the higher Reynolds number Case B ($Re_\tau = 3216$), 103 snapshots were used with $f_s^+ = f_s\nu/u_\tau^2 = 0.0489$ and $f_s h/U_b = 6.56$. As in Case A, the 2D slice at a wall distance y/h = 0.19 is considered in the analysis. Here, the sampling wavelength is $\lambda_x^+ = U^+/f_s^+$ = 409. In mPOD, the splitting frequency is set to $F_V^+ = [1.88, 7.07] \times 10^{-3}$, which corresponds to the scale separation wavelength at $\lambda_x^+ = [10613, 2830]$ and $\lambda_x/h = [3.3, 0.88]$, as shown in Tab. 2. Due to the limited number of snapshots for this case, the cutoff at $\lambda_x/h = 1.65$ used in Case A could not be applied, so that our analysis for Case B is restricted to examining the eVLSM and the combined mode of VLSM and LSM.

The first 6 spatial modes for mPOD are shown in Fig. 5, together with the reconstructed velocity fields using the first 4 modes. The 1st and 2nd modes make a pair, and show relatively straight and very long coherent structures (covering the domain length). Similar structures were seen in Case A and dubbed eVLSM, although in that case the first mode did not have a pair (this is discussed further below). The 3rd and 4th mPOD modes make up another pair, with a typical length of 8-12h (indicative of VLSM), and a slight inclination to the streamwise direction. The next pair is made up of the 5th and 6th modes, where the mPOD modes show distinct structures with a typical length less than $2h$ (indicative of small scale motions).

As noted earlier, in Case A the eVSLM mode is the 1st mPOD mode while in Case B it is a pair made up of the 1st and 2nd mPOD modes. Although it is not clear why this is so, one possibility is that this eVLSM mode evolves with Reynolds number. For example, the combined energy content of the 1st and 2nd mPOD modes in Case B is approximately 0.08 of the total energy, whereas the 1st mPOD mode in Case A contains only 0.05 of the total energy. Another possibility is the limited spanwise width of the domain, where periodic boundary conditions are enforced. In Case A, that width is $4h$ (only $2h$ is shown in Fig. 3) but in Case B it is only $2h$. Comparing Figs. 3(b) and 5(b-c), it is likely that the smaller domain width imposes constraints on the spanwise length scales which could lead to changes in the inferred modes.

The inclined structures in the 3rd and 4th mPOD modes have a typical angle of 5° to the streamwise direction. This is consistent with Fig. 4(c) in Case A for VLSM. Comparing the original flow field (especially for the low-speed streaks) with the inclined structures in the 3rd and 4th mPOD modes, similar tilt angles are observed, and this inclination of the modes could be the main reason for meandering features of the VLSM. The reconstruction using the first 4 mPOD modes combines eVLSM and the inclined VLSM. As is seen from Fig. 5(h), the inclined VLSM terminate the very long structure in the streamwise direction (see the low-speed streak) and the reconstructed field is dominated by eVLSM, which contain much more energy than the smaller scale motions.

4 Concluding Remarks

DNS for turbulent channel flows at high Reynolds numbers of Re_τ = 906 and 3216 was conducted using the computing power of JUWELS for the visualisation and analysis of VLSM. Multi-scale POD (mPOD) was found to help separate different scales and to provide a more detailed analysis of the individual contribution from different scales. In particular, a new energetic mode (eVLSM) was identified by mPOD, which is aligned in the streamwise direction with a characteristic length covering the full $20h$ domain length. This eVLSM mode was found to be the most energetic mode at both Reynolds numbers, and therefore crucial for the VLSM representation in the full flow field. Large scale structures (apart from eVLSM) inclined to the streamwise direction are consistent with the meandering features of VLSM in the original flow field. It is expected that VLSM and eVLSM play a very prominent role to explain transport processes in turbulent flows.

Acknowledgements

This work is supported by the Priority Programme SPP 1881 Turbulent Superstructures of the Deutsche Forschungsgemeinschaft (DFG). The DNS simulations were made possible thanks to the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`), providing computing time on the Supercomputer JUWELS (under project "dnschannelfspp") at Jülich Supercomputing Centre (JSC, Germany). AJS acknowledges support under ONR Grants N00014-17-1-2309 and N00014-19-1-2301 (Program Manager Dr. Peter Chang).

References

1. K. C. Kim and R. J. Adrian, *Very large-scale motion in the outer layer*, Phys. Fluids, **11**, no. 2, 417–422, 1999.
2. C. D. Tomkins and R. J. Adrian, *Energetic spanwise modes in the logarithmic layer of a turbulent boundary layer*, J. Fluid Mech., **545**, 141–162, 2005.
3. M. Guala, S. E. Hommema, and R. J. Adrian, *Large-scale and very-large-scale motions in turbulent pipe flow*, J. Fluid Mech., **554**, 521–542, 2006.
4. B. J. Balakumar and R. J. Adrian, *Large- and very-large-scale motions in channel and boundary-layer flows*, Philos. Trans. Royal Soc. A, **365**, no. 1852, 665–681, 2007.
5. N. Hutchins and I. Marusic, *Evidence of very long meandering features in the logarithmic region of turbulent boundary layers*, J. Fluid Mech., **579**, 1–28, 2007.

6. J. P. Monty, J. Stewart, R. Williams, and M. S. Chong, *Large-scale features in turbulent pipe and channel flows*, J. Fluid Mech., **589**, 147–156, 10 2007.
7. S. C. C. Bailey and A. J. Smits, *Experimental investigation of the structure of large- and very-large-scale motions in turbulent pipe flow*, J. Fluid Mech., **651**, 339–356, 2010.
8. J. P. Monty, N. Hutchins, H. C. H. Ng, I. Marusic, and M. S. Chong, *A comparison of turbulent pipe, channel and boundary layer flows*, J. Fluid Mech., **632**, 431–442, 2009.
9. J. H. Lee and H. J. Sung, *Very-large-scale motions in a turbulent boundary layer*, J. Fluid Mech., **673**, 80–120, 2011.
10. I. Marusic, R. Mathis, and N. Hutchins, *High Reynolds number effects in wall turbulence*, Int. J. Heat Fluid Flow, **31**, no. 3, 418–428, 2010.
11. A. J. Smits, B. J. McKeon, and I. Marusic, *High–Reynolds Number Wall Turbulence*, Annu. Rev. Fluid Mech., **43**, no. 1, 353–375, 2011.
12. M. Hultmark, M. Vallikivi, S. C. C. Bailey, and A. J. Smits, *Turbulent Pipe Flow at Extreme Reynolds Numbers*, Phys. Rev. Lett., **108**, 094501, Feb 2012.
13. L. H. O. Hellström, A. Sinha, and A. J. Smits, *Visualizing the very-large-scale motions in turbulent pipe flow*, Phys. Fluids, **23**, no. 1, 011703, 2011.
14. B. Ganapathisubramani, E. K. Longmire, and I. Marusic, *Characteristics of vortex packets in turbulent boundary layers*, J. Fluid Mech., **478**, 35–46, 2003.
15. A. J. Smits, M. Hultmark, M. Lee, S. Pirozzoli, and X. Wu, *Reynolds stress scaling in the near-wall region of wall-bounded flows*, J. Fluid Mech., **926**, A31, 2021.
16. Z. Liu, R. J. Adrian, and T. J. Hanratty, *Large-scale modes of turbulent channel flow: transport and structure*, J. Fluid Mech., **448**, 53–80, 2001.
17. Y. Wu and K. T. Christensen, *Spatial structure of a turbulent boundary layer with irregular surface roughness*, J. Fluid Mech., **655**, 380–418, 2010.
18. Y. Wu, *A study of energetic large-scale structures in turbulent boundary layer*, Phys. Fluids, **26**, no. 4, 045113, 2014.
19. M. Mendez, M. Balabane, and J.-M. Buchlin, *Multi-scale proper orthogonal decomposition of complex fluid flows*, J. Fluid Mech., **870**, 988–1036, 07 2019.
20. M. A. Mendez, D. Hess, B. B. Watz, and J.-M. Buchlin, *Multiscale proper orthogonal decomposition (mPOD) of TR-PIV data—a case study on stationary and transient cylinder wake flows*, Meas. Sci. Technol, **31**, no. 9, 094014, jun 2020.
21. A. Abdelsamie, G. Fru, T. Oster, F. Dietzsch, G. Janiga, and D. Thévenin, *Towards direct numerical simulations of low-Mach number turbulent reacting and two-phase flows using immersed boundaries*, Comput. Fluids, **131**, 123–141, 2016.
22. C. Chi, A. Abdelsamie, and D. Thévenin, *A directional ghost-cell immersed boundary method for incompressible flows*, J. Comput. Phys., **404**, 109122, 2020.
23. A. Abdelsamie, G. Lartigue, C. E. Frouzakis, and D. Thévenin, *The Taylor–Green vortex as a benchmark for high-fidelity combustion simulations using low-Mach solvers*, Comput. Fluids, **223**, 104935, 2021.
24. R. Mathis, N. Hutchins, and I. Marusic, *Large-scale amplitude modulation of the small-scale structures in turbulent boundary layers*, J. Fluid Mech., **628**, 311–337, 2009.
25. A. Önder and J. Meyers, *On the interaction of very-large-scale motions in a neutral atmospheric boundary layer with a row of wind turbines*, J. Fluid Mech., **841**, 1040–1072, 2018.

26. B. J. McKeon, J. Li, W. Jiang, J. F. Morrison, and A. J. Smits, *Further observations on the mean velocity distribution in fully developed pipe flow*, J. Fluid Mech., **501**, 135–147, 2004.
27. A. J. Smits, *Batchelor Prize Lecture: Measurements in wall-bounded turbulence*, J. Fluid Mech., **940**, A1, 2022.
28. L. H. O. Hellström, B. Ganapathisubramani, and A. J. Smits, *The evolution of large-scale motions in turbulent pipe flow*, J. Fluid Mech., **779**, 701–715, 2015.
29. J. R. Baltzer, R. J. Adrian, and X. Wu, *Structural organization of large and very large scales in turbulent pipe flow simulation*, J. Fluid Mech., **720**, 236–279, 2013.

Plasma Physics

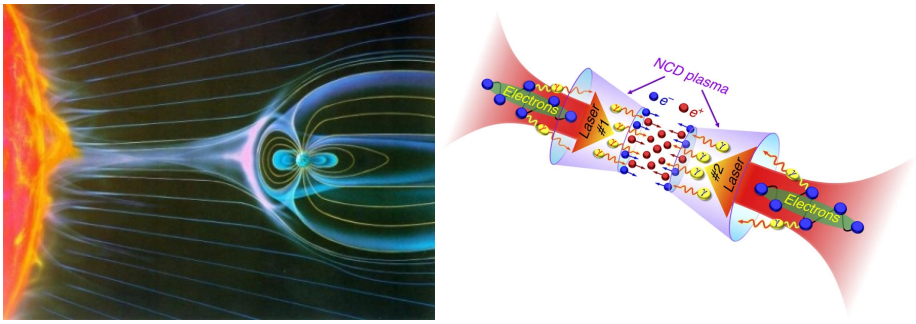
Plasma Models Covering Wide Ranges of Temporal and Spatial Scales

Karl-Heinz Spatschek

Institut für Theoretische Physik, Heinrich-Heine-Universität Düsseldorf,
40225 Düsseldorf, Germany
E-mail: khs@tp1.uni-duesseldorf.de

Two plasma scenarios are discussed in the following. They have quite different applications. The first paper belongs to extraterrestrial physics. It contributes to a better understanding of the interplay between the Earth's magnetic dynamo and the solar wind. Spatial dimension is up to one astronomical unit (AU) and times vary between fractions of seconds and hours or days. The second paper looks quite far into the future of high-intensity, short-pulse laser physics applications. Within the next decade(s) the study of QED (quantum electrodynamic) processes with high-intensity ($\geq 10^{23}$ W/cm^2) lasers seems to be realistic and achievable.

Characteristic plasma parameters vary over a wide range of magnitudes, e.g. density from 10^6 to 10^{34} m^{-3}, velocities from almost standing to relativistic velocities, and time scales from days ($\sim 10^5$ s) to femtoseconds (10^{-15} s). Even when considering quite specific physical phenomena, in many cases variations over many orders of magnitude have to be taken into account. For reliable results, different model zones are then combined, e.g. plasmadynamical equations for the slow and near-homogeneous overall behaviour together with calculations based on kinetic (partial) differential equations for the near-linear fast and very inhomogeneous substructures. The PIC (particle in cell) method, on the other hand, is the preferred kinetic approach for analysing strong-field plasma physics. The following two articles show in impressive manner how the most appropriate numerical tools can be further improved. Fig. 1 briefly visualises the physical problems under consideration.



(a) (b)

Figure 1. Sketches of the two scenarios considered in the present plasma physics HPC section: (a) Earth magnetosphere in the presence of the solar wind *[Courtesy of NASA]*. (b) Reactions triggered by two counter-propagating ultraintense laser pulses[1].

(a) Generation of Earth's Magnetic Field and Interaction with the Solar Wind[2]

A family of models that couple multifluid systems to the full Maxwell equations have been used in laboratory, space, and astrophysical plasma modelling. These models are more complete descriptions of the plasma than reduced models like magnetohydrodynamic (MHD) since they are derived more closely from the full kinetic Vlasov-Maxwell system, without assumptions like quasi-neutrality, negligible electron mass, etc.[3]

An experiment is being planned at HZDR (Helmholtz Zentrum Dresden) to better understand planetary magnetic fields powered by precession-driven fluid flows. The first part of the paper shows numerical simulations which might contribute to an optimal design of the device. It estimates relevant operating parameters. In the second part it is discussed how magnetic reconnection affects the so called space weather. Improvements in the multiphysics plasma simulation codes enabled numerical simulations with significant reduction in computational resources.

(b) Study of QED Processes with (Future) High-Intensity Lasers[4]

State-of-art laser systems are capable of delivering a laser pulse with intensity up to 10^{22} W/cm^2. The next-generation multi-PW lasers [for example, the XCELS (Exawatt Center for Extreme Light Studies) and ELI (Extreme Light Infrastructure) facilities] are expected to reach 10^{24} W/cm^2 and beyond[5]. This opens the door for studying light-matter interactions as well as QED effects in unexplored domains[1].

The quivering electrons in the ultraintense laser fields experience large radiation-reaction forces by emitting photons so that a large number of electrons are trapped in the laser fields. These trapped electrons perform extreme oscillations in the transverse direction and emit bright γ rays (red- and blue-yellow) around the laser axis. Finally, copious numbers of e^-e^+ pairs are created via the multi-photon Breit-Wheeler process[1]. The paper opens new perspectives for future studies in the regime of non-perturbative QED.

References

1. X.-L. Zhu, T.-P. Yu, Z.-M. Sheng, Y. Yin, I. C. E. Turcu, and A. Pukhov, *Dense GeV electron-positron pairs generated by lasers in near-critical-density plasmas*, Nature Commun., **7**, 13686, 2016.
2. M. Wilbert et al., *Recent Advancements in HPC Simulations of Dynamos and Multiphysics Plasmas*, this volume.
3. L. Wang, A. Hakim, J. Ng, C. Dong, and K. Germaschewski, *Exact and locally implicit source term solvers for multifluid-Maxwell systems*, J. Comp. Phys., **415**, 109510, 2020.
4. L. Reichwein et al., *Interaction of Extremely Intense Flows of Electromagnetic Energy and QED Processes in Supercritical Fields*, this volume.
5. G. Mourou and T. Tajima, *More intense, shorter pulses*, Science, **331**, 41, 2011.

Recent Advancements in HPC Simulations of Dynamos and Multiphysics Plasmas

Mike Wilbert[1], **Florian Allmann-Rahn**[1], **André Giesecke**[2], **and Rainer Grauer**[1]

[1] Ruhr University Bochum, Institute for Theoretical Physics I,
Universitätsstraße 150, 44801 Bochum, Germany
E-mail: grauer@tp1.rub.de

[2] Institut fur Fluiddynamik, Helmholtz-Zentrum Dresden-Rossendorf (HZDR),
Bautzner Landstraße 400, 01328 Dresden, Germany
E-mail: a.giesecke@hzdr.de

For a better understanding of the generation of the Earth's magnetic field and its interaction with the solar wind, numerical simulations are essential. The multiscale nature of these phenomena makes such simulations very demanding concerning computational resources. Here we present two simulation codes designed for large HPC systems: The code SpecDyn computes hydrodynamic flows in precessing vessels in order to investigate velocity fields suitable for dynamo action – a process that is expected to generate the Earth's magnetic field. The *muphy*II code is utilised to couple different plasma models in order to efficiently model reconnection and other large scale processes in the Earth's magnetosphere. Recent advancements in the codes and the latest results are discussed.

1 Introduction

Precession, i.e. the simultaneous rotation around two axes, provides an efficient natural mechanism for powering vigorous fluid flows in rotating vessels[1]. Beside their relevance for technical applications, e.g. mixing in bioreactors or the stability of rotating flying objects with liquid payloads, precession driven flows have also been proposed as a power source for the geodynamo or the ancient lunar dynamo[2].

Motivated by the question of planetary magnetic fields powered by precession driven fluid flows, an experiment is being designed and constructed at *HZDR* which consists of a cylinder that will be filled with liquid sodium and simultaneously rotates around two axes[3]. The experiment is part of the project *DRESDYN* (DREsden Sodium facility for DYNamo and thermohydraulic studies), which is a new platform for a variety of liquid sodium experiments devoted to problems of geo- and astrophysical magnetohydrodynamics[4].

For an optimal design of the device and the estimation of the relevant operating parameters, numerical simulations are inevitably required beforehand. These are performed by the code *SpecDyn* which combines the speed of pseudo-spectral methods with the ability to treat arbitrarily shaped boundaries by employing a direct forcing immersed boundary method (IBM). The latter is crucial since the precessing vessel consists of a cylinder with embedded baffles at the end caps. To obtain the necessary resolution required by the boundary layer the code is massively parallelised on CPUs on a two-dimensional processor grid. In this paper we present a brief discussion of the numerical methods and the parallelisation of the code *SpecDyn* as well as results obtained from hydrodynamic simulations. All computations presented in this paper by the code *SpecDyn* were performed on the JUWELS CPU cluster.

Plasma particles emitted from the sun form the so called solar wind and interact with the Earth and its magnetic field. The conditions in the solar wind – the space weather – vary with the activity of the sun, in particular the occurrence of coronal mass ejections and solar flares. An accurate prediction of space weather can in the future help to better prevent catastrophic damage to satellites or spacecraft and avoid radiation exposure of astronauts and persons in high altitude flights.

Space weather is affected by magnetic reconnection where magnetic field lines reconnect towards an energetically more advantageous state. The accurate modelling of magnetic reconnection typically requires a kinetic description where the velocity distribution of particles is taken into account. Due to their high computational cost, it is not possible to treat very large three-dimensional systems with fully kinetic models. An approach to overcome this problem is to use kinetic models only in those regions where kinetic effects are important and treat the rest of the domain with reduced multi-fluid and MHD models. This spatial multiphysics coupling of different plasma models is a matter of current research. Recent progress using a hierarchy of kinetic Vlasov, hybrid and multi-fluid models with the *muphy*II code on JUWELS Booster is discussed in this paper.

2 The Code SpecDyn: Incompressible Fluid Simulations in Precessing Vessels

The code *SpecDyn* currently treats the incompressible Navier-Stokes equations with no-slip boundary conditions. The basic numerical scheme is the pseudo-spectral method in which the considered equations are treated in Fourier space to compute derivatives and the presentation in real space is used for calculations of multiplications of space dependent fields, e.g. the non-linearity in the NS equation. The time stepping is performed by the strong stability preserving Runge-Kutta method of convergence order 3 (SSPRK3)[5]. This method is known for its fast performance, especially due to the efficiency of the Fast Fourier Transform (FFT)[6] and since the pressure-Poisson equation can be inverted directly in Fourier space[7].

The main drawback of pseudo-spectral Fourier methods is that they can only treat periodic boundary conditions. To overcome this issue we utilised a direct forcing *Immerse Boundary Method* (IBM). The IBM is based on a method proposed by Peskin[8] in the context of fluid flows with elastic boundaries. Here the fluid and the boundary act a feedback force on each other and the fluid is simulated on an Eulerian whereas the boundaries are described on a Lagrangian grid. Mohd-Yusof[9] introduced this idea in the context of fluid flows with solid boundaries. He suggested to treat the boundary by a force which ensures the no-slip boundary condition, so that the computation of the fluid can be carried out on simple grids. In contrast to other IB methods for rigid boundaries (e.g. Goldstein *et al.*[10]) the boundary condition is explicitly set and is not dependent on the flow itself. If the location of the boundary coincides with the Eulerian grid the boundary condition can simply be set. In the more likely case that this in not the case some kind of interpolation must be performed. We use a linear interpolation along the coordinate axes of a Cartesian grid as suggested by Fadlun *et al.*[11]. The boundary is described by a triangular mesh. To determine the solid and the fluid region a ray-tracing algorithm[12] is used and all the fluid points with at least one neighbour located in the solid are marked as IB nodes, whose velocity is interpolate between the closest fluid point and the boundary in each time step.

The only global operation in our code is the FFT. There exist several C++/FORTRAN libraries to perform parallel FFTs[13–15], which are generally all based on the same method. The simulation domain is distributed on a processor grid in one (slab decomposition) or two dimensions (pencil decomposition). To be able to utilise a greater amounts of cores, we use the slab decomposition. Here one dimension is local on each process over which a 1D FFT is then to be performed. To treat the other two dimensions the field of interest is transposed on the processor grid, so that a parallel FFT can be achieved by three 1D FFTs and two processor grid transformations.

However, it is mandatory for our purposes to transform the velocity field between different rotating frames of reference. For this purpose we use the method of Larkin *et al.*[16], who make use of the Fourier shift theorem and achieves the rotation of the field by means of three shear transformations. For this method it is necessary to be able to perform one dimensional FFTs in arbitrary directions, which is not, to the best of the authors knowledge, supported by the available parallel FFT libraries. Therefore we implemented it ourselves following the suggestions of Mortensen *et al.*[17]. They utilise the Message Passing Interface[18] and make use of *MPI datatypes* as well as the advanced MPI API routine *MPI-all2allw* to perform the processor grid transforms. By this approach a highly scalable and easy to implement parallel FFT is achieved. For the one dimensional FFTs we use the FFTW library[19].

Our implementation is denoted in this paper as *DIYFFT*. The performance of the parallel FFT as well as the overall scaling of the code SpecDyn on the JUWELS CPU cluster module are presented in Fig. 1.

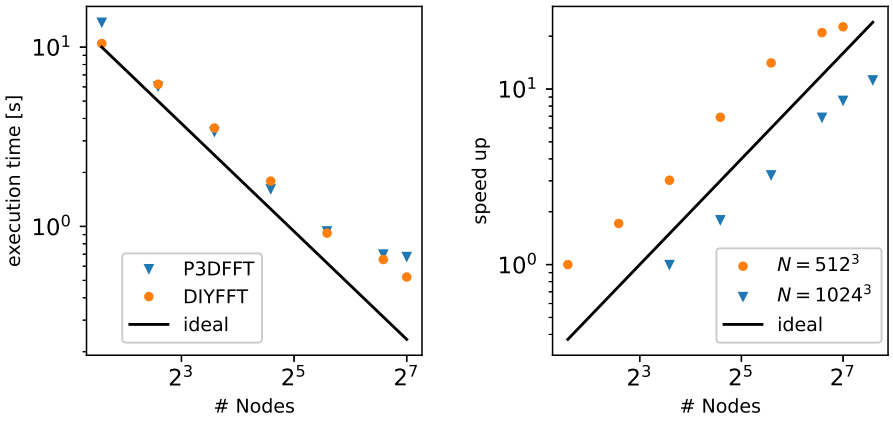


Figure 1. Performance of the *DIYFFT* and *SpecDyn* on the JUWELS CPU cluster module with 48 cores per node: a) Scaling of the *DIYFFT* compared to the *P3DFFT* library[13] performing 100 real-to-complex/complex-to-real transforms under a spatial discretisation of 512^3. b) Scaling of the code *SpecDyn* performing 100 time steps on a grid of size 512^3 and 1024^3 respectively.

3 Numerical Experiments on Precession Driven Flows in the Context of Dynamo Action

Precession driven flows are studied by the use of the code *SpecDyn*. The different observable flows are mainly dependent on the geometry of the cylinder, the Poincaré number Po, i.e. the ratio of precession to spin frequency, and the Reynolds number Re. As the aspect ratio of the cylinder in the experiment is fixed to a value of 2 to gain resonant waves, the other two parameters can be freely varied in a certain range. Depending on their value, three different flow regimes are observed: laminar, non-linear and turbulent[20]. We are interested in flows suitable for dynamo action, for which rather complex flows (particularly in the non-linear regime) are usually supposed to be beneficial. To pass behind the laminar regime, relatively high Reynolds numbers are required. If Re is sufficiently high, the transition merely depends on the Poincaré number[3]. Additionally, the Reynolds numbers achievable by the experiment are restricted to high values to guarantee a smooth rotation of the vessel. On the other hand higher Reynolds numbers cause flows with smaller boundary layers forming at the bounding walls due to viscosity whose thickness scales with $Re^{-1/2}$ [21]. In our numerical experiments we choose $Re = 6500$ and a resolution of 256 Fourier modes per dimension for which the laminar regime is readily overcome and the Ekman boundary layer is yet resolved with 2 grid cells. For the numerical experiments the Poincaré number is varied and the cylindrical base geometry of the vessel is penetrated by baffles of different depths. The cylinder has a diameter and height of 2 length units L. The dimensions of the four blades located at the end caps of the cylinder are $0.055L \times 0.040L$ and the maximum penetration depth adds up to $0.065L$.

For runs with the plain cylindrical vessel we obtain a peak in the non-axial-symmetric flow and an abrupt drop in the forced mode at $Po = 0.0925$. Further we observe a double roll structure which reaches its peak amplitude at $Po = 0.1$. This is in good agreement with other numerical experiments[22,23] (Fig. 2). The double roll structure is of particular interest since kinematic simulations of Giesecke *et al.*[23] showed the onset of dynamo action only

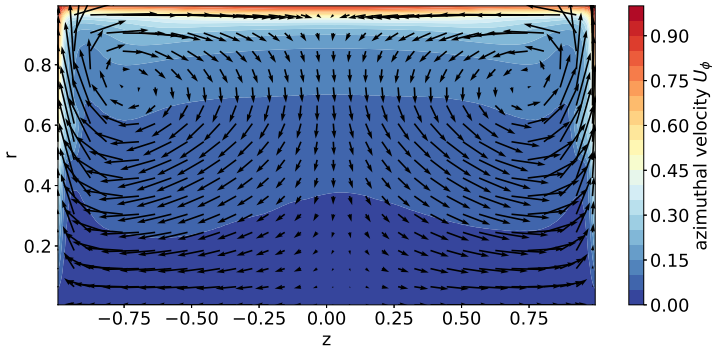


Figure 2. Time and azimuthal averaged flow field for $Po = 0.105$ and baffles extended to maximum penetration depth in cylinder coordinates. The azimuthal velocity field is colour coded whereas the arrows represent the flow in the r-z plane.

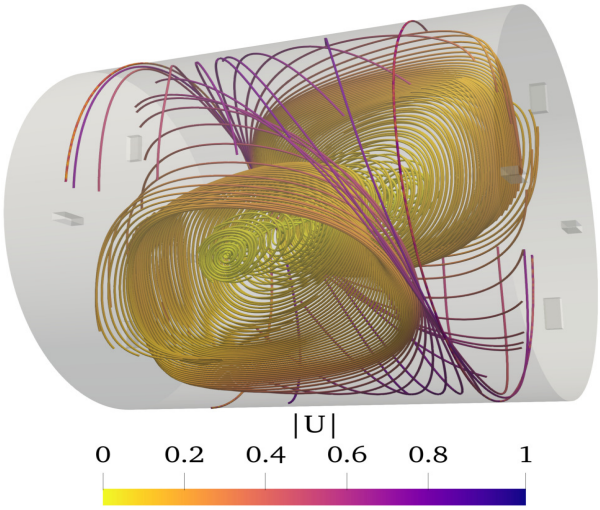


Figure 3. Velocity streamlines of the time averaged flow colour coded by the velocity magnitude for $Po = 0.105$ and the baffles extended to maximum penetration depth.

for regimes where the double roll structure was excited.

Further, for the runs with the baffles located in the flow, we observe a trend of the maximum of the non-axial-symmetric energy as well as that of the double roll structure to appear at slightly higher Poincaré numbers, although the overall flow structure is barely altered by the baffles. This effect was already predicted by Léorat *et al.*[24] with the aim to generate more energy in the forced Kelvin mode before it partially breaks down into higher modes and especially the double roll flow after some critical Po is reached. The simulations were carried out for the baffles extended to $0.5, 1.$ and 2 times the maximum penetration depth reachable in the real world experiment.

In Fig. 3 the characteristic behaviour of the flow to rotate about the axis of precession for large Poincaré numbers is presented.

4 The muphyII Code: Multiphysics Plasma Simulation on JUWELS Booster

The *muphy*II multiphysics code provides a framework for coupling physical models and is applied in the context of plasma simulations. Kinetic, hybrid and multi-fluid continuum models can be run either individually or spatially coupled to each other. The name muphy is an acronym for multiphysics and *muphy*II is the successor to the previous *muphy* code. The framework part of *muphy*II is written in C++ for flexibility and the numeric calculations utilise Fortran for optimal performance in multi-dimensional array operations and good maintainability.

Several physical models and the respective solvers are implemented in *muphy*II. The most detailed plasma model solves the Vlasov equation on a phase space grid and can utilise either a pure Vlasov solver[25,26] or an energy conserving dual Vlasov solver[27]. In

order to reduce computational cost, a plasma is commonly approximated as a system of multiple fluids, one per particle species, or a single fluid (then called magnetohydrodynamics, MHD). The ten moment multi-fluid model is the next model in the hierarchy. It includes the full pressure tensor, but needs a physically motivated approximation for heat flux. The least detailed model that is currently used in the coupling hierarchy is the five moment multi-fluid model which assumes perfectly isotropic pressure and zero heat flux. The standard form of the ten and five moment multi-fluid equations is used which is commonly written down in the literature[28,29]. A more fine-grained usage of the models can be achieved by making hybrid versions. Between the Vlasov and the ten moment model there is a hybrid kinetic model that evolves the ions according to the Vlasov equation and the electrons according to the ten moment fluid equations. Accordingly, there is a hybrid fluid model with ten moment ions and five moment electrons. All the mentioned plasma models are paired with a finite difference time domain (FDTD) solver for Maxwell's equations.

Parallelisation is done via domain decomposition. There, the spatial simulation domain is split up into several blocks and all blocks can be computed individually by separate computing processes, for example separate GPUs, CPUs or CPU cores. This is a common method which has the advantage that it scales very well to large numbers of processes. As the numerical scheme needs information on neighbouring cells, the blocks copy information into boundary cells of their neighbouring blocks. This is done via the message passing interface (MPI), the standard tool for communication within distributed memory computing systems. Since each block consists of multiple cells which can be updated independently from each other, there is a second level of parallelism inside a block where shared memory is available. In *muphy*II shared memory parallelisation is done via OpenACC in order to be able to utilise GPUs. CUDA-aware MPI is utilised for fast communication between the GPUs.

The scaling of the computing time that *muphy*II needs at increasing numbers of GPUs is shown in Fig. 4 for up to 1024 A100 GPUs. Since one A100 GPU features 6912 compute cores, 1024 GPUs correspond to more than 7 million cores. The plot shows the weak scaling where the number of cells is doubled while at the same time the computing re-

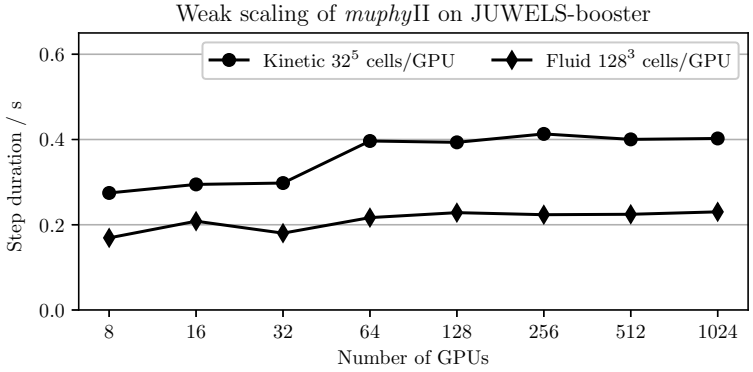


Figure 4. Computing time scaling of *muphy*II while increasing resolution and number of GPUs on JUWELS Booster at Jülich Supercomputing Centre.

sources are doubled. Ideally, this results in the same computation time per step. The scaling of *muphy*II between 64 and 1024 GPUs is close to ideal and this trend is expected to continue towards even higher numbers of GPUs. The sudden increase in step time of the kinetic solver between 32 and 64 GPUs is likely due to nodes being part of one network cell in the 32 GPUs case and spread among different network cells in the 64 GPUs case (MPI bandwidth within a network cell in JUWELS Booster is higher). This is a consequence of JUWELS Booster's network design and unrelated to the implementation. The scaling study was performed in double precision at the example of an Orszag-Tang vortex plasma turbulence simulation in two (kinetic) and three (fluid) dimensions. The kinetic solver used the energy conserving dual Vlasov method[27] and the fluid solver employed the ten moment multi-fluid model with the temperature gradient heat flux closure[27,30,29]. The FDTD Maxwell solver used 15 subcycles.

In coupled simulations there is a large gap in computational time between the Vlasov model and increasingly more reduced models. This is technically treated by utilising the CUDA Multi-Process Service (MPS) where multiple MPI processes share one GPU. With proper load balancing it is possible to utilise the GPUs to full capacity with a single (or few) Vlasov process(es) in the first step of the physical scheme. Then in the second step the GPUs are also utilised to up to full capacity by a much larger number of fluid processes that are computed on the GPU at the same time.

5 Coupled Simulation of Magnetic Reconnection in the Earth's Foreshock

Since the launch of the Magnetospheric Multiscale Mission (MMS)[31], it is possible to directly measure reconnection in the Earth's magnetosphere. A reconnection event on 10th November 2017 at 17:26:17 UT measured by MMS in the Earth's foreshock was reported in Liu *et al.*[32] and modelled with a PIC simulation. A Vlasov simulation of the event was performed in Allmann-Rahn *et al.*[27] Here, a coupled simulation is carried out where only a small fraction of the domain is treated with the fully kinetic Vlasov model.

The initial conditions are given by a Harris sheet with guide field and are identical to the ones used in Allmann-Rahn *et al.*[27] apart from a slightly different resolution; also see the original setup in Liu *et al.*[32] Position space is here resolved by 1152×435 cells, and velocity space by 36^3 cells for the electrons and 24^3 cells for the ions.

Magnetic reconnection in the magnetosphere is one of the many scenarios where spatial coupling of complete and reduced plasma models can be effectively used. Heat flux is

Scheme	Description	Criterion
V	Vlasov ions, Vlasov electrons	$j > 1.25\, n_0 v_{A,0} \;\vee\; u_e > 4.0\, v_{A,0}$
V-F10	Vlasov ions, 10 moment electrons	$j > 0.75\, n_0 v_{A,0} \;\vee\; u_e > 2.0\, v_{A,0}$
F10	10 moment ions, 10 moment electrons	$j > 0.30\, n_0 v_{A,0} \;\vee\; u_e > 1.0\, v_{A,0}$
F10-F5	10 moment ions, 5 moment electrons	$j > 0.10\, n_0 v_{A,0} \;\vee\; u_e > 0.5\, v_{A,0}$
F5	5 moment ions, 5 moment electrons	else.

Table 1. Plasma models and criteria used in the coupled simulation of foreshock reconnection.

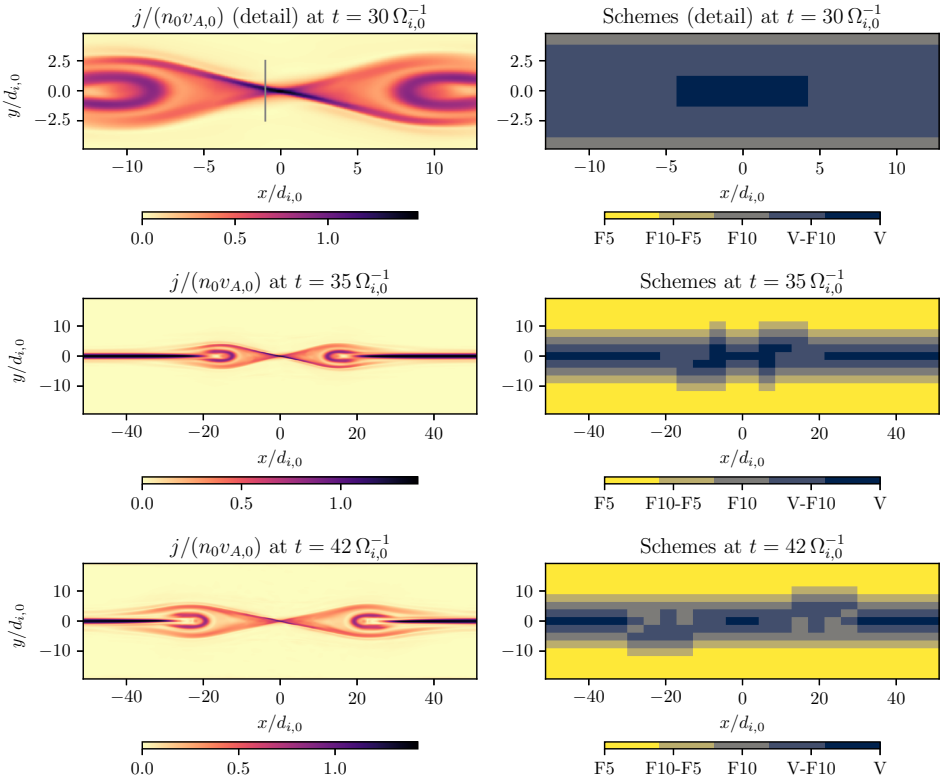


Figure 5. Current density $|\mathbf{j}|$ (left) next to the utilised plasma schemes (right) at different times in the foreshock reconnection simulation. The grey line in the top left panel visualises the virtual path along which the data in Fig. 6 is taken.

localised around the reconnection current sheet and the separatrix[27] so that a ten moment fluid model is sufficient in large parts of the physical domain. Further outside of the reconnection region there is no pressure anisotropy so that the even less detailed five moment fluid model can be applied.

A criterion based on local plasma parameters decides which of the models in the hierarchy covers the respective parts of the domain. For the reasons listed in the previous paragraph, it is suitable for magnetic reconnection to choose the magnitude of the current density j as a criterion. In addition we also consider the magnitude of the electron mean velocity u_e in order to appropriately treat the reconnection X-line and the separatrix borders. The different available models and the thresholds of j and u_e that form the criterion are given in Tab. 1.

Each block of the domain decomposition can contain a different model, but neighbouring models are restricted to differ by only one step in the hierarchy of models. For example the full Vlasov model (V) can be a neighbour to the ten moment fluid electrons, Vlasov ions hybrid model (V-F10), but not to the full ten moment fluid model (F10). In the re-

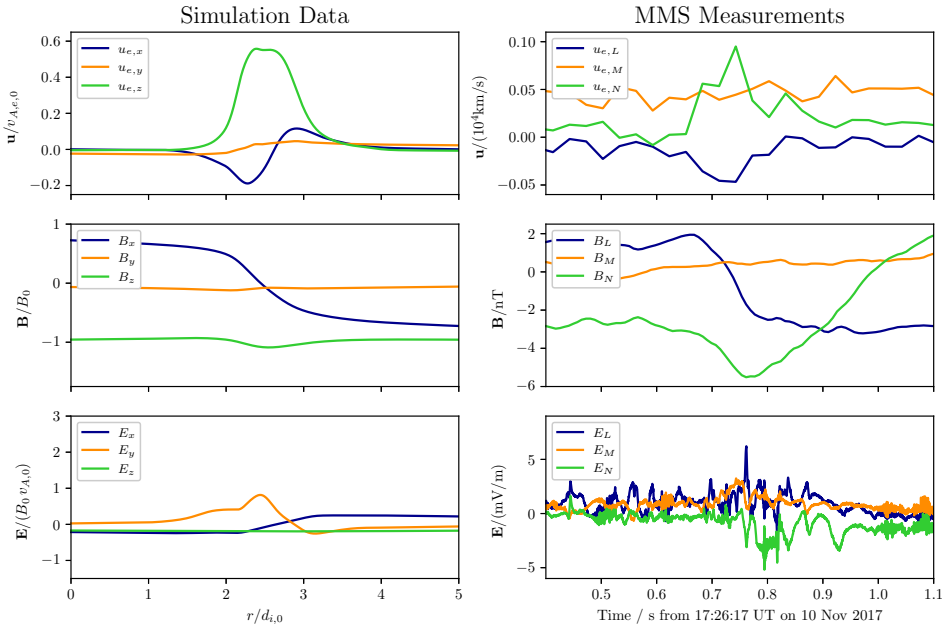


Figure 6. Electron velocity, magnetic field and electric field along a virtual path in the foreshock reconnection simulation at $t = 30\,\Omega_{i,0}^{-1}$ (left) and as measured by MMS (right).

connection simulation discussed in this section, the x-y plane is decomposed into 24×15 blocks.

For smooth coupling between the Vlasov and fluid models, a method similar to the one introduced in Rieke *et al.*[33] is employed. There the distribution function is extrapolated into the boundary cells and then adapted based on the moments prescribed by the fluid model. Here, the adaption of the distribution function is done by exchanging Maxwellians[27]. The part of the distribution function in the boundary cells stemming from the constant extrapolation is chosen do decrease linearly ($2/3$ in the first and $1/3$ in the second boundary cell). For the coupling of the ten and five moment fluid models, the pressure tensor is extrapolated into the boundary cells and then rescaled uniformly in all components such that its trace is consistent with the isotropic pressure prescribed by the fluid model.

In Fig. 5 the current density in the simulation at different times is shown together with the spatial distribution of the schemes. The chosen criterion mostly identifies the areas where a fully kinetic treatment is necessary, including the reconnection X-line. It is also clear that there is room for improvement, for example the reconnection O-lines would better be represented by the full Vlasov scheme. Despite the tiny fraction of the domain in which the fully kinetic scheme is used, the results agree exceptionally well with a simulation that uses full Vlasov everywhere[27], both concerning the reconnection rate and the spatial structure. The transition between models is mostly smooth and without problems. In Fig. 6 the simulation data is compared to MMS measurements along the virtual path

shown in the top left panel of Fig. 5. Both the path and the time in the simulation are identical to the ones chosen in Allmann-Rahn *et al.*[27] in a fully kinetic simulation. The coupled simulation here agrees quantitatively well with the measurements, just like the fully kinetic counterpart, while it required only one eighth of the computational resources.

Acknowledgements

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). Computations were conducted on JUWELS/JUWELS Booster[34] and on the DaVinci cluster at TP1 Plasma Research Department. F.A. was supported by the Helmholtz Association (VH-NG-1239). M.W. and R.G. acknowledge financial support from DFG GR 967/7-1 and A.G. from DFG GI 1405/1-1. Thanks to the MMS team for the measurement data available at the MMS Science Data Center (`https://lasp.colorado.edu/mms/sdc/`), and to the developers of the pySPEDAS software (`https://github.com/spedas/pyspedas`) as well as the developers of the SpacePy software (`https://spacepy.github.io/`).

References

1. M. L. Bars and D. Lecoanet, *Fluid Mechanics of Planets and Stars*, CISM International Centre for Mechanical Sciences. Springer International Publishing, 2019.
2. W. V. R. Malkus, *Precession of the Earth as the cause of geomagnetism: Experiments lend support to the proposal that precessional torques drive the Earth's dynamo*, Science, **160**, no. 3825, 259–264, 1968.
3. F. Stefani, T. Albrecht, G. Gerbeth, A. Giesecke, T. Gundrum, J. Herault, C. Nore, and C. Steglich, *Towards a precession driven dynamo experiment*, Magnetohydrodynamics, **51**, no. 2, 275–284, 2015.
4. F. Stefani, S. Eckert, G. Gerbeth, A. Giesecke, Th. Gundrum, C. Steglich, T. Weier, and B. Wustmann, *DRESDYN - A new facility for MHD experiments with liquid sodium*, Magnetohydrodynamics, **48**, no. 1, 103–113, 2012.
5. C. -W. Shu and St. Osher, *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, Journal of computational physics, **77**, no. 2, 439–471, 1988.
6. J. W. Cooley and J. W. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Mathematics of computation, **19**, no. 90, 297–301, 1965.
7. C. Canuto, M. Y. Hussaini, A. Quarteroni, and Th. A. Zang, *Spectral methods: fundamentals in single domains*, Springer Science & Business Media, 2007.
8. Ch. S. Peskin, *Numerical analysis of blood flow in the heart*, Journal of computational physics, **25**, no. 3, 220–252, 1977.
9. J. Mohd-Yusof, *Combined immersed boundary/B-Spline method for simulations of flows in complex geometries in complex geometries CTR annual research briefs, NASA Ames*, NASA Ames/Stanford University, 1997.
10. D. Goldstein, R. Handler, and L. Sirovich, *Modeling a no-slip flow boundary with an external force field*, Journal of computational physics, **105**, no. 2, 354–366, 1993.

11. E. A. Fadlun, R. Verzicco, P. Orlandi, and J. Mohd-Yusof, *Combined immersed-boundary finite-difference methods for three-dimensional complex flow simulations*, Journal of computational physics, **161**, no. 1, 35–60, 2000.
12. S. Woop, C. Benthin, and I. Wald, *Watertight ray/triangle intersection*, Journal of Computer Graphics Techniques (JCGT), **2**, no. 1, 65–82, 2013.
13. D. Pekurovsky, *P3DFFT: A framework for parallel computations of Fourier transforms in three dimensions*, SIAM Journal on Scientific Computing, **34**, no. 4, C192–C209, 2012.
14. N. Li and S. Laizet, *2decomp & fft-a highly scalable 2d decomposition library and fft interface*, in: Cray user group 2010 conference, 1–13, 2010.
15. M. Pippig, *PFFT: An extension of FFTW to massively parallel architectures*, SIAM Journal on Scientific Computing, **35**, no. 3, C213–C236, 2013.
16. K. G. Larkin, M. A. Oldfield, and H. Klemm, *Fast Fourier method for the accurate rotation of sampled images*, Optics communications, **139**, no. 1-3, 99–106, 1997.
17. L. Dalcin, M. Mortensen, and D. E. Keyes, *Fast parallel multidimensional FFT using advanced MPI*, Journal of Parallel and Distributed Computing, **128**, 137–150, 2019.
18. Message Passing Interface Forum, *MPI: A Message-Passing Interface Standard Version 4.0*, June 2021.
19. M. Frigo and S. G. Johnson, *The Design and Implementation of FFTW3*, Proceedings of the IEEE, **93**, no. 2, 216–231, 2005, Special issue on "Program Generation, Optimization, and Platform Adaptation".
20. F. Pizzi, A. Giesecke, J. Šimkanin, and F. Stefani, *Prograde and retrograde precession of a fluid-filled cylinder*, New Journal of Physics, **23**, no. 12, 2021.
21. F. Pizzi, A. Giesecke, and F. Stefani, *Ekman boundary layers in a fluid filled precessing cylinder*, AIP Advances, **11**, no. 3, 035023, 2021.
22. A. Giesecke, T. Vogt, Th. Gundrum, and F. Stefani, *Nonlinear large scale flow in a precessing cylinder and its ability to drive dynamo action*, Physical Review Letters, **120**, no. 2, 024502, 2018.
23. A. Giesecke, T. Vogt, Th. Gundrum, and F. Stefani, *Kinematic dynamo action of a precession-driven flow based on the results of water experiments and hydrodynamic simulations.*, Geophysical & Astrophysical Fluid Dynamics, **113**, no. 1-2, 235–255, 2019.
24. J. Léorat, P. Lallemand, J. L. Guermond, and F. Plunian, "Dynamo action, between numerical experiments and liquid sodium devices", in: Dynamo and dynamics, a Mathematical challenge, 25–33, Springer, 2001.
25. F. Filbet, E. Sonnendrücker, and P. Bertrand, *Conservative Numerical Schemes for the Vlasov Equation*, Journal of Computational Physics, **172**, no. 1, 166–187, 2001.
26. H. Schmitz and R. Grauer, *Kinetic Vlasov simulations of collisionless magnetic reconnection*, Physics of Plasmas, **13**, no. 9, 092309, 2006.
27. F. Allmann-Rahn, S. Lautenbach, and R. Grauer, *An Energy Conserving Vlasov Solver That Tolerates Coarse Velocity Space Resolutions: Simulation of MMS Reconnection Events*, Journal of Geophysical Research: Space Physics, **127**, no. 2, e2021JA029976, 2022.
28. L. Wang, A. H. Hakim, A. Bhattacharjee, and K. Germaschewski, *Comparison of multi-fluid moment models with particle-in-cell simulations of collisionless magnetic reconnection*, Physics of Plasmas, **22**, no. 1, 012108, 2015.

29. F. Allmann-Rahn, S. Lautenbach, R. Grauer, and R. D. Sydora, *Fluid simulations of three-dimensional reconnection that capture the lower-hybrid drift instability*, Journal of Plasma Physics, **87**, no. 1, 905870115, 2021.
30. F. Allmann-Rahn, T. Trost, and R. Grauer, *Temperature gradient driven heat flux closure in fluid simulations of collisionless reconnection*, Journal of Plasma Physics, **84**, no. 3, 905840307, 2018.
31. J. L. Burch, T. E. Moore, R. B. Torbert, and B. L. Giles, *Magnetospheric Multiscale Overview and Science Objectives*, Space Science Reviews, **199**, 5–21, 2016.
32. T. Z. Liu, S. Lu, D. L. Turner, I. Gingell, V. Angelopoulos, H. Zhang, A. Artemyev, and J. L. Burch, *Magnetospheric Multiscale (MMS) Observations of Magnetic Reconnection in Foreshock Transients*, Journal of Geophysical Research: Space Physics, **125**, no. 4, e2020JA027822, 2020.
33. M. Rieke, T. Trost, and R. Grauer, *Coupled Vlasov and two-fluid codes on GPUs*, Journal of Computational Physics, **283**, 436–452, 2015.
34. Jülich Supercomputing Centre, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, Journal of large-scale research facilities, **5**, no. A135, 2019.

Interaction of Extremely Intense Flows of Electromagnetic Energy and QED Processes in Supercritical Fields

Lars Reichwein[1], Marko Filipovic[1], Xiaofei Shen[1], Ke Jiang[1], Christoph Baumann[1,2], and Alexander Pukhov[1]

[1] Institut für Theoretische Physik I, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany
E-mail: lars.reichwein@hhu.de

[2] Institut für Energie- und Klimaforschung – Plasmaphysik, Forschungszentrum Jülich, 52425 Jülich, Germany

We present an overview of our recent results studying extremely intense flows of electromagnetic energy and QED processes in supercritical fields. This includes a novel solver for Maxwell's equations which is able to suppress the numerical Cherenkov instability commonly arising in this parameter range. The solver is used for simulations of electron-electron and electron-positron beam collisions that aim to produce an abundance of γ-photons and/or electron-positron pairs. An analytical model describing the disruption parameter in a radiation-reaction-dominated regime is presented.

1 Introduction

With the increase of laser intensity over the last few decades and the future aim of facilities like the Extreme Light Infrastructure (ELI) to obtain intensities $\geqslant 10^{23}\ \mathrm{Wcm}^{-2}$, the investigation of quantum electrodynamic (QED) processes in plasma is reachable in the near future. In the scope of our project at JUWELS we have investigated extremely intense flows of electromagnetic energy and QED processes in supercritical fields at several fronts: First, the optimisation of beam-beam collisions pair production yields as well as the suppression of the numerical Cherenkov instability (NCI) with our group's x-dispersionless Maxwell solver was studied[1–3]. A second sub-project examined the acceleration of electrons and protons from the interaction of a highly intense laser pulse with the edge of a high-Z tape target which i.a. allows for mono-energetic high-energy ions[4]. Femtosecond lasers were also used to amplify magnetic fields in plasma up to the Giga-Gauss level[5]. Lastly, a third project consisted of the acceleration of spin-polarised proton beams via magnetic vortex acceleration which prove to be useful for investigating the Standard Model of particle physics[6]. Each of the different projects utilises the resources of JUWELS in the form of massively parallelised particle-in-cell (PIC) simulations, in our case with the code VLPL[7,8]. In this proceeding, we will focus on the results of the first sub-project dealing with QED processes for beam-beam collisions.

The importance of QED effects can be quantified with the parameter $\chi = \sqrt{-(F_{\mu\nu}p^{\nu})^2}/(m_e c E_{\mathrm{cr}})$, which physically gives the electric field seen by a particle in its rest frame in units of the so-called Schwinger field E_{cr} (i.e. the critical field of vacuum breakdown)[9]. Further, m_e is the electron rest mass, $F_{\mu\nu}$ is the electromagnetic field tensor, c is the vacuum speed of light and p^{ν} is the momentum four-vector. If χ approaches the value of unity, QED effects like hard-photon emission due to non-linear Compton scattering and electron-positron pair production due to the multi-photon Breit-Wheeler process

become significant. Experiments in the regime $\chi \approx 1$ are planned to be realised at the Stanford Linear Accelerator Center (SLAC) as well as in the Laser Und XFEL Experiment (LUXE) at DESY. Further, if $\alpha\chi^{2/3} \approx 1$ (α being the fine-structure constant, and consequently $\chi \approx 1600$), QED is conjectured to become fully non-perturbative[10]. This regime has thus far not been studied experimentally, and even a reliable theory is still lacking. Despite these limitations, several schemes for experimental implementation and preliminary analytical studies have been proposed in the not-so-distant past. A comprehensive overview about recent advances in QED is given by Fedotov *et al.*[11].

PIC simulations are an indispensable tool in the strong-field plasma community in order to study highly dynamical many-charge systems numerically. In the following, we will give a brief introduction into the basic idea of the PIC method and how one can embed the most prominent QED effects, namely radiation reaction and pair creation, into that framework. We will then discuss the advantages of a special solver for Maxwell's equations that is able to suppress certain instabilities in simulations that would otherwise yield unphysical results. The last two sections will summarise our findings for the collision of a high-density electron beam with another electron (or positron) beam. These setups aim to reach the fields relevant for QED without the need of external intense laser fields. Here, we will specifically show how the pair and photon yield can be improved upon.

2 The Particle-In-Cell Method

The simulation of self-consistent plasma dynamics is highly resource-intensive, if one aims for a good representation of the physical reality due to the high number of particles that are involved in the processes. A first approach might be to describe the phase space in a statistical manner, e.g. with the collisionless Boltzmann-Vlasov equation which describes the time evolution of a single-particle distribution function $f_\mu(\mathbf{x}, \mathbf{p}, t)$ for a species μ.

However, the particle motion induces changes in the charge and current distribution, which lead to new electromagnetic fields described by Maxwell's equations. A self-consistent approach thus demands a simultaneous solution of particle motion and Maxwell's equations, which poses a significant numerical challenge due to the six-dimensionality of the phase space.

The PIC method allows for further simplification of the problem: we can consider the continuous distribution function f_μ to be consisting of several finite phase-fluid elements, i.e. we discretise the problem. These finite elements are often called "macro-particles" in the context of PIC simulations, but can be thought of as clouds of a certain charge density that can substitute for a large number of physical particles.

Each macro-particle obeys the Boltzmann-Vlasov equation. The characteristic motion of the particle is nothing else but the solution of the relativistic equations of motion for a particle subjected to an electromagnetic field. Due to its charge-to-mass ratio being the same as that of a physical particle, a macro-particle follows the same trajectory. The current induced by the macro-particle motion allows for self-consistent description of the plasma when used for the temporal evolution of the electromagnetic fields in Maxwell's equations. Importantly, the plasma sampling technique via macro-particles reduces the computational effort tremendously as now only the three-dimensional real space has to be discretised to solve Maxwell's equations compared to the six-dimensional phase space needed for the Vlasov equation.

In order to solve Maxwell's equations on a discrete grid, several solvers have been proposed in the past, one of the most prominent being the Yee solver. As we will discuss later on, this solver has certain limitations when compared with other solvers like our group's rhombi-in-plane (RIP) solver[12].

The PIC cycle thus consists of the following steps: First, the fields have to be interpolated to the correct particle position, as they are only known at the discretised grid points. Next, the particles are pushed by the interpolated force. After the push, the generated current is interpolated to the grid on which the fields are known. Finally, the fields are advanced with the currents on the grid. This cycle is continued throughout the simulation.

In order to deliver physically correct results, several stability criteria depending on the simulation's spatial and temporal resolution need to be fulfilled. For the Yee solver, it is required that

$$c\Delta t < \frac{1}{\sqrt{\frac{1}{h_x^2} + \frac{1}{h_y^2} + \frac{1}{h_z^2}}} \,, \tag{1}$$

where the h_i denote the grid size for the different directions, Δt denotes the time step and c is the vacuum speed of light. In other words, information cannot propagate faster than the speed of light. For the later discussed RIP solver, the stability requirement is that

$$h_x^2 \left(\frac{1}{h_y^2} + \frac{1}{h_z^2} \right) < 1 \,. \tag{2}$$

Here, a further requirement of RIP is that the time step is solely related to the grid size in main particle propagation direction. In the VLPL code, this is usually the x-direction, so that $h_x = c\Delta t$. As the time step typically sets the scale of the resolution (fast processes need to be resolved), this can lead to an increase of computational resources, but also delivers advantages we will discuss in the following section.

As the standard PIC method solves the collisionless Vlasov-Boltzmann equation, the method has to be extended in order to incorporate QED processes. Here, the inclusion of an additional Monte Carlo routine proves to be the state-of-the-art approach in the regime $\chi \gtrsim 1$. In the case of VLPL, we utilise the model of Elkina *et al.*[13]. A benchmark of the VLPL implementation can be found in Ref. 14. The Monte Carlo method aims to capture the probabilistic nature of QED processes and is implemented as an additional step in the PIC cycle after the particle push.

A first sub-routine concerns the emission of high-energy photons by ultra-relativistic particles ("non-linear Compton scattering"). Whether or not an electron or positron emits a photon is decided the following way: a uniformly distributed random number $r_1 \in [0, 1]$ is generated. This value gives a percentage of the considered particle's energy and momentum that shall be emitted. In a second step, the quantum non-linearity parameter for particle and photon is calculated. The probability for photon emission is $P_\gamma = (\mathrm{d}W_{\mathrm{rad}}/\mathrm{d}t)\, \varepsilon_e \Delta t$ for a time step Δt. Using a second random number $r_2 \in [0, 1]$, it is finally decided whether a photon is emitted or not: if $r_2 < P_\gamma$ a photon is emitted, otherwise it is not. To model the recoil induced by the emission, the momentum of the emitting particle is reduced by the emitted photon momentum, $\mathbf{p}'_e = \mathbf{p}_e - \mathbf{p}_\gamma$.

A second sub-routine deals with electron-positron pair creation through the decay of high-energy γ photons. Again, a random number $\tilde{r}_1 \in [0, 1]$ is chosen to give the electron momentum as a percentage of the photon momentum, and simultaneously for its energy.

The probability $P_{\mathrm{pair}} = (\mathrm{d}W_{\mathrm{pair}}/\mathrm{d}t)\,\varepsilon_\gamma \Delta t$ for photon decay according to Breit-Wheeler is then compared to a random number $\tilde{r}_2$ to decide whether the process occurs. The momentum of the created positron simply is calculated from momentum conservation at the vertex as $\mathbf{p}_{e^+} = \mathbf{p}_\gamma - \mathbf{p}_{e^-}$ from the electron and photon momenta.

In contrast, there is no need to introduce a stochastic Monte Carlo approach in the regime $\chi \ll 1$. There, radiation damping, also known as radiation reaction, can be considered as a continuous loss of energy. The corresponding damping force is implemented by correcting the momentum after the Lorentz force push with the characteristic radiation loss frequency ν_{RR}. In order to determine this frequency, we calculate the estimated total emitted power with the relativistic Larmor formula and assume the radiation-reaction force for ultra-relativistic particles to act anti-parallel to the current momentum,

$$\mathbf{F}_{\mathrm{RR}} = -\frac{2}{3}\alpha \frac{m_e c^2}{\hbar} \chi^2 G(\chi) \frac{\mathbf{p}}{\gamma} \equiv -\nu_{\mathrm{RR}} \mathbf{p}. \tag{3}$$

Here, $\hbar$ is the reduced Planck constant and $G(\chi)$ is the Gaunt factor, which incorporates a reduced emission of radiation by charges once QED effects become relevant[11]. Finally, the momentum of the particle is damped as $\mathbf{p}_{\mathrm{RR}} = \mathbf{p}/(1 + \nu_{\mathrm{RR}} \Delta t)$.

3 Suppression of the Numerical Cherenkov Instability

The Yee algorithm commonly used for solving Maxwell's equations is prone to numerical errors. These can result in a slow-down of a wave's propagation velocity (also known as "numerical dispersion") which can favour numerical instabilities. These numerical instabilities in turn may lead to unphysical behaviour in the simulations. The numerical Cherenkov instability (NCI) gets its name from the Cherenkov radiation emitted when charged particles move faster than the electromagnetic phase velocity in a medium. Accordingly, NCI occurs when particles move faster than the numerical phase velocity.

The rhombi-in-plane solver is an x-dispersionless Maxwell solver that successfully suppresses this instability in simulations. In a recent publication[3], we have compared the two solvers for the case of ultra-relativistic and high-density electron beam propagation in vacuum with the continuous radiation-reaction model considered. Physically, the electron beam is not expected to suffer Cherenkov radiation or radiation-reaction losses from the interaction with its own self-fields. As we will see, both is not true for the Yee solver, but can be efficiently suppressed with the RIP solver. For our comparison we use a simulation box of $200 \times 400 \times 400\ \mathrm{nm}^3$. In transverse direction, a grid size of $h_y = h_z = 1$ nm regardless of the utilised solver. The time step is chosen to be $c\Delta t = 0.5$ nm for both solvers. Due to the different stability conditions for the two solvers, we are able to set $h_x = 1$ nm for the Yee solver, but have to use $h_x = c\Delta t$ for the RIP solver.

The electron bunch has a spherical shape and is of Gaussian density profile; 8 particles per cell are used. The electrons have a uniform Lorentz factor of $\gamma_0 = \sqrt{1 + p_0^2/(m_e c)^2} = 2.5 \times 10^5$ and propagate in x-direction. The peak density is $n_0 = 5.6 \times 10^{25}\ \mathrm{cm}^{-3}$ and its radius is $\sigma_r = 10$ nm.

The particles are not pushed until approx. T_b, where $T_b = \sqrt{\gamma_0 m_e/(4\pi e^2 n_0)} \approx 7.5$ fs is the inverse of the relativistically corrected beam plasma frequency. This is done to minimise the error of the initialised beam self-fields on the grid. Subsequently, the time $t = 0$ is taken with respect to this period.

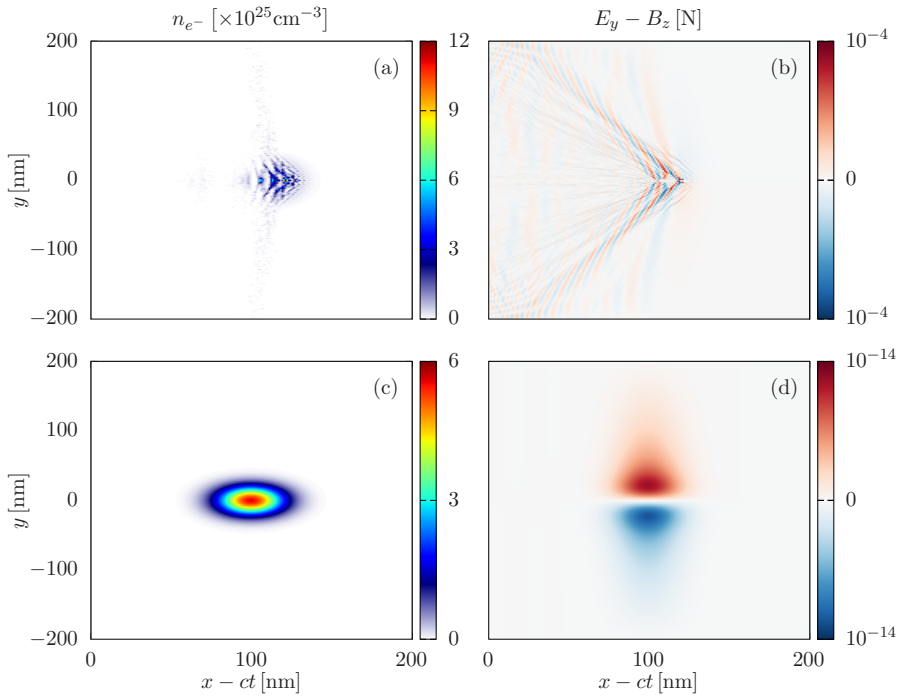


Figure 1. Propagation of an ultra-relativistic electron bunch through vacuum with radiation reaction considered. Plots (a) and (b) show the electron density and transverse force for a simulation with the Yee solver, (c) and (d) show the same for the RIP solver. Note that the force in the case of RIP is several magnitudes smaller and indeed gives the analytically correct order.

Already after $10\,T_b$ the electron bunch in the simulation with the Yee solver is so strongly deformed, that it cannot be considered a Gaussian bunch anymore (see Fig. 1). Due to compression the peak density has increased to 1.3×10^{26} cm^{-3}. In comparison to simulation results without radiation reaction, we observe significantly more distortion of the beam. In contrast, the simulations using the RIP solver do not show such behaviour: the Gaussian bunch still exhibits a Gaussian density profile.

The Fourier transform of the electric field's y-component can be used as a diagnostic for NCI. For the Yee algorithm, a mode in the $k_x - k_y$ plane can be seen regardless of whether radiation reaction is considered or not [see upper plot in Fig. 2(a)]. This excitable mode does not arise for the RIP solver [see lower plot in Fig. 2(a)], indicating that the instability is suppressed. The effect of the instability can also be seen, when considering the field energy associated with the longitudinal magnetic field, $W_{B_x} = \int B_x^2 \, \mathrm{d}V/(8\pi)$ (not shown here, see Ref. 3). For an ultra-relativistic bunch propagating in x-direction, no such field component B_x should be observed. Here, the comparison between Yee and RIP further underlines that RIP is able to suppress NCI and does not exhibit differences due to radiation reaction like they arise for the Yee scheme.

Physically, the interaction of a highly relativistic beam with its self-field is negligible and the electron beam moves with almost constant energy through space. However, the

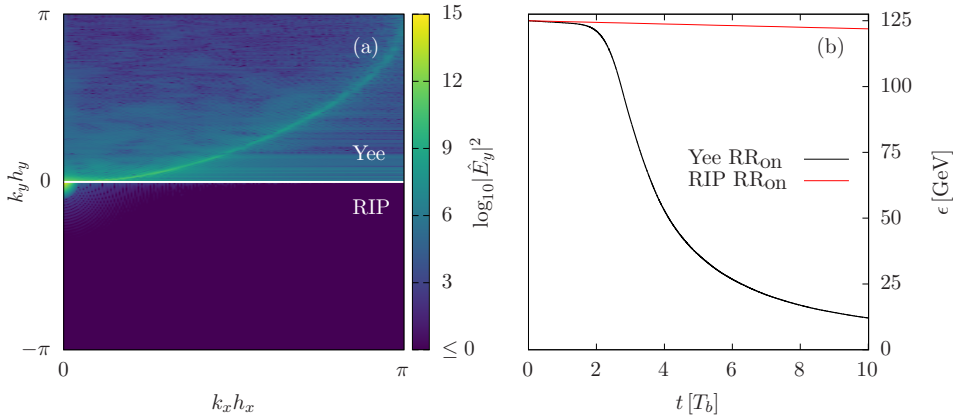


Figure 2. (a) Results of the Fourier transform for Yee and RIP solver. Yee exhibits an excited mode that is absent from the RIP plot. (b) Comparison of the energy loss for both solvers.

simulation using the Yee Maxwell solver reveals a different behaviour. The electron beam loses a large amount of its initial energy when the radiation reaction module is activated [see Fig. 2(b)]. In fact, the electron energy ϵ drops by approx. 91.4% at time $t = 10\,T_b$ for the Yee case, while it only drops by 4.4% for the RIP Maxwell solver. The onset of the NCI with the Yee solver causes the transverse force to be overestimated by orders of magnitude [see Figs. 1(b) and (d)]. The transverse force in turn results in transverse acceleration and so leads to unphysical radiation-reaction losses. Especially in scenarios where radiation losses need to be mitigated as much as possible (e.g. when one aims at reaching fully non-perturbative QED), the RIP solver turns out to be a good choice as it correctly calculates the transverse force. Finally, it is noted that the linear energy drop for the RIP solver is ascribed to the longitudinal component of the beam's self-field. In total, this means that the RIP scheme prevents non-physical radiation losses and suppresses the NCI, and therefore allows for the improved study of ultra-relativistic particle beams.

4 Beam-Beam Collisions

During the collision of two counter-propagating particle beams, the particles of one beam will experience the superintense fields of the other. Depending on the beam parameters this can lead to the regime of non-perturbative QED. Up to now, this regime has not been investigated experimentally and only analytical and numerical results have been obtained. Originally, Yakimenko *et al.* have investigated a setup with a head-on collision of the two particle beams[15]. This setup, however, can be optimised: the fields of the opposing beam are not strongest in the middle of the beam but actually slightly radially outwards [cf. Fig. 3(a)]. Thus, in a separate work, we have investigated the beam-beam collision for laterally shifted bunches[1].

Each of the presented simulations has a domain size of $200 \times 300 \times 300$ nm^3. The grid size is $h_x = 0.25$ nm and $h_y = h_z = 1$ nm. The time step is chosen in accordance with the RIP solver, i.e. $\Delta t = h_x/c$. The beams are modelled as Gaussian ellipsoids with

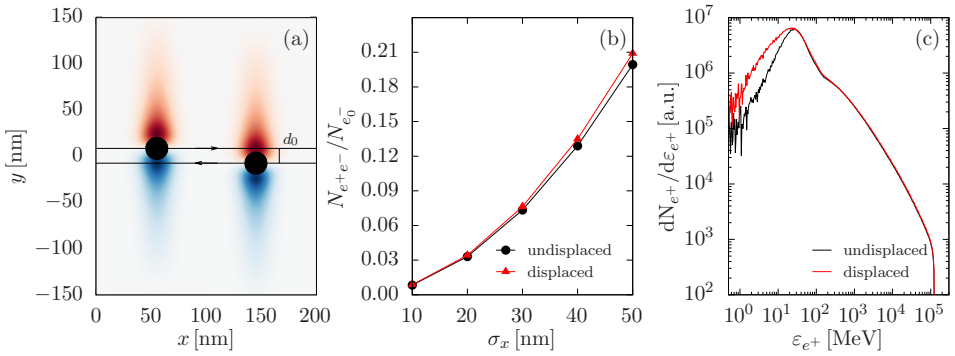


Figure 3. (a) Setup consisting of two laterally shifted electron beams (black dots). Note that the strongest fields (darkest red/blue regions) occur in the distance d_0 from the electron bunches' centre. (b) Pair production yield for several bunch lengths. (c) Positron energy spectra for the electron-electron collisions with 50 nm bunches.

length 10 nm, peak current $I_{\max} = 1.7$ MA and Lorentz factor $\gamma_0 = 2.5 \times 10^5$.

As mentioned earlier, QED effects are modelled using the Monte Carlo method. Two setups were compared, the first consisting of a head-on collision of two electron beams, while for the second one a lateral displacement was introduced. The optimal shift for the maximum field seen by the electron bunch can be calculated from Gauss' theorem; in our case it is $d_0 = 15.9$ nm [cf. Fig. 3(a)].

The interaction of the two bunches yields χ values of up to $\chi \approx 1695$, i.e. we actually consider the non-perturbative QED regime. While 34% of the electrons would reach this regime in the undisplaced case, a slightly lower amount of 33% reach it for the displacement. In contrast, the displaced setup delivers higher yields of emitted photons and electron-positron pairs. The increase in the secondary particle yield, however, can be favourable as information about the conditions during the generation will be imprinted on the particle properties. Experimental signatures may be thus be easier identified in the proposed setup.

Further, the effect of bunch length was investigated. The length of the bunches was varied in the range of 10 nm up to 50 nm in simulations. It is observed that longer bunches not only lead to a general increase in the pair and photon yields, but also slightly favour the yields in the displaced configuration. For instance, the pair yield ratio increases from approx. 4.4% for 10 nm bunches to 4.9% for 50 nm bunches. This is in part due to transverse particle dynamics: Increased transverse momentum of the electrons induces changes in the charge distribution and thus changes in the prevalent electromagnetic fields, leading to said deviations.

Lastly, the energy spectrum for the created positrons was studied. As shown in Fig. 3(c), the displaced setup yields more lower energy positrons than a head-on collision; for higher energies in the range of 0.1-100 GeV the energy spectra of the two configurations are almost indistinguishable.

5 Beamstrahlung-Enhanced Disruption

In a collaboration with Samsonov *et al.*[2], the effect of radiation reaction on particle dynamics for the collision of an electron beam with a positron beam was considered. Due to the prevalent fields during collision, the beam particles are subject to transverse betatron oscillations. The characteristic time for a particle to reach the central axis in dependence of its oscillations frequency and its energy is called disruption time; the so-called disruption parameter D can be used to quantify the distortion of the beam. Generally, a small disruption parameter is desired in order to reduce interference in precision measurements, despite the accompanying increase in luminosity. The main goal of this publication was to give an expression for the disruption parameter even for the case where radiation reaction is relevant. If a particle trajectory is bending at the point of interaction, radiation reaction losses can occur, which is referred to as beamstrahlung in the particle collider community. Additionally, the beamstrahlung photons may induce pair production in the strong electromagnetic fields. Together, both effects can lead to a significant increase in particle number, a process called QED cascading. For collider setups, beamstrahlung is usually an unwanted phenomenon.

In our study, we have considered the collisions of two cylindrical beams for which the initial beam radius r_b and the maximum quantum parameter χ_0 are varied. Here, χ_0 corresponds to the initial field maximum at the cylinder edge seen by an electron/positron. The simulations box has dimensions $116 \times 7.5 \times 7.5\ \mu\mathrm{m}^3$ with a grid size of $h_x = 0.0125\ \mu$m, $h_y = h_z = 0.05\ \mu$m. For the presented simulation both the electron and positron beam have a peak density of $7.4 \times 10^{21}\ \mathrm{cm}^{-3}$, initial radius $r_b = 1\ \mu\mathrm{m}$ and $55\ \mu\mathrm{m}$ length. The electrons are moving in $+x$-direction with momentum $10^5\, m_e c$, while the positrons move in $-x$-direction with the same momentum. The resolution of the simulation is comparatively small in order to account for the strong growth in particle number due to cascading and the accompanying computational effort. The simulation is conducted with the stochastic Monte Carlo QED routine.

The inclusion of QED effects makes a visible difference in the simulation results (cf. Fig. 4). For a value of $\chi_0 = 10$, a high yield of electron-positron pairs is observed. This indicates the formation of a QED cascade. What must further be noted, is that for the simulations with QED, a transverse kink instability arises [cf. Fig. 4(b)]. This might be due to the probabilistic nature of the QED effects that change the originally symmetric particle distribution and therefore act as a seed for instability growth.

In the paper, an analytical scaling for the ratio D/D_0 (D_0 being the uncorrected disruption parameter) was also derived for various regimes. The scaling was found to be in good agreement with the PIC simulations as long as particles are ultra-relativistic. As stated before, for the disruption parameter usually a small value is wanted. This gives some restrictions on beam length and diameter. The simulations, however, show that considering radiation reaction improves the results: if the beam energy is 125 GeV with 1.4 nC charge and radius of 0.01μm, in order to achieve $\chi \approx 1600$, disruption enhancement reaches 55%. In the QED regime with $\chi_0 \gg 1$, we find that

$$\frac{D}{D_0} \approx 30.8 \left(\varepsilon_b[100\ \mathrm{GeV}]\ n_e[10^{21}\ \mathrm{cm}^{-3}]\ r_b[\mu\mathrm{m}]^4\right)^{1/3}\ , \tag{4}$$

where ϵ_b is the beam particle energy and n_e the electron density. This ratio can be used to decide whether radiation effects are important.

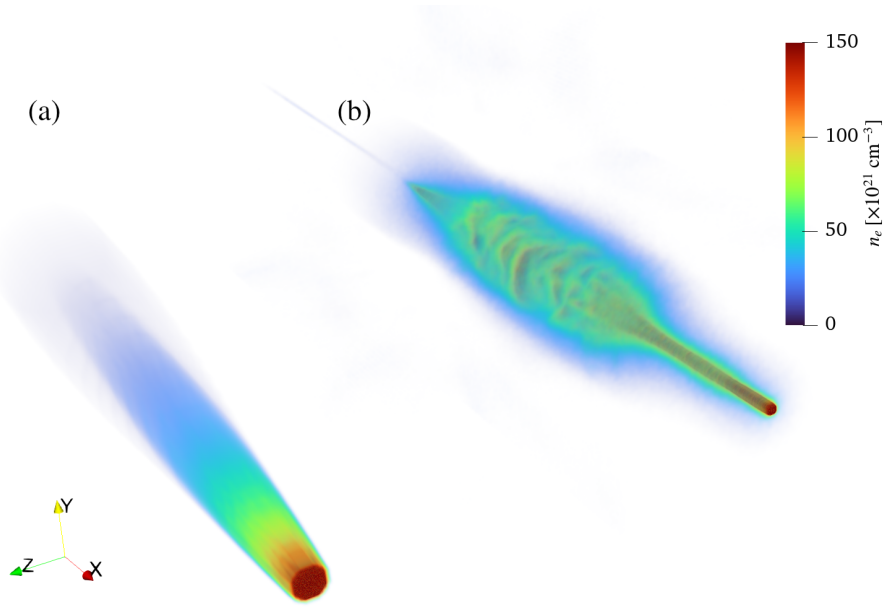


Figure 4. Electron-positron beam collision results (a) without and (b) with QED effects considered. In both cases, $r_b = 1$ μm and $\chi_0 = 10$. Note that in the case with QED effects a kink instability is observed. The colourbar is clipped for better visibility.

6 Concluding Remarks

In our studies concerning quantum electrodynamic processes in supercritical fields we were able to use a novel solver for Maxwell's equations in order to suppress the numerical Cherenkov instability. This allows for a better description of ultra-relativistic particle beams like they are used in beam-beam collisions for higher photon emission and electron-positron pair yield. We have shown that a transverse displacement of the two colliding beams increases the maximum fields observed by the beam particles and thus improves upon the yields compared to a head-on collision. Further, the incorporation of beamstrahlung leads to better beam focusing in the case of beam-beam collisions. Our results generally are in good agreement with analytical estimates. Future work will aim to improve upon those results and further study the regime of non-perturbative QED.

Acknowledgements

The authors would like to thank all the collaborators that have been involved in the various sub-projects. This project was funded in parts by the DFG under project numbers 367991447 and 430078384. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project (qed20) 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. M. Filipovic, C. Baumann, A. M. Pukhov, A. S. Samsonov, and I. Yu. Kostyukov, *Effect of transverse displacement of charged particle beams on quantum electrodynamics processes during their collision*, Quantum. Electron **51**, 807, 2021.
2. A. S. Samsonov, E. N. Nerush, I. Yu. Kostyukov, M. Filipovic, C. Baumann, and A. Pukhov, *Beamstrahlung-enhanced disruption in beam-beam interaction*, New J. Phys. **23**, 103040, 2021.
3. M. Filipovic, C. Baumann, and A. Pukhov, *Suppression of errors in simulated ultra-relativistic bunch propagation using the X-dispersionless Maxwell solver*, Phys. Rev. Accel. Beams 25, 054405, 2022.
4. X. F. Shen, A. Pukhov, and B. Qiao, *Monoenergetic High-Energy Ion Source via Femtosecond Laser Interacting with a Microtape*, Phys. Rev. X **11**, 041002, 2021.
5. K. Jiang, A. Pukhov, and C. T. Zhou, *Magnetic field amplification to gigagauss scale via hydrodynamic flows and dynamos driven by femtosecond lasers*, New J. Phys. **23**, 063054, 2021.
6. L. Reichwein, A. Hützen, M. Büscher, and A. Pukhov, *On the robustness of spin polarization for magnetic vortex accelerated proton bunches in density down-ramps*, Plasma Phys. Control. Fusion **63**, 085011, 2021.
7. A. Pukhov, *Three-dimensional electromagnetic relativistic particle-in-cell code VLPL (Virtual Laser Plasma Lab)*, J. Plasma Phys. **61**, 425-433, 1999.
8. A. Pukhov, *Particle-In-Cell Codes for Plasma-based Particle Acceleration*, CERN Yellow Reports **1**, 181-206, 2016.
9. A. Di Piazza, C. Müller, K. Z. Hatsagortsyan, and C. H. Keitel, *Extremely high-intensity laser interactions with fundamental quantum systems*, Rev. Mod. Phys. **84**, 1177, 2012.
10. N. B. Narozhny, *Expansion parameter of perturbation theory in intense-field quantum electrodynamics*, Phys. Rev. D **21**, 1176, 1980.
11. A. Fedotov, A. Ilderton, F. Karbstein, B. King, D. Seipt, H. Taya, and G. Torgrimsson, *Advances in QED with intense background fields*, arXiv:2203.00019, 2022.
12. A. Pukhov, *X-dispersionless Maxwell solver for plasma-based particle acceleration*, J. Comp. Phys. **418**, 109622, 2020.
13. N. V. Elkina, A. M. Fedotov, I. Yu. Kostyukov, M. V. Legkov, N. B. Narozhny, E. N. Nerush, and H. Ruhl, *QED cascades induced by circularly polarized laser fields*, Phys. Rev. Spec. Top. Accel. Beams **14**, 054401, 2011.
14. C. Baumann and A. Pukhov, *Influence of e^-e^+ creation on the radiative trapping in ultraintense fields of colliding laser pulses*, Phys. Rev. E **94**, 063204, 2016.
15. V. Yakimenko, S. Meuren, F. Del Gaudio, C. Baumann, A. Fedotov, F. Fiuza, T. Grismayer, M. J. Hogan, A. Pukhov, L. O. Silva, and G. White, *Prospect of Studying Nonperturbative QED with Beam-Beam Collisions*, Phys. Rev. Lett. **122**, 190404, 2019.

John von Neumann Institute for Computing
NIC Series

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

John von Neumann Institute for Computing
NIC Series

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

NIC Series Volume 51
NIC Symposium 2022 – Proceedings
29 – 30 September 2022 I Jülich, Germany
edited by M. Müller, Ch. Peter, A. Trautmann (2022), v, 450 pages
ISBN: 978-3-95806-646-5

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 11th NIC Symposium are from projects that have been supported by the supercomputers JUWELS and JURECA Booster in Jülich. They cover selected topics in the fields of Astrophysics, Biophysics, Chemistry, 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 51

ISBN 978-3-95806-646-5