John von Neumann Institute for Computing

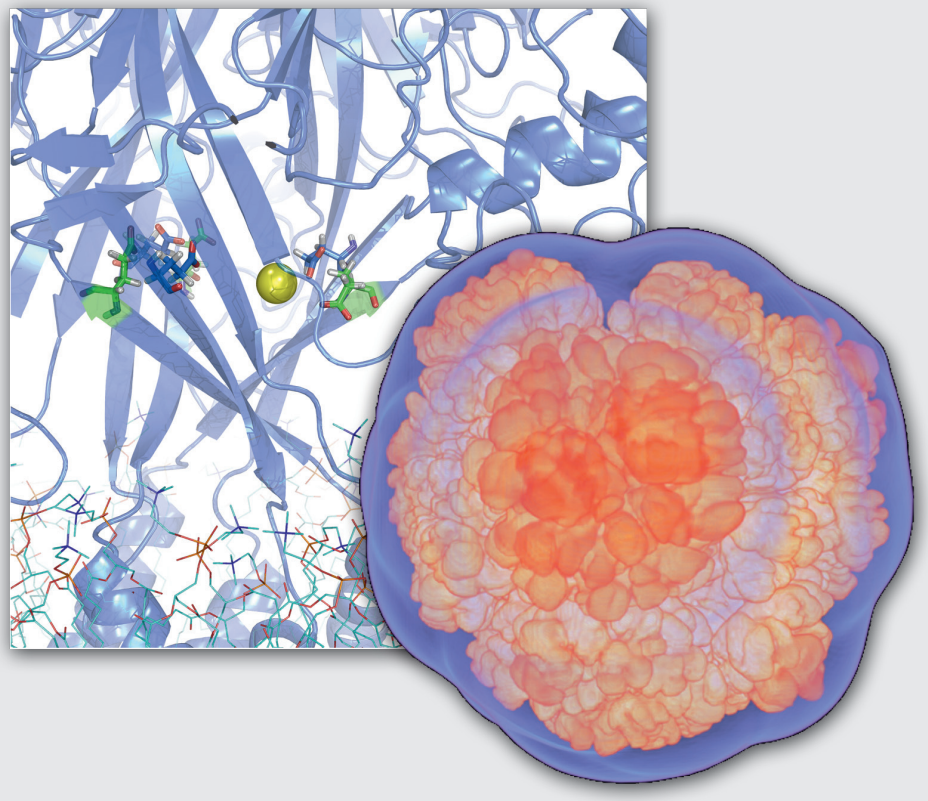
NIC Symposium 2012

25 Years HLRZ/NIC

7 – 8 February 2012 | Jülich, Germany

K. Binder, G. Münster, M. Kremer (Editors)

Proceedings



NIC Series

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 45

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

NIC Symposium 2012 – Proceedings

25 Years HLRZ / NIC

7 - 8 February 2012 | Jülich, Germany

K. Binder, G. Münster, M. Kremer (Editors)

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

NIC Series Volume 45

ISBN 978-3-89336-758-0

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

Publisher and Distributor:	Forschungszentrum Jülich GmbH Zentralbibliothek 52425 Jülich Phone +49 (0) 24 61 61-53 68 · Fax +49 (0) 24 61 61-61 03 e-mail: zb-publikation@fz-juelich.de Internet: http://www.fz-juelich.de/zb
Cover Design:	Jülich Supercomputer Centre, Forschungszentrum Jülich GmbH
Printer:	Druckerei Gehler, Düren (Cover) Grafische Medien, Forschungszentrum Jülich GmbH
Copyright:	Forschungszentrum Jülich 2012

We thank R. Böckmann (Erlangen-Nürnberg) and W. Hillebrandt (Garching) for the pictures used in the cover design.

Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 45

ISBN 978-3-89336-758-0

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

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

Preface

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

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

Manfred Kremer
Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: m.kremer@fz-juelich.de

Since 25 years the John von Neumann Institute for Computing (NIC), the former "Höchstleistungsrechenzentrum", plays a pioneering role in supporting research in computational science at the fore-front, by giving large grants of computer time to carefully selected research projects. The scope of these projects ranges from fundamental aspects of physics, such as the physics of elementary particles and nuclear physics, astrophysics, statistical physics and physics of condensed matter, computational chemistry and life sciences, to more applied areas of research, such as the modelling of processes in the atmosphere, materials science, fluid dynamics applications in engineering, etc. Use of the supercomputer resources that the Jülich Supercomputing Centre (JSC) provides for these research projects hence has an important impact both on the progress of basic science and on the needs of the society, e.g. by progress made for better materials in research addressing the needs of energy supply information technologies, or for applications in medicine, for instance.

The research projects that are funded by computer time grants on the JUGENE and JUROPA supercomputers of the JSC are selected by a peer review process, managed by a panel in which leading scientists from several European countries participate. Scientific excellence is the sole criterion which controls the ranking of the projects in all these different fields mentioned above. The present book, which appears in the framework of the biannual NIC Symposia series, continues a tradition started 10 years ago, to present selected highlights of this research to a broader audience. Due to space restrictions, only a small number of the research projects that are carried out at the NIC can be presented in this way.

Projects that stand out as particularly excellent are nominated as "John von Neumann Excellence Project" by the review board.

In 2010 this award was given to A. Muramatsu (Stuttgart) for his project on "Quantum Monte Carlo studies of strongly correlated systems".

In 2011, two such awards were given to C. Hoelbling (Wuppertal) for his project "Computing B_K with 2+1 flavours at the physical mass point", and another one to W. Paul (Halle) for "Long range correlations at polymer-solid interfaces".

The procedures adopted by the NIC to identify the scientifically best projects for the allocation of computer time are of the same character as those used by organisations founded more recently, such as (in Germany) the Gauss Centre for Supercomputing (GCS), an alliance of the three German national supercomputing centres in Jülich, Garching and Stuttgart, and (in Europe) the Partnership for High Performance Computing in Europe (PRACE). In fact, the successful experience of NIC has served as a role model for this joint European effort, which now is hosted by France, Italy and Spain in addition to Germany. Actually, by now 22 of the 27 member states of the EU have joined PRACE in order to profit from the advances that supercomputing offers for basic science and technology. In fact, in the last two years a significant fraction of the supercomputer resources at JSC have been given to PRACE, and have fertilised research on an even broader range of topics. The first PRACE projects have been finished since a couple of months, and very important results can be expected.

The John von Neumann Institute for Computing operates on the basis of a cooperation agreement between three research centres of the Helmholtz association, namely Forschungszentrum Jülich, Deutsches Elektronen-Synchrotron (DESY) Hamburg and the Gesellschaft für Schwerionenforschung Darmstadt (GSI). A very important ingredient of this research is that these institutions sponsor NIC research groups, where specific know-how on computer simulation techniques relevant for topical research at supercomputers is developed further. The present volume contains brief reports from these groups; at Jülich the NIC research group devoted to "Computational Materials Physics", headed by M. Müser from the Universität des Saarlandes at Saarbrücken, has started its work on April 1st, 2011.

We thank the authors for representing the NIC with their contributions. Special thanks go to the Jülich Supercomputing Centre JSC for providing the high quality computing environments and user service, which were vital to the successful realisation of the projects presented in this book. Thanks go also to Ms. Martina Kamps, who compiled the contributions and produced the present high quality book, and to Walter Nadler, who took care of the communication with the printing services at Forschungszentrum Jülich. Already here we want to thank Walter Nadler and Ms. Elke Bielitza for their help in organising the 6th NIC-Symposium.

Jülich, February 2012

K. Binder G. Münster M. Kremer

Contents

Supercomputing at Scale – Architectural Evolution at Jülich Supercomputing Centre
N. Eicker, St. Graf, W. Gürich, H. Kleimann, B. Orth, U. Schmidt, M. Stephan, E. Suarez, L. Wollschläger, Th. Lippert **1**

The NIC Research Groups

The Chiral Transition, Deconfinement, and the Equation of State of QCD
F. Karsch, W. Söldner **15**

Computational Materials Physics at the JSC
M. H. Müser **25**

QCD Quantum Corrections from Computer Simulations
K. Jansen, R. Sommer **33**

Astrophysics

Introduction by P. L. Biermann **41**

Generation of Strong Magnetic Fields via the Small-Scale Dynamo During the Formation of the First Stars
R. Banerjee, S. Sur, Ch. Federrath, D. R. G. Schleicher, R. S. Klessen **43**

The Progenitor Systems of Thermonuclear Supernovae
W. Hillebrandt, M. Fink, M. Kromer, R. Pakmor, F. K. Röpke, I. R. Seitenzahl, S. A. Sim **53**

High Resolution Simulations of Planet Formation Processes: Gravoturbulent Planetesimal Formation on JUGENE
H. Klahr, A. Johansen **61**

Computational Biology and Biophysics

Introduction by W. Wenzel **71**

pH Dependence of ASIC1a Dynamics on a Sub-Microsecond Time Scale
K. Pluhackova, R. Böckmann **73**

Memory Consumption of Neuronal Network Simulators at the Brain Scale
S. Kunkel, M. Helias, T. C. Potjans, J. M. Eppler, H. E. Plesser, M. Diesmann, A. Morrison **81**

Structure, Specificity and Mechanism of ATP Synthase Membrane Motors
J. D. Faraldo-Gómez **89**

Mesoscale Simulations of Human Blood Flow: From Red Blood Cell Elasticity and Interactions to Blood Rheology
D. A. Fedosov, G. Gompper **97**

Chemistry
Introduction by W. Klopper **105**

***Ab initio* Molecular Dynamics Simulations of the Adsorption on Complex Surfaces**
A. Groß **107**

Mechanochemistry of Covalent Bond Breaking from First Principles Simulations
P. Dopieralski, J. Ribas-Arino, D. Marx **115**

Internal Dynamics in Endohedral Metallofullerenes as Studied by *ab initio* Born-Oppenheimer Molecular Dynamics
A. A. Popov **123**

Simulations of Electron Transfer and Electron Transport in Molecular Systems at Metal and Semiconductor Substrates
P. B. Coto, Ó. Rubio-Pons, R. Härtle, M. Bockstedte, M. Thoss **131**

Elementary Particle Physics
Introduction by G. Münster **141**

Searching for New Physics in Neutral Kaon Mixing
Ch. Hoelbling **143**

Nuclear Physics from Lattice Simulations
U.-G. Meißner **151**

Finite Temperature Lattice QCD with Wilson Quarks
S. Borsányi, Z. Fodor, Ch. Hoelbling, S. Katz, St. Krieg, K. Szabó, B. Tóth **159**

Probing the $N_f = 2$ QCD Phase Transition with Light Wilson Fermions
B. B. Brandt, O. Philipsen, H. Wittig, L. Zeidlewicz **167**

Materials Science

Introduction by R. O. Jones **175**

Understanding Molecular Electronics and Spintronics from First Principles Simulations
N. Atodiresei, V. Caciuc, P. Lazić, St. Blügel **177**

Quantum-Chemical Calculations on Ion-Permeable Membranes
M. Lumeij, M. Gilleßen, B. Eck, R. Dronskowski **189**

Graphene Building Blocks on Si (111): Going Towards Graphene-Based Nanoelectronics with the Help of Large Scale *ab initio* Simulations
J. Kiss, Th. Gruhn, C. Felser **197**

Atomic Structure of AgInSbTe Phase Change Materials: Simulations and a Model of Crystallization
J. Akola, R. O. Jones **207**

The Running-In of Hydrogenated Diamond-Like Carbon Coatings
L. Pastewka, M. Moseler **215**

First Principles Numerical Study of the Magnetic Properties of Transition-Metal Alloy Nanoparticles
L. E. Díaz-Sánchez, G. M. Pastor **223**

Condensed Matter

Introduction by K. Binder, A. Muramatsu **231**

Monte Carlo Simulation Studies of Interfacial Free Energies in Colloidal Suspensions
D. Deb, D. Wilms, A. Winkler, P. Virnau, K. Binder **235**

Coupled Lattice Boltzmann and Molecular Dynamics Simulations on Massively Parallel Computers
J. Harting, St. Frijters, F. Janoschek, F. Günther **243**

Computational Studies of Insulating Magnetic Oxides
M. Ležaić, P. Mavropoulos, K. Z. Rushchanskii, I. Slipukhina, St. Blügel **251**

Quantum Spin Liquid in Correlated Fermions on a Graphene-Like Structure
Z. Y. Meng, T. C. Lang, S. Wessel, F. F. Assaad, A. Muramatsu **259**

KKRnano: Precise Density Functional Calculations for Thousands of Atoms
R. Zeller, A. Thiess, M. Bolten, P. H. Dederichs, St. Blügel **267**

Computational Soft Matter Science

Introduction by K. Kremer **275**

Grafted versus Non-Grafted Polymer Adsorption

M. Möddel, M. Bachmann, W. Janke **277**

Structure and Dynamics at Polymer-Solid Interfaces: A Molecular Dynamics Simulation of 1,4-Polybutadiene at Graphite

W. Paul, L. Yelash, P. Virnau, K. Binder **285**

Earth and Environment

Introduction by U. Hansen **293**

Direct Numerical Simulation of Turbulent Mixing in the Planetary Boundary Layer

J. P. Mellado, C. Ansorge **295**

Simulation of Processes in Terrestrial Systems

J. Vanderborght, O. Ippisch, M. Blatt, St. Kollet, S. Rahman, T. Aljazzar, R. Maxwell, N. Schröder, M. Javaux, H.-J. Hendricks-Franssen, W. Kurtz J. van der Kruk, M. Oberröhrmann, A. Klotzsche, H. Vereecken **303**

Modelling of Atmospheric Processes

M. König, R. Wolke, O. Knoth, E. Renner **313**

Computer Science and Numerical Mathematics

Introduction by M. Griebel **321**

Highly Parallel Geometric Multigrid Algorithm for Hierarchical Hybrid Grids

B. Gmeiner, T. Gradl, H. Köstler, U. Rüde **323**

Comparable Corpora in Wikipedia Text for Machine Translation

J. Xu, C. Kennington, Č. Przywara, L. Wanzare **331**

Hydrodynamics and Turbulence

Introduction by D. Kröner **339**

Numerical Investigation of Reacting and Non-Reacting Flows in Gas Turbine Related Configurations

M. Di Domenico, E. Ivanova **341**

Walberla: Simulations of Complex Flows on Supercomputers

J. Götz, St. Donath, Ch. Feichtinger, K. Iglberger, H. Köstler, U. Rüde **349**

A Hybrid Aeroacoustic Prediction Method for Non-Reactive and Reactive Flows
G. Geiser, M. Meinke, W. Schröder **357**

Impact of the Heterogeneity of Random Sphere Packings on Hydrodynamic Dispersion
S. Khirevich, U. Tallarek **367**

Finite Size Particles in Homogeneous Turbulence
M. Uhlmann, T. Doychev **377**

Computational Plasma Physics
Introduction by K.-H. Spatschek **385**

Plasma Formation and Ionization Dynamics in Intense Laser-Droplet Interaction
T. Liseykina, D. Bauer **387**

Thomson Scattering on Inhomogeneous Targets
P. Sperling, T. Liseykina, D. Bauer, R. Redmer **395**

Supercomputing at Scale
–
Architectural Evolution at Jülich Supercomputing Centre

Norbert Eicker, Stephan Graf, Wolfgang Gürich, Hanna Kleimann, Boris Orth, Ulrike Schmidt, Michael Stephan, Estela Suarez, Lothar Wollschläger, and Thomas Lippert

Institute for Advanced Simulation, Jülich Supercomputing Centre
Forschungszentrum Jülich, 52425 Jülich, Germany
and Jülich-Aachen Research Alliance (JARA)
E-mail: th.lippert@fz-juelich.de

Supercomputing at scale has become the decisive challenge for users, providers and vendors of leading supercomputer systems. On next-generation systems, approaching Exascale by the end of the decade, we will be confronted with millions of cores, given the apparent saturation of core frequency. We argue here that a rethinking of the interaction between users, centres and manufacturers is required, and we report on co-development activities at the Jülich Supercomputing Centre (JSC). We describe the system design of the new Blue Gene/Q and early application porting to this architecture, we present new ideas to tackle the imminent I/O gap, we report on the European DEEP cluster project and we speculate how to utilise future knowledge about the human brain's functionality for more energy efficient computing.

1 Introduction

Since mid of 2009, the IBM Blue Gene/P system named JUGENE at the JSC is pioneering scalable Petaflop/s supercomputing in Jülich, Germany and Europe. For more than two years, JUGENE held the world record in terms of scalable hardware, with a number of 294.912 cores that can be used concurrently and in a scalable manner for quite a variety of codes from science and engineering. Parallel to the launch of JUGENE, the JSC started to operate its 300 Teraflop/s cluster system JUROPA, aiming at most flexible and efficient general purpose parallel computing as well as highest possible scalability. It is a remarkable fact that JUROPA, featuring more than 25.000 Intel Nehalem CPU cores, since mid of 2009 is the most efficient cluster system in terms of Linpack efficiency ever making an appearance in the TOP500 list of the 500 fastest computers worldwide. On top of this achievement, its cluster operating system Parastation, jointly developed by JSC and the Munich HPC software company ParTec, was the first to achieve MPI-scalability up to more than 25.000 MPI tasks on a components-off-the-shelf cluster system with a full Linux OS, as has been elucidated by the successful cosmological Millennium XXL simulations[1].

Together, JUGENE and JUROPA have perfected the dual system strategy of the JSC that was started in 2005. Its goal is to most efficiently serve the application portfolio of the users of the Jülich Research Centre, the John von Neumann Institute for Computing in Germany and since mid of 2010 the Partnership for Advanced Computing in Europe (PRACE), within the given budget. In this spirit, the JSC chooses to optimise for highest scalability (JUGENE) and highest flexibility and efficiency (JUROPA), refusing to put the entire budget on one architecture just for a high TOP500 ranking.

It is certainly fair to say that the users of the Jülich facilities and the members of the support teams of JSC's simulation laboratories are among the most competent and experienced scientists as far as supercomputing at scale is concerned. However, the current challenges lie ahead: given a roughly constant speed of future compute cores[a], the desired increase in supercomputer performance can only be achieved by accepting extremely parallel architectures with millions of cores.

The Jülich Supercomputing Centre is convinced that this challenge requires a fundamental re-thinking of its traditional mission and interaction with the users. Whereas, for a while now, the new simulation laboratories at the JSC combine forces of users, application specialists, mathematicians and computer scientists for code optimisation and efficient implementation, the application driven technology development, in co-development with leading companies, is more and more becoming a crucial element of the centre's mission as well. We have to strengthen both the interaction with companies in getting early experience with new technology and the active application-driven design of novel architectures. In keeping with this, the Jülich Research Centre has launched two technology development laboratories within the JSC: the Exascale Innovation Center (EIC) is a collaboration with IBM to drive the design of IBM's future Exascale systems involving application aspects as computation and communication scalability or I/O requirements already in their early design phases. The Exa-Cluster Laboratory (ECL) together with Intel and ParTec aims at the creation of next-generation components-off-the-shelf HPC cluster technology focussing on both hardware and software.

In this contribution, we report on activities of the JSC within both labs, the EIC and the ECL. In section 2, we describe the system design of the successor of the present Blue Gene/P system JUGENE, named Blue Gene/Q, along with early application scaling on a prototype hardware. In section 3, we present new ideas fighting the imminent I/O performance bottleneck, in section 4, the concept of the Dynamical Exascale Entry Platform (DEEP) is introduced, a project funded by the EU DG INFSO with the goal to pave the way towards scalable systems of the Exascale class which should be largely based on commodity components. In section 5, we give a short outlook on more speculative ideas of creating cognitive computing devices within the planned "Human Brain Project", currently applying for a EU Flagship grant, followed by some concluding remarks.

2 Blue Gene/Q

In November 2011, IBM has announced the third generation of its ultra-scaling Blue Gene supercomputing family as a generally available product. As of per rack, the Blue Gene/Q will be 15 times faster than the existing Blue Gene/P: thus, five racks of Blue Gene/Q can replace the 72-rack JUGENE system. A first Blue Gene system consisting of eight racks will be installed in the first half of 2012 at the Jülich Supercomputing Centre. It will be followed by an upgrade of JUGENE available to German, European and JARA-users through the John von Neumann Institute for Computing (NIC), the Partnership for Advanced Computing in Europe (PRACE) and the Jülich-Aachen Research Alliance (JARA).

[a]Some applications like deterministic protein folding are intrinsically limited in scalability. Special purpose systems like ANTON by D. E. Shaw attempt to overcome these limitations.

2.1 System Design

To achieve this big leap in performance, IBM designed its new PowerPC A2 processing architecture. Each processor includes 16 compute cores (as compared to four used in Blue Gene/P, the present system) plus a core allocated to operating system administrative functions and a spare core. As every core supports 4-way simultaneous multithreading (SMT), a Blue Gene/Q node can run up to 64 independent hardware threads. Every core is assisted by a 4-pipe double precision floating point unit (SIMD).

Blue Gene/Q incorporates novel architectural advances that promise to contribute to the system's performance and will help users to simplify programming their codes on such highly scaling many-core systems.

Hardware-based speculative execution capabilities facilitate efficient multi-threading for long code sections, even those with potential data dependencies. If conflicts are detected, the hardware can backtrack and redo the work with minimal effects on the applications' performances. Hardware-based transactional memory helps programmers avoid the potentially complex integration of locks and eliminate bottlenecks caused by deadlocking when threads become stuck during the locking process. It also supports efficient and effective multi-threading while reducing the need for complicated programming.

The L1 prefetcher can run in normal stream prefetching mode which adaptively balances resources to prefetch L2 cache lines in response to observed memory traffic. But in addition, it can also use four list-based prefetching engines in order to record memory access patterns in arbitrarily long code segments on a first iteration of a loop and to play back this pattern for subsequent iterations. On subsequent passes, this list is adaptively refined for extra cache misses and can be activated by programme directives.

The network has been changed from a 3D torus to a 5D torus topology which promises to give several advantages over Blue Gene/P: (i) A reduced worst-case latency of 3 μsec for point-to-point communication and $\sim$6 μsec within collectives and barrier; (ii) an improved tradeoff of nearest neighbour and bisection bandwidths; (iii) the bisectional bandwidth of a 64k partition is 19 times higher than on a Blue Gene/P system; (iv) a much more flexible partitioning into independent, non-interfering sub-machines is now possible; (v) the hardware includes direct support for MPI collective reduce and all-reduce operations so that single pass floating point reductions can be executed with near-link bandwidth; (vi) with flexible configurability in its network, Blue Gene/Q can by-pass failed lasers without disrupting a running application, and (vii) the new networking hardware supports off-loading of the I/O traffic from the compute cores. A comparison of node and network features between the three generations of Blue Gene can be found in Table 1.

Designed with a small footprint and low power requirements, Blue Gene/Q is ranked as the number-one most energy-efficient supercomputer in the world by the Green500 list (Nov. 2011). Engineered with fewer moving parts and built in redundancy, Blue Gene/Q has proven to be extremely reliable. Fig. 1 shows the packaging hierarchy of the system, which is similar to the one of Blue Gene/P, with the most important change the compute cards being directly water-cooled.

2.2 Porting Applications to Blue Gene/Q

Application specialists from JSC's simulation laboratories had access to a Blue Gene/Q prototype system through the Exascale Innovation Center (EIC) – a collaboration with IBM

	Blue Gene/L	**Blue Gene/P**	**Blue Gene/Q**
Compute Nodes			
Processor	32-bit PowerPC 440	32-bit PowerPC 450	64-bit PowerPC (A2 Core)
Processor Frequency	700 MHz	850 MHz	1.6 GHz
Cores	2	4	16 (4-way SMT)
Peak Performance (per Node) [Gigaflop/s]	5.6	13.6	204.8
Coherency	Software Managed	SMP	SMP + transactional memory
L1 Cache (per Core)	32 KB	32 KB	16/32 KB
L2 Cache (prefetch per Core/Thread)	14 stream	14 stream	16 stream, List-based
L3 Cache size (shared, per Node)	4 MB	8 MB	32 MB
Main Store/Node (same for I/O Node)	512 MB or 1 GB	2 GB or 4 GB	16 GB
Main Store Bandwidth	5.6 GB/s (16B wide)	13.6 GB/s (2*16B wide)	43 GB/s
Torus Network			
Topology	3D	3D	5D
Bandwidth	6*2*175 MB/s = 2.1 GB/s	6*2*425 MB/s = 5.1 GB/s	32 GB/s
Hardware Latency	200 ns (32B packet)	100 ns (32B packet)	80 ns (32B packet)
(Nearest Neighbour)	1.6 μs (256B packet)	800 ns (256B packet)	640 ns (256B packet)
Hardware Latency (Worst Case)	6.4 μs (64 hops)	5.5 μs (64 hops)	3 μs
Per Rack			
Peak Performance [Teraflop/s]	5.7	13.9	209
Sustained Performance (Linpack) [Teraflop/s]	4.6	11.9	$\gtrsim$170
Power	$\sim$20 kW	$\sim$32 kW	$\sim$100 kW
Power Efficiency [Gigaflop/s/W]	0.23	0.37	2.1

Table 1. Comparison of node and network features between the three generations of Blue Gene.

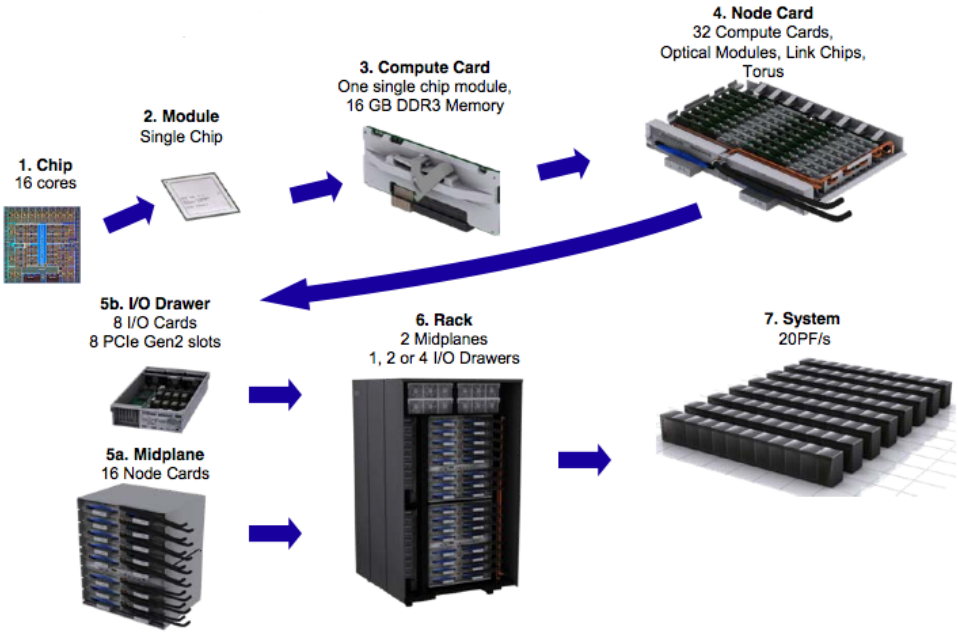


Figure 1. Blue Gene/Q packaging hierarchy[2].

to drive the design of future Exascale systems. The prototype system was procured by the JSC to gain early experience with porting a selected set of applications to the new platform. Since the step from Blue Gene/L to Blue Gene/P was merely an evolutionary one as far as system design was concerned, porting applications was relatively straightforward. Since Blue Gene/Q introduces fundamentally new design elements like the 5D torus network and the many-core processor, application porting might become more demanding.

Four applications have yet been successfully ported: PEPC (Pretty Efficient Parallel Coulomb-Solver), MP2C (Massively Parallel Multi-Particle Collision), Lattice QCD, and Lattice Boltzmann D2Q37 code[3].

The left canvas of Fig. 2 demonstrates the intra-node scaling of D2Q37. As one can see, having two or even four threads per core exploited leads to a significant speed-up of the application's performance. The major experience – also relevant for other applications – is the fact that it is not possible to utilise the full performance of a Blue Gene/Q compute node without running at least two threads per core. For D2Q37 the 5D torus shows a nearly linear scaling behaviour (see Fig. 2) up to 128 nodes which can also be expected for other applications.

Blue Gene/Q is the first step into the area of many-core systems, a development that we will also see from other vendors in the future. Although Blue Gene/Q can still be used in the MPI-only style, hybrid programming models with a mixture of MPI and OpenMP or pthreads are highly recommended and might become mandatory in the future, in particular when the application's memory footprint is a key factor. Novel features in Blue Gene/Q like hardware-based transactional memory and hardware-based thread level speculation will support users in dealing with such challenges.

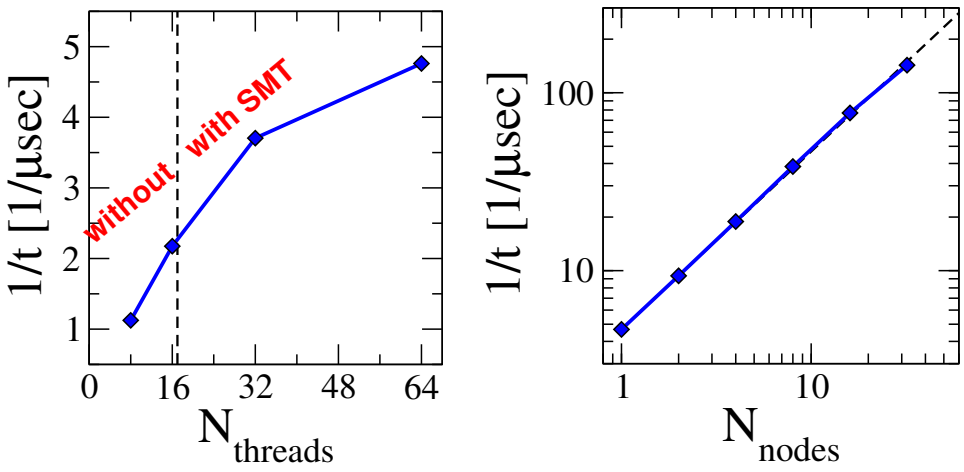


Figure 2. Left: Intra-node scaling of D2Q37. Right: Strong inter-node scaling of D2Q37.

3 New Storage Concepts for Exascale Architectures

It is a common perception that the most important properties of a supercomputer are its peak and sustained performances. The storage system in the background with its capacity and giant throughput is equally important, however. In the last decade, the gap between I/O performance and compute power is steadily increasing. At the Jülich Supercomputing Centre, for example, the relation (GB/s per Teraflop/s) has been 1:4 in 2002 (JUMP: 1.4 GB/s to 5.8 Teraflop/s) and is now 1:15 (JUGENE: 66 GB/s to 1002.7 Teraflop/s). This implies that today the jobs on the supercomputer must wait longer to save their data to the file system.

A standard criterion to estimate the bandwidth necessary for a supercomputer is to demand that a certain amount of memory of the full system is written to disk in a given time. To give an example, for the procurement of the first-phase Blue Gene/Q system requirements for the file system I/O had to be defined. There are 8 I/O nodes per rack, each featuring a dual 10GbE port with an aggregated bandwidth of 1.5 GB/s. So the maximum throughput of a rack is 12 GB/s and of the full system it is 96 GB/s. To write 50% of the main memory of one rack (1024 compute nodes with 16 GB RAM per node) on disk it will take approximately 12 minutes. Our worst-case requirement for an acceptable supercomputer storage system is to write 50% of the full system main memory to the storage in 15 minutes. For the new Blue Gene/Q system this implies that a bandwidth of $50\,\% \times 131\,\mathrm{TB}/900\,\mathrm{s} = 73\,\mathrm{GB/s}$ is required.

The standard approach to achieve this throughput rate is to use an aggregation of ordinary hard disks. In order to increase the throughput, one just has to increase the number of disks. But this will eventually lead into a dead end. Today we use about 7000 disks to get 60 GB/s I/O performance for the JUGENE, a 1-Petaflop machine. One to two disks are failing per week. An Exascale system with an adequate storage attached will need at least a hundred times more disks. It will not be manageable just to change the broken disks, not to mention the risk of loosing data.

As growth in compute power is much faster than speedup in storage technology, it is evident that new storage concepts have to be developed in order to meet the demands of future Exascale systems. The JSC together with IBM and CSCS is engaged in research and development activities to tackle this problem. So far, the JSC has started the PRACE (Partnership for Advanced Computing in Europe)-funded project JUNIORS (JUelich novel I/O Research System). In this project, Blue Gene/Q infrastructure will help to investigate using flash memory cards as a kind of "cache" between the I/O nodes and the ordinary disk storage. The advantage of the flash cards is to handle up to 200.000 I/O operations per second which is an equivalent for a very low latency, while an ordinary hard disk can only perform about 200 I/O operations per second.

So far we have installed the JUNIORS prototype cluster consisting of 4x86 systems each with two Fusion-io ioDrive Duo 320 GB SLC installed. The bandwidth of the flash card is 1.5 GB/s. The cumulated performance of these four nodes should be 12 GB/s, a similar value as given on eight Blue Gene/Q I/O nodes. In addition, there are ordinary disk-based storage subsystems attached to the prototype cluster. To manage these two kinds of devices, flash and disk, the newest version of IBM's Global Parallel File System (GPFS) is used. The software stack includes features to create different kinds of storage pools and to implement placement and migration policy rules. A concept will be modelled, where newly-created files are started on flash, and GPFS will migrate the files to disk in the background automatically. This concept will be adopted with real Blue Gene/Q hardware as soon as it is available.

But I/O hardware is not the only bottleneck I/O-dominated supercomputing is facing. Also the software stack has several weak points. To get the maximum performance from the file system, the user must know exactly how to achieve it. It depends on the best block size to write or read the data, to bundle the I/O into one or a couple of files and of course on the job's algorithm. The problem is that the file systems are not optimised for I/O patterns typical for supercomputing. Therefore, several workarounds have been created in the community. The JSC has developed the tool SIONlib[4] which is internationally recognised. With SIONlib, users can readily achieve the maximum I/O performance for their jobs without need for concern about how large the job is running on JUGENE. It compensates for the fact that the underlying file system itself is incapable of optimising the I/O for throughput.

4 A Dynamical Exascale Entry Platform (DEEP)

Today clusters are the dominant architecture in High Performance Computing (HPC). They can profit directly from advances in mainstream computing, the reason for the success of the concept. Due to their modular setup it is straightforward to replace, upgrade or add single components and thus to improve the availability or the overall performance of such systems. With the 300 Teraflop/s production system JUROPA//HPC-FF installed in 2009, cluster technology has become the backbone of general-purpose supercomputing at JSC, too.

With Petascale systems in production today, the next challenge of supercomputer numerology is to reach Exascale by the end of the decade. This goal poses enormous challenges. For example, most probably using general-purpose CPU will no longer be competitive with more specialised solutions as provided by the Blue Gene/Q processor or inte-

grated CPU-GPGPU processors[b] expected to be available until the middle of the decade. Therefore, it becomes necessary to reconsider the idea of cluster architectures for HPC in order to ensure competitiveness of the components-off-the-shelf-idea in the future.

For this reason, the requirements set by applications capable to make use of Exascale systems have to be investigated. The most crucial feature of HPC-architectures is not the highest performance of a single component but the exactly right balance of all building-blocks. For example, most HPC-systems do not rely on the highest performing CPU available at a given time but instead use a compromise that is more balanced as to the available memory and network-bandwidths, that shows higher energy-efficiency or optimises the price-performance ratio.

While 4 of the top 10 systems in the current TOP500 list[5] are already equipped with accelerator-cards, the architecture of these accelerated clusters[c] suffers from severe limitations concerning the balance between total raw compute speed and memory bandwidth or communication, and, thus, scalability. This is due to the fact that the accelerators are neither directly connected to the cluster node's main memory nor capable to directly send or receive data from their local memory. Eventually, data has to be copied back and forth frequently between main-memory and the accelerator's memory wasting the scarce resource of bandwidth on the node's system bus. At the same time, latency of communication between the actual compute elements is significantly increased.

A promising escape from these dilemmas is to build a cluster of accelerators. In such a concept, a node consists of an accelerator only, accompanied by some memory and a high-speed interconnect. In particular, programmes running on the accelerator-cores should be capable to initiate communication operations without the support of a commodity processor. An early example of this concept was the QPACE system[7] that was ranked #1 in the Green500 list[6] in June 2010. Of course, this concept has limitations, too. Besides the problem of finding accelerators that are capable to run autonomously, they might not be flexible enough to drive a high-speed interconnect efficiently. Furthermore, the gain introduced by the direct connection between the compute-element and the interconnect-fabric might be diminished by the fact that accelerators – as highly specialised devices – are not effective when running general purpose codes. Thus, a radically new concept is required for cluster systems in order to render them competitive at the Exascale level.

Talking about Exascale, it is crucial to discuss the effects of parallelism on scalability. An early theorem in this field is known as Amdahl's Law[8], stating that scalability of parallel programs is limited by its sequential part. For example, an application with a sequential portion of just 1% will have a speedup limited to 100 independent of the number of processors used. Since at least some applications are able to scale on Blue Gene-systems with $\mathcal{O}(100000)$ processors, Amdahl's description obviously has some flaws. In fact, the behaviour of such applications is better described by Gustafson's Law[9]. While in Amdahl's Law it is assumed that putting a programme on a larger machine will leave the problem-size untouched, in reality it is scaled following to the capabilities of the machine. Of course, Gustafson is only valid, if it is possible to implement the parallel part of the problem in a scalable manner.

[b] GPGPU stands for General-purpose Graphical Processing Units – these are highly optimized floating-point units.

[c] We will distinguish between accelerated clusters, i.e. classical clusters with nodes equipped with accelerators, and clusters of accelerators.

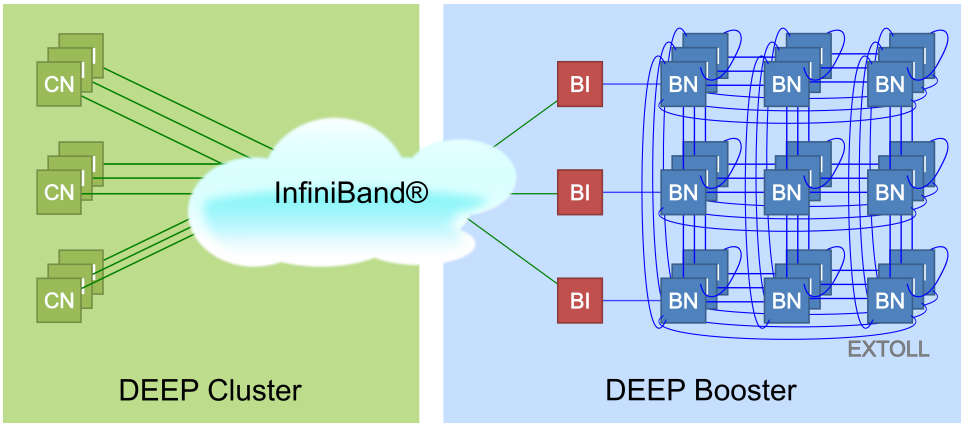


Figure 3. Sketch of the DEEP architecture.

A different view on the question on how to reach Exascale might be taken by looking at applications and their inherent scalability. Unless core speed will not be increased considerably, by the end of the decade, HPC-systems undoubtedly will show a much higher degree of parallelism than today's already massively parallel ones. Thus, the scalability of applications will play a crucial role in order to exploit the computational power of such systems. As an activity of the Exa-Cluster Lab at JSC, an analysis of JSC's application portfolio has been carried out. As a basic result, we found that the applications can be sorted in two classes: (i) Highly-scalable codes using regular communication patterns. These are the codes that are able to exploit JSC's Blue Gene/P system. (ii) Less-scalable but significantly more complex codes. Most often these codes require complicated communication patterns constraining the scalability of these applications on clusters.

A closer analysis of the second class unveils that some of them have highly scalable kernels, too. In principle, they should be able to take benefit from highly-scalable machines as well. Nevertheless, in analogy to the serial element in Amdahl's Law their scalability is limited by the least scalable kernel. In the future, the second class is expected to grow due to increasing parallelism of the systems and more complex simulations. Therefore, efficient support of such applications on future systems will become mandatory.

Given all these arguments, future cluster architectures will require a hierarchy of ingredients in order to provide enough flexibility to run complex applications. The concept of DEEP (Dynamical Exascale Entry Platform) as sketched in Fig. 3 foresees a cluster element with nodes (CN) connected by a highly flexible switched network labeled as InfiniBand here. It is accompanied by a so-called booster element. The booster element is build from booster nodes (BN). Each BN hosts an accelerator based on Intel's MIC (Many Integrated Cores) architecture[10] capable to run its own operating-system autonomously. The BNs are interconnected by a highly scalable torus-network using the EXTOLL technology[11] sketched as blue lines in the figure. Booster Interfaces (BI) connect the booster to the cluster's fabric. This proposal provides several advantages: (i) The cluster element allows to run complex kernels. There is no restriction to run parallel kernels with complicated communication-patterns or highly irregular codes. (ii) The booster element

runs highly-scalable kernels in the most energy-efficient way. Limitations of accelerated clusters are avoided by direct communication. (iii) In general, the ratio of work on the commodity CPU and the accelerator is variable for different applications and even for different kernels of one application. DEEP supports this by virtualising the accelerator in the BNs. The assignment of CNs and BNs might be done dynamically. (iv) At the same time, the dynamical assignment of CNs and BNs improves resiliency. Faulty accelerators do not affect CNs and vice versa.

Furthermore, the architecture supports programmers employing the degrees of parallelism in future machines. Today, the choice of kernels to be offloaded is very restricted, e.g. accelerators do not allow to do communication inside a kernel efficiently. In contrast, DEEP allows more complex kernels including communication to be offloaded to the booster. Such a booster might even be seen as highly-scalable system on its own. Thinking of highly-scalable codes that are able to exploit Blue Gene or QPACE today, they should be able to run on the booster alone. Concurrently, BNs might be assigned to single CNs and used in a way as in an accelerated cluster. Both use-cases – and anything in between – are possible without modifying the hardware concept just by means of software configuration.

In order to benefit from the booster element, the user is forced to identify highly-scalable kernels capable to be offloaded. A cluster-booster software-stack as sketched in Fig. 4 will help to express the different levels of parallelism of the problem. The application's kernels of limited scalability remain on the cluster part and make use of a standard software-stack comprising of standard compilers, MPI, etc. Highly scalable kernels will be offloaded to the booster with the help of a resource-management layer. A cluster-booster communication layer is required to bridge the gap between the flexible cluster interconnect and the highly scalable booster network. In order to implement the highly scalable kernels running on the booster, an additional software stack is required. Since the booster might be seen as a cluster of accelerators, its software-stack is similar to the one on the cluster. Besides native compilers supporting the processors of the booster nodes, a MPI-communication library is required, too. The application itself is formed by both types of kernels running on the cluster and booster parts of the overall system, glued together by the resource management layer.

5 Towards Brain-Inspired Computing

The architectures of Exascale systems with their foreseeable unprecedented levels of parallelism will present numerous challenges, among them energy efficiency and resilience to soft errors. An exciting and promising approach to overcome these challenges is to learn from the human brain, which can be seen as an immensely powerful, energy efficient, self-repairing, self-teaching computer. Being able to understand and mimic the way the brain functions would revolutionise supercomputing and even Information and Communications Technology (ICT) in general.

This is one of the important goals the JSC and its partners seek to achieve being engaged in the Human Brain Project (HBP), a European-led project with partners all over the world. The EU has shortlisted the HBP, and five other projects, as candidate projects for its new FET Flagship Programme. The two projects selected in 2012 will receive significant European and national funding for a period of up to ten years. The current HBP Consortium, led by the École Polytechnique Fédérale de Lausanne, Switzerland, consists of a

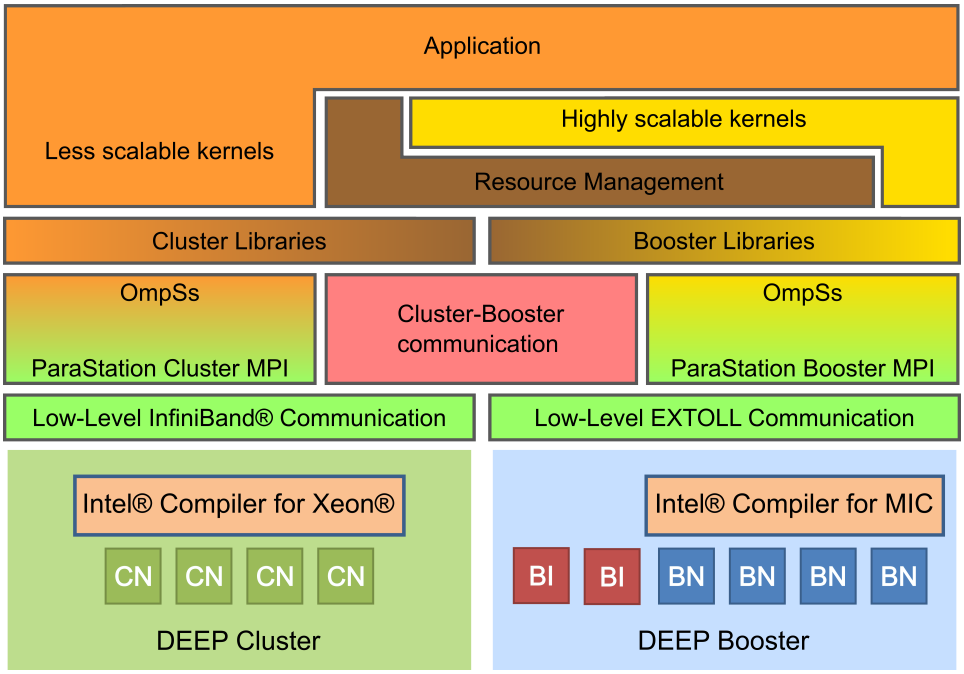


Figure 4. High-level view of the cluster-booster software-stack.

core group of 13 universities and research institutions from nine EU member and associate member states, including Forschungszentrum Jülich as a major partner. JSC as the leader of the HPC Pillar will be responsible for the design, implementation, and management of the HBP's supercomputing infrastructure, drawing on its long-standing expertise in the development and operation of leadership-class systems and its pivotal role in the creation of the European supercomputing infrastructure PRACE.

The aim of the HBP is to bring together everything that is known and everything that can be learned about the inner workings of the brain's molecules, cells and circuits, organise the knowledge in massive databases, and use it to build biologically detailed models of the complete human brain. The simulation of these integrated multi-scale brain models will serve as the basis for new diagnostic tools and treatments for brain diseases, new prosthetic technologies for people with disabilities, a new generation of intelligent robots, and a new class of low energy information technologies with brain-like intelligence, including new types of supercomputing architectures.

Achieving all this will require an international supercomputing infrastructure more powerful than any before. Ultimately, cellular-level simulations of the 100 billion neurons of the human brain will require a machine with Exascale compute power. Computers powerful enough to meet the project's initial requirements are already there or just around the corner. Jülich's IBM Blue Gene/P system JUGENE with its Petaflop performance has the capability to simulate up to 100 million neurons, roughly the number of neurons found in the mouse brain. JSC's new eight-rack Blue Gene/Q system scheduled for spring 2012,

see above, will be available to the HBP.

The procurement of Jülich's next generation large scale production supercomputers will be guided by the result of a comparison of different architectural developments from manufacturers like IBM and Cray and from European Exascale projects like DEEP, see section 4, and Mont-Blanc. JSC will select the hard- and software configuration best suited to the HBP's needs, which may also turn out to be a synthesis of several technologies from different projects and providers. As the project progresses, the capability of the systems will be upgraded from the Petascale towards the Exascale level. Neuromorphic boosters developed by the HBP, devices that emulate the circuitry of the brain to perform information processing that conventional technology cannot perform efficiently, are expected to provide further improvements in performance and efficiency of next-generation supercomputers.

Very few applications challenge supercomputing so severely as the attempt to simulate the brain with biological accuracy. The HBP is thus in the unique position to guide and drive the development of future supercomputers, whose role is twofold as both instrument and subject of study: On the one hand, they will serve as key scientific tool for obtaining new fundamental insights into the functioning of the human brain. The HBP and the JSC in particular will collaborate closely with manufacturers and European Exascale consortia to co-develop hardware and software that best meet the extreme requirements of brain simulations and the scientists' requirements for usability (interactivity!) and end-to-end performance. On the other hand, supercomputing technology will eventually be transformed fundamentally through just these insights, as they might provide the basis for new brain-inspired architectures.

6 Summary and Concluding Remarks

In the last three years, the JSC has substantially strengthened its activities in the field of HPC technology co-development. In 2010, two new labs came into operation addressing important challenges towards Exascale HPC technology. The Exascale Innovation Centre (EIC), jointly operated with IBM, has been essential for the evaluation of early Blue Gene/Q hardware and successful application porting to this new architecture. This has led to a step-wise procurement of 8 racks Blue Gene/Q at Jülich Research Centre to be followed by a larger upgrade of JUGENE. The EIC currently addresses new I/O and data storage concepts in cooperation with a EU funded PRACE 1IP prototype project.

The Exa-Cluster Lab, jointly operated with Intel and ParTec, was among a European consortium of 16 partners, successfully awarded funding of the DEEP cluster-booster project. DEEP realises a new hierarchical concurrency concept aiming at reconciling highly scalable and general purpose supercomputing. The cluster-booster approach might lead to a new type of scalable *and* flexible architecture in the 10 Petaflop/s class starting in 2014 with a huge potential for exascaling.

The JSC is convinced that only application driven technology co-development together with leading companies and with a strong feedback on applications can meet the challenges of Exascale computing. Users can benefit from these activities by getting very early access to new technology and by having a substantial impact on the design of novel architectures.

References

1. V. Springel et al., “The Millennium-XXL Project: Simulating the Galaxy Population of Dark Energy Universes”, http://inside.hlrs.de/htm/Edition_02_10/article_06.html
2. http://www-03.ibm.com/systems/deepcomputing/solutions/bluegene/
3. F. Pozzati, ”Porting and optimization of a Lattice Boltzmann D2Q37 code to Blue Gene/Q,” JSC internal report, 2011.
4. www.fz-juelich.de/jsc/sionlib
5. http://www.top500.org
6. http://www.green500.org
7. H. Baier et al., “QPACE: power-efficient parallel architecture based on IBM PowerXCell 8i”, Computer Science - R&D 25 (2010), pp. 149-154. doi:10.1007/s00450-010-0122-4.
8. Gene Amdahl (1967), “Validity of the Single Processor Approach to Achieving Large-Scale Computing Capabilities” (PDF), AFIPS Conference Proceedings (30), pp. 483-485.
9. John L. Gustafson, “Re-evaluating Amdahl’s Law”, Communications of the ACM 31(5), 1988, pp. 532-533.
10. http://www.intel.com/pressroom/archive/releases/20100531comp.htm
11. Mondrian Nüssle et al., “A resource optimized remote-memory-access architecture for low-latency communication”, The 38th International Conference on Parallel Processing (ICPP-2009), September 22-25, Vienna, Austria.

The Chiral Transition, Deconfinement, and the Equation of State of QCD

Frithjof Karsch[1,2,3] **and Wolfgang Söldner**[4]

in collaboration with

A. Bazavov[3], **T. Bhattacharya**[5], **M. Cheng**[6], **C. DeTar**[7], **H.-T. Ding**[3], **Steven Gottlieb**[8], **R. Gupta**[5], **P. Hegde**[3], **U.M. Heller**[9], **E. Laermann**[1], **L. Levkova**[7], **S. Mukherjee**[3], **P. Petreczky**[3], **C. Schmidt**[1,10], **R.A. Soltz**[6], **R. Sugar**[11], **D. Toussaint**[12], **W. Unger**[13], **and P. Vranas**[6]

(Bielefeld/GSI and HotQCD Collaboration)

[1] Fakultät für Physik, Universität Bielefeld, D-33615 Bielefeld, Germany
E-mail: karsch@physik.uni-bielefeld.de

[2] GSI Helmholtzzentrum für Schwerionenforschung, D-64291 Darmstadt, Germany

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

[4] Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany
E-mail: wolfgang.soeldner@physik.uni-regensburg.de

[5] Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

[6] Physics Division, Lawrence Livermore National Laboratory, Livermore CA 94550, USA

[7] Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA

[8] Physics Department, Indiana University, Bloomington, IN 47405, USA

[9] American Physical Society, One Research Road, Ridge, NY 11961, USA

[10] Frankfurt Institute for Advanced Studies, J.W.Goethe Universität Frankfurt
D-60438 Frankfurt am Main, Germany

[11] Physics Department, University of California, Santa Barbara, CA 93106, USA

[12] Physics Department, University of Arizona, Tucson, AZ 85721, USA

[13] Institut für Theoretische Physik, ETH Zürich, CH-8093 Zürich, Switzerland

Results for thermodynamic quantities obtained from lattice calculations within (2+1)-flavour quantum chromodynamics (QCD) are presented. Such calculations are of great interest for the phenomenology of relativistic heavy ion collisions as well as for our understanding of the early universe. We focus here on a description of recent progress obtained in the analysis of chiral and deconfinement aspects of the QCD transition as well as the QCD equation of state. We report on calculations that have been performed with $2+1$ flavours of quarks using improved staggered fermion actions. Lattices with temporal extent $N_\tau = 6$, 8 and 12 have been used to understand and control discretization errors and to reliably extrapolate estimates obtained at finite lattice spacings to the continuum limit. The chiral transition temperature is defined in terms of observables that are sensitive to the phase transition in a theory with two massless flavours. It is analyzed using $O(N)$ scaling fits to the chiral condensate and susceptibility. We find consistent estimates from the HISQ/tree and asqtad actions and our main result is $T_c = 154 \pm 9$ MeV.

1 Introduction

The thermodynamics of quantum chromodynamics (QCD) and, especially, its numerical analysis using the lattice regularized definition of QCD (lattice QCD) is of great importance to the phenomenology of heavy ion collisions and the physics of the early universe. Lattice calculations for the equation of state (EoS)[1–4] are of direct relevance for model calculations describing the expansion of dense matter formed in relativistic heavy ion collisions as well as in the early universe. In large heavy ion collision experiments at LHC, RHIC, SPS, and future experiments at FAIR, hadronic matter is studied under extreme conditions similar to those that were present milliseconds after the Big Bang.

In those experiments two heavy ions are colliding at very high energies (presumably) producing a new form of matter, the so-called quark-gluon plasma (QGP). During the expansion of this dense matter the medium cools down and eventually undergoes a chiral and confinement transition where free quarks and gluons become again confined to hadrons. The temperature regime where this dramatic change appears is the target of our studies.

While simulations at small quark mass are notoriously difficult, great progress has been made in the past years, especially with respect to a better understanding of the physics of the QCD transition. Simulations with quark mass values close to the physical point or even below allow for a detailed investigation of critical phenomena which arise in the vicinity of the chiral phase transition. Although QCD at the physical point shows a crossover transition, heavy ion experiments may still be sensitive to large correlation lengths caused by "nearby" critical phenomena. To quantify the proximity to criticality lattice simulations are absolutely essential[6]. In this paper we give a brief overview of the chiral and deconfinement aspects of QCD and present the latest results from our lattice QCD simulations. In addition, we present an update of ongoing calculations for the QCD EoS.

The paper is organized as follows. In Sec. 2 we give an overview of the simulation details. Our latest results for the chiral transition are presented in Sec. 3 and some aspects of the transition related to deconfinement are discussed in Sec. 4. In Sec. 5, finally, we give an update of the calculations of the EoS with the asqtad action that are currently performed at the John von Neumann Centre (NIC) at Jülich.

2 Numerical Details and Setup

In order to perform state-of-the-art calculations in finite temperature QCD within the framework of lattice QCD large computational resources are required. The calculations we are presenting have been performed on IBM Blue Gene/L and Blue Gene/P, where we have been using those supercomputers located at the Jülich Supercomputing Centre, the Lawrence Livermore National Laboratory, and the New York Center for Computational Sciences.

In order to calculate expectation values we generated ensembles of gauge field configurations for several temperatures T by using a Rational Hybrid Monte Carlo (RHMC) algorithm[3]. For our calculations we used the HISQ/tree[7,8] action and the asqtad action[9] which are improved actions in the sense that cutoff effects are reduced[10]. We simulated QCD with $N_f = 2+1$ flavours, i.e. we have two degenerate light quarks and one (heavier) strange quark. We have generated various ensembles with different light quark masses m_l, keeping the strange quark mass m_s fixed at its physical value. We performed simulations

on a line of constant physics (LCP) with a fixed value of the temporal extent of the lattice (N_τ), i.e. while changing the lattice spacing, and hence the temperature, by changing the coupling constant, we choose the quark masses such that their values in physical units are kept constant. In the temperature range we are simulating it turns out that the LCP is to a good approximation obtained when choosing m_l/m_s to be constant. Thus, throughout this paper, a change in the ratio of m_l/m_s is identical to a change in the light quark mass (since the strange quark mass is kept at its physical value). In order to get rid of discretization errors in our results we have to perform the continuum limit, i.e. we have to send the lattice spacing to zero or, correspondingly, send N_τ to infinity. To convert the lattice results to physical units we use either the r_1 or the f_K scale. Some computational details are summarized in Refs. 11-13. Further information can be found in Refs. 14-16.

3 Chiral Aspects of the Quark Gluon Plasma (QGP)

In the chiral limit, where the quark mass goes to zero, QCD undergoes a phase transition. The light quark chiral condensate, defined as the derivative of the QCD partition function with respect to the light quark mass, $\langle\bar{\psi}\psi\rangle_l = \frac{T}{V}\frac{\partial \ln Z}{\partial m_l}$, is an order parameter for that phase transition. Since it needs multiplicative renormalization, and for finite quark mass also additive renormalization, the combination $\Delta_{l,s}$ of light and strange quark chiral condensates at finite and zero temperature T is introduced,

$$\Delta_{l,s} = \frac{\langle\bar{\psi}\psi\rangle_{l,T} - \frac{\hat{m}_l}{\hat{m}_s}\langle\bar{\psi}\psi\rangle_{s,T}}{\langle\bar{\psi}\psi\rangle_{l,0} - \frac{\hat{m}_l}{\hat{m}_s}\langle\bar{\psi}\psi\rangle_{s,0}}. \tag{1}$$

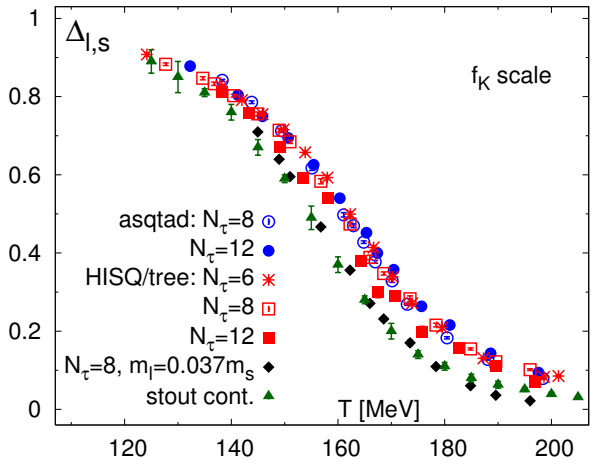


Figure 1. The subtracted chiral condensate for the asqtad and HISQ/tree actions with $m_l = m_s/20$ is compared with the continuum extrapolated stout action results[5]. The temperature T is converted into physical units using f_K. The black diamonds show HISQ/tree results for $N_\tau = 8$ lattices after an interpolation to the physical light quark mass using the $m_l/m_s = 0.05$ and 0.025 data.

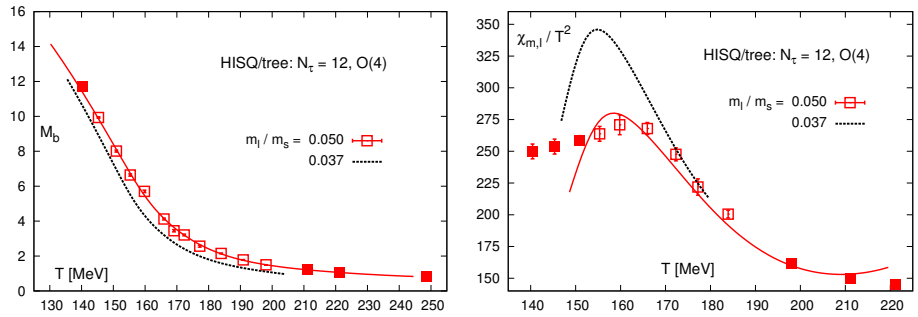


Figure 2. Scaling fits and data for the chiral condensate M_b calculated with the HISQ/tree action on lattices with temporal extent $N_\tau = 12$ (left) and the chiral susceptibility $\chi_{m,l}$(right). The data for M_b and $\chi_{m,l}$ at $m_l/m_s = 0.05$ are fitted simultaneously using the $O(4)$ scaling Ansatz. Fits using the $O(2)$ Ansatz are similar. Data used in the scaling fits are plotted using open symbols. The black, dotted lines give the data scaled to the physical quark masses.

In this combination the multiplicative renormalization constant and the additive renormalization constant term linear in the quark mass are cancelled. The results from our simulations for $\Delta_{l,s}$ at light quark mass $m_l = m_s/20$ are shown in Fig. 1. This value of the light quark mass corresponds to a Goldstone pion mass of about 160 MeV. To convert the lattice data into physical units we have set the scale by f_K. As can be seen in Fig. 1, the chiral condensate $\Delta_{l,s}$ drops rapidly in the transition region. The discretization effects that arise from a non-zero lattice spacing a (or, equivalently, from a finite temporal extent N_τ) are moderate when f_k is used to set the scale. We also show data extrapolated to the physical pion mass, where $m_l = 0.037 m_s$. For the extrapolation we made use of our data for smaller light quark masses, $m_l = 0.025 m_s$. A comparison to the results of the Wuppertal-Budapest group with stout fermions[5] yields good agreement.

In the following we will focus on the transition region. In this region physics is captured by the singular part of the QCD partition function. For sufficiently small (light) quark masses the order parameter is described by a universal scaling function f_G which depends on the critical exponents of the $O(2)$ ($O(4)$) universality class, the external fields m_l and m_s through $H = \frac{m_l}{m_s}$, and the reduced temperature $\Delta T = \frac{T-T_c}{T_c}$, see Ref. 17 for more details. On the left side of Figs. 2 and 3 we have plotted the multiplicatively renormalized chiral condensate $M_b \equiv \frac{m_s}{T^4} \langle \bar{\psi}\psi \rangle_l$ and its parameterization given by the scaling function plus additional, scaling violating terms which stem from the regular part of the partition function,

$$M_b(T, m_l, m_s) = h^{1/\delta} f_G(t/h^{1/\beta\delta}) + (a_0 + a_1 \Delta T + a_2 \Delta T^2) H, \qquad (2)$$

where a_0, a_1, and a_2 are constants, $h = \frac{H}{h_0}$ and $t = \frac{\Delta T}{t_0}$ with scales h_0 and t_0 which need to be determined together with T_c in the chiral limit[17]. The constants δ and β are critical exponents. Fits based on this parameterization yield good agreement with the data in the transition region as well as for larger temperatures. Less agreement is found at lower temperatures and for larger quark masses. The latter observation may lead to the conclusion that the scaling window in the hadronic phase is significantly smaller compared to the QGP phase. As an example we show results in Figs. 2 and 3 for the HISQ/tree and asqtad action

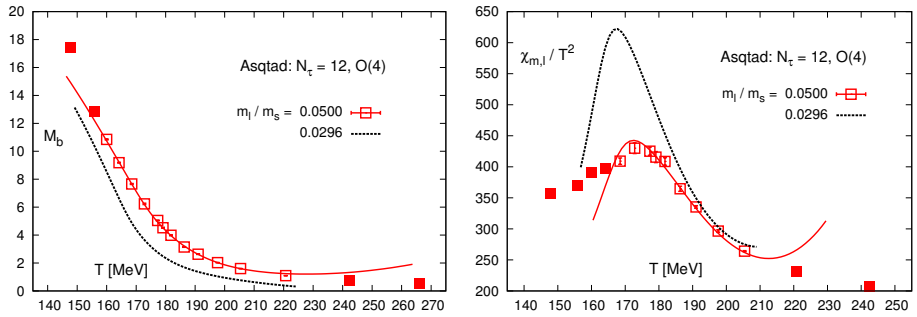


Figure 3. Scaling fits and data for the chiral condensate M_b calculated with the asqtad action on lattices with temporal extent $N_\tau = 12$ (left) and the chiral susceptibility $\chi_{m,l}$(right). The data for M_b and $\chi_{m,l}$ at $m_l/m_s = 0.05$ are fit simultaneously using the $O(4)$ scaling Ansatz. Fits using the $O(2)$ Ansatz are similar. Data used in the scaling fits are plotted using open symbols. The dotted lines give the data scaled to the physical quark masses, which in the case of our calculations with the asqtad action corresponds to $m_l = 0.0296 m_s$ as the strange quark mass turned out to be somewhat larger than its physical value[18].

for $N_\tau = 12$ at $m_l = 0.05 m_s$ using $O(4)$ scaling. From the scaling analysis we know explicitly how the crossover temperature depends on the light quark mass as it approaches the critical point. For a given N_τ value this scaling analysis is especially useful for defining the crossover temperature T_p for given m_l/m_s and for extrapolating that temperature to its value at the physical light quark mass or to the chiral limit of zero light quark mass (T_c). For example, in Figs. 2 and 3 we have also plotted M_b for the physical light quark mass $m_l = 0.037 m_s$ which is determined by the fits to the results at the larger light quark mass $m_l = 0,05 m_s$. Furthermore, we can make use of the scaling function of M_b for the chiral susceptibility $\chi_{m,l}$ which is just the derivative of the chiral condensate with respect to the light quark mass. As can be seen on the right side of Figs. 2 and 3, the data are well described by the fitting function in the transition region. Again we can determine $\chi_{m,l}$ at physical light quark mass value from this fit. The crossover temperature can then be identified with the peak in the chiral susceptibility.

The above analysis using simultaneous fits to data for the order parameter M_b and the total chiral susceptibility $\chi_{l,m}$ has also been repeated for the asqtad action on $N_\tau = 8$ lattices where we analyzed the transition on two lines of constant physics, $m_l/m_s = 1/20$ and $1/10$. We failed to find a fit that reproduced the data for the chiral susceptibility at the heavier mass. We conclude that $m_l/m_s = 1/10$ is not within the scaling region and restrict the simultaneous fit to just data with $m_l/m_s = 1/20$.

In order to determine the crossover temperature at the physical point, $m_l/m_s = 1/27$, we still have to perform the continuum limit, i.e. $N_\tau \to \infty$. In Fig. 4 we show our results from the scaling analysis for the extrapolated pseudocritical temperatures for different N_τ for the HISQ/tree and asqtad action. For the combined fit of the two actions we use an ansatz quadratic in $1/N_\tau^2$. After a detailed analysis of our data we find for the value of the crossover temperature at the physical point,

$$T_c = (154 \pm 9)\mathrm{MeV}\,. \tag{3}$$

For more details we refer the reader to Ref. 18.

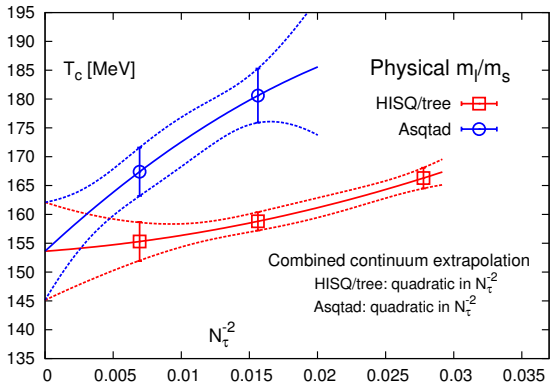


Figure 4. Data and continuum extrapolation for the transition temperature obtained from the $O(4)$ scaling analysis for $m_l/m_s = 1/27$ extrapolated to the physical value. The figure shows a combined quadratic fit to both data sets.

4 Quark Number Susceptibility and the Polyakov Loop

In the discussion of deconfinement one often investigates quantities like the light (l) and strange (s) quark number susceptibilities, $\chi_{l,s} \equiv \frac{1}{VT}\frac{\partial^2 \ln Z}{\partial(\mu_{l,s}/T)^2}$, the energy density ε, or the Polyakov loop. Aside from the Polyakov loop, where a relation (if any) to the singular part of the partition function of QCD is unknown, the other quantities are also sensitive to critical behaviour. However, criticality is less pronounced in the sense that these quantities do not diverge at T_c in the chiral limit, in contrast, e.g., to the disconnected chiral susceptibility. For example, in the chiral limit the temperature derivative of the quark number susceptibilities in the vicinity of T_c are given by,

$$\frac{\partial \chi_{l,s}}{\partial T} \sim c_r^{l,s} + A_{\pm}^{l,s} \left| \frac{T - T_c}{T_c} \right|^{-\alpha} . \tag{4}$$

Note that at the critical temperature, T_c, the slope is controlled by a contribution that arises from the regular part of the free energy, while its variation with temperature is given by the singular part. Since the critical exponent α is negative for theories belonging to the 3-d, O(N) universality class, it appears to be very difficult to reliably extract information about T_c from $\chi_{l,s}$ because contributions from the regular part of the partition function may hide critical behaviour. High statistics lattice data would be needed to discriminate the singular from the regular contributions. Therefore, critical behaviour in state-of-the-art lattice calculations is best seen in quantities like the disconnected chiral susceptibility.

Let us take a closer look at χ_s. For low temperatures the relevant degrees of freedom which contribute to χ_s are hadrons and χ_s is small. At high temperatures the degrees of freedom are quarks and gluons. The quark number fluctuations are then related to the number of quarks and increase with temperature towards the ideal gas limit. In Fig. 5 we compare χ_s for different N_τ and actions. We observe that in the vicinity of the pseudocritical temperature χ_s appears to be a rather smooth function. The discretization effects are rather small when f_K is used to set the scale and we find good agreement with the results of Ref. 5.

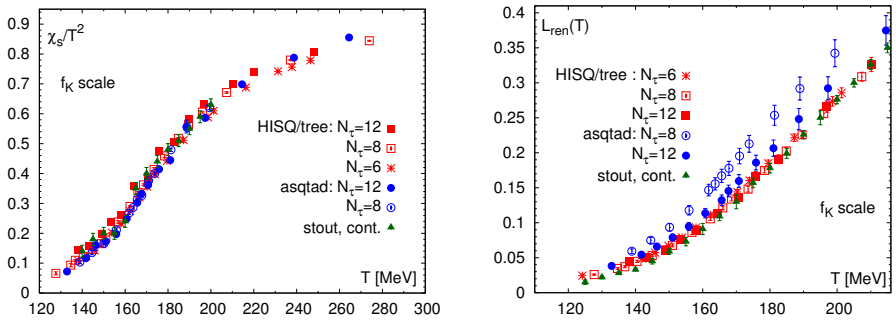


Figure 5. Data for the strange quark number susceptibility for $m_l = 0.05m_s$ with the asqtad and HISQ/tree actions is compared to the continuum extrapolated stout results[5] (left panel). The renormalized Polyakov loop calculated for the asqtad and HISQ/tree actions and compared with the continuum extrapolated stout result[5] (right panel). To set the scale we have used f_K.

The Polyakov loop is a sensitive probe of the thermal properties of the medium even though it is not a good measure of the critical behaviour as discussed above. After proper renormalization, the square of the Polyakov loop characterizes the long distance behaviour of the static quark anti-quark free energy; it gives the excess free energy needed to screen two well separated colour charges. The renormalized Polyakov loop is calculated from the bare Polyakov loop by multiplying it by the renormalization constant $z(\beta)$, see Ref. 11,

$$L_{ren} = z(\beta)^{N_\tau} L_{bare} = z(\beta)^{N_\tau} \frac{1}{3} \mathrm{tr} \prod_{x_0=0}^{N_\tau - 1} U_0(x_0, \mathbf{x}). \tag{5}$$

In the above formula $U_0(x_0, \mathbf{x})$ denotes the temporal link variables. Also for the Polyakov loop we observe a rapid change in the transition region. The discretization effects are moderate and the dependence on the quark mass in the transition region is also small which is expected because the Polyakov loop is a gluonic quantity.

5 Equation of State

The calculation of the QCD equation of state is done by means of the trace anomaly, $\varepsilon - 3p$, where ε is the energy density and p is the pressure. We work on a line of constant physics at fixed value of N_τ. In Fig. 6 we show the trace anomaly for the asqtad action for two different values of the light quark masses; the strange quark mass has been tuned close to its physical value. As can be seen in the plot, at large temperatures $\varepsilon - 3p$ decreases towards zero. This is expected because at high temperatures the conformal symmetry is gradually restored. In the transition region and above, the trace anomaly has its largest values. With respect to critical behaviour the trace anomaly shares the same functional dependence as the quark number susceptibilities, i.e.,

$$\varepsilon - 3p \sim a_r + b_r \frac{T - T_c}{T_c} + c_\pm \left| \frac{T - T_c}{T_c} \right|^{1-\alpha}. \tag{6}$$

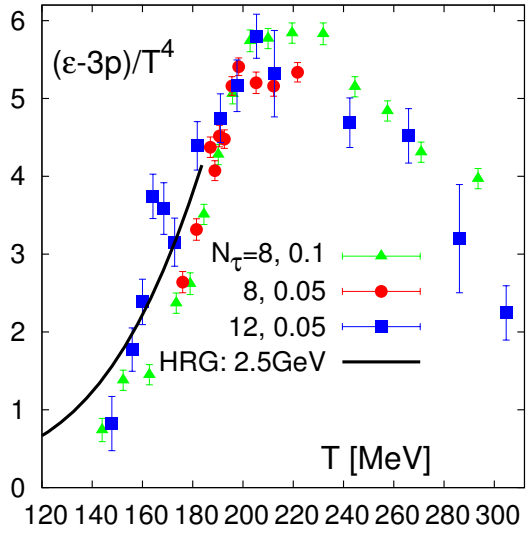


Figure 6. Trace anomaly for the asqtad action at light quark masses $m_l = 0.05m_s$ and $m_l = 0.1m_s$. We compare to the HRG model where resonances up to 2.5 GeV are included.

Critical behaviour thus is a subdominant feature in $\varepsilon - 3p$, just like in $\chi_{l,s}$. Therefore, it may not be surprising that $\varepsilon - 3p$ obtains its largest value at a temperature above the pseudocritical temperature. At temperatures below T_p the trace anomaly decreases and, eventually, follows the prediction of the HRG model[19]. Compared with our earlier calculations we observe that our new $N_\tau = 12$ asqtad data appear to be closer to the HRG curve. In the high temperature region we find qualitative agreement for our new data at $N_\tau = 12$ and $m_l = 0.05m_s$ with the old data at $N_\tau = 8$ and $m_l = 0.1m_s$. However, it is also evident that at the current stage of our calculations the data is not yet conclusive and statistics for the new data sets has to be improved.

6 Outlook

In this contribution we gave a status report of ongoing calculations within the Bielefeld/GSI and HotQCD collaboration focusing on chiral and deconfinement properties at finite temperature for QCD with $2+1$ flavours. We supplemented our earlier work with new calculations with asqtad fermions at $N_\tau = 12$ and $m_l = 0.05m_s$, and with HISQ/tree fermions at $N_\tau = 6, 8, 12$ at $m_l = 0.05m_s$. A detailed study of the universal scaling properties of the chiral transition allowed us to extrapolate our results to the physical point and to extrapolate to the continuum limit. As a result, we extracted a value for the transition temperature at the physical point, $T_p \simeq (154 \pm 9)$ MeV.

Relating critical behaviour to observables which are dominantly sensitive to deconfinement, i.e. the sudden change of degrees of freedom, is more subtle. We have shown latest results for the Polyakov loop and the strange quark number susceptibility. Furthermore, we gave an update on our ongoing calculations of the QCD EoS.

Acknowledgements

We are grateful to the Jülich Supercomputing Centre, LLNL and the New York Center for Computational Sciences for providing access to the computational resources.

References

1. Y. Aoki, Z. Fodor, S. D. Katz and K. K. Szabo, *The equation of state in lattice QCD: With physical quark masses towards the continuum limit,* JHEP **0601**, 089 (2006) [arXiv:hep-lat/0510084].
2. C. Bernard *et al.*, *QCD equation of state with 2+1 flavors of improved staggered quarks,* Phys. Rev. D **75**, 094505 (2007).
3. M. Cheng *et al.*, *The QCD Equation of State with almost Physical Quark Masses,* Phys. Rev. D **77** (2008) 014511.
4. A. Bazavov *et al.*, *Equation of state and QCD transition at finite temperature,* Phys. Rev. D **80** (2009) 014504.
5. S. Borsanyi *et al.*, *Is there still any Tc mystery in lattice QCD? Results with physical masses in the continuum limit III,* JHEP **1009**, 073 (2010).
6. M. Cheng *et al.*, *Baryon Number, Strangeness and Electric Charge Fluctuations in QCD at High Temperature,* Phys. Rev. D **79** (2009) 074505.
7. E. Follana *et al.* [HPQCD collaboration and UKQCD collaboration], *Highly Improved Staggered Quarks on the Lattice, with Applications to Charm Physics,* Phys. Rev. D **75**, 054502 (2007).
8. A. Bazavov *et al.* [HotQCD Collaboration], *Deconfinement and chiral transition with the highly improved staggered quark (HISQ) action,* J. Phys. Conf. Ser. **230**, 012014 (2010).
9. C. Bernard [MILC Collaboration], **PoS** (LAT2010)181.
10. A. Bazavov *et. al.* [HotQCD Collaboration], **PoS** (LAT2010)169.
11. M. Cheng *et al.*, *Equation of State for physical quark masses,* arXiv:0911.2215 [hep-lat].
12. R. Gupta, *The EOS from simulations on Blue Gene L Supercomputer at LLNL and NYBlue*, PoS **LAT2008** 170 (2008).
13. W. Soldner [HotQCD collaboration], *Chiral Aspects of Improved Staggered Fermions with 2+1-Flavors from the HotQCD Collaboration,* PoS **LATTICE2010** (2010) 215 [arXiv:1012.4484 [hep-lat]].
14. C. DeTar, *Recent Progress in Lattice QCD Thermodynamics*, PoS **LAT2008** (2008).
15. F. Karsch, *Recent lattice results on finite temperature and density QCD*, PoS **LAT2007** 015 (2007).
16. C. DeTar, R. Gupta [HotQCD Collaboration], *Towards a precise determination of Tc with 2+1 flavors of quarks*, PoS **LAT2007** 179 (2007) [arXiv:hep-lat/0710.1655].
17. S. Ejiri *et al.*, *On the magnetic equation of state in (2+1)-flavor QCD,* Phys. Rev. D **80**, 094505 (2009).
18. A. Bazavov *et al.*, *The chiral and deconfinement aspects of the QCD transition,* arXiv:1111.1710 [hep-lat].
19. P. Braun-Munzinger, K. Redlich and J. Stachel, *Particle production in heavy ion collisions,* arXiv:nucl-th/0304013.

Computational Materials Physics at the JSC

Martin H. Müser

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

New materials have always been central to technological breakthroughs and the increase of living standards. Their design and processing is ever more influenced by the knowledge of the ways in which interactions between the constituents of a material affect its macroscopic properties. Computer simulations increasingly contribute towards this end, thereby furthering our fundamental understanding as well as providing useful guidelines for real-life applications. This is the mission of the newly established NIC research group Computational Materials Physics. Here, I will summarize selected previous results by the head of the research group, lay out some of the anticipated research, which will benefit tremendously from the computational resources at the Jülich Supercomputing Centre, and present some first results. The simulations discussed here will evolve around (i) the possibility of piezoelectronic materials, (ii) avenues for (force-field based) simulations of (non-equilibrium) redox reactions, and (iii) simulations of the leakage between a seal and a rigid surface with self-affine roughness.

1 Introduction

As is the case in virtually all branches of science and technology, computer simulations play a crucial role in materials science. Historically, materials simulators can be divided into two groups: Engineers as well as applied mathematicians tend to start from continuum mechanics, while scientists, typically chemists and physicists, rather invoke a particle-based (atoms, ions, electrons) picture. At first sight, the two communities tackle different questions and use disjunct methods to solve them. In continuum mechanics, one assumes certain *constitutive laws* that a material (or a fraction of it) obeys, and investigates how a given object (of rather concise and usually "unpleasant" geometry) responds to externally imposed boundary conditions. An example is the question of how much a Space Shuttle heats up when it enters the atmosphere or how resilient a bone with osteoporosis may be to mechanical stress. Conversely, computational chemists, physicists, and an increasing number of biologists would like to predict or explain the property of matter, all the way from molecules to materials, as a function of their atomic composition.

The frontiers between the two camps disappear more and more for a variety of reasons. An important one is the ongoing miniaturization of structures, which can make it important for the engineer to consider atomic scale properties, while scientists can no longer idealize boundary conditions in the way they used to or assume a clear separation of length and time scales for many problems of contemporary interest. Moreover, the mathematical and computational methodologies employed in the two camps can be very much related. Quite a number of problems can be reduced to second-order differential equations (the engineer's heat equation and the chemist's Schrödinger equation are formally identical), and traditional continuum problems, such as the Navier Stokes equation in fluid mechanics, are frequently solved with particle-based methods. Using the points of view of scientists and engineers and bridging between the large and the small is at the very heart of materials physics. A central task of computational materials physicists is to derive, explain and parameterize the constitutive equations that practitioners need to know for their applications.

Many times this means that the comforting idea of universality has to be given up and that material-specific details need to be considered. In the following sections, three examples are given that reflect the interests of the new NIC Computational Materials Physics group.

2 Are Piezoelectronic Materials Possible?

Rewritable optical media such as DVD-RW rely on the strong contrast of the optical properties that phase change materials (PCMs) exhibit between their ordered (metallic or semimetallic) and disordered (insulating or semi-conducting) phase, representing 0 and 1 bits respectively[1]. Owing to many beneficial properties, e.g., fast switching and low costs, PCMs are promising candidates for improved non-volatile memory cells and for other applications such as programmable switches.

PCMs can be switched within a few dozen nanoseconds from crystalline to amorphous through heating that is followed by a quench and back to crystalline by annealing the glass at moderately high temperatures. Due to their fast switching, PCMs are particularly amenable for molecular dynamics (MD) simulations so that density functional theory (DFT) based MD can be directly compared to experiments[2]. The simulations confirm the idea that both long-range and local (!) order differ substantially between crystal and glass. This would explain the unusual large contrast in optical and electronic properties. However, experimental conjectures on the structure of the glass needed to be corrected: Kolobov et al.[3] had suggested that the phase switch was intimately linked to the motion of certain group 14 elements from local tetrahedral bonding in the glass to (distorted) octahedral in the crystal. In their DFT based MD simulations of the PCM $Ge_2Sb_2Te_5$, Akola and Jones[2] found that only one third of the Ge atoms assumed tetrahedral configurations in the glass and instead identified additional systematic pattern differences in the local order of crystals and glasses as being important.

In a hard-sphere picture for atoms, the two different sites that some germanium atoms occupy in the amorphous and the crystalline structure can be associated with different amounts of available volume, namely small tetrahedral and large (distorted) octahedral sites. This made us conclude that it should be possible in principle to change the *local order* through pressure[4]. Under compression a germanium atom would be squeezed from its tetrahedral coordination into a large, adjacent, octahedral void, while a tensile stress would direct it back into tetrahedral coordination. To study this hypothesis, we investigated the PCM $Ge_{0.15}Sb_{0.85}$, which unlike most commercial PCMs is free of stoichiometric vacancies in the crystal, both experimentally and by computer simulation. The compression was easiest done in a diamond anvil cell, and indeed a transformation from the amorphous state to a (nano-) crystalline state was observed. Applying a large tensile load would necessitate a tremendous experimental effort, which was our motivation to study the tensile-stress induced transformation with DFT-based MD. These simulations showed the anticipated transformation to the amorphous insulating structure as summarized in Fig. 1.

An interesting side aspect of the simulations was the observation that the glass – which had been obtained through a quench from the liquid – did not relax towards the crystal but instead seemed to be relaxing away from it[5]; when temperature and volume were fixed, the bare ionic interaction $V_{\mathrm{ion-ion}}$ (assuming unscreened Coulomb interaction between the nuclei in a constant charge-neutralizing background) increased with time during relaxation thereby moving away from the crystalline structure, where $V_{\mathrm{ion-ion}}$ is much smaller than in the fluid or the glass phase.

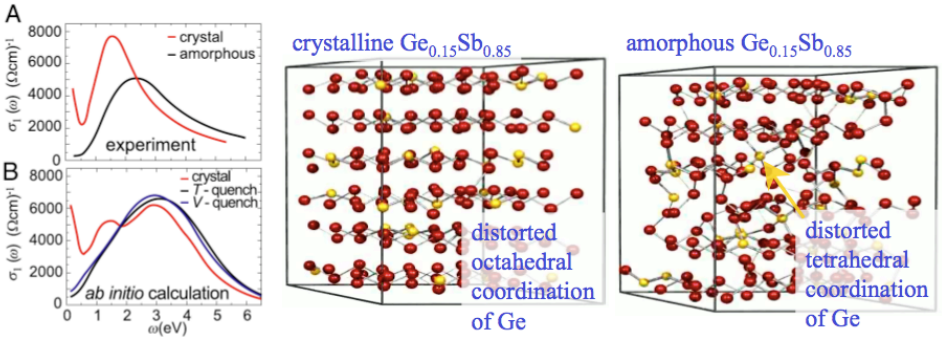


Figure 1. Left graph: Conductivity for crystalline and amorphous $Ge_{0.15}Sb_{0.85}$ measured experimentally (A) and as obtained from simulations (B). The glasses in the simulations were obtained by a temperature and a "volume" (pressure) quench. DFT tends to underestimate band gaps, which leads to an overestimation of conductivity. The calculations yet reproduce experimental trends and moreover show that a glass, which had been produced by a tensile load, shows a similar frequency dependent conductivity as a glass that was obtained by quenching the liquid. Two right graphs: Structure of $Ge_{0.15}Sb_{0.85}$ in its crystalline and amorphous form. The latter is obtained by first applying a tensile load on the simulation cell and then releasing the stress again.

The combination of experiment and simulation shows that it is possible to switch some (vacancy-free) PCMs reversibly through stress. If it were possible to make such a "piezoelectronic" switching mechanism happen, e.g., through nanoscale piezoelectric transducers, many advantages could arise: (i) Switching times could decrease by an order of magnitude, as the relevant kinetics would be related to the speed of sound rather than the thermal diffusion constant. (ii) Thermal degradation and phase separation would be a much smaller issue than if the alloy is melted during switching. (iii) Less heat per volume element would be required to induce the phase change.

Computer simulations could certainly guide the search for the most appropriate switching material and provide approximate switching pressures, latent volumes, and other quantities that are important in the design process. A full set of tests for one given alloy would require one to do a pressure cycle as well as a temperature cycle. This would take about four to six weeks on a Blue Gene/P for a system of 256 atoms. The compression of amorphous material and the thermal switching is more easily done in experiment, unlike the tensile loading, which is technically very challenging to achieve. It would only be left to experiments to investigate how many rewrite processes a material can withstand, once a promising alloy has been identified in the simulations.

3 Towards the Force Field Based Simulation of Non-Equilibrium Redox Reactions

In many cases, one is interested in studying the dynamics of molecular systems for longer times than feasible with DFT or *ab initio* based MD. If the problem at hand is material specific, reasonably accurate force fields are needed. They have been derived for a variety of systems, but a bottleneck remains the realistic handling of electrostatic interactions. Assigning fixed charges to atoms is prohibitive in an abundance of cases. For example, when studying interfaces of bulk silicon and silicon dioxide, all silicon atoms should be

treated equally, however, their oxidation state and thus their partial charge depends on their chemical environment. Likewise, whenever a redox reaction occurs, the partial charge of the atom will be changed, sometimes in a quasi-discontinuous fashion. Similar comments apply for almost all chemically heterogeneous systems.

The general strategy in approaches that assign partial atomic charges on the fly, is to minimize an expression, which one could interpret as a "poor man's Hamiltonian", with respect to the charges. The most popular "charge equilibration" (QE) model, also used in the so-called reactive force field ReaxFF[6], includes terms related to the Coulomb interaction, $V_{\rm C}$, between atoms, the electronegativity, χ, and the chemical hardness of atoms, κ. The latter is quadratic in an atomic charge and reflects its self-interaction on a given atom. An alternative approach is the atom-atom charge transfer (AACT) model[7]. It is based on the idea that charge is transferred through chemical bonds. Now the transfer charges are penalized, i.e., a bond hardness, $\kappa_{\rm b}$ is introduced, which is essentially inversely proportional to the polarizability of the chemical bond. The split charge equilibration (SQE) model is a hybrid of these two approaches[8]. Let q_{ij} be the split charge - or transfer charge - from atom j to atom i. Then the charge on atom i is given by

$$Q_i = \sum_{ij} q_{ij} + n_i \cdot e, \tag{1}$$

where n_i is an integer number, reflecting the number of "external" excess charges on atom i, and e is the elementary charge. The expression to minimize (for example, self consistently) becomes

$$V = V_{\rm C}(\{\mathbf{R}, Q\}) + \sum_i \left(\frac{\kappa_i}{2} Q_i^2 + \chi_i Q_i \right) + \sum_{i,j>i} \frac{\kappa_{{\rm b},ij}}{2} q_{ij}^2, \tag{2}$$

where the $\kappa_{{\rm b},ij}$ may also depend on the atomic configuration $\{\mathbf{R}\}$, e.g., such that the proper dissociation limit is obtained. The QE model arises in the limit where all $\kappa_{\rm b}$ are set to zero, while AACT is equivalent to neglecting the atomic κ.

In principle, atomic hardnesses and electronegativities are element-specific properties that one can obtain through finite differences from the free-atom electron affinity and the first ionization energy. However in practice, it is convenient to work with values that differ slightly from these values, i.e., by $\mathcal{O}(10\%)$. This adjustment reduces errors on atomic charges in molecules and solids, when compared to quantum chemical results, from roughly 80% to 30%. In the original SQE paper[8], we found that optimizing bond and atomic hardnesses simultaneously reduces the error in the SQE charges to approximately 10%, as compared to about 30% in pure QE or pure AACT approaches – see also Ref. 9 for a thorough comparison between SQE and regular QE. Fig. 2 contains more precise numbers along with a brief explanation of how the model parameters can be found. For a formal, DFT-based justification of the originally *ad hoc* introduced parameters used in SQE, the reader is referred to a work by Verstraelen *et al*[10].

What may be more important than mean errors of predicted charges, are the generic trends that the various charge equilibration schemes produce. For example, AACT does not show the proper scaling of the polarizability for oligomers[11] in the limit of small degrees of polymerization P, and the skin depth for external fields in solids is always less than an atomic spacing[12]. Conversely, QE produces the wrong dissociation limit of molecules. For example, if a CH_4 molecule were separated from an H_2O molecule, the molecules would carry unphysical charges of about $\pm 0.4\, e$, unless artificial constraints were imposed.

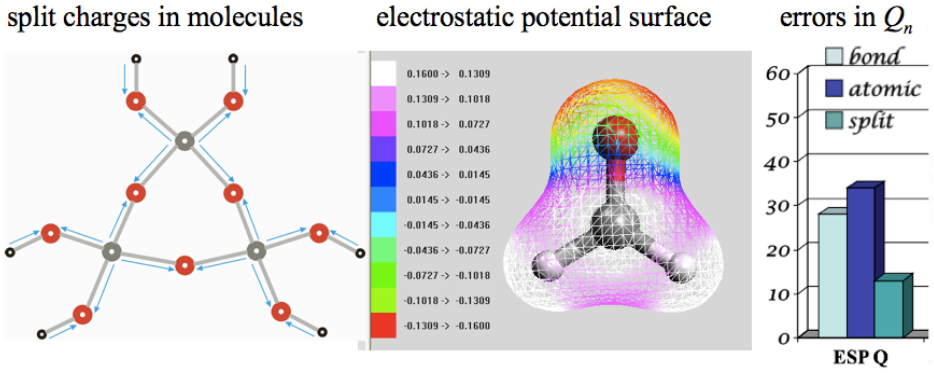


Figure 2. Left: Sketch of the allowed exchange of partial charges between two chemically bonded atoms in the split charge model. Charge build-up is penalized in the bond and on the atom, while former models were exclusively either atom or bond based. The represented molecule is a three-membered hydrogen terminated silica ring - $[(HO)_2SiO]_3$. Centre: Electrostatic potential surface (EPS) of formaldehyde. Once the EPS of a variety of molecules is known, free parameters in the various charge equilibration schemes can be adjusted such that the computed charges fit directly the ESP or the partial charges deduced from it. Right: Errors in partial charges for different charge equilibration models, when compared to charges obtained from quantum chemical calculations.

Solids automatically correspond to ideal metals in the QE formalism[12], making the claim questionable that reaxFF describes redox reactions, and (polar) molecules show the wrong scaling of the polarizability[11] (and dipole[13]) with P. None of these artifacts arise with SQE. They are due to QE being too polarizable at long distances or wavelengths (unless the system is metallic), while AACT, which intrinsically cannot handle metals, is too rigid or insulating at short wavelengths[12].

One of the problems in previous formulations of SQE, which did not yet contain the term $n_i \cdot e$ on the r.h.s. of Eq. 1, is that true ions cannot be formed for any realistic parameterization of the bond hardnesses. They should diverge once orbitals between neighboured atoms stop overlapping at large interatomic distances so that electropositive atoms automatically cease to donate their electrons to electronegative atoms. This shortcoming would make it impossible to simulate charging or decharging in galvanic cells, or more generally any type of processes involving redox reactions, as well as electrostatic fields of zwitterionic molecules, which violate the principle of local charge neutrality. Moreover, the partial charge of an atom cannot be made history dependent, because for any given atomic configuration, there will always be a unique minimum of V, which is linear and bilinear in the split charges.

Once *integer* transfer is introduced on top of the split charges, the just-mentioned problems should disappear. Atoms could now be assigned a formal (integer) oxidation state, whose dynamics could be parameterized so that they mimic a Landau-Zener process. The resulting dynamics can involve "radiation", i.e., a discontinuous (energy) change of the system. This could happen because the new oxidation state of the system necessitates a reoptimization of all split charges. However, a change of the oxidation state in the SQE model does not have to involve such discontinuities. For example, when the split charge q_{ij} starts to exceed $e/2$, the transformation $Q_i \to Q_i + 1$, $Q_j \to Q_j - 1$, $q_{ij} \to q_{ij} - 1$, would leave V invariant. This invariance makes possible the description of radiation-free redox reactions. They play an important role in Marcus theory[14], which was the first generally

accepted theory of electron transfer.

Implementing the just-mentioned ideas and extending the split charge model to integer charge transfer is certainly the most challenging task that the new NIC group faces. However, if this endeavor were successful, even on the modest level of qualitative demonstrators, force-field based MD could address many more problems than is currently possible. For example, to date, there is not a single force-field based MD method with which one could simulate (even on a qualitative level) the voltage-driven charging and decharging of galvanic elements, tribo-charging, or any other process in which the charge distribution of a system depends on its history.

4 Simulation of Fractal Contacts: Leakage through Seals

Seals are devices for closing a gap to prevent fluid leakage. Although seals play a crucial role in many devices, the design of seals is mostly based on experimental investigations, because theory has failed to reliably predict leakage rates. Current estimates are frequently based on the height distribution function of the rigid, rough substrate, against which the flat, elastically deformable seal is pressed. Given an estimated relative contact area (regular solids in mechanical contact only touch microscopically at a miniscule fraction of the apparent contact area), this distribution is assumed to be cut and shifted and the leakage rate is computed from the distribution function of the gap between substrate and seal.

The by far most problematic point in this procedure is the implicit assumption that the seal first forms contact where the substrate has its highest peak, next at the second highest point, and so on. In reality, the seal undergoes long-range elastic deformation. Therefore the distribution of the gap between seal and substrate is strongly perturbed from the simple shifted-and-cut substrate height distribution function.

Recent advances in the theoretical description of contact mechanics[15] and consequently the leakage through seals[16] were made by Persson. His theories are too involved to be described in much detail; only this much be said here: Assuming the laws for an elastic half space, Persson analyzes by how much the pressure distribution broadens when ever finer details of the substrate's surface topography are considered. This leads to a diffusion equation for the pressure distribution function, from which the real contact area can be derived as well as the gap distribution function. Once the gap distribution function, the pressure difference across the junction, and the viscosity of the fluid are given, one can estimate the flow with the effective medium theory by Bruggeman, which results in a self-consistent equation similar to the better known Clausius Mossotti relation for dielectrics.

While Persson theory appears to yield satisfactory results when compared to experiment[16], no test of his leakage theory has been conducted yet that would be free of uncontrolled approximations, experimental uncertainties, or that would be able to detect fortuitous cancellation of errors. Some crucial steps towards such a test, which were conducted by Lücke within the 2011 JSC guest student programme, shall be sketched here. Final results will be presented elsewhere[17,18].

The numerical modeling consists of two steps. First, an elastic calculation, which is based on the elastic energy V_{ela} of a half-space needs to be conducted. For a solid with Poisson number, $\nu = 0$, (the precise value of ν should not matter here), the elastic problem reduces to a scalar theory and V_{ela} can be written in the continuum limit as

$$V_{\mathrm{ela}} = \sum_{\mathbf{q}} \frac{qE'}{4} |\tilde{u}(\mathbf{q})|^2, \tag{3}$$

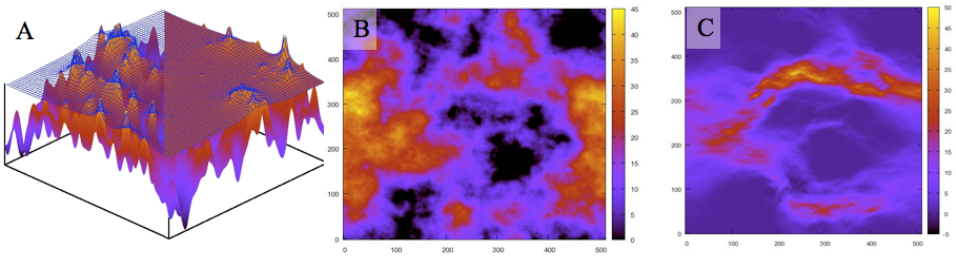


Figure 3. A: Small fraction (1/6400) of a Green's function molecular dynamics simulation in which a two-dimensional elastic manifold is pressed against a rough corrugated substrate. The elastic Green's function of the manifold are designed such that it mimics a full three-dimensional solid. B: Height map of a contact geometry that results from the contact between solids whose surfaces have self-affine roughness. C: Low Reynolds number flow through a geometry that is identical to the configuration shown in B.

where E' is the indentation modulus, $\tilde{u}(\mathbf{q})$ the Fourier transform of the normal surface displacement, and q is the magnitude of the in-plane wave vector $\mathbf{q}$. Given this expression, adding an external load, damping, and boundary conditions, the elastic half-space can be relaxed similar to the way in which it is done with Green's function MD[19], and the gap $g(x, y)$ be found as a function of the interfacial coordinates x, y (see also Fig. 3 A and B). Second, Reynolds lubrication equation needs to be solved, i.e., $\nabla \mathbf{j}(x, y) = 0$, where the surface current density $\mathbf{j}(x, y)$ is given by

$$\mathbf{j}(x, y) = \frac{g^3(x, y)}{12\eta} \nabla p, \tag{4}$$

where p is the fluid pressure in the interface. A solution for the x-component of the current density is shown in Fig. 3 C.

Simulations similar to the ones shown in Fig. 3 indeed confirm that Persson's prediction for the gap distribution function are quite accurate and that Bruggeman theory can be applied reasonably well to estimate the fluid flow. Thus, similar to Persson, we find a broad regime in which the total current decreases exponentially with load (and thus with contact area), while the simpler and frequently employed shift-and-cut approach underestimates the rate by more than a factor of two.

While the presented simulations are in the continuum limit, it is straightforward to include and parameterize system-specific properties, such as the slip length of the fluid at surfaces, shear thinning, and the pressure dependence of viscosity. The long-term goal is to make quantitative predictions for tribological contacts, where the used parameters are determined from the bottom up.

Acknowledgements

The calculations were carried out at the Jülich Supercomputing Centre and at Sharcnet, Canada.

References

1. S. R. Ovshinsky, *Reversible electrical switching phenomena in disordered structures*, Phys. Rev. Lett. **21**, 1450, 1968.
2. J. Akola and R. O. Jones, *Structural phase transitions on the nanoscale: The crucial pattern in the phase-change materials Ge(2)Sb(2)Te(5) and GeTe*, Phys. Rev. B **76**, 235201, 2007.
3. A. V. Kolobov et al., *Understanding the phase-change mechanism of rewritable optical media*, Nat. Mater. **3**, 703-708, 2004.
4. D. Shakhvorostov et al., *Evidence for electronic gap-driven metal-semiconductor transition in phase change materials*, Proc. Natl. Acad. Sci. USA **106**, 10907–10911, 2009.
5. M. H. Müser, *Elementary considerations on the local symmetry in optoelectronic materials and their phase change behavior*, Eur. Phys. J. B **74**, 291–302, 2010.
6. A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard III, *ReaxFF: A reactive force field for hydrocarbons*, J. Chem. Phys. A **105**, 9396–9409, 2001.
7. R. Chellii, P. Procacci, R. Righini, and S. Califano, *Electrical response in chemical potential equalization schemes*, J. Chem. Phys. **111**, 8569–8575, 1999.
8. R. A. Nistor, J. G. Polihronov, M. H. Müser, and N. J. Mosey, *A generalization of the charge equilibration method for non-metallic materials*, J. Chem. Phys. **125**, 094108, 2006.
9. T. Verstraelen, V. Van Speybroeck, and M. Waroquier, *The electronegativity equalization method and the split charge equilibration applied to organic systems: Parametrization, validation, and comparison*, J. Chem. Phys. **131**, 044127, 2009.
10. T. Verstraelen *et al.*, *The Significance of parameters in charge equilibration models*, J. Chem. Theo. Comp. **7**, 1750–1764, 2011.
11. G. L. Warren and J. E. Davis and S. Patel, *Origin and control of superlinear polarizability scaling in chemical potential equalization methods*, J. Chem. Phys. **128**, 144110, 2008.
12. R. A. Nistor and M. H. Müser, *Dielectric properties of solids in the regular and split-charge equilibration formalisms*, Phys. Rev. B **79**, 104303, 2009.
13. P. T. Mikulski, M. T. Knippenberg, and J. A. Harrison, *Merging bond-order potentials with charge equilibration*, J. Chem. Phys. **131**, 241105, 2009.
14. R. A. Marcus, *Chemical and Electrochemical Electron-Transfer Theory*, Annu. Rev. Phys. Chem. **15**, 155–196, 1964.
15. B. N. J. Persson, *Theory of rubber friction and contact mechanics*, J. Chem. Phys. **115**, 3840–3861, 2001.
16. B. Lorenz and B. N. J. Persson, *Leak rate of seals: Effective-medium theory and comparison with experiment*, Eur. Phys. J. E **31**, 159–167, 2010.
17. A. Lücke, *Viscous flow through fractal contacts*, M. Winkel (Ed.): *2011 Proceedings of the JSC Guest Student Programme on Scientific Computing*, FZJ-JSC-IB-2011-06, December 2011.
18. A. Lücke, M. H. Müser, N. Prodanov, and B. N. J. Persson (in preparation).
19. C. Campana and M. H. Müser, *Practical Green's function approach to the simulation of elastic, semi-infinite solids*, Phys. Rev. B **74**, 075420, 2006.

QCD Quantum Corrections from Computer Simulations

K. Jansen and R. Sommer

NIC, DESY Zeuthen, Platanenallee 6, D-15738 Zeuthen, Germany
E-mail: {karl.jansen, rainer.sommer}@desy.de

We report on activities of the NIC research group on elementary particle physics. After a brief overview, we focus on a determination of the hadronic vacuum polarization, needed in the theoretical computation of the anomalous magnetic moment of the muon. At present the comparison of theory and the ultraprecise experiment reveals a more than three standard deviations difference.

1 Overview

Particle Physics experiments, hand in hand with theory, have lead to the Standard Model (SM) of particle physics. It incorporates relativistic invariance and quantum mechanics in a quantum field theory. Fundamental principles dictate the possible form of interactions, leaving us with a theory with very few parameters and many predictions. Many laboratory experiments involving energy scales up to 10 TeV (the Large Hadron Collider at CERN) have been carried out, but only small deviations between predictions from the theory and experimental results have been found. All in all, these deviations are statistically insignificant. On the other hand, from astrophysical observations we know that there is a form of matter, which is not contained in the SM, dark matter, seen indirectly in the rotation speed of stars in distant galaxies. It also seems very hard to explain the abundance of matter over anti-matter in the Universe with just the SM and standard cosmology. There is too little CP-violation in the SM.

Consequently intense searches for small but significant deviations from the SM are being carried out. On the theoretical side of these searches, precise predictions of the model are needed. The strong interactions ("QCD") of quarks and gluons render it very difficult to obtain reliable predictions in many cases. Lattice gauge theory contributes to solving this riddle through numerical simulations of discretized QCD. The elementary particle physics group of the NIC carries out research in this field. Here we can't describe all activities of the group. Instead we first give a broad overview and then discuss one activity in detail.

Apart from the study of Higgs-Yukawa models[1–3], the NIC group in particular focuses on three aspects. The first are precise determinations of the fundamental parameters of QCD which are the basis for making predictions[4,5]. The second concerns the area of flavour physics, in particular heavy flavours, because particles with heavy quark flavours can decay through many different decay channels and these processes are sensitive to the structure and the fundamental parameters of the SM. In particular the group has developed a theoretical tool, non-perturbative HQET[6], which enables the simulation of B-mesons on the lattice. These boundstates of b quark and a light antiquark have particularly interesting decays where one may find physics beyond the SM. The NIC group has developed the tool to some perfection[7–9] and now is applying it to more and more realistic discretizations of the theory[10]. In the process, mathematical tools are being developed for the analysis of the

simulation data, which are applicable throughout lattice QCD[11,12]. Also more conventional paths for treating heavy flavours are applied[13].

In this report we focus on the third topic: indirect QCD effects through quantum fluctuations on the properties of leptons, which have no direct strong interactions. For the understanding of ultraprecise experiments this does play a role. We explain it in some detail in the following section.

Since the large-scale QCD simulations rely on the efficient use of supercomputer facilities, the group also carries out research on algorithms[14], as well as software and hardware aspects related to new computer architectures[15,16]. The numerical simulations of QCD are carried out by large collaborations. Two of these are coordinated by us, the ETM collaboration and the ALPHA collaboration.

2 The Anomalous Magnetic Moment of the Muon

In Dirac's theory the electron is described as a spin $1/2$ particle with the positron as its anti-particle. Like in the classical theory the angular momentum, the spin $\vec{S}$ leads to a magnetic moment of the electron

$$\vec{\mu}_m = -g_e \mu_0 \vec{S} \tag{1}$$

with $\mu_0 = e/4m_e$, e being the electric charge and m_e the mass of the electron. The quantity g_e is the gyro-magnetic ratio of the electron and assumes in Dirac's theory a value of $g_e = 2$.

When a quantum field theory is considered, deviations from $g_e = 2$ appear, originating from virtual photon exchange which leads to an *anomalous magnetic moment*

$$a_e = \frac{g_e - 2}{2} \ . \tag{2}$$

The most simple correction to g_e is depicted in Fig. 1(a). Here, one of the real photons, which represent the magnetic field, emits during the interaction with the charged electron a *virtual* photon. This correction, first considered by Schwinger[17], can be directly computed from quantum electrodynamics (QED), the quantum field theory of electromagnetic interaction. The seminal work by Schwinger revealed that the quantum correction $\delta(a_e)$ due to the diagram in Fig. 1(a) is

$$\delta(a_e) = \frac{\alpha}{2\pi} \ . \tag{3}$$

Experimental measurements[18] found a value for $g_e = 2.00238(10)$ and Schwinger's calculation yielded $g_e = 2.00232$. The agreement of these results established the theoretical and experimental fact that quantum corrections lead indeed to deviations from the classical value of $g_e = 2$. Subsequent more precise theoretical calculations and experimental measurement confirmed the agreement between QED predictions and observation, see Ref. 19. This agreement has been one of the first triumphs of quantum field theory giving much trust in this concept to describe the interactions of elementary particles in Nature.

Today, the theoretical efforts have been driven to enormously complicated calculations involving the computation of thousands of diagrams such as the ones shown in Fig. 1(b)

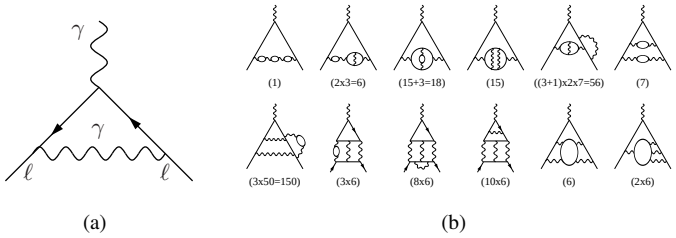


Figure 1. In (a) the leading order contribution to the electron anomalous magnetic moment is shown. In (b) a number of subdiagrams at the 8th order involving 4-loop contributions is shown.

which involve three loops. The calculations of Ref. 20 involve nowadays diagrams including four of such loops, we speak of a 4-loop order calculation and even a 5-loop order calculation is in sight. The theoretical computations are based on a systematic expansion of the theory in the fine structure constant α. As we saw above, the diagram in Fig. 1(a) appears at the order α, see Eq. 3, whereas the diagrams in Fig. 1(b) is a contribution at order α^4. Employing such high loop order quantum field theoretical calculations and extremely accurate and subtle experimental measurements, one finds an impressive agreement between experiment and and standard model predictions for the anomalous magnetic moment of the electron, i.e. (see e.g. Ref. 21)

$$\begin{aligned} a_e(\mathrm{th}) &= 1159652201.1(2.1)(27.1)\cdot 10^{-12} \\ a_e(\mathrm{exp}) &= 1159652188.4(4.3)\cdot 10^{-12}\ . \end{aligned} \tag{4}$$

This establishes the electron magnetic moment as one of the most precisely determined quantity in nature and the agreement between the two numbers in Eqs. 4 achieved with such an amazing accuracy is one of the cornerstone results to make us confident that indeed quantum field theories are the correct method to describe particle interactions.

However, the story does not end here. Nature has decided, for an unknown reason, that it comprises three particle generations. Besides the electron, which belongs to the first generation, there are further so-called leptons, the muon with a mass of about $105.7\mathrm{MeV}$ and the τ having a mass of $1777\mathrm{MeV}$. Comparing these values to the mass of the electron of $0.511\mathrm{MeV}$, we see that the leptons of the second (the muon) and the third (the τ) generations are much larger than the electron mass. This fact will play an important role in the discussion below. Being leptons, also the muon and τ carry spin leading to magnetic moments which become, due to the discussed quantum corrections, anomalous. For example, a comparison for the case of the muon anomalous magnetic moment, a_μ, between experiment and theory one finds[19]

$$\begin{aligned} a_\mu^{\mathrm{exp}} &= 1.16592080(63)\times 10^{-3} \\ a_\mu^{\mathrm{theory}} &= 1.16591790(65)\times 10^{-3}\ . \end{aligned} \tag{5}$$

Here the agreement between theory and experiment is not as nice as in the case of the electron and in fact, one finds $a_\mu^{\mathrm{exp}} - a_\mu^{\mathrm{th}} = 2.90(91)\times 10^{-9}$ which leads to a larger than 3σ level discrepancy.

However, this is a very interesting result. It means that either in the theoretical calculation something has been neglected or has not properly been included. Or, much more exciting, the discrepancy points to a breakdown of the standard model of particle interactions and the inconsistency stems from effects of some yet unknown new physics beyond the standard model.

Indeed, calculations show that these new physics effects would lead to a correction to the anomalous magnetic moment of size

$$\delta(a_l^{\rm newphysics}) \propto \alpha^2 m^2_{\rm lepton}/M^2_{\rm newphysics}\ . \tag{6}$$

Here $m_{\rm lepton}$ is the mass of one of the leptons and $M_{\rm newphysics}$ represents the mass (or scale) of a particle originating from the (unknown) new physics beyond the standard model. akingThe formula in Eq. 6 shows that in the case of the muon anomalous magnetic moment the effect of new physics would show up about $(m_\mu/m_e)^2 \approx 4\cdot 10^4$ times stronger than in the case of the electron. In principle, the τ-lepton would be even more suitable to detect these new physics effects, but unfortunately due to the very short lifetime of the τ lepton the experimental measurements of the anomalous magnetic moment of the τ are presently much too imprecise to unveil a possible new physics contribution. This leaves us then with the muon anomalous magnetic moment as the ideal place to look for new physics and indeed a large number of works has been devoted to explore this possibility, see Ref. 19.

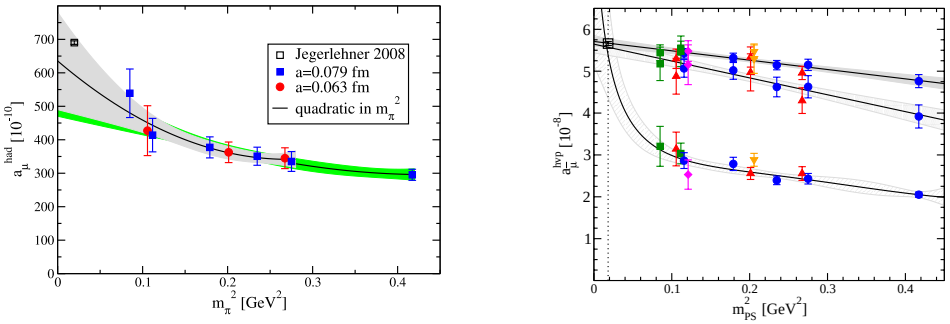
3 When the Lattice Enters the Game

As we discussed for the case of the electron, theoretically the anomalous magnetic moment is computed in an expansion of the theory in the fine structure constant to a given loop order. In the loops, in principle all particles of the standard model can appear, if allowed by symmetries and quantum numbers. Therefore, besides leptons and photons, also the vector bosons, W and Z, of the weak interaction and the quarks and gluons of the strong interaction, described by quantum chromodynamics (QCD) can and will contribute. It can be shown that the contribution of the W and Z bosons are first of all rather small and can be, secondly, well controlled by perturbation theory[19].

However, the strong interaction of quarks and gluons in QCD lead to so-called hadronic contributions $a_\mu^{\rm had}$ which are intrinsically of non-perturbative nature. Taking these contributions into account by perturbation theory is therefore rather doubtful. Employing additional model assumptions to estimate $a_\mu^{\rm had}$ will not provide a fully controlled and reliable calculation of the hadronic contributions and hence an unambiguous and stringent test whether the standard model is correct or must be extended by some new physics cannot be performed.

The standard way of taking $a_\mu^{\rm had}$ into account is by its relation to the probability of creating hadronic final states in the annihilation of electron and positron in so-called e^+e^- collisions. One then uses the experimental results for these probabilities (cross sections) depending on the energy of the e^+e^- pair, performs interpolations of these data and carries out the necessary integral over energy with a know weight function. At large energies the experimental data is replaced by the result obtained in perturbation theory. This phenomenological way of determining $a_\mu^{\rm had}$ will be called $a_\mu^{\rm had,exp}$.

It is exactly at this point where lattice field methods applied to quantum chromodynamics can help – at least in principle[22]. In lattice QCD the theory is formulated on a discrete



(a) (b)

Figure 2. In (a) we show a result for the hadronic contribution to the muon anomalous magnetic moment computed on the lattice in 2008. In (b) the present day status is represented using the improved observables introduced by members of the NIC group. The curves correspond to the defintions r_1, r_2 and r_3 in Eqs. 9 from bottom to top.

4-dimensional euclidean space-time lattice and the theory is approached by means of numerical simulations. It goes beyond the scope of this article to explain the mathematical concepts of lattice QCD but such numerical simulations allow then to compute physical quantities in a fully non-perturbative fashion without relying on any model assumptions or approximations.

Of course, the discretization itself induces a systematic error which must be removed by making the lattices finer and finer until the continuum of space time points is recovered by some suitable extrapolation process, a procedure which is called the *continuum limit*. In addition, the simulations necessarily demand a finite number of lattice points which can lead to *finite size effects* when the lattice is not large enough in physical units. Finally, often the simulations need to be performed at values of hadron masses that are larger than observed in nature. The reason is that for smaller and smaller hadron masses the computational costs increase rapidly such that one is restricted to values of, say, pion masses that are a factor of about two larger than the ones observed in nature.

All these systematic effects that appear in lattice simulations need to be controlled in a quantitative way. For example, to reach physical values of the pion masses, an extrapolation to the physical point where the pion mass assumes its physical value needs to be performed. This appeared to be very problematic in the past. It is illustrated in Fig. 2(a). The figure shows that the lattice results for $a_\mu^{\rm had}$ are significantly below the experimental number. An extrapolation to the physical point, reconciling the lattice data with experiment becomes in this situation very difficult and even needs some additional model assumptions. This all leads to an error of $a_\mu^{\rm had}$ as obtained from lattice simulations that is about a factor of 10 larger than the phenomenological one[23]. The lattice community have been therefore rather sceptical in the past that lattice QCD can provide a significant contribution to our understanding of the discrepancy in $g_\mu - 2$.

One additional suspicion has been that so-called dis-connected (singlet) contributions could be substantial. In all existing lattice calculations these contributions were neglected,

however. The reason is simply that these contributions are very noisy and therefore hard to compute reliably. Members of the NIC group[24] undertook a dedicated effort to calculate for the first time these contributions. As a result, it could be established that the dis-connected contributions are in fact small and can be safely neglected. In addition, also the effects of non-zero values of the lattice spacing and the finite volume turned out to be small. Thus the difficulty to reconcile lattice data with the experimental result, shown in Fig. 2(a), is rather puzzling.

A resolution of this puzzle was only given this year again by members of the NIC group[24]. They observed that by a suitable redefinition of the lattice observable needed to compute $a_\mu^{\rm had}$ a much smoother and much better controlled approach to the physical point can be achieved.

To illustrate the idea, let us give the definition of $a_\mu^{\rm had}$,

$$a_\mu^{\rm had} = \alpha^2 \int_0^\infty dQ^2 \frac{1}{Q^2} \omega(r) \Pi_R(Q^2) \; . \tag{7}$$

Here α is the electromagnetic coupling and $\Pi_R(Q^2)$ the renormalized vacuum polarization function, $\Pi_R(Q^2) = \Pi(Q^2) - \Pi(0)$. The functional form of $\omega(r)$ is analytically known and the argument r is given by $r = Q^2/m_\mu^2$ where m_μ denotes the mass of the muon and Q a generic momentum. The key observation is now that on the lattice there is a large freedom to choose a definition of r. The only requirement is that in the limit of reaching a physical pion mass the continuum definition of $r = Q^2/m_\mu^2$ is recovered. Hence, on may define

$$r_{\rm latt} = Q^2 \cdot \frac{H^{\rm phys}}{H} \tag{8}$$

with possible choices for H

$$\begin{aligned} r_1 : \quad & H = 1 \qquad H^{\rm phys} = 1/m_\mu^2 \\ r_2 : & H = m_V^2(m_{\rm PS}) \; H^{\rm phys} = m_\rho^2/m_\mu^2 \\ r_3 : & H = f_V^2(m_{\rm PS}) \; H^{\rm phys} = f_\rho^2/m_\mu^2 \; . \end{aligned} \tag{9}$$

Here, $m_V(m_{\rm PS})$ is the mass of the ρ-meson and $f_V(m_{\rm PS})$ the ρ-meson decay constant as determined on the lattice at unphysical pion masses $m_{\rm PS}$. Furthermore, m_ρ and f_ρ denote the corresponding ρ-meson mass and decay constant at the physical point. All the definitions in Eqs. 9 guarantee that indeed the desired definition of r is recovered in the limit of a physical pion mass since then by definition $m_V(m_{\rm PS})$ and $f_V(m_{\rm PS})$ assume their physical values. In Fig. 2(b) we show the results for $a_\mu^{\rm had}$ for all three definitions of r. Clearly, for the definitions r_2 and r_3 the behaviour of the lattice data towards physical pion masses is simply linear and allows for a controlled extrapolation to the physical point. As a result, one finds using definition r_2 in Eq. 9 values from the lattice computations and experiment

$$\begin{aligned} a_{\mu,N_f=2}^{\rm had,latt} &= 5.72\,(16) \cdot 10^{-8} \\ a_{\mu,N_f=2}^{\rm had,exp} &= 5.660\,(47) \cdot 10^{-8} \; . \end{aligned} \tag{10}$$

In the equations above, the index $N_f = 2$ indicates that in the lattice QCD calculations only a mass-degenerate pair of up and down quarks were used. Since therefore the strange and charm quark flavours were neglected the simulations do not correspond to a fully physical situation, a shortcoming that still needs to be overcome in the future.

In conclusion, using the modified and improved definitions of $a_\mu^{\rm had}$ on the lattice it is not only possible to recover the experimental result. As the comparison in Eq. 10 shows it is now also possible to come significantly closer to the experimental accuracy. The idea of the improved observables which led to a much reduced error for an important quantity such as $g_\mu - 2$ has been recognized in the lattice community by providing this work of the NIC group with the first Wilson lattice award stating: "*This is really a new application of lattice methods, applied in a timely fashion, and making an impact on an important (current) discrepancy for the muon magnetic moment*"[25].

As already mentioned above, the here discussed results for $a_\mu^{\rm had}$ have been achieved for the case of two mass-degenerate quark flavours. What is needed in the future is the inclusion of the strange and the charm quarks to allow for a direct comparison to the experimental results. In addition, newly planned experiments at Fermilab[26] and J-PARC[27] are aiming at an accuracy of below 0.5% for the hadronic contribution to the muon anomalous magnetic moment. To match this accuracy, lattice simulations have possibly to include explicit effects of isospin breaking and electromagnetism. All this is in principle reachable within lattice QCD but, it constitutes a real challenge for the lattice community. An even larger challenge are contributions to $g_\mu - 2$ that appear at higher order of the electromagnetic coupling, most notably the so-called light-by-light contributions. However, the lattice community is ready to attack these challenges and the NIC group will certainly contribute to help understanding whether the anomalous magnetic moment of the muon can provide signs of new physics beyond the standard model.

Acknowledgements

It is a pleasure to thank the members of the NIC group of the last two years, A. Athenodorou, F. Bernardoni, J. Bulava, S. Dinter, R. Donnellan, V. Drach, X. Feng, E. Garcia Ramos, O. Haas, G. Herdoiza, D. Hesse, J. Kallarackal, P. Kessel, J. Lopez Gonzalez, F. Mantovani, A. Nagy, A. Nube, D. Renner, H. Simma, F. Virotta, G. von Hippel and R. Welsing. They have made the progress described here – and much more – possible through their enthusiasm, collaborative attitude and creativity.

The calculations were carried out at the Jülich Supercomputing Centre.

References

1. P. Gerhold and K. Jansen. Lower Higgs boson mass bounds from a chirally invariant lattice Higgs-Yukawa model with overlap fermions. *JHEP*, 0907:025, 2009.
2. P. Gerhold and K. Jansen. Upper Higgs boson mass bounds from a chirally invariant lattice Higgs-Yukawa model. *JHEP*, 1004:094, 2010.
3. P. Gerhold, K. Jansen, and J. Kallarackal. Higgs boson mass bounds in the presence of a very heavy fourth quark generation. *JHEP*, 01:143, 2011.
4. Fatih Tekin, Rainer Sommer, and Ulli Wolff. The Running coupling of QCD with four flavors. *Nucl.Phys.*, B840:114–128, 2010.

5. R. Sommer. in: Workshop on Precision Measurements of alphas. 2011.
6. Jochen Heitger and Rainer Sommer. Non-perturbative heavy quark effective theory. *JHEP*, 02:022, 2004.
7. Benoit Blossier, Michele della Morte, Nicolas Garron, and Rainer Sommer. HQET at order $1/m$: I. Non-perturbative parameters in the quenched approximation. *JHEP*, 1006:002, 2010.
8. Benoit Blossier et al. HQET at order $1/m$: II. Spectroscopy in the quenched approximation. *JHEP*, 1005:074, 2010.
9. Benoit Blossier et al. HQET at order 1/m: III. Decay constants in the quenched approximation. *JHEP*, 1012:039, 2010.
10. Benoit Blossier et al. B meson spectrum and decay constant from Nf=2 simulations. *PoS*, LATTICE2010:308, 2010.
11. Benoit Blossier, Michele Della Morte, Georg von Hippel, Tereza Mendes, and Rainer Sommer. On the generalized eigenvalue method for energies and matrix elements in lattice field theory. *JHEP*, 04:094, 2009.
12. John Bulava, Michael Donnellan, and Rainer Sommer. On the computation of hadron-to-hadron transition matrix elements in lattice QCD. 2011.
13. B. Blossier et al. A proposal for B-physics on current lattices. *JHEP*, 04:049, 2010.
14. Stefan Schaefer, Rainer Sommer, and Francesco Virotta. Critical slowing down and error analysis in lattice QCD simulations. *Nucl.Phys.*, B845:93–119, 2011.
15. H. Simma. Compute engines for lattice QCD with commodity processors. *Nuovo Cim.*, B123:969–971.
16. Y. Nakamura, A. Nobile, D. Pleiter, H. Simma, T. Streuer, et al. Lattice QCD Applications on QPACE. 2011.
17. Julian S. Schwinger. On Quantum electrodynamics and the magnetic moment of the electron. *Phys.Rev.*, 73:416–417, 1948.
18. H.M. Foley and P. Kusch. On the Intrinsic Moment of the Electron. *Phys.Rev.*, 73:412–412, 1948.
19. Fred Jegerlehner and Andreas Nyffeler. The Muon g-2. *Phys.Rept.*, 477:1–110, 2009.
20. T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio. Revised value of the eighth-order QED contribution to the anomalous magnetic moment of the electron. *Phys.Rev.*, D77:053012, 2008.
21. F. Jegerlehner. The muon g-2: Status and perspectives. *Lect.Notes Phys.*, 745:9–34, 2008.
22. T. Blum. Lattice calculation of the lowest order hadronic contribution to the muon anomalous magnetic moment. *Phys.Rev.Lett.*, 91:052001, 2003.
23. Dru B. Renner and Xu Feng. Hadronic contribution to g-2 from twisted mass fermions. *PoS*, LATTICE2008:129, 2008.
24. Xu Feng, Karl Jansen, Marcus Petschlies, and Dru B. Renner. Two-flavor QCD correction to lepton magnetic moments at leading-order in the electromagnetic coupling. *arXiv:1103.4818*, 2011.
25. *https://kwla.llnl.gov/awards/2011.html*, 2011.
26. B.L. Roberts. Results and prospects for the measurement of muon (g-2). *Nucl.Phys.Proc.Suppl.*, 131:157–161, 2004.
27. Susumu Sato. J-PARC status, nuclear and particle physics. *AIP Conf.Proc.*, 1336:560–564, 2011.

The Early and Late Universe

Peter L. Biermann

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

Inst. Exp. Nucl. Phys., Karlsruher Institut für Technologie KIT, Germany

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

Dept. of Phys., Univ. of Alabama at Huntsville, AL, USA

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

Where do we come from and where do we go, is one of the oldest questions in human civilization. Here we highlight the search for the origin of magnetic fields in the early universe; the search for an understanding of the proof for the existence of dark energy, that drives the universe apart ever faster; and the search for an understanding of the origin of planets like Earth.

The universe is observed to be full of magnetic fields, and its distribution is extremely inhomogeneous. There are three possible classes of ideas where they come from, and all three must contain an element of truth: First of all, magnetic fields may and should derive from phase transitions in the very early universe; these magnetic fields ought to be enormously diluted by the time stars begin to form. Second, as soon as vorticity begins to appear in the large scale flows in the early universe electric current systems form, driving the creation of weak seed fields. Finally, as soon as stars form, tidal forces give rotation, and as soon as surfaces of constant density and constant density no longer coincide in a forming and young star, electric current systems appear, and also form seed fields. Then convection in a rotating star drives a dynamo, so strengthens the magnetic fields, and when the star ejects mass or even blows up, these magnetic fields are ejected to magnetize all the surroundings. Galaxies full of such stars then drive a magnetic wind from their magnetic interstellar medium, that in turn ejects the magnetic fields throughout the universe; central super-massive black hole also help through their relativistic jets and winds. This concept relying on stars is one of the oldest concepts, the simplest and appears to be quite sufficient to explain all what we know today. This last mechanism is explored in the first contribution.

The Nobel prize of 2011 was given for the discovery of dark energy, a force that pushes the universe into accelerated expansion. The critical step in the discovery was the empirical use of the properties of a certain class of supernovae (SN), called SN Ia, thought to be the end-point of white dwarf stars in binary systems, that exceed a famous stability limit, the Chandrasekhar limit. The light curve of these supernovae Ia can be empirically calibrated so as to yield a standard candle, allowing to measure distances, and so to determine the expansion of the universe. Using other data, especially from the very precise observations of the angular fluctuations in the residual microwave background, allows now to determine the parameters of the universe to very high precision. So the proof using these super-novae type Ia is a very critical ingredient. And yet, we do not yet understand why they are so readily calibrated, and what exactly the explosion mechanism is. The data on these supernovae seem to suggest many contradictions in their interpretation, and so it has become one

of the biggest challenges to comprehend their origin. The second contribution explores all the current ideas using very detailed modelling of of the evolution of white dwarf stars in binary systems.

We live on a planet in the Solar system, and by now many planets have been discovered, usually more massive than Earth, due to observational difficulties and selection effects; the question remains how planets form, and what their time scale is to form and sustain life. So there are many efforts to discover more planets, to understand our own Solar system better, to observe planetary disks around very young stars, and to model the formation and evolution of planets. Disks form around young stars due to angular momentum conservation and cooling, as noted already several centuries ago, with a first mathematical theory about sixty years ago. The material in these disks contains dust, that can agglomerate and so forms small clumps. How such clumps then evolve into small and large planets, solid bodies such as Earth and gaseous planets such as Jupiter, is still not fully understood. The third contribution explores these processes.

All the quests require the use of supercomputers.

Generation of Strong Magnetic Fields via the Small-Scale Dynamo During the Formation of the First Stars

Robi Banerjee[1], Sharanya Sur[2,3], Christoph Federrath[3,4,5], Dominik R. G. Schleicher[6], and Ralf S. Klessen[3]

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

[2] Inter-University Centre for Astronomy and Astrophysics
Post Bag 4, Ganeshkhind, Pune-411007, India

[3] Zentrum für Astronomie der Universität Heidelberg, Institut für Theoretische Astrophysik
Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany
E-mail: {sur, federrath, klessen}@uni-heidelberg.de

[4] Ecole Normale Supérieure de Lyon, CRAL, 69364 Lyon Cedex 07, France

[5] Monash University, Victoria 3800, Australia

[6] Institut für Astrophysik, Georg-August-Universität Göttingen
Friedrich-Hund-Platz 1, 37077 Göttingen, Germany
E-mail: dschleic@astro.physik.uni-goettingen.de

Here we summarize our recent results of high-resolution computer simulations on the turbulent amplification of weak magnetic seed fields showing that such fields will be exponentially amplified also during the gravitational collapse reminiscent to the situation during primordial star formation. The exponential magnetic field amplification is driven by the turbulent small-scale dynamo that can only be observed in computer simulations if the turbulent motions in the central core are sufficently resolved. We find that the Jeans length, which determines the central core region, has to be resolved by at least 30 grid cells to capture the dynamo activity. We conclude from our studies that strong magnetic fields will be unavoidably created already during the formation of the first stars in the Universe, potentially influencing their evolution and mass distribution.

1 Introduction

Magnetic fields are ubiquitous in the local Universe[1] and there is growing evidence of their presence also at high redshifts[2–4]. The seeds for these fields could be a relic from the early Universe, possibly arising due to inflation or some other phase transition process[5]. Alternatively, they could be generated by the Biermann battery[6,7] or the Weibel instability[8,9]. Regardless of their physical origin, most models predict weak field strengths and/or have large uncertainties[10,11]. Consequently, magnetic fields are thought to be irrelevant for the formation of the first stars and galaxies[12]. Here we show that strong and dynamically important magnetic fields will be generated from weak initial magnetic seed fields in the presence of the small-scale turbulent dynamo during the collapse of primordial halos[13,14]. The properties of the small-scale dynamo have been explored both in computer simulations of driven turbulence without self-gravity and in analytic models[15–18] as well as in the context of magnetic fields in galaxy clusters[19]. Analytic estimates show that the small-scale dynamo could be important already during the formation of the first stars and galaxies[20–22].

Field amplification via the small-scale dynamo requires both turbulent gas motions and high magnetic Reynolds numbers. The presence of such turbulence is suggested by cosmological hydrodynamical simulations of first star formation[23–25] where it also plays an important role in regulating the transport of angular momentum. High magnetic Reynolds numbers are expected in primordial gas as it follows closely the conditions of ideal magnetohydrodynamics (MHD).

2 Numerical Method and Initial Conditions

In this study, we focus on the gravitational collapse and magnetic field amplification of the inner parts of a contracting primordial halo. The initial conditions for our computer simulation were motivated from larger-scale cosmological models[23–25]. Thus, we set up a super-critical Bonnor-Ebert (BE) sphere with a core density of $\rho_{\rm BE} \simeq 4.68 \times 10^{-20}\,{\rm g\,cm^{-3}}$ ($n_{\rm BE} = 10^4\,{\rm cm^{-3}}$) and a small amount of rotation, $E_{\rm rot}/E_{\rm g} = 0.04$. We use a random initial velocity field with transonic velocity dispersion, and a weak random magnetic field with $B_{\rm rms} \sim 1{\rm nG}$. Both the turbulent energy and magnetic field spectra were initialized with the same power law dependence, $\propto k^{-2}$, with most power on large scales. Consistent with previous studies that follow the thermodynamics during the collapse, we adopt an effective equation of state with $\gamma = d\log T/d\log\rho + 1 = 1.1$ for all densities relevant to this study[26,27]. For this setup, we solve the equations of ideal magnetohydrodynamics (MHD) including self-gravity with an adaptive-mesh refinement technique[28], which guarantees that the critical length scale of gravitational collapse is always resolved with the same number of cells regardless of density. This scale, the local Jeans length

$$\lambda_{\rm J} = \left(\frac{\pi\, c_{\rm s}^2}{G\,\rho}\right)^{1/2}, \qquad (1)$$

where $c_{\rm s}$ and G are sound speed and gravitational constant, is set by the competition between gravity and thermal pressure. This is also the turbulent injection scale below which the small-scale dynamo is active. We note that the efficiency of the dynamo process depends on the Reynolds number and is thus related to how well the turbulent motions are resolved[29]. Higher resolution results in larger field amplification. To demonstrate this effect, we perform five computer simulations where we resolve $\lambda_{\rm J}$ by either 8, 16, 32, 64 or 128 cells. Our results indicate that a minimum resolution of 30 cells per Jeans length is required to see an exponential growth of the magnetic field.

3 Results

3.1 Physical Properties

We find that the dynamical evolution of the system is characterized by two distinct phases. First, as the initial turbulent velocity field decays the system exhibits weak oscillatory behaviour and contracts only slowly. Soon, however, the run-away collapse sets in. Fig. 1 shows a snapshot of the central region of the collapsing core in our highest-resolution simulation at a time when the central density has increased by a factor of $\sim 10^5$. The magnetic field strength has grown by a factor of 10^6, reaching peak values of about $1\,{\rm mG}$.

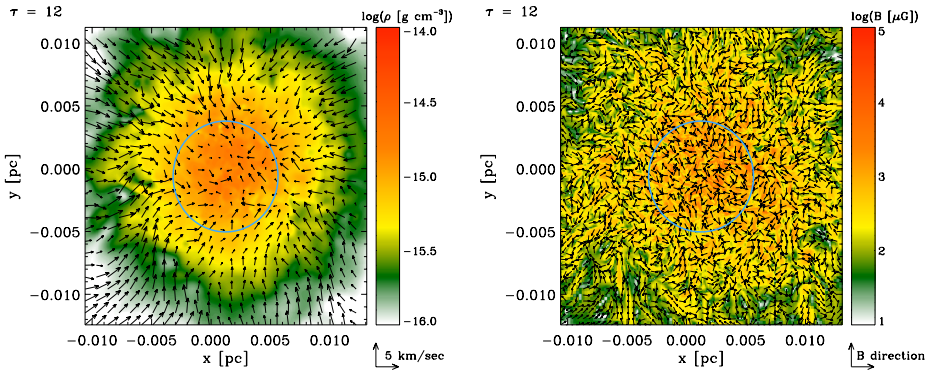


Figure 1. Two-dimensional slices through the centre of the collapsing core at the time when the initial field strength has increased by a factor of $\sim 10^6$, showing the central region of about $0.02 \times 0.02\,\mathrm{pc}^2$ in size for our highest-resolution simulation (λ_J resolved by 128 cells). The circle indicates the central averaging volume V_J. The left image shows the density and the velocity component in the xy-plane, indicating radial infall in the outer regions and turbulent motions in the inner core. The right image depicts the total magnetic field amplitude and direction. [Figure taken from Ref. 13]

The left image shows density and velocity structure, and the right image shows magnetic field strength and morphology.

To understand the behaviour of the system more quantitatively, we need to account for its dynamical contraction. First, we note that the physical time scale becomes progressively shorter during collapse. We therefore define a new time coordinate τ,

$$\tau = \int dt/t_\mathrm{ff}(t)\,, \tag{2}$$

based on the local free-fall time, $t_\mathrm{ff}(t) = \sqrt{3\pi/(32G\rho_\mathrm{m}(t))}$, where $\rho_m(t)$ is the mean density of the contracting central region. We can also define a critical volume for gravitational collapse, the so-called Jeans volume, $V_\mathrm{J} = 4\pi(\lambda_\mathrm{J}/2)^3/3$, with λ_J given by Eq. 1. We obtain all dynamical quantities of interest as averages within the central Jeans volume. This approach ensures that we always average over the relevant volume for collapse and field amplification.

Gravitational compression during the collapse of a primordial gas cloud can at most lead to an amplification of the magnetic field strength by a factor of $\sim \rho^{2/3}$ in the limit of perfect flux freezing (i.e., ideal MHD). A stronger increase implies the presence of an additional amplification mechanism. Starting from an initial field strength of $\sim 1\mathrm{nG}$, our simulations show a total magnetic field amplification by six orders of magnitude, leading to a field strength of about $\sim 1\,\mathrm{mG}$ for the case where we resolve the local Jeans length by 128 cells. This is illustrated in Fig. 2a. Fig. 2b shows that the obtained field amplification is indeed stronger than what is expected for purely adiabatic compression, which demonstrates that the small-scale turbulent dynamo provides significant additional field amplification over compression. A closer comparison of Fig. 2a and 2b shows that the amplification of the field by compression and by the turbulent dynamo are roughly comparable. The time evolution of the central density ρ_m is depicted in Fig. 2c, while Fig. 2d shows the corresponding rms velocity. The presence of turbulence delays the collapse until

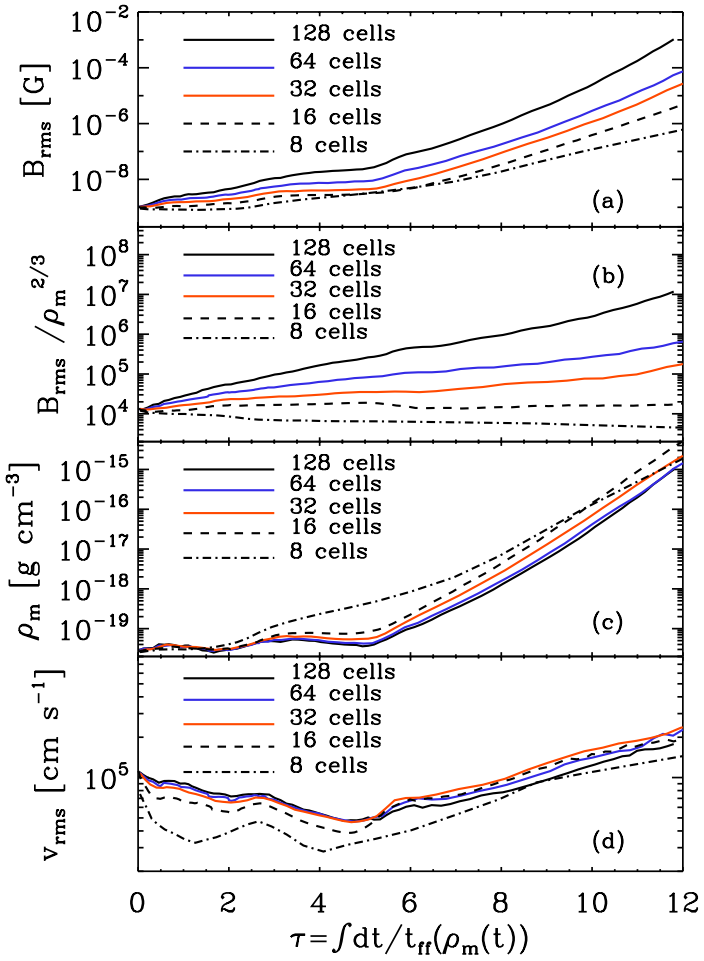


Figure 2. Evolution of the dynamical quantities as a function of $\tau = \int dt/t_{\rm ff}$, defined in Eq. 2 for five runs with different number of cells to resolve the local Jeans length. Panel (a) - the rms magnetic field strength $B_{\rm rms}$, amplified to 1 mG from an initial field strength of 1 nG, (b) - the evolution of $B_{\rm rms}/\rho_{\rm m}^{2/3}$, showing the turbulent dynamo amplification by dividing out the maximum possible amplification due to pure compression of field lines, (c) - the evolution of the mean density $\rho_{\rm m}$ and (d) - the rms velocity $v_{\rm rms}$. The onset of runaway collapse commences at about $\tau \sim 6$. [Figure taken from Ref. 13]

$\tau \sim 6$. During this time, the mean density shows some oscillations while the rms velocity decreases as the turbulence decays. A comparison between Fig. 2b and Fig. 2d indicates that dynamo amplification takes place throughout the entire duration of the simulation, that is during the initial phase of turbulent decay, i.e., for $\tau \lesssim 6$, as well as during the run-away collapse phase ($\tau \gtrsim 6$). We note that the time constant for the field amplification is roughly the same in both regimes.

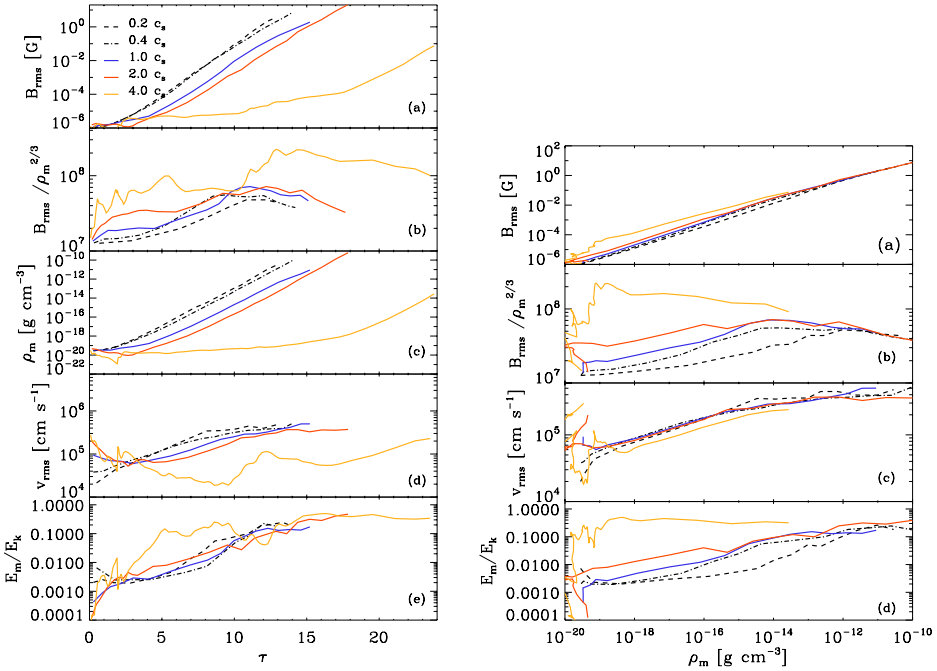


Figure 3. Evolution of the dynamical quantities in the case of stronger initial fields ($B = 10^{-6}$ G) within the central Jeans volume as a function of τ (left) and the mean density ρ_m (right) for runs with different initial Mach numbers. Shown are the rms magnetic field strength $B_{\rm rms}$, the ratio $B_{\rm rms}/\rho_m^{2/3}$, the mean density (only in the left panel), the rms velocity $v_{\rm rms}$ and the ratio of magnetic to kinetic energy, $E_{\rm m}/E_{\rm k}$. All the simulations correspond to a resolution of 128 cells per Jeans length. [Figures taken from Ref. 30]

3.2 Implications for Numerical Resolution

We also point out, that sufficient numerical resolution is a crucial issue when studying the small-scale turbulent dynamo. We see that a minimum of 30 cells per Jeans length is required to unambiguously identify the exponential growth of the magnetic field[14,18]. Seeing the dynamo process is computationally extremely demanding, and even our 128^3 high-resolution run is by no means numerically converged. This reflects the fact that no numerical scheme to date is able to reach the enormous Reynolds numbers, which determines the dynamo growth rate, of star-forming, turbulent gas. Therefore, these simulations show no signs of saturation, thus we simply stop the calculation when the numerical cost becomes prohibitively high. In our subsequent study[30] we investigate the saturation behaviour when back-reactions of the Lorentz force become important. Here we find typical values of $E_{\rm mag}/E_{\rm kin}$ of $0.2 - 0.4$ (see Fig. 3) in agreement with calculations of non-selfgravitating MHD turbulence[31,18]. The physical dissipation scales are much smaller than the Jeans length and thus the growth rates obtained in our simulations are lower limits on the physical growth rates. In Fig. 4 we visualize various quantities of our 128^3 resolution run to give an impression of the complex structure within the central core.

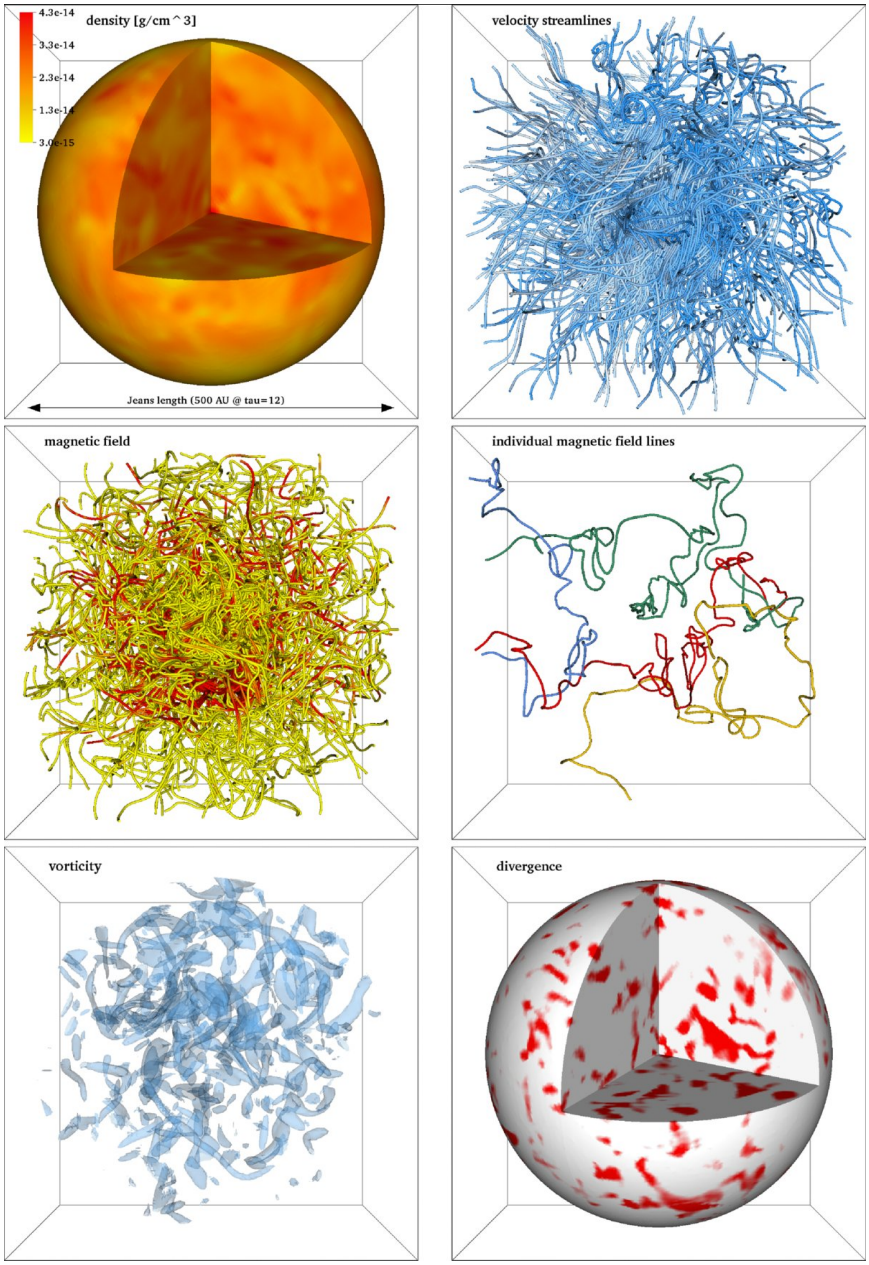


Figure 4. (a) Spherical slice of the gas density inside the Jeans volume at $\tau = 12$ for our run with 128 cells per Jeans length. (b) Velocity streamlines on a linear colour scale ranging from dark blue (0 km s^{-1}) to light grey (5 km s^{-1}). (c) Magnetic field lines, showing a highly tangled and twisted magnetic field structure typical of the small-scale dynamo; yellow: 0.5 mG, red: 1 mG. (d) Four randomly chosen, individual field lines. The green one, in particular, is extremely tangled close to the centre of the Jeans volume. (e) Contours of the vorticity modulus, showing elongated, filamentary structures typically seen in subsonic turbulence. (f) Spherical slice of the divergence of the velocity field, white: compression, red: expansion. [Figure taken from Ref. 14]

4 Conclusions

Taken all together, our results strongly indicate that dynamically important magnetic fields are generated during the formation of the first stars. This has important consequences for our understanding of how the first stars form and how they influence subsequent cosmic evolution. We know from modeling galactic star-forming clouds that the presence of magnetic fields can reduce the level of fragmentation, and by doing so strongly influences the stellar mass spectrum[32]. Furthermore, we also know that dynamos in accretion disks produce jets and outflows[33], which remove a signicant fraction of the mass and angular momentum. Again, this influences the stellar mass spectrum. There are first attempts to study this process in the context of first star formation[34], but more sophisticated initial conditions and more appropriate magnetic field geometries need to be considered. Once the first stars have formed, they are likely to produce a copious amount of ionizing photons, which drive huge HII regions, bubbles of ionized gas, expanding into the low-density gas between the halos. These dynamics could be substantially different if magnetic outflows drive a cavity-wave into the surrounding gas. The magnetic field may further affect fluid instabilities near the ionization front.

The mechanisms discussed here are likely to work not only during the formation of the first stars, but in all types of gravitationally bound, turbulent objects. Highly magnetized gas is thus expected already in the first galaxies, which subsequently needs to become more coherent to form the typical field structures that are observed in the present-day Universe.

Acknowledgements

R.B. acknowledges fundings from the German Science Foundation (DFG) through the grant BA 3706. S.S. thanks the DFG for financial support via the priority program 1177 "Witnesses of Cosmic History: Formation and Evolution of Black Holes, Galaxies and their Environment" (grant KL 1358/10). C.F. has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 247060) and from a Discovery Projects Fellowship (DP110102191) of the Australian Research Council. C.F., R.B., and R.S.K. acknowledge subsidies from the Baden-Württemberg-Stiftung (grant P-LS-SPII/18) and from the German Bundesministerium für Bildung und Forschung via the ASTRONET project STAR FORMAT (grant 05A09VHA). D.S. thanks for funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No 229517. Supercomputing time at the Forschungszentrum Jülich (projects hhd20 and hhd14) are gratefully acknowledged. The software used in this work was in part developed by the DOE-supported ASC / Alliance Center for Astrophysical Thermonuclear Flashes at the University of Chicago. Fig. 4 was produced with the open-source visualization software VISIT.

References

1. Beck, R., Brandenburg, A., Moss, D., Shukurov, A. & Sokoloff, D. Galactic Magnetism: Recent Developments and Perspectives. *Ann. Rev. Astron. Astrophys.* **34**, 155–206 (1996).

2. Bernet, M. L., Miniati, F., Lilly, S. J., Kronberg, P. P. & Dessauges-Zavadsky, M. Strong magnetic fields in normal galaxies at high redshift. *Nature* **454**, 302–304 (2008). 0807.3347.
3. Robishaw, T., Quataert, E. & Heiles, C. Extragalactic Zeeman Detections in OH Megamasers. *Astrophys. J.* **680**, 981–998 (2008). 0803.1832.
4. Murphy, E. J. The Far-Infrared-Radio Correlation at High Redshifts: Physical Considerations and Prospects for the Square Kilometer Array. *Astrophys. J.* **706**, 482–496 (2009). 0910.0011.
5. Turner, M. S. & Widrow, L. M. Inflation-produced, large-scale magnetic fields. *Phys. Rev. D* **37**, 2743–2754 (1988).
6. Biermann, L. Über den Ursprung der Magnetfelder auf Sternen und im interstellaren Raum (miteinem Anhang von A. Schlüter). *Z. Naturforsch* **5**, 65–+ (1950).
7. Xu, H. *et al.* The Biermann Battery in Cosmological MHD Simulations of Population III Star Formation. *Astrophys. J.* **688**, L57–L60 (2008). 0807.2647.
8. Schlickeiser, R. & Shukla, P. K. Cosmological Magnetic Field Generation by the Weibel Instability. *Astrophys. J.* **599**, L57–L60 (2003).
9. Medvedev, M. V., Silva, L. O., Fiore, M., Fonseca, R. A. & Mori, W. B. Generation of Magnetic Fields in Cosmological Shocks. *Journ. Korean. Astron. Soc.* **37**, 533–541 (2004).
10. Grasso, D. & Rubinstein, H. R. Magnetic fields in the early Universe. *Phys. Rep.* **348**, 163–266 (2001). arXiv:astro-ph/0009061.
11. Banerjee, R. & Jedamzik, K. Evolution of cosmic magnetic fields: From the very early Universe, to recombination, to the present. *Phys. Rev. D* **70**, 123003–+ (2004). arXiv:astro-ph/0410032.
12. Bromm, V., Yoshida, N., Hernquist, L. & McKee, C. F. The formation of the first stars and galaxies. *Nature* **459**, 49–54 (2009). 0905.0929.
13. Sur, S., Schleicher, D. R. G., Banerjee, R., Federrath, C. & Klessen, R. S. The Generation of Strong Magnetic Fields During the Formation of the First Stars. *Astrophys. J., Letters* **721**, L134–L138 (2010). 1008.3481.
14. Federrath, C., Sur, S., Schleicher, D. R. G., Banerjee, R. & Klessen, R. S. A New Jeans Resolution Criterion for (M)HD Simulations of Self-gravitating Gas: Application to Magnetic Field Amplification by Gravity-driven Turbulence. *Astrophys. J.* **731**, 62–+ (2011). 1102.0266.
15. Haugen, N. E. L., Brandenburg, A. & Dobler, W. High-Resolution Simulations of Nonhelical MHD Turbulence. *Astrophysics and Space Science* **292**, 53–60 (2004). arXiv:astro-ph/0306453.
16. Schekochihin, A. A., Cowley, S. C., Taylor, S. F., Maron, J. L. & McWilliams, J. C. Simulations of the Small-Scale Turbulent Dynamo. *Astrophys. J.* **612**, 276–307 (2004). arXiv:astro-ph/0312046.
17. Brandenburg, A. & Subramanian, K. Astrophysical magnetic fields and nonlinear dynamo theory. *Phys. Rep.* **417**, 1–209 (2005). arXiv:astro-ph/0405052.
18. Federrath, C. *et al.* Mach Number Dependence of Turbulent Magnetic Field Amplification: Solenoidal versus Compressive Flows. *Phys. Rev. Lett.* **107**, 114504–+ (2011). 1109.1760.
19. Xu, H., Li, H., Collins, D. C., Li, S. & Norman, M. L. Turbulence and Dynamo in Galaxy Cluster Medium: Implications on the Origin of Cluster Magnetic Fields. *Astrophys. J., Letters* **698**, L14–L17 (2009). 0905.2196.

20. Arshakian, T. G., Beck, R., Krause, M. & Sokoloff, D. Evolution of magnetic fields in galaxies and future observational tests with the Square Kilometre Array. *Astron. & Astrophys.* **494**, 21–32 (2009). 0810.3114.
21. Schleicher, D. R. G. *et al.* Small-scale dynamo action during the formation of the first stars and galaxies. I. The ideal MHD limit. *Astron. & Astrophys.* **522**, A115+ (2010). 1003.1135.
22. de Souza, R. S. & Opher, R. Origin of magnetic fields in galaxies. *Phys. Rev. D* **81**, 067301–+ (2010). 0910.5248.
23. Abel, T., Bryan, G. L. & Norman, M. L. The Formation of the First Star in the Universe. *Science* **295**, 93–98 (2002). arXiv:astro-ph/0112088.
24. Bromm, V., Coppi, P. S. & Larson, R. B. The Formation of the First Stars. I. The Primordial Star-forming Cloud. *Astrophys. J.* **564**, 23–51 (2002). arXiv:astro-ph/0102503.
25. Yoshida, N., Omukai, K. & Hernquist, L. Protostar Formation in the Early Universe. *Science* **321**, 669– (2008). 0807.4928.
26. Omukai, K., Tsuribe, T., Schneider, R. & Ferrara, A. Thermal and Fragmentation Properties of Star-forming Clouds in Low-Metallicity Environments. *Astrophys. J.* **626**, 627–643 (2005). arXiv:astro-ph/0503010.
27. Glover, S. C. O. & Savin, D. W. Is H_3^+ cooling ever important in primordial gas? *Mon. Not. R. Astron. Soc.* **393**, 911–948 (2009). 0809.0780.
28. Fryxell, B. *et al.* FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes. *Astrophys. J. Suppl. Ser.* **131**, 273–334 (2000).
29. Haugen, N. E., Brandenburg, A. & Dobler, W. Simulations of nonhelical hydromagnetic turbulence. *Phys. Rev. E* **70**, 016308–+ (2004). arXiv:astro-ph/0307059.
30. Sur, C., S. Federrath, Schleicher, D. R. G., Banerjee, R. & Klessen, R. S. Magnetic field amplification by gravity-driven turbulence – Influence of initial conditions and saturation field strength. *submitted to Mon. Not. R. Astron. Soc.* (2011).
31. Subramanian, K. Unified Treatment of Small- and Large-Scale Dynamos in Helical Turbulence. *Phys. Rev. Lett.* **83**, 2957–2960 (1999). arXiv:astro-ph/9908280.
32. Hennebelle, P. & Teyssier, R. Magnetic processes in a collapsing dense core. II. Fragmentation. Is there a fragmentation crisis? *Astron. & Astrophys.* **477**, 25–34 (2008). 0709.2887.
33. von Rekowski, B., Brandenburg, A., Dobler, W., Dobler, W. & Shukurov, A. Structured outflow from a dynamo active accretion disc. *Astron. & Astrophys.* **398**, 825–844 (2003).
34. Machida, M. N., Omukai, K., Matsumoto, T. & Inutsuka, S. The First Jets in the Universe: Protostellar Jets from the First Stars. *Astrophys. J.* **647**, L1–L4 (2006). arXiv:astro-ph/0605146.

The Progenitor Systems of Thermonuclear Supernovae

Wolfgang Hillebrandt[1], **Michael Fink**[1,2], **Markus Kromer**[1], **Rüdiger Pakmor**[3], **Friedrich K. Röpke**[1,2], **Ivo R. Seitenzahl**[1,2], **and Stuart A. Sim**[4]

[1] Max-Planck Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany
E-mail: wfh@mpa-garching.mpg.de

[2] Universität Würzburg, Emil-Fischer-Str. 31, D-97074 Würzburg, Germany
E-mail: fritz@mpa-garching.mpg.de

[3] Heidelberger Institut für Theoretische Studien
Schloss-Wolfsbrunnenweg 35, D-69118 Heidelberg, Germany
E-mail: ruediger.pakmor@h-its.org

[4] Research School of Astronomy and Astrophysics, Mount Stromlo Observatory
Weston Creek, ACT 2611, Australia
E-mail: ssim@mso.anu.edu.au

Recent progress in modeling thermonuclear ("Type Ia") supernovae by means of 3-dimensional simulations as well as several of the still open questions are discussed. It is shown that the new models have considerable predictive power allowing us to study observable properties of supernovae such as their light curves and spectra without adjustable non-physical parameters. This is a basic prerequisite to improve our understanding of the explosion mechanism and the progenitor systems and to settle the question of the applicability of SNe Ia as distance indicators for cosmology. We explore the capabilities of the models by comparison with observations and show how such models can be applied to study the origin of the observed diversity of SNe Ia, which could be a source of considerable systematic errors in their distances.

1 Introduction

During the past couple of years Type Ia supernovae (SNe Ia), i.e., supernovae that lack lines of hydrogen in their spectra and which are thermonuclear explosions of rather compact stars, presumably White Dwarfs (WDs), have become a major tool in observational cosmology through their role as distance indicators. The now generally accepted view that the Universe entered a phase of accelerated expansion a few billion years ago, caused by either a cosmological constant or an unknown form of "dark energy" with negative pressure, rests mainly on the interpretation of supernova luminosities. While, however, the number of supernovae observed at cosmological distances is steadily increasing (see Goobar & Leibundgut[1] for a recent review), thus reducing the statistical errors of the cosmological parameters derived from them, a sound theoretical understanding of these objects – justifying in particular the calibration techniques applied in distance measurements – is still lacking.

Major uncertainties in using supernova distances (as determined from their apparent luminosity) are effects caused by extinction and evolution. Extinction through absorption by dust is usually corrected for by means of their observed colour: dust absorption makes the spectrum redder. However, if the more distant (and therefore "older") supernovae have intrinsic colours different from their local counterparts, we would be fooled. Similarly, if there are different kinds of progenitors with somewhat different luminosity and a relative

rate changing with time this would lead to wrong (average) distances. Another potentially important effect that may modify the luminosity-distance relationship is the chemical composition of the progenitor star that also may, on average, change with time.

Here we give a progress report on a project, started at the Max-Planck-Institut für Astrophysik several years ago, with the objective to contribute to the understanding of SNe Ia by developing a theoretical framework consistent with prototypical observed explosions, but also capable of reproducing their observed diversity.

In the sections that follow we first discuss the so-called "Chandrasekhar mass" models (and their observable predictions) in some detail, and in a second part alternative models that have become more popular again recently. The latter are based on either WDs exceeding the Chandrasekhar mass, for instance by merging two WDs dynamically, or on WDs accreting helium-rich material from a companion and being disrupted by a "double detonation" before they even reach the Chandrasekhar mass.

2 Chandrasekhar Mass Explosion Models

The most popular progenitor model for the average Type Ia supernova is a massive white dwarf, consisting of carbon and oxygen, which approaches the Chandrasekhar mass ($M_{Chan} \simeq 1.4\ M_\odot$) by a yet unknown mechanism, presumably accretion from a companion star, and is disrupted by a thermonuclear explosion (see, e.g., Hillebrandt & Niemeyer[2] for a review). At high densities explosive carbon burning mostly produces radioactive ^{56}Ni. At lower densities intermediate-mass nuclei, like ^{28}Si, are produced. These elements give rise to the typical observed spectra of SNe Ia, which are dominated by lines of Fe, Si and S.

A generic property of such an explosion is that first carbon fuses rather quietly to intermediate-mass nuclei in the core of the contracting WD. Because this core is convectively unstable, temperature fluctuations will be present, which may locally reach run-away values, above 2×10^9K. After ignition, the "flame" is thought to propagate through the star as a sub-sonic turbulent deflagration wave that may or may not change into a detonation at low densities (around 10^7g/cm^3), disrupting the star in the end in both cases, the physics of which is very similar to pre-mixed chemical combustion.

Early attempts to model SNe Ia were based on one-dimensional numerical simulations. Such models gave valuable insight into the basic mechanism of the explosions. However, their predictive power was limited due to the fact that underlying physical processes enter the models in a parametrized way. In particular, the velocity of the thermonuclear flame front entered as a free parameter. However, recent models are in 3D and largely avoid such parameters.

2.1 Modeling Turbulent Thermonuclear Combustion in White Dwarfs

In practice, computing the propagation of a nuclear flame in dense matter is not quite as complicated as it appears. If the turbulence is sufficiently strong the turbulent flame speed becomes independent of the laminar speed, and therefore of the microphysics of burning and diffusion, and scales only with the velocity of the largest turbulent eddy. At high densities, near the centre of the WD, the thermal width of the thermonuclear combustion front is very small, of the order of a fraction of a millimeter, and the turbulence intensity is moderate. The flame is said to be in the "flamelet"-regime and can be modelled by means of front

capturing/tracking schemes, coupled to a turbulence model[3,4]. In this approach the front is tracked by a level set, and the thermodynamic properties of the matter before and after the front are reconstructed from the jump conditions for mass, energy, and momentum.

As the density of the WD material declines and the laminar flamelets become slower and thicker, it is plausible that at some point turbulence significantly alters the thermal flame structure. So far, modeling this so-called distributed burning regime in exploding white dwarfs has not been attempted explicitly since neither nuclear burning and diffusion nor turbulent mixing can be properly described by simplified prescriptions. However, it is this regime where the transition from deflagration to detonation is assumed to happen in certain phenomenological models.

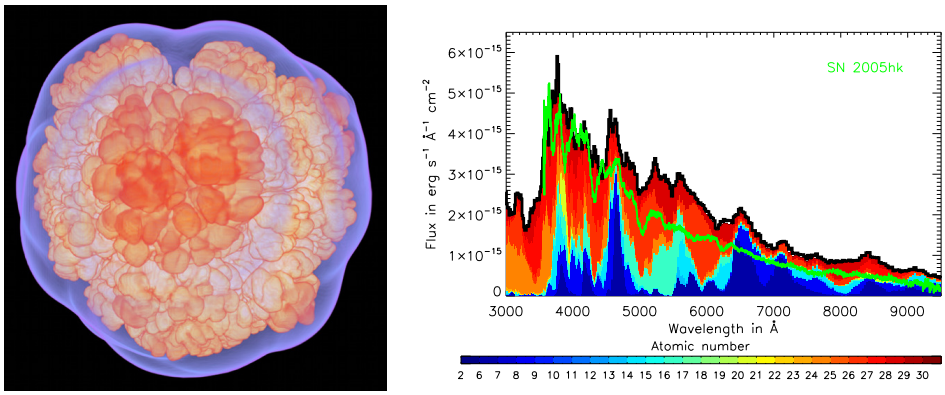


Figure 1. *left:* Snapshot from a full star SN Ia pure deflagration simulation at the end of the burning phase (1.6 s after ignition in a multi-spot scenario)[5]. The WD star is shown in blue and the isosurface corresponds to the thermonuclear flame. *right:* A synthetic spectrum (black solid line) derived from this model with the Monte Carlo radiative transfer code ARTIS[6] compared to the spectrum of SN 2005hk near maximum brightness. The region below the synthetic spectrum is colour coded to indicate the fraction of escaping quanta in each wavelength bin which last interacted with a particular element (for the associated atomic numbers see the colour bar).

2.2 Some Results of Numerical Simulations

The outcome of a prototypical example of a simulation modeling turbulent deflagration in a Chandrasekhar mass WD is shown in Fig. 1. Initially, the deflagration flame is ignited in multiple sparks around the centre of the WD. The mushroom-shaped features due to the buoyancy instability are clearly visible as well as the wrinkling of the flame at later times. This model produces $0.32\,\mathrm{M}_\odot$ of ^{56}Ni – a value that is typical for centrally ignited turbulent deflagrations, but insufficient for the explanation of normal SNe Ia. The large-scale buoyancy instabilities lead to strong mixing of the chemical composition of the ejecta that is not observed in normal SNe Ia [7]. Thus, the simulations show that only faint, low energy, peculiar events are expected from pure deflagrations in Chandrasekhar mass WDs (see also Sect. 4).

A modification that brings the models closer to the observations is a transition of the mode of thermonuclear burning from a subsonic deflagration to a supersonic detonation

in the course of the explosion, called the "delayed detonation" scenario[8]. Since here the detonation propagates through a pre-expanded star that is brought out of hydrostatic equilibrium by the preceding deflagration phase, it produces a layer of intermediate-mass elements that encompasses almost all of the outer ejecta, in agreement with observations[9,10]. An open question is, however, whether or not such deflagration-to-detonation transitions (DDTs) occur in SNe Ia, but recent studies indicate that they may be possible[11,12].

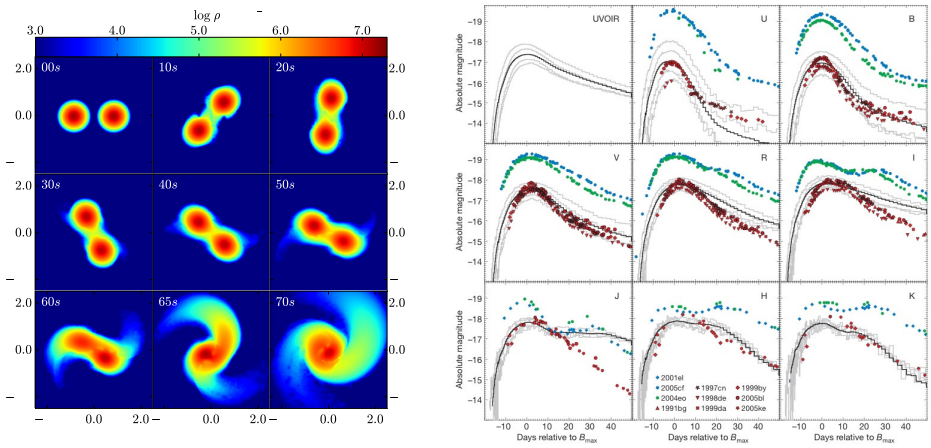


Figure 2. *left:* Snapshots of the evolution of a white-dwarf binary system during the inspiral[19]. The system consists of two $0.9 M_\odot$ WDs. Colour-coded is the logarithm of the density. *right:* Synthetic light curves (grey lines, black solid line: average over all viewing angles) derived from this model[19] with the radiation transfer code ARTIS in comparison with light curves of normal SNe Ia (green, blue) and SN 1991bg-like objects (red).

3 Non-Chandrasekhar Mass Explosion Models

Non-Chandrasekhar mass explosion models come in two flavours. In sub-Chandrasekhar mass models, a white dwarf accretes helium from a binary companion (which may be a non-degenerate He star or a He WD). Above the C+O core, a He layer builds up and when it becomes sufficiently massive, a detonation triggers by compression. This detonation burns the He layer and drives a shock wave into the core. If strong enough, the shock wave may initiate a secondary detonation there (either at the interface between the core and the He shell or near the centre of the core). Sub-Chandrasekhar mass models were first studied by Woosley & Weaver[13] and Livne & Arnett[14] but were not investigated further because the lightcurves and spectra predicted by those models seemed to contradict observations. From a population synthesis point of view, however, sub-Chandrasekhar mass explosions could perhaps even account for the bulk of normal SNe Ia[15,16]and, as recent multi-dimensional simulations have shown, some of the earlier concerns may not be justified[17,18] (see also Sect. 4).

In contrast, super-Chandrasekhar mass explosions can result if two C+O WDs merge. If the masses of the WDs are unequal the lighter star is disrupted by tidal forces and may

form an accretion disk around the heavier companion. Such WD-WD mergers are not expected to lead to thermonuclear explosions in general, but the parameter space for WD-WD mergers is rich and has not been fully explored yet.

In contrast, explosions are more likely in the special case of the merger of two WDs of nearly equal masses. Due to the symmetry in the initial setup, a break-up of the lighter star is avoided. Instead, the stars merge violently within a few orbits. In the violent merger scenario[19,20], this leads to a thermonuclear explosion and disruption of both WDs (see Fig. 2).

4 A Few Observable Predictions

Again, we start with a discussion of Chandrasekhar mass models. Apart from the initial conditions, simulations like those described above contain no free parameters. Therefore we may ask whether such models are capable of reproducing observations without any fitting. The explosion energies achievable in the outlined scenarios reach up to $\sim 8 \times 10^{50}\,\mathrm{erg}$ for pure deflagrations and can be up to 50% higher for delayed detonation models, and the models produce up to 0.4 $M_\odot$ of ^{56}Ni in case of pure deflagrations and up to about 1 $M_\odot$ for DDTs. This falls into the range of observational expectations. However, both classes of models cannot explain the sub-luminous nor the super-luminous SNe Ia. Sub-luminous explosions typically have Ni-masses around 0.1 $M_\odot$, about a factor of two below the minimum value obtained from Chandrasekhar mass models. In contrast, in super-luminous explosions the Ni-mass may even exceed M_{Chan} (e.g., SN 2009dc[21]).

Nonetheless, synthetic lightcurves derived from models of these classes fit the observations in the B and V bands around maximum luminosity reasonably well[5,22]. A much harder constraint on the explosion model is posed by spectral observations, since spectra are particularly sensitive to the chemical composition of the ejecta. In general terms, DDT models are able to fit both lightcurves and spectra of normal SNe Ia quite well, and even their observed width-luminosity relation[22], whereas pure-deflagration models can only fit the low-luminosity end of the normal SNe Ia[23].

Recently, alternative progenitor channels have received renewed attention, mainly because the observed SNe Ia rate seems to be in conflict with population-synthesis models predicting a M_{Chan} (C+O) white dwarf formation rate about a factor of 10 too low[15,16]. In contrast, the rate of double-degenerate mergers as well as of He accreting sub-M_{Chan} C+O white dwarfs and merging pairs of C+O and He white dwarfs can account for the rate more easily.

As far as sub-M_{Chan} explosions are concerned, recent multi-dimensional simulations show that once the He shell detonates, the initiation of a core detonation is a very robust phenomenon for C+O core masses above, say, 0.8 $M_\odot$, and is relatively independent of the He shell ignition geometry[17,18]. As in previous 1-D models, however, the nucleosynthesis products from the detonation of a massive He shell above a C+O core is inconsistent with the observations[25]. Also models with larger core masses (around 1 $M_\odot$) have been studied. They have the advantage that a thin He shell may trigger a detonation[24]. The He shell detonation then does not produce significant amounts of ^{56}Ni but other iron-group isotopes[18]. In particular, absorption by Cr and Ti leads to strong flux redistribution towards the red parts of the spectrum making the models inconsistent with observations. This problem may be avoided if the He shell is polluted by carbon[25], for instance by instabilities in the

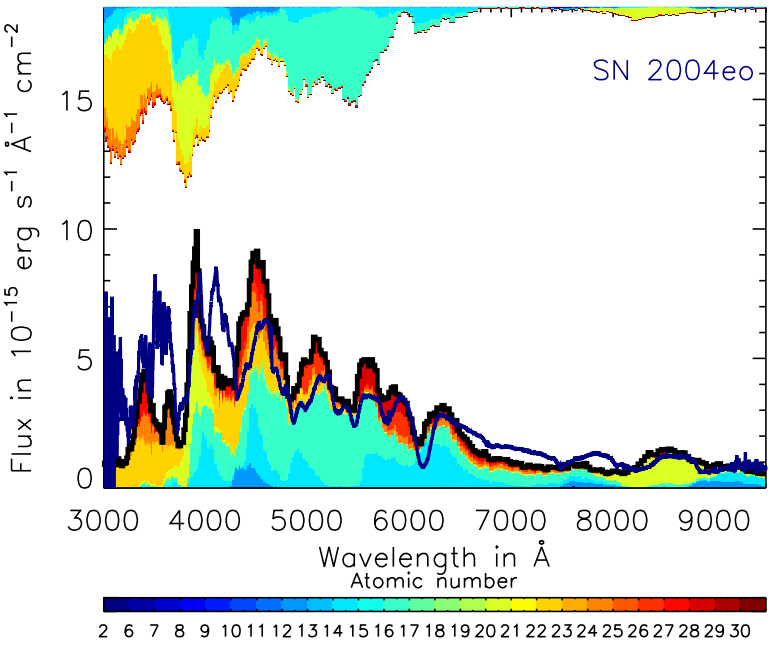


Figure 3. Angle-averaged (thick black line) spectra at three days before B-band maximum for a sub-M$_{Chan}$ model with M$_{core}$ = 1.025 M$_\odot$ and M$_{shell}$ = 0.055 M$_\odot$ and carbon enriched to 33%. For comparison the blue line shows the observed spectrum of SN 2004eo at the corresponding epoch. The colour coding indicates the element(s) responsible for both bound-bound emission and absorption of quanta in the Monte Carlo simulation. The region below the synthetic spectrum is colour coded to indicate the fraction of escaping quanta in each wavelength bin which last interacted with a particular element (the associated atomic numbers are illustrated in the colour bar). Similarly, the coloured regions above the spectra indicate which elements were last responsible for removing quanta from the wavelength bin (either by absorption or scattering/fluorescence)[25].

accretion or by He burning prior to the onset of the detonation. A carbon enrichment reduces the mass number of α elements that are synthesized, and with a $\sim$33% admixture, the spectrum agrees very well with the observations (see Fig. 3).

In contrast to sub-M$_{Chan}$ models, which, in principle, can explain the full range of SN Ia (leaving aside the super-luminous ones for which no convincing model exists) violent mergers of two C+O white dwarfs of roughly equal mass can explain the class of sub-luminous events only for masses around 1 M$_\odot$ (or less)[19,20]. More massive white dwarf mergers are likely to give rise to more luminous explosions, but they are rare and may not be able to account for the bulk of the SNe Ia.

5 Conclusion

Type Ia supernovae remain an important tool for observational cosmology. The measurement of cosmic distances based on these objects holds promise to constrain dark energy models – in particular, if the data is analyzed in a model-independent way.

Nonetheless, the theoretical understanding of SNe Ia is still an open question. Although fundamental scenarios (thermonuclear explosions of C+O WDs in binary systems)

seem compelling, a level of realism is required from the models that challenges astrophysical theory and numerical modeling. Besides the complex explosion mechanism, the lack of observations of the progenitor systems is a serious obstacle to astrophysical modeling. Numerous scenarios have been proposed over the last decades. To clarify which of these contribute to the observed SN Ia sample we follow the explosion, nucleosynthesis, and radiative transfer in a pipeline of simulations. This allows for a direct comparison of predictions with observations.

Apart from pure detonations of Chandrasekhar mass WDs it is difficult to rule out any of the explosion scenarios discussed here (pure turbulent deflagrations, delayed detonations, sub-Chandrasekhar mass models, and violent mergers of two WDs) based on the predicted observables, although none of them matches the observations perfectly. This could imply that all the scenarios are indeed realized in Nature. However, it is also conceivable that the observables around peak brightness are rather insensitive to details of the explosion mechanism as they originate from the outer layers of the ejecta. To break this degeneracy, the comparison between models and observations has to be extended. In particular nebular spectra and spectropolarimetry data seem promising to provide complementary information.

Acknowledgements

The calculations were carried out at the Jülich Supercomputing Centre and at the Rechenzentrum Garching (RZG).

References

1. A. Goobar, and B. Leibundgut, *Supernova cosmology: legacy and future*, arXiv:1102.1431, Annual Review of Nuclear and Particle Science (in print), 2011.
2. W. Hillebrandt, and J. C. Niemeyer, *Type Ia Supernova Explosion Models*, ARA&A **38**, 191–230, 2000.
3. J. C. Niemeyer, and W. Hillebrandt, *Turbulent Nuclear Flames in Type Ia Supernovae*, Astrophys. J. **452**, 769, 1995.
4. W. Schmidt, J. C. Niemeyer, and W. Hillebrandt, *A localised subgrid scale model for fluid dynamical simulations in astrophysics. I. Theory and numerical tests*, Astron. Astrophys. **450**, 265–281, 2006.
5. F. K. Röpke, W. Hillebrandt, W. Schmidt, J. C. Niemeyer, S. I. Blinnikov, and P. A. Mazzali, *A Three-Dimensional Deflagration Model for Type Ia Supernovae Compared with Observations*, Astrophys. J. **668**, 1132–1139, 2007.
6. M. Kromer, and S. A. Sim, *Time-dependent three-dimensional spectrum synthesis for Type Ia supernovae*, MNRAS **398**, 1809–1826, 2009.
7. P. A. Mazzali, D. N. Sauer, A. Pastorello, S. Benetti, and W. Hillebrandt, *Abundance stratification in Type Ia supernovae - II. The rapidly declining, spectroscopically normal SN2004eo*, MNRAS **386**, 1897–1906, 2008.
8. A. M. Khokhlov, *Delayed detonation model for type Ia supernovae*, Astron. Astrophys. **245**, 114–128, 1991.
9. F. K. Röpke, and J. C. Niemeyer, *Delayed detonations in full-star models of type Ia supernova explosions*, Astron. Astrophys. **464**, 683–686, 2007.

10. P. A. Mazzali, F. K. Röpke, S. Benetti, and W. Hillebrandt, *A Common Explosion Mechanism for Type Ia Supernovae*, Science **315**, 825, 2007.
11. F. K. Röpke, *Flame-driven Deflagration-to-Detonation Transitions in Type Ia Supernovae?* Astrophys. J. **668**, 1103–1108, 2007.
12. S. E. Woosley, *Type Ia Supernovae: Burning and Detonation in the Distributed Regime*, Astrophys. J. **668**, 1109–1117, 2007.
13. S. E. Woosley, and T. A. Weaver, *Sub-Chandrasekhar mass models for Type I supernovae*, Astrophys. J. **423**, 371–379, 1994.
14. E. Livne, W. D. Arnett, *Explosions of Sub–Chandrasekhar Mass White Dwarfs in Two Dimensions*, Astrophys. J. **452**, 62, 1995.
15. A. J. Ruiter, K. Belczynski, and C. Fryer, *Rates and Delay Times of Type Ia Supernovae*, Astrophys. J. **699**, 2026–2036, 2009.
16. A. J. Ruiter, K. Belczynski, S. A. Sim, W. Hillebrandt, C. Fryer, M. Fink, M. Kromer, *Delay times and rates for Type Ia supernovae and thermonuclear explosions from double-detonation sub-Chandrasekhar mass models*, arXiv:1011.1407, 2011.
17. M. Fink, W. Hillebrandt, and F. K. Röpke, *Double-detonation supernovae of sub-Chandrasekhar mass white dwarfs*, Astron. Astrophys. **476**, 1133–1143, 2007.
18. M. Fink, F. K. Röpke, W. Hillebrandt, I. R. Seitenzahl, S. A. Sim, and M. Kromer, *Double-detonation sub-Chandrasekhar supernovae: can minimum helium shell masses detonate the core?* Astron. Astrophys. **514**, id.A53, 2010.
19. R. Pakmor, M. Kromer, F. K. Röpke, S. A. Sim, A. J. Ruiter, and W. Hillebrandt, *sub-luminous type Ia supernovae from the mergers of equal-mass white dwarfs with mass 0.9* M_{solar}, Nature **463**, 61–64, 2010.
20. R. Pakmor, S. Hachinger, F. K. Röpke, and W. Hillebrandt, *Violent mergers of nearly equal-mass white dwarf as progenitors of subluminous Type Ia supernovae*, Astron. Astrophys. **528**, id.A117, 2011.
21. S. Taubenberger, et al., *High luminosity, slow ejecta and persistent carbon lines: SN 2009dc challenges thermonuclear explosion scenarios*, MNRAS **412**, 2735–2762, 2011.
22. D. Kasen, F. K. Röpke, and S. E. Woosley, *The diversity of type Ia supernovae from broken symmetries*, Nature **460**, 869–872, 2009.
23. F. K. Röpke, et al., *Modeling Type Ia supernova explosions*, PrPNP **66**, 309-318, 2011.
24. L. Bildsten, K. J. Shen, N. N. Weinberg, and G. Nelemans, *Faint Thermonuclear Supernovae from AM Canum Venaticorum Binaries*, Astrophys. J. **662**, L95–L98, 2007.
25. M. Kromer, S. A. Sim, M. Fink, F. K. Röpke, I. R. Seitenzahl, and W. Hillebrandt *Double-detonation Sub-Chandrasekhar Supernovae: Synthetic Observables for Minimum Helium Shell Mass Models*, Astrophys. J. **719**, 1067–1082, 2010.

High Resolution Simulations of Planet Formation Processes: Gravoturbulent Planetesimal Formation on JUGENE

Hubert Klahr[1] **and Anders Johansen**[2]

[1] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
E-mail: klahr@mpia.de

[2] Lund Observatory, Box 43, 221 00 Lund, Sweden
E-mail: anders@astro.lu.se

We study a new scenario for the rapid formation of kilometer sized planetary building bricks, called planetesimals, based on the combined effects of magneto hydrodynamical turbulence and gravity. Meter-sized objects get concentrated by turbulence; these over-densities start to get enhanced even further due to a clumping instability and finally densities are large enough that heaps of boulders fit into their combined Roche Lobe. On JUGENE we performed the highest resolution simulations to date of dust dynamics and planetesimal formation in turbulence generated by the magnetorotational instability. We present a new domain decomposition algorithm for particle-mesh schemes. Particles are spread evenly among the processors and the local gas velocity field and assigned drag forces are exchanged between a domain-decomposed mesh and discrete blocks of particles. We obtain good load balancing on up to 4096 cores even in simulations where particles sediment to the mid-plane and concentrate in pressure bumps.

1 Introduction

Planetesimals are several-kilometer-sized objects in the solar nebula, which are believed to be the building bricks for planets. Whereas there is no observational evidence for their existence in circumstellar disks, it is generally accepted that asteroids and comets are the last survivors from this class of objects. The behaviour of planetesimals is quite distinct from the behaviour of smaller, e.g. meter-sized objects, as their dynamical behaviour is determined by the gas motion. The first defining property of planetesimals is that they move on Keplerian orbits and do not drift radially due to gas drag. The second property is that they can accrete and accumulate smaller objects by gravity, e.g. the gravitational pull at the surface of a planetesimal is stronger than the force exerted by the head wind. While we have a good understanding how meter-sized objects grow via collisions and sticking, e.g. coagulation, and how planetesimals continue to form larger objects up to planetary cores, it is still a mystery how planetesimals form from sub-meter-sized boulders[1–3].

As gravity is the binding force for planetesimals it seems logical to invoke a formation scenario around a gravitational collapse of the dust phase in the solar nebula[4]. The modern picture of planetesimal formation[5] invokes gas turbulence to concentrate cm to dm sized particles until a gas-solid "streaming instability" is triggered, thus eventually a critical density of particles is reached and the heaps collaps under their own weight and form many kilometer-sized planetesimals right away.

In a recent paper[6] we presented high-resolution and long-time-integration simulations of planetesimal formation in turbulence caused by the magnetorotational instability (MRI) obtained on JUGENE. We find that the large-scale geostrophic pressure bumps that are

responsible for particle concentration are sustained when going from moderate (256^3) to high (512^3) resolution. Particle concentration in these pressure bumps is also relatively independent of resolution. We also study the initial burst of planetesimal formation at 512^3 resolution.

Finding the Initial Mass Function of planetesimals forming from the gravitationally bound clumps will ultimately require an improved algorithm for the dynamics and interaction of bound clumps as well as the inclusion of particle shattering and coagulation during the gravitational contraction.

The paper is organised as follows. In Sec. 2 we describe the dynamical equations for gas and particles. Sec. 3 contains descriptions of a number of improvements made to the Pencil Code in order to be able to perform particle-mesh simulations at up to at least 4096 cores. Simulations including particle self-gravity are presented in Sec. 4 (512^3 resolution). We summarise the paper and discuss the implications of our results in Sec. 5.

2 Dynamical Equations

We perform simulations solving the standard shearing box MHD/drag force/self-gravity equations for gas defined on a fixed grid and solid particles evolved as numerical superparticles. We use the Pencil Code, a sixth order spatial and third order temporal symmetric finite difference code[a].

We model the dynamics of a protoplanetary disc in the shearing box approximation. The coordinate frame rotates at the Keplerian frequency Ω at an arbitrary distance r_0 from the central star. The axes are oriented such that the x points radially away from the central gravity source, y points along the Keplerian flow, while z points vertically out of the plane.

2.1 Gas Velocity

The equation of motion for the gas velocity $\boldsymbol{u}$ relative to the Keplerian flow is

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} + u_y^{(0)}\frac{\partial \boldsymbol{u}}{\partial y} = 2\Omega u_y \boldsymbol{e}_x - \frac{1}{2}\Omega u_x \boldsymbol{e}_y + 2\Omega\Delta v \boldsymbol{e}_x + \frac{1}{\rho}\boldsymbol{J}\times(\boldsymbol{B}+B_0\hat{\boldsymbol{z}}) - \frac{1}{\rho}\boldsymbol{\nabla}P - \frac{\rho_{\rm p}/\rho_{\rm g}}{\tau_{\rm f}}(\boldsymbol{u}-\overline{\boldsymbol{v}}) + \boldsymbol{f}_\nu(\boldsymbol{u})\,. \quad (1)$$

The left hand side includes advection both by the velocity field $\boldsymbol{u}$ itself and by the linearised Keplerian flow $u_y^{(0)} = -(3/2)\Omega x$. The first two terms on the right hand side represent the Coriolis force in the x- and y-directions, modified in the y-component by the radial advection of the Keplerian flow, $\dot{u}_y = -u_x \partial u_y^{(0)}/\partial x$. The third term mimics a global radial pressure gradient which reduces the orbital speed of the gas by the positive amount Δv. The fourth and fifth terms in Eq. 1 are the Lorentz and pressure gradient forces. The current density is calculated from Ampère's law $\mu_0 \boldsymbol{J} = \boldsymbol{\nabla}\times\boldsymbol{B}$. The Lorentz force is modified to take into account a mean vertical field component of strength B_0. The sixth term is a drag force term is described in Sect. 2.4.

The high-order numerical scheme of the Pencil Code has very little numerical dissipation from time-stepping the advection term[7], so we add explicit viscosity through the

[a]See `http://code.google.com/p/pencil-code/`.

term $\boldsymbol{f}_\nu(\boldsymbol{u})$ in Eq. 1. We use sixth order hyperviscosity with a constant dynamic viscosity $\mu_3 = \nu_3\rho$,

$$\boldsymbol{f}_\nu = \frac{\mu_3}{\rho_{\rm g}}\nabla^6\boldsymbol{u}\,. \tag{2}$$

This form of the viscosity conserves momentum. The ∇^6 operator is defined as $\nabla^6 = \partial^6/\partial x^6 + \partial^6/\partial y^6 + \partial^6/\partial z^6$. It was shown by Ref. 8 that hyperviscosity simulations show zonal flows and pressure bumps very similar to simulations using Navier-Stokes viscosity.

2.2 Gas Density

The continuity equation for the gas density ρ is

$$\frac{\partial\rho}{\partial t} + (\boldsymbol{u}\cdot\boldsymbol{\nabla})\rho + u_y^{(0)}\frac{\partial\rho}{\partial y} = -\rho\boldsymbol{\nabla}\cdot\boldsymbol{u} + f_D(\rho)\,. \tag{3}$$

The diffusion term is defined as

$$f_{\rm D} = D_3\nabla^6\rho\,, \tag{4}$$

where D_3 is the hyperdiffusion coefficient necessary to suppress Nyquist scale wiggles arising in regions where the spatial density variation is high. We adopt an isothermal equation of state with pressure $P = c_{\rm s}^2\rho$ and (constant) sound speed $c_{\rm s}$.

2.3 Induction Equation

The induction equation for the magnetic vector potential $\boldsymbol{A}$ is[9]

$$\frac{\partial\boldsymbol{A}}{\partial t} + u_y^{(0)}\frac{\partial\boldsymbol{A}}{\partial y} = \boldsymbol{u}\times(\boldsymbol{B} + B_0\hat{\boldsymbol{z}}) + \frac{3}{2}\Omega A_y\hat{\boldsymbol{x}} + \boldsymbol{f}_\eta(\boldsymbol{A})\,. \tag{5}$$

The resistivity term is

$$\boldsymbol{f}_\eta = \eta_3\nabla^6\boldsymbol{A}\,, \tag{6}$$

where η_3 is the hyperresistivity coefficient. The magnetic field is calculated from $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}$.

2.4 Particles and Drag Force Scheme

The dust component is treated as a number of individual superparticles, the position $\boldsymbol{x}$ and velocity $\boldsymbol{v}$ of each evolved according to

$$\frac{{\rm d}\boldsymbol{x}}{{\rm d}t} = \boldsymbol{v} - \frac{3}{2}\Omega x\boldsymbol{e}_y\,, \tag{7}$$

$$\frac{{\rm d}\boldsymbol{v}}{{\rm d}t} = 2\Omega v_y\boldsymbol{e}_x - \frac{1}{2}\Omega v_x\boldsymbol{e}_y + \boldsymbol{\nabla}\Phi - \frac{1}{\tau_{\rm f}}(\boldsymbol{v} - \overline{\boldsymbol{u}})\,. \tag{8}$$

The particles feel no pressure or Lorentz forces, but are subjected to the gravitational potential Φ of their combined mass. Particle collisions are taken into account as well (see Sect. 2.6 below). Two-way drag forces between gas defined on a fixed grid and Lagrangian

particles are calculated through a particle-mesh method[10]. First the gas velocity field is interpolated to the position of a particle, using second order spline interpolation. The drag force on the particle is then trivial to calculate. To ensure momentum conservation we then take the drag force and add it with the opposite sign among the 27 nearest grid points, using the Triangular Shaped Cloud scheme to ensure momentum conservation in the assignment[11].

2.5 Self-Gravity

The gravitational attraction between the particles is calculated by first assigning the particle mass density $\rho_{\rm p}$ on the grid, using the Triangular Shaped Cloud scheme described above. Then the gravitational potential Φ at the grid points is found by inverting the Poisson equation

$$\nabla^2\Phi = 4\pi G\rho_{\rm p} \tag{9}$$

using a Fast Fourier Transform (FFT) method (see online supplement of our nature publication[5]). Finally the self-potential of the particles is interpolated to the positions of the particles and the acceleration added to the particle equation of motion (Eq. 8).

2.6 Collisions

Particle collisions become important inside dense particle clumps. Therefore, we have implemented a superparticle collision algorithm in the Pencil Code. The algorithm gives each particle a chance to interact with all other particles in the same grid cell. The characteristic time-scale $\tau_{\rm coll}^{(ij)} = 1/(n_j\sigma_{ij}\delta v_{ij})$ for a representative particle in the swarm of superparticle i to collide with any particle from the swarm represented by superparticle j is calculated by considering the number density n_j represented by each superparticle, the collisional cross section σ_{ij} of two swarm particles, and the relative speed δv_{ij} of the two superparticles. For each possible collision a random number P is chosen. If P is smaller than $\delta t/\tau_{\rm coll}$, where δt is the time-step set by magnetohydrodynamics, then the two particles collide. The collision outcome is determined as if the two superparticles were actual particles with radii large enough to touch each other. By solving for momentum conservation and energy conservation, with the possibility for inelastic collisions to dissipate kinetic energy to heat and deformation, the two colliding particles acquire their new velocity vectors instantaneously. All simulations include collisions with a coefficient of restitution of $\epsilon = 0.3$[12], meaning that each collision leads to the dissipation of approximately 90% of the relative kinetic energy to deformation and heating of the colliding boulders.

3 Computing Resources and Code Optimisation

The use of the Pencil Code on several thousand cores like on the "JUGENE" Blue Gene/P system at the Jülich Supercomputing Centre required both trivial and more fundamental changes to the code. We describe these technical improvements in detail in the following.

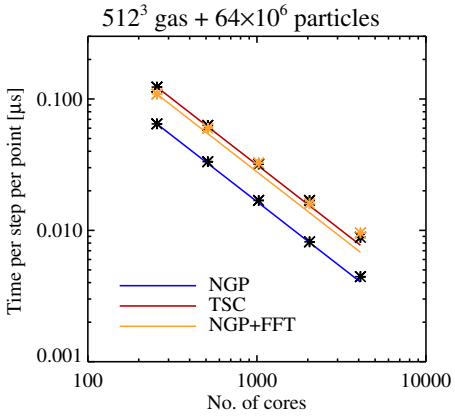


Figure 1. Scaling test for particle-mesh problem with 512^3 grid cells and 64×10^6 particles. The particles are distributed evenly over the grid, so that each core has the same number of particles. The inverse code speed is normalised by the number of time-steps and by either the total number of grid points and particles (top panel) or by the number of grid points and particles per core (bottom panel).

3.1 Changes Made to the Pencil Code

3.1.1 Memory Optimisation

We had to remove several uses of *global* arrays, i.e. 2-D or 3-D arrays of linear size equal to the full grid[6]. This mostly affected certain special initial conditions and boundary conditions.

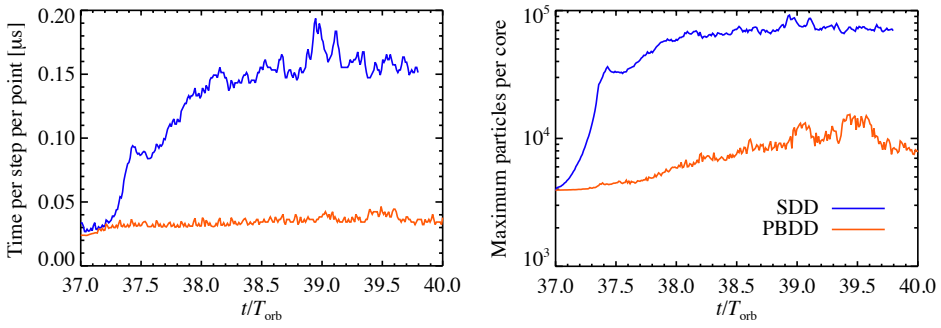


Figure 2. Code speed as a function of simulation time (left panel) and maximum particle number on any core (right panel) for 256^3 resolution on 2048 cores. Standard domain decomposition (SDD) quickly becomes unbalanced with particles and achieves only the speed of the particle-laden mid-plane cores. With the Particle Block Domain Decomposition (PBDD) scheme the speed stays close to its optimal value, and the particle number per core (bottom panel) does not rise much beyond 10^4.

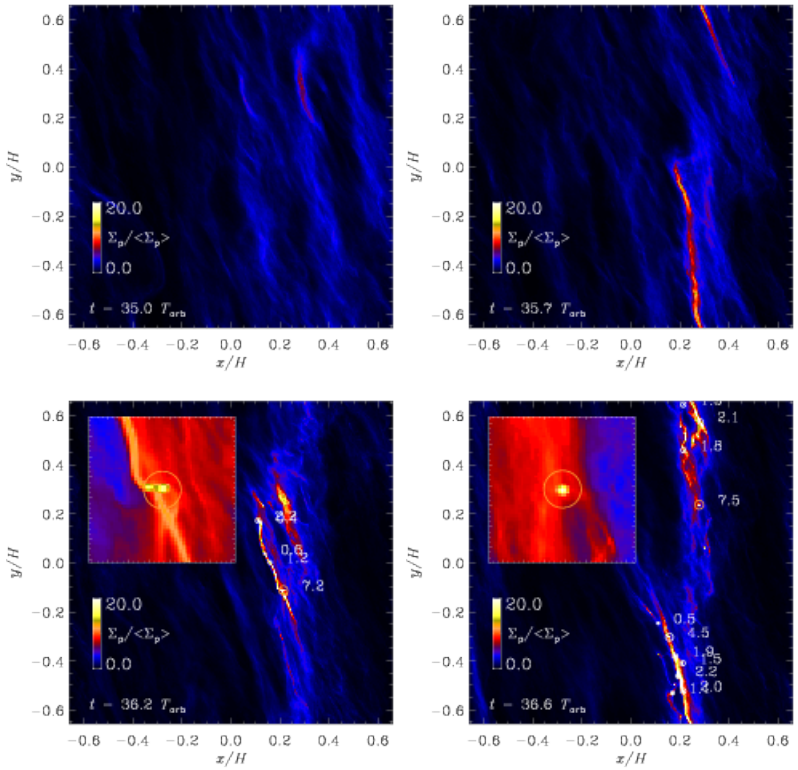


Figure 3. The particle column density as a function of time after self-gravity is turned on after $t = 35.0 T_{\rm orb}$ in the high-resolution simulation (run H with 512^3 grid cells and 64×10^6 particles). Two clumps condense out within each other's Hill spheres and quickly merge. At the end of the simulation bound clumps have masses between 0.5 and $7.5\ M_{\rm Ceres}$.

3.1.2 Particle Migration

At the end of a sub-time-step each processor checks if any particles have left their spatial domains. Information about the number of migrating particles, and the processors that they will migrate into, is collected at each processor. The Pencil Code would then let all processors exchange migrating particles with all other processors. In practice particles would of course only migrate to neighbouring processors. However, at processor numbers of 512 or higher, the communication load associated with each processor telling all other processors how many particles it wanted to send (mostly zero) was so high that it dominated over both the MHD and the particle evolution equations. Since particles in practice only migrate to the neighbouring processors any way, we solved this problem by letting the processors only communicate the number of migrating particles to the immediate neighbours. Shear-periodic boundary conditions require a (simple) algorithm to determine the three neighbouring processors over the shearing boundary in the beginning of each sub-time-step.

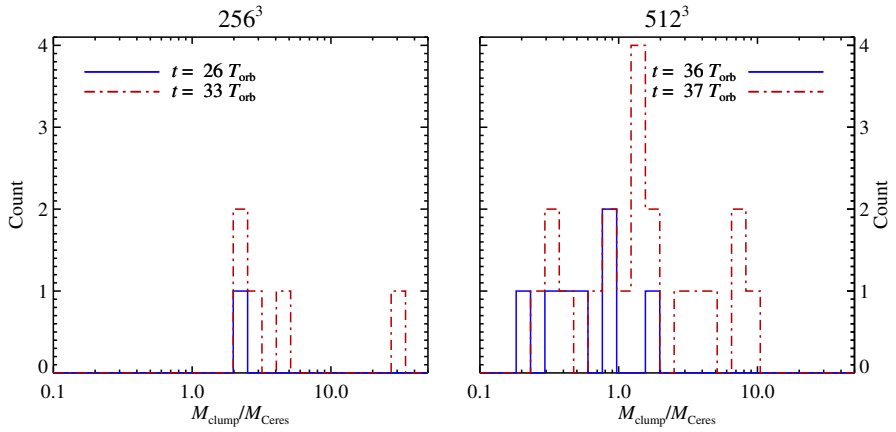


Figure 4. Histogram of clump masses after first production of bound clumps and at the end of the simulation. At moderate resolution (left panel) only a single clump condenses out initially, but seven orbits later there are five clumps, including the 30+ $M_{\rm Ceres}$ object formed by merging. At high resolution (right panel) the initial planetesimal burst leads to the formation of many sub-Ceres-mass clumps. The most massive clump is similar to what forms initially in the moderate-resolution run.

4 Self-Gravity – High Resolution

In this section we show that the inclusion of self-gravity at high resolution leads to rapid formation of bound clumps similar to what we observe at moderate resolution. Given the relatively high computational cost of self-gravity simulations we start the self-gravity at $t = 35T_{\rm orb}$ in the 512^3 simulation, three orbits after inserting the particles. The evolution of particle column density is shown in Fig. 3[b]. Due to the smaller grid size bound particle clumps appear visually smaller than in the 256^3 simulation. The increased compactness of the planetesimals can potentially decrease the probability for planetesimal collisions, which makes it important to do convergence tests.

The increased resolution of the self-gravity solver allows for a number of smaller planetesimals to condense out as the bands reach the local Roche density at smaller radial scales (panel 3). Two of the clumps are born within each other's Hill spheres. They merge shortly after into a single clump (panel 5). This clump has grown to $7.5\ M_{\rm Ceres}$ at the end of the simulation, which is the most massive clump in a population of clumps with masses between 0.5 and $7.5\ M_{\rm Ceres}$.

Fig. 4 shows histograms of the clump masses for moderate resolution (left panel) and high resolution (right panel).

5 Summary and Discussion

With the changes described in Sect. 3.1.1 and Sect. 3.1.2 the Pencil Code can be run with gas, particles and gravity efficiently at several thousand processors.

[b]Movies can be found at: `http://pc292.astro.lu.se/~anders/research.php`

Recently[6] we published the first 512^3 grid cell simulation of dust dynamics in turbulence driven by the magnetorotational instability. Perhaps the most important finding is that large-scale pressure bumps and zonal flows in the gas appear relatively independent of the resolution. The same is true for particle concentration in these pressure bumps. Resulting size spectra are essentially broader in high resolution cases, which is essential to match the observational constraints of asteroids in the solar system.

In future simulations strongly bound clusters of particles should be turned into single gravitating sink particles, in order to prevent planetesimals from having artificially large sizes. An additional N-body solver is also under development.

Acknowledgements

This project was made possible through a computing grant at the JUGENE Blue Gene/P supercomputer at Jülich Supercomputing Centre. HK has been supported in part by the Deutsche Forschungsgemeinschaft DFG through grants DFG Forschergruppe 759 "The Formation of Planets: The Critical First Growth Phase" and DFG-SPP 1385 "The First 10 Million Years of the Solar Systems - a Planetary Materials Approach". AJ was supported by the Netherlands Organization for Scientific Research (NWO) through Veni grant 639.041.922 and Vidi grant 639.042.607 during part of the project.

References

1. S. J. Weidenschilling, *Dust to planetesimals - settling and coagulation in the solar nebula.* Icarus **44**, 172–189, 1980.
2. S. J. Weidenschilling, *Evolution of grains in a turbulent solar nebula.* Icarus **60**, 553–567, 1984.
3. S. J. Weidenschilling, *Can gravitational instability form planetesimals?* Icarus **116**, 433–435, 1995.
4. P. Goldreich and W. R. Ward, *The formation of planetesimals.* Astrophysical Journal **183**, 1051–1062, 1973.
5. A. Johansen, J. Oishi, M. Mac Low, H. Klahr, Th. Henning, and A. Youdin, *Rapid planetesimal formation in turbulent circumstellar discs* nature **448**, 1022-1025, 2007.
6. A. Johansen, H. Klahr, and Th. Henning, *High-resolution simulations of planetesimal formation in turbulent protoplanetary discs* Astrophysical Journal **529**, A62 (16 p.), 2011.
7. A. Brandenburg, *Computational aspects of astrophysical MHD and turbulence* Advances in Nonlinear Dynamos , 269-+, 2003.
8. A. Johansen, A. Youdin, and H., Klahr, *Zonal Flows and Long-lived Axisymmetric Pressure Bumps in Magnetorotational Turbulence.* Astrophysical Journal **697**, 1269–1289, 2009.
9. A. Brandenburg, Å. Nordlund, R. F. Stein, and U. Torkelsson, *Dynamo-generated Turbulence and Large-Scale Magnetic Fields in a Keplerian Shear Flow.* Astrophysical Journal **446**, 741–754, 1995.
10. A. Youdin and A. Johansen, *Protoplanetary Disk Turbulence Driven by the Streaming Instability: Linear Evolution and Numerical Methods* Astrophysical Journal **662**, 613-626, 2007.

11. R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles.* in Computer Simulation Using Particles, New York: McGraw-Hill, 1981.
12. J. Blum and M., Muench, *Experimental investigations on aggregate-aggregate collisions in the early solar nebula.* Icarus **106**, 151–167, 1973.

High Performance Computing for Life Science Applications

Wolfgang Wenzel

Institute of Nanotechnology, Karlsruhe Institute of Technology
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany
E-mail: wolfgang.wenzel@kit.edu

Continuing efforts in method development and a steady increase of the available computational resources has led to an ever more prominent role of simulation and modeling in the life sciences. Methods that operate at various time and length scales are now routinely used to complement, rationalize and even predict experiments and have thus emerged as a valuable tool in to understand complex problems in the life sciences. Development of special-purpose computers and increasingly scalable methods has allowed inroads into the protein folding and structure prediction problems, while coupling classical and quantum techniques have elucidated intricate protein functions, such as energy harvesting and ion transport across membranes. However, progress is not confined to the atomistic or microscopic scales, as increasingly powerful methods help to understand biological phenomena at the length scales of organs or the body as a whole. Many of these simulation problems are characterized by a large complexity in the constituents and/or a multi-scale coupling of different time and length scales, requiring linkage of different methods to adequately model the underlying systems. Often such challenges can be addressed only with high performance computational resources, which in turn necessitate adaptation of the codes in order to achieve sufficient scalability on these systems. In this issue several contributions offer interesting insights into this ongoing trend:

In their study on the pH dependence of an important acid sensing channel, Pluhackova et. al. demonstrate how molecular dynamics simulations can help to understand membrane protein function. Membrane proteins constitute an important protein class that is difficult to elucidate experimentally in structure, let alone in function. In this work high performance computational resources were used to study the complex dynamics of systems with up to 300,000 atoms, comprising the ion channel, the membrane in which it is embedded and its physiological environment that needs to be presented in great detail in order to understand how the protein performs its intricate function, reacting to changes in the pH of its environment.

In a similar vein Faraldo-Gómez investigates one of the most fascinating energy generation mechanisms in cells, detailing how molecular motion, driven by an ion flux across a physiological membrane, is used to generate ATP, one of the generic energy carriers of eukaryotic cells. Simulations of this large protein complex, which accomplishes this feat with high efficiency, require a detailed understanding of transitions between many local equilibria, the characterization of the relative free energies and transition rates. Not surprisingly such simulations are best performed on high performance supercomputers that are able to represent such complex systems in atomistic detail for high-quality predictions.

Since experiments yielded presently only snapshots of relevant intermediate states of the underlying processes molecular simulations can be used, as reported in this study, to connect the states and propose new mechanisms.

High performance simulation is not confined to the molecular scale, but can be used to understand many physiologically important processes, such as the blood flow on the basis of increasingly sophisticated models. Fedosov et. al. report an interesting study on the flow properties of human blood in a mesoscopic model that can elucidate the effect of microscopic cell-cell interactions on macroscopic properties, such as blood viscosity and shear rates. Such studies can be used to model physiological properties of the cardiovascular system that is relevant for many diseases and pharmaceutical applications.

Moving to one of the most complex conceivable problems, Kunkel at al. report on the likely computational requirements that need to be met by emerging computational architectures in order to be able to model neuronal networks at the scale of the human brain. The European Union, as well as other countries, aims at the development of exascale computational architectures before the end of this decade. These architectures will provide unprecedented computational resources and enable simulations far beyond the scale that can presently be realized and challenge our creativity to address problems far beyond our present horizon. However, provisioning of these resources will call for new paradigms in high performance computing that challenge the way we think about programming today. It is therefore important, as reported in this study, to focus on possible bottlenecks facing interesting challenges, such as large-scale neuronal network simulation, and suggest avenues of these challenges can be overcome before the new hardware is ready.

The examples document realization of fruitful interdisciplinary efforts between application groups and high performance computing centres to enable investigations that are beyond the scope of the locally available computational resources. While naturally focusing on the computational aspects of this work it should not be forgotten that the computation application groups now form an important bridge to the experimental life science community that increasingly relies on the results of simulations to interpret and to guide the experiments in order to understand and influence biological function in increasingly complex problems.

pH Dependence of ASIC1a Dynamics on a Sub-Microsecond Time Scale

Kristyna Pluhackova and Rainer Böckmann

Computational Biology, Department of Biology, University of Erlangen-Nürnberg
91054 Erlangen, Germany
E-mail: {*kristyna.pluhackova, rainer.boeckmann*}*@biologie.uni-erlangen.de*

Acid sensing ion channels (ASICs), members of epithelial sodium channels (ENaC) family, are sodium-selective, ligand-gated ion channels present in most neurons. ASICs are emerging as non-classical excitotoxic therapeutic targets for combating stroke, their inhibition has antidepressant-like effects and helps to protect the brain against ischemic injury during stroke and axonal injury and demyelination in multiple sclerosis. ASICs open rapidly upon decrease in extracellular pH, allowing flux of sodium and partially calcium ions inside the cell for about 1-2seconds before they completely desensitize.

By performing molecular dynamics studies of a recently crystallized functional form of a chicken ASIC1a at low pH embedded in a phosphatidylcholine (POPC) membrane and surrounded by a physiological salt concentration it is shown that the protein can be inserted at two different but equally stable insertion depths in the membrane. We further address the filling mechanism of the upper chamber by sodium ions depending on the protonation states of the aminoacids involved as well as on the free sodium concentration and protein insertion depth. The dynamics of the ion channel pore radius and its dependance on the protonation state of key residues is studied.

1 Introduction

Acid sensing ion channels (ASICs) are neuronal nonvoltage-gated, sodium-selective and ligand-gated ion channels that are activated by a drop in extracellular pH. ASICs belong to the epithelial sodium channels (ENaCs)/Degenerins family of ion channels whose members – although sharing a similar structure – are activated in different ways, thereby enabling the cell to respond similarly to various stimuli. Upon activation, sodium cations may pass the channel into the cell. ASIC1a contributes to the excitatory postsynaptic current in many neurons[1] and as such plays an important role in the fear circuit[2]. Inhibition of ASIC1a has antidepressant-like effects, thus making ASIC1a a new possible target to combat depressions[3]. ASIC1a further acts as a pain receptor[4]. As activation of ASIC1a in the brain as a consequence of stroke leads to ischemic cell death, its inhibition ameliorates stroke-induced cell damage and death[5]. The upregulation of ASIC1a was moreover discovered in lesions from patients suffering from active multiple sclerosis[6] suggesting that its blockade has a potential to provide both neuro- and myelo-protective benefits in multiple sclerosis.

ASIC1a opens rapidly upon decrease in extracellular pH, with a pH_{50} (pH of half-maximal activation) of 6.5 and a Hill coefficient of approximately 3[7]. Subsequently, the ion channel desensitizes completely within 1-2 seconds[7] at low pH. Recently, a functional form of chicken ASIC1a, being the most abundant ASIC family member in the mammalian central nervous system and possessing the highest proton affinity, was crystallized at low pH[3], most probably representing the desensitized state of the ion channel.

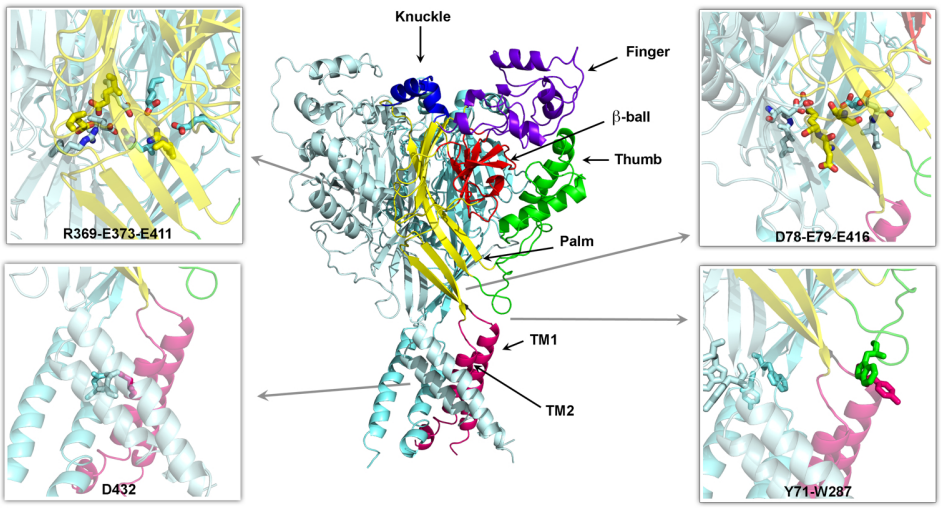


Figure 1. Structure of chicken ASIC1. One monomer unit coloured according to the main structural domain scheme as suggested by Gründer and Chen[1]. Panel *left up* shows the residues at the ceiling of the upper chamber, Arg^{369}, Glu^{373} and Glu^{411}. The nitrogens of the Arg sidechains are coloured in blue, oxygens of glutamic acid sidechains in red. The panel in the lower left corner shows the triade of Asp^{432} forming the neck of the pore, the panel in the upper right corner shows residues at the bottom of the upper chamber, Asp^{78}, Glu^{79} and Glu^{416} (oxygen atoms in red, nitrogen atoms in blue). The panel in the lower right corner shows the ball-and-socket joint formed by Tyr^{71} and Trp^{287}.

The ASIC1a is a homotrimeric assembly of ASIC subunits 1a. Each subunit possesses a large extracellular domain (EC), a symmetric transmembrane domain (TM) formed by two transmembrane helices (TM1 and TM2) and short intracellular N and C termini (IC), which, however, were not possible to get resolved in the crystal structures probably due to their large mobility. The ion channel has the shape of an hourglass with two transmembrane (TM) helices contributed by every subunit, symmetrically ordered around the three-fold axis. The desensitization gate was found as a sterical gate formed by the three TM2 α-helices, located in the centre of the TM domain. The extracellular side of the gate is formed by the three Asp^{432} residues (see Fig. 1), which according to a recent experimental study[8] have to be protonated at low pH to allow for the channel opening and sodium ion passage. At high pH, on the contrary, at least two aspartic acids of this triade are supposed to be charged and to bind an calcium ion thus blocking the pore[9].

The large negative charge in the upper chamber at high pH suggests that this part of the protein serves as a sodium ion reservoir. Two triades of highly conserved glutamic acids (Glu^{373} and Glu^{411}) among ASICs are found in the upper part of the chamber, in the lower part there is an acidic ring formed out of residues Glu^{79} and Glu^{416} contributed from each chain. The negative electrostatic potential inside the chamber is slightly shielded by the positively charged Arg^{369}. The upper chamber is separated from the rest of the pore by a hydrophobic lock formed by the residues Thr^{75} and Leu^{77}. The size of Leu^{77} proved as indespensable as an exchange by glutamine in this position led to a proton-insensitive protein[8].

2 Methods

For all molecular dynamics simulations the program GROMACS (Groningen MAchine for Chemical Simulations) was applied[10]. The advantage of GROMACS over other programs is its extremely high performance and its scalability to large numbers of processors. For the latest GROMACS-4.5-beta4 version a 2D decomposition support for PME was implemented with a significant improvement of load balancing for large systems. After modeling of parts of the termini to the 3IJ4 crystal structure[3] we have embedded the protein to a POPC lipid bilayer containing 500 lipids. The Berger force field[11] was chosen for the lipids, the channel was described by the OPLS-AA foce field[12]. We have simulated our system in the NPT ensemble, with the pressure coupled to a pressure bath at 1 bar by semi-isotropic Parrinello-Rahman pressure coupling and the temperature at 310 K by using the v-rescale thermostat. A timestep of 2 fs was applied, the Coulombic forces were calculated using the particle mesh Ewald method with a short-range cutoff of 1.5 nm, Lennard-Jones interactions were treated with a switch function between 1.1 nm and 1.3 nm. A long range dispersion correction was applied both for the energy and the pressure.

3 Code Performance

The code performance was tested for a typical system studied. It consists of an ASIC1a homotrimer embedded in a phospholipid bilayer (500 POPC molecules), surrounded by water molecules and ions at physiological concentration. The total system size amounts

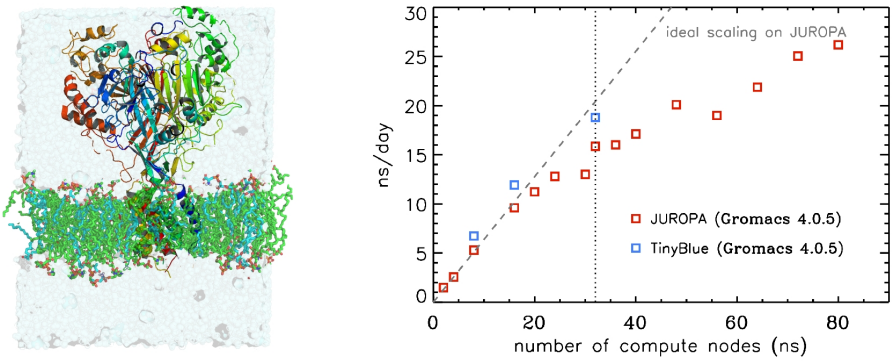


Figure 2. Scaling behaviour of GROMACS 4.0.5. The scaling was tested with the ASIC1a protein embedded in a phospholipid bilayer surrounded by water molecules (*left*). The total system size is approx. 300,000 atoms. The scaling was tested on the JUROPA system and on the TinyBlue system of the computing centre RRZE in Erlangen (*right*).

to approx. 300,000 atoms. Fig. 2 shows benchmarks that were computed both on the JUROPA system (2,208 nodes, each with two Xeon 5570 Nehalem chips at 2.93 GHz) using the GROMACS 4.0.5 module and on the local TinyBlue system at the RRZE in Erlangen

(84 nodes, each with two Xeon 5550 Nehalem chips at 2.66 GHz) using GROMACS 4.0.5. Different to JUROPA, hyperthreading was used on the TinyBlue cluster. A good scaling was obtained for up to 32 compute nodes (i.e. 256 compute cores), enabling for simulation times of more than 15 ns per day.

4 Results

4.0.3 Membrane Positioning

ASIC1a can be anchored in the POPC membrane at two different insertion depths which both proved to be stable over simulation times of a few hundred nanoseconds. Both starting configurations showed very similar interaction energies with the membrane. In the position suggested originally by Jasti *et al.*[13] the protein is anchored in the POPC membrane by Trp^{46} and Phe^{69}, while at increased insertion the Tyr^{287} and Phe^{50} are positioned at the membrane interface. The advantage of the latter positioning is the closed neck part of the protein, thus hindering a free flux of ions from/into the solution.

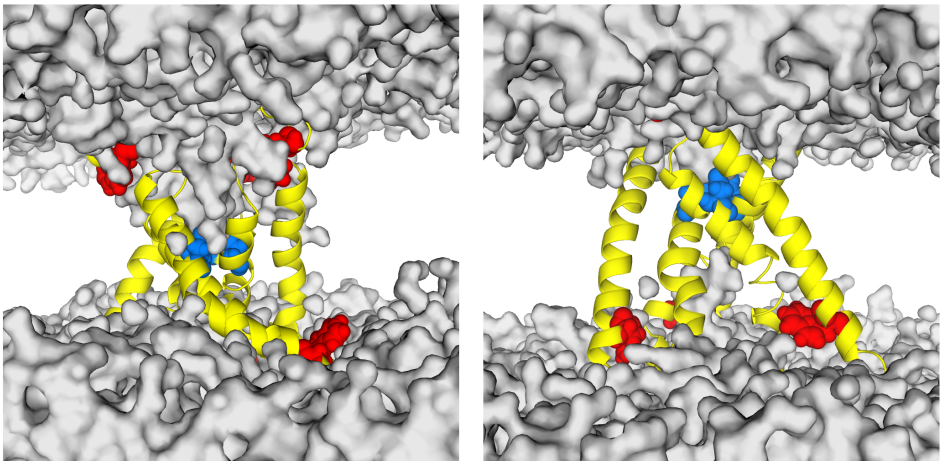


Figure 3. Depth of insertion. The *left* panel shows the insertion of ASIC1a in the POPC membrane at normal depth, the *right* panel a deeper variant. Trp residues 46 and 287 are coloured in red, water in grey, Asp^{432} as a lock of the pore in blue. POPC molecules are not displayed for clarity. Water can penetrate on both sides of the pore up to the central pore lock formed by Asp^{432}.

As can be seen in Fig. 4 the interaction energies of ASIC1a with the POPC lipid bilayer for normal and increased insertions are very similar. At pH 6.5 the interaction energy decreases for about 100 ns for both configurations and stays approximately constant for longer simulation times. At pH 8.0, the ion channel interacts more strongly with the membrane due to an increased number of charged residues that strongly interact with the lipid headgroups. At increased insertion depth the protein exhibits larger attraction to the membrane until 200 ns when the extracellular (EC) head of the protein lifts up from the membrane, thus disrupting some of the attractive interactions between charged amino acids in

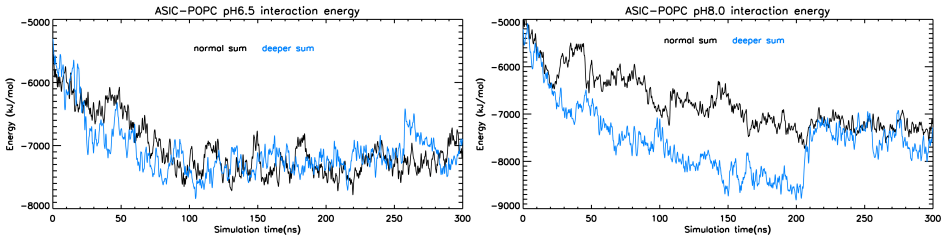


Figure 4. Interaction energies of ASIC1a and the membrane for normal and increased insertion depths for both investigated pH values (6.5 and 8.0).

the EC and the membrane but decreasing the bending tension between TM and EC parts of the protein. The similarity in interaction energies and the stability of the relative orientations of ASIC1a relative to the membrane for the two different insertion depths suggests a possible transition mechanism between both orientations which is likely dependant also on the membrane composition.

Functionally, the opening of the hydrophobic lock of the upper chamber might be triggered by this transition and will be subject of future studies.

4.0.4 Upper Chamber Filling

If both triades of residues Glu^{79} and Glu^{416} forming the acidic ring and moreover Asp^{78} are charged, one of the Glu^{79} flips out to the solution (see Fig. 5 for the detailed mechanism) and pulls up to three sodium ions into the chamber within tens of nanoseconds thus filling the upper chamber with sodium ions. This occurance is clearly electrostatically driven as the first ion enters rapidly after Glu^{79} flips out while it takes a few tens of nanoseconds for the third ion to enter the chamber. The whole order of events is moreover depending on the concentration of free sodium ions in the solution. Because this procedure is considerably slowed down if the protein is inserted deeper in the membrane we hypothesize that the protein at high pH will be anchored by Trp^{46} and Phe^{69} residues.

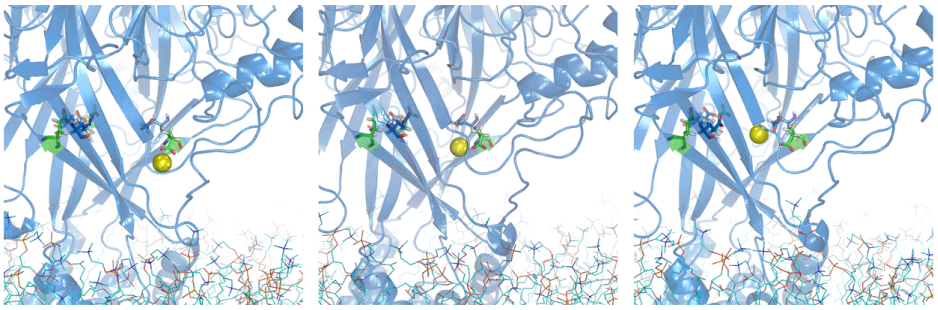


Figure 5. Glu^{79} (green) pulling in a sodium ion (yellow). In the *left* panel the first contact of a sodium ion with Glu^{79} is shown, in the *middle* panel the sodium ion is shared between Glu^{79} and Glu^{416} represented as blue sticks. In the *right* panel the sodium ion resides inside the upper chamber interacting with the anionic ring formed by both triades of residues Glu^{79} and Glu^{416}. Only one Glu^{79} flipped out of the pocket.

A decrease in extracellular pH modelled by partial protonation of the negatively charged amino acids in this chamber will decrease the electronegative potential in the chamber significantly, which will allow the sodium ions to move further on the way through the ion channel. After passing (the now partially protonated) acidic ring, the sodium ions have to cross the hydrophobic lock at the bottom of the chamber formed by the triades of residues Thr^{75} and Leu^{77}. A partial opening of this lock was observed in the simulations at decreased pH, but since no concentration gradient was applied so far no sodium ion was observed to pass, yet.

4.0.5 Ion Channel Dynamics

In order to understand the differences in pore dynamics between the two protonation states (at pH 6.5 and pH 8.0) and the influence of the insertion depth, we have calculated the pore radius by using the program HOLE[14]. As seen in Fig. 6, both protonation states and depth insertions share a similar pore shape, namely an inflexible, closed part between -40 Å and -45 Å corresponding to the pore gate of charged Asp^{432} and a flexible neck region between -15 Å and -25 Å. The closed pore lock of Asp^{432} is expected since in all simulations the whole triade was charged and as such it is supposed to block the pore. First upon protonation, Asp^{432} releases the cation, either calcium to the extracellular or sodium to the intracellular part. However, in both simulations of the protein at increased insertion depth, a broadening of the intracellular side of Asp^{432} is visible together with an increased flexibility of this part.

The upper chamber (between 0 Å and -15 Å) is broader for simulations at pH 8.0 which is due to repulsion of the negative amino acids inside. Thereby the accessible volume for sodium ions is significantly increased.

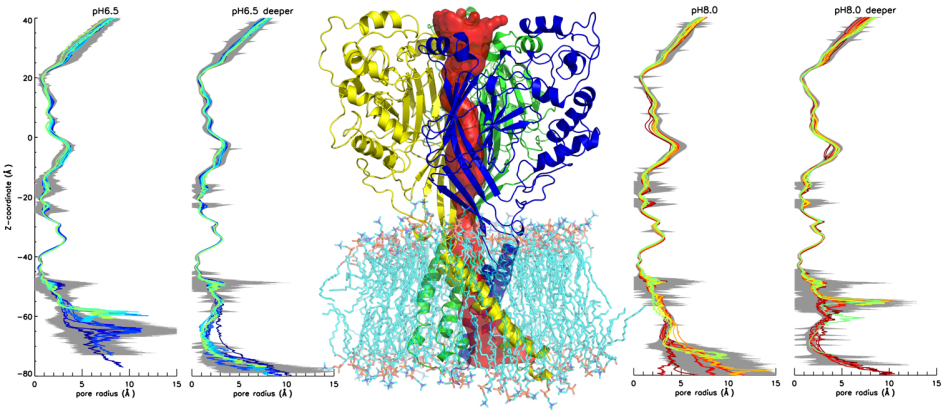


Figure 6. Pore radius. The central figure shows the ASIC1a at protonation state pH 6.5 and normal insertion depth after a 100 ns long simulation. Monomers are depicted in cartoon representation in yellow, blue and red, the POPC membrane (just closest lipids shown for clarity) are displayed using sticks and the pore is depicted as red surface. Pore radii averages (20 ns windows) of simulations are given for different protonation states corresponding to pH 6.5 and pH 8.0 and both insertion depths. The error bars are shown in grey (displayed always just for 1 example average curve) as a measure of the flexibility of the respective region.

5 Concluding Remarks

We studied the equilibrium properties of the ASIC1a ion channel and the first steps of its activation upon decrease in the extracellular pH. The simulation results showed that ASIC1a can be inserted into phospholipid bilayers at two different depths displaying similar interaction energies with the surrounding phospholipid bilayer. Subsequent studies are required to estimate the free energy barrier between these two configurations and to test for the functional implications. At pH 8.0 (inactive state of ASIC1a) and normal insertion depth the upper chamber of the protein fills within tens of nanoseconds with three sodium ions. This electrostatically driven action requires Glu^{79} to flip out to attract (together with Asp^{78}) sodium ions from the solution. The hydrophobic lock at the bottom of the upper chamber separates the chamber from the rest of the pore. It stayed closed in all simulations not allowing sodium ions to pass through without applying a concentration gradient on the system. The Asp^{432} triade, forming a neck at the transmembrane part of the protein binds up to two sodium ions in its charged form. At decreased pH (activation of ASIC1a) the channel is expected to open and allow passage of sodium ions into the cell. In simulations at low pH without applying a concentration gradient, however, the ions stay at the triade not passing through the pore. Ion concentration gradients across the membrane will be applied in subsequent studies to enable the full ion passage starting from the entrance to the ion chamber to their passage through the hydrophobic lock and the channel neck region.

Acknowledgements

We gratefully acknowledge support with computing time from the Jülich Supercomputing Centre.

References

1. S. Gründer and X. Chen *Structure, function, and pharmacology of acid-sensing ion channels (ASICs): focus on ASIC1a.* Int. J. Physiol. Pathophysiol. Pharmacol. **2**, 73-94, 2010.
2. M. W. Coryell, A. M. Wunsch, J. M. Haenfler, J. E. Allen, J. L.McBride, B. L. Davidson and J. A.Wemmie *Restoring Asid-sensing ion channel-1a in the amygdala of knock-out mice rescues fear memory but not unconditioned fear responses.* J. Neurosci. **28**, 13738-13741, 2008.
3. E. B. Gonzales, T. Kawate and E. Gouaux *Pore architecture and ion sites in acid-sensing ion channels and P2X receptors.* Nature **460**, 599-604, 2009.
4. S. P. Sutherland, C. J. Benson, J. P. Adelman and E.W. McCleskey *Acid-sensing ion channel 3 matches the acid-gated current in cardiac ischemia-sensing neurons.* Proc. Natl. Acad. Sci. U.S.A. **98**, 711-716, 2001.
5. Z. G. Xiong, X. M. Zhu, X. P. Chu, M. Minami, J. Hey, W. L. Wei, J. F. MacDonald, J. A. Wemmie, M. P. Price, M. J. Welsh and R. P. Simon *Neuroprotection in ischemia: Blocking calcium-permeable acid-sensing ion channels.* Cell **118**, 687-698, 2004.
6. S. Vergo, M. J. Craner, R. Etzensperger, K. Attfield, M. A. Friese, J. Newcombe, M. Esiri and L. Fugger *Acid sensing ion channel 1 is involved in both axonal injure and demyelination in multiple aclerosis and its animal model.* Brain **134**, 571-584, 2011.

7. R. Waldmann, G. Champigny, F. Bassilana, C. Heurteaux and M. Lazdunski *A proton-gated cation channel involved in acid-sensing.* Nature **386**, 173-177, 1997.
8. T. Li, Y. Yang and C. M. Canessa *Asp433 in the closing gate of ASIC1 determines stability of the open state without changing properties of the selectivity filter or Ca2+ block.* J. Gen. Physiol. **137**, 289–297, 2011.
9. M. Paukert, X. Chen, G. Polleichtner, H. Schindelin and S. Gründer *Candidate amino acids involved in H^+ gating of acid-sensing ion channel 1a.* JBC **283**, 572-581, 2008.
10. E. Lindahl, B. Hess and D. van der Spoel *GROMACS 3.0: a package for molecular simulation and trajectory analysis.* Journal of Molecular Modeling **7**, 306-317, 2001.
11. O. Berger, O. Edholm and F. Jähnig *Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature.* Biophys. J. **72**, 2002-2013, 1997.
12. G. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen *Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides.* J. Phys. Chem. B. **104**, 6474-6487, 2001.
13. J. Jasti, H. Furukawa, E. B. Gonzales and E. Gouaux *Structure of acid-sensing ion channel 1 at 1.9Å resolution and low pH.* Nature **489**, 316-323, 2007.
14. S. O. Smart, J. M. Goodfellow and B. A. Wallace *The pore dimensions of gramicidin A.* Biophys J. **65**, 2455-2460, 1993.

Memory Consumption of Neuronal Network Simulators at the Brain Scale

Susanne Kunkel[1,2], Moritz Helias[3], Tobias C. Potjans[3], Jochen M. Eppler[3], Hans Ekkehard Plesser[4], Markus Diesmann[2,3,5], and Abigail Morrison[1,2]

[1] Functional Neural Circuits Group, Faculty of Biology
Albert-Ludwig University of Freiburg, Germany
E-mail: {kunkel, morrison}@bcf.uni-freiburg.de

[2] Bernstein Center Freiburg, Albert-Ludwig University of Freiburg, Germany

[3] Institute of Neuroscience and Medicine (INM-6), Computational and Systems Neuroscience
Forschungszentrum Jülich, Germany
E-mail: {m.helias, t.c.potjans, j.eppler, m.diesmann}@fz-juelich.de

[4] Department of Mathematical Sciences and Technology
Norwegian University of Life Sciences, Aas, Norway
E-mail: hans.ekkehard.plesser@umb.no

[5] Medical Faculty, RWTH Aachen University, Germany

Simulations of neuronal networks involve the representation of numerous objects and frequent interactions between them. As this requires extensive software infrastructure, a neuronal network simulator optimized for clusters of several hundred processors can fail to run on supercomputers with of the order of $100,000$ nodes due to massive serial memory overhead of insufficiently parallelized data structures. To systematically investigate this issue we developed a mathematical model of the memory usage of a neuronal network simulator. The model can be instantiated for a particular software in order to analyze the memory utilization of the constituent components for different regimes of network size and number of processes. In its original formulation the model is based on the assumption of random network connectivity, but it can be adapted to account for more structured networks, such as systems of cortical columns, which are complexes of up to $100,000$ neurons with high intra-connectivity. Based on model predictions, we investigate to what extent exploiting the structure by gathering the neurons of a column on a subset of processors results in improved memory utilization.

1 Introduction

The human brain comprises about 10^{11} neurons that are sparsely and specifically connected through synapses, which enable the transmission of electrical pulses from one neuron to another, and hence the exchange of information within the neuronal network. Simulators which are tailored to investigate the dynamics of the communication between neurons in large-scale networks rather than the biochemical or biophysical processes within single neurons need to make simplifying assumptions at the level of neurons and synapses (see Brette et al.[1] for a review on available neuronal network simulators). Typically, neurons are abstracted to single-compartment or few-compartment models that interact by discrete electrical pulses, which are commonly referred to as spikes. These simplifications not only bring about technical advantages such as low per-object memory usage and facilitate the parallelization of algorithms but also impose reasonable constraints on the parameter space of neuronal network models.

Whenever a neuron fires a spike, each of the neuron's outgoing connections needs to be notified in order to transmit the signal to the referring target neuron, which then integrates the incoming spike into its dynamics. Neurons in the human cortex generally show sustained spiking activity and maintain up to around $10,000$ outgoing connections. Therefore, parallel simulation of cortical networks involves inter-process communication at short intervals to ensure that spikes, which have been emitted by the neurons of a particular process, will also reach the targets residing on other processes. Moreover, infrastructure for efficient spike delivery is required on each process.

Today networks of the order of 10^5 neurons can be routinely simulated on clusters of about 100 processors. The technology which enables such simulations is freely available and well documented in the literature[2]. This is unfortunately not the case for the few simulation studies on brain size networks published so far[3–5]. The portability of a neuronal network simulator from smaller clusters to supercomputers critically depends on how well the memory utilization of the constituent software components scales with the number of processes. When attempting to simulate a scaled up network of 10^8 neurons on supercomputers of up to $100,000$ nodes, different parts of the infrastructures may overload the available memory.

To investigate this issue systematically, we developed a mathematical model for analyzing the memory usage of a neuronal network simulator for different regimes of network size and number of processes[6]. The model terms and parameters account for the contributions of different software components to the total memory usage and they can be instantiated for a particular neuronal network simulator. Once the model parameters have been determined, the model enables the identification of memory bottlenecks, and potential improvements to the simulator design can be tried out within the model before any implementation work takes place.

The original model is based on the assumption of random network connectivity, which constitutes a worst case scenario, as the human brain shows a hierarchical organization spanning multiple scales[7]. For example, the cortex exhibits complexes of up to $100,000$ neurons, which are commonly referred to as cortical columns and characterized by high intra-column connectivity and specific inter-column connectivity. In order to reduce the memory overhead of neuronal network simulators the existence of such structures can be exploited by storing the neurons of each column on a subset of the available processes.

In this chapter we give a concise description of our model of memory usage and show how the model term describing network connectivity can be adapted to account for a localized storage of neuronal columns (Sec. 2). In Sec. 3 we instantiate the model for the particular case of the Neural Simulation Tool NEST[8], and in Sec. 4 we use the parameterized model to predict how well the memory utilization of NEST scales for networks of 10^7 and 10^8 neurons in the regime of 1000 to $100,000$ processes. We compare the memory usage of columnar networks where the neurons of each column are stored on a subset of the available processes to the worst case scenario of randomly connected networks. The conceptual and algorithmic work described here is a module in our long-term collaborative project to provide the technology for neural systems simulations[8].

2 Adaptation of the Memory Usage Model to Structured Networks

The memory usage of a simulation software typically depends on the number of objects to be simulated and the number of available processes. Objects are distributed across processes but on each process infrastructure is needed that provides efficient access to the objects themselves or to information about the objects. Scalability of these structures with respect to memory usage becomes a critical issue when a parallel software, which was optimized for the simulation of moderate numbers of objects on small clusters, is used to run an accordingly scaled up simulation on supercomputers of more than $10,000$ processors. Serial parts of the infrastructure, i.e. data structures which grow proportionally with the size of the simulated problem, can rapidly saturate the available memory resources.

The model introduced in Kunkel et al.[6] describes the memory usage of a neuronal networks simulator per process in general terms as a function of the number of processes M, the number of neurons N, and the expected number of incoming connections per neuron K. It takes into account contributions of the simulator's basic maintenance structure $\mathcal{M}_0(M)$, of the neuron objects and neuronal infrastructure $\mathcal{M}_\mathrm{n}(M,N)$, and of the connection objects and infrastructure $\mathcal{M}_\mathrm{c}(M,N,K)$, such that the total memory consumption per process is given by

$$\mathcal{M}(M,N,K) = \mathcal{M}_0(M) + \mathcal{M}_\mathrm{n}(M,N) + \mathcal{M}_\mathrm{c}(M,N,K) \tag{1}$$

where the last two terms can be further broken down to

$$\mathcal{M}_\mathrm{n}(M,N) = Nm_\mathrm{n}^0 + (N - N_M)\, m_\mathrm{n}^\emptyset + N_M \left(m_\mathrm{n}^+ + m_\mathrm{n}\right) \tag{2}$$

$$\mathcal{M}_\mathrm{c}(M,N,K) = Nm_\mathrm{c}^0 + N_\mathrm{c}^\emptyset m_\mathrm{c}^\emptyset + \left(N - N_\mathrm{c}^\emptyset\right) m_\mathrm{c}^+ + K_M m_\mathrm{c} \tag{3}$$

in order to clearly distinguish between the per-object memory usage of neurons m_n and synapses m_c and the memory taken up by infrastructure. $N_M = N/M$ and $K_M = N_M K$ denote the expected number of locally stored neuron and synapse objects on each process, respectively. In the neuronal infrastructure, each local neuron causes an overhead of m_n^+ and each remote neuron causes an overhead of $m_\mathrm{n}^\emptyset$. In the connection infrastructure, each neuron with local targets causes an overhead m_c^+ and each neuron without local targets causes an overhead $m_\mathrm{c}^\emptyset$, where $N_\mathrm{c}^\emptyset$ denotes the number of neurons which do not connect to any of the N_M local neurons. The parameters m_n^0 and m_c^0 represent infrastructure overheads which accrue for any of the N neurons.

The definition of $N_\mathrm{c}^\emptyset$ depends on the network connectivity. When assuming random connectivity and a fixed number of K incoming synapses per neuron, such that on a particular process each of the N_M local neurons draws K sources from the set of all N neurons in the network, the probability that a particular neuron is not picked as a source by any of the local neurons is given by $p_\emptyset = (1 - 1/N)^{K_M}$. In this worst case scenario, $N_\mathrm{c}^\emptyset = p_\emptyset N$ is the expected number of neurons without local targets on each process[6].

Alternatively, we can adapt the model parameter $N_\mathrm{c}^\emptyset$ to account for a network which exhibits a columnar structure, where the neurons of each column are stored in a localized way on a subset of the available processes. Each process stores N_M of the N_col neurons of a particular column and their incoming synapses. K_col of the K incoming synapses of each neuron are from sources within the same column, and conversely $K - K_\mathrm{col}$ synapses are from external neurons. The probability that a particular neuron in that same column is not picked as a source by any of the N_M local neurons is $p_\emptyset^\mathrm{intra} = (1 - 1/N_\mathrm{col})^{N_M K_\mathrm{col}}$, and

the probability that a particular external neuron is not picked as a source by any of the N_M local neurons is $p_\emptyset^{\text{inter}} = (1 - 1/(N - N_{\text{col}}))^{N_M(K-K_{\text{col}})}$. Hence, $N_{\text{c}}^\emptyset = p_\emptyset^{\text{intra}} N_{\text{col}} + p_\emptyset^{\text{inter}} (N - N_{\text{col}})$ is the expected number of neurons without local targets on each process.

3 Instantiation of the Model for the Neural Simulation Tool NEST

Our model explicitly considers parallel and serial parts of the infrastructures. Hence, the components of a given neuronal network simulator can be mapped to the model terms and parameters in order to analyze the scalability with respect to memory utilization. Here, we instantiate the model for a particular branch of the Neural Simulation Tool NEST, which is used to explore simulations on supercomputers such as JUGENE and K. Fig. 1 illustrates the major objects and infrastructures of the simulator which contribute to the memory consumption on each process. In NEST, every neuron obtains a unique global identifier (GID). In the particular version of NEST analyzed here, neuron and connection infrastructure each maintain a sparse table[9] of length N that when given a specific GID provides access to the corresponding information or objects. In our memory usage model $m_{\text{n}}^0 = m_{\text{c}}^0 = 2.67$ bits account for the low per-neuron overheads of these two sparse tables.

On each process, the neuronal infrastructure provides information about whether a neuron object resides on this process or on a different process (see Fig. 1A), and in the former

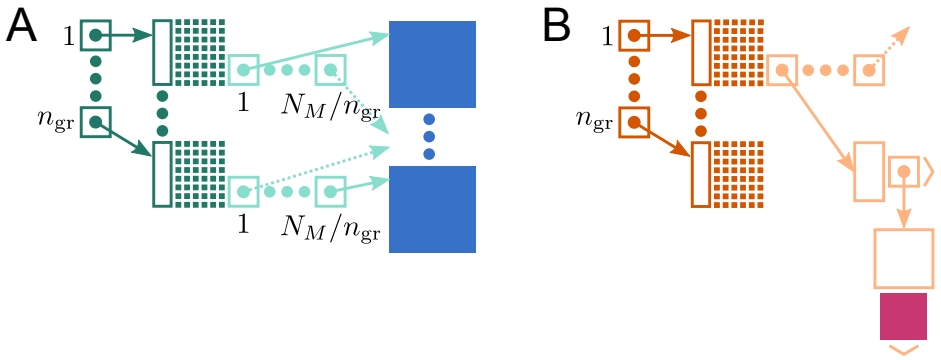


Figure 1. Design of neuronal and connection infrastructure in a specific branch of NEST which is presently employed on supercomputers. Figure adapted from Kunkel et al.[6]. (A) Neuronal infrastructure. Persistent pointers (dark green arrows) to sparse groups are stored in a vector of length n_{gr} (dark green squares). Sparse groups correspond to equally sized groups of neurons with successive global index ranges. Each sparse group has a maintenance overhead including a bit field (dark green rectangle and tiny squares), which indicates whether a neuron is local or remote, and maintains persistent pointers (light green arrows) to local neurons (filled blue squares). (B) Connection infrastructure. Persistent pointers to sparse groups are stored in a vector of length n_{gr} (dark orange arrows and squares). Sparse groups correspond to equally sized groups of neurons with successive global index ranges. Each sparse group has a maintenance overhead including a bit field (dark orange rectangle and tiny squares), which indicates whether a neuron has local targets or not, and it maintains persistent pointers (light orange arrows) to further structures for neurons with local targets. If a neuron has local targets, connector objects are created for all synapse types which the neuron actually uses and a vector (light orange rectangle, square and chevron) is established that holds persistent pointers (light orange arrow) to existing connector objects. Each connector object has a maintenance overhead including the overhead of a vector (big light orange square and chevron) that holds the connection objects (pink square).

case, it holds the persistent pointer to that neuron to facilitate access to local neurons. In addition, the neuronal infrastructure comprises another two persistent pointers per local neuron, the first one is needed to assign the neuron to a sub-network, and the second one is needed during simulation. These three persistent pointers per local neuron yield $m_{\mathrm{n}}^{+} = 24$ B, when assuming that a pointer takes up 8 B of memory, however there is no further overhead for non-local neurons, such that $m_{\mathrm{n}}^{\emptyset} = 0$ B.

Connection objects are stored on the same process as their post-synaptic targets, but are accessed by the GID of their pre-synaptic source and the index of their synapse model. When a neuron emits a spike, the signal is communicated to all processes at the next synchronization step and then on each process it will be delivered to the local targets through the corresponding connection objects. For this purpose, the sparse table in the connection infrastructure provides information about whether a neuron has local targets or not, and in the former case it holds persistent pointers to further structures that enable efficient access to each neuron's locally residing outgoing connections (see Fig. 1B). The function of the inner structure is to facilitate the storage of connection objects sorted by synapse type. If a neuron has local targets, the inner vector holds pointers to type-specific connector objects. Connectors maintain the matching connection objects in a vector, however the connector of a specific synapse type is established only if the neuron actually uses outgoing connections of that type. For simplicity, our analysis is based on the assumption that only one of the possible synapse models is used. In the model the memory usage for neurons with local targets amounts to $m_{\mathrm{c}}^{+} = 104$B, which includes the overhead of an inner vector and its persistent pointer, and the memory usage of a connector object and its persistent pointer. There is no further overhead for neurons without local targets, such that $m_{\mathrm{c}}^{\emptyset} = 0$ B.

The amount of memory taken up by one neuron or one synapse is model dependent. For the analysis performed here we assume per-object memory consumptions of $m_{\mathrm{n}} = 1072$ B and $m_{\mathrm{c}} = 48$ B, which refer to the neuron model `iaf_neuron` and the synapse model `stdp_synapse_hom`, respectively.

A couple of simple simulations that use at most two processes are needed in order to estimate the model parameters. Once the model is parametrized it allows predictions of the memory consumption for different regimes of network size N and number of processes M. We provide detailed information on the parameter estimation in Kunkel et al.[6]

4 Analysis of Component Wise and Total Memory Usage

The model parametrized for NEST enables us to distinguish between the contributions of objects and infrastructures to the total memory usage of the software. Fig. 2 shows the predicted memory usage of the different software components and the total memory usage as a function of the number of processes for a strong scaling scenario, which is for a fixed total number of objects. For the investigated network of size $N = 10^7$ contributions of neurons and neuronal infrastructure to the total memory usage can be neglected throughout. The model predicts that in a regime of several thousand processes the connection objects overload the 2 GB of RAM which are available per node on the JUGENE supercomputer and prevent the simulator from running. About 5000 processes are needed to represent a columnar network of $N = 10^7$ neurons, whereas about 6000 processes are needed to represent the corresponding randomly connected network. The connection infrastructure starts to dominate the total memory usage at about $10,000$ processes in the case of colum-

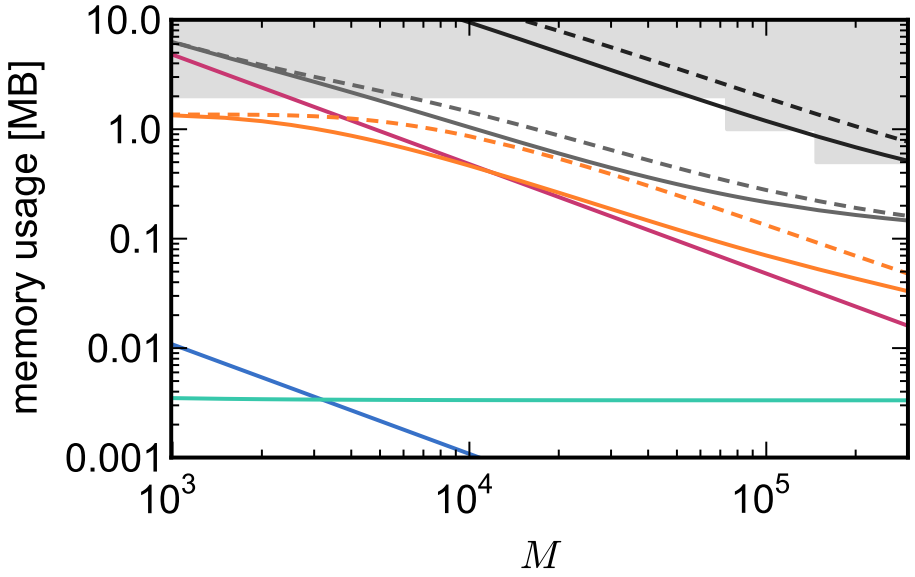


Figure 2. Predicted memory usage per process of neurons (blue), synapses (pink), neuronal infrastructure (green) and connection infrastructure (orange curves), and total memory usage (grey curves) as a function of the number of processes M for a network of size $N = 10^7$ with $K = 10,000$ incoming connections per neuron. Memory usage of connection infrastructure and total memory usage are shown for the case of columnar connectivity with $K_{\mathrm{col}} = 6000$ (solid curves) and for the case of random connectivity (dashed curves). The solid black curve and the dashed black curve correspond to the predicted total memory usage for a columnar and a randomly connected network of size $N = 10^8$, respectively. The shaded area indicates a regime that lies beyond the capacity of the JUGENE supercomputer.

nar connectivity, and already at about 4000 in the case of random connectivity. Fig. 2 also shows the predicted total memory usage of a network of $N = 10^8$ neurons. A columnar network of this size just fits on $M = 147,456$ processes with 1 GB of RAM each, which corresponds to using two cores per node on the entire JUGENE supercomputer, whereas the corresponding randomly connected network cannot be represented on JUGENE.

In the version of NEST analyzed here, infrastructures are implemented as sparse tables, such that non-local neurons and neurons without local targets cause very low overhead. There is no additional memory usage for local neurons in the neuronal infrastructure, however the memory taken up for neurons with local targets in the connection infrastructure is quite substantial. An overhead of $m_{\mathrm{c}}^{+} = 104\mathrm{B}$ accrues as soon as a neuron establishes a single local synapse. Hence, on supercomputers with more than $10,000$ processes random connectivity constitutes a worst case scenario as the outgoing connections of each neuron require the additional m_{c}^{+} on up to about K processes. Storing each column of a structured network on a subset of the available processes also centralizes the intra-column connections and prevents them from causing an overhead in the connection infrastructure on external processes. Compared to the random network fewer connections are distributed across all available processes. This effectively increases the number of neurons without local targets $N_{\mathrm{c}}^{\emptyset}$ in the connection infrastructure. In a regime of few thousand processes,

exploiting the structure of a columnar network of $N = 10^7$ does not result in an effective decrease in memory usage compared to the randomly connected network, as many of the neurons still have targets on all processes. Moreover, in the strong scaling scenario investigated here, the strategy becomes less effective as soon as the number of processes reaches a regime where each column is distributed across a few thousand processes, which then also results in massively distributed intra-column connections. Consequently, storing a columnar network without gathering the neurons of each column on a subset of the available processes takes up almost the same amount of memory as the corresponding randomly connected network (data not shown).

Our analysis shows that exploiting the structure of a columnar network helps to improve the memory utilization of a neuronal network simulator for the case that the software enables high load per process in terms of number of neurons, in particular when there is a considerable overhead for neurons with local targets. We demonstrate that the model presented in Kunkel et al.[6] is a useful tool to investigate the effect of such strategies without any changes to code, as the model is adaptable to different connectivity scenarios.

Acknowledgements

Partially funded by JUGENE Grant JINB33 of the Jülich Supercomputing Centre, BMBF Grant 01GQ0420 to BCCN Freiburg, EU Grant 15879 (FACETS), EU Grant 269921 (BrainScaleS), DIP F1.2, the Helmholtz Alliance on Systems Biology (Germany), the Next-Generation Supercomputer Project of MEXT (Japan), the Research Council of Norway under grant 178892/V30 (eNeuro), Neurex, and the Junior Professor Program of Baden-Württemberg.

References

1. Romain Brette, Michelle Rudolph, Ted Carnevale, Michael Hines, David Beeman, James M. Bower, Markus Diesmann, Abigail Morrison, Philip H. Goodman, Frederick C. Harris, Jr., Milind Zirpe, Thomas Natschläger, Dejan Pecevski, Bard Ermentrout, Mikael Djurfeldt, Anders Lansner, Olivier Rochel, Thierry Vieville, Eilif Muller, Andrew P. Davison, Sami El Boustani, and Alain Destexhe, *Simulation of networks of spiking neurons: A review of tools and strategies*, J Comput Neurosci, **23**, no. 3, 349–398, 2007.
2. Abigail Morrison, Carsten Mehring, Theo Geisel, Ad Aertsen, and Markus Diesmann, *Advancing the boundaries of high connectivity network simulation with distributed computing*, Neural Comput, **17**, no. 8, 1776–1801, 2005.
3. Rajagopal Ananthanarayanan and Dharmendra S. Modha, *Anatomy of a Cortical Simulator*, in: Supercomputing 2007: Proceedings of the ACM/IEEE SC2007 Conference on High Performance Networking and Computing, Association for Computing Machinery, New York, NY, 2007.
4. Eugene M. Izhikevich and Gerald M. Edelman, *Large-scale model of mammalian thalamocortical systems*, Proc Natl Acad Sci U S A, **105**, no. 9, 3593–3598, 2008.
5. Rajagopal Ananthanarayanan, Steven K. Esser, Horst D. Simon, and Dharmendra S. Modha, *The Cat is Out of The Bag: Cortical Simulations with* 10^9 *neurons and* 10^{13}

synapses, in: Supercomputing 09: Proceedings of the ACM/IEEE SC2009 Conference on High Performance Networking and Computing, Portland, OR, 2009.
6. Susanne Kunkel, Tobias C. Potjans, Jochen M. Eppler, Hans Ekkehard Plesser, Abigail Morrison, and Markus Diesmann, "Meeting the memory challenges of brain-scale network simulation", submitted, 2011.
7. Olaf Sporns, Giulio Tononi, and Rolf Kötter, *The Human Connectome: A Structural Description of the Human Brain*, PLoS Comput Biol, **1**, no. 4, e42, September 2005.
8. Marc-Oliver Gewaltig and Markus Diesmann, *NEST (NEural Simulation Tool)*, Scholarpedia, **2**, no. 4, 1430, 2007.
9. Craig Silverstein, "google-sparsehash: An extremely memory-efficient hash_map implementation", `http://google-sparsehash.googlecode.com`, 2005.

Structure, Specificity and Mechanism of ATP Synthase Membrane Motors

José D. Faraldo-Gómez

Theoretical Molecular Biophysics Group, Max Planck Institute of Biophysics
Max-von-Laue Str. 3, 60438 Frankurt am Main, Germany
E-mail: jose.faraldo@biophys.mpg.de

1 Introduction

Adenosine triphosphate (ATP) is the most important energy source in all living cells. Using the energy released upon ATP hydrolysis, the cell is able to sustain many and diverse processes essential for life. From microbes to humans, ATP synthases are the most prominent source of ATP – hence this protein family is of great importance, both from a biomedical (cancer inhibitors, antibiotics) and a biotechnological (metabolic engineering, bioenergetics) standpoint.

ATP synthases are large multi-protein complexes, and are found in the membranes of mitochondria, chloroplasts and bacteria. They produce ATP by catalyzing a chemical reaction between ADP and inorganic phosphate. This reaction, however, requires an input of energy. To power ATP production, ATP synthases facilitate the movement of ions across the membrane, down an electrochemical gradient (Fig. 1). This ion flow causes a rotary mechanism in the protein, which ultimately results in ATP synthesis and release. The mechanism of ATP synthases is therefore akin that of a turbine, or a watermill.

The architecture of the ATP synthase protein complex can be broadly divided into two units, mechanically coupled to each other. ATP synthesis takes place in the so-called F_1 subcomplex[1,2], which lies in the cell interior. The F_o subcomplex, by contrast, is embedded in the membrane, and facilitates ion flow. A ring-like structure in F_o

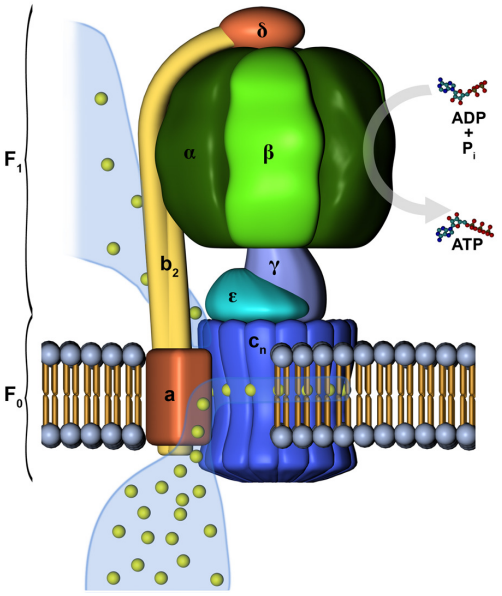


Figure 1. **F_1F_o-ATP synthase complex.** Adapted from Ref. 8. The F_1 unit (α, β) catalyzes the reaction between ADP and P_i. Rotation of the c_n-ring, embedded in the membrane is concurrent with ion flow across the F_o unit (a, c_n), provides the mechanical energy for this reaction.

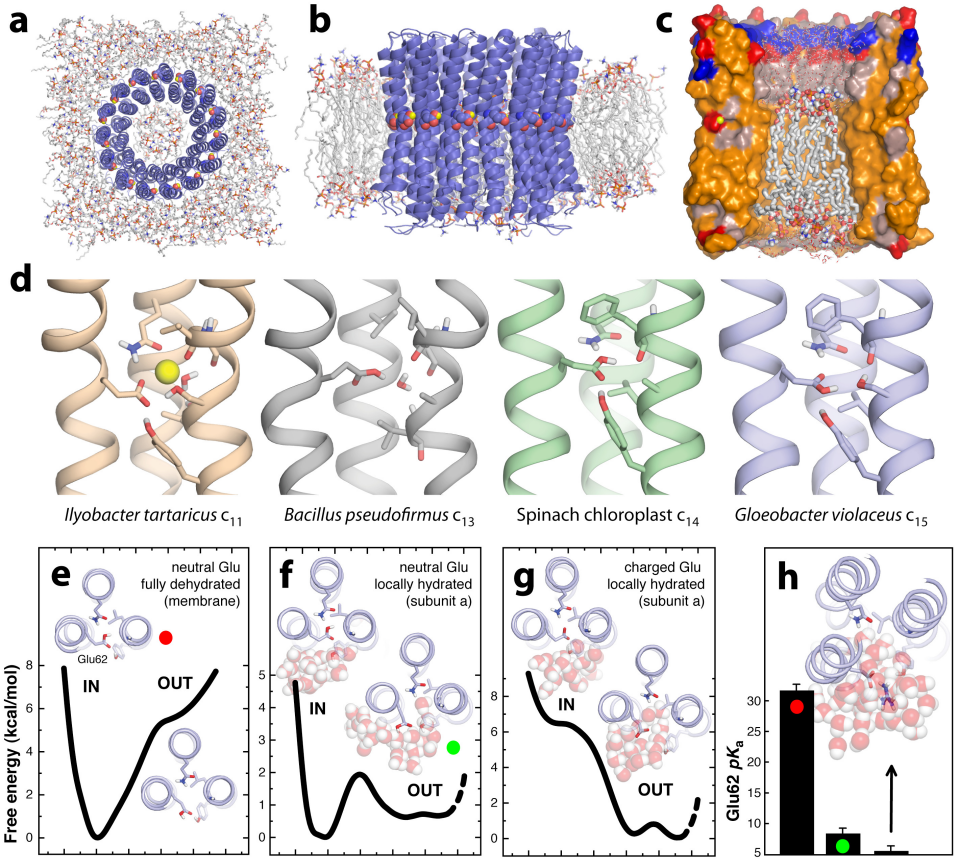


Figure 2. **Insights into the structure and mechanism of the ATP synthase rotor ring from molecular dynamics simulation.** Adapted from Ref. 3. (a, b) Simulation model of a H^+-driven rotor ring in a phospholipid membrane. The ring is an oligomer of helix hairpins (cartoons), or c-subunits, which vary in number for different organisms. The ion-binding sites (spheres) are formed at the outer face of the ring, in between c-subunits, and halfway across the membrane. (c) A hydrophobic (orange surface) pore is formed inside all rotor rings; this pore, however, is occluded by lipid molecules, thus precluding uncontrolled ion diffusion. (d) Atomic structures of Na^+ and H^+ binding sites derived from molecular simulations based on low-resolution or ambiguous X-ray crystallography data. The stoichiometry and origin of each ring are indicated. (e f, g, h) Environmental control of the conformation and occupancy of the c-subunit binding sites, in a representative H^+-driven rotor. When a binding site faces the membrane, the closed conformation of the conserved carboxyl side chain is energetically preferred (e), and the proton is inextricably bound (h, red dot). As the binding site enters the a-subunit interface, local hydration facilitates exchange between closed and open states (f), and allows for deprotonation to occur (h, green dot). Subsequent engagement of a conserved arginine in subunit a stabilizes the deprotonated state (h). After disengagement of this ion pair (now formed with the trailing c-subunit), the open conformation is still energetically preferred, because of local hydration (g). Therefore, rotation of the ring can only proceed once a proton rebinds to the empty site; this enables the site to close, and the c-subunit to face the membrane again.

functions as the rotor. This is an assembly of identical proteins, known as c-subunits, each of which carries an ion-binding site (Fig. 2a-c). These sites can bind either Na^+ or H^+, but other cations, and all anions, are excluded. A second important element in F_o is

subunit a, which lies adjacent to the c-ring rotor. Together, the a/c subcomplex provides access pathways for ions into the membrane; as they cross the membrane, the c-ring rotates against subunit a. This rotation is mechanically transduced to F_1, where it is used to power the synthesis of ATP.

The c-ring rotary motor is therefore the key element that enables the direct coupling between selective, transmembrane ion flow and the enzymatic mechanism of F_1. Understanding its properties and mechanism is therefore of great importance. This is, however, also difficult through experimental means. In addition to the challenges characteristic of membrane protein biochemistry and structural biology, mechanistic and electrophysiological studies of ion transport and ATP synthesis are problematical, since they require re-constitution of a multi-protein complex in artificial membranes that also need to be energized. Similarly, questions relating to the mechanism of assembly and membrane organization are challenging and require novel single molecule biophysical assays. Thus, research in ATP synthases, and more generally, in the area of membrane bioenergetics and transport, has great potential for a meaningful application of computational methods, to help interpret current experimental data and help design novel work.

In this paper, we provide an outline of our recent studies of the structure and mechanism of the ATP synthase membrane rotor, carried out in part with the support of the Jülich Supercomputing Centre. These studies are based upon atomistic molecular dynamics simulations and free energy calculations of various c-rings in model lipid membranes - which are very costly from a computational standpoint. As the record shows, this effort has paid off, and our investigations have had a significant impact in this research field. The availability of computer time at JUROPA has been instrumental for the success of this research.

2 Atomic Structure of Na^+ and H^+-Coupled Rotors

The first calculations in this project were focused on the c-ring of the Na^+ dependent F-type ATP synthase from *Ilyobacter tartaricus*, which includes 11 c-subunits. Although an atomic-resolution crystal structure was available for this ring[10], a number of preliminary observations led us to hypothesize that (a) the binding sites for Na^+ in this structure were incomplete; and (b) that a buried structural water molecule provides the missing coordination ligand for Na^+. Computations were thus used to complete the ion-coordination structure, to ascertain the presence of this additional ligand, and to assess its structural and energetic influence (Fig. 2d). Specifically, using molecular dynamics simulations of the full c_{11} ring in a lipid membrane, we showed that the stability of the binding site structure as measured by X-ray crystallography requires this additional water ligand; in addition, using free energy perturbation simulations, we calculated the dissociation constant of this key water molecule and its occupancy; finally, we showed that this water molecule is also required to rationalize the experimentally measured selectivity of this c-ring for Na^+, against K^+ or Li^+. Subsequent inspection of improved electron density maps from X-ray diffraction validated the computational result. This joint computational-experimental study was published in the *Journal of Molecular Biology*[9].

In a second study, we resolved an outstanding controversy with regard to the H^+-dependent ATP-synthase c-rings from the cyanobacterium *Spirulina platensis* and from spinach chloroplast. At the time, these were the only two H^+-coupled c-rings for

which structural data existed at atomic resolution, including 15 and 14 c-subunits respectively. A most puzzling observation, however, was that while the amino acid sequences of these c-ring rotors are essentially identical in the vicinity of the H^+ binding sites, their reported X-ray structures differed substantially. Again using molecular dynamics simulations of these c-rings in lipid membranes, we showed that the apparent difference in their structures is an artifact of the refinement protocol used for the chloroplast c_{14} ring, added to the underlying low resolution of the diffraction data. Specifically, using a range of structural descriptors, our simulations showed how the reported structure of the c_{14} ring transforms, quickly and reproducibly, into that reported for the c_{15} ring from *S. platensis*. By contrast, we found excellent agreement between the X-ray data for the c_{15} ring and the structural ensembles obtained from the simulations - thus helping to establish the true H^+-coordination structure in this class of ATP synthases (Fig. 2d). A second publication in the *Journal of Molecular Biology* derives from this computational work[6]; our preliminary observations also contributed to the article where the c_{15} structure was first reported, in *Nature Structural & Molecular Biology*[13].

In a third study in this subproject, we addressed the long-standing hypothesis that some ATP synthases that are physiologically coupled to the proton-motive-force do so by binding hydronium ions (H_3O^+) to their membrane rotors, rather than by providing side chains susceptible to protonation (glutamate, aspartate). This notion, first advanced by Nobel laureate Paul Boyer over 20 years ago, has since been used to interpret a number of experimental observations in the ATP synthase field as well as other areas of membrane transport biology. Our study, published in *Biophysical Journal*[7] unequivocally demonstrates that the hydronium hypothesis is implausible on structural and energetics grounds. In this study, we carried out both classical and *ab initio* molecular dynamics simulations of the c_{13} rotor from *Bacillus pseudofirmus* OF4. This rotor is the most suitable to address this question, because in its binding sites it contains what could be interpreted as either a plain water molecule or a hydronium ion, alongside the canonical glutamate side chain (Fig. 2d); this ambiguity cannot be resolved from the X-ray crystallography data[14]. Our simulations demonstrate that the binding site conformation and hydrogen-bonding network observed in the crystal structure are inconsistent with a stable hydronium species bound to the deprotonated site; instead, the structure reflects a protonated glutamate (as in the c_{15} and c_{14} rings above) partially hydrated by a plain H_2O molecule (and seen for Na^+ in the c_{11} ring).

3 Structural Basis for the Ion Specificity of Rotor Rings

From a cell-bioenergetics standpoint the fundamental characteristic of c-ring rotors is their ability to selectively bind H^+ or Na^+. This specificity allows the ATP synthase to operate either under the proton-motive force or the sodium-motive force[11], which are produced and sustained by the metabolism of the organism. Having established the complete binding site atomic structures of the c_{11} and c_{15} rings, we set out to use molecular dynamics simulations and free energy calculations of wild-type and mutagenized rotors to elucidate the structural and energetic basis for their distinct ion specificity. From our analysis we concluded that H^+ selectivity is most likely a robust property of all c-ring rotors, arising from the prominent presence of a conserved carboxylic acid and its intrinsic chemical propensity for protonation - as well as from the structural plasticity of the binding sites. In H^+-coupled rotors, the incorporation of hydrophobic side chains to the binding sites

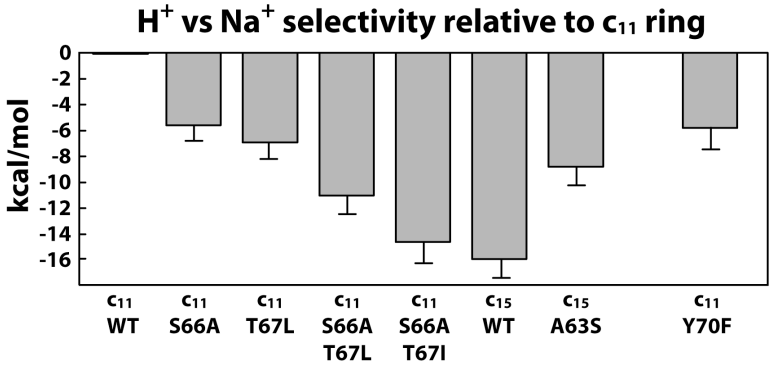


Figure 3. **Ion selectivity and amino acid sequence of the binding sites of two representative c-rings.** Adapted from Ref. 5. The plot shows the free energy of selectivity for H^+ and against Na^+ in the c_{15} and c_{11} rings, as a function of the hydrophobicity / hydrophilicity / size of the ion first-coordination shell in the c-subunit binding sites. All values are relative to the wild-type c_{11} ring. The change in selectivity upon Y70F mutation in the c_{11} site is shown as control (the experimental value is $\sim$ 4.5 kcal/mol).

enhances dramatically this inherent H^+ selectivity. Size restriction may also favour H^+ over Na^+, but increasing size alone does not confer Na^+ selectivity. Rather, the degree to which F_o rotors may exhibit Na^+ coupling relies on the presence of a sufficient number of suitable coordinating side chains and/or structural water molecules. These accomplish a shift in the relative binding energetics, which under some physiological conditions may be sufficient to provide Na^+ dependence (Fig. 3). This study, which also includes experimental data, is now published in *BBA-Bioenergetics*[5]; it is the first of its kind for any protein in the mitochondrial/bacterial respiratory chain, or for any membrane protein selectively coupled to H^+ or Na^+.

4 Mechanism of Ion Transport and Coupled Rotation of the F_o Motor

The rotary mechanism of the c-ring in response to a transmembrane electrochemical gradient, postulated almost two decades ago[4,15], has thus far been supported by electrophysiological and imaging data. In its essence, this rotary mechanism postulates that the occupancy of the ion binding sites in the c-ring is dependent on their local environment. This environment is believed to change as the c-ring rotates inside the membrane, and individual c-subunits become sequentially exposed to the interface with the so-called a-subunit, also in the F_o domain. Ion exchange with the surrounding media at either side of the membrane is thought to occur at this interface, through separate access pathways into the c-ring binding sites; by contrast, the bound ions remain so for those binding sites facing the lipid membrane.

Notwithstanding its wide acceptance, a structure-based microscopic model for this mechanism had not been put forward until now. In this sub-project, we succeeded to utilize molecular simulations and free energy calculations, based on high- and low-resolution structural data, to formulate a microscopic description of the mechanism of coupling be-

tween ion transport and c-ring rotation. This microscopic model was validated by subsequent crystallography and biochemical data. Underscoring the importance and novelty of this work, our calculations and the experiments that followed are now published in *Nature Chemical Biology*[12].

In the proposed mechanism, the environment of the rotor-ring binding sites modulates both their conformation and protonation state (Fig. 2e-h). Specifically, through free energy calculations using the Adaptive-Biasing-Force technique, we identified an open-to-closed conformational equilibrium that is enhanced in the partially hydrated environment of the interface between the rotor c-ring and subunit a. Free energy perturbation simulations show that in this open, partially hydrated state, reversible protonation becomes also energetically viable; so is the reversible formation of a salt-bridge between the key glutamate side chain in the c-ring binding sites and a conserved arginine side chain in the a-subunit. Based on these insights, a three-step mechanistic model was formulated, which can be generalized to H^+ and Na^+ driven motors in all ATP synthase subfamilies (Fig. 4). Crystallographic data and biochemical assays designed to assess this model have so far provided supporting evidence (in preparation).

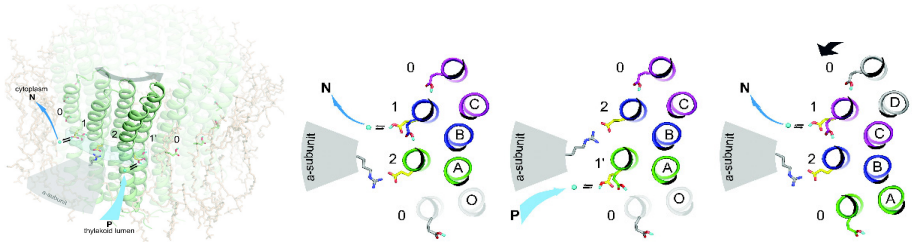


Figure 4. **Hypothetical mechanism of the F_o motor of the ATP synthase.** Adapted from Ref. 12. The mechanism comprises three steps, resulting in sequential ion loading and release, and coupled c-ring rotation. The model derives from extensive free energy molecular dynamics simulations of the H^+-driven c_{15} ring in a pure phospholipid membrane, and in the context of a reduced model of the a-subunit interface.

References

1. Abrahams, J. P., Leslie, A. G. W., Lutter, R., and Walker, J. E. (1994). Structure at 2.8-Å resolution of F_1-ATPase from bovine heart-mitochondria. Nature 370, 621-628.
2. Boyer, P. D. (1993). The binding change mechanism for ATP synthase - some probabilities and possibilities. Biochim. Biophys. Acta 1140, 215-250.
3. Faraldo-Gómez, J. D., and Forrest, L. R. (2011). Modeling and simulation of ion-coupled and ATP-driven membrane proteins. Curr. Opin. Struct. Biol. 21, 173-179.
4. Junge, W., Lill, H., and Engelbrecht, S. (1997). ATP synthase: an electrochemical transducer with rotary mechanics. TIBS 22, 420-423.
5. Krah, A., Pogoryelov, D., Langer, J. D., Bond, P. J., Meier, T., and Faraldo-Gómez, J. D. (2010a). Structural and energetic basis for H(+) versus Na(+) binding selectivity in ATP synthase F(o) rotors. Biochim Biophys Acta 1797, 763-772.

6. Krah, A., Pogoryelov, D., Meier, T., and Faraldo-Gómez, J. D. (2010b). On the structure of the proton-binding site in the F_o rotor of chloroplast ATP synthases. J. Mol. Biol. 395, 20-27.
7. Leone, V., Krah, A., and Faraldo-Gómez, J. D. (2010). On the question of hydronium binding to ATP-synthase membrane rotors. Biophys. J. 99, L53-L55.
8. Meier, T., Faraldo-Gómez, J. D., and Boersch, M. (2011). ATP synthase: a paradigmatic molecular machine. In Molecular Machines in Biology: Workshop of the Cell, J. Frank, ed. (Cambridge, UK: Cambridge Univerisity Press).
9. Meier, T., Krah, A., Bond, P. J., Pogoryelov, D., Diederichs, K., and Faraldo-Gómez, J. D. (2009). Complete ion-coordination structure in the rotor ring of Na^+-dependent F-ATP Synthases. J. Mol. Biol. 391, 498-507.
10. Meier, T., Polzer, P., Diederichs, K., Welte, W., and Dimroth, P. (2005). Structure of the rotor ring of F-type Na^+-ATPase from *Ilyobacter tartaricus*. Science 308, 659-662.
11. Mulkidjanian, A. Y., Galperin, M. Y., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2008). Evolutionary primacy of sodium bioenergetics. Biol. Direct 3, 13-32.
12. Pogoryelov, D., Krah, A., Langer, J. D., Yildiz, Ö., Faraldo-Gómez, J. D., and Meier, T. (2010). Microscopic rotary mechanism of ion translocation in the F-0 complex of ATP synthases. Nat. Chem. Biol. 6, 891-899.
13. Pogoryelov, D., Yidiz, Ö., Faraldo-Gómez, J. D., and Meier, T. (2009). High-resolution structure of the rotor ring of a proton dependent ATP synthase. Nat. Struct. Mol. Biol. 16, 1068-1073.
14. Preiss, L., Yildiz, Ö., Hicks, D. B., Krulwich, T. A., and Meier, T. (2010). A new tpe of proton coordination in an F_1F_o-ATP synthase rotor ring. PLOS Biology 8, -.
15. Vik, S. B., and Antonio, B. J. (1994). A mechanism of proton translocation by F_1F_o ATP synthases suggested by double mutants of the a subunit. J. Biol. Chem. 269, 30364-30369.

Mesoscale Simulations of Human Blood Flow: From Red Blood Cell Elasticity and Interactions to Blood Rheology

Dmitry A. Fedosov and Gerhard Gompper

Theoretical and Soft Matter Biophysics
Institute of Complex Systems and Institute for Advanced Simulation
Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {d.fedosov, g.gompper}@fz-juelich.de

The rheological properties of human blood govern its flow characteristics in the body's cardiovascular system and may serve as an indicator of various blood pathologies and diseases. Using a mesoscopic blood flow model, we accurately predict the dependence of blood viscosity on shear rate and hematocrit. We explicitly represent cell-cell interactions which lead to reversible rouleaux structures and a tremendous increase of blood viscosity at low shear rates. Our results suggest that the presence of a yield stress is associated with cell aggregation, and we present the first quantitative estimates of the magnitude of aggregation forces in blood. This cell model is general enough and can easily be adapted to tune the properties of a much wider class of complex fluids including capsule and vesicle suspensions.

1 Introduction

Rheological and material properties of cell and capsule suspensions have many applications in biology, engineering, and materials science. Blood is one of such examples, which consists of red blood cells (RBCs), white blood cells (WBCs), platelets, and plasma with various suspended molecules. Its main functions are the transport of oxygen and nutrients to cells of the body, removal of waste products, and circulation of molecules and cells which mediate the organism's immune response and haemostasis. RBCs constitute approximately 45% of the total blood volume, WBCs around 0.7%, and the rest is taken up by blood plasma. Healthy human RBCs have a biconcave shape of approximately $8\mu m$ in diameter and are highly deformable. Due to a high volume fraction of RBCs, the rheological properties of blood are mainly determined by their properties and interactions.

Modern rheometry techniques are able to reliably measure macroscopic properties of cell suspensions, for instance the bulk viscosity of blood[1–3]. At low shear rates the RBCs in whole blood have been observed to aggregate into structures called "rouleaux", which resemble stacks of coins[1,4,5]. The aggregation process appears to be strongly correlated to the presence of the plasma proteins[4,5]. Experiments with washed RBCs re-suspended in pure saline to which fibrinogen was added progressively[4] showed a tremendous viscosity increase at low deformation rates with respect to fibrinogen concentration. In addition, such suspensions exhibit a yield stress[1,6,7], i.e., a threshold stress for flow to begin.

These experimental advances have not been accompanied by theoretical developments which can yield quantitative predictions of rheological and flow properties of blood. Recent theoretical and numerical studies focused mostly on the behaviour of a single RBC in various flows[8–10]. Several studies have also been performed to simulate a suspension of multiple cells[11–13] in tube flow.

The aim of our simulation study is to examine blood rheological properties by employing a three-dimensional, deformable RBC model[10]. In particular, we investigate the effect

of RBC aggregation on blood viscosity, reversible rouleaux formation, and yield stress in a RBC suspension[14]. We also present the first quantitative estimates of the magnitude of aggregation forces in whole blood.

2 Models and Methods

Dissipative Particle Dynamics (DPD)[15,16] is a mesoscopic particle method, where each particle represents a collection of atoms or molecules rather than an individual atom, and can be thought of as a soft lump of fluid. The DPD system is constructed by N point particles, which interact through soft pairwise forces. The DPD system is kept at equilibrium temperature with a local thermostat. The time evolution of velocities and positions of particles is determined by Newton's second law of motion and is integrated using the velocity-Verlet algorithm.

2.1 Red Blood Cell Model

The RBC membrane is represented by a collection of $N_v = 500$ DPD particles $\{\mathbf{x}_{i=1\ldots N_v}\}$, which are the vertices of a two-dimensional triangulated network on the RBC surface[8,10,17,18]. The network has a fixed connectivity with the energy

$$U(\{\mathbf{x}_i\}) = U_{spring} + U_{bend} + U_{area} + U_{volume}, \tag{1}$$

where U_{spring} is the spring's potential energy, U_{bend} is the bending energy, and U_{area} and U_{volume} impose the area and volume conservation constraints, respectively. The U_{spring} contribution provides membrane shear elasticity similar to that of a spectrin network of RBC membrane. A "dashpot" is attached to each spring, and therefore, the spring forces are a combination of conservative elastic forces and dissipative forces, which provide network viscous response similar to RBC membrane viscosity. The bending energy mimics bending resistance of the RBC lipid bilayer membrane, while the area and volume conservation constraints mimic area-incompressibility of the lipid bilayer and incompressibility of a cytosol, respectively.

Linear analysis for a regular hexagonal network allows us to uniquely relate the model parameters and the network macroscopic properties, see Refs. 10, 19 for details. Thus, in practice, the given macroscopic RBC properties serve as an input to be used to calculate the necessary mesoscopic model parameters without any manual adjustment. We also employ a "stress-free" model[10,19], which eliminates existing artifacts of irregular triangulation. It is obtained by computational annealing such that each spring assumes its own equilibrium spring length adjusted to be the edge length after triangulation. More details on the RBC model can be found in Refs. 10, 19.

2.2 RBC Aggregation Interactions

For blood, the attractive cell-cell interactions are crucial for simulation of aggregation into rouleaux. These forces are approximated phenomenologically with a Morse potential,

$$U_M(r) = D_e \left[e^{2\beta(r_0 - r)} - 2e^{\beta(r_0 - r)} \right], \tag{2}$$

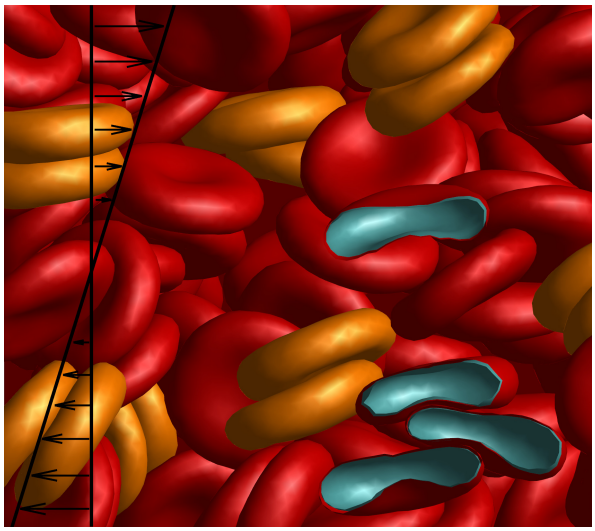


Figure 1. Simulation of whole blood under shear flow. RBCs are shown in red and in orange, where orange colour depicts the rouleaux structures formed due to aggregation interactions between RBCs. The image also displays several cut RBCs with the inside drawn in cyan to illustrate RBC shape and deformability.

where r is the separation distance, r_0 is the zero force distance, D_e is the well depth of the potential, and β characterizes the interaction range. The Morse potential interactions are implemented between every two vertices of separate RBCs if they are within a defined potential cutoff radius r_d. Eventhough the Morse potential in Eq. (2) contains a short-range repulsive force when $r < r_0$, such repulsive interactions cannot prevent two RBCs from an overlap. To guarantee no overlap among RBCs we employ a short range Lennard-Jones potential and specular reflections of RBC vertices on membranes of other RBCs. The specular reflections of RBC vertices on surfaces of other RBCs are necessary due to coarseness of the triangular network which represents the RBC membrane.

2.3 Simulation Setup and Parameters

The RBC suspension was subjected to linear shear flow with periodic Lees-Edwards boundary conditions[20] as shown in Fig. 1. The computational domain had the size of $5.6D_0 \times 4.0D_0 \times 3.4D_0$, where D_0 is the RBC diameter which is equal to about $7.82\ \mu m$ for a healthy RBC. 168 RBCs and 117599 solvent particles were placed in the computational domain. The RBC membrane Young's modulus was set to $Y_0 = 269924\ k_BT/D_0^2$ (k_BT is the energy unit), which corresponds to $Y_0 = 18.9\ \mu N/m$ at physiological temperature of $T = 37^o\ C$. The RBC bending rigidity was assumed to be $\kappa = 3 \times 10^{-19} J$, which is equal to approximately $70k_BT$ at $T = 37^o\ C$. The corresponding Föppl-von Kármán number $0.25Y_0D_0^2/\kappa$ is therefore equal to approximately 963. The membrane viscosity was set to be $12\eta_0$, where η_0 is the suspending fluid viscosity. The coefficients for the area and volume constraints were set large enough in order to closely approximate membrane and cytosol incompressibility. Coupling between the solvent and RBCs was performed through a dissipative force between fluid particles and membrane vertices.

Interactions between different RBCs included the short range repulsive Lennard-Jones potential with parameters $\epsilon = 10.0\ k_BT$ and $\sigma_{LJ} = 0.037\ D_0$. These repulsive interactions result in a thin layer next to a RBC membrane which cannot be accessed by other cells. This layer can be interpreted as a slight increase of the RBC volume. Therefore, the RBC volume was assumed to be about 10% larger than that of the triangulated network. The concentration of RBCs is called hematocrit and denoted as H_t. RBC aggregation interactions were mediated by the Morse potential with parameters $D_e = 3.0\ k_BT$, $r_0 = \sigma_{LJ}$, $\beta = 0.45\ \sigma_{LJ}^{-1}$, and $r_d = 3.7\ \sigma_{LJ}$. For more details see Ref. 14.

3 Results

We present simulation results for the blood viscosity with and without aggregation, rouleaux formation and magnitude of aggregation forces, and yield stress.

3.1 Blood Viscosity

Blood viscosity was computed, with and without aggregation, as a function of the shear rate $\dot{\gamma}$ over the range $0.005s^{-1}$ to $1000.0s^{-1}$ in plane Couette flow. The shear rate and the cell density in our simulations were verified to be spatially uniform. Fig. 2 shows the relative viscosity (RBC suspension viscosity normalized by η_0) against shear rate $\dot{\gamma}$ at

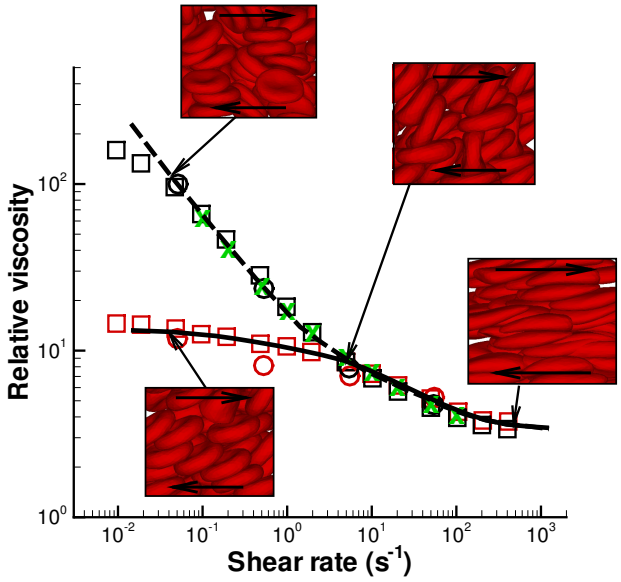


Figure 2. Validation of simulation results for whole blood and non-aggregating RBC suspension. Plot of non-Newtonian relative viscosity (the cell suspension viscosity normalized by η_0) as a function of shear rate at $H_t = 0.45$ and 37^oC: *simulated* curves are in black, and *experimental* points: Whole blood: green crosses - Merril et al.[1]; black circles - Chien et al.[2], black squares - Skalak et al.[3]. Non-aggregating RBC suspension: red circles - Chien et al.[2]; red squares - Skalak et al.[3].

hematocrit $H_t = 0.45$. The blood model predictions are in excellent agreement with the blood viscosities measured in three different laboratories[1–3]. The blood model, consisting only of RBCs in suspension, clearly captures the effect of aggregation on the viscosity at low shear rates, and suggests that cells and molecules other than RBCs have little effect on the viscosity, at least under healthy conditions. At intermediate shear rates, where aggregation is no longer relevant, shear thinning is due to a transition from tumbling to tank-treading motion, accompanied by strong cell deformations[14].

Recent attempts in modeling and simulation of two-cell[21,22], and of multiple-cell aggregates[11] simulated only a single rouleau of a few RBCs in flow. Specifically, in Ref. 11 the link of viscosity to RBC aggregation was investigated, but the viscosity predictions failed to capture the steep rise of that function at low shear rates.

3.2 Reversible Rouleaux Formation

The formation of rouleaux in blood occurs in equilibrium and at sufficiently small shear rates, while large shear rates result in immediate dispersion of gentle RBC structures. Experimentally, aggregation is observed[1] to be a two-step process: the formation of short linear stacks with few RBCs, followed by their coalescence into long linear and branched rouleaux. As the shear rate increases the large rouleaux break up into smaller ones, and at higher values the suspension ultimately becomes one of mono-dispersed RBCs[23]. This process then reverses as the shear rate is decreased. This typical formation-destruction behaviour of rouleaux is consistent with the results of our simulations as shown in Fig. 3. At low shear rates (left plot), the initially dispersed RBCs aggregate into large rouleaux

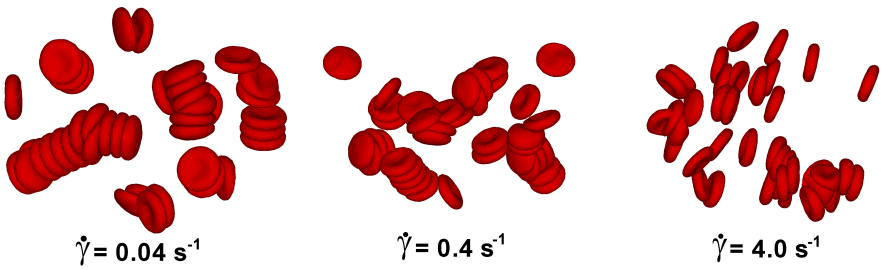


Figure 3. Visualization of aggregation. Simulated reversible rouleaux are formed by RBCs at $H_t = 0.1$. The left plot corresponds to low shear rates, middle plot to moderate share rates, and right plot to high shear rates as indicated with the shear rate values.

of up to about 20 RBCs; as the shear rate is increased to moderate values (middle plot), these structures are reduced in size until at high rates (right plot) they are dispersed almost completely into individual RBCs. Reversibility is demonstrated by reduction of the shear rate to the formation value at which point individual RBCs begin to re-aggregate.

3.3 Yield Stress and Aggregation

Whole blood is believed to exhibit a yield stress, i.e. a threshold stress for flow to begin[1,6,7], which is often estimated by the extrapolation of measured shear stress to the zero shear rate

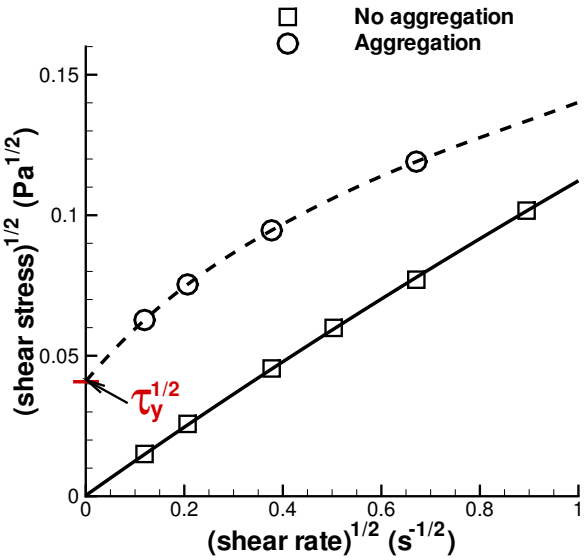


Figure 4. Correlation of aggregation with yield stress. Casson plots with a polynomial fit showing the extrapolated intercept τ_y for simulated suspensions with, dashed lines, and without aggregation, solid lines, at $H_t = 0.45$.

on the basis of the Casson's equation[24],

$$\tau_{xy}^{1/2} = \tau_y^{1/2} + \eta^{1/2}\dot{\gamma}^{1/2}, \tag{3}$$

where τ_y is a yield stress and η is the suspension viscosity at large $\dot{\gamma}$. The assumptions of Casson's relation are likely to hold at least at low shear rates, which was successfully demonstrated for pigment-oil suspensions[24], Chinese ovary hamster cell suspensions[25], and blood[4]. Fig. 4 is a polynomial fit in Casson coordinates ($\dot{\gamma}^{1/2}$,$\tau_{xy}^{1/2}$) to the simulated data for a $H_t = 0.45$ suspension, which shows clearly that τ_y is non-zero for the aggregating RBC suspension, while τ_y is absent without cell aggregation. The yield stress for blood has previously been attributed to the presence of rouleaux in experiments reported in Refs. 1, 6, 7. Merrill et al.[1] found τ_y of healthy human blood to lie between 0.0015 and $0.005\ Pa$ at $H_t = 0.45$. Our simulation results in Fig. 4 fall into this range of the yield stress of whole blood.

3.4 Magnitude of Aggregation Forces

The predictions of Fig. 2 show that a suspension of modelled RBCs captures the viscosity of healthy whole blood with cell aggregation. The plausibility of the aggregation strength was checked by calculation of the maximum force needed to break up two aggregated RBCs. The break-up pulling force in the normal direction is about $3.0\ pN$ - $7\ pN$, where the lower value corresponds to a peeling breakup. Tangential or sliding breakup requires a force in the range of $1.5\ pN$ - $3\ pN$. These forces are much smaller than those imposed on single RBCs in stretching tests with optical tweezers[26], and are consistent with

observations of rouleaux which do not show any large cell deformations. In addition, measurements of a disaggregation force in shear flow by Chien et al.[27] indicate that the shear stress required to break up rouleaux structure lies approximately between $0.01\ Pa$ and $0.1\ Pa$, while the analogous simulations yield about $0.02\ Pa$.

4 Conclusions

We have presented simulation results on the rheological properties of human blood. The RBC suspension model is able to accurately predict shear-dependent viscosity of blood with and without aggregation interactions between RBCs. The RBC aggregation model was able to properly capture the assembly of RBCs into rouleaux structures. These simulations also confirmed that whole blood is a fluid with a non-zero yield stress. The predictive capability of the current cell/capsule suspension model can readily be extended to a variety of engineering and material science applications, which may aid in the development of new soft materials. Finally, such simulations of soft capsule suspensions are computationally demanding and are only feasible on massively parallel computers.

Acknowledgements

D.A.F. acknowledges funding by the Humboldt Foundation through a postdoctoral fellowship. The simulations have been performed on JUROPA of the Forschungszentrum Jülich within a grant provided by NIC-VSR.

References

1. E. W. Merrill, E. R. Gilliland, G. Cokelet, H. Shin, A. Britten, and R. E. Wells Jr, *Rheology of human blood near and at zero flow*, Biophys. J., **3**, 199–213, 1963.
2. S. Chien, S. Usami, H. M. Taylor, J. L. Lundberg, and M. I. Gregersen, *Effects of hematocrit and plasma proteins on human blood rheology at low shear rates*, J. Appl. Physiol., **21**, 81–87, 1966.
3. R. Skalak, S. R. Keller, and T. W. Secomb, *Mechanics of blood flow*, J. Biomech. Eng., **103**, 102–115, 1981.
4. E. W. Merrill, E. R. Gilliland, T. S. Lee, and E. W. Salzman, *Blood Rheology: Effect of Fibrinogen Deduced by Addition*, Circ. Res., **18**, 437–446, 1966.
5. S. Chien, S. Usami, R. J. Kellenback, and M. I. Gregersen, *Shear-dependent interaction of plasma proteins with erythrocytes in blood rheology*, Am. J. Physiol., **219**, 143–153, 1970.
6. G. Cokelet, E. W. Merrill, E. R. Gilliland, H. Shin, A. Britten, and R. E. Wells Jr, *The rheology of human blood-measurement near and at zero shear rate*, Trans. Soc. Rheol., **7**, 303–317, 1963.
7. A. L. Copley, C. R. Huang, and R. G. King, *Rheogoniometric studies of whole human blood at shear rates from 1,000-0.0009* $\sec^{-1}$. *Part I. Experimental findings*, Biorheology, **10**, 17–22, 1973.
8. H. Noguchi and G. Gompper, *Shape transitions of fluid vesicles and red blood cells in capillary flows*, Proc. Natl. Acad. Sci. USA, **102**, 14159–14164, 2005.

9. M. Abkarian, M. Faivre, and A. Viallat, *Swinging of red blood cells under shear flow*, Phys. Rev. Lett., **98**, 188302, 2007.
10. D. A. Fedosov, B. Caswell, and G. E. Karniadakis, *A multiscale red blood cell model with accurate mechanics, rheology, and dynamics*, Biophys. J., **98**, 2215–2225, 2010.
11. Y. Liu and W. K. Liu, *Rheology of red blood cell aggregation by computer simulation*, J. Comp. Phys., **220**, 139–154, 2006.
12. J. L. McWhirter, H. Noguchi, and G. Gompper, *Flow-induced clustering and alignment of vesicles and red blood cells in microcapillaries*, Proc. Natl. Acad. Sci. USA, **106**, 6039–6043, 2009.
13. J. B. Freund and M. M. Orescanin, *Cellular flow in a small blood vessel*, J. Fluid Mech., **671**, 466–490, 2011.
14. D. A. Fedosov, B. Pan, W. Caswell, G. Gompper, and G. E. Karniadakis, *Predicting human blood viscosity in silico*, Proc. Natl. Acad. Sci. USA, **108**, 11772–11777, 2011.
15. P. J. Hoogerbrugge and J. M. V. A. Koelman, *Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics*, Europhys. Lett., **19**, 155–160, 1992.
16. P. Espanol and P. Warren, *Statistical mechanics of dissipative particle dynamics*, Europhys. Lett., **30**, 191–196, 1995.
17. D. E. Discher, D. H. Boal, and S. K. Boey, *Simulations of the erythrocyte cytoskeleton at large deformation. II. Micropipette aspiration*, Biophys. J., **75**, 1584–1597, 1998.
18. G. Gompper and D. M. Kroll, "Triangulated-surface models of fluctuating membranes", in: Statistical Mechanics of Membranes and Surfaces, D. R. Nelson, T. Piran, and S. Weinberg, (Eds.), pp. 359–426. World Scientific, Singapore, 2nd edition, 2004.
19. D. A. Fedosov, B. Caswell, and G. E. Karniadakis, *Systematic coarse-graining of spectrin-level red blood cell models*, Comput. Meth. Appl. Mech. Eng., **199**, 1937–1948, 2010.
20. A. W. Lees and S. F. Edwards, *The computer study of transport processes under extreme conditions*, J. Phys. C, **5**, 1921–1928, 1972.
21. P. Bagchi, A. S. Popel, and P. C. Johnson, *Computational Fluid Dynamic Simulation of Aggregation of Deformable Cells in a Shear Flow*, J. Biomech. Eng., **127**, 1070–1080, 2005.
22. T. Wang, T.-W. Pan, Z. W. Xing, and R. Glowinski, *Numerical simulation of rheology of red blood cell rouleaux in microchannels*, Phys. Rev. E, **79**, 041916, 2009.
23. Q. Zhao, L. G. Durand, L. Allard, and G. Cloutier, *Effects of a sudden flow reduction on red blood cell rouleau formation and orientation using RF backscattered power*, Ultrasound Med. Biol., **24**, 503–511, 1998.
24. N. Casson, *Rheology of Disperse Systems*, Pergamon Press, New York, 1992.
25. A. Iordan, A. Duperray, and C. Verdier, *Fractal approach to the rheology of concentrated suspensions*, Phys. Rev. E, **77**, 011911, 2008.
26. S. Suresh, J. Spatz, J. P. Mills, A. Micoulet, M. Dao, C. T. Lim, M. Beil, and T. Seufferlein, *Connections between single-cell biomechanics and human disease states: gastrointestinal cancer and malaria*, Acta Biomater., **1**, 15–30, 2005.
27. S. Chien, L. A. Sung, S. Kim, A. M. Burke, and S. Usami, *Determination of aggregation force in rouleaux by fluid mechanical technique*, Microvasc. Res., **13**, 327–333, 1977.

Chemistry: Molecular Dynamics and Electron Dynamics from First Principles

Wim Klopper

Theoretical Chemistry Group, Department of Chemistry and Biosciences
Karlsruhe Institute of Technology, KIT Campus South
Kaiserstraße 12, 76131 Karlsruhe, Germany
E-mail: klopper@kit.edu

Atomistic simulations from first principles have developed into both prevalent and indispensable computational tools of modern chemistry. They provide researchers with concepts and insights that are central to the understanding of both electron dynamics (*e.g.*, electron transfer/transport) and molecular dynamics (*e.g.*, rotations/vibrations, reaction dynamics) in chemistry. Powerful computer programs have been developed and tailored toward use on state-of-the-art supercomputers. In many areas of chemistry, these programs and supercomputers allow for atomistic simulations on physically relevant length- and time scales that otherwise would be out of reach. Examples of such cutting-edge simulations are provided by the four "chemistry" contributions in the present proceedings of the NIC Symposium 2011.

In his contribution *Ab Initio Molecular Dynamics Simulations of the Adsorption on Complex Surfaces*, A. Groß reports on computational investigations of the adsorption dynamics of small molecules (H_2, O_2) on prestructured nanosurfaces. In particular, H_2 adsorption on Pd surfaces that are precovered with H, S, or Cl atoms is studied, as well as the interaction dynamics of molecular oxygen with stepped Pt surfaces. *Ab initio* molecular dynamics (AIMD) simulations were conducted using the *Vienna Ab initio Simulation Package (VASP)*.

In their contribution *Mechanochemistry of Covalent Bond Breaking from First Principles Simulations*, P. Dopieralski, J. Ribas-Arino, and D. Marx report on *ab initio* metadynamics simulations performed in the presence of a constant external force, thereby computing force-transformed free-energy surfaces. Exploratory studies were conducted on ring opening reactions of a cyclopropane derivative that has been a key mechanophore in recent experiments. The simulations were conducted using the *CPMD* code.

In his contribution *Internal Dynamics in Endohedral Metallofullerenes as Studied by Ab Initio Born-Oppenheimer Molecular Dynamics*, A. A. Popov reports on *ab initio* Born-Oppenheimer molecular dynamics simulations concerning the internal rotational motion of a series of endohedral metallofullerenes. Extended simulation times of up to 50 ps allowed to reveal the details of the rotations of the endohedral metallofullerenes at room temperature and to fill the gap between spectroscopy and (static) density functional electronic structure theory. The simulations were carried out with the *CP2K* code.

Finally, in their contribution *Simulations of Electron Transfer and Electron Transport in Molecular Systems at Metal and Semiconductor Substrates*, P. B. Coto, Ó. Rubio-Pons,

R. Härtle, M. Bockstedte, and M. Thoss report on a novel approach that combines *ab initio* electronic structure methods with quantum dynamics to study electron-transport processes in different molecular systems adsorbed on metal or semiconductor substrates. For example, electron transport was studied in a nanoscale molecular junction, whose bridge can exist in two tautomeric states with different conductance. The program packages *TURBOMOLE* and *VASP* were used.

In conclusion, the four "chemistry" contributions in the present proceedings provide wonderful examples of the variety of applications of first principles atomistic simulations that are feasible today with respect to both electron and molecular dynamics.

Ab initio Molecular Dynamics Simulations of the Adsorption on Complex Surfaces

Axel Groß

Institut für Theoretische Chemie, Universität Ulm, 89081 Ulm, Germany
E-mail: axel.gross@uni-ulm.de

Ab initio molecular dynamics (AIMD) simulations represent a versatile tool to study dynamical processes of molecules, solids and/or surfaces as will be demonstrated focusing on the adsorption dynamics of molecules on nanostructured surfaces such as precovered or stepped surfaces. Performing AIMD simulations corresponds to unbiased computer experiments allowing the identification of mechanistic processes whose sequence is sometimes unexpected. This will be illustrated using the H_2 adsorption on a precovered surface without any dimer vacancies as an example. In addition, first results related to the adsorption of molecular oxygen on stepped platinum surfaces will be presented.

1 Introduction

In *ab initio* molecular dynamics simulations, the forces necessary to integrate the classical equations of motions are calculated *on the fly* using some suitable first principles total energy electronic structure method[1]. This makes the method very versatile since no fitting of multi-dimensional potential energy surfaces is required which can be rather cumbersome[2]. Yet, in particular for the adsorption dynamics on metal surfaces, such AIMD simulations have been prohibitively expensive just a few years ago allowing only the determination of a few trajectories[3,4]. However, due to the development of efficient periodic electronic structure codes[5,6] based on density functional theory (DFT) and the improvement in the available computer power this situation has now changed. It has become possible to determine thousands of trajectories of molecules impinging on surfaces[7–9]. Note that the statistical uncertainty of reaction probabilities does not depend on the complexity of the system making the determination of statistically significant reaction probabilities on structured surfaces possible.

Furthermore, the AIMD simulations provide an unbiased view on dynamical processes on surfaces. This can lead to the detection of unexpected processes which might still be highly relevant. This is not only of technological, but also of fundamental relevance, as will be illustrated using the interaction of hydrogen with precovered surfaces as an example. Furthermore, the technologically and environmentally important interaction of oxygen with platinum surfaces will also be addressed.

2 Motivation

The interaction of molecules with surfaces is of high technological relevance. All materials interact with their environment through their surfaces. Sometimes this interaction can be harmful and lead, e.g., to corrosion or wear. However, also beneficial processes occur at surfaces. Many chemical reactions are considerably facilitated at catalytic surfaces. This is

of particular interest considering the importance of (electro-)chemical energy storage and conversion for the future energy technology. In order to improve the chemical activity of catalysts, it is important to understand how reactions proceed on surfaces. For decades, surface science studies, both experimental and theoretical, have been extremely successful in elucidating atomistic structures and processes on well-defined clean surfaces. Yet, many technologically relevant processes occur on precovered, structured surfaces. Therefore it is crucial to take into account the full complexity when studying such processes on surfaces.

3 Computational Method

The AIMD simulations were performed using the periodic Vienna *Ab initio* Simulation Package (VASP)[5] within the supercell approach. Electronic exchange and correlation were described within the generalized gradient approximation (GGA)[10,11]. The ionic cores have been represented either by ultrasoft pseudopotentials[12,13] or by projected augmented wave (PAW) potentials[14,15]. The convergence of the electronic structure calculations with respect to the energy cutoff representing the size of the plane-wave basis set and the **k**-point sampling for the integration over the first Brillouin zone was carefully tested.

The AIMD simulations were performed using the Verlet algorithm[16] with a time step of 1 fs within the microcanonical ensemble, i.e., the total energy was conserved during the simulations. This energy conservation was typically fulfilled to within ± 5 meV along a AIMD run. Sticking probabilities for each considered structure and incident kinetic energy were evaluated by averaging over 100 to 150 trajectories which were started with random initial lateral positions and orientations of the impinging molecules far enough above the surface, The statistical error of the sticking probabilities amounts to $\sigma = \sqrt{S(1-S)}/\sqrt{N}$ where S is the sticking probability and N the number of trajectories[17]. For $N \geq 150$, the statistical error is $\sigma \leq 0.04$, and for $N \geq 100$, $\sigma \leq 0.05$. Note that the statistical error does not depend on the complexity of the system, i.e., on the number of considered dynamical degrees of freedom, but only on the number of calculated trajectories.

4 Hydrogen Adsorption on H-Precovered Pd(100)

The adsorption of H_2 on Pd(100) represents one of the model systems in surface science for the study of adsorption phenomena. In addition to its importance in hydrogenation reaction this system is also interesting because of the fact that palladium can absorb huge amounts of hydrogen[18]. Because of the high sticking probability of H_2 on Pd(100)[19], the surface becomes quickly covered by hydrogen once it is exposed to a H_2 gas. In order to model the adsorption of H_2 on H-precovered Pd(100), we have used (2×2), (4×2), (3×2), and (3×3) unit cells with various amounts or preadsorbed hydrogen atoms. Note that there are typically several different arrangements of the adsorbed hydrogen atoms for a given coverage unless the surface is fully covered.

The results of more than 6,000 AIMD trajectory calculations with respect to the sticking probability at various hydrogen coverages are summarized in Fig. 1 where the sticking probability of H_2 on H-precovered Pd(100) is plotted as a function of the coverage. The different values of the sticking probability for the same hydrogen coverage are due to the fact that several arrangements of the pre-adsorbed hydrogen atoms were considered which

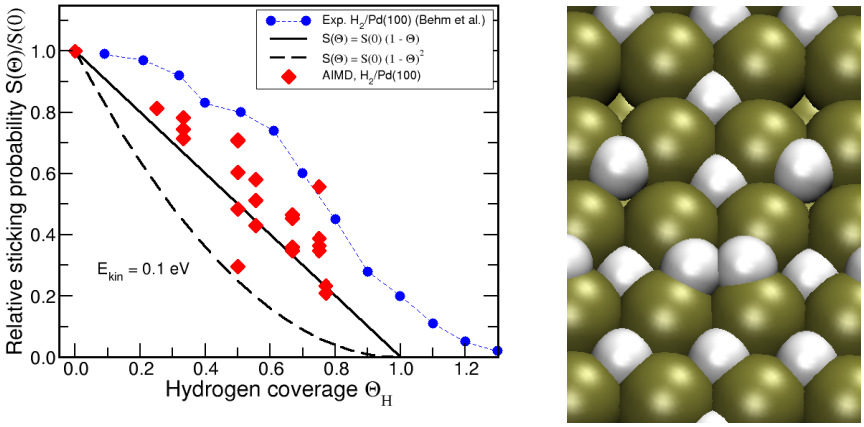


Figure 1. Relative dissociative adsorption probability $S(\Theta)/S(0)$ of hydrogen on hydrogen-covered Pd(100) as a function of coverage. The theoretical results are obtained for an initial kinetic energy of 0.1 eV at different configurations of the adsorbed hydrogen atoms whereas the experimental results by by Behm *et al.*[20] are measured using a H_2 gas. In addition, a snapshot of an AIMD trajectory is shown for a hydrogen precoverage of $\Theta_H = 3/4$ in a 2×4 structure.

in general exhibit different sticking probabilities. In addition, the solid and the dashed line in Fig. 1 correspond to the sticking probability in the case of pure site-blocking without any dynamical effects if jut one or two empty sites, respectively, are required for the H_2 adsorption. The fact that the AIMD sticking probabilities are larger than the results for pure site-blocking indicates that there are important dynamical effects such as steering or dynamical trapping in the adsorption process which enhance the sticking probability.

Experimental results[20] are also included in Fig. 1. Note that the calculations have been done for a mono-energetic molecular beam with a initial kinetic energy of $E_{\rm kin} = 0.1\,$eV whereas the experiment was performed with a H_2 gas. Still it is satisfying that the theoretical results trail the experimental ones.

In addition, in Fig. 1 two curves $S(\Theta) = S(0)(1-\Theta)$ and $S(\Theta) = S(0)(1-\Theta)^2$ are included which correspond to the sticking probability at a precovered surface if it is governed by pure site-blocking requiring one or two empty adsorption sites, respectively. The measured and most of the calculated sticking probabilities are above these two curves indicating that there are steering and trapping effects that lead to a higher sticking probability than expected for pure site-blocking although there is some slight repulsive interaction between pre-adsorbed hydrogen atoms and the impinging hydrogen molecule, as DFT calculations show. Interestingly enough, for the adsorption of H_2 on hydrogen-precovered Pd(111) both experimental and theoretical results find a dependence of the sticking probability on the coverage which is close to the $S(\Theta) = S(0)(1-\Theta)^2$ curve[21].

This significant difference between Pd(100) and Pd(111) is related to the fact that on Pd(100) the hydrogen atoms in the four-fold hollow adsorption site are located almost at the same height as the surrounding Pd atoms whereas on Pd(111) the adsorbed hydrogen atoms in the smaller three-fold hollow sites are positioned higher above the Pd atoms. Consequently, on Pd(100) the adsorbed hydrogen atoms are effectively shielded by the Pd

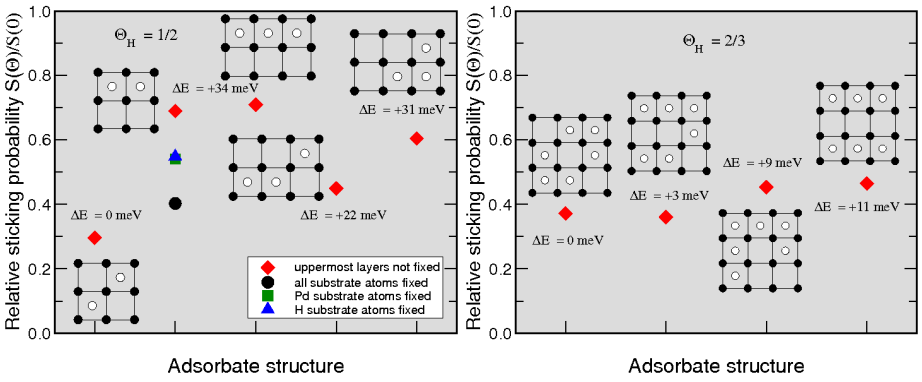


Figure 2. Calculated relative dissociative adsorption probability $S(\Theta)/S(0)$ of H_2 impinging on hydrogen-covered Pd(100) with a coverage of $\Theta_H = 1/2$ and $\Theta_H = 2/3$, respectively, with an initial kinetic energy of $E_{kin} = 0.1$ eV for three different adsorbate structure within a (3×3) periodicity as indicated in the insets. In addition, the stability ΔE in meV of the adsorbate structures per hydrogen atom with respect to the most favourable structure is given in the figure.

atoms leading to a rather small mutual repulsion between the hydrogen atoms. On Pd(111), there is a stronger H-H repulsion which effectively leads to the much lower sticking probability than on Pd(100) at the same hydrogen coverage.

The fact that there is still some admittedly small mutual hydrogen repulsion on Pd(100) is illustrated in Fig. 2 where for several hydrogen adsorbate structures their stability ΔE per hydrogen atom with respect to the most favourable structure is given. Typically, the more compact the H-adsorbate structures, i.e., the higher the number of nearest-neighbour H atoms, the less stable the structure is. For the corresponding sticking probability of H_2 molecules impinging on these particular structures that is also plotted in Fig. 2, we find an anti-correlation with respect to the stability. The more compact H structures which as a consequence also exhibit compact vacancy structures lead to higher sticking probabilities. Still, it is important to realize that also on substrates where there is no dimer vacancy present, as for the most stable structures included in Fig. 2, still H_2 dissociative adsorption is possible. The dissociation starts with one H atom entering the four-fold hollow site with the other H atom remaining at a neighbouring bridge site. From there, it propagates in an exchange-like mechanism until it finds an empty hollow site.

For one particular $\Theta_H = 1/2$ structure within a 2×2 geometry, simulations with different subsets of substrate atoms kept fixed have been performed. The corresponding sticking probabilities also plotted in Fig. 2 show that there is a strong dependence of the sticking probability on the recoil of the substrate atoms, much stronger than on flat surfaces[7]. This again indicates the importance of dynamical effects in these strongly corrugated, non-activated system.

5 Hydrogen Adsorption on S- and Cl-Precovered Pd(100)

Both the study of sulfur as well as chlorine co-adsorption is technologically relevant and interesting. Sulfur is known as a poison for the car exhaust catalyst, and on Pd(100), it

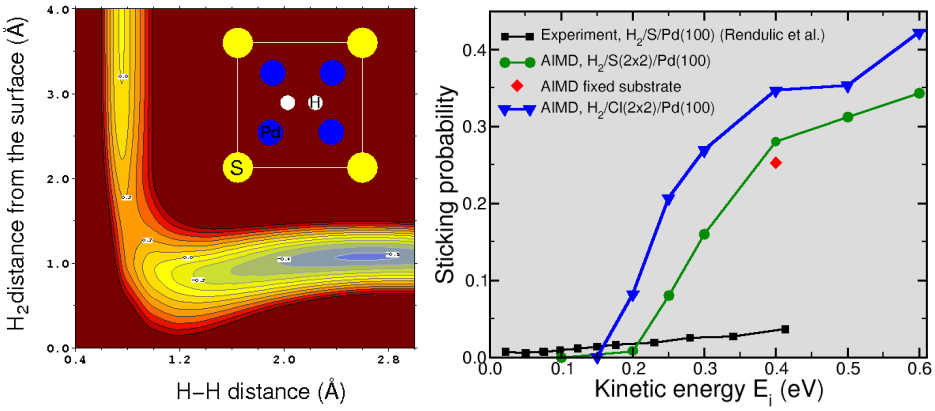


Figure 3. Left panel: Two-dimensional cut through the potential energy surface of H_2/S(2×2)/Pd(100) as a function of the H-H distance and the H_2 distance from the surface calculated by periodic DFT calculations. The inset illustrates the molecular orientation and lateral centre of mass position. The contour spacing is 0.1 eV. Right panel: Sticking probability determined in AIMD simulations for H_2 impinging on S(2×2)- and Cl(2×2)-covered Pd(100). For $E_{\mathrm{kin}} = 0.4\,\mathrm{eV}$ and H_2/S(2×2)/Pd(100), the AIMD result for a fixed substrate is also shown. In addition, the measured sticking probability of H_2 on S-covered Pd(100)[19] is included.

leads to a significant reduction of the hydrogen dissociation probability[19,22]. The specific adsorption of chlorine is a typical process occuring at electrochemical metal-electrolyte interfaces. Whereas there is already some mutual repulsion between adsorbed hydrogen atoms, preadsorbed sulfur or chlorine atoms lead to much stronger repulsion, turning the non-activated H_2 adsorption on Pd(100) into an activated none, i.e., there are no longer any paths without any barrier towards dissociative adsorption[23].

The minimum barrier for H_2 dissociation on S(2×2)/Pd(100) is illustrated Fig. 3 where a two-dimensional cut through the potential energy surface is shown. Whereas the particular path shown in Fig. 3 is non-activated on clean Pd(100)[24], it becomes hindered by a barrier of about 0.25 eV. In fact, chlorine adsorption has a rather similar effect as sulfur on the H_2-Pd(100) interaction, the barriers are just 50 meV smaller.

Such a poisoning of the substrate of course also changes the adsorption dynamics significantly from non-activated to activated behaviour, i.e., the sticking probability typically increases monotonically with the incident kinetic energy, as Fig. 3 shows, where the results of AIMD simulations are also shown.The difference in the minimum barrier heights between H_2/S(2×2)/Pd(100) and H_2/Cl(2×2)/Pd(100) are reflected in the corresponding shift between the sticking curves. Interestingly, keeping the substrate fixed modifies the sticking probability only to a small extent. Apparently for an activated system the recoil of the substrate atoms matters less than for a strongly corrugated non-activated system (see Fig. 2).

Still, for H_2/S(2×2)/Pd(100) there is a large discrepancy between AIMD and experimental results. Although shortcomings of DFT calculations can never be excluded, it should be noted that experimentally the preparation of an ordered (2×2) sulfur overlayer on Pd(100) has not been fully confirmed[19,25]. Hence the discrepancy might be due to different sulfur overlayers realized in the calculations and in the experiment.

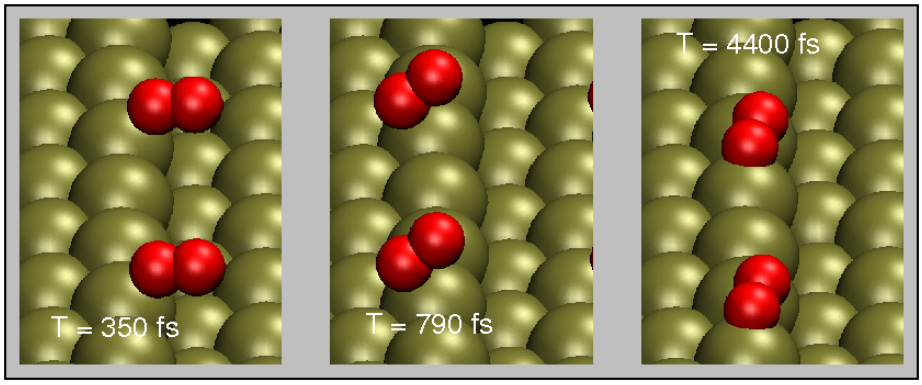


Figure 4. Snapshots of an AIMD trajectory of O_2 impinging on Pt(211) with an initial kinetic energy of $E_{\mathrm{kin}} = 0.1\,\mathrm{eV}$.

6 Interaction Dynamics of O_2 with Stepped Pt

The adsorption of O_2 on Pt is one of the crucial steps occurring in the car exhaust catalyst. Since realistic catalysts do not correspond to low-index single crystal surfaces but rather consists of small particles with high step densities, it is important to clarify the role of steps in catalytically important reactions. On Pt(111), direct O_2 dissociation does not occur due to steric hindrance[26,27], but it is still unclear whether it might happen at stepped surfaces where the energy gain upon adsorption is much larger[28].

From a technical point of view, DFT simulations including O_2 are much more demanding than simulations involving H_2 due to the fact that O_2 is in the triplet state in the gas phase which requires a spin-polarized treatment. This does not only mean that twice as many basis functions have to be taken into account for both the spin up and the spin down components, but typically the convergence of the self-consistent field iteration to solve the Kohn-Sham equations is much worse compared to spin-unpolarized calculations.

Still, it is possible to run AIMD trajectories in the O_2/Pt(211) system. First preliminary results are shown in Fig. 4 where snapshots of a trajectory of O_2 impinging on Pt(211) are depicted. After 350 fs, the O_2 molecule hits the surface with one atom above the lower side of the step. Still, the O_2 molecule is steered towards the upper side of the step ($T = 790\,\mathrm{fs}$) where the most favourable adsorption site is located. Here, the molecule becomes molecularly trapped in the most favourable molecular adsorption site ($T = 4.4\,\mathrm{ps}$) that has already been identified in DFT calculations[28].

The AIMD simulations thus indicate that indeed steps do attract impinging O_2 molecules. Still, so far no dissociative adsorption event has been observed for O_2/Pt(211). Therefore it remains to be seen whether O_2 dissociation on Pt(211) is a two-step process as on Pt(111)[27] or whether it rather corresponds to a direct event.

7 Concluding Remarks

This study has illustrated that *ab initio* molecular dynamics studies are a powerful tool to study the adsorption of molecules on complex surface structures, such as precovered or

stepped surfaces. The unbiased nature of the AIMD simulations allows important insights into the reaction dynamics on surfaces as long as several hundred trajectories are sufficient to obtain statistically meaningful results. Since the statistical error of the AIMD simulations with respect to the sticking probability does not depend on the complexity of the considered systems, AIMD simulations are also well-suited to address the adsorption dynamics of larger molecules such as e.g. small organic molecules. Simulations along these lines are planned for the future.

Acknowledgements

Many fruitful discussions with Fabio Busnengo and Ariel Lozano are gratefully acknowledged. The simulations presented in this work were made possible by a grant of computer time provided by the NIC at the Research Centre Jülich.

References

1. D. Marx and J. Hutter, *Ab initio* Molecular Dynamics: Theory and Implementation, in *Modern Methods and Algorithms of Quantum Chemistry*, edited by J. Grotendorst, volume 3 of *NIC series*, pages 329–477, John von Neumann-Institute for Computing, Jülich (2000).
2. S. Lorenz, M. Scheffler, and A. Groß, Descriptions of surface chemical reactions using a neural network representation, Phys. Rev. B **73**, 115431 (2006).
3. A. Groß, M. Bockstedte, and M. Scheffler, *Ab initio* molecular dynamics study of the desorption of D_2 from Si(100), Phys. Rev. Lett. **79**, 701 (1997).
4. L. C. Ciacchi and M. C. Payne, "Hot-Atom" O2 Dissociation and Oxide Nucleation on Al(111), Phys. Rev. Lett. **92**, 176104 (2004).
5. G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).
6. B. Hammer, L. B. Hansen, and J. K. Nørskov, Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals, Phys. Rev. B **59**, 7413 (1999).
7. A. Groß and A. Dianat, Hydrogen dissociation dynamics on precovered Pd surfaces: Langmuir is still right, Phys. Rev. Lett. **98**, 206107 (2007).
8. A. Groß, *Ab initio* molecular dynamics study of hot atom dynamics after dissociative adsorption of H_2 on Pd(100), Phys. Rev. Lett. **103**, 246101 (2009).
9. A. Groß, *Ab initio* molecular dynamics simulations of the adsorption of H_2 on palladium surfaces, ChemPhysChem **11**, 1374 (2010).
10. J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation, Phys. Rev. B **46**, 6671 (1992).
11. J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77**, 3865 (1996).
12. D. Vanderbilt, Soft self-consistent pseudopotentials in a generalized eigenvalue formalism, Phys. Rev. B **41**, 7892 (1990).

13. G. Kresse and J. Hafner, Norm-conserving and ultrasoft pseudopotentials for the first-row and transition elements, J. Phys.: Condens. Matter **6**, 8245 (1994).
14. P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).
15. G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B **59**, 1758 (1999).
16. L. Verlet, Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules, Phys. Rev. **159**, 98 (1967).
17. P. M. Agrawal, A. N. A. Samadh, L. M. Raff, M. T. Hagan, S. T. Bukkapatnam, and R. Komanduri, Prediction of molecular-dynamics simulation results using feed-forward neural networks: Reaction of a C_2 dimer with an activated diamond(100) surface, J. Chem. Phys. **123**, 224711 (2005).
18. A. Pundt and R. Kirchheim, Hydrogen in metals: Microstructural aspects, Annu. Rev. Mater. Res. **36**, 555–608 (2006).
19. K. D. Rendulic, G. Anger, and A. Winkler, Wide range nozzle beam adsorption data for the systems H_2/nickel and H_2/Pd(100), Surf. Sci. **208**, 404 (1989).
20. R. J. Behm, K. Christmann, G. Ertl, and M. A. Van Hove, Adsorption of CO on Pd(100), J. Chem. Phys. **73**, 2984 (1980).
21. Y. Xiao and W. Dong, Molecular dynamics simulation of a complex surface reaction: The effect of coverage on H_2 dissociation on Pd(111), Phys. Rev. B **83**, 125418 (2011).
22. M. L. Burke and R. J. Madix, Hydrogen on Pd(100)-S: The effect of sulfur on precursor mediated adsorption and desorption, Surf. Sci. **237**, 1 (1990).
23. C. M. Wei, A. Groß, and M. Scheffler, *Ab initio* calculation of the potential energy surface for the dissociation of H_2 on the sulfur-covered Pd(100) surface, Phys. Rev. B **57**, 15572 (1998).
24. S. Wilke and M. Scheffler, Potential-energy surface for H_2 dissociation over Pd(100), Phys. Rev. B **53**, 4926 (1996).
25. M. Rutkowski, D. Wetzig, H. Zacharias, and A. Groß, Rotational and vibrational population of D_2 desorbing from sulfur-covered Pd(100), Phys. Rev. B **66**, 115405 (2002).
26. C. T. Rettner and C. B. Mullins, Dynamics of the chemisorption of O_2 on Pt(111): dissociation via direct population of a molecularly chemisorbed precursor at high-incidence kinetic-energy, J. Chem. Phys. **94**, 1626 (1991).
27. A. Groß, A. Eichler, J. Hafner, M. J. Mehl, and D. A. Papaconstantopoulos, *Ab initio* based tight-binding molecular dynamics simulation of the sticking and scattering of O_2/Pt(111), J. Chem. Phys. **124**, 174713 (2006).
28. P. Gambardella, Ž. Šljivančanin, B. Hammer, M. Blanc, K. Kuhnke, and K. Kern, Oxygen dissociation at Pt steps, Phys. Rev. Lett. **87**, 056103 (2001).

Mechanochemistry of Covalent Bond Breaking from First Principles Simulations

Przemyslaw Dopieralski[1,2], **Jordi Ribas-Arino**[1,3], **and Dominik Marx**[1]

[1] Lehrstuhl für Theoretische Chemie, Ruhr–Universität Bochum, 44780 Bochum, Germany
E-mail: {*przemyslaw.dopieralski, dominik.marx*}*@theochem.rub.de*

[2] Permanent Address: Faculty of Chemistry, University of Wroclaw
Joliot–Curie 14, 50-383 Wroclaw, Poland
E-mail: mclar@elrond.chem.uni.wroc.pl

[3] Present Address: Universitat de Barcelona, Av. Diagonal 647, 08028, Barcelona, Spain
E-mail: jribasjr@yahoo.es

Thanks to a series of novel experimental developments, it is now well established that external mechanical forces can be applied to specific covalent bonds within molecular systems, thus opening the doors to the emerging field of "covalent mechanochemistry" or "molecular nanomechanics". Herein we report on our recent progress on the generalization of our static methodology (established in Bochum in 2009) aimed at exploring force–transformed potential energy surfaces to finite temperatures in order to explore both thermal activation and solvation effects. Our novel formalism, which is based on *ab initio* metadynamics simulations performed in the presence of a constant external force, has enabled us to compute force–transformed *free energy* surfaces. This new framework has been first used to explore the mechanochemistry of a small model system, namely, a cyclopropane derivative, which has been a key mechanophore in recent experiments. Building upon the experience acquired in the study of the cyclopropane derivative, we have made significant progress toward the elucidation of the mechanism of disulfide bond reduction reactions in aqueous solution.

1 Introduction

Only during the last decade became it possible to exert nano–Newton forces on molecules in rather controlled ways as required for triggering and influencing site–specific chemical reactions[1,2]. There is now an avalanche of breathtaking experimental results using different experimental techniques (force–clamp AFM experiments[3–5], sonochemical experiments[6,7], and "molecular force probes"[8]) which are desperately waiting for theoretical explanation of the observed phenomena. Thanks to all these new experiments, it is now well appreciated that mechanical forces can not only promote but also steer truly complex chemical processes, including pericyclic, redox and enzymatic reactions, even if such reactions might be regarded as impossible in the context of "standard" thermal chemistry[6,2,9]. Therefore there is currently an increasing interest in covalent mechanochemistry (CMC)[1,2] with wide-ranging interest in different fields such as synthetic chemistry, materials science, and also biophysics. At the same time the theoretical and mechanistic understanding of the underpinnings of CMC are largely unknown.

Gathering information on the possibilities offered by the use of forces in the realm of reactivity is not only extremely valuable for a fundamental understanding of chemical reactions under external mechanical forces, but also has the potential to dramatically influence both synthetic chemistry[7] and materials science[10,2]. With that in mind, in the last

period of activity we have explored the covalent mechanochemistry of cyclopropanes. Cyclopropane derivatives constitute an important family of mechanophores which have been experimentally studied by Craig and coworkers in the last two years[11,9,12]. The force–induced electrocyclic ring opening of these compound classes has furnished several striking results. The application of a transient tensile force on *gem*–difluorocyclopropanes has been shown to lead to an unexpected isomerization of *trans* species into their less stable *cis* isomers via a mechanochemical trapping of a diradical transition state[9]. On the other hand, *gem*–dichlorocyclopropanes, *g*DCCs, feature a counterintuitive lack of selectivity in the mechanically assisted ring opening reactions of *cis* versus *trans* isomers: whereas it would be expected that the external forces would promote the ring opening of *cis* *g*DCCs more efficiently (due to a better coupling between the mechanical coordinate and the reaction coordinate), the experimental observations indicate that both isomers undergo force–induced ring opening processes with approximately the same probability.

In the last activity period, we have scrutinized the mechanochemical reactivity and force–induced stereochemistry of *g*DCCs based on extensive *ab initio* molecular dynamics simulations (AIMD). As a result of our studies, not only have we provided a rationale for the lack of selectivity but we have also unveiled an unprecedented complex force dependent mechano–stereochemical behaviour[13].

Our studies on cyclopropane derivatives have also been very useful to develop and to validate the methodology we have employed to investigate another extremely important system in the realm of CMC, namely, the mechanically promoted disulfide bond reduction in aqueous environments.

These disulfide bond reduction reactions, which involve the breaking of a disulfide bond (S–S), are an important class of reactions that occurs not only in bulk solution in combination with a wide variety of R–OH and R–SH type nucleophilic reagents, but also in enzymes where it is the basis for breaking up disulfide bridges in proteins which permanently fix tertiary structures via *covalent* interactions. Most surprisingly, recent AFM stretching experiments carried out at *constant* force (aka force–clamp spectroscopy) on model system of disulfide bond unambiguously show that the rate of the reduction reaction exhibits an abrupt change at some critical mechanical force which is about 500 pN when using OH^- as the nucleophilic reagent. Most clearly, two perfectly linear Arrhenius regimes are found which are characterized by vastly different activation energies. It has been speculated that this unexpected phenomenon, which is unknown from analogues thermally activated such reactions, should be connected to different reaction mechanisms in the low and high force regimes[4] but no convincing mechanistic explanation is available yet. The simulations run in the last period have been aimed precisely at elucidating such mechanistic scenario.

2 Explicit External Force Approach

The simulations on both the cyclopropane and the disulfide systems draw on our previously devised conceptual framework based on force transformed potential energy surfaces (FT–PES)[14–16]. In our formalism, the FT–PES, given an external constant force $\mathbf{F}_0$, is rigorously defined as

$$V_{\mathrm{EFEI}}(\mathbf{x}, \mathbf{F}_0) = V_{\mathrm{BO}}(\mathbf{x}) - \mathbf{F}_0\mathbf{q}(\mathbf{x}), \tag{1}$$

where $V_{\mathrm{BO}}(\mathbf{x})$ is the usual Born–Oppenheimer PES as a function of all nuclear cartesian coordinates $\mathbf{x}$, and $\mathbf{q}$ is the mechanical coordinate, i.e. a structural parameter (a generalized coordinate in terms of $\mathbf{x}$) on which the force acts. By locating the stationary points of this function, in which the "External Force is Explicitly Included" (EFEI), one can evaluate, without invoking any approximation, properties such as reactant and transition state structures as functions of $\mathbf{F}_0$. The novelty of the simulations on the *g*DCCs and the disulfide systems lies in the fact that for these studies, the static "0 K perspective" underlying our previous analysis of FT–PESs[14–16] has been rigorously generalized to finite temperatures, thereby including thermal, entropic and dynamical effects by virtue of AIMD techniques. This generalization, in practical terms, has been made possible through the implementation of the EFEI approach in the `CPMD` code[17]. Besides, upon combining the EFEI-AIMD technique with the metadynamics technique[18,19], we have been able for the first time to properly obtain force–transformed free energy surfaces (FT–FES). These new objects have been the basis for the exploration of possible reaction pathways as a function of applied tensile external constant force at a finite temperature.

3 Results on Mechano–Stereochemistry of Cyclopropane Ring Opening Reactions

The model system chosen to explore the mechanochemistry of *g*DCCs is 1,1–dichloro–2,3–dimethylcyclopropane: its cis and trans isomers and the four possible distinct reaction products of the corresponding ring opening processes are depicted in Fig. 1. In the first stage of our study, the thermal chemistry (i.e. the chemistry at zero force) of *g*DCCs was dissected. Our simulations[20] (both at 0 K and at 300 K) have revealed that the ring opening of these molecules to yield the corresponding 2,3–dichloroalkenes proceeds via a concerted disrotatory mechanism, whereby the breaking of the C–C bond takes place in concert with the C–Cl bond cleavage and the subsequent Cl migration. For cis *g*DCC there are two possible pathways as indicated in Fig. 1: the "outward pathway" which passes through TS–I and the "inward pathway" via TS–III. On the basis of the difference between the activation energies of both pathways (about 5 kcal/mol, see Fig. 1), it has been concluded that the thermal ring opening of cis *g*DCC occurs via a "disrotatory outward mechanism", whose BO–PES features a TS of C_s symmetry (TS–I) and a bifurcation point along the IRC after the TS is left behind. By virtue of this topological feature, the migrating Cl atom can move either to the C atom on the right side (thus yielding the Z,R–alkene) or to the left (leading to Z,S–alkene), see Fig. 1. Given the topology of the underlying PES, the ring opening of cis *g*DCC is expected to yield the two enantiomeric alkenes with equal probability. The disrotatory ring opening of trans *g*DCC at zero force, in its turn, implies either the TS–II or the TS–IV according to Fig. 1, with neither of these two TSs featuring symmetry. Trajectory shooting simulations initiated from TS–II have brought to light the fact that the ring opening of trans *g*DCC can yield either the E,S–alkene or the Z,S–alkene in the absence of external forces. The specific calculated probabilities of obtaining these two diastereomers were calculated to be 0.76 and 0.24, respectively. The branching ratio of obtaining E,R– versus Z,R–alkenes via TS–IV are identical because of symmetry. The computed activation free energies were found to be lower (by about 4 kcal/mol) for the ring opening of cis *g*DCC than trans *g*DCC, as shown in Fig. 1.

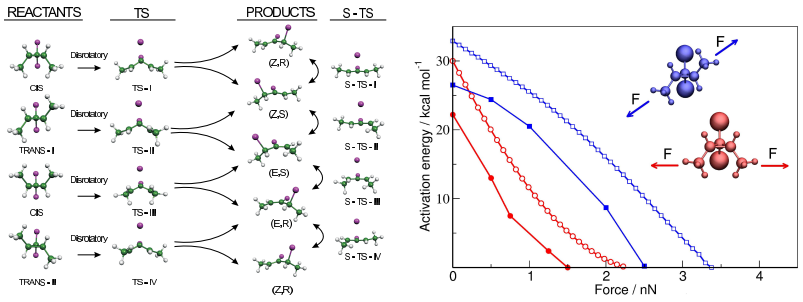


Figure 1. **Left**: Scheme showing the involved chemical species, i.e. all reactants (cis; trans–I and trans–II being enantiomers), transition states (TS–I to TS–IV; S–TS–I to S–TS–IV), and products (Z,R; Z,S; E,S; E,R). The arrows connecting the reactants with the distinct products via the corresponding TSs represent the reaction paths obtained from IRC mapping and *ab initio* trajectory shooting starting from the TSs (see text). The second set of TSs (S–TS) belongs to interconversion reactions between selected products as indicated. For simplicity all structures correspond to the stationary points at zero force, the Z,R product is reproduced twice for clarity, and the Cl atoms are coloured violet. **Right**: Force dependence of activation energies $\Delta E^{\ddagger}(F_0)$ (open symbols) and free energies $\Delta A^{\ddagger}(F_0)$ at 300 K (filled symbols) of the disrotatory ring opening of cis (red circles for the "outward" pathway) and trans (blue squares) 1,1–dichloro–2,3–dimethylocyclopropanes. The stretching force is applied to the C atoms of the two terminal methyl groups as indicated in the inset.

After having set the stage by describing the thermal chemistry of *g*DCCs, we will now focus on the mechanochemical behaviour of these molecules. As justified in Ref. 13, of all the reaction pathways plotted in Fig. 1, only two are relevant for fully describing the mechanochemistry of *g*DCCs: the disrotatory outward mechanism of cis *g*DCC and the disrotatory ring opening of trans *g*DCC passing through TS–II. As shown in Fig. 1, the force dependence featured by the activation free energy, $\Delta A^{\ddagger}$, and the activation energy, $\Delta E^{\ddagger}$, for the two considered pathways follows a similar general trend. The values of $\Delta A^{\ddagger}(F_0)$ have been obtained from the FT-FESs displayed in Fig. 2. Despite the similar trends, the values of $\Delta A^{\ddagger}(F_0)$ are much lower by about 6–7 kcal/mol over the whole range of forces, which results in a dramatic rate acceleration due to finite–temperature and entropy effects not accounted for by $\Delta E^{\ddagger}(F_0)$. With reference to the selectivity with which the tensile load might promote preferentially the ring opening of one isomer, both the qualitatively different shapes of the energy curves associated with the cis and trans reactants and the corresponding rupture forces (1.5 nN and 2.5 nN, respectively, at 300 K) reflect the fact that the external forces enhance the ring opening of cis *g*DCCs more efficiently. This is somehow contradictory with the experimentally found lack of selectivity[11]. As argued in Ref. 13, the most plausible suggestion for resolving such apparent contradiction would be to speculate that sufficiently large forces (forces on the order of 2 nN would suffice, according to the data in Fig. 1) were generated in the sonochemical experiments as to reach the barrierless regime for the ring opening reactions of both isomers.

Last but not least, our exploration of the mechanochemistry of *g*DCCs has unveiled a most surprising feature: the force dependent selectivity of the ring opening of *trans* reactant to yield either the Z– or the E–diastereomer of the corresponding dichloroalkene. As clearly displayed in Fig. 3, the corresponding branching ratio changes dramatically and *non–monotonically* as a function of F_0. This intricate switching behaviour obviously implies that a ratio of about 50:50 is expected to be observed only close to some specific

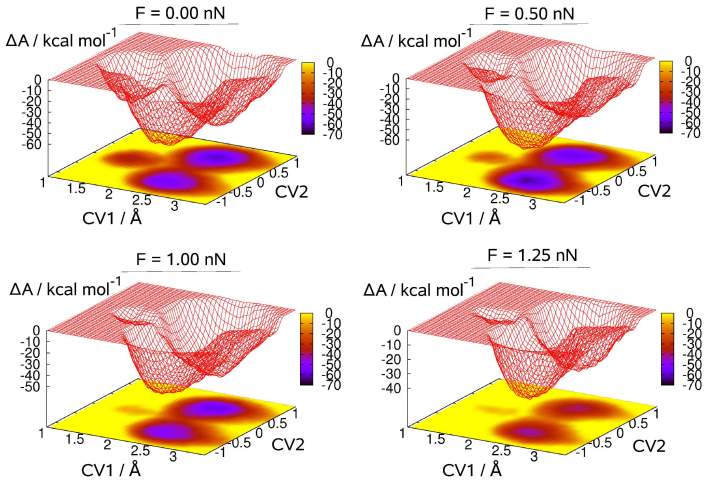


Figure 2. Force–transformed free energy surfaces (FT–FESs) for ring opening for the cis reactant of 1,1-dichloro-2,3-dimethylcyclopropane in the absence of any external force and with three different external forces ($F_0 = 0.5$, 1.0, and 1.25 nN as indicated). These FT–FESs have been obtained by means of metadynamics simulations performed in a reaction subspace spanned by two collective variables: CV1 is the C$\cdots$C distance associated with the bond that yields upon the ring opening process. CV2 is the difference $CN_1 - CN_2$ of the coordination numbers of both chlorine atoms with respect to the left (CN_1) or right (CN_2) carbon atom in the cyclopropane ring.

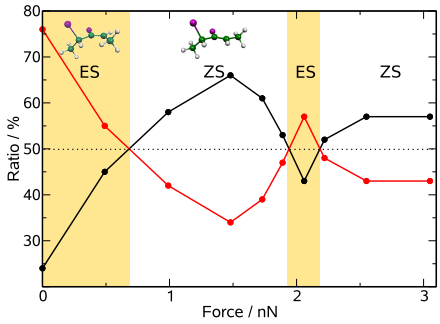


Figure 3. Force dependence of the probability of obtaining the E,S–products (red) versus Z,S–products (black) upon ring opening of trans reactant as computed from dynamical trajectory shooting simulations. The range of forces with yellow and white backgrounds correspond to the forces at which the majority product of the ring opening process is the E,S–alkene or the Z,S–alkene, respectively.

critical forces (0.7 nN, 1.9 nN and 2.2 nN). Importantly, in the limit of large forces the majority product is the Z,S–product, which is just opposite to the situation encountered at zero force. To the best of our knowledge, this is the first time that such an intricate mechanochemical behaviour of a reaction resulting from intrinsically dynamical effects by determining its stereochemistry is reported.

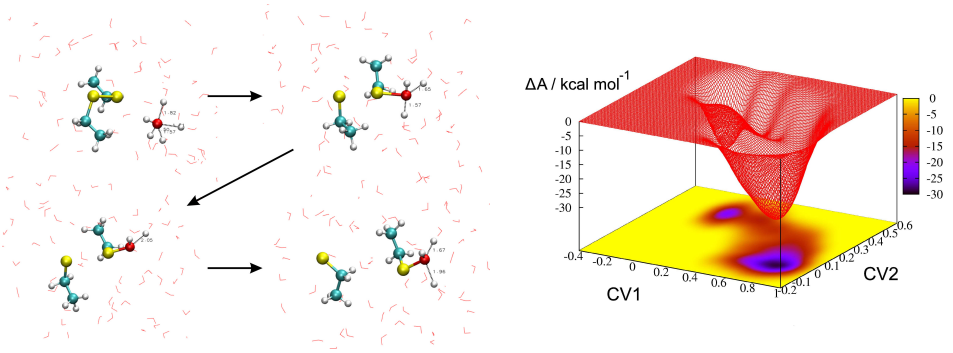


Figure 4. **Left**: Model system with disulfide bridge in aqueous solution: CH_3-CH_2-S-S-CH_2-CH_3 + OH^-. All observed steps of the enforced S–S bond breaking reaction are also displayed. The solvation shell of the attacking OH^-(aq) is indicated by marking the nearest hydrogens using grey spheres to reveal its solvation changes during this reaction. **Right**: Reference free energy surface (FES) for disulfide bond reduction by hydroxide in bulk aqueous solution, OH^-(aq), without applying any external force.

4 Diethyl Disulfide in Aqueous Solution

The model system of our first simulations on the disulfide bond reduction in water consists of a diethyl disulfide molecule (CH_3-CH_2-S-S-CH_2-CH_3), an OH^- ion, which acts as the nucleophile, and 71 explicit water molecules solvating both species (see Fig. 4).

The mechanism for the nucleophilic attack of the OH^- to the disulfide bridge has been investigated by means of metadynamics simulations. The following collective variables have been chosen to properly drive the reaction: (a) coordination number between both sulfurs of the disulfide bridge, and (b) coordination number between the hydroxyl oxygen OH^-(aq) and both sulfur atoms. The free energy surface associated with the reaction at zero force in the subspace spanned by these two variables is shown in Fig. 4. Such free energy landscape and the snapshots of the crucial steps of the reaction featured in Fig. 4 are indicative of a S_N2 reaction mechanism. The value of $\Delta A^{\ddagger}$ that can be estimated from the free energy landscape of Fig. 4 is *ca.* 20 kcal/mol, which agrees very well with the experimental finding that the barrier is in the range of 17.6 to 20.6 kcal/mol[4]. One crucial aspect of the studied reaction is the time evolution of the solvation pattern around (OH^-) during its nucleophilic attack to the disulfide bridge. At the beginning of the simulation, the hydroxide ion was mostly tetracoordianted $OH^-(H_2O)_4$ (in which the hydroxide oxygen accepts four hydrogen bonds) or tricoordinated $OH^-(H_2O)_3$, which is in agreement with previous calculations on OH^- in aqueous solution[21] and experimental findings[22]. Upon approaching one of the sulfur atoms, the hydroxide ion is solvated to a lesser extent by the water molecules. In particular, in the TS region only 2 water molecules are coordinated to OH^-. Once the products are formed (R-S-OH and R-S^-) and the S–S bond is fully broken, the oxygen atom from R-S-OH accepts only one or two H-bonds with water.

5 Concluding Remarks

Upon introducing a theoretical framework which greatly extends static approaches by providing rigorous access to multi–dimensional force–transformed free energy surfaces and genuinely dynamical effects on branching ratios, the mechanisms of force–induced ring opening of *gem*–dichlorocyclopropanes have been dissected. We have also unveiled a complex non–monotonic force dependence of the probabilities for obtaining different stereoisomers of *gem*–dichlorocyclopropanes.

Additionally the first steps to unveil the mechanism of various disulfide bond reduction reactions in aqueous solution have been made. In particular, we have explored the mechanism of hydrolysis reaction (S_N2) and have discarded one of the postulated scenarios in which the reduction of disulfide bond is describe by an energy profile composed of two consecutive energy barriers.

Acknowledgements

It gives us pleasure to thank Martin Beyer, Motoyuki Shiga, Janos Kiss, Padmesh Anjukandi, Martin Krupicka, and Marcus Böckmann for fruitful discussions. We are grateful to Deutsche Forschungsgemeinsschaft (Reinhart Koselleck Grant to D.M.), Alexander von Humboldt Stiftung (Humboldt Fellowships to J.R.A), the Catalan Government (Beatriu de Pinós Fellowship to J.R.A.), as well as Fonds der Chemischen Industrie (to D.M.) for partial financial support. The calculations were carried out at the Jülich Supercomputing Centre.

References

1. M. K. Beyer and H. Clausen-Schaumann, *Mechanochemistry: The mechanical activation of covalent bonds*, Chem. Rev., **105**, 2921–2948, 2005.
2. M. M. Caruso, D. A. Davis, Q. Shen, S. A. Odom, N. R. Sottos, S. R. White, and J. S. Moore, *Mechanically–Induced Chemical Changes in Polymeric Materials*, Chem. Rev., **109**, 5755–5798, 2009.
3. A. P. Wiita, R. Perez-Jimenez, K. A. Walther, F. Gräter, B. J. Berne, A. Holmgren, J. M. Sanchez-Ruiz, and J. M. Fernandez, *Probing the chemistry of thioredoxin catalysis with force*, Nature, **450**, 124–127, 2007.
4. S. Garcia-Manyes, J. Liang, R. Szoszkiewicz, T. L. Kuo, and J. M. Fernandez, *Force–activated reactivity switch in a bimolecular chemical reaction*, Nat. Chem., **1**, 236–242, 2009.
5. J. Liang and J. M. Fernandez, *Mechanochemistry: One Bond at a Time*, ACSNANO, **3**, 1628–1645, 2009.
6. C. R. Hickenboth, J. S. Moore, S. R. White, N. R. Sottos, J. Baudry, and S. R. Wilson, *Biasing reaction pathways with mechanical force*, Nature, **446**, 423–427, 2007.
7. A. Piermattei, S. Karthikeyan, and R. P. Sijbesma, *Activating catalysts with mechanical force*, Nat. Chem., **1**, 133–137, 2009.
8. Q. Z. Yang, Z. Huang, T. J. Kucharski, D. Khvostichenko, J. Chen, and R. Boulatov, *A molecular force probe*, Nat. Nanotech., **4**, 302–306, 2009.

9. J. M. Lenhardt, M. T. Ong, R. Choet, R. E. Christian, T. J. Martinez., and S. L. Craig, *Trapping a Diradical Transition State by Mechanochemical Polymer Extension*, Science, **329**, 1057–1060, 2010.
10. D. A. Davis, A. Hamilton, J. Yang, L. D. Cremar, D. Van Gough, S. L. Potisek, M. T. Ong, P. V. Braun, T. J. Martinez andS. R. White, J. S. Moore, and N. R. Sottos, *Force–induced activation of covalent bonds in mechanoresponsive polymeric materials*, Nature, **459**, 68–72, 2009.
11. J. M. Lenhardt, A. L. Black, and S. L. Craig, *gem–Dichlorocyclopropanes as Abundant and Efficient Mechanophores in Polybutadiene Copolymers under Mechanical Stress*, J. Am. Chem. Soc., **131**, 10818–10819, 2009.
12. D. Wu, J. M. Lenhardt, A. L. Black, B. B. Akhremitchev, and S. L. Craig, *Molecular Stress Relief through a Force–Induced irreversible Extension in Polymer Contour Length*, J. Am. Chem. Soc., **132**, 15936–15938, 2010.
13. P. Dopieralski, J. Ribas-Arino, and D. Marx, *Force–Transformed Free Energy Surfaces and Trajectory Shooting Simulations Reveal the Mechano–Stereochemistry of Cyclopropane Ring–Opening Reactions*, Angew. Chem. Int. Ed., **50**, 7105–7108, 2011.
14. J. Ribas-Arino, M. Shiga, and D. Marx, *Understanding Covalent Mechanochemistry*, Angew. Chem. Int. Ed., **48**, 4190–4193, 2009.
15. J. Ribas-Arino, M. Shiga, and D. Marx, *Unravelling the Mechanism of Force–Induced Ring–Opening of Benzocyclobutenes*, Chem. Eur. J., **15**, 13331–13335, 2009.
16. J. Ribas-Arino, M. Shiga, and D. Marx, *Mechanochemical Transduction of Externally Applied Forces to Mechanophores*, J. Am. Chem. Soc., **132**, 10609–10614, 2010.
17. "`CPMD`, J. Hutter et al., IBM Corp 1990-2004, MPI für Festkörperforschung Stuttgart 1997-2001, see also `http://www.cpmd.org`".
18. A. Laio and M. Parrinello, *Escaping free-energy minima*, Proc. Natl. Acad. Sci., **99**, 12562–12566, 2002.
19. M. Iannuzzi, A. Laio, and M. Parrinello, *Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics*, Phys. Rev. Lett., **90**, 238302–1–4, 2003.
20. "All *ab initio* molecular dynamics, metadynamics, and trajectory shooting simulations were performed on the ft–pes using our in–house version of `cpmd`, the blyp density functional, and a plane wave basis set together with norm–conserving pseudopotentials. all static calculations were performed using `turbomole`, the blyp density functional, and the tzvp basis set".
21. M. E. Tuckerman, D. Marx, and M. Parrinello, *The nature and transport mechanism of hydrated hydroxide ions in aqueous solution*, Nature, **417**, 925–929, 2002.
22. Robert Vacha, Tunde Megyes, Imre Bako, Laszlo Pusztai, and Pavel Jungwirth, *Benchmarking Polarizable Molecular Dynamics Simulations of Aqueous Sodium Hydroxide by Diffraction Measurements*, J. Phys. Chem. A, **113**, no. 16, 4022–4027, 2009.

Internal Dynamics in Endohedral Metallofullerenes as Studied by *ab initio* Born-Oppenheimer Molecular Dynamics

Alexey A. Popov

Leibniz Institute for Solid State and Materials Research, 01069 Dresden, Germany
E-mail: a.popov@ifw-dresden.de

We report on the DFT-based Born-Oppenheimer molecular dynamics (BOMD) studies of the internal dynamics in a series of endohedral metallofullerenes, including carbide clusterfullerenes $Sc_2C_2@C_{82}$, $Sc_3C_2@C_{80}$, $Sc_4C_2@C_{80}$, nitride cluisterfullerenes $Sc_3N@C_{80}$, $ScY_2N@C_{80}$, $TiY_2N@C_{80}$, and one oxide clusterfullerene, $Sc_4O_2@C_{80}$. Extended propagation time of up to 50 ps enabled us to reveal the details of the rotation of endohedral clusters at room temperature and fill the gap between static DFT and experimental spectroscopic studies.

1 Introduction

Endohedral metallofullerenes are molecules comprising closed-shape carbon cages (fullerenes) and metal atoms or clusters encapsulated inside the carbon cage[1]. One of the fascinating properties of endohedral metallofullerenes (EMFs) is their ability to stabilize the endohedral species which are not stable or not known to exist otherwise[1–5]. The most apparent example is the family of nitride clusterfullerenes, $M_3N@C_{2n}$ (M=Sc, Y, lanthanides; $2n$ = 68–96), in which trimetallic nitride clusters are encapsulated inside the carbon cages of different size and symmetry[6]. The most prominent member of this family, $Sc_3N@C_{80}$, is the third most abundant fullerene after C_{60} and C_{70}[7]. Another important class of EMFs is carbide clusterfullerenes in which endohedral cluster also comprises acetylide group C_2^{2-}; typical representatives of this class are $Sc_3C_2@C_{80}$[8], $Sc_4C_2@C_{80}$[9], and $Sc_2C_2@C_{82}$[10]. Recently, a new type of clusterfullerenes has been discovered in which interaction of endohedral metal atoms is mediated by oxygen as in $Sc_4O_2@C_{80}$[11].

Depending on the carbon cage as well as the type of the encapsulated species, the latter can exhibit different types of bonding with dynamic regimes varying from the fixed position inside the carbon cage to nearly free rotation. The cluster dynamics seriously affects spectroscopic properties of the EMFs such as, for instance, ESR spectra of their radicals[12]. In case of rotation of the cluster, experimentally measurable properties correspond to the time-averaged situation (the time scale can reach nanosecond diapason as in ESR and NMR spectroscopies) and prediction of such properties by static DFT computations can be qualitatively incorrect. In most cases nowadays, the information on the dynamics of endohedral clusters is revealed either through the ordering of single-crystal X-ray structures (the EMFs with "free rotation" of the cluster exhibit disordered structures[7]) or through DFT computations of the possible cluster configurations, which show many almost isoenergetic minima for EMF molecules with free rotation and noticeable difference of the energies of such conformers in the case of hindered rotation[13,14]. These methods provide only qualitative conclusions and do not give detailed information on the cluster behaviour.

Molecular dynamics (MD) simulations, which allow direct "observation" of the dynamic behaviour of the species on the microscopic scale, appear to be an ideal tool for addressing the problems of the cluster dynamics and its influence on spectroscopic properties. However, the wide use of such methods is still hindered by the high computational costs, especially when long propagation times are necessary. For this reason, first principle MD studies of endohedral fullerenes are still very rare and in most cases they were either performed at low level of theory or the propagation time was limited to only few picoseconds[4,5,12,15–19]. In this work we apply DFT-based molecular dynamics simulations to study internal dynamics of several representative endohedral metallofullerenes different families with different number of metal atoms.

2 Computational Details

Refined studies of the internal dynamics require on the one hand a reliable theoretical approach (such as DFT) and on the other hand the system should be propagated for a rather long time with precise temperature control. Such calculations require more powerful computational resources than used in our previous works and specially optimized software especially suitable for such purposes. In this project we will apply CP2K[20–23] code installed on the JUROPA supercomputer to advance the level of MD simulations of EMFs to a new level. DFT level of theory is PBE[24]/DZVP, which provides reasonable description of the electronic and molecular structure but is still rather economic. The EMF molecules were placed in the canter of a periodic box with the size of 26×26×26 Å. Computations were performed on 16-32 nodes, the optimal number of processors determined in testing runs to provide insignificant degradation of scalability. Born-Oppenheimer molecular dynamics (BOMD) studies of the EMF molecules were performed at 300 K in canonical ensemble using Nose-Hoover thermostat; each molecule was propagated for 50 ps after 5 ps equilibration using velocity Verlet algorithm with the step of 0.5 fs.

3 Results and Discussion

3.1 Studied Molecules

For the first extended study of internal dynamics in EMFs, we have chosen three groups of clusterfullerenes. First, carbide clusterfullerenes $Sc_3C_2@C_{80}$, $Sc_4C_2@C_{80}$, and $Sc_2C_2@C_{82}$ were studied. In this group we were interested in the dynamics of the whole metal-carbide cluster with respect to the carbon cage and in the internal motions of the acetylide unit with respect to the metal atoms. The influence of the cluster composition (Sc_3C_2 versus Sc_4C_2 inside the same icosahedral carbon cage, C_{80}-I_h) as well as the carbon cage (C_{80}-I_h versus C_{82}-C_{3v}) were also considered. Another group of EMFs comprised nitride clusterfullerenes. The study was started with the archetypical nitride clusterfullerene, $Sc_3N@C_{80}$, in which the Sc_3N cluster is known to rotate freely at room temperature at the NMR time scale[7]. The first goal was to analyze the influence of the cluster composition by comparing dynamics of $Sc_3N@C_{80}$ to that of $Sc_3C_2@C_{80}$ with the same carbon cage (the difference between the two EMFs is in the non-metal component in their endohedral clusters, namely, nitride N^{3-} in Sc_3N and acetylide C_2^{2-} in Sc_3C_2). Besides, we have analyzed the influence of the cluster size on its dynamics by substituting

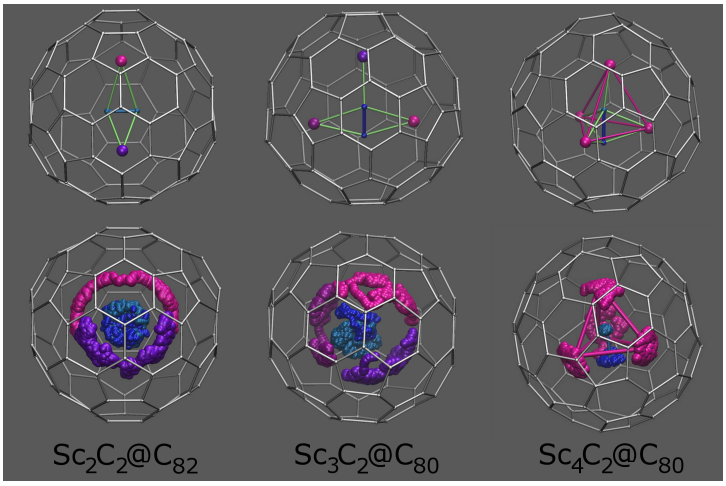


Figure 1. DFT-optimized molecular structurs (upper row) and BOMD trajectories (bottom row, 50 ps at 300 K) of carbide clusterfullerenes. Sc atoms are shown in magenta and violet, cage carbon atoms are light grey (their vibrational displacements in BOMD trajectories are not shown), carbon atoms in acetylide unit are cyan and blue.

two scandium atoms by yttrium atoms which have noticeably larger radius. Then, localization of the spin on the cluster was imposed via substitution of scandium in $ScY_2N@C_{80}$ by a titanium atom, and the effect of such a Sc-to-Ti substitution on the cluster dynamics was followed. Finally, the computations were also performed for the oxide clusterfullerene $Sc_4O_2@C_{80}$. Information on the internal dynamics in this EMF is not available so far, and hence the first insight is provided in this work. Besides, we were interested in how rigid the cluster is (i.e. if migration of the oxygen atoms between the faces of the Sc_4 tetrahedron is possible) and how dynamics of the Sc_4O_2 cluster is different from that of Sc_4C_2, another cluster with tetrahedral arrangement of Sc atoms.

3.2 Carbide Clusterfullerenes

Before the analysis of the internal dynamics in EMFs, their "static" structures in the lowest energy configuration are briefly reviewed. The upper row in Fig. 1 shows DFT-optimized molecular structures of $Sc_3C_2@C_{80}$, $Sc_4C_2@C_{80}$, $Sc_2C_2@C_{82}$. In $Sc_2C_2@C_{82}$, the Sc_2C_2 cluster has a butterfly-shape with non-equivalent Sc atoms and the C_2 unit located perpendicular to the Sc-Sc axis so that two Sc-C-C planes are crossing at the angle of 133 °. The overall symmetry of this configuration is C_s, which is significantly lower than the C_{3v} symmetry of the carbon cage. In $Sc_3C_2@C_{80}$, three Sc atoms form a triangle, and the C_2 unit in the lowest energy configuration of the molecule is tilted from the plane of the metal atoms by 14 °. In this configuration, one Sc atom is bonded only to one carbon atom located close to the centre of the Sc_3 triangle, while two other Sc atoms are equally bonded to each of the carbon atoms of the acetylide unit, making thus a "sub-cluster" Sc_2C_2 with almost identical Sc-C bond lengths. In contrast to the Sc_2C_2 cluster in $Sc_2C_2@C_{82}$, the Sc_2C_2 "sub-cluster" in $Sc_3C_2@C_{80}$ is virtually planar (the angle between two Sc–C–C planes is 171 °). In $Sc_4C_2@C_{80}$, Sc atoms form a tetrahedron, while the C_2 unit is cross-

ing the tetrahedron face so that one of the C atoms is located almost in the centre of the carbon cage[9]. As a result, not all four Sc atoms in Sc_4C_2 cluster are equivalent: the atom which is opposite to the face crossed by the acetylide is bonded to only one carbon atoms, while three other Sc atoms form Sc–C–C triangles with almost equal Sc–C bonds.

The second row of the Fig. 1 shows the BOMD trajectories of three carbide cluster-fullerenes propagated at 300 K for 50 ps. Obviously, cluster dynamics is very different in three EMFs showing that the cluster composition and the carbon cage are both the factor of vital importance. Experimental room temperature NMR studies show that the time average symmetry of the carbon cage in $Sc_2C_2@C_{82}$ is C_{3v} and two Sc atoms are equivalent (*i.e.* only one line is found in the ^{45}Sc NMR spectrum)[10,25]. Thus, at the nanosecond timescale one can conclude that the Sc_2C_2 cluster is rotating freely in side the cage. Although the time scale of the BOMD studies in this work is significantly shorter, it appears sufficient to give the details of the cluster rotation. In particular, it is clearly seen in Fig. 1 that over the period of 50 ps the cluster accomplished a half turn around the axis coinciding with the C_3 axis of the carbon cage. It should be noted that C_{3v} symmetry of carbon cage implies 3 equivalent configurations of C_s symmetry resulting totally in 6 sites (three pairs) which can be occupied by Sc atoms. More detailed insight into the cluster dynamics shows that Sc atoms exhibit librations of rather large amplitude close to one of these minimum-energy positions till at some point the amplitude appears to be sufficient to reorient the cluster to another minimum. While the Sc_2C_2 cluster as a whole rotates around the C_3 axis of the carbon cage, the C_2 unit also rotates with respect to the metal atoms. On a time scale of few picoseconds when Sc atoms resides close to one of the minima, dynamics of the C_2 unit can be described as a large-amplitude rocking motion around the Sc–Sc axis. The motion is chaotic, and the figure formed by the trajectories of the carbon atoms in the C_2 unit collected over the time range of 50 ps has an almost spherical shape (slightly elongated along the C_3 axis). Hence, the quantized rotational states found earlier for $Sc_2C_2@C_{84}$[19] are hardly to be expected in $Sc_2C_2@C_{82}$.

ESR spectroscopic studies of $Sc_3C_2@C_{80}$ show that at room temperature three Sc atoms are equivalent[26], while the DFT-optimized structure has only C_s symmetry (Fig. 1). This fact indicates that the cluster may rotate as a whole inside the cage and that the C_2 unit should also experience rotational motions with respect the Sc atoms to make the Sc atoms equal on a long time scale. BOMD trajectory obtained at 300 K is in principle consistent with these expectations. On a sub-picosecond time scale, the motion of Sc_3 cluster can be described as large-amplitude vibrations of Sc atoms near one of the energy minima which at some moments is switching to another minimum. Over the period of 50 ps, each Sc atom exhibited long displacement trajectory without any preferable rotation direction. This indicates that the rotations of the Sc_3C_2 cluster as a whole is not hindered and complete 3D rotation is expected at longer time scales. In due turn, the acetylide unit exhibits out-of-plane vibrations and displacements with respect to the Sc_3 cluster. The maximum out-of-plane distortion is achieved in the configuration when the C_2 unit is located in the middle of the Sc_3 triangle and is perpendicular to its plane. In this configuration the cluster has an effective D_{3h} symmetry, *i.e.* all three Sc atoms are equivalent. The acetylide unit has then several possibilities to relax back to the in-plane position, which at the long time scale results in the equalizing of all three Sc atoms seen in the ESR spectrum.

Dynamical effects are the least enhanced in $Sc_4C_2@C_{80}$: over the 50 ps time scale, the cluster Sc_4C_2 was oscillating close to one position, while the C_2 unit remained almost

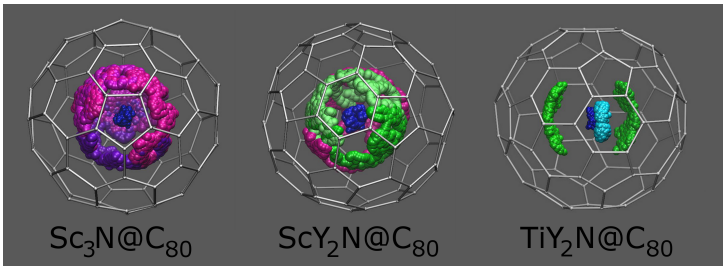


Figure 2. BOMD trajectories (50 ps at 300 K) of nitride clusterfullerenes. Sc atoms are shown in magenta and violet, Ti atom is cyan, Y is green, cage carbon atoms are light grey, nitrogen atom is blue.

fixed with respect to the Sc_4 tetrahedron. ^{13}C NMR spectrum of $Sc_4C_2@C_{80}$ exhibits only two carbon cage signals (*i.e.* effective I_h symmetry of the carbon cage is observed), which means that the cluster is rotating at the NMR time scale[9]. Yet, it can be seen that rotation of the endohedral in $Sc_4C_2@C_{80}$ is much slower than in $Sc_2C_2@C_{82}$ and $Sc_3C_2@C_{80}$. Besides, it is probable that Sc atoms are not equalized even over the long time scale, which should be verified by ^{45}Sc NMR spectroscopy.

3.3 Nitride Clusterfullerenes

Static DFT computations show that in all nitride clusterfullerenes the nitride cluster adopts triangular shape with the nitrogen atoms in the middle and three metal atoms on the vertices. First, we have analyzed internal dynamics of $Sc_3N@C_{80}$. As can be seen in Fig. 2, the Sc_3N cluster rotates rapidly and Sc atoms cover significant part of the sphere over the 50 ps. DFT-based BOMD simulations are in good agreement with earlier experimental and theoretical studies. Interestingly, rotations of the Sc_3N cluster appears to be comparable or somewhat faster than that of Sc_3C_2, which can be concluded from the extensions of rotational trajectories exhibited by both clusters over the same period of time.

Static DFT calculations showed that increase of the cluster size by replacement of Sc by metals with larger ionic radii (Y or lanthanides) results in the increase of the cluster rotation barriers[13,14]. To analyze if this effect can have a noticeable influence on the cluster dynamics over the intermediate time scale, we have performed BOMD simulations for $ScY_2N@C_{80}$. Its 50 ps trajectory shown in Fig. 2 proves that the effect of the metal size on the cluster dynamics is rather weak. At least, rotations of the ScY_2N cluster does not seem to be slowed down in comparison to the rotation of the Sc_3N cluster.

Substantially different dynamical situation evolved when the scandium in $ScY_2N@C_{80}$ was replaced by titanium. At 300 K, the Ti atom remains fixed close to the centre of one hexagon and exhibits vibrations between two pentagon/hexagon edges (Fig. 2). Thus, 3D rotation of the TiY_2N cluster is hindered, at least at the 50 ps time scale. Instead, the cluster tends to rotate around the Ti–N bond, which can be concluded based on the prolonged curved trajectories of both yttrium atoms. However, this rotation also appears to be hindered (*i.e.* it should be better described as "libration"), and the full turn around the Ti–N bond was not accomplished over the 50 ps pointing to the existence of the energy barrier precluding further rotation of the cluster. Similar behaviour (albeit on a shorter time

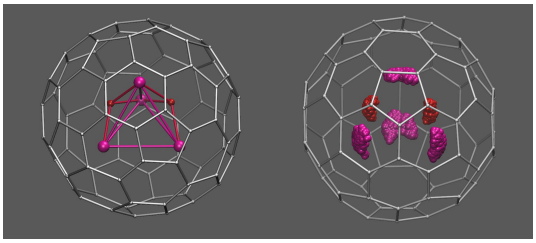


Figure 3. DFT-optimized structure (left) and BOMD trajectory (right, 50 ps at 300 K) of oxide clusterfullerene $Sc_4O_2@C_{80}$. Sc atoms are shown in magenta, oxygen atoms are red, cage carbon atoms are light grey.

scale) was recently described for $TiSc_2N@C_{80}$[17]. Substitution of Sc by Ti can be seen as an addition of one electron to the system. Our BOMD studies of $Sc_3N@C_{80}^-$ did not show noticeable changes of the cluster dynamics in the anion in comparison to that of the neutral molecule[17]. Thus, it is the localization of the spin density on Ti atom and presumably an increased strength of the Ti–C bonding in comparison to that of Sc–C which dramatically change the internal dynamics of the nitride cluster in $TiY_2N@C_{80}$.

3.4 Oxide Clusterfullerene

BOMD trajectory of $Sc_4O_2@C_{80}$ followed for 50 ps is shown in Fig. 3. Dynamics of the cluster is similar to that of $Sc_4C_2@C_{80}$ in that the cluster does not show rapid rotation at this time scale, although vibrational displacements of the metal atoms are rather large. Thus, it appears that the presence of four metal atoms bonding to the carbon is the factor which significantly hinders rotation of the endohedral metal cluster. It is still to be seen how the cluster behaves itself at the longer time scales relevant for ESR and NMR spectroscopy.

4 Conclusions

In this work, we applied DFT-based Born-Oppenheimer molecular dynamics to perform comparative study of the internal dynamics of a series of endohedral clusterfullerenes at 300K. The use of computational facilities of the Jülich Supercomputing Centre along with the fast and efficient computational procedures implemented in CP2K code enabled us to extend the propagation time up to 50 ps while pertaining reasonably rather high level of theory (namely, PBE/DZ2P). At this time scale, dynamics of the endohedral clusters inside carbon cages can be followed with great details and the influence of the cluster composition and the cage size can be revealed. In particular, we have found completely different rotational behaviour of the cluster in a family of carbide clusterfullerenes, which ranges from the rotation around the one axis in $Sc_2C_2@C_{82}$ to free 3D rotation in $Sc_3C_2@C_{80}$ to the hindered librational motions in $Sc_4C_2@C_{80}$. In nitride clusterfullerenes, substitution of two Sc atoms by two larger yttrium atoms did not change the free rotation of the M_3N cluster, but when the Sc atom in ScY_2N cluster was substituted by Ti, rotation of the cluster was changed form 3D to the librational motions around the Ti–N bond. Finally, in the oxide clusterfullerene $Sc_4O_2@C_{80}$ rotation of the cluster is hindered in the same way as it is found in $Sc_4C_2@C_{80}$, which indicates that tetrahedral metal clusters tend to

exhibit similar dynamics properties despite the different electronic and geometrical properties of the non-metal atoms. Our results thus partially fill the gap between the results of static DFT computations and experimental ESR and NMR studies which give information on the internal dynamics on the nanosecond timescale. Our results can be also used to fit and validate cheaper semiempirical approaches which can be then employed for molecular dynamics studies with much longer propagation time.

Acknowledgements

The author is thankful for a computational time on JUROPA granted by Jülich Supercomputing Centre. This work was supported by DFG (PO 1602/1-1).

References

1. H. Shinohara, *Endohedral Metallofullerenes*, Rep. Prog. Phys. **63**, 843–892, (2000).
2. M. N. Chaur, F. Melin, A. L. Ortiz, and L. Echegoyen, *Chemical, Electrochemical, and Structural Properties of Endohedral Metallofullerenes*, Angew. Chem.-Int. Edit. **48**, 7514-7538, (2009).
3. A. A. Popov, *Metal-Cage Bonding, Molecular Structures and Vibrational Spectra of Endohedral Fullerenes: Bridging Experiment and Theory*, J. Comput. Theor. Nanosci. **6**, 292-317, (2009).
4. L. Dunsch, S. Yang, L. Zhang, A. Svitova, S. Oswald, and A. A. Popov, *Metal Sulfide in a C_{82} Fullerene Cage: A New Form of Endohedral Clusterfullerenes*, J. Am. Chem. Soc. **132**, 5413-5421, (2010).
5. A. A. Popov, L. Zhang, and L. Dunsch, *A Pseudoatom in a Cage: Trimetallofullerene $Y_3@C_{80}$ Mimics $Y_3N@C_{80}$ with Nitrogen Substituted by a Pseudoatom*, ACS Nano **4**, 795-802, (2010).
6. L. Dunsch, and S. Yang, *Metal Nitride Cluster Fullerenes: Their Current State and Future Prospects*, Small **3**, 1298–1320, (2007).
7. S. Stevenson, G. Rice, T. Glass, K. Harich, F. Cromer, M. R. Jordan, J. Craft, E. Hadju, R. Bible, M. M. Olmstead, K. Maitra, A. J. Fisher, A. L. Balch, and H. C. Dorn, *Small-Bandgap Endohedral Metallofullerenes in High Yield and Purity*, Nature **401**, 55–57, (1999).
8. Y. Iiduka, T. Wakahara, T. Nakahodo, T. Tsuchiya, A. Sakuraba, Y. Maeda, T. Akasaka, K. Yoza, E. Horn, T. Kato, M. T. H. Liu, N. Mizorogi, K. Kobayashi, and S. Nagase, *Structural Determination of Metallofuilerene Sc_3C_{82} Revisited: A Surprising Finding*, J. Am. Chem. Soc. **127**, 12500–12501, (2005).
9. T.-S. Wang, N. Chen, J.-F. Xiang, B. Li, J.-Y. Wu, W. Xu, L. Jiang, K. Tan, C.-Y. Shu, X. Lu, and C.-R. Wang, *Russian-Doll-Type Metal Carbide Endofullerene: Synthesis, Isolation, and Characterization of $Sc_4C_2@C_{80}$*, J. Am. Chem. Soc. **131**, 16646–16647, (2009).
10. Y. Iiduka, T. Wakahara, K. Nakajima, T. Tsuchiya, T. Nakahodo, Y. Maeda, T. Akasaka, N. Mizorogi, and S. Nagase, *^{13}C NMR Spectroscopic Study of Scandium Dimetallofullerene, $Sc_2@C_{84}$ vs. $Sc_2C_2@C_{82}$*, Chem. Commun., 2057–2059, (2006).

11. S. Stevenson, M. A. Mackey, M. A. Stuart, J. P. Phillips, M. L. Easterling, C. J. Chancellor, M. M. Olmstead, and A. L. Balch, *A Distorted Tetrahedral Metal Oxide Cluster inside an Icosahedral Carbon Cage. Synthesis, Isolation, and Structural Characterization of* $Sc_4(\mu^3\text{-}O)_2@I_h\text{-}C_{80}$, J. Am. Chem. Soc. **130**, 11844–11845, (2008).
12. A. A. Popov, and L. Dunsch, *Charge Controlled Changes in the Cluster and Spin Dynamics of* $Sc_3N@C_{80}(CF_3)_2$*: The Flexible Spin Density Distribution and Its Impact on ESR Spectra*, Phys. Chem. Chem. Phys. **13**, 8977–8984, (2011).
13. A. A. Popov, and L. Dunsch, *Hindered Cluster Rotation and* ^{45}*Sc Hyperfine Splitting Constant in Distonoid Anion Radical* $Sc_3N@C_{80}$*, and Spatial Spin Charge Separation as a General Principle for Anions of Endohedral Fullerenes with Metal-Localized Lowest Unoccupied Molecular Orbitals*, J. Am. Chem. Soc. **130**, 17726–17742, (2008).
14. R. Valencia, A. Rodriguez-Fortea, A. Clotet, C. de Graaf, M. N. Chaur, L. Echegoyen, and J. M. Poblet, *Electronic Structure and Redox Properties of Metal Nitride Endohedral Fullerenes* $M_3N@C_{2n}$ *(M=Sc, Y, La, and Gd; 2n=80, 84, 88, 92, 96)*, Chem.–Eur. J. **15**, 10997–11009, (2009).
15. A. Reich, M. Panthofer, H. Modrow, U. Wedig, and M. Jansen, *The Structure of* $Ba@C_{74}$, J. Am. Chem. Soc. **126**, 14428–14434, (2004).
16. S. Taubert, M. Straka, T. O. Pennanen, D. Sundholm, and J. Vaara, *Dynamics and Magnetic Resonance Properties of* $Sc_3C_2@C_{80}$ *and Its Monoanion*, Phys. Chem. Chem. Phys. **10**, 7158–7168, (2008).
17. A. A. Popov, C. Chen, S. Yang, F. Lipps, and L. Dunsch, *Spin-Flow Vibrational Spectroscopy of Molecules with Flexible Spin Density: Electrochemistry, ESR, Cluster and Spin Dynamics, and Bonding in* $TiSc_2N@C_{80}$, ACS Nano **4**, 4857–4871, (2010).
18. W. Andreoni, and A. Curioni, *Freedom and Constraints of a Metal Atom Encapsulated in Fullerene Cages*, Phys. Rev. Lett. **77**, 834–837, (1996).
19. M. Krause, M. Hulman, H. Kuzmany, O. Dubay, G. Kresse, K. Vietze, G. Seifert, C. Wang, and H. Shinohara, *Fullerene Quantum Gyroscope*, Phys. Rev. Lett. **93**, 137403, (2004).
20. The CP2K Developers Group, 2010, `http://cp2k.berlios.de`.
21. J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach*, Comput. Phys. Commun. **167**, 103–128, (2005).
22. S. Goedecker, M. Teter, and J. Hutter, *Separable Dual-Space Gaussian Pseudopotentials*, Phys. Rev. B **54**, 1703–1710, (1996).
23. G. Lippert, J. Hutter, and M. Parrinello, *The Gaussian and Augmented-Plane-Wave Density Functional Method for Ab Initio Molecular Dynamics Simulation*, Theor. Chem. Acc. **103**, 124–140, (1993).
24. J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **77**, 3865–3868, (1996).
25. Y. Miyake, S. Suzuki, Y. Kojima, K. Kikuchi, K. Kobayashi, S. Nagase, M. Kainosho, Y. Achiba, Y. Maniwa, and K. Fisher, *Motion of Scandium Ions in* Sc_2C_{84} *Observed by* ^{45}*Sc Solution NMR*, J. Phys. Chem. **100**, 9579–9581, (1996).
26. H. Shinohara, H. Sato, M. Ohkohchi, Y. Ando, T. Kodama, T. Shida, T. Kato, and Y. Saito, *Encapsulation of a Scandium Trimer in* C_{82}, Nature **357**, 52–54, (1992).

Simulations of Electron Transfer and Electron Transport in Molecular Systems at Metal and Semiconductor Substrates

Pedro B. Coto[1], Óscar Rubio-Pons[1,2], Rainer Härtle[1], Michel Bockstedte[1], and Michael Thoss[1]

[1] Institut für Theo. Physik und Interdisziplinäres Zentrum für Molekulare Materialien (ICMM) Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany
E-mail: michael.thoss@physik.uni-erlangen.de

[2] Lehrstuhl für Theoretische Chemie, Technische Universität München, München, Germany

We have applied a methodology which combines first principles electronic structure methods and with quantum dynamical and transport approaches to study electron transfer and electron transport processes in different molecular systems adsorbed at metal and semiconductor substrates. In particular, we have investigated the molecular factors controlling the electron transfer in a model of a nitrile-substituted alkanethiolate self-assembled monolayer adsorbed at the Au(111) metal surface and the case of N,N'-diallyl-1,6,7,12-tetraphenoxyperylene-3,4:9,10-tetracarboxylic acid diimide dye adsorbed at the Si(111)-H hydrogen-terminated semiconductor surface. In addition, we have also investigated electron transport in a nanoscale molecular junction, where the molecular bridge exists in two tautomeric states having different conductance.

1 Introduction

Electron transfer (ET) processes in condensed phases are of fundamental importance in many areas of chemistry, biology, and physics[1–4]. Important examples include ET processes in biological systems, such as photosynthetic reaction centres of bacteria and plants[4] or heterogeneous ET reactions, such as injection of electrons from an excited state of a molecular chromophore into the conduction band of a metal (see Fig. 1) or semiconductor surfaces,[5–10], which is a key step in photonic energy conversion in dye-sensitized solar cells[11,12]. Moreover, heterogeneous ET is a fundamental process in the transport of electrons through single-molecule junctions (i.e. molecules which are chemically bound to metal electrodes) in the field of molecular electronics[13]. A detailed understanding of the basic mechanisms underlying the ET processes is essential to design and optimize more efficient advanced materials and molecular opto-electronic devices.

In this work, we discuss some of the results obtained in the theoretical modeling of ET processes in molecular systems at different substrates. Specifically, we have investigated ET processes that take place in a model of a nitrile-substituted alkanethiolate self-assembled monolayer (SAM) adsorbed at the Au(111) surface and in the N,N'-diallyl-1,6,7,12-tetraphenoxyperylene-3,4:9,10-tetracarboxylic acid diimide dye (PA-PTCDI) adsorbed at the Si(111)-H surface. As another example of heterogeneous ET processes, we have analyzed electron transport in a molecular junction where the molecular bridge can exist in two different isomeric forms with different conductance.

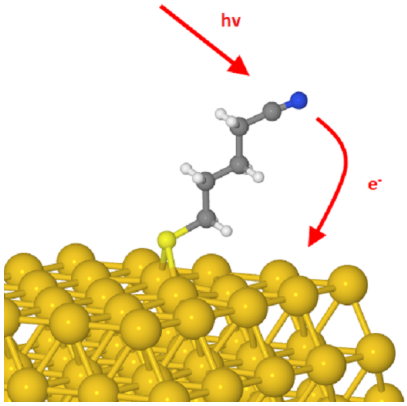


Figure 1. Schematic representation of the photoinduced ET process in the system CN$-(CH_2)_4-$S/Au(111). The molecule is adsorbed at the fcc-bridge like position.

2 Photoinduced Electron Transfer in Molecule-Metal and Molecule-Semiconductor Interfaces

In this section, we report the results obtained in the theoretical modelling of the photoinduced ET processes in molecules adsorbed at metal or semiconductor surfaces. After a brief outline of the theoretical protocol used, we discuss the results obtained in the simulation of the electron injection dynamics and the electronic characterization of the ET process in a model of a nitrile-substituted alkanethiolate SAM adsorbed at the Au(111) and in the case of the PA-PTCDI dye adsorbed at the Si(111)-H semiconductor surface.

2.1 Theoretical Methods

To study the heterogeneous ET dynamics in molecules adsorbed at metal or semiconductor substrates, we use a first principles based model motivated by the Anderson-Newns approach to the description of chemisorption[14]. Within this model the Hamiltonian is represented in a basis of localized diabatic electronic states relevant for the photoreaction. This basis contains the donor state of the ET process $|\psi_d\rangle$ (in the limit of vanishing coupling between chromophore and substrate this state corresponds to the product of an electronically excited state localized in the chromophore and an empty band of the metal or semiconductor substrate) and the (quasi)continuum of acceptor states of the ET reaction $|\psi_k\rangle$ (corresponding in the zero coupling limit to the product of the cationic state of the chromophore and an additional electron in the conduction band of the metal or semiconductor substrate). In this basis, the Hamiltonian for the electronic dynamics takes the form (we use units with $\hbar = 1$)

$$H = |\psi_d\rangle E_d \langle\psi_d| + \sum_k |\psi_k\rangle E_k \langle\psi_k| + \sum_k \left(|\psi_d\rangle V_{dk} \langle\psi_k| + |\psi_k\rangle V_{kd} \langle\psi_d|\right) \qquad (1)$$

Here, E_d and E_k are the energies of the electronic donor and acceptor states, respectively, while the non-diagonal elements, V_{dk}, characterize the donor-acceptor ET coupling. The parameters used in the ET Hamiltonian for the donor and acceptor energies and the coupling terms, see eq (1) have been calculated using a partitioning method. The details of this method have been discussed elsewhere[15] and here we will just mention the most important characteristics. The method is conceptually based on the projection-operator approach to resonant electron–molecule scattering[16,17]. In short, the donor and acceptor diabatic electronic states used in eq (1) are constructed following a procedure consisting in three steps: (i) a partitioning of the Hilbert space associated to the problem into donor and acceptor subspaces using a localized basis, (ii) a partitioning of the Hamiltonian according to the donor–acceptor separation, and (iii) a separate diagonalization of the donor and acceptor blocks of the Hamiltonian. In this work, we have used a mean-field single electron approach based on density functional theory (DFT) calculations for finite-size cluster models to compute the Hamiltonian and the molecular orbitals needed in the partitioning method.

To carry out the partitioning method, the first step is the selection of a localized basis set able to separate the donor and the acceptor spaces. As the atomic orbitals used as basis set in the DFT calculations, $|\phi_j\rangle$, are centred on the atoms, it is straightforward to split them into those centred on the donor moiety, $|\phi_j^d\rangle$, and those centred on the acceptor one, $|\phi_j^k\rangle$. Since working with orthogonal orbitals is advantageous[18], the $|\phi_j\rangle$ set is orthogonalized according to Löwdin's method of symmetric orthogonalization[19]. This method ensures that the transformed orthogonal orbitals $|\widetilde{\phi}_i\rangle$ are the closest in the least squares sense to the original non-orthogonal ones, thus preserving the localization of the basis and therefore maintaining the donor–acceptor separation.

In the following step, the Kohn–Sham matrix F obtained from the DFT calculation is transformed to the new orthogonal basis, $\widetilde{F}$, and rearranged in a donor–acceptor block structure.

The last step of the partitioning method is the separate diagonalization of the donor $\widetilde{F}_{dd}$ and acceptor $\widetilde{F}_{kk}$ blocks and the transformation of the off-diagonal $\widetilde{F}_{kd}$ terms to the new basis. This generates the matrix

$$\bar{F} = \left(\begin{array}{ccc|ccc} \varepsilon_{d,1} & 0 & \cdots & & & \\ 0 & \varepsilon_{d,2} & \cdots & & \bar{F}_{dk} & \\ \vdots & \vdots & \ddots & & & \\ \hline & & & \varepsilon_{k,1} & 0 & \cdots \\ & \bar{F}_{kd} & & 0 & \varepsilon_{k,2} & \cdots \\ & & & \vdots & \vdots & \ddots \end{array}\right) \tag{2}$$

The diagonal blocks of this matrix contain the energies of the donor (d) and acceptor (k) states whilst the off-diagonal blocks contain the electronic couplings between donor and acceptor sites. The eigenvectors of the diagonal blocks are the localized molecular orbitals of the donor ($|\bar{\phi}_d\rangle$) and acceptor $|\bar{\phi}_k\rangle$) moieties, respectively. In this manner, we identify these eigenvectors with the donor and acceptor states of the ET process[15].

To describe the effects of the infinite surface with the cluster model approach we use a simplified version of the surface Green's function technique[20] and approximate the self-energy matrix, which describes the effect of the infinite surface, by a constant diagonal imaginary part added to the boundary atoms of the cluster structure. The above mentioned

protocol provides the information needed for the calculation of the time evolution of the donor state population

$$P_d(t) = |\,\langle\bar{\phi}_d|\,e^{-i\hat{H}t}\,|\bar{\phi}_d\rangle\,|^2. \tag{3}$$

This quantity characterizes the photoinduced electron injection dynamics and can be compared with the experiment.

2.2 Photoinduced Electron Transfer in Molecule-Metal Interfaces

The method described above has been applied to the investigation of the photoinduced ET process between a photoexcited donor electronic state and a quasi-continuum set of acceptor electronic states in a model of a nitrile-substituted alkanethiolate SAM ($CN-(CH_2)_4-S$, **C4**) adsorbed at the Au(111) surface. In particular, we have considered the injection dynamics from the donor π_1^* and π_2^* states (see Fig. 2) corresponding to the two N[1s]π^* resonances of the CN chromophore that can be selectively monitored using core-hole clock spectroscopic techniques[21]. To characterize the single-molecule aspects of

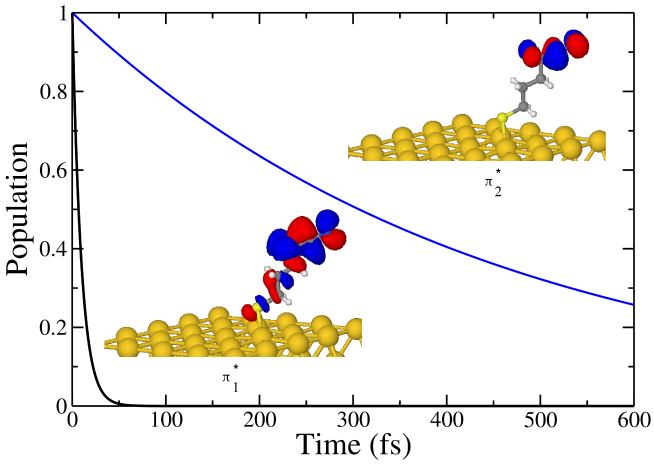


Figure 2. Time evolution of the population of the π_1^* (black) and π_2^* (blue) donor states of **C4**.

the ET process and to get a realistic description of the system we have determined the most stable adsorption geometry for the molecule at the Au(111) surface in the very dilute adsorption limit. To this end, we performed a slab model calculation consisting of five layers of Au atoms (comprising 125 atoms) and the **C4** molecule. The computations were carried out at the DFT level using periodic boundary conditions, the projected augmented wave method and the PBE correlation-exchange functional[22] as implemented in VASP[23]. The geometry optimization of the system carried out in JUROPA took 5 days using 32 nodes and 128 cores. We found that the most stable adsorption position at the Au(111) surface was the bridge-fcc like position with the molecule slightly tilted against the surface normal (see Fig. 1), in agreement with similar systems studied in previous works[24].

In order to keep the electronic structure calculations manageable from a computational point of view, the simulation of the electron injection dynamics was carried out for a cluster model representing the most stable adsorption geometry of the molecule at the Au(111) surface. The size of the cluster was the same as the one of the slab model employed in the periodic computations. The data needed to apply the partitioning method (the Kohn–Sham matrix and the molecular orbitals) and obtain the parameters for the model Hamiltonian describing the electron injection dynamics (the donor and acceptor energies and the coupling terms) were calculated at the PBE/def-SV(P) level[22,25]. To account for the effects of the infinite Au(111) surface, a constant diagonal imaginary part of 1 eV was added to the boundary atoms of the cluster model. All the calculations for the cluster model were carried out using TURBOMOLE[26].

The results obtained (see Fig. 2) show that the π_1^* and π_2^* states have different injection times, with π_1^* displaying an ultrafast ($\tau_{1/e} = 9.7$ fs) injection rate whilst π_2^* shows a significantly slower rate ($\tau_{1/e} = 442.0$ fs). This difference in electron injection times can be explained in terms of the different donor-acceptor interactions characterizing both states. While the π_1^* state shows non-negligible contributions at the aliphatic chain and the sulphur atom of the thiolate bond (see Fig. 2) the π_2^* state is essentially localized at the CN chromophore and as a consequence of the exponential decay of the ET coupling terms with the donor–acceptor distance[27] the π_2^* state is characterized by a slower electron injection time, in agreement with the experiment[21].

2.3 Photoinduced Electron Transfer in Molecule-Semiconductor Interfaces

As a second example for a photoinduced interfacial ET process we consider ET in dye-semiconductor interfaces. Elucidating the mechanisms underlying photoinduced ET processes in such systems is important for application in organic photovoltaics. Here, we report the results obtained for the ET process between the dye molecule PA-PTCDI covalently bonded to the Si(111)-H semiconductor surface. After optimizing the structure at the DFT level with the PBE correlation-exchange functional using VASP, we have analyzed the electronic properties and energy levels of the PA-PTCDI/Si(111)-H interface employing the HSE06 functional, which provides a more accurate description of the semiconductor band gaps and molecular HOMO-LUMO gaps. Fig. 3 depicts the energy levels of the isolated and the adsorbed molecule. The analysis of the energy levels shows that the resonance associated with the HOMO level of the PA-PTCDI is located 1.03 eV below the valence band (VB) edge and 5.94 eV below the vacuum level, which is in good agreement with the experimental photoemission values of 1.07 eV and 6.39 eV obtained for Si(111)-H[28]. The calculation also shows a highly accurate description of the Si band gap (1.16 eV compared to the 1.20 eV experimental value). On other hand, the LUMO level is slightly lower than the conduction band (CB) minimum. However, time-dependent DFT calculations of the excitation energy of the absorbing state of the isolated molecule (which is mainly characterized by a HOMO-LUMO like excitation) indicate that this state lies close to the Si CB (see Fig. 3). As the HOMO-LUMO gap hardly changes upon adsorption, the excitation energy of the adsorbed molecule may be estimated using the gas phase value[28]. This locates the photoexcited state of the adsorbed molecule, which is relevant for the photoinduced electron injection process, at -3.89 eV and thus close to the Si CB edge (see Fig. 3). Despite the fact that the calculated LUMO resonance is slightly off

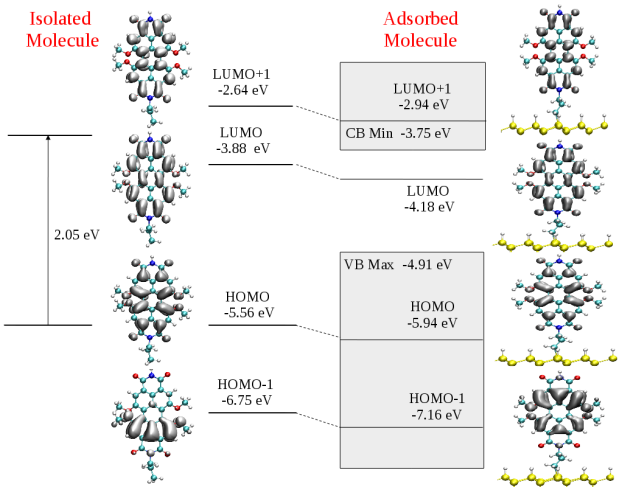


Figure 3. Calculated energy levels and frontier orbitals of PTCDI in the gas phase and chemisorbed at the Si(111)-H surface. The left panel shows the excitation energy for the isolated molecule. (Adapted from A. Decker *et al.* J. Phys. Chem. C., accepted, 2011).

the CB, the photoexcited state is expected to lie within the CB, which facilitates the charge transfer process from the molecule into the Si CB (see Fig. 3). The uncertainty of the LUMO position is related to the well known difficulty of DFT to accurately predict virtual level positions.

The VASP calculations were performed on parallel nodes with 64, 96 and 128 cores, depending of the input parameters. The VASP code uses the discrete Fourier Transform and subroutines from Fastest Fourier Transform in The West (FFTW) libraries[29] together with Message Passing Interface (MPI) libraries[30] to optimize the performance. For our systems, calculations with VASP at JUROPA show linear scaling up to 96 processors.

3 Electron Transport in Single-Molecule Junctions

In this section, we report the results obtained in the simulation of the electron transport in single molecule junctions. After a brief outline of the theory, we discuss the electron transport properties of a model for a molecular switch based on the enol-keto tautomerism.

3.1 Electron Transport Theory

To study electron transport in molecular junctions, we use the parametrized model Hamiltonian

$$H_{\rm el} = \sum_{i\in {\rm M}} \epsilon_i c_i^\dagger c_i + \sum_{k\in\{{\rm L},{\rm R}\}} \epsilon_k c_k^\dagger c_k + \sum_{k,i} (V_{ki} c_k^\dagger c_i + h.c.) \qquad (4)$$

that describes the coupling of the electronic states of the molecular bridge, with energies ϵ_i, to electronic states in the left (L) and right (R) leads, with energies ϵ_k, by interaction matrix elements V_{ki}.

To determine the parameters in the Hamiltonian we employ first principles electronic structure calculations based on a partitioning procedure similar to that described in Section 2.1. Thereby, the overall system is partitioned into a finite part that is treated explicitly in terms of electronic structure methods and an infinite part that is treated implicitly in terms of surface self-energies. The finite part comprises the central molecule and a left and a right cluster of gold atoms, representing the contacts of the molecule to the respective electrodes. A detailed description of the strategy used in the electronic structure calculations can be found in Ref. 34. Transport properties, in particular the current-voltage characteristic $I(V)$, are obtained employing a nonequilibrium Green's function (NEGF) method[31–34]. Within this approach, the current $I(V)$ can be expressed as:

$$I(V) = \frac{2e}{\hbar} \int \frac{\mathrm{d}E}{2\pi}\, tr\{\Sigma_{\mathrm{L}}^{<}\mathbf{G}^{>} - \Sigma_{\mathrm{L}}^{>}\mathbf{G}^{<}\}, \tag{5}$$

where $\mathbf{G}^{</>}(E)$ denotes the lesser/greater Green's function matrix associated with the electronic states of the molecular bridge and $\Sigma_{\mathrm{L}}^{</>}(E)$ is the corresponding lesser/greater self-energy matrix that reflects the coupling of the molecule to the left (L) lead.

3.2 Switching the Conductance Properties of Enol-Keto Tautomers

An important element for the design of molecular memory or logic devices is a molecular switch. A molecular junction may be used as a nanoswitch if the molecule can exist in two or more different conducting states that are sufficiently stable and can be reversibly converted into each other. Most mechanisms for optical switches considered so far are based on light-induced conformational changes, in particular isomerization[35,36] or ring opening[37] reactions of the molecular bridge. We have previously reported an alternative mechanism for optical switching of a molecular junction between different conductance states based on photoinduced excited state hydrogen transfer[38]. The basic idea of the mechanism is the use of a tautomeric reaction between two different conformations of the molecular bridge, the enol and the keto forms. Although both tautomers differ only in the position of a few hydrogen atoms, their conductance properties are very different. The keto form exhibits a rather low current, while the enol form shows a significant current and an almost Ohmic current-voltage characteristic, see Fig. 4. The very different conductance properties of the two tautomeric forms can be rationalized in terms of the valence bond structure. The enol tautomer has a fully conjugated π bond system which in general facilitates conduction through the molecule, whereas in the keto form the conjugation is broken at two sites (benzoquinone and amino moieties). Therefore, for a given voltage the two tautomers with different conductance can be used as the "on" (enol form) and the "off" (keto form) states of a molecular switch.

4 Conclusions

We have shown that a theoretical protocol based on a combination of first principles electronic structure calculations with quantum dynamics and electron transport approaches can be used for the modelling and rationalization of electron transfer and electron transport processes in molecular systems at metal and semiconductor substrates. The general

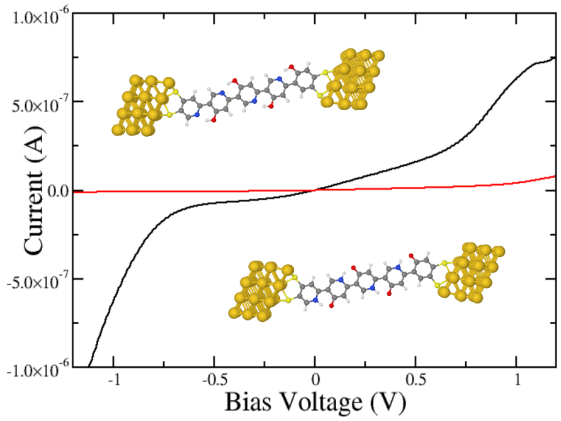


Figure 4. Current-voltage characteristics of the two tautomeric forms of a molecular junction. The black line shows the current for the enol form (upper structure), while the red line depicts the result for the keto form (lower structure).

applicability of the approach used in this work can be employed to elucidate the molecular factors underlying electron transfer processes which may pave the way for the design of more efficient optoelectronic devices.

Acknowledgements

We acknowledge the computer time provided by the Jülich Supercomputing Centre. We thank P. Feulner, A. Decker, T. Mayer, A. L. Sobolewski, S.-L. Suraru, F. Würthner, and W. Jägermann for helpful discussions. This work was supported by the BMBF within the project "OPV-Hybride" and the DFG within the cluster of excellence "Munich Center of Advanced Photonics" and a research grant.

References

1. A.M. Kuznetsov, J. Ulstrup, *Electron Transfer in Chemistry and Biology: An Introduction to the Theory (Wiley Series in Theoretical Chemistry)* (WILEY Seris in Theoretical Chemistry, John Wiley and Sons, New York, 1999).
2. P.F. Barbara, T.J. Meyer, M.A. Ratner, J. Phys. Chem. **100**, 13148 (1996).
3. J. Jortner, M. Bixon, *Electron Transfer: From Isolated Molecules to Biomolecules, Dynamics and Spectroscopy Adv. Chem. Phys. Vols. 106-107* (Wiley, New York, 1999).
4. M.E. Michel-Beyerle, *The Reaction Center of Photosynthetic Bacteria: Structure and Dynamics* (Springer, Berlin, 1996).

5. J. Schnadt, P.A. Brühwiler, L. Patthey, J.N. O'Shea, S. Södergreen, M. Odellus, R. Ahuja, O. Karis, M. Bässler, P. Persson, H. Siegbahn, S. Lunell, N. Martenson, Nature **418**, 620 (2002).
6. R. Huber, J.E. Moser, M. Grätzel, J. Wachtveitl, Chem. Phys. Lett. **285**, 39 (2002).
7. A. Fölisch, P. Feulner, A. Fink, D. Menzel, D. Sanchez-Portal, P. Echenique, W.Wurth, Nature (London) **436**, 373 (2005).
8. R. Ernstorfer, L. Gundlach, S. Felber, W. Storck, R. Eichberger, F. Willig, J. Phys. Chem. B **110**, 25383 (2006).
9. K. Tominaga, D.A.V. Kliner, A.E. Johnson, N.E. Levinger, P.F. Barbara, J. Chem. Phys. **98**, 1228 (1993).
10. C. Wang, B.K. Mohney, R.D. Williams, V. Petrov, J.T. Hupp, G.C. Walker, J. Am. Chem. Soc. **120**, 5848 (1998).
11. M. Grätzel, Nature **414**, 338 (2001).
12. A. Hagfeldt, M. Grätzel, Chem. Rev. **95**, 49 (1995).
13. J. Cuevas, E. Scheer, *Molecular Electronics An introduction to Theory and Experiment* (World Scientif Co. Pte. Ltd., Singapore, 2010).
14. Y.G. Boroda, A. Calhoun, G.A. Voth, J. Chem. Phys. **107**, 8940 (1997).
15. I. Kondov, M. Cizek, C. Benesch, H. Wang, M. Thoss, J. Phys. Chem. C **111**, 11970 (2007).
16. H. Feshbach, Ann. Phys. **19**, 287 (1962).
17. W. Domcke, Phys. Rep. **208**, 97 (1991).
18. L. Kurnikov, D. Beratan, J. Chem. Phys. **105**, 9561 (1996).
19. P.O. Löwdin, J. Chem. Phys. **18**, 365 (1950).
20. Y. Xue, M. Ratner, Phys. Rev. B **68**, 115406 (2001).
21. M. Zharnikov, private communication (2011).
22. J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
23. *Vienna Ab-intio Simulation Package (VASP), V5.2.11*, `http://cms.mpi.univie.ac.at/CMSPage/main/`, (2006).
24. M. Vargas, P.Giannozzi, A. Selloni, G. Scoles, J. Phys. Chem. B **105**, 9509 (2001).
25. R.A. F. Weigend, Phys. Chem. Chem. Phys. **7**, 3297 (2005).
26. M.V. Arnim, R. Ahlrichs, J. Comput. Chem. **19**, 1746 (1998).
27. M. Paddon-Row, Adv. Phys. Org. Chem. **38**, 1 (2011).
28. A. Decker, S. Suraru, O. Rubio-Pons, E. M., M. Bockstedte, M. Thoss, F. Wüthner, T. Mayer, W. Jaegermann, J. Phys. Chem. C (Accepted (2011)).
29. `http://www.fftw.org/`.
30. `http://www.mpi forum.org/`.
31. M. Galperin, M. Ratner, A. Nitzan, Phys. Rev. B **73**, 045314 (2006).
32. R. Härtle, C. Benesch, M. Thoss, Phys. Rev. Lett. **102**, 146801 (2009).
33. R. Härtle, C. Benesch, M. Thoss, Phys. Rev. B **77**, 205314 (2008).
34. C. Benesch, M. Cizek, M. Thoss, W. Domcke, J. Phys. Chem. C **112**, 9880 (2008).
35. B.Y. Choi, S.J. Kahng, S. Kim, H. Kim, H. Kim, Y. Song, J. Ihm, Y. Kuk, Phys. Rev. Lett. **96**, 156106 (2006).
36. M. Del Valle, R. Gutierrez, C. Tejedor, G., Cuniberti, Nature Nanotechnology **2**, 176 (2007).
37. J. Li, G. Speyer, O. Sankey, Phys. Rev. Lett. **93**, 248302 (2004).
38. C. Benesch, M. Rode, M. Cizek, R. Härtle, O. Rubio-Pons, M. Thoss, A. Sobolewski, J. Phys. Chem. C **113**, 10315 (2009).

Elementary Particle Physics: With Supercomputers Towards New Destinations

Gernot Münster

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

One of the aims of elementary particle physics is to gain an understanding of the fundamental constituents of matter and their mutual interactions. Recent decades have brought enormous progress in this direction. Today we have a Standard Model of Elementary Particle Physics, which is able to describe a wealth of phenomena in the submicroscopic world. The particles, which form the building blocks of matter in the Standard Model, are quarks and leptons. The forces between them are of three types: the electromagnetic, the weak and the strong interactions. The Standard Model very successfully describes many properties and reactions of elementary particles, observed in colliders and other experiments.

The theory of elementary particle physics is, however, not in a completely satisfactory state. One the one hand, there are strong indications for physics beyond the Standard Model. The existence of dark matter and the asymmetry between matter and antimatter in the universe are two prominent examples. One approach towards new physics is to look for deviations from Standard Model predictions in precision experiments. For example, the so-called CP symmetry relates matter to antimatter. The Standard Model contains a source for small violations of this symmetry. Such violations are observed in the decays of K^0 mesons. To check whether the experimental results are consistent with the Standard Model is, however, a very difficult task. The reason for this is the fact that CP violation is tied to the quarks, which are strongly bound and do not exist as free particles. To investigate the amount of CP violation theoretically with sufficient precision requires numerical simulations with a very large amount of computing time. This research is the content of the project described by Christian Hoelbling in this volume. For the research of his collaboration, using supercomputer time provided by NIC, Hoelbling's project was awarded the predicate "John von Neumann Excellence Project 2011".

Another area of extensive activities in the theory of elementary particles concerns the non-perturbative properties of strongly interacting systems. One of the characteristic features of the theory of the strong interactions, Quantum Chromodynamics (QCD), is the fact that the strength of interactions increases with decreasing energy. In the low-energy regime the coupling gets so strong that analytical methods finally fail and non-perturbative methods are required. This applies for example to the calculation of the masses of particles. The method of choice for the study of non-perturbative problems of QCD are numerical simulations on supercomputers. For this purpose a volume of space-time is discretised on a finite four-dimensional lattice and the variables representing the degrees of freedom are associated to the points and links of the lattice. In this way calculations of physically important observables from first principles become possible.

The main sources of systematic errors, which have to be controlled in these calculations, are: (1) most importantly the finite lattice spacing a, which must be small enough to be near the continuum physics, (2) the size L of the lattice, which must be large enough to comprise an adequate part of the system under study, (3) the masses of the light up- and down-quarks, which must be small enough to represent their values in nature, and (4) the sampling of configurations. To meet these conditions requires large lattices and substantial amounts of computer time. In order to achieve the highest possible efficiencies in these calculations, much effort has been invested in the development of algorithms and the tuning of the corresponding codes for lattice QCD. As a result, they scale extremely well up to high numbers of processors, as can e. g. be seen from the performance plot in the article of Fodor *et al.*.

QCD predicts that nuclear matter at extreme conditions undergoes a phase change to a plasma of quarks and gluons. This is the subject of current experiments at e. g. the LHC and FAIR experiments. On the theoretical side the nature of the phase change and the properties of the quark-gluon-plasma are problems of a non-perturbative nature. Different collaborations investigate these in the framework of lattice QCD. In this volume, the articles by Fodor, Szabo *et al.* and by Philipsen, Wittig *et al.* present work of two internationally leading groups in this field.

From the point of view of QCD, atomic nuclei are highly complex systems, which are not yet accessible to *ab initio* calculations. Therefore the group of Ulf Meißner took an approach which combines approximate analytical and numerical methods in a clever way, as explained in their article. The nucleons and the forces between them are described by an effective theory, which originates from QCD. The properties and reactions of nuclei, on the other hand, are then studied by means of numerical simulations on a supercomputer. In this way Meißner and collaborators have been able to identify the long predicted Hoyle state of carbon ^{12}C, which plays a central role in the production of carbon in the early universe, providing the precursor for the later formation of life on earth.

Searching for New Physics in Neutral Kaon Mixing

Christian Hoelbling

Department of Physics, Wuppertal University, Gaussstr. 20, D-42119 Wuppertal, Germany
E-mail: hch@physik.uni-wuppertal.de

1 Introduction

The standard model of elementary particle physics has been enormously successful in describing a large range of high energy physics phenomena. Although it is seen as a conceptually unsatisfactory effective theory and not as a valid candidate for a fundamental theory, experiments so far have failed to come up with any firm evidence against its validity (the only exception being the existence of neutrino masses which are however a trivial extension).

In the search for physics beyond the Standard Model there are two principal avenues being pursued: On the one hand there are direct detection experiments, currently most prominently pursued at CERN's LHC. On the other hand, one can look for effects of physics beyond the Standard Model at the so-called "precision frontier". Processes that are extremely well known or very rare within the Standard Model may reveal small contributions from new physics if one is able to both precisely measure them as well as predict the rate at which they occur within the Standard Model.

While there is a large number of standard model predictions for which no precise measurement exists, there are on the other hand some accurate measurements for which there is no firm standard model prediction. Typically these are weak interaction processes with hadronic initial and final states. While it is easy to account for the weak interaction between quarks, quarks are not free particles and are therefore never observed as final states. The strong interaction, described by Quantum Chromodynamics (QCD), is responsible for permanently confining quarks into hadrons (such as e.g. the proton). Due to the strong nature of that force, it is not possible to treat it as a small correction to the quark level description. One rather has to treat QCD in a nonperturbative fashion in order to understand how the quark level weak interaction manifests itself in hadronic processes.

One specifically prominent example is the CP violation parameter ϵ in the mixing of neutral K^0 mesons which is experimentally measured to better than a percent precision[1]. CP is the matter-antimatter symmetry and its violation plays a crucial role in the understanding why our universe is composed of matter and almost free of antimatter.

Neutral K^0 mesons are not the only place where CP violation has been detected (for an overview of experimental results see e.g. the CKMfitter review[2]). In the Standard Model there is however only one mechanism for CP violation in the quark sector, hence all experimentally observed CP violation needs to be explained by one fundamental quark-level parameter only. It is therefore a nontrivial consistency check of the Standard model to see if all experimental observations are indeed described by this one parameter. If they are, one can set stringent limits on new physics contributions towards the experimentally

observed results. If on the other hand they are not, one could gain some hints on new physics outside the known Standard Model.

It is the aim of the current project to check whether the experimentally observed CP violation parameter ϵ in the mixing of neutral K^0 mesons is compatible with standard model expectations and to explore possible consequences for conjectured physics beyond the Standard Model.

2 Lattice QCD

Our current understanding of the strong nuclear force is based on the theory of Quantum Chromodynamics (QCD). QCD may be viewed as a straightforward generalization of Quantum Electrodynamics (QED) where the electromagnetic gauge symmetry group $U(1)$ is replaced by the "colour group" $SU(3)$. After a Wick rotation to Euclidean space, QCD is described by the action

$$S = \int d^4x \bar{\psi}(x) \left(\gamma_\mu D_\mu + \mathrm{m}\right) \psi(x) + \frac{1}{4} F(x)_{\mu\nu} F(x)_{\mu\nu} \tag{1}$$

with the field strength tensor $F_{\mu\nu}$ and the covariant derivative D_μ given by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu] \qquad D_\mu = \partial_\mu + igA_\mu \tag{2}$$

One can regulate QCD by introducing a finite space-time lattice with a lattice spacing a by putting fermion fields on the lattice points and gauge transporters $U_\mu(x) = e^{igA_\mu(x)}$ on the links between them. Kaons may be introduced by putting a pseudoscalar source $P = \bar{\psi}_d \gamma_5 \psi_s$ at a timeslice t_i and going far enough in Euclidean time t' to have all excited modes die out exponentially. One may then put an effective weak operator at t'

$$O_1 = \left(\bar{\psi}_s \gamma_\mu (1-\gamma_5) \psi_d\right) \left(\bar{\psi}_s \gamma_\mu (1-\gamma_5) \psi_d\right) \tag{3}$$

that describes the CP violating transition of a $\bar{K}^0$ into its antiparticle K^0 and eventually, after another sufficient time separation at t_f a pseudoscalar sink P that can annihilate the antiparticle. Numerically performing a path integral over all gauge and fermion fields in this setup, one can extract the desired transition matrix element of neutral K^0 mixing and thus relate the experimentally observed hadronic transition amplitudes to the fundamental quark level parameter. It is customary to give the matrix element in form of the ratio B_K that can be computed as

$$B_K = \frac{\langle \bar{P}(t_f) | O_1(t') | P(t_i) \rangle}{\frac{8}{3} \langle \bar{P}(t_f) | A_\mu(t') \rangle \langle \bar{A}_\mu(t') | P(t_i) \rangle} \tag{4}$$

with $A_\mu = \bar{\psi}_d \gamma_\mu \gamma_5 \psi_s$ and for sufficiently large time separations between t_i resp. t_f and t'.

For a detailed discussion of general lattice techniques see e.g. the standard textbook by Montvay and Münster[3].

3 Technical Details

One reason why CP violating neutral K^0 mixing is a promising candidate to find new physics is the fact that the standard model operator (3) is chirally suppressed compared to

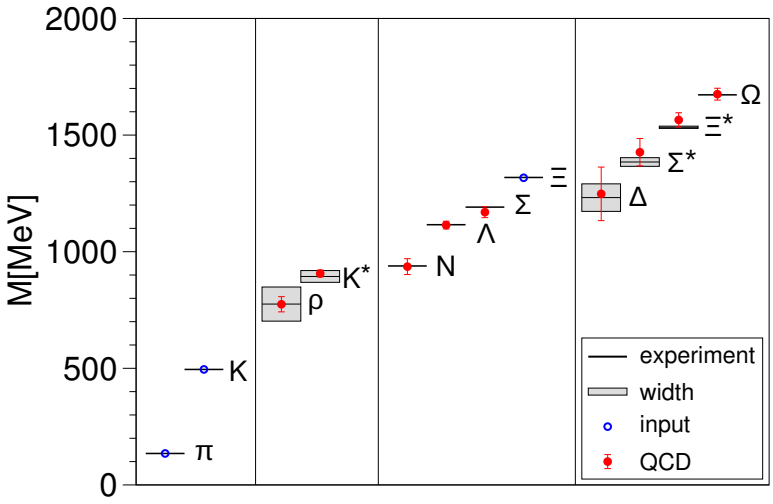


Figure 1. The light hadron spectrum[9].

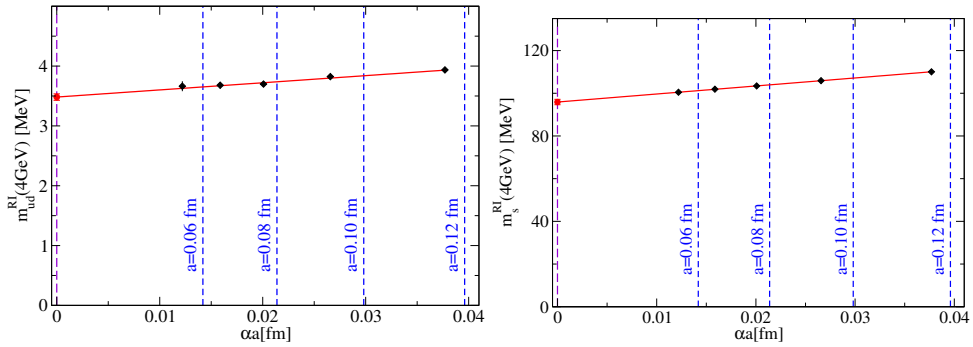


Figure 2. Continuum limit of the average up-down and the strange quark masses[10].

other operators $O_2, \ldots, O_5$ that generically might appear in Standard Model extensions[a]. The mixing of these other operators with the Standard Model one is forbidden by chiral symmetry. The downside of this feature is the need for a lattice discretization that has good chiral symmetry properties since otherwise the chirally dominant operators $O_2, \ldots, O_5$ give huge additive renormalizations that need to be subtracted properly.

For this reason, recent attempts at a precision calculation of B_K were restricted to fermion formulations that preserve some remnant of chiral symmetry on the lattice.[5–8] These formulations however tend to have other problems as they either do break the flavour symmetry or are computationally expensive.

In an alternative approach, one can start with Wilson-type fermions that explicitly break chiral symmetry but retain full flavour symmetry. By a combination of suppression of un-

[a]The exact definition of these operators can be found in the literature[4]

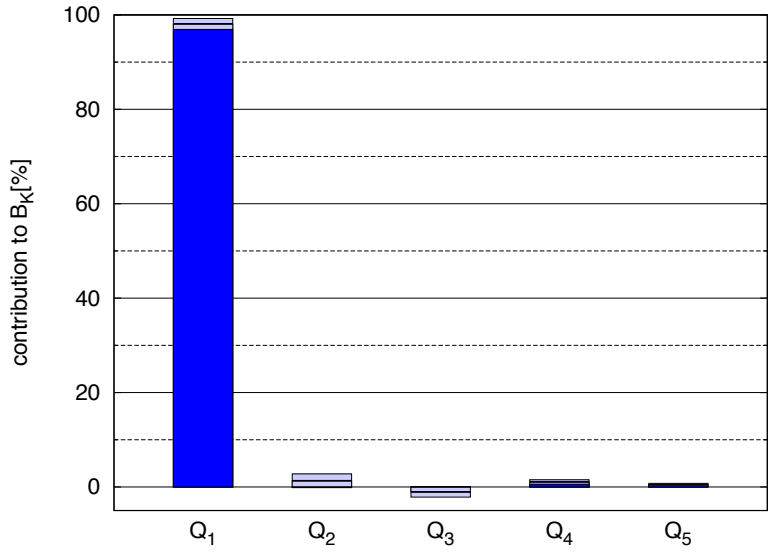


Figure 3. Total contribution of individual operator components $Q_1, \Delta_{1i}Q_i$ to B_K in % for $a{\approx}0.077$ fm and $M_\pi{\sim}120$ MeV with their errors. One can see that even at the physical point the contribution from the continuum allowed operator Q_1 operators Q_2 to Q_5 are small or even compatible with zero.

physical UV modes and classical improvement of the fermion operator one can show[11] that approximate chiral symmetry can be obtained with moderate computational cost. In prior efforts[9,10] it was demonstrated that with actions of this type one could go to very small quark masses or even to the physical point[12] and obtain the ground state light hadron spectrum (see Fig. 1) and the light quark masses (see Fig. 2) with fully controlled systematic errors.

These results suggested that with this type of action the chiral symmetry violations might be small enough to enable a precision computation of B_K. As detailed in a recent paper[13] the operator mixing is indeed small. Fig. 3 shows the contributions of the individual operators towards the final value of B_K on the ensemble with the smallest quark mass, which is below the physical value. It can be seen that even in this case the dominant contribution comes from the Standard Model operator O_1.

We were thus able to extract the renormalized value of B_K on a number of ensembles with five different lattice spacings and many different pion and kaon masses. In Fig. 4 an overview of the ensembles used is given. One can see that lattices were chosen sufficiently large so that finite volume corrections are small (they were corrected for nonetheless).

In order to obtain a physical result, we interpolated our results to the physical point (see Fig. 5) and at the same time extrapolated to the continuum. While it turned out that the coarsest of our five lattice spacings was not fine enough, the other four were in perfect agreement which allowed us to perform a precise extrapolation to the continuum theory (see Fig. 6).

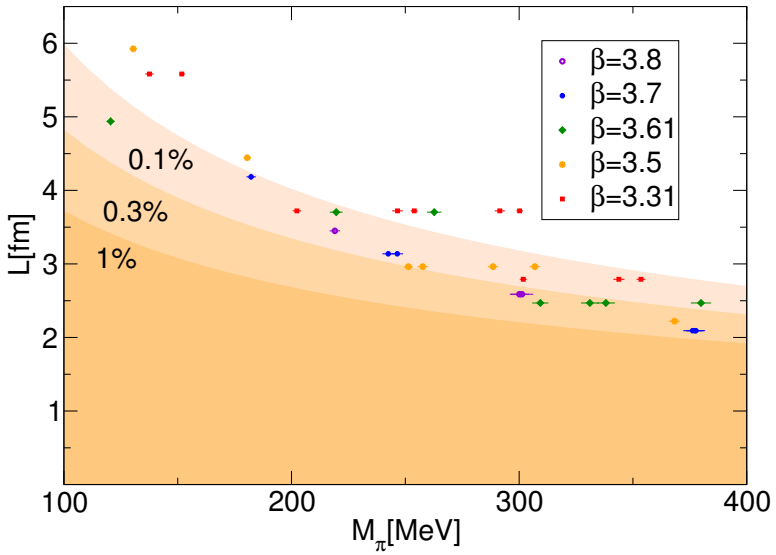


Figure 4. Pion mass vs. lattice extent of the ensembles used in the present calculation. The shaded areas correspond to regions where finite volume effects on the pion mass exceed 1%, 0.3% and 0.1% according to Ref. 14.

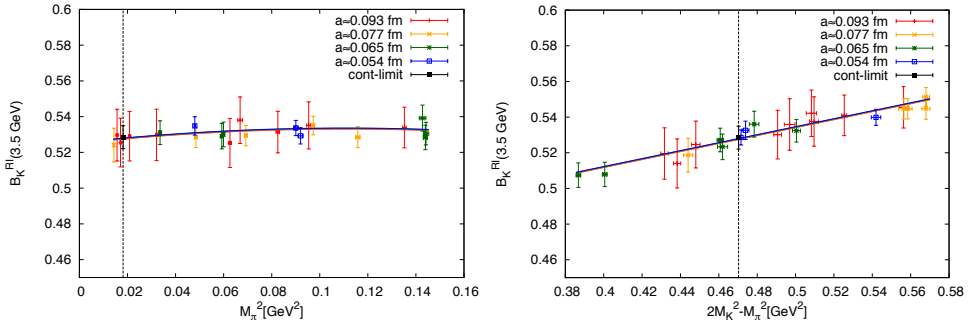


Figure 5. Combined continuum extrapolation and interpolation to physical quark mass for one of our 5760 fits. The filled square represents our final result, the dashed vertical line the corresponding physical mass. The interpolation in the light quark mass (left) is mild while the slope of the interpolation in m_s (right) is somewhat steeper. Both interpolations as well as the continuum extrapolation are fully under control.

4 Results

In order to quote a final result, we performed our analysis 5760 times, combining all reasonable variations of every piece of our analysis procedure. The central value of all these analyses represents our final number while the spread gives the systematic error. Statistical errors are computed in a standard way by 2000 bootstrap samples.

Fig. 7 displays the results of this error analysis procedure. As one can see, the systematic error is clearly subdominant. Converting our lattice results into continuum renormal-

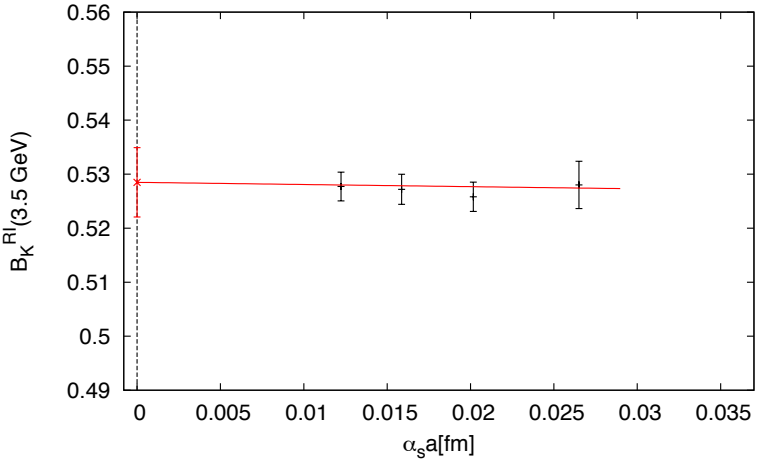


Figure 6. Continuum extrapolation of B_K as obtained from the fit in Fig. 5.

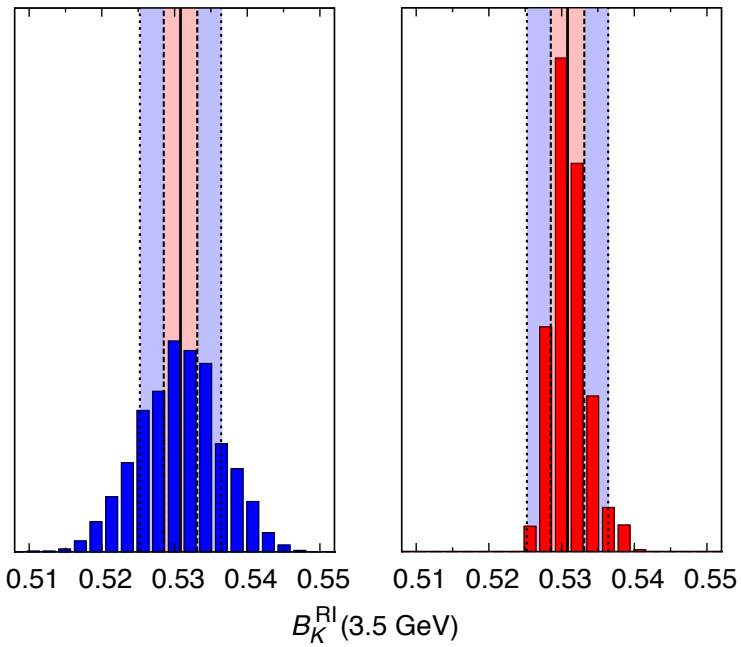


Figure 7. Histogram of the statistical (left) and systematic (right) uncertainties of our result. The solid vertical line denotes our final value and the light and dark bands our statistical and systematic errors.

ization frameworks we finally obtain

$$B_K^{\overline{\text{MS}}\text{-NDR}}(2\,\text{GeV}) = 0.5644(59)_{\text{stat}}(25)_{\text{sys}}(56)_{\text{PT}} \tag{5}$$

$$\hat{B}_K = 0.7727(81)_{\text{stat}}(34)_{\text{sys}}(77)_{\text{PT}} \tag{6}$$

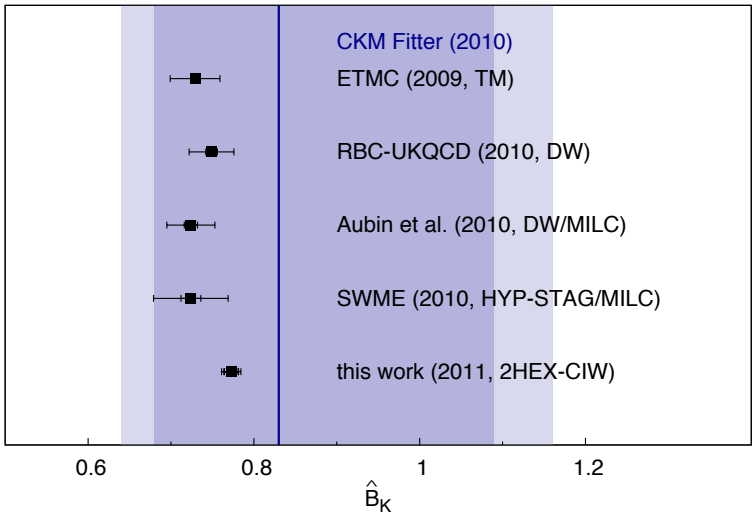


Figure 8. Comparison of our result (6) with those of other lattice groups (ETMC[8], RBC-UKQCD[5], Aubin et. al.[6], SWME[7]) and with the Standard Model prediction by the CKMfitter group[2].

where the errors are statistical, systematical and finally perturbative conversion errors.

In Fig. 8 we compare our result with other recent results in the literature and with Standard Model expectations according to the CKMfitter group. As one can see, our result is nicely in agreement with the Standard Model value. The large uncertainty on the standard model value is partly due to the very conservative approach of the CKMfitter group that adds systematic errors linearly. While the UTfit group cites smaller errors[15], there is no tension between our result and their slightly more aggressive Standard Model expectations either.

5 Conclusion

The lattice determination of B_K has been pushed to percent level precision. This precision is tiny compared to other uncertainties in the determination of CP violation in neutral K^0 meson mixing within the Standard Model. One can therefore hope that this precise lattice result will trigger new efforts to determine the other components to a comparable accuracy, which would allow a percent level consistency test of the Standard Model in this promising channel.

Acknowledgements

I wish to thank all my colleagues in the Budapest-Marseille-Wuppertal collaboration with whom this work has been performed. This work is supported in part by DFG grant SFB-TR 55. The calculations were carried out at the Jülich Supercomputing Centre.

References

1. K. Nakamura, et al., Review of particle physics, J. Phys. G37 (2010) 075021.
2. J. Charles, et al., CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories, Eur. Phys. J. C41 (2005) 1–131.
3. I. Montvay, G. Munster, Quantum fields on a latticeCambridge, UK: Univ. Pr. (1994) 491 p. (Cambridge monographs on mathematical physics).
4. R. Babich, N. Garron, C. Hoelbling, J. Howard, L. Lellouch, et al., K0 - anti-0 mixing beyond the standard model and CP-violating electroweak penguins in quenched QCD with exact chiral symmetry, Phys.Rev. D74 (2006) 073009.
5. C. Aubin, J. Laiho, R. S. Van de Water, The neutral kaon mixing parameter B_K from unquenched mixed-action lattice QCD, Phys. Rev. D81 (2010) 014507.
6. Y. Aoki, et al., Continuum Limit of B_K from 2+1 Flavor Domain Wall QCD, Phys. Rev. D84 (2011) 014503.
7. T. Bae, et al., B_K using HYP-smeared staggered fermions in $N_f = 2+1$ unquenched QCD, Phys. Rev. D82 (2010) 114509.
8. M. Constantinou, et al., BK-parameter from Nf = 2 twisted mass lattice QCD, Phys. Rev. D83 (2011) 014505.
9. S. Durr, et al., Ab-Initio Determination of Light Hadron Masses, Science 322 (2008) 1224–1227.
10. S. Durr, et al., Lattice QCD at the physical point: light quark masses, Phys. Lett. B701 (2011) 265–268.
11. S. Capitani, S. Durr, C. Hoelbling, Rationale for UV-filtered clover fermions, JHEP 11 (2006) 028.
12. S. Durr, et al., Lattice QCD at the physical point: Simulation and analysis details, JHEP 08 (2011) 148.
13. S. Durr, Z. Fodor, C. Hoelbling, S. Katz, S. Krieg, et al., Precision computation of the kaon bag parameter, arXiv:1106.3230 [hep-lat].
14. G. Colangelo, S. Durr, C. Haefeli, Finite volume effects for meson masses and decay constants, Nucl. Phys. B721 (2005) 136–174.
15. M. Bona, et al., The unitarity triangle fit in the standard model and hadronic parameters from lattice QCD: A reappraisal after the measurements of Delta(m(s)) and BR(B $\rightarrow$ tau nu/tau), JHEP 10 (2006) 081.

Nuclear Physics from Lattice Simulations

Ulf-G. Meißner[1,2]
representing the NLEFT collaboration

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

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

I give a brief introduction into the simulation of atomic nuclei based on forces derived from chiral effective field theory. As an example, I discuss the recent *ab initio* calculation of the Hoyle state in the spectrum of ^{12}C. Future lines of research are briefly discussed.

1 Introduction

Nuclear physics has entered a new and exciting era as *ab initio* calculations[a] of nuclei based on forces derived from the chiral effective Lagrangian of Quantum Chromodynamics (QCD) become possible. For few-nucleon system, the chiral effective field theory (EFT) for the forces between two, three and four nucleons has been worked out up to next-to-next-to-next-to-leading order (N^{3}LO) in the chiral power counting. As stressed by Weinberg, this power counting applies to the nuclear potential[1]. This potential consists of long-range pion exchange(s) and shorter ranged multi-nucleon contact interactions. Bound and scattering states are calculated based on exact solutions of the Lippmann-Schwinger or Faddeev-Yakubowsky equations. This method has passed many tests as reviewed in Ref. 2. There are two different venues to tackle the nuclear many-body problem, that is nuclei with atomic number $A \geq 5$. Either one utilizes the forces from EFT within a conventional, well established many-body technique (no-core-shell-model, coupled cluster approach, etc.) or one develops a novel scheme that combines these forces with Monte Carlo methods that are so successfully used in lattice QCD. This novel scheme is termed "nuclear lattice simulations" and has recently enjoyed wide recognition in the popular press as the first ever *ab initio* calculation of the Hoyle state in ^{12}C has been performed, see Sec. 5. In the following, I will give a short introduction into this novel nuclear many-body technique. The foundations of the method and its early applications are reviewed in Ref. 4.

2 Nuclear Lattice Simulations: General Remarks

Space-time is discretized in Euclidean time on a torus of volume $L_s \times L_s \times L_s \times L_t$, with $L_s(L_t)$ the side length in spatial (temporal) direction. The minimal distance on the lattice, the so-called lattice spacing, is a (a_t) in space (time). This entails a maximum momentum

[a] An *ab initio* calculation is defined as follows: First, one determines the parameters that appear in the nuclear force *entirely* from a fit to two-, three and four-nucleon observables. Then, the so-determined forces are used without modification in an exact solution of the A-body problem.

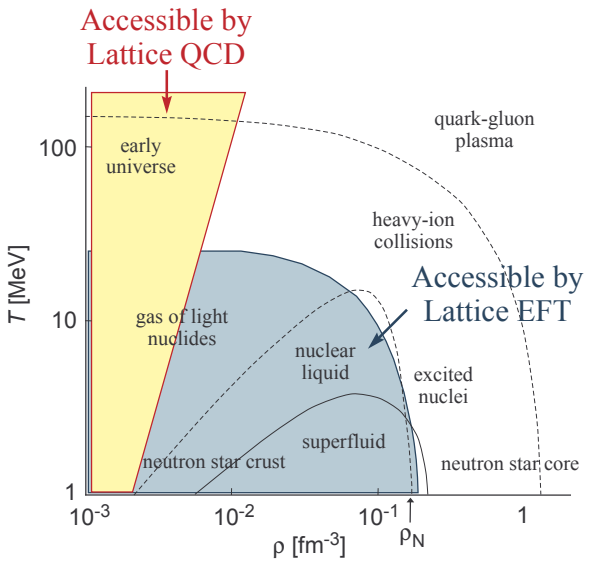


Figure 1. Nuclear phase diagram as accessible by lattice QCD and by nuclear lattice EFT. Figure courtesy of Dean Lee.

on the lattice, $p_{\max} = \pi/a$, which serves as an UV regulator of the theory. The nucleons are point-like particles residing on the lattice sites, whereas the nuclear interactions (pion exchanges and contact terms) are represented as insertions on the nucleon world lines using standard auxiliary field representations. The nuclear forces have an approximate spin-isospin SU(4) symmetry (Wigner symmetry)[5] that is of fundamental importance in suppressing the malicious sign oscillations that plague any Monte Carlo (MC) simulation of strongly interacting fermion systems at finite density. For this reason, nuclear lattice simulations allow access to a large part of the phase diagram of QCD, see Fig. 1, whereas calculations using lattice QCD are limited to finite temperatures and small densities (baryon chemical potential). In this contribution, I will concentrate on the calculation of the ground state properties and excited states of atomic nuclei with $A \leq 16$.

3 Simulation Method - A Short Glimpse

We simulate the interactions of nucleons using the MC transfer projection method. Each nucleon evolves as a single particle in a fluctuating background of pion and auxiliary fields, the latter representing the multi-nucleon contact interactions. We also perform Gaussian smearing of the LO contact interactions which is required by the too strong binding of four nucleons on one lattice site. To leading order, we start with a Slater determinant of single-nucleon standing waves in a periodic cube for Z protons and N neutrons (with $Z + N = A$). We use the SU(4) symmetric approximation of the LO interaction as an approximate inexpensive filter for the first t_0 time steps – this suppresses dramatically the

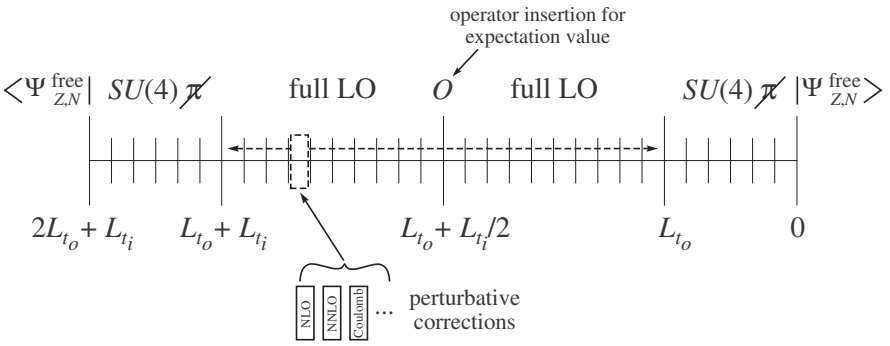


Figure 2. Schematic diagram for the transfer matrix calculation. For details, see the text.

sign oscillations. Then we switch on the full LO interaction and calculate the ground state energy and other properties from the correlation function $Z(t) = \langle\Psi_A|\exp(-tH)|\Psi_A\rangle$, letting the Euclidean time t go to infinity. Higher order contributions, the Coulomb repulsion between protons and other isospin-breaking effects (due to the light quark mass difference) are computed as perturbative corrections to the LO transfer matrix. This is symbolically depicted in Fig. 2. Excited states are calculated from a multi-channel projection MC method. As a first step, we use various improvements in our LO lattice action, including $\mathcal{O}(a^4)$ improvements for the nucleon kinetic energy and the Gaussian smearing factors of the contact interactions. Moreover, all lattice operators at $\mathcal{O}(Q^3)$ are included, in particular also the ones related to the breaking of rotational symmetry. Their strengths can be tuned to eliminate unphysical partial wave mixing like e.g. between the ${}^3S_1 - {}^3D_1$ and the ${}^3D_3$ partial waves. We then use a set of Slater determinants of 24 different single standing nucleon waves. For the calculation of the Hoyle state, we have constructed from these three states with total momentum zero and $J_z = 0$ mod 4 and one state with $J_z = 2$ mod 4. We are presently enlarging the number of basis states to avoid the orientation bias of the initial states encountered in the simulations performed so far.

4 Computational Scaling and Performance of the Simulations

The very low memory and trivially parallel structure of our lattice MC codes allows jobs to scale ideally with several thousand processors. In the left panel of Fig. 3 we show the computational time for each processor on JUGENE/P to generate one hybrid Monte Carlo (HMC) trajectory. The performance is entirely independent of the number of processors. The computational time scales with the number of nucleons A as $A^{1.7}$ at fixed volume V and with $V^{1.5}$ for fixed A. The average sign – which is a measure of the severity of the sign oscillations – scales approximately as $\exp(-0.11A)$. The right panel of Fig. 3 shows the statistical error for the LO energy calculation of ^{12}C and ^{16}O versus the number of HMC trajectories for a periodic cube with length 13.8 fm. Each data set shows one-half rack-day running on JUGENE. The scaling for the ^{12}C (^{16}O) statistical error indicates that a warm-up period of 500 (800) trajectories is sufficient to collect data. After this initial

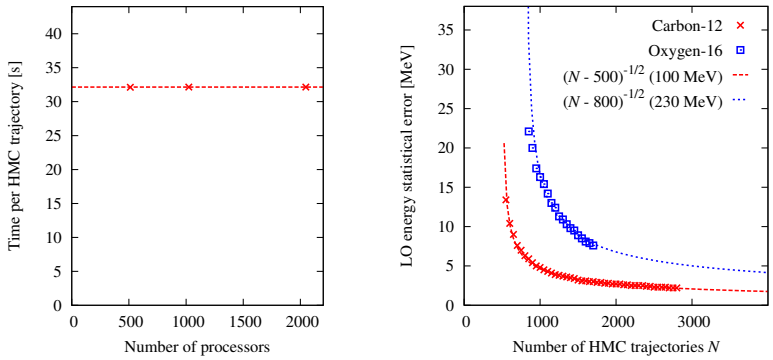


Figure 3. Left: Computational time for each JUGENE processor to generate one HMC trajectory versus the number of parallel processors. Right: Statistical errors in the LO energy calculation versus the number of HMC trajectories for ^{12}C and ^{16}O.

warm-up independent configurations are generated at a rate of at least one per 50 HMC trajectories.

5 Recent Results

So far, we have performed calculations at next-to-next-to-leading order in the chiral expansion of the nuclear potential. In the two-nucleon system, we have nine parameters that are determined from a fit to the S- and P-waves in np scattering. Two further isospin-breaking parameters are determined from the pp and nn scattering lengths. The three-nucleon force features only two low-energy constants, that can be determined from the triton binding energy and low-energy neutron-deuteron scattering in the doublet channel[8]. The first non-trivial predictions are the energy dependence of the pp $^1S_0$ partial wave, which agrees with the Nijmegen partial wave analysis up to momenta of about the pion mass and the binding energy difference between the triton (^{3}H) and ^{3}He. We find $E(^3\mathrm{H}) - E(^3\mathrm{He}) = 0.78(5)$ MeV, in good agreement with the experimental value of 0.76 MeV[9,10]. The ground state energies of nuclei with $A = 4, 6, 12$ where also calculated in these papers.

Our collaboration has recently completed *ab initio* lattice calculations of the low-energy spectrum of ^{12}C using chiral nuclear EFT[3]. In addition to the ground state and the excited spin-2 state[b], our calculation finds a resonance with angular momentum zero and positive parity at $-85(3)$ MeV, very close to the ^{4}He+^{8}Be threshold at $-86(2)$ MeV. This is nothing but the Hoyle state. It plays a crucial role in the hydrogen burning of stars heavier than our sun and in the production of carbon and other elements necessary for life. This excited state of ^{12}C was postulated by Hoyle[6] as a necessary ingredient for the fusion of three α-particles to produce carbon at stellar temperatures. For this reason,

[b]The orientation bias in our initial state ensemble leads to the splitting of the $J_z = 0$ and $J_z = 2$ components of the spin-2 excited state.

	0_1^+	0_2^+	$2_1^+, J_z = 0$	$2_1^+, J_z = 2$
LO [$O(Q^0)$]	$-110(2)$	$-94(2)$	$-92(2)$	$-89(2)$
NLO [$O(Q^2)$]	$-93(3)$	$-82(3)$	$-87(3)$	$-85(3)$
IB + EM [$O(Q^2)$]	$-85(3)$	$-74(3)$	$-80(3)$	$-78(3)$
NNLO [$O(Q^3)$]	$-91(3)$	$-85(3)$	$-88(3)$	$-90(4)$
Experiment	-92.16	-84.51	-87.72	

Table 1. Lattice results for the ground state 0_1^+ and the low-lying excited states of ^{12}C. For comparison the experimentally observed energies are shown. All energies are in units of MeV.

the Hoyle state plays also a very important role in the context of the anthropic principle, although such considerations did not play any role when this state was predicted[7]. The Hoyle state has been an enigma for nuclear structure theory since decades, even the most successful Green's function MC methods based on realistic two- and three-nucleon forces[11] or the no-core-shell-model employing modern (chiral or $V_{\mathrm{low\ k}}$) interactions[12,13] have not been able to describe this state. In Tab. 1 we show results for the ground state and the low-lying excited states of ^{12}C at leading order (LO), next-to-leading order (NLO), next-to-leading order with isospin-breaking and electromagnetic corrections (IB + EM), and next-to-next-to-leading order (NNLO). All energies are in units of MeV. For comparison we list the experimentally observed energies. The error bars in Tab. 1 are one standard deviation estimates which include both Monte Carlo statistical errors and uncertainties due to extrapolation at large Euclidean time. Systematic errors due to omitted higher-order interactions can be estimated from the size of corrections from $O(Q^0)$ to $O(Q^2)$ and from $O(Q^2)$ to $O(Q^3)$. In Fig. 4 we show lattice results used to extract the excited state energies at leading order. For each excited state we plot the logarithm of the ratio of the projection amplitudes, $Z(t)/Z_{0_1^+}(t)$, at leading order. $Z_{0_1^+}(t)$ is the ground state projection amplitude, and the slope of the logarithmic function at large t gives the energy difference between the ground state and the excited state. As seen in Tab. 1 and summarized in the right panel of Fig. 4, the NNLO results for the Hoyle state and spin-2 state are in agreement with the experimental values. While the ground state and spin-2 state have been calculated in other studies, these results are the first *ab initio* calculations of the Hoyle state with an energy close to the phenomenologically important ^{8}Be-alpha threshold. Experimentally the ^{8}Be-alpha threshold is at -84.80 MeV, and the lattice determination at NNLO gives $-86(2)$ MeV. We also note the energy level crossing involving the Hoyle state and the spin-2 state. The Hoyle state is lower in energy at LO but higher at NLO. One of the main characteristics of the NLO interactions is to increase the repulsion between nucleons at short distances. This has the effect of decreasing the binding strength of the spinless states relative to higher-spin states. We note the 17 MeV reduction in the ground state binding energy and 12 MeV reduction for the Hoyle state while less than half as much binding correction for the spin-2 state. This degree of freedom in the energy spectrum suggests that at least some fine-tuning of parameters is needed to set the Hoyle state energy near the ^{8}Be+^{4}He threshold. It would be very interesting to understand which fundamental parameters in nature control this fine-tuning. At the most fundamental level there are only a few such parameters, one of the most interesting being the masses of the up and down quarks. Such investigations have already been performed to unravel the quark mass dependence

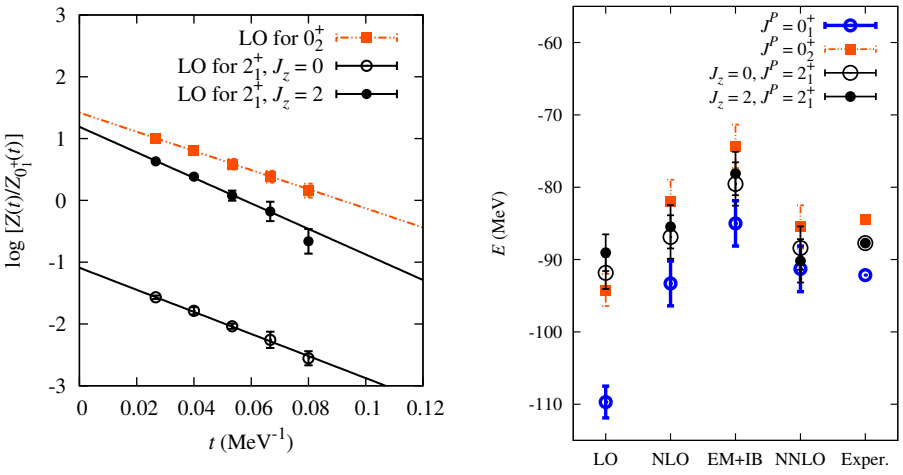


Figure 4. Left: Extraction of the excited states of ^{12}C from the time dependence of the projection amplitude at LO. The slope of the logarithm of $Z(t)/Z_{O_1^+}(t)$ at large t determines the energy relative to the ground state. Right: Summary of the results for the ^{12}C spectrum and comparison with experiment. For each order in chiral EFT labeled on the left. results are shown for the ground state (blue circles). the Hoyle state (red squares) and the $J_z = 0$ (open black circles) and $J_z = 2$ (filled black circles) projections of the lowest-lying spin-2 state.

of the deuteron binding energy and of the S-wave nucleon-nucleon scattering lengths[14,15]. The impact on the primordial abundances of light elements created by a variation of the quark masses at the time of Big Bang nucleosynthesis was also studied in Ref. 16.

To gain insight into the structure of the ground and excited states, we have computed the proton-proton radial distribution function $f_{pp}(r)$ at LO. Using any given proton as a reference point, the function $f_{pp}(r)$ is proportional to the probability of finding a second proton at a distance r,

$$f_{pp}(\vec{n}) = \frac{1}{Z^2} \sum_{\vec{n}'} \langle \Psi_A | : \rho_p^{a^\dagger,a}(\vec{n}+\vec{n}') \rho_p^{a^\dagger,a}(\vec{n}') : | \Psi_A \rangle \,. \tag{1}$$

For macroscopic liquids the radial distribution function is normalized to 1 at asymptotically large distances. In our finite system we instead normalize the integral of $f_{pp}(r)$ over all space to equal $1 - Z^{-1}$, where Z is the total number of protons. On the lattice, one only has a finite number of values for the radial distance r. Therefore, we define $f_{pp}(r)$ as an interpolating function that takes the values of $f_{pp}(\vec{n})$ averaged over all points $\vec{n}$ with length closest to r. Evaluating $f_{pp}(r)$, we find that the ground state is very compact with a large central core. The Hoyle state and spin-2 state look qualitatively similar, though the Hoyle state has a slightly larger central core. A secondary maximum near $r \simeq 4$ fm is visible in the ground state and each of the excited states. This secondary maximum seems to arise from configurations where three α–clusters are arranged approximately linearly. More calculations are planned to confirm whether this configuration is physically important or just a lattice artifact. Higher order corrections to the radial distribution function, charge radii, as well as electromagnetic transition strengths are currently under investigation. In particular,

it will be challenging to compare the theoretical results with the precision data on electromagnetic transitions from the Hoyle state obtained with the SDALINAC at Darmstadt[17].

6 Summary and Outlook

Nuclear lattice simulations are a tool to investigate the structure of atomic nuclei based on the forces derived from the chiral Lagrangian of QCD. Besides the nuclei discussed here, our collaboration has also investigated dilute neutron matter in the unitary limit[18] and is presently addressing the issue of possible P-wave pairing in neutron matter, that would modify the nuclear equation of state and the cooling rate of neutron stars. Nuclear reaction dynamics can also be addressed within this framework, however, more conceptual developments are necessary. A first step in this direction was recently performed. In Ref. 19 it was shown that bound states moving in a finite periodic volume have an energy correction which is topological in origin and universal in character. The topological volume corrections contain information about the number and mass of the constituents of the bound states. These phase corrections should be considered when determining scattering phase shifts for composite objects at finite volume. Such investigations pave the way for a model-independent calculation of the unsatisfactorily determined reaction $^{12}C(\alpha,\gamma)^{16}O$ at stellar energies. This process plays an important role in the element synthesis in stars and in the dynamics underlying supernova explosions. It is often called the "holy grail" of nuclear astrophysics. The framework presented here holds the promise to solve this outstanding problem and thus might contribute significantly to our understanding of the generation of the elements needed for life on Earth.

Acknowledgements

I thank my NLEFT colleagues Dean Lee, Evgeny Epelbaum and Herrman Krebs for sharing their insights and performing the work reported here. Work supported in part by DFG, HGF and BMBF. The simulations reported here were performed on the JUGENE at the Jülich Supercomputing Centre.

References

1. S. Weinberg, *Effective chiral Lagrangians for nucleon - pion interactions and nuclear forces*, Nucl. Phys. **B363** (1991) 3-18.
2. E. Epelbaum, H.-W. Hammer, U.-G. Meißner, *Modern Theory of Nuclear Forces*, Rev. Mod. Phys. **81** (2009) 1773-1825. [arXiv:0811.1338 [nucl-th]].
3. E. Epelbaum, H. Krebs, D. Lee, U.-G. Meißner, *Ab initio calculation of the Hoyle state*, Phys. Rev. Lett. **106** (2011) 192501. [arXiv:1101.2547 [nucl-th]].
4. D. Lee, *Lattice simulations for few- and many-body systems*, Prog. Part. Nucl. Phys. **63** (2009) 117-154. [arXiv:0804.3501 [nucl-th]].
5. E. Wigner, *On the Consequences of the Symmetry of the Nuclear Hamiltonian on the Spectroscopy of Nuclei*, Phys. Rev. **51** (1937) 106-119.
6. F. Hoyle, *On Nuclear Reactions Occuring in Very Hot Stars. 1. The Synthesis of Elements from Carbon to Nickel*, Astrophys. J. Suppl. **1** (1954) 121-146.

7. H. Kragh, *An anthropic myth: Fred Hoyle's carbon-12 resonance level*, Arch. Hist. Exact Sci. **64** (2010) 721.
8. E. Epelbaum, H. Krebs, D. Lee, U.-G. Meißner, *Lattice chiral effective field theory with three-body interactions at next-to-next-to-leading order*, Eur. Phys. J. **A41** (2009) 125-139. [arXiv:0903.1666 [nucl-th]].
9. E. Epelbaum, H. Krebs, D. Lee, U.-G. Meißner, *Lattice effective field theory calculations for A = 3,4,6,12 nuclei*, Phys. Rev. Lett. **104** (2010) 142501. [arXiv:0912.4195 [nucl-th]].
10. E. Epelbaum, H. Krebs, D. Lee, U.-G. Meißner, *Lattice calculations for A=3,4,6,12 nuclei using chiral effective field theory*, Eur. Phys. J. **A45** (2010) 335-352. [arXiv:1003.5697 [nucl-th]].
11. S. C. Pieper, *Quantum Monte Carlo calculations of light nuclei*, Riv. Nuovo Cim. **31** (2008) 709-740. [arXiv:0711.1500 [nucl-th]].
12. P. Navratil, V. G. Gueorguiev, J. P. Vary, W. E. Ormand, A. Nogga, *Structure of A=10-13 nuclei with two plus three-nucleon interactions from chiral effective field theory*, Phys. Rev. Lett. **99** (2007) 042501. [nucl-th/0701038].
13. R. Roth, J. Langhammer, A. Calci, S. Binder, P. Navratil, *Similarity-Transformed Chiral NN+3N Interactions for the Ab Initio Description of 12-C and 16-O*, Phys. Rev. Lett. **107** (2011) 072501. [arXiv:1105.3173 [nucl-th]].
14. S. R. Beane, M. J. Savage, *Variation of fundamental couplings and nuclear forces*, Nucl. Phys. **A713** (2003) 148-164. [hep-ph/0206113].
15. E. Epelbaum, U.-G. Meißner, W. Gloeckle, *Nuclear forces in the chiral limit*, Nucl. Phys. **A714** (2003) 535-574. [arXiv:nucl-th/0207089 [nucl-th]].
16. P. F. Bedaque, T. Luu, L. Platter, *Quark mass variation constraints from Big Bang nucleosynthesis*, Phys. Rev. **C83** (2011) 045803. [arXiv:1012.3840 [nucl-th]].
17. M. Chernykh, H. Feldmeier, T. Neff, P. von Neumann-Cosel, A. Richter, *Structure of the Hoyle State in C-12*, Phys. Rev. Lett. **98** (2007) 032501.
18. E. Epelbaum, H. Krebs, D. Lee, U.-G. Meißner, *Ground state energy of dilute neutron matter at next-to-leading order in lattice chiral effective field theory*, Eur. Phys. J. **A40** (2009) 199-213. [arXiv:0812.3653 [nucl-th]].
19. S. Bour, S. König, D. Lee, H.-W. Hammer, U.-G. Meißner, *Topological phases for bound states moving in a finite volume*, [arXiv:1107.1272 [nucl-th]].

Finite Temperature Lattice QCD with Wilson Quarks

Szabolcs Borsányi[1], **Zoltán Fodor**[1,2,3], **Christian Hoelbling**[1], **Sándor Katz**[3], **Stefan Krieg**[1,2], **Kálmán Szabó**[1], **and Bálint Tóth**[1]

[1] Theoretische Physik, Bergische Universität, 42119 Wuppertal, Germany

[2] Institute for Advanced Simulation, Jülich Supercomputing Centre
Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

[3] Elméleti Fizika Tanszék, Eötvös Loránd Tudományegyetem, 1131 Budapest, Hungary
E-mail: szaboka@general.elte.hu

We propose to push forward calculations in finite temperature QCD with Wilson quarks, which are free of the problems of commonly used staggered quarks. This approach also faces challenges, which require supercomputer resources to be tackled. We have generated configurations at four lattice spacings and four pion masses. Results show, that fine lattices are necessary. The results from fine lattices at the largest mass are in agreement with previous staggered results

1 Introduction and Motivation

Motivation for Finite Temperature QCD

Understanding the behaviour of the matter under extreme conditions, such as extremly high temperature or density, contributes to our knowledge about the early Universe and the structure of neutron stars. This type of matter is produced in heavy-ion collision experiments: eg. in the ALICE experiment of the recently started Large Hadron Collider or in the heavy ion program of the FAIR accelerator to be started at the GSI in Darmstadt. In these experiments one is probing a temperature region, where a transition happens: from a phase, where protons and nucleons are stable, to another phase, where these particles fall apart to their more elementary constituents, called quarks. Providing a firm theoretical basis for the above experiments is essential. The theory, which describes these experiments is called Quantum Chromo Dynamics (QCD). It describes how quarks interact with each other. Mathematically it is equivalent to a nonlinear system of equations.

Motivation for Lattice QCD

There are many ways to solve these complicated equations, but lattice QCD is the only known way, which is able to solve these equations, so that all systematic errors can be kept under control. Lattice QCD approximates the continuous spacetime volume, where the experiments take place, with a discrete grid, a lattice. By taking finer and finer lattices an upper bound on the systematic error can be given.

Motivation for Supercomputers

The core of these calculations is the numerical solution of non-linear differential equations with dimension of around 10 millon. The differential operator is local, which enables a very efficient parallelization. The large dimension and the good parallelizability makes the Blue Gene supercomputer our ideal candidate to carry out these calculations.

Motivation for Wilson Quarks

Previous large scale projects for finite temperature lattice QCD dominantly used *staggered quarks* (see for example Refs. 1,2) . The staggered formulation has three major drawbacks:

1. it is not an *a priori* proper discretization of the strong interaction (due to the so-called rooting procedure).

2. it has a delicate discretization problem (called taste violation), which distorts the results at small temperatures. The analysis of these discretization errors shows that there is no hope to get them under control below $T < 130$ MeV temperature with reasonable amount of computer time in this formulation.

3. it has problems with its definition at non-zero baryon density.

In the light of these facts a formulation without such drawbacks would be highly welcomed both from the lattice practicioners and from experimentalists. There have also been projects with other quark discretizations, however these were more like exploratory studies due to the very large discretization artefacts.

Wilson quarks offer a solution for all the shortcomings of the staggered discretization, listed above:

1. it provides a theoretically sound discretization of the strong interaction.

2. it does not suffer from the taste violation discretization error, therefore it is particularly well suited to study the low temperature regime, where staggered taste breaking has a significant distortive effect on the results.

3. it is well defined even at finite baryon density.

One has to keep in mind, that these advantages have their price: calculations with Wilson quarks are considerably more expensive, than with staggered quarks. Moreover, due to the different characteristics of Wilson quarks new kind of methods are required, only a few of which have been already tested in the case of staggered quarks. This means, that results with Wilson quarks will need some time to reach the same level of quality as those with staggered fermions.

2 Physics Goals

Our long term goal is to push thermodynamic simulations with Wilson quarks ahead replacing previous calculations with staggered quarks. This project is an important step to reach this goal.

The continuum limit is a very crucial point in lattice calculations. Some preliminary investigations at a large pion mass and a coarse lattice spacing ($a = 0.139$ fm) have shown large lattice artefacts. In order to avoid this problem, we push simulations to the continuum limit aggressively: we use four lattice spacings down to $a = 0.057$ fm. Since the cost increases with a high power of the inverse lattice spacing and of the inverse pion mass, we leverage the cost of the fine lattices by using pion masses which are somewhat larger

than the physical (down to $m_\pi = 280$ MeV, which is approximately two times the physical pion mass). We use four pion masses for each lattice spacing, which will allow us a chiral extrapolation to the physical pion. We plan to measure standard thermodynamical observables (renormalized chiral condensate, quark number susceptibility, Polyakov-loop) and compare them to staggered results.

3 Methods

In this section we highlight the methods, on which we build our investigations.

Action and Algorithm

We use the same lattice action, that was already used by us to study the properties of the hadrons in several publications. The exact definition and the properties of our action, with particular focus on the excellent scaling behaviour, were published in Ref. 3.

We use the Hybrid Monte Carlo (HMC) algorithm to generate gauge field configurations. The light quark flavours are preconditioned ala Hasenbusch, the strange quark is treated using rational functions (Rational-HMC). The evolution of the gauge field is integrated with multiple timescale integrator in the Omelyan scheme. By analyzing the contribution of different forces into the HMC energy violation, we are currently developing a new and more efficient scheme to set the timescales of the integrator.

Mixed Precision Inverter

The most time consuming part of the simulation is the inversion of the Wilson fermion matrix, based on an iterative procedure involving many fermion matrix-vector multiplications. Due to bandwidth and cache occupancy, the matrix multiplication in single precision is about 50% faster than in double precision; however it is well known that double precision results are necessary. Therefore we use a mixed precision method which yields double precision results but uses single precision for almost all of the multiplications.

Ensemble Generation

We are working in the so-called *fixed-scale scheme* of thermodynamics. The temperature is changed by changing the number of lattice points in the temporal direction. This technique is especially useful for Wilson fermions, where one has a fine tuning problem as compared to a normal tuning problem in case of staggered quarks.

We are performing calculations at four different lattice spacings. The target is to have four different pion masses at each ($m_\pi = 280, 350, 420$ and 490 MeV). The spatial lattice volumes are chosen so that at the critical temperature the aspect ratio is around 3. This also means, that even for our lightest pion mass the pion correlation length fits at least four times into the box. For all pion masses the bare strange mass is set to its numerical value at the physical point (ie. at $m_\pi = 135$ MeV). Tab. 1 lists the lattice spacing, pion mass and spatial lattice size of the ensembles. We have collected at least 600 trajectories at around 10 different temperature values and also at zero temperature.

a[fm]	m_π[MeV]	N_s
0.139(1)	485(5)	32
	410(5)	32
	350(5)	32
	275(5)	32
0.093(1)	495(5)	32
	425(5)	48
	360(5)	48
	290(5)	48
0.070(1)	500(5)	48
	425(10)	48
0.057(1)	500(5)	64
	441(15)	64

Table 1. Ensembles for Wilson thermodynamics.

Code Performance

Our code is the `dynqcd` program suite, which incorporates all algorithms, techniques and tricks, which we have used in the previous years on the JUGENE supercomputer, on the JUROPA cluster and on various other clusters.

The code on JUGENE has optimized one node performance, internode communication, and input/output routines. The good scaling properties of our code are apparent from the scaling plot Fig. 1. Shown there is a strong scaling analysis of the most time critical part of our code from one Rack of the Blue Gene/P in Jülich to the whole machine. The scaling is almost perfect and peaks at a sustained performance of 37.5% of machine peak.

The code on JUROPA uses persistent and asynchronous MPI calls for internode com-

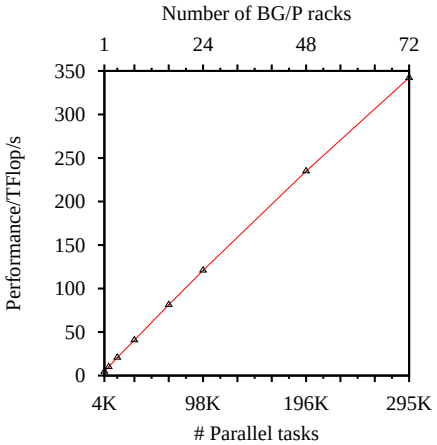


Figure 1. Performance scaling analysis for Wilson fermion solvers, which is the bottleneck in our code.

munication. We did not optimize for the one node performance, since our application is bandwidth limited and using all the eight computer cores of a node saturates the bandwidth even with an unoptimized code. It performs at 18.4 Gflops per JUROPA node.

Improving the Random Vector Technique

The evaluation of the traces of the fermion matrix to obtain the quark number susceptibility turned out to be really challenging. Even the connected part of the operator was much noisier than in the staggered case, the calculation of the disconnected part was hopeless in a reasonable amount of time. When calculating traces with the usual random vector technique, the noise is proportional to the strength of the off-diagonal elements. We implemented a new random vector basis, that is constructed to suppress the off-diagonal terms of the fermion matrix. With this improvement we could achieve to calculate even the disconnected part with a reasonable accuracy.

Renormalization of the Chiral Condensate

In order to investigate the restoration of chiral symmetry with Wilson fermions in the continuum limit a careful renormalization of the bare chiral condensate

$$\langle\bar{\psi}\psi\rangle = \frac{1}{Z}\frac{dZ}{dm} \tag{1}$$

is needed. Unlike in the staggered case there is no chiral symmetry with Wilson fermions, which would come handy when defining a renormalized chiral condensate. For staggered fermions one can define the renormalized chiral condensate multiplied with the renormalized mass as

$$m^{\mathrm{ren}}\langle\bar{\psi}\psi\rangle_T^{\mathrm{ren}} \equiv m\left(\langle\bar{\psi}\psi\rangle_T - \langle\bar{\psi}\psi\rangle_{T=0}\right), \tag{2}$$

which has a proper continuum limit. The Wilson case is somewhat more complicated. In this case one has to define the chiral condensate through Ward-identities (for staggered quarks this would yield the same definition as above). This definition of the bare chiral condensate uses the PP correlator

$$\langle\bar{\psi}\psi\rangle' = 2Z_A m_{PCAC}\sum_x\langle P(x)P(0)\rangle. \tag{3}$$

The additive divergences can be removed by the usual zero temperature subtraction and the remaining multiplicative divergences can be renormalized away by Z_P. Finally we also multiply it with the renormalized PCAC quark mass:

$$m^{\mathrm{ren}}\langle\bar{\psi}\psi\rangle_T^{\mathrm{ren}} \equiv \frac{Z_A}{Z_P}m_{PCAC}\cdot Z_P\cdot\left(\langle\bar{\psi}\psi\rangle'_T - \langle\bar{\psi}\psi\rangle'_{T=0}\right) \tag{4}$$

In order to evaluate this prescription we need to determine the m_{PCAC} and Z_A at zero temperature.

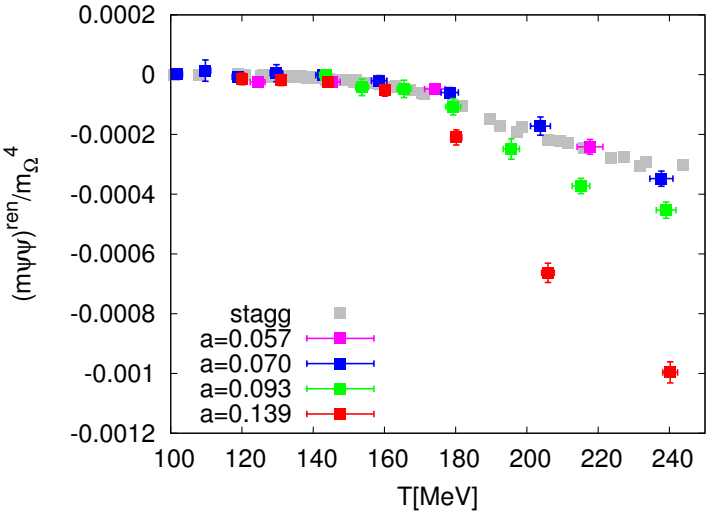


Figure 2. Renormalized chiral condensate.

Renormalization of the Polyakov Loop

The real part of the bare Polyakov loop also needs to be renormalized in order to have a quantity with a finite continuum limit. Since there is an additive divergence in the free energy, a convenient choice of renormalization prescription is demanding a fixed value P_* for the renormalized Polyakov loop at a fixed but arbitrary temperature $T_* > T_c$. Then the renormalized Polyakov loop P_R is given by

$$P_R(T) = \left(\frac{P_*}{P_0(T_*)}\right)^{\frac{T_*}{T}} P_0(T) \tag{5}$$

in terms of the bare Polyakov loop $P_0(T)$. We choose $T_* = 0.143 m_\Omega$ and $P_* = 1.2$. Other choices would simply correspond to other renormalization schemes.

4 Results

We show results for the highest pion masses for all four lattice spacings, the smaller pion masses will follow soon. We have measured three quantities. The renormalized chiral condensate is sensitive to the chiral transition whereas the renormalized Polyakov loop and the strange quark number susceptibility are sensitive to the confinement-deconfinement transition. Each quantity is renormalized properly so that in the continuum limit finite and regularization scheme independent values are obtained. These can be compared with results from other regularizations such as the staggered formulation.

Fig. 2 shows the renormalized chiral condensate at the 4 different lattice spacings together with staggered results at 3 fixed N_t values, 8, 10 and 12 (the staggered results were obtained using the fixed N_t scheme). As is apparent from the figure the two finer lattice

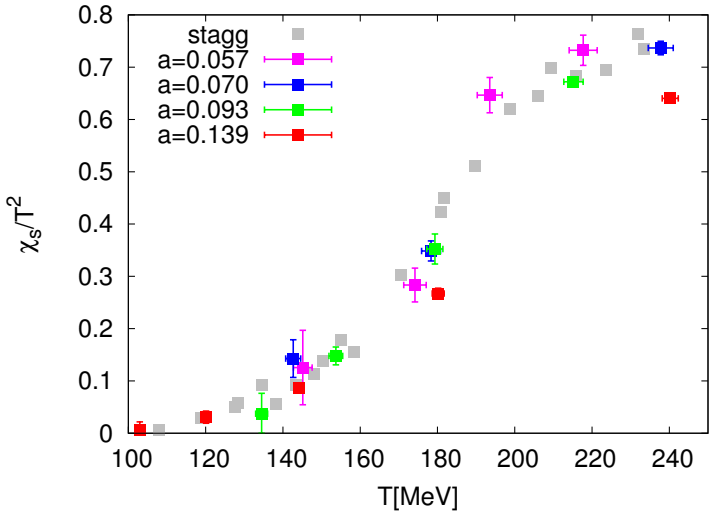


Figure 3. Strange quark number susceptibility.

spacing values are small enough to observe only a small change in the chiral condensate at any temperature. The roughest lattice spacing on the other hand is seen to include large discretization effects and is probably not in the scaling region. What is also clear is that results at $N_t < 8$ are not trustworthy at any of the lattice spacings due to a too rough resolution of the imaginary time direction (temperature). This is also true for the finest lattice spacing ($a = 0.057$ fm). This constraint limits our ability to achieve very high temperature values without introducing too large discretization effects. On the other hand at only moderately high and small temperatures, particularly around T_c, the three finest lattice spacings are probably in the scaling region and continuum extrapolation is possible.

The strange quark number susceptibility is a sum of two contributions, the connected and disconnected terms. The disconnected part is a very noisy quantity (as usual) and a large number of random vectors are needed in order to evaluate it precisely. The results for the four lattice spacings are shown in Fig. 3 together with the staggered results. What is plotted is the measured quark number susceptibility divided by the massless free value (Stefan-Boltzmann limit) in both cases. This combination has a smoother continuum limit than the measured susceptibility itself but the continuum limit value is of course the same.

The third quantity we measured is the renormalized Polyakov loop which is sensitive to the confinement-deconfinement transition similarly to the quark number susceptibility. The results are shown in Fig. 4.

5 Concluding Remarks

In this work we have started our study of QCD thermodynamics with Wilson quarks. Special emphasis is placed on obtaining results in the continuum limit, i.e. the simulations are performed at four lattice spacings. In this first study the light quark masses were kept

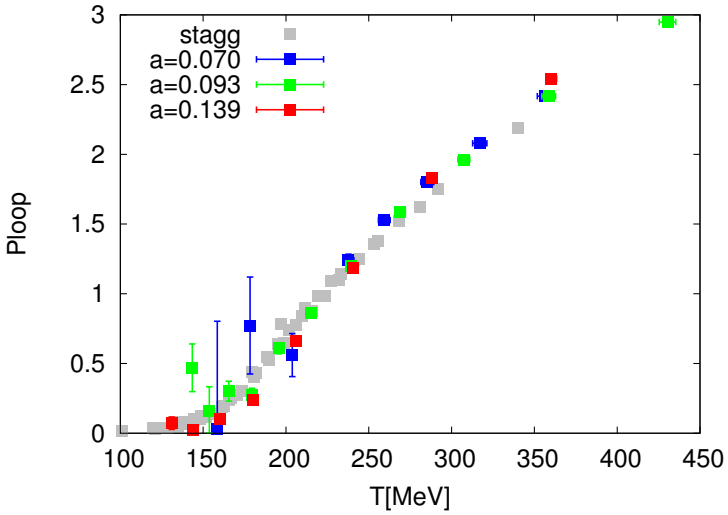


Figure 4. Renormalized Polyakov loop.

above the physical value but since continuum results are universal even at non-physical quark masses a comparison is possible with continuum staggered results. At our largest pion mass we found nice agreement between the finest lattice spacings of the Wilson and staggered formulations. Whether this agreement also holds at smaller pion masses, we will report on it soon.

Acknowledgements

Computations were performed on the JUGENE and JUROPA supercomputer at FZ Jülich and on clusters at Wuppertal and also at the Eötvös University, Budapest. This work is supported in part by European Union (EU) grant I3HP; Deutsche Forshungsgemeinschaft grants FO 502/2 and SFB-TR 55 and (FP7/2007-2013)/ERC no. 208740.

References

1. A. Bazavov, T. Bhattacharya, M. Cheng, N. H. Christ, C. DeTar, S. Ejiri, S. Gottlieb, R. Gupta *et al.*, Phys. Rev. **D80** (2009) 014504. [arXiv:0903.4379 [hep-lat]].
2. S. Borsanyi, G. Endrodi, Z. Fodor, A. Jakovac, S. D. Katz, S. Krieg, C. Ratti, K. K. Szabo, JHEP **1011** (2010) 077. [arXiv:1007.2580 [hep-lat]].
3. S. Durr, Z. Fodor, C. Hoelbling, R. Hoffmann, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch *et al.*, "Scaling study of dynamical smeared-link clover fermions", Phys. Rev. **D79** (2009) 014501.

Probing the $N_f = 2$ QCD Phase Transition with Light Wilson Fermions

Bastian B. Brandt[1], **Owe Philipsen**[2], **Hartmut Wittig**[1,3], **and Lars Zeidlewicz**[2]

[1] Institut für Kernphysik, Johannes Gutenberg-Universität Mainz
Johann Joachim Becher-Weg 45, 55099 Mainz, Germany
E-mail: {*brandt, wittig*}*@kph.uni-mainz.de*

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

[3] Helmholtz Institut Mainz, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany

We report on the status of our current lattice QCD study on the deconfinement transition with two degenerate flavours of non-perturbatively $\mathcal{O}(a)$-improved Wilson fermions. We focus on the unknown transition temperature and order of the transition in the chiral limit. To simulate sufficiently close to the continuum and the chiral limits, we work on lattices as large as 16×64^3 and aim at pion masses down to 200 MeV. We also show first results obtained employing an extended multi-histogram method for β-dependent fermion actions.

1 Introduction

The properties of matter at very high temperatures and densities are the subject of intense experimental and theoretical investigations. Such extreme conditions are believed to have occurred during the immediate aftermath of the big bang. Moreover, it is expected that extremely high matter densities are realised in the interior of compact or neutron stars. A big effort is currently undertaken by the nuclear and particle physics communities to recreate these conditions in the laboratory, which is signified by the large-scale experimental programmes of heavy-ion collisions at RHIC, CERN and the future FAIR facility. In order to interpret the experimental findings, these efforts must be complemented by intense theoretical research. Under extreme conditions the conventional picture of matter consisting of nuclei and their constituents – protons and neutrons – is no longer valid. Instead one believes that a plasma-like phase exists which contains the fundamental building blocks of matter, i.e. quarks and gluons. The conjectured form of the phase diagram of "ordinary" hadronic matter and the quark-gluon plasma in the plane defined by the temperature and the baryon density is shown in Fig. 1[a]. By definition, the mechanism which is responsible for the "confinement" of quarks and gluons into hadrons is no longer at work in the plasma phase. In addition to deconfinement, another interesting feature of the quark-gluon plasma is the restoration of chiral symmetry, which is spontaneously broken in the hadronic phase.

The interaction between quarks and gluons is described by the theory of Quantum Chromodynamics (QCD). Theoretical and experimental studies have shown that the critical temperature T_c at which the transition to the plasma phase occurs, lies in a regime where the coupling strength between quarks and gluons is still large. Therefore, analytic techniques

[a] See `http://www.gsi.de/forschung/fair_experiments/CBM/1intro_e.html`.

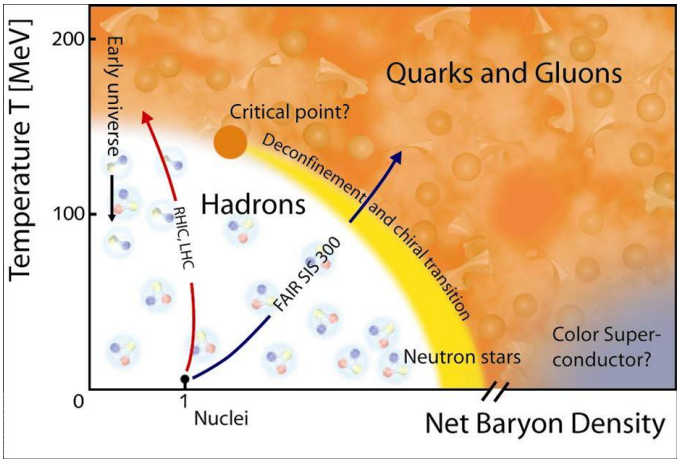


Figure 1. The phase diagram of QCD.

based on a perturbative expansion in the QCD coupling constant fail, and non-perturbative methods such as lattice QCD must be employed instead. However, in order to obtain detailed and reliable information on the QCD phase diagram from lattice simulations, a large amount of computer time must be invested.

QCD is parameterised in terms of the coupling constant and the masses of each of the six known quarks. As far as the properties of the phase diagram are concerned, the light quarks – up, down and strange – play a special role. In order to study the phase diagram in lattice QCD, one has to include the temperature T and the baryon chemical potential μ. Lattice simulations are performed by employing a Euclidean version of space-time, where the temporal extent of the lattice can be interpreted as the inverse temperature of the system. Non-zero values of the temperature are thus easy to realise. By contrast, the inclusion of non-zero values of μ is difficult as a result of the so-called sign-problem, which arises from the fact that the QCD action becomes complex. In this case the interpretation of the exponentiated action as a probability measure breaks down. In our project we are concerned with the properties of the QCD phase diagram at vanishing chemical potential. Nonetheless, as will be explained below, our results may also have important consequences for the phase transition to the plasma phase at non-zero μ.

In lattice QCD the bare quark masses, $m_{u,d}$ and m_s for the light quarks, are free parameters. Our simulations are focused on investigating the order of the phase transition, as well as studying the interplay between chiral symmetry restoration and deconfinement in different limits in the enlarged parameter space $\{m_{u,d}, m_s, T\}$. Fig. 2 shows the phase diagram for vanishing chemical potential in the $m_{u,d}$-m_s plane. In the opposite limits of pure gauge theory and QCD with massless quarks, there are true first-order phase transitions associated with the breaking of centre symmetry[1] and restoration of chiral symmetry[2], respectively. For intermediate quark masses the finite temperature transition is a smooth and analytic crossover, and the boundary lines separating first order and crossover regions are second-order critical lines.

Of particular interest for our study is the boundary line which separates the first-order

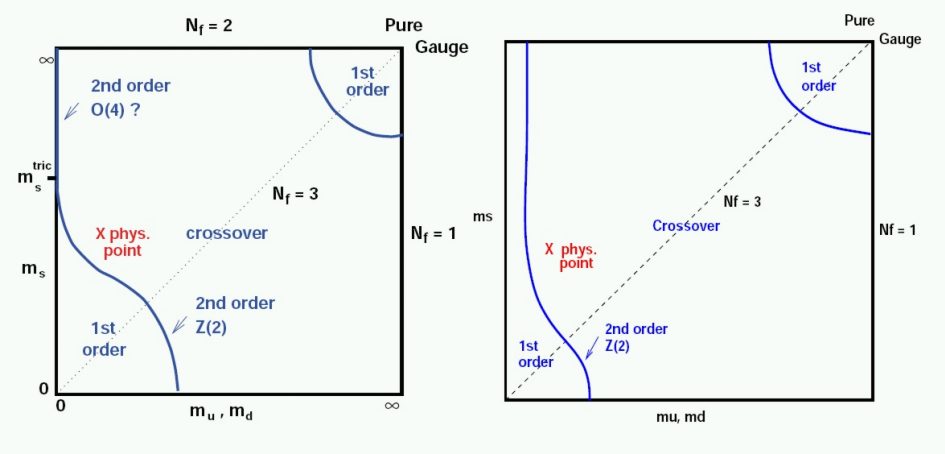


Figure 2. Two different scenarios for the quark mass dependence of the phase structure of QCD at zero chemical potential. The left panel depicts the case of a second-order transition in the chiral limit of two-flavour QCD, implying the existence of a tri-critical point. The right panel illustrates the alternative scenario of a first-order transition.

chiral phase transition from the crossover region, i.e. the so-called chiral critical line. On that line the transition is in the Z(2) universality class of the 3d Ising model[3,4]. The situation in the limit of two degenerate massless quarks is still largely unknown. Here, chiral symmetry breaking or restoration must be accompanied by a non-analytic phase transition, since the chiral condensate constitutes a true order parameter. There are two possible scenarios: In the first scenario, depicted in the left panel of Fig. 2, the chiral critical line reaches the $m_{u,d} = 0$ axis at some tri-critical point $m_s = m_s^{tric}$. In this case the transition in the chiral limit of the $N_f = 2$ theory is second order and the restoration of chiral symmetry corresponds to the $\mathrm{O}(4)$ universality class[5]. In the alternative scenario the Z(2) chiral critical line never reaches the $m_{u,d} = 0$-axis, implying that the chiral transition at $m_{u,d} = 0$ is first order for all values of the strange quark mass.

The nature of the transition in the two-flavour theory is also of fundamental importance for the phase diagram at non-zero baryon chemical potential. If the first of the above scenarios is realised, one would expect another tri-critical point in the (T, μ)-plane. When the quarks acquire a small mass, chiral symmetry is broken explicitly, and there is no true order parameter connected with chiral symmetry restoration. As a result the second-order phase transition becomes a smooth crossover, while the tri-critical point becomes a critical endpoint. The search for this endpoint is a major research activity, both in numerical studies and experiment.

2 Goals and Strategy

The primary goal of our study is the extraction of the order of the transition in the chiral limit and to discriminate between the two scenarios discussed above. Benchmark calculations for the free pressure[6] and the recently resolved controversy on the results for the

transition temperature from simulations with staggered fermions [7,8] have shown that it is mandatory to be close enough to the continuum and chiral limits to extract reliable results. For improved Wilson fermions we expect to be in the a^2-scaling region, provided that lattices with a temporal extent of $N_t \geq 16$ are used.[6] Thus we are forced to simulate on lattices as large as 16×48^3 and 16×64^3.

The standard procedure for determining T_c and the order of the transition is to monitor the behaviour of a suitable set of observables whilst varying the temperature. The temperature T of the system is related to the temporal extent N_t via $T = 1/(N_t\, a)$, where a is the lattice spacing. Since the bare coupling $\beta \equiv 6/g_0^2$ controls the value of a, one can perform a temperature scan by varying β for fixed N_t and keeping the bare quark mass constant. As discussed in section 4, one may also employ a generalised multi-histogram method to locate the transition more reliably. As observables we use the chiral condensate $\langle\bar{\psi}\psi\rangle$ and the Polyakov loop

$$L = \frac{1}{N_c}\mathrm{Tr}\prod_{x_4} U_4(\mathbf{x}, x_4)\ . \tag{1}$$

For the Polyakov loop we also employ an APE-smeared[9] version L_S, since the smearing procedure is expected to enhance the signal for the deconfinement transition[10]. Detailed information on the nature of the transition is carried by the associated susceptibilities χ_O and Binder cumulants B_4, defined by

$$\chi_O = N_s^3\,\left(\langle O^2\rangle - \langle O\rangle^2\right) \quad \text{and} \quad B_4(O) = \frac{\langle \delta O^4\rangle}{\langle \delta O^2\rangle^2}\ . \tag{2}$$

Here O denotes any linear combination of our observables, and $\delta O = O - \langle O\rangle$ is the associated fluctuation. The susceptibilities develop a peak in the region of the transition whose maximum is identified with the transition point. The scaling of the peak with the spatial volume encodes information on the order of the transition. When evaluated on the phase boundary, i.e. at the critical coupling for the particular quark mass, the Binder cumulant reveals information on the order of the transition, too. In the infinite volume limit, B_4 approaches unity in case of a first-order phase transition, 3 for a crossover, 1.604 for a $Z(2)$ second-order transition and 1.092 for an $O(4)$ second-order transition.

3 Simulation Setup and Algorithm

We work with two degenerate flavours of non-perturbatively $\mathcal{O}(a)$-improved Wilson fermions, using the Sheikholeslami-Wohlert lattice Dirac operator[11]

$$D_{SW} = \sum_{\mu=0}^{3}\frac{1}{2}\,\left[\gamma_\mu\left(\nabla_\mu^* + \nabla_\mu\right) - \nabla_\mu^*\nabla_\mu\right] + c_{SW}\,\frac{i\,a}{4}\,\sigma_{\mu\nu}\,\hat{F}^{\mu\nu} + m \tag{3}$$

for the fermionic action, together with the standard plaquette action for the gauge part. Here ∇_μ and ∇_μ^* denote the forward and backward gauge-covariant lattice derivative, $\sigma_{\mu\nu}$ is the totally antisymmetric tensor and $\hat{F}^{\mu\nu}$ the "clover leaf" representation of the gluonic field strength tensor on the lattice. Instead of the bare quark mass m one uses the hopping parameter $\kappa = 1/(2am + 8)$. To ensure full $\mathcal{O}(a)$-improvement for spectral quantities, the clover coefficient c_{SW} has to be tuned properly with the bare gauge coupling β, using

the interpolation formula[12] for $5.2 < \beta < 12.0$. In our simulations we use the deflation-accelerated DDHMC algorithm, introduced by Lüscher[13,14], which has been used extensively in large-scale lattice simulations and is known to work well for pion masses down to 250 MeV[15,16]. The excellent scaling properties and the dedicated optimisation for both cluster architectures like JUROPA and the Blue Gene/P allows us to make efficient use of the facilities provided by NIC. In fact, large-scale machines are needed to be able to perform a temperature scan on lattices as large as 16×48^3 and 16×64^3. In particular, the 16×48^3 lattices can be run efficiently on 27 nodes on JUROPA while the 16×64^3 lattice is easily accommodated by one rack of JUGENE.

4 Optimized Analysis – Extended Multi-Histogram Method

One of the main limitations is the necessarily limited knowledge on the shape of the peak of the susceptibilities around the critical point from direct simulations. The common strategy to date is to use simulation points closest to the peak region and perform a Gaussian fit. The maximum of the fit is then identified with the transition point (see e.g. Ref. 17, 18). However, this procedure relies strongly on the validity of the Gaussian approximation for the shape of the susceptibility peak. Improving on the density of measurement points in the vicinity of the transition is costly and also limited by available computer time. Multi-histogram methods[19,20] (MH) offer the possibility to combine the knowledge on the density of states, obtained from the measurements in the vicinity of the transition, for a quasi-continuous interpolation of the results at the stage of data analysis.

For our chosen lattice action a complication arises, since the explicit β-dependence of the Dirac operator, introduced by the tuning of c_{SW} in Eq. 3, spoils the applicability of the common MH method. We have therefore developed and implemented a generalised MH method for fermion actions with an explicit β-dependence[21] (see also our upcoming publication[22]). For testing purposes one would like to apply the method in a region where the signal of the transition is still strong, while the density of states receives significant contributions from the fermionic part of the action. To meet these requirements we have performed a dedicated temperature scan on an 8×16^3 lattice located near the upper right corner in Fig. 2. The results for $|L_S|$ from the extended MH method are shown in Fig. 3. The MH analysis maps out the shape of the peak as implied by the available data points and thus allows for the extraction of its maximum without approximation. Moreover, the height and the width of the peak can be extracted accurately, which is particularly useful to study finite-volume scaling.

5 Current Status

The simulation parameters for the three existing temperature scans with $N_t = 16$ are listed in Tab. 1. Scans C_2 and B_3 are new, while scan B_1 was already discussed in Ref. 21, 23. The runs for scan C_2 were done on JUROPA while the simulations for the additional volume of scan B were performed on JUGENE. With the help of the scale determination in Ref. 16, the pion mass of scan B at the transition point is estimated to be above 500 MeV. Scan C was designed to achieve a pion mass of 300 MeV at $\beta = 5.39$. The behaviour of the real part of the Polyakov loop and its susceptibility for both scans are shown in Fig. 4.

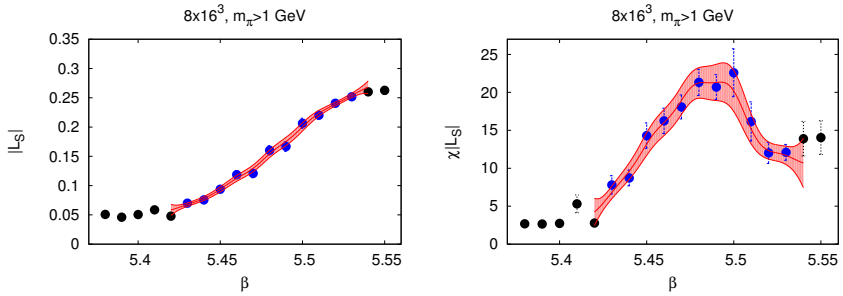


Figure 3. Results for absolute value of the APE-smeared Polyakov loop for a test scan on an 8×16^3 lattice as described in the text. The blue points are the measurements included in the extended multi-histogram analysis outlined in the text and the red filled area is the corresponding MH result.

The transition point for scan B_1 is estimated to be $\beta_c = 5.497(3)$, which in physical units corresponds to

$$T_c(m_\pi = 500\,\text{MeV}, a = 0.053\,\text{fm}) \approx 233\,\text{MeV}. \tag{4}$$

This result lies above the transition temperatures reported in Ref. 17 but is consistent with the results obtained by the tmft collaboration [24]. In addition, we show results from the larger volume of scan B, i.e. from the 16×64^3 lattice. We see good agreement in the results for the Polyakov loop, while the susceptibility for B_3 seems to favour a transition at a somewhat smaller β-value.

For scan C a first estimate for the transition point yields $\beta = 5.40(1)$. Using the results for the Sommer scale r_0 [25], obtained on the CLS ensembles in Ref. 26, to interpolate the lattice spacing, this result corresponds to a transition temperature of

$$T_c(m_\pi = 330\,\text{MeV}, a \approx 0.062\,\text{fm}) \approx 199\,\text{MeV}. \tag{5}$$

Finally, the bottom panels of Fig. 4 show our results for the Binder cumulant B_4, evaluated for the Polyakov loop for scans B (left) and C (right). We see that all runs performed so far are consistent with $B_4(L) = 3$, i.e. a crossover, as expected for these relatively large quark masses. Also, both volumes for scan B agree within error bars, which suggests that the crossover scenario might also be realised in the infinite volume limit. While these results do not yet allow for a definite conclusion, they demonstrate that we are able to extract Binder cumulants with the required accuracy.

scan	Lattice	κ	β-range	$\tau_{int}[P]$	f_{meas}	Statistics
B_1	16×32^3	0.136500	$5.40 - 5.55$	$\mathcal{O}(10)$	2	$\mathcal{O}(5000)$
B_3	16×64^3	0.136500	$5.4875 - 5.5125$	$\mathcal{O}(30)$	2	$\mathcal{O}(4000)$
C_2	16×48^3	0.136575	$5.35 - 5.42$	$\mathcal{O}(10)$	2	$\mathcal{O}(3500)$

Table 1. Run parameters for available scans at $N_t = 16$. The subscripted number labels the volume of the run, while the letters labels the pion mass at the transition point. We show the measurement frequency f_{meas} and the integrated autocorrelation time τ_{int} of the plaquette P, in units of the HMC trajectory length $\tau = 2$.

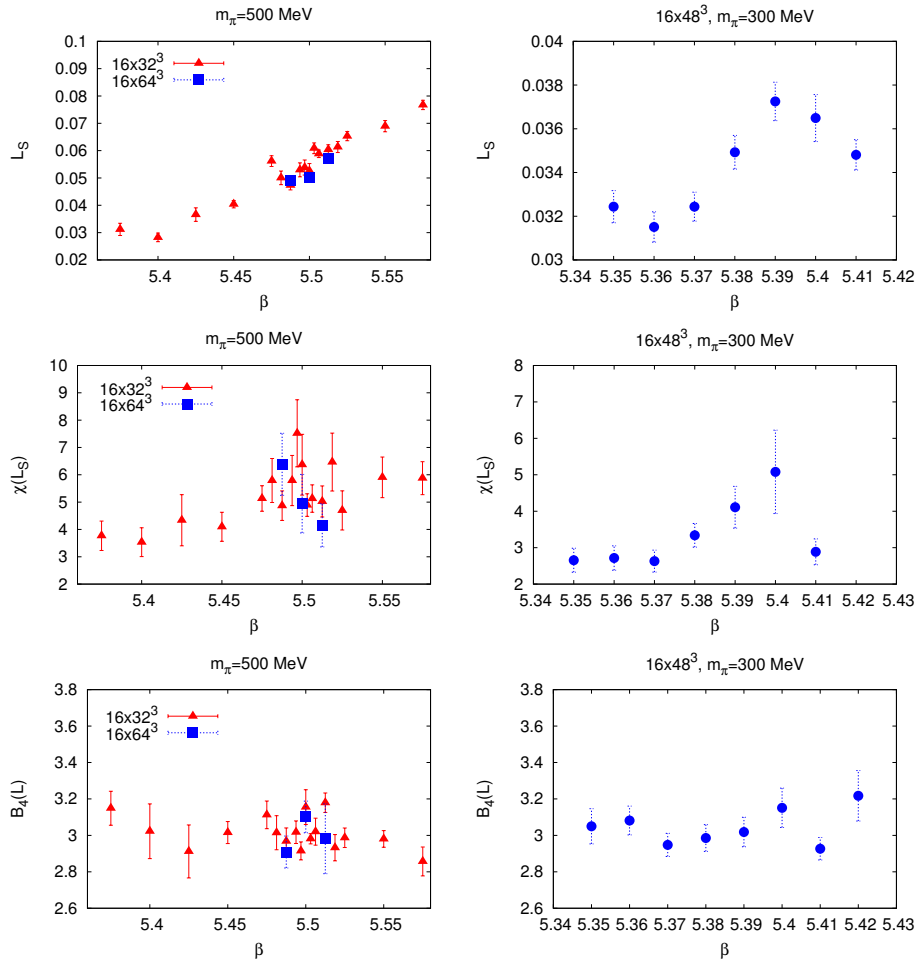


Figure 4. Results for scans B (left) and scan C_2 (right). We show the real part of the Polyakov loop (top), its susceptibility (middle) and the Binder cumulant of the bare Polyakov loop (bottom).

6 Conclusions and Outlook

In these proceedings we have presented a status report on our project to determine the critical temperature and the order of the QCD deconfinement transition in the chiral limit of two-flavour QCD. For the first time we have employed lattices as large as 16×48^3 and 16×64^3 in a study on QCD at finite temperature with Wilson fermions. We have shown results for the behaviour of the Polyakov loop, its susceptibility and Binder cumulant and extracted the associated transition temperatures. In addition, we have outlined a new generalised multi-histogram analysis, which can be used with fermionic discretisations which contain additional coupling-dependent parameters. Our priority for future runs is the extension of the simulated parameter range, in particular the realisation of smaller pion masses. In the long run, this shall allow for reliable chiral and continuum extrapolations.

Acknowledgements

Our simulations were performed on JUROPA and JUGENE as NIC project HMZ21, and on the dedicated QCD cluster "Wilson" at the Institute for Nuclear Physics, University of Mainz. We thank Harvey Meyer for useful discussions.

References

1. L.G. Yaffe and B. Svetitsky, Phys. Rev. **D26** (1982) 963.
2. R.D. Pisarski and F. Wilczek, Phys. Rev. **D29** (1984) 338.
3. F. Karsch, E. Laermann and C. Schmidt, Phys. Lett. **B520**, 41 (2001) [hep-lat/0107020].
4. P. de Forcrand and O. Philipsen, Nucl. Phys. **B673**, 170 (2003) [hep-lat/0307020].
5. K. Rajagopal and F. Wilczek, Nucl. Phys. **B399** (1993) 395 [hep-ph/9210253].
6. O. Philipsen, L. Zeidlewicz, Phys. Rev. **D81** (2010) 077501. [arXiv:0812.1177 [hep-lat]].
7. S. Borsanyi *et al.* [Wuppertal-Budapest Collaboration], JHEP **1009**, 073 (2010). [arXiv:1005.3508 [hep-lat]].
8. A. Bazavov *et al*, Phys. Rev. **D80** (2009) 014504 [arXiv:0903.4379 [hep-lat]].
9. M. Teper, Phys. Lett. **B183** (1987) 345.
10. H. Wittig, Nucl. Phys. **B325** (1989) 242.
11. B. Sheikholeslami, R. Wohlert, Nucl. Phys. **B259** (1985) 572.
12. K. Jansen and R. Sommer [ALPHA collaboration], *Nucl. Phys.* **B530** (1998) 185, [arXiv:hep-lat/9803017].
13. M. Lüscher, JHEP **0707** (2007) 081 [arXiv:0706.2298]; JHEP **0712** (2007) 011 [arXiv:0710.5417].
14. M. Lüscher, Comput. Phys. Commun. **165** (2005) 199 [hep-lat/0409106].
15. S. Capitani *et al*, PoS **LAT2009** (2009) 095 [arXiv:0910.5578].
16. B. B. Brandt *et al*, PoS **LATTICE2010** (2010) 164 [arXiv:1010.2390].
17. V.G. Bornyakov *et al.*, Phys. Rev. **D82** (2010) 014504 [arXiv:0910.2392].
18. E. -M. Ilgenfritz *et al*, PoS **LAT2007** (2007) 238 [arXiv:0710.0569].
19. A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63** (1989) 1195.
20. S. Z. Huang, K. J. M. Moriarty, E. Myers and J. Potvin, Z. Phys. C **50** (1991) 221.
21. B.B. Brandt, O. Philipsen, H. Wittig, L. Zeidlewicz, PoS **LAT2010** (2010) 172 [arXiv:1008.2143].
22. B.B. Brandt, O. Philipsen, H. Wittig and L. Zeidlewicz, in preparation.
23. B.B. Brandt, O. Philipsen, H. Wittig, L. Zeidlewicz, AIP Conf. Proc. **1343** (2011) 516 [arXiv:1011.6172].
24. F. Burger *et al*, arXiv:1102.4530.
25. R. Sommer, Nucl. Phys. **B411** (1994) 839 [hep-lat/9310022].
26. B. Leder *et al.* [ALPHA Collaboration], PoS **LATTICE2010** (2010) 233 [arXiv:1012.1141].

Materials Science

Robert O. Jones[1,2]

[1] Peter Grünberg Institut, Forschungszentrum Jülich, D-52425 Jülich, Germany
E-mail: r.jones@fz-juelich.de

[2] German Research School for Simulation Sciences
FZ Jülich and RWTH Aachen University, D-52425 Jülich, Germany

The main focus in the category "Materials Science" is on families of materials with technological applications or prospects of such in the future. While most may be viewed as a subgroup of "Condensed Matter", it is striking that the overwhelming majority of the projects in "Materials Science" and of those presented in the proceedings of recent NIC-Symposia involve calculations based on the density functional (DF) formalism. This approach allows one to calculate structures and energy differences (including reaction paths) without requiring the input of experimental information. This is a crucial advantage in studies of *families* of related compounds, e.g. what is the effect of exchanging one element with another of the same group? The availability of leadership computers, such as JUGENE, has expanded greatly the range of materials *and* the length and time scales accessible to such calculations. Nevertheless, such calculations are still limited to some hundreds of atoms and simulation times of a few hundred picoseconds, and this is a serious restriction for some applications.

The discovery of giant magnetic resonance by Peter Grünberg and Albert Fert led to the award of the 2007 Nobel Prize for Physics and the new field of "spintronics". Here the focus is on changes in magnetic (spin) properties, and it has been a very active area in computational materials science for well over a decade. In the contribution of Atodiresei *et al.* it is coupled with another important development, namely the use of single molecules as "molecular rectifiers", leading to the new field of *molecular electronics*. Calculations have been performed for various organic molecules and graphene on metallic surfaces and provide us with a picture of the the bonding and magnetic properties of these systems. It is interesting that the description of the exchange-correlation energy, an essential ingredient in DF calculations, must be extended to include so-called "van der Waals" attractions between closed shell systems.

Graphene is also an essential part of the contribution by Kiss *et al.* The discovery of single layers of hexagonally bonded carbon atoms led to the award of the 2010 Nobel Prize for Physics to Andre Geim and Konstantin Novoselov and to much excitement among materials scientists. Applications to the real world of electronics still lie in the future, and this paper studies the deposition of graphene-like building blocks in silicon (111) surfaces and predicts that well defined two-dimensional patterns will result.

The ground state magnetic properties of transition metal alloy nanoparticles have been studied using DF calculations, and the results are described in the article by Díaz-Sánchez and Pastor. The size of the clusters (CoRh and CoPt alloys with up to 531 atoms and a diamter of 2 nm) emphasize both the dramatic improvements that have resulted from the availability of computers such as JUGENE *and* the limitations that remain. An applica-

tion of DF calculations in another context, ion-permeable membranes, is described in the contribution of Lumeij *et al.* We see here how the energy barrier for oxygen hopping in a five-element alloy of Ba, Sr, Co, Fe, and oxygen can be understood, and an additional application is to proton-conducting membranes.

Phase change materials form the basis of many familiar rewriteable storage devices, including DVD-RW, DVD-RAM, and the Blu-ray Disc. The information is stored in very thin layers of an alloy in microscopic "bits" that can switch rapidly and reversibly between the amorphous and crystalline states. The article by Akola and Jones describes a detailed DF study of the amorphous structure of an alloy of Ag/In/Sb/Te that has found widespread use in the DVD-RW world. It is still not possible to simulate the crystallization for such an alloy for samples with hundreds of atoms, but the work has led to a plausible picture of the process that can be studied further.

The remaining application described here requires far larger samples and longer time scales than are currently accessible to DF calculations. Pastewka and Moseler describe molecular dynamics simulations of friction properties of hydrogenated diamond-like carbon films under high pressure. The hydrogen concentration of such films plays a crucial role in determining these properties.

These short articles provide just a sample of the wide range of materials properties being studied using the facilities of NIC. Much of this work seeks answers to problems originating in technological applications, and there will be many challenging problems in this field for years to come. There is no doubt that the coming generations of supercomputers, as well as the development of new algorithms and programming techniques, will play crucial roles in finding the solutions.

Understanding Molecular Electronics and Spintronics from First Principles Simulations

Nicolae Atodiresei[1], Vasile Caciuc[1], Predrag Lazić[2], and Stefan Blügel[1]

[1] Peter Grünberg Institut and Institute for Advanced Simulation, Forschungszentrum Jülich
52425 Jülich, Germany
E-mail: n.atodiresei@fz-juelich.de

[2] Massachusetts Institute of Technology, Cambridge, 02139 Massachusetts, USA

The ability to reliably describe the electronic properties of molecules adsorbed on metallic surfaces[1] is essential to understand and design the functionality of hybrid organic molecular devices. This strongly depends on the accuracy of the state-of-the-art theoretical methods used to assess the interaction between a molecule or a molecular layer and a substrate of choice. Our theoretical studies show that, with its predictive power, the density functional theory (DFT)[2] provides a framework where a realistic understanding of these hybrid systems can be expected and can be further used to tailor both: the electronic and magnetic properties at hybrid organo-metallic interfaces and the clever design of such molecular-based electronic and spintronic devices[3–5].

1 Introduction

Nowadays it is almost impossible to imagine the structure of modern human civilization without computers controlled by logical operations carried out by microprocessors currently designed in the framework of the complementary metal-oxide-semiconductor (CMOS) technology. Nevertheless, the CMOS technology will inherently reach its physical limits due to the requirement to gradually downscale the active elements of the microchips like transistors or diodes up to a scale where the tunneling processes, for instance, will significantly deteriorate the performance of the CMOS microprocessors[6].

The most promising avenue to overcome the intrinsic limits of the CMOS technology is to use organic molecules as basic functional units in electronic devices. So far, the organic molecules have been successfully used in organic electronic devices like solar cells[7], photovoltaic cells[8], organic light-emitting diodes (OLEDs)[9] or organic field-effect transistors (OFETs)[10] that are already present in our daily life as sensors or flat computer screens. In such organic electronic devices the electronic structure of individual molecules is not directly addressed since their electronic properties are usually related to those of thin film organic crystals.

In 1974 Aviram and Ratner[11] suggested for the first time that a single molecule with a specific functional structure could be used as a molecular rectifier in a way similar to a $p - n$ semiconductor diode. This theoretical prediction is a milestone moment in the development of the molecular electronics field that aims at using the electronic properties of each individual molecule to design and construct functional molecular-based devices.

2 Molecular Electronics

Density functional theory (DFT) is the theoretical tool of choice to analyse and to predict the electronic properties of systems characterized by strong chemical bonds. However, the loosely bounded physical systems[12] represent one of the major challenges for DFT because the effective Kohn-Sham (KS) potential does not exhibit the correct asymptotic behaviour. In particular, the currently used exchange correlation energy functionals like the local density approximation (LDA) or the generalized gradient approximation (GGA) do not properly describe the long range van der Waals (vdW) interactions. For instance, the GGA fails to predict a bonding ground state for a van der Waals system like graphite[13].

One way to circumvent this limitation is to include the dispersion effects in DFT in a semiempirical fashion, in which the total energy of the physical system is a sum of the self-consistent Kohn-Sham energy as obtained from the DFT calculations and semiempirical dispersion correction (DFT-D) which depends on the interatomic distance[14].

Another way to accurately account for the vdW interactions is a first principles approach (vdW-DF)[15] which requires the calculation of a nonlocal correlation energy functional as a post perturbation procedure.

2.1 Chemisorption versus Physisorption at π-Conjugated Molecule-Metal Surface Interfaces

We investigated the role of the heteroatoms on the chemical and the van der Waals interactions for a flat adsorption geometry on the Cu(110) surface of three prototype-electron systems as benzene [Bz, (C_6H_6)], pyridine [Py, (C_5NH_5)], and pyrazine [Pz, ($C_4N_2H_4$)] molecules. Our *ab initio* simulations reveal that, for the Bz, the long-range dispersion effects are basically important only for the adsorption energy. This is not the case of the Py and Pz, where due to their low lying-orbitals, the inclusion of the long-range correlations drastically influence the adsorption geometry and electronic structure; i.e., Py becomes *chemisorbed* on the surface while the Pz binds to the surface mostly due to the van der Waals interaction. To evaluate the dispersion effects, we used the *ab initio* (vdW-DF)[15] as well as a semiempirical (DFT-D)[14] method.

In order to deeply understand how the correlation effects contribute to the molecule-metal interaction, in Fig. 1 we plotted the binding energies due to the nonlocal (NL) correlation effects as calculated within the vdW-DF theory[15] implemented in JuNoLo code[16]. For the Bz and Py molecules, the nonlocal contribution is mainly localized in the regions where the C atoms are situated on top of Cu atoms of the surface. In the case of Pz, the contribution of the nonlocal correlation effects is much more delocalized over the entire molecule. From the electronic structure point of view, this peculiar feature is related to a larger contribution to the NL part of the binding energy of the σ_1 molecular orbital in the case of the Pz with respect to the Py molecule. Note that for Bz, this contribution is not significant since the σ_1 lies deep in energy with respect to the Fermi energy of the molecule-surface system. Therefore, which molecular orbitals will contribute to the vdW interactions clearly depends on the specific alignment of these orbitals at the molecule-surface interface[17]. Our study also clearly pointed out that the van der Waals interactions depends on the alignment of the molecular orbitals at adsorbate-substrate interface and can involve not only π-like orbitals (perpendicular to the molecular plane) but also σ-like orbitals (in the molecular plane).

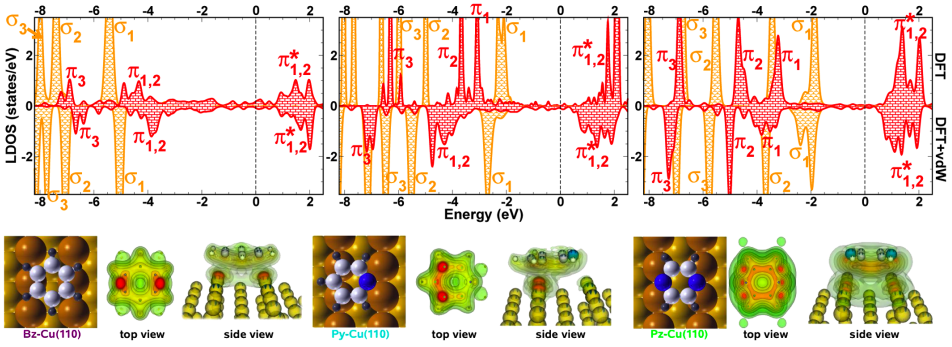


Figure 1. Upper panel: Local density of states (LDOS) of the Bz, Py, and Pz molecules adsorbed on Cu(110). In the case of Bz molecule, the electronic structure of the molecule-surface system is not significantly modified by the inclusion of the dispersion effects. This is not the case of the Py- and Pz-Cu(110) systems, where the van der Waals attractive interaction brings the molecule closer to the surface and allows a stronger hybridization between the p_z-orbitals and the Cu d-bands. Lower panel: Binding energies due to nonlocal correlation effects evaluated with the vdW-DF method. Red (dark grey) colour indicates a strong nonlocal (NL) contribution to the binding energy while green (light grey) denotes a much weaker contribution. In the case of the Pz molecule, this contribution is delocalized over the whole heterocycle ring while in the case of the Bz and Py, it mainly arises from the two carbon sites that directly bind to the surface.

2.2 Graphene on Ir(111): Physisorption with Chemical Modulation

Epitaxial growth on metals is a key method to produce high quality graphene on large scales[18]. Due to the strength of the C-C bonds, large (incommensurate or weakly commensurate) superstructures are found for lattice-mismatched systems. The extent of superperiodicity in the electronic structure[19] and buckling of the carbon layer[20] depends on the strength and local variation of the interaction between graphene and substrate, which ranges from strong chemisorption to weak physisorption[21].

In density functional theory (DFT) a quantitative description of the interaction between graphene and metal is a challenge, because the most commonly used exchange-correlation functionals [local density approximation (LDA) and generalized gradient approximation (GGA)] are (semi)local[22] and lack the nonlocal-correlation effects responsible for van der Waals (vdW) interaction[23].

Using a realistically in-plane large supercell [(10×10) unit cells of graphene on (9×9) cells of Ir(111)] and applying the nonlocal van der Waals density functional approach, we calculated the binding of graphene to Ir(111). We find bonding of graphene to Ir(111) to be due to the van der Waals interaction with an antibonding average contribution from chemical interaction. Despite its globally repulsive character, in certain areas of the large graphene moire unit cell charge accumulation between Ir substrate and graphene C atoms is observed, signalling a weak covalent bond formation.

Our theoretical results are in excellent agreement with x-ray standing wave (XSW) experiments and allows us to unambiguously specify the average bond distance $\overline{h}$ and the amplitude of buckling Δh. The geometry resulting from this calculation is shown in Fig. 2. The largest height of 3.62 Å is found in regions where the centre of the C hexagon is located on top of an Ir atom (top region) and the lowest height of 3.27 Å in the hcp region (centre of C hexagon above threefold coordinated hcp site), slightly lower than the

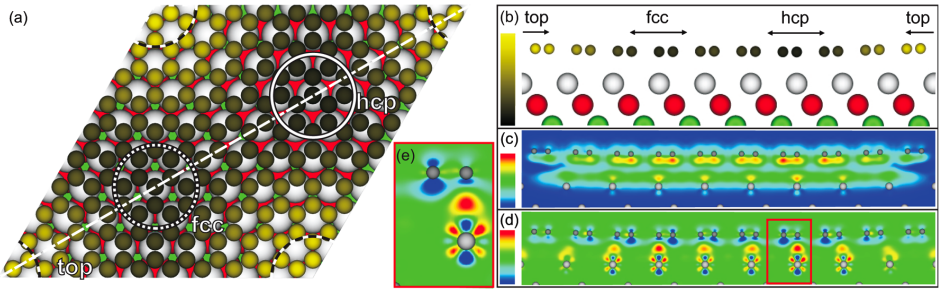


Figure 2. (a) Top view and (b) side view [cut along the dashed line in (a)] of the relaxed structure of graphene/Ir(111) obtained by density functional theory including van der Waals interactions. Regions of high-symmetry stacking (fcc, hcp, top) are marked by circles (a) or arrows (b)(d). The colour scale in (b) and (a) ranges from $\overline{h}$ 3.27 Å (dark) to $\overline{h}$ 3.62 Å [light grey (yellow)]. (c) Visualization of the nonlocal-correlation binding-energy density caused by adsorption. The colour scale ranges from 0.0 meV Å $^{-3}$ [dark grey (blue)] over light grey (green) to -28.3 meV Å $^{-3}$ [medium grey (red)]. (d) Charge transfer upon adsorption. The colour scale ranges from -0.0138 Å $^{-3}$ [dark grey (blue)] over light grey (green) to +0.0130 Å $^{-3}$ [medium grey (red)]. A negative value indicates loss of electron density. (e) Magnified view of red box in (d).

one of 3.29 Å in the fcc region. The mean height $\overline{h}$ is 3.41 Å is in excellent agreement with the experimental value. The corrugation resulting from DFT including vdW is Δh is 0.35 Å and thus safely within the upper limits determined by experiment for several coverages. Therefore, the comparison of experiment and theory allows us to benchmark the functionals used to describe bonding of graphene on Ir(111)[24] and the bonding of π-conjugated systems on metals in general.

3 Molecular Spintronics

Combining molecular electronics with spintronics represents one of the most exciting avenues in building future molecular nanospintonics devices[25,26]. For example, widely used in spintronic applications, the spin valve[27] is a layered structure of two ferromagnetic electrodes separated by a nonmagnetic spacer to decouple the two electrodes and allows spin-polarized electrons to travel through it. The efficiency of a spin valve depends on the spin transport throughout and the spin injection into the nonmagnetic spacer. Organic molecules are made of light elements (as C and H) with weak spin-orbit coupling. Therefore, spin transport properties (spin coherence with time and distance) are improved compared to conventional semiconductors present in todays devices[28]. A crucial issue that remains is the spin injection efficiency at the ferromagnetic-organic interface[29]. Understanding the relevance of interface hybrid states for the transmission of spin information to tailor the electronic and magnetic structure of such interfaces will help to overcome the significant spin loss observed in spintronic devices[30].

3.1 Adsorption of Non-Magnetic Molecules on Ferromagnetic Surfaces

We have performed a theoretical conceptual study to understand the local spin polarization present at the interface for single organic, even diamagnetic molecules adsorbed on a

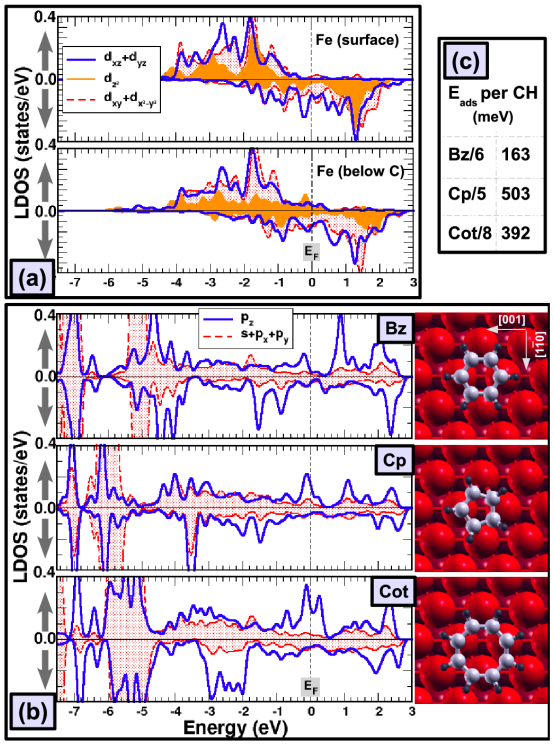


Figure 3. (a) Spin-resolved local density of states of an Fe atom of the clean surface (upper panel) and an Fe atom below the C atom of the Cot molecule (lower panel); (b) adsorption geometries and the spin-resolved local density of states of the Bz, Cp and Cot molecules adsorbed on the 2ML Fe/W(110); (c) the adsorption energies of the Bz, Cp and Cot are given in meV per CH group of atoms. Compared to Bz, Cp and Cot molecules interact strongly with the surface due to an effective hybridization between out-of-plane orbitals (p_z of C and d_{z^2} , d_{xz}, d_{yz} of Fe), while the in-plane orbitals are weakly interacting (s, p_x, p_y of C and/or $d_{x^2-y^2}$, d_{xy} of Fe). All adsorbed molecules show a general characteristic: the energy dependent spin polarization, i.e., in a given energy interval the number of spin-up and spin-down electrons is unbalanced. For this specific interval the molecule has a net magnetic moment delocalized over the molecular plane, although the total magnetic moment of the molecule is $0.0\mu_B$. Note also the increased weight of the states crossing the Fermi level situated in the spin-down (↓) channel at metal site and in the spin-up (↑) channel at molecular site. Therefore, at the molecular site, an inversion of the spin polarization occurs with respect to the ferromagnetic surface.

ferromagnetic surface. By means of spin-polarized *ab initio* simulations we calculated the adsorption of prototype conjugated organic molecules on a ferromagnetic 2 ML Fe/W(110) surface, a well established system in spin-polarized scanning tunnelling microscopy (SP-STM) experiments[31]. We selected organic molecules containing π-electron systems like benzene [Bz, (C_6H_6)], cyclopentadienyl radical [Cp, (C_5H_5)], and cyclooctatetraene [Cot, (C_8H_8)] because they are representative of classes of organic molecules with significantly different reactivities[32]. The Bz molecule adsorbs with two C atoms on top of two neighbouring Fe atoms along the [001] direction while the other four C atoms sit in bridge positions between Fe atoms of adjacent [001] rows. The Cp molecule adsorbs in a similar

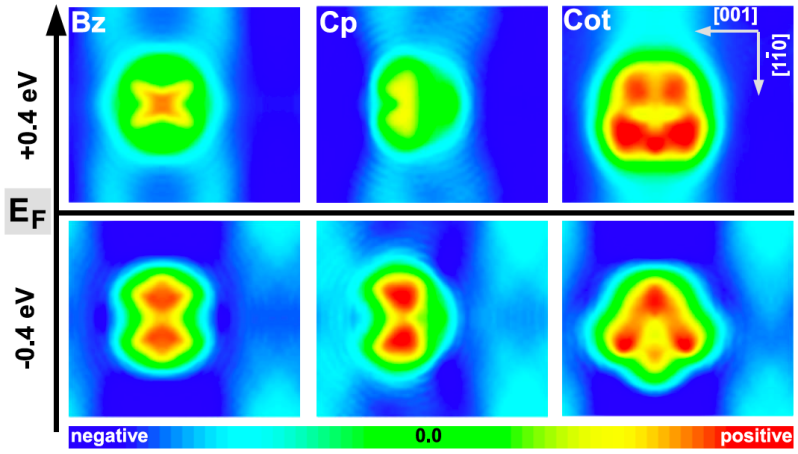


Figure 4. The spin polarization at 2.5 Å above the Bz, Cp and Cot molecules adsorbed on the 2ML Fe/W(110) surface plotted for occupied [0.4 : 0.0] eV and unoccupied [0.0 : +0.4] eV energy intervals around the Fermi level. All the organic molecules show a high, locally varying spin polarization ranging from attenuation to inversion with respect to the ferromagnetic Fe film.

geometry. With one C atom less than the Bz, Cp binds to the surface with a C atom and a C-C bond on top two neighbouring Fe atoms of the [001] row. The Cot molecule binds to the surface with two C atoms and a C-C bond on top of Fe atoms, while the other C atoms sit in bridge sites between Fe atoms of adjacent [001] rows. As a general characteristic of the adsorption geometry, we note that all adsorbed molecules have a nonplanar structure in which the H atoms are situated above C atoms (0.35 up to 0.45 Å). The shortest C-Fe bond is about 2.1S Å and corresponds to the C atoms sitting directly on top of Fe while the C atoms situated in bridge positions are 2.3 up to 2.5 Å away from Fe atoms. Furthermore, each of the molecules is nonmagnetic[33] upon adsorption on the ferromagnetic surface.

A general picture of the binding mechanism between Bz, Cp, and Cot molecules and the ferromagnetic surface can be extracted from the analysis of the spin-resolved local density-of-states (LDOS) of the calculated molecule surface systems shown in Fig. 3. In the spin-up channel, the p_z atomic type orbitals which originally form the π-molecular orbitals hybridize with the majority d-states of the Fe atoms forming molecule-metal hybrid states with bonding and antibonding character. The bonding states are situated at low energies while the antibonding states appear at much higher energies, more precisely in an energy window situated around the Fermi level. Because of the $p_z - d$ interaction in the spin-up ($\uparrow$) channel, the spin-down ($\downarrow$) p_z atomic type orbitals are lowered in energy and slightly hybridize with the minority d states of Fe. As a consequence, the states with large weight around the Fermi level are in the spin-up channel at the molecule site and in the spin-down channel at clean metal sites.

The spin-polarized LDOS of the adsorbed Bz, Cp, and Cot molecules show a very intriguing feature: the energy dependent spin polarization. As clearly visible in Fig. 3(b) for a given energy interval the number of spin-up and spin-down states is unbalanced. For that specific interval the molecule has a net magnetic moment delocalized over the molecular plane since it is carried mostly by the p_z states. Even more interesting, as depicted in the

LDOS, around the Fermi level the states with large weight are in the spin-up channel at the molecular site, while on the clean Fe surface these states are in the spin-down channel. As a consequence, at the molecular site an inversion of the spin polarization[34] occurs with respect to the ferromagnetic surface. This effect is clearly seen in Fig. 4 which illustrates the spin polarization at 2.5 Å above the molecule in the energy intervals below [0.4 : 0.0] eV and above [0.0 : +0.4] eV the Fermi level. A common characteristic for all the molecule-ferromagnetic surface systems is the high and locally varying spin polarization ranging from attenuation to inversion with respect to the ferromagnetic surface. Since the carbons p_z atomic orbitals of the Cp and Cot molecules are strongly interacting with the d states of the iron atoms, as compared to Bz molecule, an amplification of the molecules local spin polarization below Fermi level occurs. Besides this, the presence of a large number of pz electrons in the Cot molecule causes an amplification also above the Fermi level because a larger number of antibonding $p_z - d$ hybrid states are generated. The experimental observation of the local spin polarization above an industrially relevant diamagnetic macromolecule demonstrates our theoretical predictions[4].

3.2 Adsorption of Magnetic Molecules on Ferromagnetic Surfaces

The magnetic organic molecule are very promising candidates as working units for molecular spintronics[28]. We show that a significant spin polarization exists when a cobalt-phthalocyanine (CoPc) molecule is brought in contact with a ferromagnetic Fe thin film, although due to molecule-substrate hybridization the molecule loses its net spin. As confirmed by spin-polarized scanning tunnelling microscopy (SP-STM) experiments[5], an energy- and site-dependent spin polarization from inversion to amplification is resolved on the submolecular scale. State-of-the-art density functional theory theory (DFT), which includes the decisive role of van der Waals (vdW) interactions, reveals both the magnetic and electronic nature of the molecule coupled to the ferromagnetic substrate. Even though the net spin of the molecule is lost due to a transfer of one electron, spin splitting is recovered through the local bonding of molecular orbitals with Fe $3d$-bands.

Fig. 5 depicts the change in electronic structure of a CoPc with increasing molecule-substrate interaction. Fig. 5(a) shows the spin-polarized density of states (SP-LDOS) for the free molecule. In agreement with previous work the origin of spin splitting is an unpaired electron in a molecular orbital with d_{z^2} contribution situated at the Co site and the total molecular spin S is $1/2$. Fig. 5(b) presents the effect of the surface, when van der Waals forces are neglected during the relaxation process. The molecule adsorbs 3.1 Å above the surface remaining flat. The SP-LDOS shows a hybridization of molecular orbitals (MOs) and the substrate $3d$ states. According to the spatial extension of the MOs perpendicular to the molecular plane (i.e., π with character), the hybridization of MOs containing the Co d_{z^2} atomic orbital is the strongest, followed by those including d_π and p_z atomic contributions. The d_σ and σ MOs are only slightly broadened, compared to the free molecule case, as they are localized in the molecular plane. The spin splitting of the molecule-surface hybrid states is reduced due to a transfer of an electron from the substrate to the CoPc. As a result, the formerly unoccupied d_{z^2} type MO becomes occupied and the total molecular spin is quenched ($0\mu_B$).

The role of van der Waals forces is crucial as it brings the molecule 0.5 Å closer to the surface [Fig. 5(b)] and distorts the molecular geometry [Fig. 5(d)]. This new adsorption geometry has a drastically different electronic structure due to the overall changes of

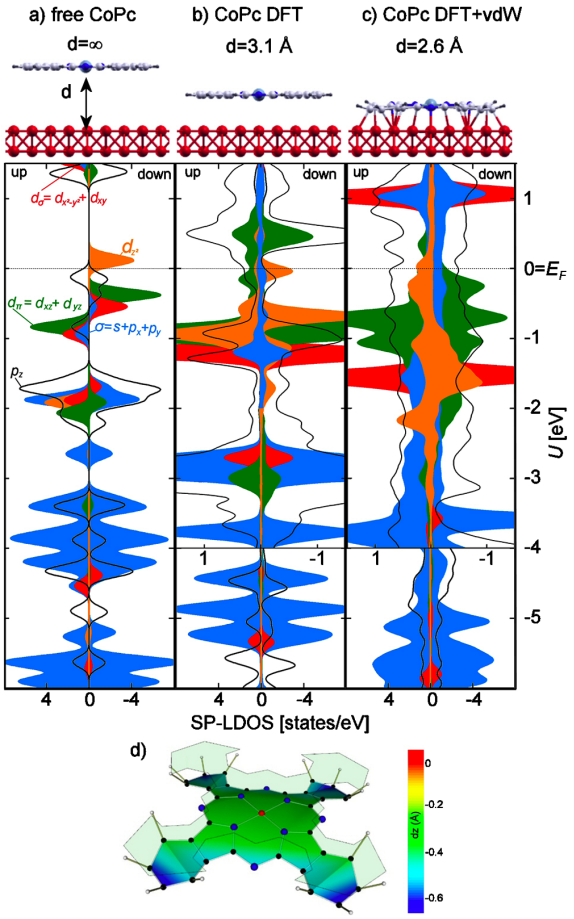


Figure 5. The geometry and electronic structure for a free CoPc (a) and adsorbed on an Fe surface (b) without and (c) with vdW forces included during the relaxation. (d) Molecular deformation due to vdW forces. The C_6 rings and two of the outer N are twisted by up to 0.3 Å toward the surface and away from the plane defined by the Co ion and four inner N. All H are pointing away from the surface.

hybridization between molecule and substrate. The SP-LDOS [Fig. 5(c)] shows not only MOs with a π character strongly hybridize with spin-polarized Fe $3d$ states of the same symmetry to form broad spin-split bands, but also σ type MOs are significantly affected by the interaction with the surface. Again, a transfer of one electron from the surface to the molecule annihilates the molecular spin, but spin splitting is recovered due to the local molecule-surface bonding at different parts of the molecule. The newly formed molecule-surface hybrid states have, within a specific energy intervals an unbalanced, locally varying electronic charge in the spin-up and spin-down channels. For a direct comparison of the first principles calculations with constant-current SP-STM images, isocharge surfaces above the CoPc are extracted from the spatial variation of the energy integrated SP-LDOS. This approach mimics the experimental situation of a local and spin-sensitive tip probing

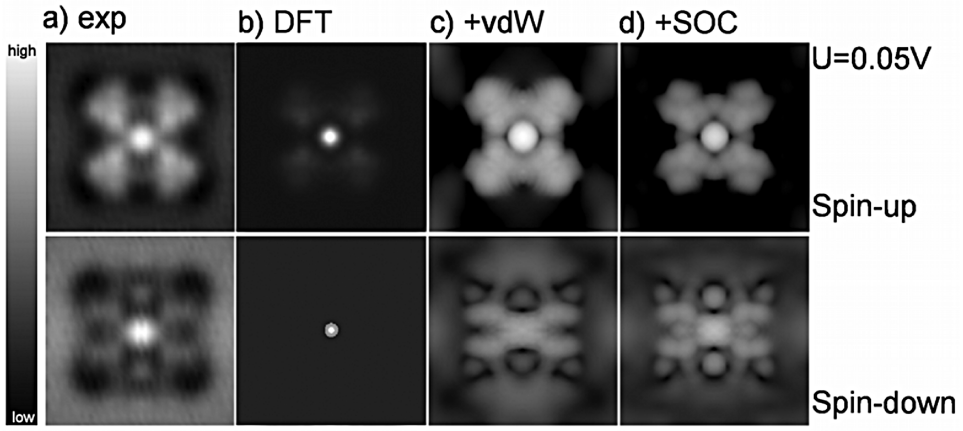


Figure 6. Experimental and simulated SP-STM images in an energy interval [0.0 : 0.05] eV for both spin-up and spin-down directions. (a) Averaged experimental images[5]. (b)(d) Isocharge surfaces; (b) conventional DFT (GGA) (without vdW), (c) DFT+vdW, (d) DFT+vdW and including SOC.

the charge density above the surface at constant current and thereby, accounts for the variation of decay lengths of the different states into vacuum. Fig. 6 gathers calculated spin-dependent isocharge surfaces and experimental SP-STM images for an energy interval [0.0 :0.05] eV. The isocharge surfaces reproduce all features of the experimental SP-STM images[5], only if van der Waals forces are taken into account [Fig. 6(c)]. Corrections due to the inclusion of spin-orbit coupling effects are minor [Fig. 6(d)].

4 Conclusions

To conclude, we prove that the van der Waals dispersion effects together with the corresponding vdW attractive forces are crucial to reliably calculate in a self-consistent manner the proper equilibrium adsorption geometry and the corresponding electronic structure of π-conjugated systems on metal surfaces.

The inclusion of the long-range dispersion effects changes qualitatively the adsorption process of pyridine on Cu(110) surface from $physisorption$ to $chemisorption$, while for the pyrazine, the van der Waals interactions represent the driving forces which bind the molecule to the surface. Our study also clearly pointed out that the vdW interactions depends on the alignment of the molecular orbitals at adsorbate-substrate interface and can involve not only π-like orbitals (perpendicular to the molecular plane) but also σ-like orbitals (in the molecular plane).

In the case of graphene adsorption on Ir(111) surface, the bonding mechanism is dominated mostly by the van der Waals interaction, but surprisingly we find that locally, a polar covalent-like chemical interaction occurs. This local bonding arises from the charge transfer from the graphene to metal surface and the subsequent charge accumulation at the interface in the fcc- and hcp- regions of the moiré superstructure.

Furthermore, we demonstrated that the spin polarization of a ferromagnetic surface can be locally tailored by flat adsorbing organic molecules containing $\pi(p_z)$-electrons onto it.

Although adsorbed molecules are nonmagnetic, due to an energy dependent spin polarization, in a given energy interval the molecules have a net magnetic moment delocalized over the molecular plane. The generality of the concept predicted by the theoretical calculations on model systems is supported our SP-STM experiments on phthalocyanine and cobalt-phthalocyanine molecules adsorbed on a ferromagnetic surface like 2ML Fe/W(110) surface which clearly show the inversion of the local spin polarization near the Fermi level even at larger distances. Our observations imply that further progress in the field of molecular spintronics does not exclusively rely on tailoring specific magnetic molecular properties but also a clever design of spin-active hybrid interfaces is equally important. Our first principles studies demonstrate that electrons of different spin [i.e., up ($\uparrow$) and down ($\downarrow$)] can selectively be injected from the same ferromagnetic surface by locally controlling the inversion of the spin polarization close to the Fermi level, an effect which can be exploited to increase the efficiency of future molecular spintronic devices.

Acknowledgements

This work is funded by the DFG SPP1243. The computations were performed on JUROPA and JUGENE supercomputers at the Jülich Supercomputing Centre, Forschungszentrum Jülich (Germany).

References

1. G. Heimel, L. Romaner, J.-L. Brédas and E. Zojer, Phys.Rev. Lett. **96**, 196806, 2006.
2. W. Kohn and L. J. Sham Phys. Rev. **140**, A1133, 1965.
3. N. Atodiresei, P. H. Dederichs, Y. Mokrousov, L. Bergqvist, G. Bihlmayer and S. Blügel, Phys. Rev.Lett. **100**, 117207, 2008.
4. N. Atodiresei, J. Brede, P. Lazić, V. Caciuc, G. Hoffmann, R. Wiesendanger and S. Blügel Phys. Rev.Lett. **105**, 066601, 2010.
5. J. Brede, N. Atodiresei, S. Kuck, P. Lazić, V. Caciuc, Y. Morikawa, G. Hoffmann, S. Blügel and R. Wiesendanger, Phys. Rev.Lett. **105**, 047204, 2010.
6. D. J. Frank and Y. Taur, Solid-State Electronics **46**, 325, 2002.
7. J. Peet, J. Y. Kim, N. E. Coates, W. L. Ma, D. Moses, A. J. Heeger and G. C. Bazan, Nature Mat. **6**, 497, 2007.
8. L. Schmidt-Mende, A. Fechtenktter, K. Mllen, E. Moons, R. H. Friend and J. D. MacKenzie, Science **293**, 1119, 2001.
9. H. Sirringhaus, N. Tessler and R. H. Friend, Science **280**, 1741, 1998.
10. H. Yan, Z. Chen, Y. Zheng, C. Newman, J. R. Quinn, F. Dtz, M. Kastler and A. Facchetti, Nature **457**, 679, 2009.
11. A. Aviram and M. A. Ratner, Chem. Phys. Lett. **29**, 277, 1974.
12. L. W. Bruch, R. D. Diehl, and J. A. Venables, Rev. Mod. Phys. **79**, 1381, 2007.
13. N. Ooi, A. Rairkar, and J. B. Adams, Carbon **44**, 231, 2006.
14. S. Grimme, J. Comput. Chem. **27**, 1787, 2006.
15. M. Dion, H. Rydberg, E. Schröder, D.C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401, 2004.
16. P. Lazić, N. Atodiresei, M. Alaeia, V. Caciuc, St. Blügel, R. Brako, Comput. Phys. Commun. **181**, 371, 2010.

17. N. Atodiresei, V. Caciuc, P. Lazić and S. Blügel, Phys. Rev. Lett. **102**, 136809, 2009.
18. J. Coraux, A. T. NDiaye, C. Busse, and T. Michely, Nano Lett. **8**, 565, 2008.
19. I. Pletikosić, M. Kralj, P. Pervan, R. Brako, J. Coraux, A. T. NDiaye, C. Busse and T. Michely, Phys. Rev. Lett. **102**, 056808, 2009.
20. A. T. NDiaye, S. Bleikamp, P. J. Feibelman, and T. Michely, Phys. Rev. Lett. **97**, 215501, 2006.
21. A. B. Preobrajenski, M. L. Ng, A. S. Vinogradov, and N. Martensson, Phys. Rev. B **78**, 073401, 2008.
22. In LDA, the functional depends on the charge density $\rho(r)$, whereas in GGA it also depends on its gradient $\nabla\rho(r)$.
23. R. Brako, D. Sók'cević, P. Lazić, and N. Atodiresei, New J. Phys. **12**, 113016, 2010.
24. C. Busse, P. Lazić, R. Djemour, J. Coraux, T. Gerber, N. Atodiresei, V. Caciuc, R. Brako, A. T. NDiaye, S. Blügel, J. Zegenhagen and Thomas Michely, Phys. Rev. Lett. **107**, 036101, 2011.
25. G. A. Prinz, Science **282**, 1660, 1998.
26. V. A. Dediu, L. Hueso, I. Bergenti and C. Taliani, Nature Mater. **8**, 707, 2009.
27. A. J. Drew, J. Hoppler, L. Schulz, F. L. Pratt, P. Desai, P. Shakya, T. Kreouzis, W. P. Gillin, A. Suter, N. A. Morley, V. K. Malik, A. Dubroka, K. W. Kim, H. Bouyanfif, F. Bourqui, C. Bernhard, R. Scheuermann, G. J. Nieuwenhuys, T. Prokscha and E. Morenzoni, Nature Mater. **8**, 109, 2009.
28. A. R. Rocha, V. M. García-suárez, S. W. Bailey, C. J. Lambert, J. Ferrer and S. Sanvito, Nature Mater. **4**, 335, 2005.
29. C. Barraud, P. Seneor, R. Mattana, S. Fusil, K. Bouzehouane, C. Deranlot, P. Graziosi, L. Hueso, I. Bergenti, V. Dediu, F. Petroff and A. Fert, Nature Physics **6**, 615, 2010.
30. S. Majumdar, H. S. Majumdar, R. Laiho and R. Österbacka, New J. Phys. **11**, 013022, 2009.
31. A. Kubetzka, M. Bode, O. Pietzsch and R. Wiesendanger, Phys. Rev. Lett. **88**, 057201, 2002.
32. (a) I. Fleming, *Orbitals and Organic Chemical Reactions* (JohnWiley & Sons, New York, 1978); (b) F. A. Carey, *Organic Chemistry*, (McGraw-Hill, New York, 2000).
33. Integrating the spin-resolved LDOS for each molecule yields very small magnetic moments, i.e., $0.076\mu_B$ for Bz, $0.031\mu_B$ for Cp and $0.003\mu_B$ for Cot because of a slightly higher number of electrons in spin-down channel($\downarrow$). These magnetic moments can be practically neglected and the adsorbed molecules can be considered nonmagnetic (i.e., $0.0\mu_B$).
34. The spin polarization is defined as $(n^{\uparrow} - n^{\downarrow})/(n^{\uparrow} + n^{\downarrow})$, where $n^{\uparrow}$ and $n^{\downarrow}$ are the spin-up ($\uparrow$) and spin-down ($\downarrow$) charge densities in a given energy interval.

Quantum-Chemical Calculations on Ion-Permeable Membranes

Marck Lumeij, Michael Gilleßen, Bernhard Eck, and Richard Dronskowski

Institute of Inorganic Chemistry, RWTH Aachen University, 52056 Aachen, Germany
E-mail: drons@HAL9000.ac.rwth-aachen.de

1 Introduction

The HGF Alliance "MEM-BRAIN"[1] started in October 2007 with the aim to develop, test and optimize novel gas separation membranes for coal-fired power plants. MEM-BRAIN is an integrated research project in which fundamental researchers, membrane developers and engineers combine their expertise to reach this goal. Our task in the project is to fundamentally understand the processes which take place at an atomistic scale and to help optimizing the membrane material preparation through quantum-chemical calculations. A prototype of an oxygen-conducting material is, for example, $Ba_{0.5}Sr_{0.5}Co_{0.8}Fe_{0.2}O_{3-\delta}$ which is also called BSCF5582. In particular, we investigated the energy barrier for oxygen-hopping processes. Moreover, we will concentrate on the description of the proton-hopping mechanisms within complex rare-earth/yttrium tungsten oxides.

2 Results and Discussion

2.1 Oxygen Conductors: Vacancies and Energy Barriers

BSCF5582 has a rather complex stoichiometry which originates from empirical optimization of the oxygen flux and the structural stability. The structure of the material follows the common perovskite type as shown in Fig. 1. The regular perovskite structure ABO_3 is relatively simple and contains not more than five atoms in its crystallographic unit cell. In contrast, the construction of a unit cell which corresponds to the BSCF5582 composition necessitates the use of significantly larger unit cells as a reasonable computational starting point. To deal with the complexity of BSCF and to generate a suitable input model for our calculations we use the so-called supercell approach. In this case the small perovskite unit cell is doubled in all three directions. This supercell contains eight times more atoms than the crystallographic unit cell. The supercell in Fig. 1 contains 40 atoms while Sr^{2+}/Ba^{2+} and Co^{4+}/Fe^{4+} share the same crystallographic site. Also, the octahedral surrounding of the "*B*-atoms" and their connectivity through shared corners is emphasized in Fig. 1. In contrast to common crystallographic structure refinements, where one site can be occupied by different atoms, we have to specify each single atom on one position in order to obtain a computable starting configuration in which the composition is close to the experiment. Therefore we used the supercell from Fig. 1, shifted the origin by $(\frac{1}{4},\frac{1}{4},\frac{1}{4})$ and specified the atoms. In addition BSCF5582 requires vacancies within the oxygen sublattice

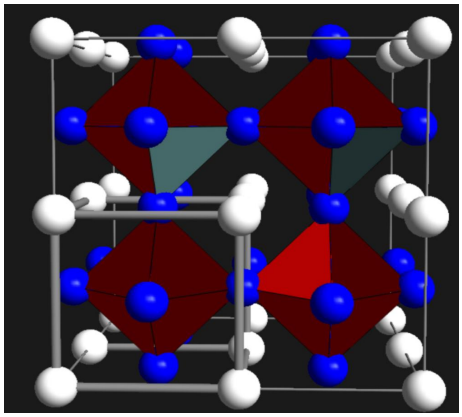


Figure 1. Example of a perovskite supercell, constructed from eight small unit cells. White atoms indicate the position of strontium and barium, the iron/cobalt positions are in red (centre of octahedron) and oxygen is in blue.

to facilitate oxygen conductivity. In a first approximation we removed one oxygen atom in the computational supercell (white square in Fig. 2) which leads to a composition of $Ba_{0.5}Sr_{0.5}Co_{0.875}Fe_{0.125}O_{2.875}\square_{0.125}$. Since there are four different cations in BSCF5582 and oxygen is always octahedrally surrounded by 2+4 cations, there are many possible positions for the oxygen vacancy. Since the high symmetry of the starting perovskite structure is broken by the removal of one oxygen atom, significant structural deformations occur after computational relaxation. This is also visualized in Fig. 2 where the largest distortion is, always found as expected, close to the oxygen vacancy. Two results of these calculations are shown in Tab. 1. The total energy calculations, however, clearly show that the oxygen is strongly attracted to iron and, therefore, the vacancies are always near the cobalt atoms.

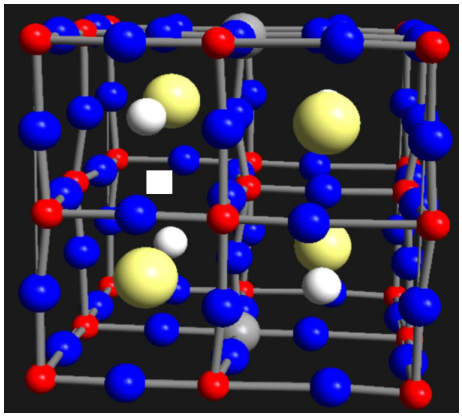


Figure 2. Example of a BSCF supercell ($Ba_{0.5}Sr_{0.5}Co_{0.875}Fe_{0.125}O_{2.875}\square_{0.125}$), containing 39 atoms after structural relaxation; strontium atoms are given in white, barium in yellow, cobalt in red, iron in grey and oxygen in blue. The white square indicates the oxygen vacancy.

Configuration	Energy (eV/cell)
Co–□–Co	−254.81
Fe–□–Co	−254.53

Table 1. Calculated total energies for $Ba_{0.5}Sr_{0.5}Co_{0.875}Fe_{0.125}O_{2.875}$ with one oxygen vacancy, located at two different positions.

The influence of barium and strontium in the cation arrangement, as shown in Fig. 2, is minor[2].

The next step is to let one oxygen atom move directly into the created vacancy. Therefore the starting and the final configurations for such a hopping process need to be structurally optimized. In Fig. 3 the octahedral environment of the oxygen atom at the beginning and at the final point of the atomic migration has been depicted. Within the used method the path is split into several images, and the energy of each image is calculated. In the here implemented migration path the oxygen atom moves through a trigonal plane which is shared by two adjacent octahedra. In order to calculate the transport properties of BSCF, a set of seven images between the starting and final point was generated. A more detailed view on this motion, especially focusing on the penetration of the oxygen atom through the trigonal plane into the neighbouring octahedral void, is also shown in Fig. 3. The energy of each single image is correctly calculated without any human intervention but the images are coupled with each other through so-called artificial spring forces which simply prevent that an oxygen atom moving through a high-energy structural configuration falls back into either the starting or the final configuration. The calculated energy profiles corresponding to the oxygen atom moving through these saddle-point configurations are shown in Fig. 4. The comparison of the transition barriers for two different saddle-point configurations, as displayed in Fig. 4, yields that the barrier is clearly larger when iron is

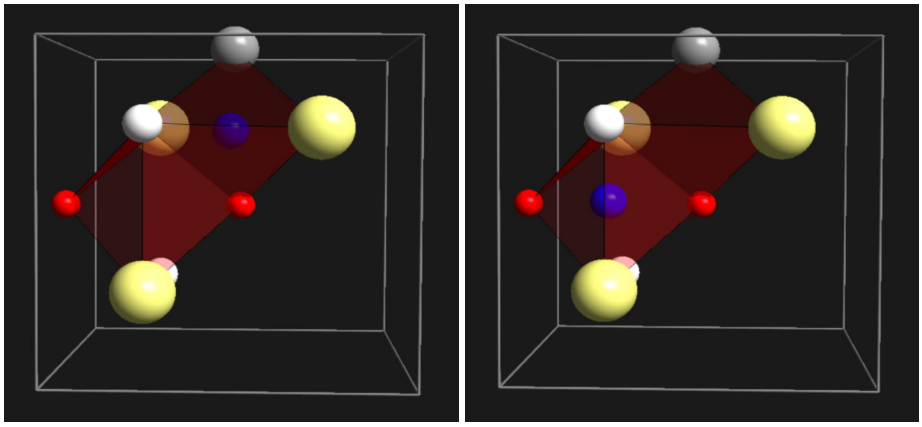


Figure 3. Oxygen hopping from an O-containing octahedron (left) into a free neighbouring vacancy (right) of BSCF5582; iron in grey, cobalt in red, oxygen in blue, strontium in white and barium in yellow. The O atom becomes first coordinated by one Fe and one Co atom (Fe-O-Co), after the hopping by two Co atoms (Co-O-Co).

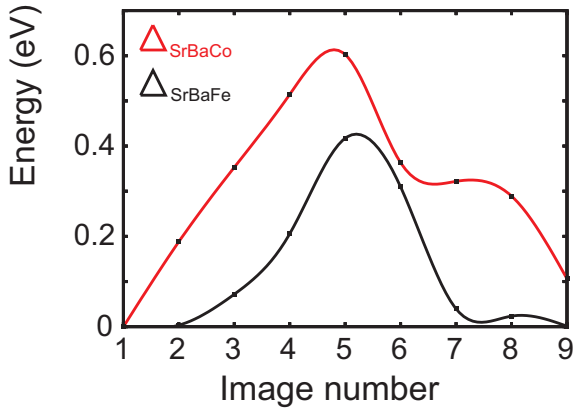


Figure 4. Calculated energies for a jump of an oxygen atom into a vacancy inside $Ba_{0.5}Sr_{0.5}Co_{0.875}Fe_{0.125}O_{2.875}$ for two different saddle-point configurations, designated as Δ_{SrBaFe} and Δ_{SrBaCo}; see also text.

replaced by cobalt; note that the transition energy for Δ_{SrBaCo} is 0.6 eV while for Δ_{SrBaFe} it arrives at 0.4 eV[2]. This is in good agreement with experimental results in which the energy barrier for oxygen-hopping processes in BSCF is reported to be around 0.5 eV[3,4]. The shape of the curve can be explained by an evaluation of the effective coordination number of the O atom according to the recipe of Brunner and Schwarzenbach[5] as this atom passes through the triangle. The maximum in the curves at image 8 is caused by unphysical stress or strain and is an artefact of a pure NEB calculations. Therefore we slightly modified our approach in a way that we perform a structural optimization by means of relaxing the lattice parameters and the atomic sites except those which are free to move during the NEB calculation.

About two years ago a Japanese group[6] proposed an orthorhombic structure of BSCF. The main difference between the common cubic perovskite structure and this orthorhombic structure is, however, not the metric of the unit cell but, as a direct consequence of the lower symmetry of the orthorhombic cell, the nonequivalence of the oxygen sites. Moreover, they suggested a difference in mobility between these two oxygen positions. We took this opportunity to investigate all the possible oxygen pathways in orthorhombic BSCF and to test our available computational methodology for calculating the energy barriers. From the calculations on cubic BSCF we already know that "single" NEB calculations will not lead to accurate activation barriers. In every case a structural relaxation is necessary. The results are shown in Fig. 5. On the left the activation energies depending on the methodology is plotted and the on right the possible oxygen pathways are displayed. The colours of the pathways (red, green, white dotted) correspond to the colours of the calculated activation energies. Note that it was impossible to obtain any results for the white-dotted pathways. We therefore conclude that these specific paths are unlikely. Nonetheless our calculations could, at least partially, confirm the results of Itoh et al.[6] and suggest that some oxygen hopping processes are more likely than others, depending on the cationic surrounding.

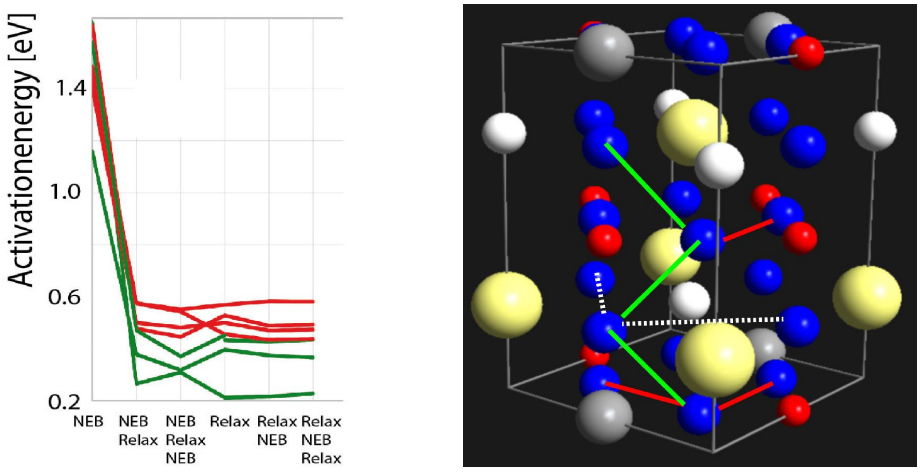


Figure 5. Calculated energies for a jump of an oxygen atom into a vacancy in $Ba_{0.5}Sr_{0.5}Co_{0.875}Fe_{0.125}O_{2.875}$ for two different saddle-point configurations, designated as Δ_{SrBaFe} and Δ_{SrBaCo}; see also text.

2.2 Hydrogen Conductors

The MEM-BRAIN[1] work package "Ceramic Proton Conducting Membranes" aims to develop and optimize new proton-conducting materials and membranes by enhancing the electronic and protonic conduction. Proton-conducting membranes can be used for separating H_2 from CO_2 of the precombustion process in coal-fired power plants. The integrated research on protonic conductors has focused on a new class of crystalline oxide-based conductors, i.e., the system $Ln_{6-x}Ca_xWO_{12}$ (Ln = La, Nd, Eu, Er) which reveals significant protonic and electronic conductivity at high temperatures. Since the exact crystal structures of these compounds have not been fully determined yet, we started our research by taking the archetypical structure of Y_6WO_{12} (see Fig. 6) into account. The calculation of activation energies by using the NEB method did not turn out to be straightforward,

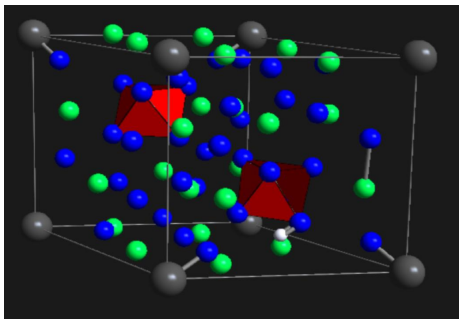


Figure 6. Crystal structure of Y_6WO_{12} and one inserted hydrogen atom which is connected to a tungsten-neighboured oxygen. Yttrium is in green, tungsten in grey, oxygen in blue and hydrogen in white. The octahedral coordination of tungsten by oxygen is also shown.

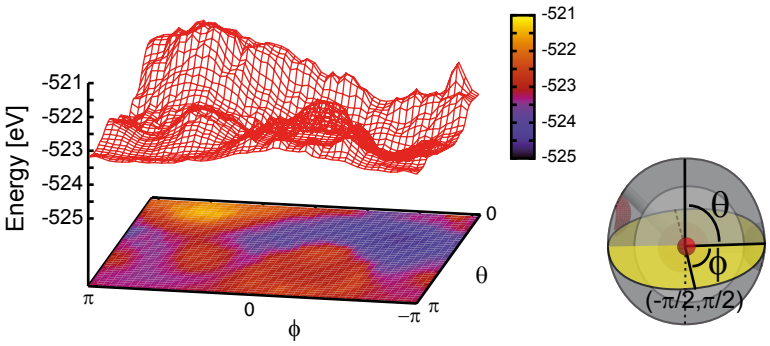


Figure 7. Energy hypersurface of hydrogen rotation around oxygen, which is connected to tungsten.

probably because of the mismatch between weak van der Waals or dispersion forces and the artificial spring forces which connect the individual images. Therefore we used a quite simple, nonetheless computationally demanding methodology where sampling techniques are used to scan real space by single-point calculations. For the calculation of the H–O rotational barrier a sphere, with a fixed radius ($R_{\mathrm{O-H}}$ = 0.99 Å), is put around one oxygen. Hydrogen atoms are arbitrarily put on the surface of this sphere one at a time. This leads to structural models which only differ in hydrogen position. The total single-point energies of all these modifications are calculated and shown as a rotational energy landscape in Fig. 7. To obtain smooth hypersurfaces, a dense sampling grid must be used. In our case we used 500 individual modifications. In Fig. 7 we see areas of high energies, at which the hydrogen atom points towards the cations, and a broad low-energy region in between. From these calculations we can conclude that hydrogen can not freely rotate around oxygen and that the low-lying states are always directed towards neighbouring oxygen atoms. In other cases the energy profile looks quite different, mostly influenced by the surrounding of the oxygen to which hydrogen is connected. These results together with the calculation of hydrogen-hopping barriers, for which we also used this methodology but which are not shown here, can give us detailed insights into the hydrogen pathways through Y_6WO_{12}.

3 Theoretical Methodology

The simulation of chemical reactions and diffusion processes requires the knowledge of the transition states and the minimum energy path (MEP). An efficient way to find the minimum energy path is given by the nudged elastic band (NEB) method[7,8] where a set of several images is constructed between the initial and the final state. The images are connected by spring forces, but the essential feature of the NEB method is its force projection. It ensures that the spring forces do not interfere with the convergence of the elastic band. It is therefore important to estimate the tangent to the path at each image, where the true force and the spring force are decomposed into a parallel and perpendicular element along the path. Only the perpendicular components of the true force and only the parallel components of the spring forces are included. This specific approach of force projection

is termed “nudging”. Since the spring forces between the images are the same, the images are equally distributed along the MEP. In case this distribution is not fine enough, especially around the transition state, the interpolation of the energy becomes inaccurate and the exact transition state cannot be found. A solution to this problem is the climbing image (CI)-NEB method[9]. Within this method, the image with the highest energy (climbing image) is identified and moved up the potential surface, thereby achieving a maximum energy. The climbing image is not affected by the spring forces and, therefore, the spacing between the images may not be the same. The CI-NEB method, when converged, will give a good approximation of the reaction coordinate around the saddle point and is implemented in the Vienna *ab initio* Simulation Package (VASP)[10–13]. The total energies and electronic structures of all phases were calculated in the context of density functional theory. We used plane-wave basis sets and ultra-soft pseudopotentials. The contributions of inter-electronic correlation and exchange to the total energies were treated using the generalized-gradient approximation (GGA)[14]. The kinetic energy cut-off of the plane waves was set to 500 eV.

Acknowledgements

We would like to thank the computing centre at RWTH Aachen University and the Jülich Supercomputing Centre (Forschungszentrum Jülich) for providing us with large amounts of CPU time. Financial support provided by the Helmholtz Association of German Research Centres (Initiative and Networking Fund) through the Helmholtz Alliance MEMBRAIN is gratefully acknowledged.

References

1. HGF-Alliance MEM-BRAIN, Gas separation membranes for zero-emission fossil power plants, Coordinator: Univ. Prof. Dr. Detlev Stöver, Forschungszentrum Jülich, `http://www.mem-brain-allianz.eu`.
2. C. Wessel, M. Lumeij, R. Dronskowski, J. Membr. Sci. **366**, 92, 2011.
3. Z. Shao, S. M. Haile, Nature **435**, 795, 2005.
4. Z. P. Shao, W. S. Yang, Y. Cong, H. Dong, J. H. Tong, G. X. Xiong, J. Membr. Sci. **172**, 177, 2000.
5. G. O. Brunner, D. Schwarzenbach, Z. Kristallogr. **133**, 127, 1971.
6. T. Itoh, Y. Nishida, A. Tomita, Y. Fujie, N. Kitamura, Y. Idemotob, K. Osakac, I. Hirosawac, N. Igawa, Solid Stat. Commun. **149**, 41, 2009.
7. H. Jònsson, G. Mills, K. W. Jacobsen, *Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions, in Classical and Quantum Dynamics in Condensed Phase Simulations*, Ed. B. J. Berne, G. Ciccotti and D. F. Coker, 385 (World Scientific, 1998).
8. G. Henkelman, H. Jònsson, J. Chem. Phys. **113**, 9978, 2000.
9. G. Henkelman, B. P. Uberuaga, H. Jònsson, J. Chem. Phys. **113**, 9901, 2000.
10. G. Kresse, J. Hafner, Phys. Rev. B **47**, 558, 1993.
11. G. Kresse, J. Hafner, Phys. Rev. B **49**, 14251, 1994.
12. G. Kresse, J. Furthmüller, Comput. Mat. Sci. **6**, 15, 1996.
13. G. Kresse, J. Furthmüller, Phys. Rev. B **54**, 11169, 1996.

14. J. P. Perdew, J. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Sing, Phys. Rev. B **46**, 6671, 1992.
15. A. Togo, *fropho V1.1*, (RWTH Aachen University, 2008).
16. R. Stoffel, C. Wessel, M.-W. Lumey, R. Dronskowski, Angew. Chem. Int. Ed. **49**, 5242, 2010.

Graphene Building Blocks on Si (111): Going Towards Graphene-Based Nanoelectronics with the Help of Large Scale *ab initio* Simulations

Janos Kiss, Thomas Gruhn, and Claudia Felser

Institut für Anorganische Chemie und Analytische Chemie, Johannes Gutenberg-Universität
55099 Mainz, Germany
E-mail: felser@uni-mainz.de

Recently the first coronene-based molecular transistor resembling graphene was synthesized with atomic precision. Although this nanodevice has shown outstanding electrical properties similar to its graphene-based counterparts. the switching characteristics were severely degraded by the breaking of chemical bonds between the electrodes and the coronene molecule. Based on large scale first principle calculations performed on parallel computers we show, that by using the Si(111) surface as a substrate for the deposition of molecular nanographene building blocks such problems could be ameliorated. Our results indicate that individual graphene-like building blocks will selectively adsorb on the binding sites of the Si(111) face, leading to well defined two-dimensional patterns. These patterns are controlled by the chemical composition of the adsorbate molecules, which could be transformed to an ordered grid of nanotransistors.

1 Introduction

Graphene[1] is a very promising material which has received prominent scientific attention in the last three years[2]. Due to its outstanding transport properties[3–7], graphene and porous graphene[8–10] has been widely studied for electronic applications. Theoretical studies indicate that graphene-based field-effect transistors (FETs) may be able to operate at very high frequencies up to the terahertz range[11,3]. Therefore, graphene-based FETs are very attractive for industrial applications compared to circuits using complementary metal-oxide semiconductors (CMOS) based on silicon. Because of its improved properties it is expected that graphene, porous graphene and graphene-like molecules will replace silicon[13,3,6,14].

The electronic devices based on graphene or graphene nanoribbons[15–17] (GNRs) so far are using Si or SiC as semiconducting substrate and SiO_2 as insulator[18,11,19,4,1], very similar to CMOS. Due to the uniformity and flatness of graphene, it seems that it can be combined excellently with silicon fabrication technology[3]. Also, one can take full advantage of the already well established procedures from CMOS technology to develop revolutionary new semiconductor devices based on graphene. Despite its superior material properties, however, graphene has a zero bandgap. Furthermore, in case of GNRs the rugged edges pick up undesired contaminants. Due to this problems FETs that are built directly via a top-down approach by cutting out patches from graphene or GNRs have low on/off ratios, which is quite unfavourable in logic circuits.

On the other hand, a molecular nanotransistor with a precisely known atomic structure resembling graphene was build only very recently[20]. The molecular transistor was engineered via a bottom-up approach, where hexa-*peri*-benzocoronene (HBC) molecules were synthesized with atomic precision having four side chains and two linker groups. Still, in

spite of the specifically designed bridges on the HBC molecule, the electric contacts between HBC and the electrodes were rather unstable[20]. The quick attachment/detachment of coronene to/from the electrodes lead to an abrupt switching behaviour, which is very detrimental both in analog and digital electronic circuits. These shortcomings were mostly caused by the nature of the liquid environment, where chemical bonds could be easily broken and new bonds dynamically formed.

2 Motivation

In order to make the industrial fabrication of molecular nanotransistors feasible, individual nanographene building blocks should be first adsorbed on an adequate substrate, where graphene-like molecules can be firmly anchored and also the electrodes can be deposited. Thus, by immobilizing the molecules on a solid surface the electrode-molecule contacts would be stabilized, and the electrodes could be realized via the means of economically viable CMOS technology. Moreover, a selective adsorption of individual graphene-like molecules on the substrate would be extremely beneficial. By following the template of the substrate, the molecules self-assemble to a well defined two-dimensional grid. These molecules could be later transformed to an array of FETs constructed by design with atomic precision. This offers a huge leap forward from conventional fabrication techniques, and it would open a new avenue in nanoelectronics for future applications, where graphene-based devices are mass produced. Based on the results of large scale density functional theory (DFT) calculations performed on parallel computers of the Jülich Supercomputing Centre we will demonstrate that the Si(111) surface is very promising to serve as substrate for the fabrication of coronene-based molecular transistors.

3 Computational Setup

The calculations were carried out in the framework of density functional theory (DFT). For the approximate treatment of electron exchange and correlation we employed the gradient-corrected functional of Perdew, Burke and Ernzerhof (PBE)[26] together with Vanderbilt type ultrasoft pseudopotentials[27] and a plane wave basis set. In all simulations we have performed spin-polarized calculations using a plane wave cutoff energy of 25 Ry together with a density cutoff of 100 Ry. The overall systems contained up to 700 atoms, thus, these calculations belong to the field of large scale simulations.

The Si(111) surface was modelled by a periodically repeated hexagonal slab. The slab consisted of 512 Si atoms arranged in thirteen silicon layers in a $27.13\,\text{Å}\times27.13\,\text{Å}\times35.48\,\text{Å}$ hexagonal simulation cell. The top three layers were build such to represent the 7×7 reconstructed surface (the topmost layer was hosting the 12 adatoms). The other ten layers were build according to the ideal bulk Si structure. The silicon atoms at the bottom of the slabs were passivated by 49 hydrogen atoms with no strain on the Si–H bonds. In order to mimic the underlying bulk, in each calculation the bottom six Si layers were kept fixed at their theoretical bulk positions. For the optimization of the graphene building blocks on the surface the top seven Si layers were allowed to relax together with the adsorbate. The configurations formed via the adsorption of coronene, CHP and HBC on the Si(111) substrate were relaxed by minimizing the

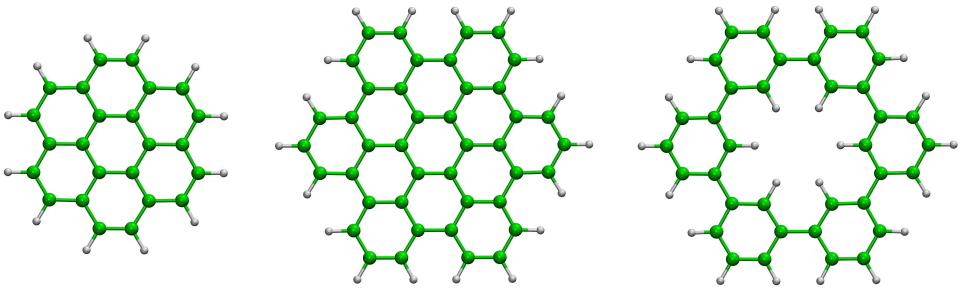
atomic forces acting on the nuclei. Convergence was assumed when the largest component of the residual forces dropped below 0.01 eV/Å. The structural optimizations were carried out employing the CPMD[24,25] software package, which is an efficient electronic structure and molecular dynamics codes based on a plane wave pseudopotential implementation of DFT. CPMD was developed from the ground up as a supermassively parallel code, employing a coarse-grained distributed memory parallelization and a fine grained parallelization by using OpenMP compiler directives. In this way it can achieve outstanding performance on large supercomputer platforms up to the terascale. We found that due to the large system size (700 atoms) and because of the efficient implementation of CPMD the scaling for our model systems was close to linear up to 180 CPU cores.

4 Results and Discussion

Using the setup described above we have analyzed the adhesion behaviour of coronene, hexa-*peri*-benzocoronene (HBC) and cyclohexa-*m*-phenylene (CHP) (see Fig. 1) on the Si(111) substrate. The results presented here correspond to a full year of runtime on 100 CPU cores, which was possible via the grant HMZ24 at the Jülich Supercomputing Centre.

Coronene and HBC are the building blocks of graphene, and CHP is considered to be the building unit of porous graphene. HBC consist of thirteen aromatic rings arranged into a well defined planar honeycomb structure resembling graphene. Due to the extended size of HBC relative to coronene, the former molecule is a good model system for the study of nanographenes. In contrast to the aromatic coronene and HBC, in CHP the C atoms from the inner ring are substituted with hydrogen atoms leading to hexagonal pores in a graphitic skeleton (see Fig. 1c). To reduce the inner tension induced by these crowded H atoms, the skeleton built up by the phenylene groups will adopt a chair conformation in the gas phase, analogous to cyclohexane[21].

In its ground state the Si(111) surface is 7×7 reconstructed[22], consisting of hexagonal surface unit cells. These can be separated into two triangular subcells along the short diagonal. One subcell has a stacking fault compared to the underlying ideal Si(111) layer. We will refer to this subcell as the faulted half (FH) of the unit cell. The other subcell will



(a) (b) (c)

Figure 1. Ball-and-stick representation of the graphene-like adsorbate molecules. Structure of coronene (a), hexa-*peri*-benzocoronene (b), and cyclohexa-*m*-phenylene (c). Carbon and hydrogen atoms are represented with green and white spheres, respectively.

be called the unfaulted half (UH). In the topmost layer there are six Si adatoms located both on the UH and on the FH of the unit cell, arranged in a local 2×2 structure. Based on their position on the surface, the twelve Si adatoms can be divided in four groups: unfaulted corner adatoms (UCo), unfaulted centre adatoms (UCe), faulted corner adatoms (FCo) and faulted centre adatoms (FCe). The UCo and UCe adatoms together with the FCo and FCe adatoms form two distinct triangular islands (see Fig. 2). Due to their dangling bonds, the twelve Si adatoms act as the main binding sites on the 7×7 reconstructed Si(111). In the triangular islands the interatomic Si–Si distances on the Si(111) surface are also very similar to the C–C interatomic distances from graphitic carbon rings. Hence, based on geometrical considerations there are several different structural possibilities how molecular graphene building blocks could potentially bind to the Si(111) surface.

The chemical nature of the interaction between extended graphene sheets or graphene-like molecules and silicon surfaces is rather complex. Particularly interesting results have been found for the graphene-like coronene molecules being adsorbed on the Si(111) 7×7 surface. Via an STM experiment[23] it was proved that individual coronene molecules chemisorb selectively on the UH of the surface unit cells. However, the mechanism inducing this type of selective undissociative chemisorption could not be clearly identified. Also, based on the STM images the exact arrangement of the molecules on the UH of the unit cell could not be unambiguously determined. In the aforementioned experimental work there were two inequivalent structures proposed, where coronene is bound to three UCe adatoms. In one of these configurations the C atoms from coronene were expected to sit directly on-top of the underlying UCe atoms such, that there are three strong C–Si bonds formed between the adsorbate and substrate. In case of the alternative structure the coronene molecule is rotated by 15 degrees relative to the on-top configuration. In the following discussion we will refer to these two alternative configurations as the `on-top` and `R-15` structures, respectively.

In order to unravel the effect behind this selective chemisorption and to untangle the

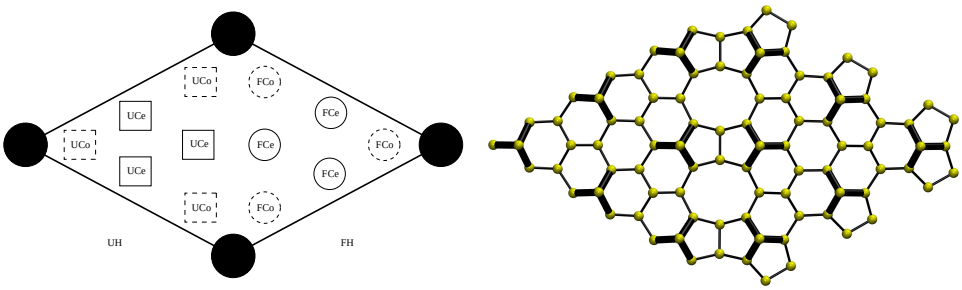


Figure 2. Position of possible binding sites on Si(111) 7×7. Left panel: cartoon representation of the position of the twelve Si adatoms on the unfaulted half (UH) and the faulted half (FH) of the unit cell. Unfaulted corner silicon adatoms (UCo) and unfaulted centre adatoms (UCe) are represented with squares and dashed squares, faulted corner adatoms (FCo) and faulted centre adatoms (FCe) are schematized with circles and dashed circles, respectively. The corner holes are depicted with large solid circles. Right panel: ball-and-stick model of Si(111) 7×7 surface unit cell. Si atoms are plotted as yellow spheres, Si–Si bonds are shown in black. The bonds between the twelve Si adatoms and the underlying Si atoms are represented with thick black cylinders. In all figures in this paper the unfaulted half of the Si(111) 7×7 surface unit cell is always shown in the left side, and the faulted half on the right side, respectively.

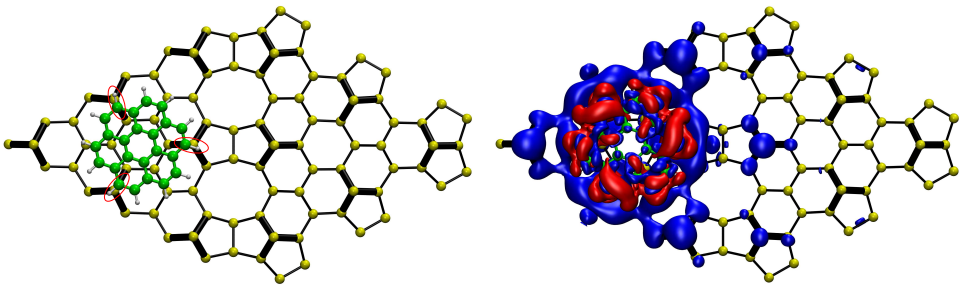


Figure 3. Left panel: top view of the most stable binding arrangement of coronene on the unfaulted half of the Si(111) 7×7 surface. The position of chemical bonds formed between the substrate and the adsorbate are highlighted with red ellipses. Right panel: charge redistribution upon adsorption of coronene superimposed on the most stable structure. Blue and red isosurfaces correspond to regions where charge is depleted and accumulated, respectively.

binding arrangement of coronene on the Si(111) surface, we have investigated the nature of the interaction via DFT calculations. To do so we have determined the adsorption energy of coronene for all reasonable binding arrangements on the UH and FH of the Si(111) 7×7. More than 20 different structures were investigated, but here only those will be presented which were found to be the most stable configurations.

We found, that the chemisorption of coronene on the FH of the surface unit cell is energetically unfavourable, and on the UH in the `R-15` configuration the adsorption of coronene on the surface is endothermic with −0.20 eV. In striking contrast to this, our calculations show that in the `on-top` configuration coronene is bound to the UH of the Si(111) 7×7 surface unit cell with a binding energy of 0.06 eV (see Fig. 3). Thus, at low coverages –up to an upper bound of one molecule per unit cell– thermodynamically it is highly favourable for coronene to adopt an `on-top` binding arrangement on the UH, leading to a selective chemisorption of the adsorbate molecules on the substrate. Considering that the long-range electron correlations, which are responsible for Van der Waals type interactions, are not treated explicitly in our computational setup, the theoretical results are in good qualitative agreement with the experimental findings.

In order to investigate the effect of the coronene adsorption upon the electronic structure of the system, we have looked into the redistribution of the bond charge density. The flow of the bond charge is plotted in Fig. 3, where blue and red isocontour surfaces show the regions where charge is depleted and accumulated, respectively. The isovalues at $1.42{\cdot}10^{-5}\ e/\text{Å}^3$ and $-3.71{\cdot}10^{-5}\ e/\text{Å}^3$ were selected such, that the respective blue and red isocontours enclose 50% of the redistributed charge. Via integrating the flow of the bond charge density over the full volume of the simulation cell we have computed that there is one electron transferred from the substrate to the adsorbate. The UCe and UCo adatoms have the main contribution to this electron transfer, which is well localized over the UH of the Si(111) 7×7 surface unit cell (see the blue isocontour in Fig. 3). Compared to the electronic structure of the isolated molecule, the electron transferred from the substrate is delocalized on the LUMO of the chemisorbed coronene (see the red isocontour in Fig. 3). This electron transfer can considerably boost the conduction properties of n-type FET devices such that the maximum current switched by such a molecular nanodevice would be adequate for real case applications.

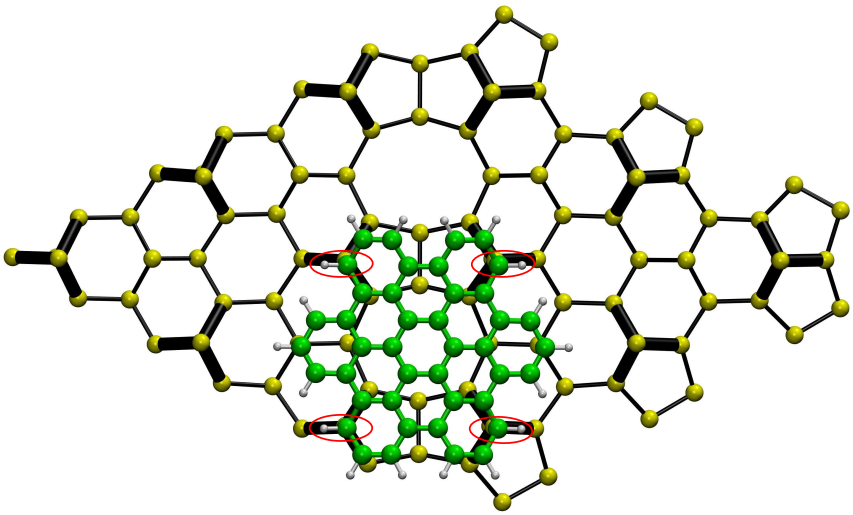


Figure 4. Top view of hexa-*peri*-benzocoronene anchored to the substrate via four Si–C bonds located along the short diagonal of the Si(111) 7x7 unit cell.

Coronene is just one among a class of graphene-like molecules that might be used to build nanotransistors. Therefore, in the following we discuss the interaction of HBC and CHP with the 7×7 reconstructed Si(111) surface, in order to inspire and support future experimental studies. Because of the good geometrical compatibility between the substrate and the adsorbate both HBC and CHP can adopt a plethora of different configurations on the surface. To compare HBC and CHP to coronene, first we have looked into the adsorption of these molecules on the UH and FH of the surface. In case of HBC the binding energies were in the same range as what we have calculated for coronene. Concerning the nature of the interaction between adsorbate and substrate, HBC shows a similar behaviour as coronene. The chemisorption of HBC on the FH of the surface unit cell is endothermic, and in the `on-top` configuration HBC is weakly bound to three UCe atoms with a binding energy of 0.13 eV. By searching for other possible configurations, as an interesting new feature we found, that HBC can be strongly anchored on Si(111) along the short diagonal of the surface unit cell with a binding energy of 0.48 eV (see Fig. 4). All Si–C bond lengths are about 2.01 Å, so in this case there is no palpable difference between the Si atoms located in the UH and FH. In this arrangement HBC is bound to the surface via four strong Si–C bonds created with four Si adatoms sitting in UCo, UCe, FCo and FCe positions, respectively. Remarkably, such geometries are analogous with the HBC derivatives used in Ref. 20, where four C atoms are blocked by side chains to prevent the aggregation of the molecules and the electrodes are attached to the two C atoms located in *para* positions. So, by supporting HBC on Si(111) the four C atoms would be blocked by Si–C bonds, and the two C atoms in *para* positions are free to serve as the source and gate electrodes of a nanographene FET.

Regarding the CHP adsorbate molecule, we have calculated that the chair conformer is 0.22 eV lower in energy compared to the boat conformer in the gas phase. Still, our calcu-

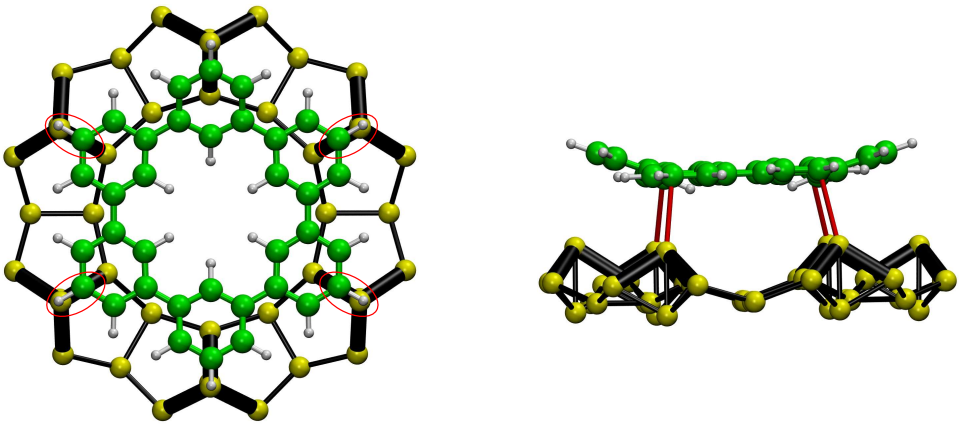


Figure 5. Top and perspective view of cyclohexa-*m*-phenylene adsorbed on the corner hole of the Si(111) 7x7 surface.

lations show, that due to the geometry of the binding sites on the Si(111) 7×7, CHP will have to undergo a configurational inversion –i. e. from chair to boat– in order to be attached to surface Si adatoms. For CHP we could not identify a binding site on the Si(111) surface where the chemisorption would be exothermic. The interaction energies between the CHP molecule on the UH and FH of the surface unit cell are similar to the values computed for coronene. Interestingly, we found that CHP is preferentially adsorbed on the corner hole of the 7×7 reconstructed Si(111) (see Fig. 5). This is because in the boat conformer the C atoms from CHP are well commensurate with the underlying Si atoms from the substrate. The Si–C interatomic distances are about 2.85 Å, which corresponds to a weak interaction between substrate and adsorbate. In this configuration the chemisorption of CHP is endothermic with −0.18 eV calculated relative to the chair conformer, which means that relative to the less stable boat conformer the adsorption would be slightly exothermic with 0.04 eV. Thus, we would expect that by using electron-withdrawing ligands on CHP the strength of the Si–C bonds could be tuned such, to allow for a firm binding of CHP molecules on Si(111).

5 Concluding Remarks

As a general trend among graphene-like molecules, we found that mostly the C atoms located at the edges of their outer rings contribute to the formation of Si–C bonds. However, in contrast to coronene, both HBC and CHP are anchored on the surface via four Si–C bonds. The various molecules prefer different specific binding sites of the substrate, resulting in selective chemisorption, leading to 1×1 patterns built up from individual adsorbate molecules. This pattern formation of the studied nanographene molecules is extremely helpful for the creation of regularly structured nanodevices. Due to the high adsorption energy and the beneficial charge transfer from the Si (111) substrate, HBC should be particularly suited for building arrays of individual nanotransistors.

Acknowledgements

The authors gratefully acknowledge financial support from the Center for Complex Matter (COMATT) at the Universität Mainz and supercomputer support from the Jülich Supercomputing Centre of the John von Neumann Institute for Computing (Project HMZ24).

References

1. K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666–669, 2004.
2. M. J. Allen, V. C. Tung, and R. B. Kaner, Chem. Rev. **110**, 132–145, 2010.
3. G. Brumfiel, Nature **458**, 390–391, 458.
4. Y. Zhang, T.-T. Tang, C. Girit, Z. Hao, M. C. Martin, A. Zettl, M. F. Crommie, Y. R. Shen, and F. Wang, Nature Lett. **459**, 820–823, 2009.
5. L. A. Ponomarenko, F. Schedin, M. I. Katsnelson, R. Yang, E. W. Hill, K. S. Novoselov, and A. K. Geim, Science **320**, 356–358, 2008.
6. A. K. Geim, and K. S. Novoselov, Nature Mater. **6**, 183–191, 2007.
7. K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, Nature Lett. **438**, 197–200, 2005.
8. A. Du, Z. Zhu, and S. C. Smith, J. Am. Chem. Soc. **132**, 2876–2877, 2010.
9. M. Hatanaka, Chem. Phys. Lett. **488**, 187–192, 2010.
10. M. Bieri, M. Treier, J. Cai, K. Ait-Mansour, P. Ruffieux, O. Groning, P. Groning, M. Kastler, R. Rieger, X. Feng, K. Müllen, and R Fasel, Chem. Commun. **45**, 6919–6921, 2009.
11. Y. Lin, C. Dimitrakopoulos, K. A. Jenkins, D. B. Farmer, H. Chiu, A. Grill, and P. Avouris, Science **327**, 662–662, 2010.
12. W. H. Lee, J. Park, S. H. Sim, S. Lim, K. S. Kim, B. H. Hong, and K. Cho, J. Am. Chem. Soc. **133**, 4447–4454, 2011.
13. P. Sutter, Nature Mater. **8**, 171–172, 2009.
14. J. Wu, W. Pisula, and K. Müllen, Chem. Rev. **107**, 718–747, 2007.
15. K. A. Ritter, and J. W. Lyding, Nature Mater. **8**, 235–242, 2009.
16. K. A. Ritter, and J. W. Lyding, Nanotechnology **19**, 015704–1–6, 2008.
17. X. Yang, X. Dou, A. Rouhanipour, L. Zhi, H. J. Räder, and K. Müllen, J. Am. Chem. Soc. **130**, 4216–4217, 2008.
18. J. Hofrichter, B. N. Szafranek, M. Otto, T. J. Echtermeyer, M. Baus, A. Majerus, V. Geringer, M. Ramsteiner, and H. Kurz, Nano Lett. **10**, 36–42, 2010.
19. X. Wang, X. Li, L. Zhang, Y. Yoon, P. K. Weber, H. Wang, J. Guo, and H. Dai, Science **324**, 768–771, 2009.
20. I. Diez-Perez, Z. Li, J. Hihath, J. Li, C. Zhang, X. Yang, L. Zang, Y. Dai, X. Feng, K. Müllen, and N. Tao, Nature Commun. **1**, 1–5, 2010.
21. W. Pisula, M. Kastler, C. Yang, V. Enkelmann, and K. Müllen, Chem. Asian J. **2**, 51–56, 2007.
22. K. Takayanagi, Y. Tanishiro, S. Takahashi, and M. Takahashi, Surf. Sci. **164**, 367–392, 1985.
23. J. Martínez-Blanco, M. Klingsporn, and K. Horn, Surf. Sci. **604**, 523–528, 2010.

24. Hutter, J. *et al.* CPMD, V3.13 IBM Corp 1990–2011, MPI für Festkörperforschung Stuttgart 1997–2001 http://www.cpmd.org.
25. D. Marx, and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: Cambridge, 2009.
26. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865–3868, 1996.
27. D. Vanderbilt, Phys. Rev. B **41**, 7892–7895, 1990.

Atomic Structure of AgInSbTe Phase Change Materials: Simulations and a Model of Crystallization

Jaakko Akola[1,2,3] **and Robert O. Jones**[1,4]

[1] Peter Grünberg Institut, Forschungszentrum Jülich, D-52425 Jülich, Germany
E-mail: {j.akola, r.jones}@fz-juelich.de

[2] Department of Physics, Tampere University of Technology, FI-33101, Tampere, Finland

[3] Nanoscience Center, Department of Physics, FI-40014 University of Jyväskylä, Finland

[4] German Research School for Simulation Sciences
Forschungszentrum Jülich and RWTH Aachen University, D-52425 Jülich, Germany

Phase change (PC) optical memories are based on the extremely rapid (nanosecond scale) crystallization of nanosized amorphous "marks" in a polycrystalline layer. Until recently, models of crystallization existed for the PC alloy $Ge_2Sb_2Te_5$ (GST), but not for the important class of Sb-Te based alloys. We have combined density functional simulations on amorphous and crystalline $Ag_{3.5}In_{3.8}Sb_{75.0}Te_{17.7}$ (AIST) with experiment (XRD, EXAFS, hard x-ray photoelectron spectroscopy) to determine their structures and how they differ from GST[1]. Amorphous (a-) AIST has a range of ring sizes, while a-GST has many small rings and cavities, and the local environment of Sb in both forms of AIST is a distorted 3+3 octahedron. We propose a "bond interchange" model (a sequence of small displacements of Sb atoms accompanied by interchanges of short and long bonds) as the origin of the rapid (slow nucleation) crystallization of a-AIST. It differs profoundly from crystallization in a-GST.

1 Introduction

Phase change (PC) materials are chalcogenide (Se, Te) alloys that switch very rapidly between the amorphous (a-) and crystalline (c-) phases. They are used extensively in rewritable high-density data storage, especially in optical recording [Digital Versatile Disc (DVD), Blu-ray Disc][2]. Information is stored as rows of nanosized amorphous marks in a polycrystalline layer and accessed via the different optical or electrical properties of the two phases. The most common materials are GeTe-Sb_2Te_3 pseudobinary compounds (Group 1 in Fig. 1a) and Sb-Te binary compounds with small amounts of In, Ag, and/or Ge, e.g. $Ag_{3.5}In_{3.8}Sb_{75.0}Te_{17.7}$ (AIST, Group 2 in Fig. 1a). Recrystallization in the two groups is strikingly different: In Group 1 it proceeds mainly via nucleation inside the marks (Fig. 1b), in Group 2 via crystal growth from the rim (Fig. 1c)[2].

Materials in both groups have superior rewrite speeds, and their amorphous phases are stable at room temperature (RT) for long periods, indispensable characteristics of PC memories. Elements in groups 15 and 16 (including As, Sb, S, Se, and Te) readily form disordered structures that are often short-lived at RT, but compounds containing other elements, such as Ag and In or Ge, can remain amorphous at RT for several decades. High stability implies a low tendency to crystallize, but this process can be accelerated by laser irradiation or electric heating. Crystallization of amorphous $Ge_2Sb_2Te_5$ (GST) and AIST materials can occur within tens of nanoseconds[3], so they have both the long-term stability and the rapid switching needed for effective PC memories. An atomic understanding of crystallization is essential, as it is the rate-limiting process in all PC materials.

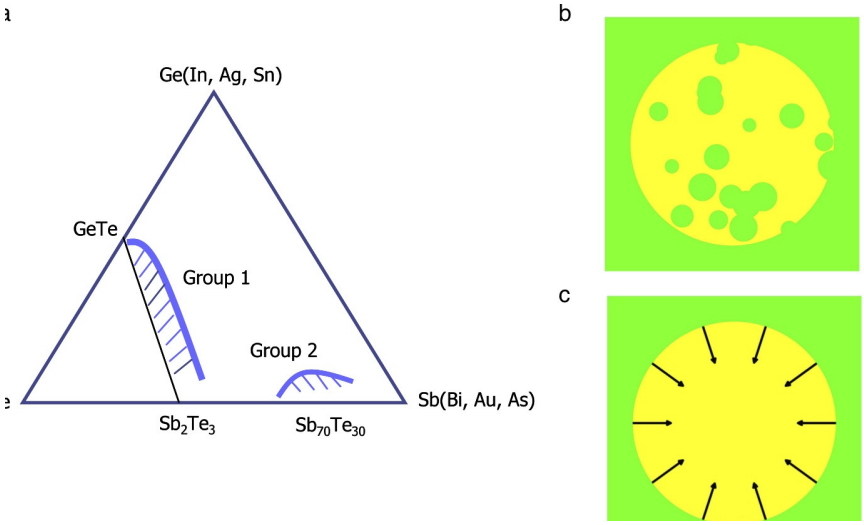


Figure 1. PC materials and their crystallization patterns. (a) Commonly used materials for optical recording are in Groups 1 and 2. (b) Nucleation-dominated (GST) and (c) growth-dominated recrystallization (AIST).

There have been many studies of crystalline GST (rock salt) and AIST (A7) materials[2]. The average number of *p*-electrons is near 3 in all compositions, suggesting orthogonal bonds with neighbouring atoms and octahedral coordination. Such structures are spatially isotropic or have near cubic symmetry, examples being high-temperature structures of AIST crystals[4]. Much less is known about the structure of the amorphous phases, and the PC mechanism remains the subject of speculation. Recent studies on a-GST include the combination of density functional (DF) / molecular dynamics (MD) and reverse Monte Carlo (RMC) simulations we used to reproduce x-ray diffraction (XRD) and hard x-ray photoelectron spectroscopy (HXPS) data. The structure of a-GST has many small rings with *AB* alternation (*A*: Ge, Sb; *B*: Te), which act as nuclei in the crystallization process[5]. However, little is known about the amorphous structures of Group 2 materials, including AIST. The technique used for a-GST [experiment (XRD, HXPS, and EXAFS) combined with DF/MD simulations] has been applied to determine the a-and c-structures of an AIST alloy, $Ag_{3.5}In_{3.8}Sb_{75.0}Te_{17.7}$. The differences between the structures of GST (Group 1) and AIST (Group 2) have wider implications.

2 Density Functional Calculations and RMC Refinement

The combined DF/MD calculations were performed with the CPMD package (Born-Oppenheimer mode)[6], and details of the exchange-correlation functional, the pseudopotential, the plane wave basis, and time steps are given in Ref. 1. The unit cell contained 23 Ag, 25 In, 480 Sb, and 112 Te atoms, and the simulations started at 3000 K in a cubic box (side 27.09 Å), followed by cooling (30 ps) to the melting point (850 K) and data collection

(20 ps)[7]. Cooling to 300 K over 100 ps with gradually decreasing density was then followed by data collection (30 ps) and quenching. The geometry was optimized at the final density (box size 27.57 Å) and refined using the RMC++ code[8] with the XRD structure factor $S(Q)$ and constraints on (a) the average coordination number of all atoms and (b) the distributions of bond angles Sb-Sb-Sb, Sb-Te-Sb, Te-Sb-Sb and Te-Sb-Te. Crystalline AIST (A7 structure) was simulated using 648 atoms in a hexagonal unit cell (a=25.826 Å, c=33.43 Å). The lattice sites were occupied by 23 Ag, 25 In, 486 Sb, and 114 Te atoms by introducing correlations between Ag/In and Te atoms (Ag/In attracted by Te, Te avoids Te) and optimizing the structure. The cohesive energy is 35 meV/atom lower in a-AIST than in c-AIST (4.07 eV/atom). Further details are given in Ref. 1.

The presence of four elements and the dominance of two (Sb, Te) with similar atomic numbers rule out determining the local structure of AIST by XRD alone, so that we have performed a melt-quenched DF simulation of a-AIST (640 atoms) and refined the structure using a reverse Monte Carlo (RMC) method. The criteria were: (a) agreement with XRD data, (b) a total energy close to the DF minimum, and (c) agreement with the HXPS measurements of electronic structure. The second and third requirements required carefully chosen constraints on bond distances and bond angles, but the resulting $S(Q)$ and $T(r)$ agreed very well with experiment (blue lines in Figs. 2a,b). AIST and GST geometries are shown in Figs. 2c and 2d, respectively. The total energy of RMC-refined a-AIST is only 68 meV/atom above the lowest DF energy. The minority atoms (48 Ag and In atoms in the simulation cell) seldom bind to each other (see Fig. 2c), but often bind to Te (Te avoids itself), suggesting that segregation into Sb and $AgInTe_2$ may occur[9]. The local environment of Sb (75% of all atoms) is very similar to that in the crystal (A7, a distorted octahedron). Cavities are small, and their total volume (7%) is much less than in a-GST (14%)[10]. Crystalline AIST has no cavities, c-GST 10%. The difference between the cavities in a-AIST and a-GST is clearly visible in Figs. 2c and 2d, which also show that *ABAB* squares[10] are common in a-GST, but insignificant in a-AIST.

For the present systems and basis sets, we found satisfactory scaling up to 4096 cores (one rack) of JUGENE. The CPMD program package scales better than competing DF/MD programs, but algorithmic changes are clearly needed before it can take full advantage of massively parallel machines.

3 Results and Discussion

Fig. 2a shows the structure factors $S(Q)$ of AIST and GST obtained using XRD[11]. The c-forms have sharp Bragg peaks (red lines), and the a-forms (black lines) have typical halo patterns. However, oscillations up to the maximum Q value indicate well defined short-range order in a-AIST. Fourier transformation of the $S(Q)$ leads to the total correlation functions $T(r)$ for c-AIST and c-GST (Fig. 2b), which are very similar beyond 4 Å. Small differences exist between the two crystalline forms at shorter distances, e.g. the double peak in c-AIST (2.93 Å and 3.30 Å) and a single peak in c-GST (2.97 Å). The $T(r)$ for the *amorphous* materials, however, differ significantly: the first peak in a-AIST (2.86 Å) is at a slightly shorter distance than in c-AIST (2.93 Å), but the first peak in a-GST (2.79 Å) is much shorter than in c-GST (2.97 Å). The shoulder on the second peak in a-AIST (3.5 Å, arrowed) is near that observed in the crystalline form (3.30 Å). These differences imply

different crystallization mechanisms; the atomic motion and/or diffusion accompanying the phase change are larger in GST than in AIST, where the bond lengths change little.

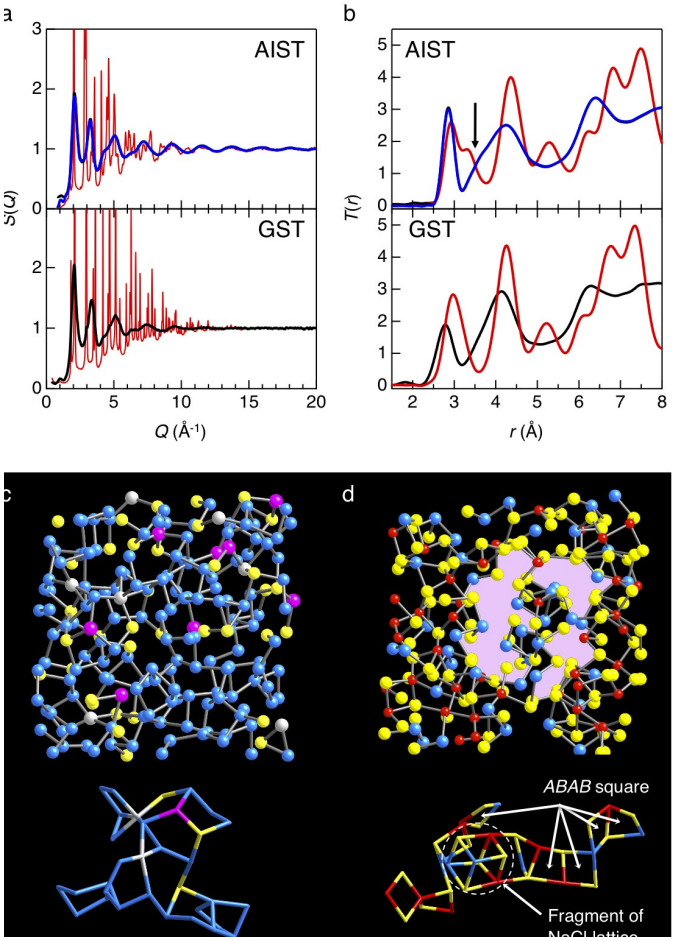


Figure 2. (a,b) $S(Q)$ and $T(r)$ of AIST and GST [Ref. 5]. Red: XRD data of c-phase, black: XRD data of a-phase, blue: DF-RMC model of a-AIST. DF-RMC and XRD results are almost indistinguishable. (c) Part of 640-atom model of a-AIST (24 Å x 24 Å x 12 Å). Ag, silver; In, magenta; Sb, blue; Te, yellow. (d) Part of 460-atom model of a-GST (24 Å x 24 Å x 12 Å), Ge, red; Sb, blue; Te, yellow, large cavity, pink.

The local DF-RMC structure for each element in a-AIST [coordination numbers N, bond lengths r, and bond orders[1]] is compared with EXAFS measurements (26 K) in Tab. 1. The coordination numbers from EXAFS measurements for Sb and Te agree well with the calculated values. EXAFS and calculated values for Ag and In show only small differences between the bond lengths.

The DF-RMC structure is consistent with both HEXRD and EXAFS data, and the corresponding density of states (DOS) has been compared with HXPS measurements in

Atom	N_{EXAFS}	r_{EXAFS}	$N_{\mathrm{DF/MD}}$	$r_{\mathrm{DF/MD}}$	N_{bond}
Ag	3.3 ± 0.5	2.768 ± 0.006	4.4	2.80 ± 0.05	1.9 (2.0)
In	4.3 ± 0.6	2.826 ± 0.006	3.1	2.85 ± 0.05	2.5 (2.9)
Sb	3.7 ± 0.3	2.872 ± 0.006	3.3	2.85 ± 0.05	3.1 (3.2)
Te	2.4 ± 0.4	2.827 ± 0.006	2.5	2.85 ± 0.05	2.5 (2.6)

Table 1. Local structure of a-AIST from EXAFS at 26 K and from DF-RMC geometry. N: coordination number, r: nearest neighbour bond lengths, N_{bond}: chemical coordination number (brackets: c-AIST value).

Ref. 1. As in GST[12], the amorphous and crystalline phases of AIST have similar HXPS spectra. The calculated DOS of c-AIST differs from the HXPS measurements in that there is a non-zero value at the Fermi energy, i.e. the small band gap semiconductor appears to be "metallic". DF calculations of this type focus on total energy differences, and the Kohn-Sham eigenvalues shown here commonly underestimate gaps between occupied and unoccupied states. If orthogonal bonds dominate in both phases, the phase change in AIST could occur by rapid bond interchange involving smaller atomic movements than in GST. This is consistent with the smaller change in $T(r)$ (XRD) for the two phases of AIST.

The above comparison with XRD, EXAFS, HXPS data shows that our DF-RMC model of a-AIST describes both the atomic arrangement and the electronic structure. A comparison of bonding in a-AIST and c-AIST is essential to understand the differences in PC mechanism, and DF calculations provide information about the bond order (the number of chemical bonds between pairs of atoms) and the effective atomic charge[1]. The chemical coordination numbers N_{bond} for Sb in AIST, calculated by adding the bond orders to the coordination numbers, provide additional insight. The values for Ag, Sb, and Te of a- and c-AIST are almost the same, as are the bond order profiles of Sb-Sb and Sb-Te[1]. Angular disorder in a-GST breaks the delocalized ("resonant") bonds in c-GST and provides sufficient optical contrast[13]. In AIST, this contrast is due to differences between the covalently bonded amorphous state and the ordered state with delocalized bonds.

The calculated PDF show that: (1) The first peak of the Sb-Sb and Sb-Te distributions are much sharper in a-AIST, and the bond lengths are near those of c-AIST. Both a- and c-AIST have a second maximum at 4.3 Å, which is consistent with a (pseudo)cubic medium range order. (2) Te-Te bonds are rare in a-AIST, while the PDF of c-AIST shows medium-range order with an enhanced peak at 6.2 Å. (3) Ag and In prefer to bind to Te, as seen in the stronger first peak for a-AIST after relaxation. There are few In-Sb bonds.

The RMC-refined DF calculations lead to the local environment around Sb atoms shown in Fig. 3, where both a-AIST (left) and c-AIST (right) show distorted 3+3 octahedra. Each vector shown is the sum of (typically three) short bonds. The norm of a vector is largest for three such bonds, and the vectors are randomly oriented in a-AIST. Laser irradiation or electric heating can lead to successive small atomic shifts that align the vectors along the c_{H}-axes of crystalline (A7, hexagonal) cells in the rim, so that the amorphous mark takes on the A7 structure. The essential PC process is a small displacement of the central atom that switches one short (red) and one long (dashed) bond. The close correspondence between the first peaks in $T(r)$ in a- and c-AIST is further evidence that the structural change is dominated by bond interchanges among near neighbours.

The atomic interchange alters the orientation of the octahedron, which can adapt to the

crystalline surroundings. The lower panels of Fig. 3 show how the vectors, randomly oriented in a-AIST, align along the c_{H}-axis. The octahedra order (A7 structure), the cavities disappear, and the density increases by 7%. The small atomic displacement simply interchanges a strong bond and a weak bond, without requiring bond breaking or diffusion. The PC process is then a sequential, collective motion of Sb atoms, and an avalanche of bond interchanges is consistent with both the instability of amorphous Sb and the rapid phase change in a-AIST and other Group 2 materials.

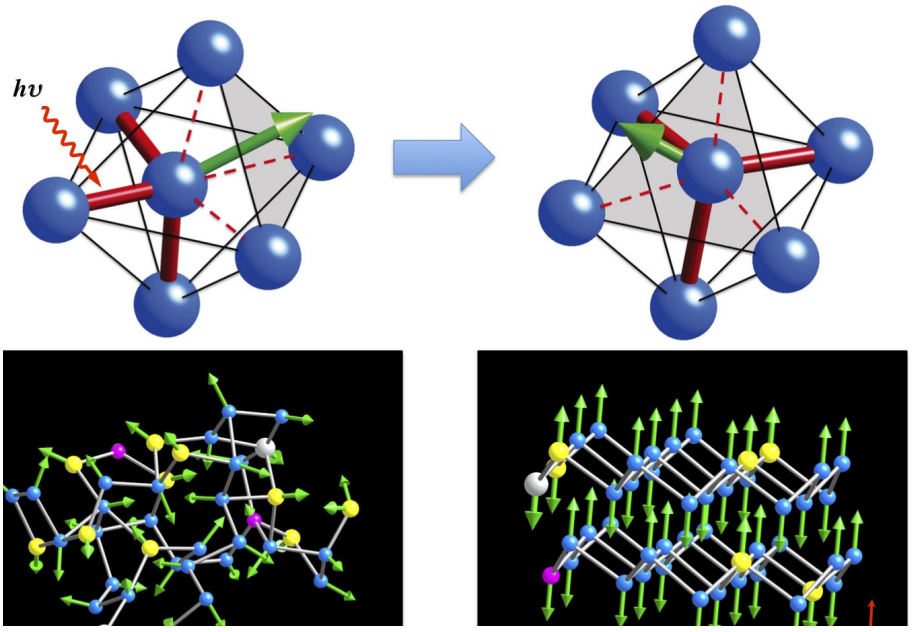


Figure 3. PC mechanism in a-AIST. Incident laser light causes atomic motion, and the central atom with three short (red) and three long (dashed) bonds moves across the distorted octahedron, interchanging short and long bonds. Green: resultant vector of short bonds. The grey sticks (lower right) correspond to the red bonds (upper right). Atom colours as in Fig. 2c.

The structure and electronic properties of a-AIST and a-GST have been determined using identical techniques, and we note first that the distributions of closed loops ("ring statistics") differ significantly[1]. In a-GST, 40 % of the rings are 4-fold or 6-fold, while the peak in the broader distribution in a-AIST is for 5-fold rings, which make up only 15 % of the total. Amorphous AIST shows links between rings of various sizes (Fig. 2c), while small rings dominate in a-GST (Fig. 2d). Differences in the topological constraints of ring statistics and cavity concentrations in a- and c-GST imply crystallization processes with changes in bond lengths and coordination numbers (see Fig. 2b). In a-AIST, we propose that a sequence of ring reconstructions via bond interchanges results in 6-fold rings with short Sb bonds, accompanied by small changes in the bond lengths (see Fig. 2b). Crystallization proceeds either from the crystalline area surrounding an amorphous mark (growth-

dominated crystallization) or from a seed that exceeds the critical size for nucleation (slow nucleation for Group 2). The latter occurs in as-deposited AIST films where the amorphous area is larger, and such films crystallize poorly if the surroundings of an amorphous mark are not crystalline. By contrast, many 4- and 6-fold rings in a-GST act as nuclei for crystallization and require larger atomic displacements than in a-AIST. Crystallization starts simultaneously from many such nuclei in the amorphous mark (NaCl fragments, Fig. 2d) and results in small crystal grains. This process is aided by the higher concentrations of cavities (Fig. 2d) and Te atoms, which favour low coordination. The structures suggest a "bond interchange" model as the origin of "growth-dominated" crystallization of a-AIST, whereas the large fraction of crystalline nuclei in a-GST is the origin of the "nucleation-driven" crystallization in GST. Nucleation is slower in AIST, because larger crystalline seeds are needed to impose an A7-type director on the neighbouring amorphous atoms.

4 Conclusions

Density functional simulations (640 atoms in the unit cell) combined with experimental (XRD, EXAFS, HXPS) studies of the AIST alloy $Ag_{3.5}In_{3.8}Sb_{75.0}Te_{17.7}$ have revealed its structure. The local environment of Sb atoms is a distorted 3+3 octahedron in both c- and a-phases, and the bonds in semiconducting a-AIST are slightly shorter (stronger) than in c-AIST (Fig. 2b, Tab. 1). Crystalline AIST has a distorted octahedral structure and more metallic features, as seen in the electronic DOS. The similar DOS in a- and c-phases is consistent with small atomic displacements during the phase change.

Our model suggests that crystallization of a-AIST is a rapid succession of diffusionless events where the 3+3 octahedra are aligned along the c_{H}-axis imposed by the surrounding crystal (see Fig. 1c). Heating or photon excitation causes alignment of the octahedra near the matrix boundary, and crystallites can grow along the laser-scanning direction. The lack of cavities and chemical alternation in a-AIST favour smooth growth. The roles of the dopant atoms (Ag, In) are unclear, but they impede atomic motion near RT and stabilize the amorphous marks. This is consistent with the dopant dependence of the crystallization temperature found in AIST[14]. In a-GST, small rings ("*ABAB* squares") act as nuclei for crystallization[5,10,15,16], aided by a higher cavity concentration and the presence of low-coordination Te atoms. Large scale DF simulations of the bond-interchange model should soon be possible.

Acknowledgements

The calculations were performed on IBM Blue Gene/P (JUGENE) and Intel Xeon 5570 (JUROPA) computers in the FZ Jülich, with grants from the FZ Jülich and the John von Neumann Institute for Computing (NIC).

References

1. T. Matsunaga, J. Akola, S. Kohara, T. Honma, K. Kobayashi, E. Ikenaga, R. O. Jones, N. Yamada, M. Takata, and R. Kojima, *From local structure to nanosecond recrystallization dynamics in AgInSbTe phase change materials*, Nature Mater. **10**, 129, (2011).

2. M. Wuttig and N. Yamada, *Phase-change materials for rewriteable data storage*, Nature Mater. **6**, 824, 2007 and references therein.
3. Y. Fukuyama *et al. Time resolved investigation of nanosecond growth in rapid phase-change materials – Correlations with the recording speed of digital versatile disc media*, Appl. Phys. Expr. **1**, 045001, 2008.
4. T. Matsunaga, Y. Umetani, and N. Yamada, *Structural study of a* $Ag_{3.4}In_{3.7}Sb_{76.4}Te_{16.5}$ *quadruple compound utilized for phase-change optical disks*, Phys. Rev. B **64**, 184116, 2001.
5. J. Akola, R. O. Jones, S. Kohara, S. Kimura, K. Kobayashi, M. Takata, T. Matsunaga, R. Kojima, and N. Yamada, *Experimentally constrained density-functional calculations of the amorphous structure of the prototypical phase-change material* $Ge_2Sb_2Te_5$, Phys. Rev. B **80**, 020201(R), 2009.
6. CPMD version 3.13. © IBM Corporation (1990-2010), MPI für Festkörperforschung, Stuttgart (1997-2001) (`http://www.cpmd.org/`).
7. J. Akola and R. O. Jones, *Structure of liquid phase change material AgInSbTe from density functional/ molecular dynamics simulations*, Appl. Phys. Lett. **94**, 251905, 2009.
8. O. Greben, P. Jóvári, L. Temleitner, and L. Pusztai, *A new version of the RMC++ Reverse Monte Carlo programme, aimed at investigating the structure of covalent glasses*, J. Optoelectron. Adv. Mater. **9**, 3021, 2007.
9. M. L. Lee, L. P. Shi, Y. T. Tian, C. L. Gan, and X. S. Miao, *Crystallization behavior of* $Sb_{70}Te_{30}$ *and* $Ag_3In_5Sb_{60}Te_{32}$ *chalcogenide materials for optical media applications*, phys. stat. sol. (a) **205**, 340, 2008.
10. J. Akola and R. O. Jones, *Structural phase transitions on the nanoscale: The crucial pattern in the phase change materials* $Ge_2Sb_2Te_5$ *and GeTe*, Phys. Rev. B **76**, 235201, 2007; *Binary alloys of Ge and Te: order, voids, and the eutectic composition*, Phys. Rev. Lett. **100**, 205502, 2008; *Density functional study of amorphous, liquid and crystalline* $Ge_2Sb_2Te_5$*: Homopolar bonds and/or AB alternation?* J. Phys.: Condens. Matter **20**, 465103, 2008.
11. S. Kohara *et al.*, *Structural basis for the fast phase change of* $Ge_2Sb_2Te_5$*: Ring statistics analogy between the crystal and amorphous states*, Appl. Phys. Lett. **89**, 201910, 2006.
12. J.-J. Kim *et al.*, *Electronic structure of amorphous and crystalline* $(GeTe)_{1-x}(Sb_2Te_3)_x$ *investigated using hard x-ray photoemission spectroscopy*, Phys. Rev. B **76**, 115124, 2007.
13. B. Huang and J. Robertson, *Bonding origin of optical contrast in phase-change memory materials*, Phys. Rev. Lett. **81**, 081204(R), 2010.
14. U.-C. Her, H. Chen, and Y.-S. Hsu, *Effects of Ag and In addition on the optical properties and crystallization kinetics of eutectic* $Sb_{70}Te_{30}$ *phase-change recording film*, J. Appl. Phys. **93**, 10097, 2003.
15. S. Caravati, M. Bernasconi, T. D. Kühne, M. Krack, and M. Parrinello, *Coexistence of tetrahedral- and octahedral-like sites in amorphous phase change materials*, Appl. Phys. Lett. **91**, 171906, 2007.
16. J. Hegedüs and S. R. Elliott, *Microscopic origin of the fast crystallization ability of Ge–Sb–Te phase-change memory materials*, Nature Mater. **7**, 399, 2008.

The Running-In of Hydrogenated Diamond-Like Carbon Coatings

Lars Pastewka[1] **and Michael Moseler**[1,2]

[1] Fraunhofer-Institut für Werkstoffmechanik IWM, Wöhlerstraße 11, 79108 Freiburg, Germany
E-mail: lars.pastewka@iwm.fraunhofer.de

[2] Institute of Physics, University of Freiburg, Hermann-Herder-Straße 3, 79104 Freiburg, Germany

The tribological performance of hydrogenated diamond-like carbon (DLC) coatings is studied by molecular dynamics simulations. DLC films with 45 % and 30 % H content as well as 50 % and 30 % sp^3 hybridization are investigated. We study the frictional properties of symmetric pairings of these films at realistic sliding speed of $20\,\mathrm{m\,s^{-1}}$ and loads of 1 and 5 GPa. While the hydrogen-rich system shows a pronounced drop of the friction coefficient for both loads, the hydrogen-depleted system exhibits such kind of running-in for 1 GPa, only. Chemical passivation of the DLC/DLC interface explains this running-in behaviour.

1 Introduction

Tuning the sp^2/sp^3 ratio and the hydrogen concentration of amorphous hydrocarbon (a-C:H) films allows to tailor the mechanical and electrical properties[1] of these materials to specific applications. Here we focus on the frictional properties of such films[2] that become important when these films are used for the protection of magnetic and optical storage discs[3], bearings[4] or micro-electro-mechanical-systems (MEMS)[5].

In addition to their technological importance, a-C:H coatings are scientifically exciting since these seemingly simple two-elemental materials exhibit already many tribological features of much more complex multi-elemental metallic tribo-systems. Therefore, insights into processes that govern the performance of sliding a-C:H interfaces can pave the path to a deeper understanding of general tribological phenomena, such as running-in, lubrication, the formation of transfer films, wear and final failure.

For instance, during the sliding of a-C:H against a-C:H, friction coefficients can approach values as low as 0.001[6–8]. This "superlubricity" is sometimes linked to another remarkable property of these films, their ultrasmoothness[9]. Interestingly, the ultralow friction coefficient requires some time to develop[6] — a phenomenon that is called *running-in*. An initially high friction coefficient (> 0.1) decreases over a couple of hundred frictional cycles until a low steady-state value is established. Although running-in occurs in many tribo-contacts (including most metallic systems), the details of the underlying mechanisms are not even understood for simple a-C:H contacts.

Molecular dynamics simulations are ideally suited to gain insight into the processes that govern the performance of a-C:H tribo-contacts. Previous theoretical investigations in the tribology of hydrocarbons have focused on systems in which the surfaces were completely saturated with hydrogen and no chemical reaction or just single hydrogen abstraction events occurred (for reviews see the recent articles by Harrison and co-workers, Refs. 10 and 11). Examples include the friction between self-assembled monolayers[12,13] as well as amorphous carbon films[14]. The latter work considered hydrogen-free DLC films

sliding with $100\,\mathrm{m\,s^{-1}}$ against a hydrogen-terminated diamond (111) surface. During the sliding over a distance of $12\,\mathrm{nm}$ the number of bond-breaking and bond-formation events *within* the amorphous carbon film was monitored. The frequency of structural rearrangements decreased over time and dropped to zero after approximately $60\,\mathrm{ps}$. Although no accompanying decrease in the friction coefficient was reported in Ref. 14, this observation was interpreted as a manifestation of running-in.

Recently, we found a different running-in mechanism in our molecular dynamics simulations[15,16], where we observed that the decrease of the friction coefficient was exactly correlated with a reduction of the chemical bonding *between* the two surfaces. The bonding even vanished after sliding for a couple of nanoseconds corresponding to a sliding distance of a couple of 100 nanometers. In these works — as in the present work —, we employed a modified incarnation of the second generation reactive empirical bond-order potential[17], that has been particularly tailored to allow for a proper description of bond-breaking and bond-formation processes as well as amorphous phases[18]. We validated our simulations by comparing the increase in sp^2 hybridization to experimental data[16] obtained from Raman measurements yielding reasonable agreement. The difference in time-scale observed in Refs. 14 and 15,16 already indicates that different relaxation processes might occur during running-in, an observation that will be discussed in this article.

2 Sliding at 1 GPa

Fig. 1a displays the temporal evolution of the friction coefficient for the $30\,\%$ H and $45\,\%$ H tribo-contacts during sliding at a load of $1\,\mathrm{GPa}$. We compute this friction coefficient by calculating the ratio of shear stress to normal pressure that are both averaged over time intervals of $10\,\mathrm{ps}$. For both, the H-rich and H-depleted system, the friction coefficient μ (Fig. 1a) starts at values of above 1.0 and then drops to a value of around 0.1. This manifestation of *running-in* is much faster for the H-rich sample where the saturated value of the friction coefficient establishes already after $5\,\mathrm{ns}$.

The friction coefficient directly correlates with the amount of time the two sliding partners stay chemically connected (Fig. 1b). We define two carbon atoms to be bound if their distance is closer than $1.85\,\mathrm{Å}$. By following the network of carbon bonds it is straightforward to determine if a connection between the two sliding partners exists. As a measure of connectivity, we compute the relative amount of time the two surfaces exhibit chemical bonding during time intervals of $0.5\,\mathrm{ns}$. In the hydrogen-(H)-depleted case we observe a steady decrease in the connected time reaching zero after approximately $30\,\mathrm{ns}$. In the hydrogen-(H)-rich case, the surfaces are also chemically connected during the first $10\,\mathrm{ns}$, but the total time this chemical connection persists is much smaller than in the $30\,\%$ hydrogen case. After running-in the two sliding partners remain disconnected in both cases, leading to a steady and low friction coefficient. The fact that chemical bonding and friction coefficient are directly correlated is nicely demonstrated at around $20\,\mathrm{ns}$ in the H-rich case. Here, a connection is established for a short time which instantly raises the friction coefficient to a value of 0.3.

During sliding, the local structure of the a-C:H changes slightly. Fig. 1c shows the sp^2 and sp^3 content in the interface region. In the H-depleted as well as the H-rich sample sp^2 increases by a few percent accompanied by a corresponding decrease in sp^3. This "phase-transformation" is another manifestation of *running-in* and is presumably also reflected in

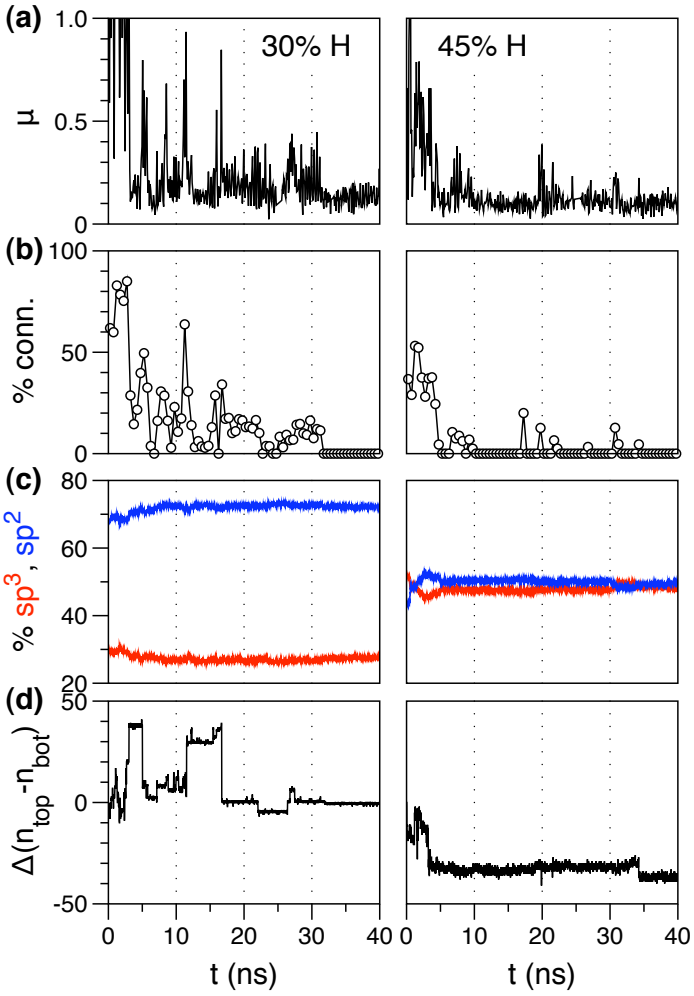


Figure 1. Dry sliding at a load of 1 GPa for a H-depleted (30 % H) and a H-rich (45 % H) a-C:H sample. The figure shows (a) the friction coefficient, (b) the amount of time a chemical bond exists between the sliding partners in an interval of 0.5 ns, (c) the sp^2 and sp^3 content in a region ± 1 nm around the sliding interface, and (d) the number of atoms transferred between the sliding partners. Quantities (a), (c) and (d) are averaged over intervals of 10 ps.

the bond-formation and bond-breaking process observed in Ref. 14. As already stressed in our earlier publications[15,18] this "graphitization" reflects a minor rearrangement in the amorphous carbon's local network structure and not the formation of graphitic sheets.

This microscopic running-in, which is evident in the friction coefficient, sample connection, and hybridization, is also seen in the rate at which atoms are transferred from one sliding partner to the other. Fig. 1d shows the difference in the number of atoms belonging to the two sliding partners, shifted to zero at $t = 0$. Each jump corresponds to a transfer

of atoms between the two slabs. Initially the transfer rate is large as indicated by frequent jumps in Fig. 1d, however, it drops to zero around the time the friction coefficient saturates in both the H-depleted and the H-rich case.

Let us summarize our main observation. The strong reduction of friction during the first few nanoseconds is accompanied by a drop in chemical connectivity of the two tribo-surfaces. Such a lowering of the surface reactivity requires some kind of passivation mechanism. By examining trajectories we are able to identify individual passivation events: Typically, a reactive surface has unsaturated carbon atoms exposing dangling bonds into the sliding contact. When the atoms approach each other a bond is established, linking the two sliding partners. In general, this bond or one of the neighbouring bonds have to break again during further sliding. The weakest bond is determined by the chemical environment of the atoms and to some extent by thermal fluctuations that can carry an atom over energy barriers[19]. In a representative passivation event, we observe that after breaking of that bond a CH_2 group has moved to the other sliding partner. An unsaturated carbon is still exposed to the sliding interface. Subsequently, this exposed unsaturated carbon finds a binding partner within the surface. Note, that such a passivation process can only occur if a suitable binding partner within the tribo-surface is available. However, once passivated such an atom is unlikely to become unsaturated again.

3 Sliding at 5 Gpa

At a load of 5 GPa the differences between the H-depleted and the H-rich sample become more pronounced. Now, the H-depleted sample shows no running-in. The friction coefficient μ (Fig. 2a) stays at an average value of 0.5 in stark contrast to the H-rich system which still exhibits a behaviour very similar to that at 1 GPa. For both loads the friction coefficient has dropped to about 0.1 after 5 ns. Again, this behaviour correlates with the amount of time the samples stay chemically connected (Fig. 2b). The H-depleted sample maintains bonds between the two sliding partners during the whole simulation, while in the H-rich sample the surfaces passivate after about 20 ns. Subsequently, only minimal re-connection between the two sliding partners is observed.

The evolution of the hybridization in the high load H-rich system is practically identical to the corresponding low load case: the sp^2 contents jumps initially by about $5\,\%$ and saturates at this final value (Fig. 2c). On the other hand, the change in hybridization is more pronounced in the high load H-depleted case. Here the increase in sp^2 is about $15\,\%$ – considerably higher than the $2\,\%$ increase in the corresponding low load simulation (Fig. 1c).

4 Conclusions

Our results clearly demonstrate that the hydrogen concentration is crucial for the running-in behaviour and the resulting steady state friction coefficient of a-C:H tribo-contacts. Films with lower H-concentration require longer for the running-in and even weld together at higher loads. This theoretical observation correlates with UHV experiments of Ref. 6 that revealed a steady state friction coefficient of 0.003 and 0.53 for an a-C:H with 40% and 34% hydrogen, respectively. Fontaine et al.[6] explained the anti-correlation between μ and

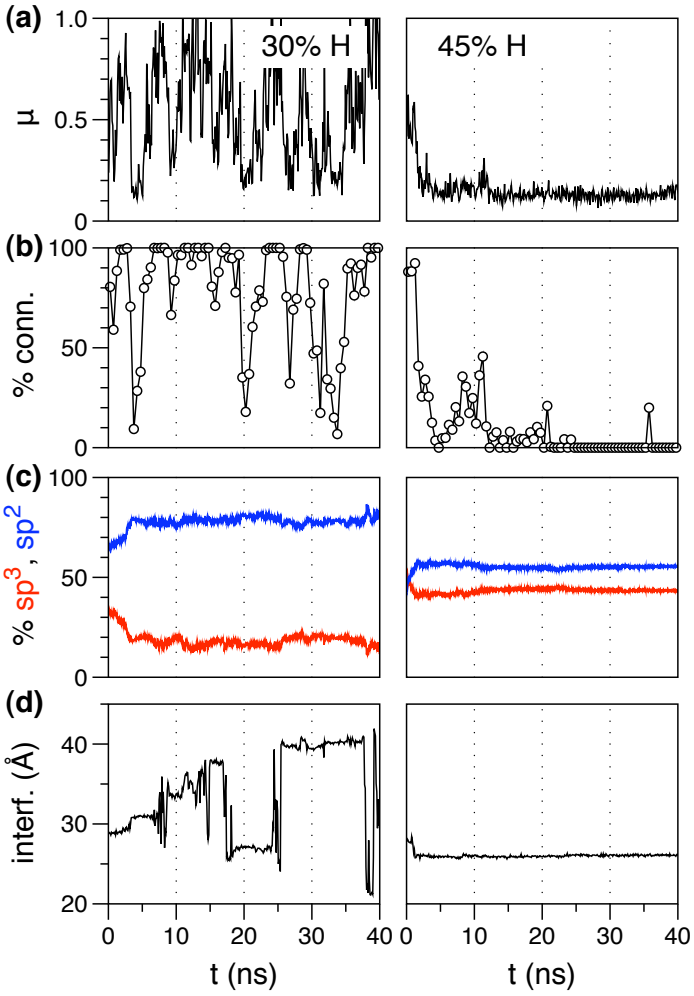


Figure 2. Dry sliding at a load of 5 GPa for a H-depleted (30 % H) and a H-rich (45 % H) a-C:H sample. The figure shows (a) the friction coefficient, (b) the amount of time a chemical bond exists between the sliding partners in an interval of 0.5 ns, (c) the sp^2 and sp^3 content in a region ± 1 nm around the sliding interface, and (d) the position of the interface, obtained from the position where the velocity drops from 20 m s^{-1} to zero. Quantities (a), (c) and (d) are averaged over intervals of 10 ps.

H concentration by the increased viscoplasticity of the H-rich films leading to weaker intra-film interactions and hence different relaxation of the asperities. Our simulations indicate that not only the bulk response of the DLC to shearing is important, but also the ability of the surfaces to passivate.

The main passivation event that can be extracted from our simulations is the tendency of dangling bonds to saturate by finding a binding partner with other dangling bonds. If this is not possible, single radicals can be abstracted from the surface and moved to a location

where such a passivation is possible. Similar observations on running-in have been made previously by Gao et al.[14]. They observed bond-breaking events *within* the amorphous carbon film which saturated on a time-scale an order of magnitude below the saturation of surface termination chiefly discussed in this work. Bond-breaking within the film is related to the transformation of sp^3 to sp^2 which is most pronounced during the first 5 ns of sliding in our simulations. The time-scale clearly depends on the specific structure of the films and the sliding velocity explaining the slight discrepancy to the work of Gao et al.[14]

However, the presence of two time-scales for surface and bulk relaxation indicates that tribology in general, and running-in in particular, is a multi-timescale problem. At longer times and lengths, for example, the surface *roughness* typically flattens as observed in experiments (e.g. diamond[20], Al[21]). As already mentioned above, also in a-C:H visco-plastic topography evolution seems to be of importance. The underlying processes are clearly suppressed by the small simulation cell size used here and in previous works[15,16,14]. We will hence focus on increasing the spatial extend of our simulation domain to capture realistic surface roughness. Initially, it will be possible to only capture the short wavelength limit of the roughness which typically extends over many length scales[22]. However, the use of appropriate multi-scaling techniques, such as using a harmonic elastic manifold for the simulation boundaries[23], will help to extend the accessible length scales even further. Future development will therefore also focus on the development of efficient parallelization techniques for such multi-scale approaches.

Acknowledgements

We thank the BMBF for funding this work within project OTRISKO. The calculations were carried out at the Jülich Supercomputing Centre.

References

1. J. Robertson, *Diamond-like amorphous carbon*, Mater. Sci. Eng. R, **37**, 129, 2002.
2. A. Erdemir and C. Donnet, *Tribology of diamond-like carbon films: recent progress and future prospects*, J. Phys. D: Appl. Phys., **39**, R311, 2006.
3. A. C. Ferrari, *Diamond-like carbon for magnetic storage disks*, Surf. Coat. Tech., **180–181**, 190, 2004.
4. J. Brand, C. Beckmann, B. Blug, G. Konrath, and T. Hollstein, *Diamond-like carbon coatings – a new design element for tribological applications*, Ind. Lubr. Tribol., **54**, 291, 2005.
5. R. Maboudian, *Adhesion and friction issues associated with reliable operation of MEMS*, MRS Bulletin, **23**, 47, 1998.
6. J. Fontaine, T. Le Mogne, J. L Loubet, and M. Belin, *Achieving superlow friction with hydrogenated amorphous carbon: some key requirements*, Thin Solid Films, **482**, 99, 2005.
7. A. Erdemir, O. L. Eryilmaz, and G. Fenske, *Synthesis of diamondlike carbon films with superlow friction and wear properties*, J. Vac. Sci. Technol. A, **18**, 1987, 2000.
8. C. Donnet, M. Belin, J. Augé, J. Martin, A. Grill, and V. Patel, *Tribochemistry of diamond-like carbon coatings in various environments*, Surf. Coat. Tech., **68–69**, 626, 1994.

9. M. Moseler, P. Gumbsch, C. Casiraghi, A.C. Ferrari, and J. Robertson, *The ultrasmoothness of diamond-like carbon surfaces*, Science, **309**, 1545, 2005.
10. J. A. Harrison, J. D. Schall, M. T. Knippenberg, G. Gao, and P. T. Mikulski, *Elucidating atomic-scale friction using molecular dynamics and specialized analysis techniques*, J. Phys.: Condens. Matter, **20**, 354009, 2008.
11. J. A. Harrison, G. Gao, J. D. Schall, M. Knippenberg, and P. T. Mikulski, *Friction between solids*, Phil. Trans. Roy. Soc. A, **366**, 1469, 2008.
12. M. Knippenberg, P. Mikulski, B. Dunlap, and J. Harrison, *Atomic contributions to friction and load for tip–self-assembled monolayers interactions*, Phys. Rev. B, **78**, 235409, 2008.
13. P. T. Mikulski, L. A. Herman, and J. A. Harrison, *Odd and even model self-assembled monolayers: Links between friction and structure*, Langmuir, **21**, 12197, 2005.
14. G. T. Gao, P. T. Mikulski, and J. A. Harrison, *Molecular-scale tribology of amorphous carbon coatings: Effects of film thickness, adhesion, and long-range interactions*, J. Am. Chem. Soc., **124**, 7202, 2002.
15. L. Pastewka, S. Moser, and M. Moseler, *Atomistic insights into the running-in, lubrication, and failure of hydrogenated diamond-Like carbon coatings*, Tribol. Lett., **39**, 49, 2010.
16. L. Pastewka, S. Moser, M. Moseler, B. Blug, S. Meier, T. Hollstein, and P. Gumbsch, *The running-in of amorphous hydrocarbon tribocoatings: a comparison between experiment and molecular dynamics simulations*, Int. J. Mat. Res., **99**, 1136, 2008.
17. D. W. Brenner, O. A. Shenderova, J. A. Harrison, S. J. Stuart, B. Ni, and S. B. Sinnott, *A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons*, J. Phys.: Condens. Matter, **14**, 783, 2002.
18. L. Pastewka, P. Pou, R. Pérez, P. Gumbsch, and M. Moseler, *Describing bond-breaking processes by reactive potentials: Importance of an environment-dependent interaction range*, Phys. Rev. B, **78**, 161402, 2008.
19. L. Pastewka, S. Moser, P. Gumbsch, and M. Moseler, *Anisotropic mechanical amorphization drives wear in diamond*, Nat. Mater., **10**, 34, 2011.
20. B. Gupta, A. Malshe, B. Bhushan, and V. Subramaniam, *Friction and wear properties of chemomechanically polished diamond films*, J. Tribol., **116**, 445, 1994.
21. S. W. Lo and S. D. Tsai, *Real-time observation of the evolution of contact area under boundary lubrication in sliding contact*, J. Tribol., **124**, 229, 2002.
22. B. B. Mandelbrot, D. E. Passoja, and A. J. Paullay, *Fractal character of fracture surfaces of metals*, Nature, **308**, 721, 1984.
23. C. Campañá and M. H. Müser, *Practical Green's function approach to the simulation of elastic semi-infinite solids*, Phys. Rev. B, **74**, 075420, 2006.

First Principles Numerical Study of the Magnetic Properties of Transition-Metal Alloy Nanoparticles

L. E. Díaz-Sánchez and G. M. Pastor

Institut für Theoretische Physik, Universität Kassel, 34132 Kassel, Germany
E-mail: pastor@uni-kassel.de

The ground-state magnetic properties of CoRh and CoPt alloy nanoparticles having up to 531 atoms (about 2 nm) are investigated in the framework of density functional theory by using a fixed-moment method and fully-unconstrained structural relaxation. The role of chemical order on the magnetic behaviour is investigated by considering core-shell nanoalloys with different phases, compositions and chemical environments. The importance of electron correlation effects is explored by comparing calculations based on the local spin density and generalized-gradient approximations to the exchange and correlation functional. All considered nanoalloys are found to be magnetic with an average spin moment that is larger than in the macroscopic alloys having similar compositions. Besides the usual finite-size effects on the electronic structure, which reduce the effective d band width and thus enhance the local spin moments, we observe important induced moments at the Rh and Pt atoms due to the proximity with the strongly magnetic Co atoms. These $4d$ or $5d$ moments are a very important contribution to the average cluster moment per atom. An analysis of the spin density distribution from a local perspective reveals an interesting dependence of the magnetic order on the local and chemical environment of the atoms. The results are interpreted in terms of the local densities of electronic states and of the structural relaxations around atoms at bulklike, interface, and surface environments. Finally, the calculations are discussed in connection with recent experiments on nanoparticles having similar sizes.

1 Introduction

In past years the magnetism of $3d$ transition metal (TM) nanoparticles (NPs) has been the subject of numerous experimental and theoretical studies[1–3]. While the emphasis of early studies has been on monometallic clusters, the interest on binary alloy nanoparticles has been growing steadily. Still, our present understanding of nanoalloys remains rather primitive[4]. Systems composed by a magnetic $3d$ metal (e.g., Cr, Fe, Co and Ni) and a highly-polarizable $4d$ or $5d$ metal (e.g., Rh, Pd and Pt) are particularly interesting in this context[5–14]. Indeed, one expects that an appropriate choice of the alloy parameters should allow one to tune the electronic, magnetic and structural behaviour in order to develop new materials with optimized properties. From the point of view of magnetism, one of the main challenges is to take advantage of the reduced dimensionality and of the large spin-orbit coupling at the $4d$ or $5d$ atomic shells in order to enhance and control the magnetic anisotropy energy (MAE) of the NPs. The MAE describes the dependence of the energy of a magnetic material as a function of the orientation of the magnetization with respect to lattice structure. It defines the direction or directions of the magnetization at equilibrium, the so called easy axes, as well as the stability of the magnetization direction with respect to temperature-induced fluctuations or external fields. Consequently, tailoring the MAE is very useful for applications in recording and memory devices, where the magnetization needs to be pinned to a given direction in space. Previous studies have already demonstrated the remarkable properties of $3d$-$4d$ and $3d$-$5d$ interfaces not only in alloy particles

but also in $3d$ clusters and thin films deposited on highly polarizable substrates[15–19]. It is therefore very important to understand and control the magnetic properties of nanomaterials as a function of structure and composition for different arrangements of the chemical species within the NPs.

Considerable efforts have been devoted to the study of nanoalloys with sizes ranging from a few atoms up to diameters of about 20 nm (e.g., FePt, CoPt, CoRh, FeRh, and CoMn)[5,8–12,14]. To this aim a variety of production methods have been developed involving, for example, organometallic chemistry synthesis, low-energy cluster-beam deposition, or diffusion-controlled aggregation on surfaces. As a consequence of the size reduction and of the elaboration process new alloy phases are synthesized and new segregation phenomena can often be observed, which are not available in bulk equilibrium conditions. Taking into account the subtle interplay between magnetism, electronic correlations and structure, it is important to investigate the magnetic properties of nanoalloys by treating electronic and structural degrees of freedom rigorously on the same footing. This requires to perform very demanding first principles calculations in the experimentally relevant size range of 300–600 atoms. The access to high-performance computer infrastructures is therefore a remarkable asset to the progress in this field.

2 Computational Method

The calculations reported in this work have been carried out by using the *Vienna ab initio* simulation package[20] (VASP) on the supercomputer JUGENE. The VASP is an efficient implementation of Hohenberg-Kohn-Sham's density functional theory, which allows to perform first principles electronic, magnetic and structural optimizations. An augmented plane-wave basis set is considered for the valence s and d electrons with periodic boundary conditions in a supercell. The changes in the wave functions of the valence electrons due to the interactions with the inner shell core electrons are described by using the projected augmented wave (PAW) method[21]. This is a very successful approach for transition metals, since it is intended to combine the accuracy of all-electron methods with the efficiency of pseudopotentials. Exchange and correlation effects are treated by using the semilocal functional given by the generalized-gradient approximation (GGA)[22]. In addition the local spin density approximation (LSDA) has also been considered for the sake of comparison.

The wave functions are expanded in a plane-wave basis set with the kinetic energy cut-off $E_{\max} = 268$ eV. In order to improve the convergence of the solution of the selfconsistent Kohn-Sham equations the discrete energy levels are broadened by using a Gaussian smearing $\sigma = 0.01$ eV. The validity of the present choice of computational parameters has been verified. The PAW sphere radii for Co, Rh and Pt are 2.393 Å, 2.536 Å, and 2.590 Å, respectively. Higher cutoff energies have been considered for test proposes but no significant changes in the magnetization values or in any other property reported in this work have been observed. A simple cubic supercell with periodic boundary conditions is used. The linear size of the supercell is such that distance between the clusters on neighbouring cells is always larger than 15 Å. This separation is more than enough in order to avoid any spurious interactions between different cluster images. Since we are interested in finite systems, the reciprocal space summations are restricted to the Γ point. The Born-Oppenheimer approximation and a conjugate gradient method are used in order to perform fully unconstrained structural relaxations. In this context convergence is assumed when

the energy difference between two consecutive line searches is less than 0.1 meV, which corresponds to interatomic forces smaller than or about 10 meV/Å.

A fixed-moment method is used in order to determine the ground state of each cluster. This is a scalar relativistic approximation in which the total spin polarization $S_z = (\nu_\uparrow - \nu_\downarrow)/2$ is kept fixed throughout the self-consistent calculation, where ν_σ refers to the number of electrons with spin σ. For each S_z an unconstrained structural relaxation is performed by using a conjugate gradient scheme. The value of S_z is then varied systematically in each system so that the lowest energy spin state corresponding to the ground state is identified. Notice that, for simplicity, the spin-orbit interactions are not taken into account in the present calculations, since they are not expected to affect neither the dominant spin moments nor the structural relaxation. In fact, we have performed calculations for representative small CoRh clusters by including spin-orbit contributions within VASP and found no significant changes in the spin density distribution or in the optimal geometries[23]. Similar conclusions have been drawn from calculations on small FeRh clusters[24]. The typical computer production runs were performed on 512 processors with a torus memory configuration, which yields an efficient scaling on JUGENE[13]. Under these conditions the unconstrained structural relaxations for clusters of $N = 273$ ($N = 531$) atoms usually require 3–5 (8–14) hours wall-clock time. In the following sections we review the main conclusions of studies on two experimentally relevant systems, namely, CoRh and CoPt nanoparticles.

3 Magnetic Properties of Core-Shell CoRh Nanoalloys

The ground-state magnetic properties of Co_xRh_{1-x} nanoparticles having sizes $N = 273$ and 531 atoms and Co concentrations $x \simeq 0$, 0.25, 0.5, 0.75 and 1 have been investigated. In order to quantify the role of chemical order on the magnetic behaviour we consider different fcc-like structures as starting configurations of the clusters, which are composed by a central atom and the successive shells of its nearest neighbours (NNs). First, we focus on core-shell distributions of the Co and Rh atoms, namely, Co core with Rh outer shell and vice versa. This is consistent with the conclusions of previous experimental studies of the geometrical structure of these nanoclusters[5]. In addition, we have also explored the properties of particles having nearly planar CoRh interfaces. It is interesting to observe that after performing an unconstrained structural relaxation the point-group symmetry of all considered clusters is reduced from the cubic to D_{4h}. Despite this symmetry reduction, the local DOS and the local magnetic moments remain in practice nearly the same for all the atoms that where equivalent in the starting fcc cubic geometry, i.e., for all atoms in the same NN shell of the central atom. For instance, the differences in the local moments are of the order of $5 \times 10^{-3}\mu_B$. Consequently, only the average results within each NN shell around the central atom will be discussed in the following.

For clusters having $N = 273$ atoms we consider eight different core-shell configurations corresponding to a Co concentration $x = 0$, 0.25, 0.5, 0.75, and 1. In addition, we have performed calculations for particles having a planar interface and $x \simeq 0.5$. The resulting relaxed geometries corresponding to the lowest energy total magnetic moment are illustrated in Fig. 1. As example of larger nanoparticles we have investigated core-shell clusters having $N = 531$ atoms and concentrations $x = 0.25$, 0.5, and 0.75. An illustration of the corresponding relaxed geometries is shown in Fig. 2. Results for intermixed interfaces and random distributions of the Co and Rh atoms will be reported elsewhere.

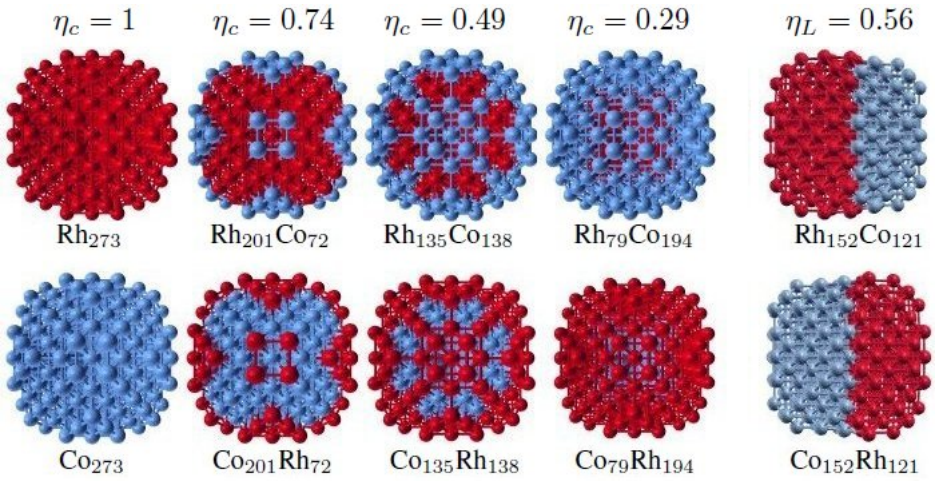


Figure 1. Illustration of the relaxed geometries of fcc-like core-shell CoRh clusters having $N = 273$ atoms. Blue (red) balls represent Co (Rh) atoms. The corresponding relative core sizes $\eta_c = N_c/N$ are indicated, where N_c refers to the number of atoms in the core. At the right the relaxed structures of clusters having a planar CoRh interface are shown for $x \simeq 0.5$. Here $\eta_L = N_L/N$ is the proportion of atoms in the left side of the clusters.

All considered CoRh systems are found to be magnetic with an average spin moment that is larger than in macroscopic alloys having similar concentrations. The enhancement of the local magnetic moments can be ascribed in part to the reduction of the local coordination number, which reduces the effective d band width and thus increases the stability of local moment formation. However, the dominant contribution to the average magnetization in $3d$-$4d$ nanoalloys stems from the spin moments at the Rh atoms close to the CoRh interface, which couple parallel to the Co moments. This polarization of the Rh atoms is induced by the proximity of the Co atoms, which show a large spin polarization and a large d-band exchange splitting even in a Rh-rich environment. As in any transition metal, the

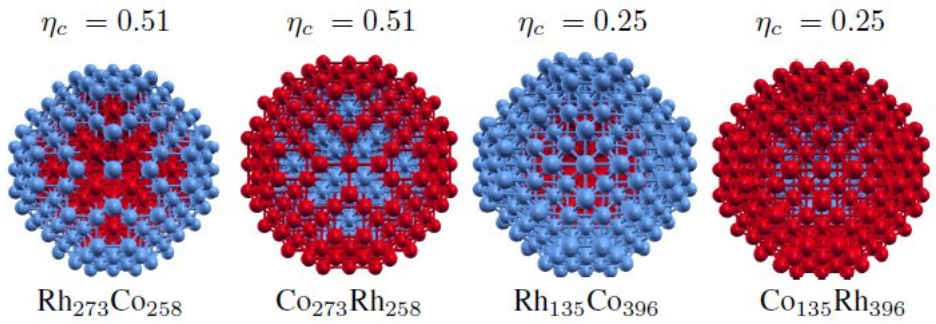


Figure 2. Illustration of the relaxed geometries of fcc-like core-shell CoRh clusters having $N = 531$ atoms. Blue (red) balls represent Co (Rh) atoms. The corresponding relative core sizes $\eta_c = N_c/N$ are indicated.

Cluster	η_c	$\bar{\mu}$	$\bar{\mu}^{\mathrm{WS}}$	$\bar{\mu}^{\mathrm{WS}}_{\mathrm{Rh}}$	$\bar{\mu}^{\mathrm{WS}}_{\mathrm{Co}}$
Rh_{273}	1.0	0.19	0.18	0.18	
$Rh_{201}Co_{72}$	0.74	0.62	0.60	0.13	1.92
$Rh_{135}Co_{138}$	0.49	1.07	1.06	0.22	1.88
$Rh_{79}Co_{194}$	0.29	1.36	1.35	0.19	1.82
Co_{273}	1.0	1.75	1.74		1.74
Rh_{273}	1.0	0.19	0.18	0.18	
$Co_{79}Rh_{194}$	0.29	0.45	0.46	-0.06	1.74
$Co_{135}Rh_{138}$	0.49	1.12	1.11	0.45	1.78
$Co_{201}Rh_{72}$	0.74	1.65	1.61	1.15	1.78
Co_{273}	1.0	1.75	1.74		1.74

Table 1. Average spin moments in fcc-like core-shell Co_nRh_m clusters having $N = n + m = 273$ atoms. Results are given for the average total moment $\bar{\mu} = (\nu_\uparrow - \nu_\downarrow)/N$, the moment within the Wigner-Seitz (WS) spheres $\bar{\mu}^{\mathrm{WS}} = (\sum_{l=1}^{N} \mu_l^{\mathrm{WS}})/N$, the WS Rh moment $\bar{\mu}^{\mathrm{WS}}_{\mathrm{Rh}} = (\sum_{l=1}^{m} \mu_l^{\mathrm{WS}})/m$, and the WS Co moment $\bar{\mu}^{\mathrm{WS}}_{\mathrm{Co}} = (\sum_{l=1}^{n} \mu_l^{w\mathrm{WS}})/n$. All values are given in Bohr magnetons per atom.

magnetic moments at the Rh atoms tend to be larger when the local coordination is smaller. Thus, the largest induced Rh moments are found at the surface atoms which combine low local coordination and Co-rich environment. Similar conclusions have been obtained in self-consistent tight-binding studies[11,25].

An analysis of the Co and Rh moments from a local perspective reveals an interesting strong dependence of the magnetic order on the local and chemical environment of the atoms. In Tab. 1 results are given for several averages of the spin moments in CoRh clusters with $N = 273$ atoms. Starting from pure Rh_{273}, which is weakly magnetic ($\overline{\mu} = 0.19\mu_B$), and increasing the Co content, we observe that the magnetic moment per atom increases monotonously as we replace Rh by the strongly magnetic Co atoms. However, the enhancement is far more important than the expected difference in the size of the local Co and Rh moments. Indeed, as already mentioned, the highly polarizable Rh atoms acquire a spin polarization due to the core Co atoms, which multiplies the effect. Comparing core-shell clusters with similar compositions, one observes that particularly large magnetic moments occur when the Rh atoms are at the outer shells. In fact, the nanoalloy which shows, among the considered ones, the largest Rh moments is the one having a large Co core with 201 atoms (relative core size $\eta_c = N_c/N = 0.74$) and 72 atoms in the Rh outer shell. In this case, reduced coordination and proximity effects work together.

In order to analyse the cluster magnetization profile in more detail we have determined the local magnetic moments μ_i^{WS} within the Wigner-Seitz (WS) spheres of each atom i. The dependence of the spin-polarized electronic structure on the chemical order and on the type of interface is given by the local densities of states reported in Fig. 3. First of all one observes, concerning the spatial distribution of the spin polarization, that more than 90% of the total cluster moment originates in the WS spheres of the atoms. This can be easily verified by comparing the results for $\bar{\mu}$ and $\bar{\mu}^{\mathrm{WS}}$ in Tab. 1 and confirms the dominant

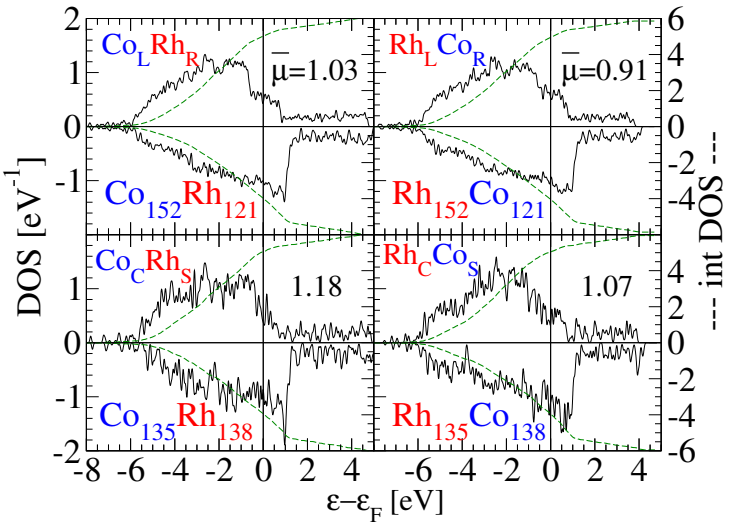


Figure 3. Spin-polarized average density of states (DOS) as a function of the Kohn-Sham eigenenergy ε relative to the Fermi energy ε_F of Co_nRh_m clusters having $N = n + m = 531$ atoms. Results are given for both core-shell spherical CoRh interfaces and left-right planar CoRh interface as obtained in the GGA. The dashed curves refer to the integrated DOS (see right vertical axis). Positive (negative) values correspond to majority (minority) spin. The corresponding average magnetic moments per atom.

role of the localized d-electron contributions. In the case of clusters with a Co core and a Rh outer shell we observe a nearly monotonous increase of the local Co moments as we move from the cluster centre to the CoRh interface (see Figs. 1 and 2). Significant local moments are induced at the Rh atoms close to the interface, which remain approximately constant or eventually tend to decrease as we move away from the interface. For thick Rh outer shells (e.g., in $Co_{135}Rh_{396}$) we find oscillations and antiparallel alignment of the Rh surface moments relative to average cluster magnetization[23]. In the case of clusters with a Rh core and a Co outer shell we observe an enhancement of the Co surface moments. The induced Rh local moments are important for $x \simeq 0.5$ (e.g., in $Rh_{273}Co_{258}$) but decrease significantly for lower Rh content, for example, for $x \simeq 0.25$ in $Rh_{135}Co_{396}$. As expected, the Rh moments decrease as we move away from the interface. In this case one also observes oscillations and antiparallel alignment of the Rh moments. Further details will be published elsewhere[23].

The larger atomic radius of Rh, as compared to Co, implies that the size of the clusters increases with increasing Rh concentration. This trend is independent of the type of core-shell or left-right arrangement of the atoms. For example, the calculated diameter D of fcc CoRh clusters with $N = 273$ atoms and $x \simeq 0.5$ is $D = 1.87$ nm for $Co_{135}Rh_{138}$ (Co core) and $D = 1.83$ nm for $Rh_{135}Co_{138}$ (Rh core). These differences seem too small to be detectable in experiment. Notice that for these cluster sizes the diameter is larger when the Co atoms are in the core. Moreover, we observe a general trend to expand the cluster by increasing the interatomic CoCo distances due to the steric effect of the Rh atoms. The trend for clusters with $N = 531$ atoms is somewhat different. In the case of $Co_{273}Rh_{258}$ (Co core) the calculated cluster diameter is $D = 2.23$ nm, while in

$Rh_{273}Co_{258}$ (Rh core) we have $D = 2.27$ nm. In contrast to the $N = 273$ case, the larger diameter now corresponds to the Rh core. These seemingly contradictory results can be understood by noting that at larger core sizes the binding between Co atoms is stronger. Therefore, the surface Rh shells have a weaker effect on the bigger Co cores. The percentage of expansion in the Co core of $Co_{273}Rh_{258}$, due to the presence of the Rh outer shell, is nearly 50% smaller than the expansion of the Co core in $Co_{135}Rh_{138}$. These results show that it is not possible to safely infer whether Co or Rh atoms are in the core, simply on the basis of experimentally determined cluster diameters[5].

4 Discussion on CoPt Nanoparticles

Recent experiments have revealed detailed quantitative information on the local structure of size-selected CoPt nanoparticles having a diameter of 2–4 nm and prepared by a low-energy cluster-beam deposition technique in an amorphous carbon matrix[10]. From Co K-edge EXAFS experiments and simulations the first-shell CoCo and CoPt coordination numbers and atomic distances have been determined as a function of cluster size. For "as prepared" samples (i.e., without any post-deposition treatment) a lattice parameter $a = 3.58$ Å is found for the largest particles by assuming an fcc chemically disordered A1 model. We have therefore performed *ab initio* calculations of the structure of 2 nm CoPt clusters in the disordered A1 phase and found similar lattice parameter as in experiment. This corresponds to a CoCo interatomic distance $d_{\mathrm{CoCo}} = 2.53$ Å. The obtained cluster structures are illustrated Fig. 4. This should be compared with the larger lattice parameter $a = 3.80$ Å of the bulk A1 phase.

Experiment has also provided clear evidence for the appearance of an $L1_0$ chemically ordered structure in these particles after thermal treatment[10]. However, in contrast to the bulk phase which shows a contracted $c/a = 0.97$ ratio, the $L1_0$ nanoalloy phase shows an expanded $c/a = 1.05$ value. These experimental findings call for theoretical calculations on $L1_0$ CoPt nanoclusters. One of the first aims of theory would be to verify if this unusually large $c/a > 1$ is energetically stable. Moreover, it would be interesting to elucidate if these special structural properties are at the origin of or related to the surprisingly low MAE that is generally found in these $L1_0$ nanoparticles.

In order to investigate this problem we have determined the relaxed geometries of a variety of CoPt particles with different initial structures and lattice parameters, from which

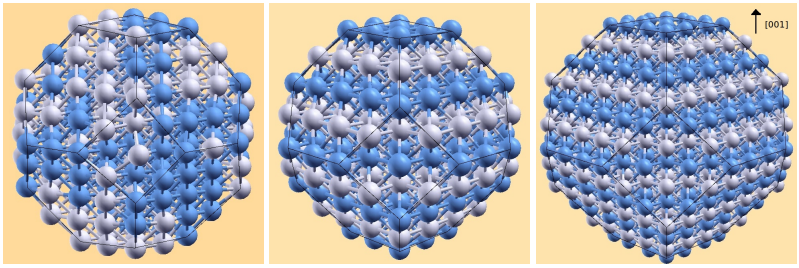


Figure 4. Illustration of the relaxed geometries of CoPt nanoparticles with disordered A1 and ordered $L1_0$ structures. Blue (white) spheres represent Co (Pt) atoms.

the experimentally relevant distribution of the CoCo, CoPt and PtPt interatomic distances has been inferred. It is most encouraging that the value of c/a calculated from d_{CoCo} and d_{CoPt} —which corresponds to the EXAFS measurements at the Co K-edge— is equal to about 1.02 for all the $\mathrm{L1_0}$ cases. Independent layer relaxations explain the unusually large observed c/a ratio in the $\mathrm{L1_0}$ structure of CoPt nanoparticles. Further investigations of the magnetic properties of these clusters are currently in progress.

Acknowledgements

It is pleasure to thank Prof. J. Dorantes-Dávila, Dr. V. Dupuis and Dr. F. Tournus for helpful discussions and useful comments. The calculations reported in this work have been performed at the supercomputer JUGENE through a computer time grant of the John von Neumann Institute for Computing. Financial support by the Deutsche Forschungsgemeinschaft and by the COST Action "Nanoalloys" are also gratefully acknowledged.

References

1. J. P. Bucher, D. C. Douglass, and L. A. Bloomfield, Phys. Rev. Lett. **66**, 3052 (1991).
2. I. M. L. Billas, A. Chatelain, and W. E. de Heer, Science **265**, 1682 (1994).
3. G. M. Pastor and J. Dorantes Dávila, The Chemical Physics of Solid Surfaces, Vol. **12**, ed. by D. P. Woodruff, (Elsevier, Amsterdam, 2007), p. 536.
4. See, for instance, Faraday Discuss. **138** (2008).
5. D. Zitoun *et al.*, Phys. Rev. Lett., **89**, 37203 (2002).
6. H. Ebert *et al.*, Comput. Mater. Sci. **35**, 279 (2006).
7. G. Rossi, R. Ferrando and C. Mottet, Faraday Discuss. **138**, 193 (2008).
8. G. Rollmann *et al.*, Phys. Rev. B **78**, 134404 (2008); M. E. Gruner *et al.*, Phys. Rev. Lett. **100**, 087203 (2008).
9. S. Ganguly *et al.*, Phys. Rev. B **78**, 014402 (2008).
10. F. Tournus *et al.*, Phys. Rev. B **77**, 144411 (2008).
11. M. Muñoz-Navia *et al.*, Appl. Phys. Lett., **95** 233107 (2009).
12. J. Honolka *et al.*, Phys. Rev. Lett. **102**, 067207 (2009).
13. M. E. Gruner and P. Entel, J. Phys.: Condens. Matter **21**, 064228 (2009).
14. F. Tournus *et al.*, Phys. Rev. B **81**, 220405 (2010).
15. P. Gambardella *et al.*, Nature (London) **416**, 301 (2002).
16. P. Gambardella *et al.*, Phys. Rev. Lett. **93**, 077203 (2004).
17. J. Dorantes-Dávila *et al.*, Phys. Rev. Lett. **91**, 197206 (2003).
18. S. M. Valvidares *et al.*, Phys. Rev. B **81**, 024415 (2010).
19. R. Félix-Medina *et al.*, Phys. Rev. B **67**, 094430 (2003).
20. G. Kresse and J. Furthmüller, Phys. Rev. B 54, 11169 (1996); G. Kresse and D. Joubert, Phys. Rev. B 59, 1758 (1999).
21. P. E. Blöchl, Phys. Rev. B 50, 17953 (1994).
22. J. P. Perdew *et al.*, Phys. Rev. B 46, 6671 (1992).
23. L. E. Díaz-Sánchez *et al.*, to be published.
24. J. H. Mokkath and G. M. Pastor, arXiv:1110.2669v1 [physics.atm-clus].
25. M. Muñoz-Navia *et al.*, Faraday Discuss. **138**, 181 (2008).

Condensed Matter: Ordering Phenomena from the Atomistic Scale to the Mesoscale

K. Binder[1] **and A. Muramatsu**[2]

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

[2] Institut für Theoretische Physik III, Universität Stuttgart, 70550 Stuttgart, Germany
E-mail: mu@theo3.physik.uni-stuttgart.de

"Condensed matter" stands for the liquid and solid phases considered in the framework of statistical thermodynamics, where matter is formed from a very large number of atoms. Such systems are extremely diverse and appear often in very different contexts: e.g., small clusters of molecules condensed in the atmosphere are responsible for cloud formation and hence an ingredient for weather prediction and climate modeling; nanostructured multicomponent metallic alloys are an important topic of materials science; folding of proteins in aqueous solution is a central field of research in molecular biology; polymer solutions under shear is a classical problem in the science of nonequilibrium soft matter; etc. All such problems (and related ones) can be considered as "condensed matter" in a wide sense, but are outside the scope of the present chapter in this book (since they are covered in different section such as "Computational Biology and Biophysics", "Materials Science", "Computational Soft Matter Science", "Earth and Environment", as is evident from this book or previous proceedings in the same series). The physics of condensed matter is certainly of basic relevance for these neighbouring disciplines, which are more specialized and often application-driven. Of course, the boundaries between "condensed matter" and such neighbouring disciplines are not sharp but rather smooth, and hence bringing all these fields together in the NIC Symposia allows interactions across the boundaries of all these disciplines, and this is particularly true for the groups working in condensed matter defining the field in a more narrow sense.

When one starts out from a very fundamental level, condensed matter amounts to deal with many-body quantum mechanics (many nuclei and electrons, all interacting via Coulomb potentials), to understand the ground state plus excitations of such highly correlated matter. The second step would be to use this information as an input into the theoretical framework of quantum statistical mechanics, predicting then macroscopic properties of matter from "first principles", as well as the outcome of experiments that use scattering of light, x-rays or neutrons, or spectroscopic methods, mechanical deformation, etc. This general task was formulated already by Paul Dirac in 1929, but even with petascale (and exascale, in the future) computers one is far from this goal, if one addresses it in full rigor and generality. Nevertheless, when one is somewhat less ambitious, one can obtain very valuable insight on condensed matter systems from large-scale computation, and the present chapter testifies for this fact.

One strategy to simplify the problem is to resort to density functional theory (DFT). In this approach, the ground state properties of condensed matter are in the focus, using the fact that the problem (via the Kohn-Sham theorem) can be reduced to dealing with the one-electron density in the framework of an exact functional. Although this functional cannot be written down explicitly in exact form, approximations have been developed which are impressively successful, despite their somewhat heuristic character. The present chapter gives two beautiful examples for the progress that can be obtained along such lines: the article by Zeller et al. describes a code called KKR nano (the name refers to the use of the full-potential Kohn-Korringa-Rostoker Green's function method) which allows to study the electronic structure of systems containing up to 8000 atoms, i.e. extending over the scale of several nanometers already. This code can run in parallel efficiently using all 294 912 processors of the JUGENE computer! First pioneering applications of this code include the role of interstatial defects occurring when GaN is doped with Gd, and effects of structural disorder on electronic properties of the phase change material $GeSb_2Te_4$. Another challenge for DFT are the so-called "multiferroic materials", where electrons lead to the simultaneous occurrence of electric and magnetic dipole moments, and the computations could guide the practical development of a suitable material, namely the mixed crystal (Eu, Ba) TiO_3, for which the prediction that this system indeed is multiferroic subsequently was confirmed experimentally.

Of course, the approximate character of DFT precludes its use for the study of strong electronic correlations in condensed matter, that may even lead to novel, exotic phases of matter. In order to address such issues, it is necessary to study generic effective Hamiltonians, which omit unnecessary details of the electronic structures but allow methods to be used that are in principle exact, such as quantum Monte Carlo (QMC); of course, this approach is practical only when the famous "minus sign" problem is absent, and even then it addresses rather small systems only, and hence requires a careful extrapolation to the thermodynamic limit.

Meng at al. demonstrate the potential of this approach by considering the Hubbard model on the honeycomb lattice (i.e. a system resembling graphene, a two-dimensional phase of carbon, the recent discovery of which was honored by the 2010 physics Nobel prize). Meng et al. establish the phase diagram of this system as a function of on-site Coulomb repulsion, and find an unexpected complexity, such as the occurrence of a resonating valence bond (RVB) liquid, a new exotic state of matter.

Effective Hamiltonians are also useful for problems of classical statistical physics, such as colloid-polymer mixtures, in which the colloids are simply approximated as hard spheres (and polymers are integrated out, leading to an effective attraction among the colloids). This system is considered in the article by Deb et al., which deals with confinement effects due to repulsive walls. The challenge that is addressed is to develop methods from which the "wall tension" (i.e., excess free energy of the system due to a wall) can be computed. This problem requires the use of new methods together with Monte Carlo data of very high accuracy, and hence large scale simulations on JUROPA were indispensable to reach this goal. The interfacial properties of the system near its liquid to solid transition are of fundamental interest for an understanding of crystal nucleation from fluid phases.

Also the final article of the present chapter by Harting *et al.* deals with a colloidal fluid, namely emulsion formation in a three-component mixture of immiscible fluids such as oil + water + surfacant, which are not treated in a chemically realistic way, of course, but rather

by the lattice Boltzmann method coupled with Molecular Dynamics simulation. Using lattice sizes as large as $1024 \times 1024 \times 1152$ and 454502 particles, the code parallelizes perfectly on the whole JUGENE (294 912 cores). In this way the possible breakup of droplets stabilized by particles but subjected to shear could be studied, as well as the "phase diagram" describing the transition from a Bîjel emulsion to a Pickering emulsion.

Thus, this chapter again gives a demonstration of the wide range of subjects for which large scale simulations on the JUGENE and JUROPA computer lead to a very significant progress on the forefront of science.

Monte Carlo Simulation Studies of Interfacial Free Energies in Colloidal Suspensions

Debabrata Deb, Dorothea Wilms, Alexander Winkler, Peter Virnau, and Kurt Binder

Institut für Physik, Johannes Gutenberg-Universität Mainz
Staudinger Weg 7, D-55099 Mainz, Germany
E-mail: kurt.binder@uni-mainz.de

In colloid-polymer mixtures the strength and range of the (effective) attraction between the colloidal particles can be varied via the choice of the polymer concentration and the size of the polymer coils, and hence one obtains a prototypical model system for the study of cooperative phenomena in condensed matter. Being interested in the study of homogeneous nucleation of crystals from the metastable fluid phase, and of heterogeneous nucleation at walls, various interfacial free energies need to be computed. New Monte Carlo methodologies to achieve this task have been developed, and we test the accuracy of these methods by applying them to the limiting case of a colloidal suspension formed from hard spheres, where other studies are available. Also a comparison of some of our numerical results with density functional theories is made. It is demonstrated that wall-attached crystalline nuclei with nonzero contact angle occur for the Asakura-Oosawa model of colloid-polymer mixtures.

1 Introduction

Spherical colloidal particles can be prepared with diameters in the μm range and almost monodisperse size distribution[1,2]. In suitable suspensions their interaction is essentially hard sphere like[1]. If polymers are dissolved in the suspension as well, an entropic depletion attraction between the colloids is created[3,4]. According to the Asakura-Oosawa model[3], polymers are described as penetrable spheres which can overlap each other with zero energy cost, while polymer-colloid overlap is strictly forbidden. For size ratios $q = \sigma_p/\sigma < 0.155$ (σ_p is the diameter of the polymer coils, σ the diameter of the colloidal particles) the polymers need not be considered explicitly, they can be integrated out to yield an effective attractive interaction[5].

$$V_{attr}(\eta_p^r, r) = -(q^{-1}+1)^3 \eta_p^r [1 - \frac{3r}{2\sigma(1+q)} + \frac{r^3}{2\sigma^3(1+q)^3}],\ \sigma < r < \sigma + \sigma_p\ , \quad (1)$$

r being the distance between the particles, and the strength of the attraction is controlled by the "polymer reservoir packing fraction" $\eta_p^r = (\pi/6)\sigma_p^3 \exp(\frac{\mu_p}{k_B T})$ where μ_p is the chemical potential of the polymers. Fig. 1 compares the equation of state of the hard sphere system to that of the Asakura-Oosawa (AO) model; choosing the parameters $q = 0.15, \eta_p^r = 0.1$ in the latter[6]. Plotting the reduced pressure pk_BT/σ^3 versus the colloid packing fraction $\eta = (\pi/6)\sigma^3\rho$, ($\rho = N_c/V$ being the density of the colloids, N_c being their number, V the system volume), we recognize two branches: a fluid branch and a solid branch, where the system crystallizes in the face-centred-cubic (fcc) structure. There occurs pronounced hysteresis, but the pressure p_{co} where phase coexistence of fluid (with $\eta = \eta_f$) and crystal (with $\eta = \eta_m$) occurs can be very accurately found from simulations

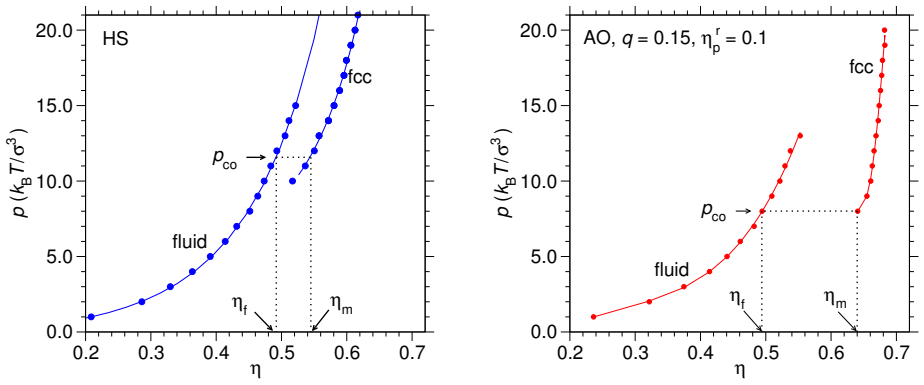


Figure 1. Phase diagram of the hard sphere (HS) model (left) and the AO model with $q = 0.15, \eta_p^r = 0.1$ (right), plotting pressure (in units k_BT/σ^3) vs. colloid packing fraction η. Freezing of the fluid occurs for $\eta_p = 0.492$ (HS) or 0.494 (AO) and melting at $\eta_m = 0.545$ (HS) or 0.64 (AO), respectively. Coexistence pressures are $p_{co} = 11.576$ (HS) or 8.00 (AO), respectively. From Zykova-Timan et al.[6].

of phase coexistence in a slab geometry[6], studying the variation of the interface velocity with pressure (for $p = p_{co}$ the interface velocity changes its sign). This method also yields well-equilibrated fluid-crystal interfaces, and analyzing the capillary-wave spectrum of these interfaces (or the broadening of the interfacial profile caused by such excitations) one can obtain reasonably accurate estimates of the interfacial tensions γ for both models[6].

In a real system, hysteresis is not as broad as in Fig. 1, since macroscopic systems are bounded by the walls of a container, where heterogeneous nucleation[7,8] of the stable phase (the crystal) from the metastable fluid is facilitated. Heterogeneous nucleation is believed to be ubiquitous in nature (e.g., ice crystals are nucleated from under cooled supersaturated water vapor in the atmosphere; etc.) but nevertheless is not well understood; the purpose of the present project is to provide reliable simulation methodologies to obtain the parameters needed to predict the free energy barrier ΔF^*_{het} to be overcome by heterogeneous nucleation.

The standard theory (due to Turnbull[7]) describes heterogeneous nucleation at a surface by a sphere-cap shaped droplet, with radius of curvature R^* and contact angle θ at the substrate surface. Then

$$\Delta F^*_{het} = \Delta F^*_{hom}(R^*) f(\theta), \quad f(\theta) = (1-\cos\theta)^2(2+\cos\theta)/4\ , \tag{2}$$

where $\Delta F^*_{hom}(R^*)$ is the free energy barrier against homogeneous nucleation. At supersaturation, this free energy barrier can be expressed as $\Delta F^*_{hom}(R^*) = 4\pi R^{*2}\gamma/3$, where a possible curvature-dependence of the interfacial tension γ is neglected. The contact angle θ is given by Young's equation in terms of the wall-fluid (γ_{wf}) and wall-crystal (γ_{wc}) interfacial tensions:

$$\gamma_{wf} - \gamma_{wc} = \gamma\cos\theta \quad . \tag{3}$$

Of course, γ_{wf} and γ_{wc} and hence θ will depend on the nature of the interactions be-

tween the wall and the colloidal particles: we assume planar, structureless walls with either hard-core repulsion or a smoother repulsion, described by the Weeks-Chandler-Andersen (WCA)-potential[10]

$$V_{WCA}(z) = 4\epsilon[(\sigma_w/z)^{12} - (\sigma_w/z)^6 + 1/4], \quad 0 \leq z \leq \sigma_w 2^{1/6} \, , \tag{4}$$

with its strength (ϵ) and range (σ_w) providing two parameters to control θ.

In the present paper, we focus on Monte Carlo methods by which the wall free energies γ_{wf} and γ_{wc} of the fluid and crystal phases of the present model system, Eq. 1, can be computed, and hence estimates for the contact angle obtained, if $\gamma_{wf} - \gamma_{wc} < \gamma$ (i.e., in the region of incomplete wetting).

2 Wall Free Energies of the Hard Sphere Model

As is well known, free energies are not straightforward outputs of simulation methods[12], and hence several methods have been applied and compared[11], in order to assert that the desired accuracy is reached, which is necessary to allow meaningful statements on the wetting behaviour.

One technique relies on the anisotropy of the pressure tensor $p_{\alpha\beta}(z)$ in a system bounded by two walls at $z = 0$ and $z = D$ (α, β are Cartesian coordinates). Defining the normal pressure $P_N = P_{zz}$ (note that p_{zz} in equilibrium is independent of the normal distance z) and $P_T = \int_0^D dz[p_{xx}(z) + p_{yy}(z)]/(2D)$, γ_{wf} becomes[13,14]

$$\gamma_{wf} = (P_N - P_T)D/2 \quad , \tag{5}$$

provided the two walls can be considered as non-interacting. As an example, Fig. 2 shows the density profile $\rho(z)$ for typical choices of system parameters[11]. One sees that the walls affect the liquid structure over an appreciable distance (the density oscillations mean a stacking of parallel rather densely packed layers). For a packing fraction $\eta = 0.47$ this layering effect has practically decayed to zero in the centre of the thin film, and hence it is a reasonable approximation to treat the walls as non-interacting. But this is no longer the case for $\eta = 0.49$.

A second difficulty for the application of Eq. 5 is that it is nontrivial to "measure" the pressure tensor for a hard sphere system (for systems with continuous potentials, this can be done more easily applying the virial theorem). However in this case, one can consider virtual volume changes by a factor ξ and compute the probability $P_{nov}(\xi)$ that there are no molecular pair overlaps when the volume is decreased from V to $V' = V(1 - \xi)$. For small ξ it can be shown that $P_{nov}(\xi) = \exp(-b\xi)$, where $b > 0$ is related to the bulk pressure by a relation[14] $P/(\rho k_B T) = 1 + b/N$, where $\rho = N/V$ is the bulk density in the system. For a thin film geometry, this method can be extended to sample P_N, P_T separately (changing D to $D' = D(1 - \xi)$ keeping L constant one obtains P_N, while changing L to $L' = L(1 - \xi)$ at fixed D yields P_T)[14].

An alternative method[11] is the "ensemble switch method", where one gradually switches between a system without walls, applying periodic boundary conditions throughout, described by Hamiltonian $\mathcal{H}_1(\vec{X})$, and a system with the same particle number and

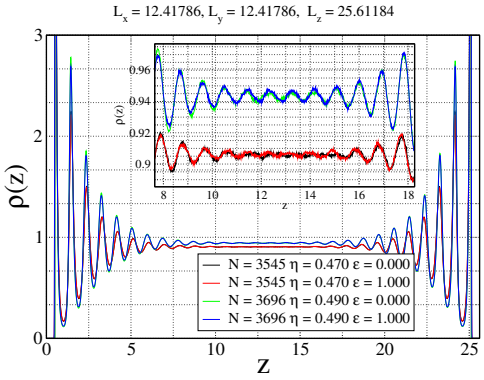


Figure 2. Density profile $\rho(z)$ vs. z (in units $\sigma = 1$) for a $L \times L \times D$ box with periodic boundary conditions in x-and y-directions and linear dimensions $L = 12.41786$, $D = 25.61184$, and a number N of hard spheres of $N = 3545$ or $N = 3696$, respectively, leading to packing fractions $\eta = 0.47$ and $\eta = 0.49$ (remember that bulk crystallization occurs for $\eta_c = 0.492$). Two choices of wall potential are shown, simple hard walls (denoted as $\epsilon = 0$ here) and WCA walls {Eq. 4, with $\epsilon = 0$}. Inset shows $\rho(z)$ in the centre of the film on magnified ordinate scale. From Deb et al.[11].

volume but with walls, $\mathcal{H}_2(\vec{X})$[15]. Here $\vec{X}$ stands symbolically for a microstate in the phase space of the system. At an intermediate step the total Hamiltonians $\mathcal{H}(\vec{X})$ is

$$\mathcal{H}(\vec{X}) = (1-\kappa)\mathcal{H}_1(\vec{X}) + \kappa\mathcal{H}_2(\vec{X}). \tag{6}$$

When the parameter κ varies from zero to unity, $\mathcal{H}(\vec{X})$ switches from $\mathcal{H}_1(\vec{X})$ to $\mathcal{H}_2(\vec{X})$. In a simulation κ is typically discretized and the system is allowed to move from κ_i to κ_{i-1} or κ_{i+1} with a Metropolis acceptance criterion ruling such a move. The free energy difference between the two states is given by $k_BT(\ln P(i) - \ln P(i-1, i+1))$ where $P(i)$ is the relative probability that the system is in state i. Since this probability varies over a wide range when κ changes, we employ a variant of Wang-Landau sampling to simulate each state with equal probability. In this way we sample free energy differences relative to the free energy of the system with periodic boundary conditions as a function of κ. For $\kappa = 1$, the wall free energy then follows as

$$\gamma_{wf}(\rho, T) = \lim_{D\to\infty} \Delta F(D)/(2k_BTA), \tag{7}$$

A being the area of the wall. As an example, Fig. 3 shows the wall-fluid surface free energy of the hard sphere model both for the case of hard walls and for walls where a WCA potential {Eq. 4} acts.

While the ensemble switch method has also been found to be very useful to compute the wall-crystal free energy γ_{wc}, none of the methods described so far would be suitable to obtain the surface excess free energy due to crystalline clusters coexisting with surrounding fluid. For this purpose, the technique recently proposed by Schilling and Schmid[17] should be useful, by which the absolute free energy of a (disordered) system is computed. This is a variant of the well-known thermodynamic integration method due to Frenkel and Ladd[18]

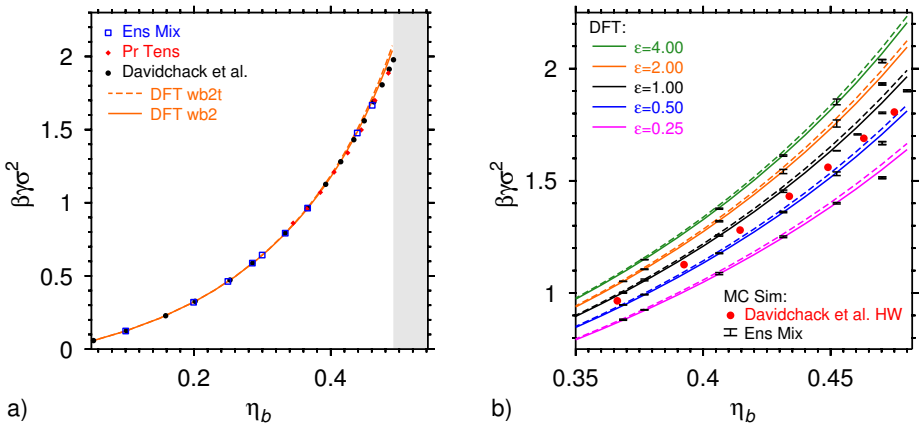


Figure 3. Plot of the wall-fluid surface free energy density γ_{wf} versus bulk packing fraction η_b of the colloids for a hard sphere system. (a) Hard walls: Both the pressure anisotropy method {Eq. 5} and the ensemble switch method {Eq. 7} are used and compared to previous simulation data due to Laird and Davidchack[16] and to advanced density functional theory (DFT) calculations[11]. (b) Same as (a), but for walls with a WCA potential. Several choices for the parameter ϵ in Eq. 4 are included. Again a comparison to DFT results is presented. From Deb et al.[11].

to compute the free energy of crystals, where $\mathcal{H}_1(\vec{X})$ is the Hamiltonian of an "Einstein crystal" of non-interacting particles bound to the crystal sites by harmonic springs, and $\mathcal{H}_2(\vec{X})$ the Hamiltonians of the considered crystal, a thermodynamic integration is performed. For the Einstein crystal, the free energy can be computed analytically. The new method[17] simply rests on the observation, that one can use the "Einstein crystal" as well as the "Einstein fluid" as a reference state, i.e. at a snapshot configuration of a fluid the interactions are turned off and replaced by springs fixing the particles at the positions of this particular (but representative) configuration.

As a first test of this method for an inhomogeneous system bounded by walls, Fig. 4 (left part) shows the free energy of a hard sphere system per particle, using a $L \times L \times D$ system with two WCA walls (with $\epsilon = 1$), plotted vs. $2/D$ to extract both the bulk free energy $f_b(\eta_b)$ and γ_{wf} from the "first principles" expression[13]

$$f(\eta_b, D) = f(\eta_b) + (2/D)\gamma_{wf} \tag{8}$$

Fig. 4 demonstrates that both $f(\eta_b)$ (from the intercept at the ordinate) and γ_{wf} (from the slope) can be extracted with meaningful accuracy, although the computational effort is considerable. The next step is the use of this method for the AO model. For the latter model, we have shown that it is possible to stabilize wall-attached crystalline "droplets" which have a large contact angle (Fig. 4, right part). These "droplets" are produced in the following way: We prepare a system at a packing fraction η inside the two-phase coexistence region, $\eta_f < \eta < \eta_m$. In the thermodynamic limit, we would expect in equilibrium that the crystal phase takes a volume fraction $x = (\eta - \eta_f)/(\eta_m - \eta_f)$ while the fluid phase takes a volume fraction $1 - x$ (according to the standard lever rule). To facilitate equilibration of such a two-phase state, an artificial initial condition is created,

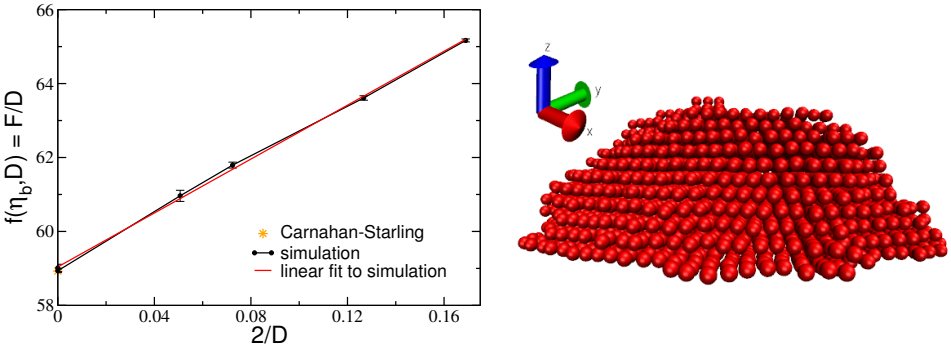


Figure 4. (a) Extrapolation of the free energy $f(\eta_b, D)$ with WCA walls ($\epsilon = 1$) as obtained by the Schmid-Schilling method[17] vs. $2/D$, for a hard sphere fluid at $\eta = 0.3686$. Note that $f_b(\eta_b)$ is known accurately from the Carnahan-Starling equation[19] (left part). (b) Side view of a typical snapshot picture of an equilibrated wall-attached (crystalline) droplet in the AO model at $\eta_b = 0.509914$ ($> \eta_f = 0.494$). Particles are classified as belonging to the fluid or crystal phases using the $q_6 q_6$-parameter. Fluid particles are not shown.

where a perfect face-centred cubic crystallite containing a fraction x of all the particles in the simulation box is placed into the system, with closely packed planes parallel to the substrate. The packing fraction inside of the crystal is chosen as η_m. The remaining volume is filled by fluid with packing fraction η_f, avoiding forbidden overlaps between particles in the crystal-fluid interfacial region. This state then is carefully equilibrated (note that if x is too small, the crystal melts, in analogy to the fluid droplet evaporation at the vapor-liquid transition[20]). It is possible to identify a favourable range of volume fractions η and associated droplet sizes, where "measurements" of the contact angle and volume as well as shape of these droplets are possible.

3 Wall Free Energies of the Asakura-Oosawa (AO) Model

From the comparative study of methods for the hard sphere model one can conclude that the ensemble switch method {Eqs. 6,7} is particularly suitable to obtain both γ_{wf} and γ_{wc}. Hence we focus on the use of this method here. In addition to the WCA walls {Eq. 4} a wall potential LJP1 was used, similar to a shifted Lennard-Jones (LJ) potential, where one can control the range and strength of the attractive tail by a parameter ϵ

$$U_w(z,\epsilon) = 4\epsilon[(\sigma/z)^{12} - (\sigma/z)^6] + 1, \; z \le z_c = \sigma \sqrt[6]{2(\epsilon + \sqrt{\epsilon^2 - \epsilon})}, \qquad (9)$$

$U_w(z,\epsilon) = 0$ for $z > z_c$. Fig. 5 shows a typical example of the application of the ensemble switch method, using this potential. One chooses $L \times L \times L_z$ geometries with $L = 8.05$ and L_z is varied between $L_z = 8.05$ and $L_z = 64.4$, to allow an extrapolation towards $1/L_z \to 0$. Particular care is needed for the wall-crystal free energy, since the geometries must be chosen such that the crystal is strain free. Since the closely packed (111) planes of the crystal are parallel to the walls, linear dimensions L_x, L_y must be compatible with an integer number of rows of particles at a spacing resulting from the

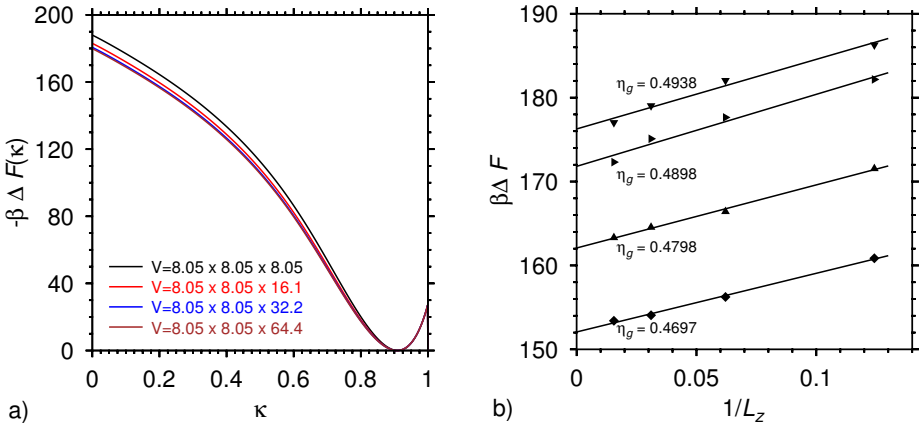


Figure 5. Typical example showing how wall tensions γ_{wf} are extracted using the ensemble mixing method, using the AO model and the LJP1 wall potential {Eq. 9}. Part (a) shows the free energy difference $\Delta F(\kappa)/k_BT$ vs. κ for $\epsilon = 2$, $\eta_b = 0.4697$. Part (b) shows the extrapolations to $1/L_z \to 0$ for several packing fractions.

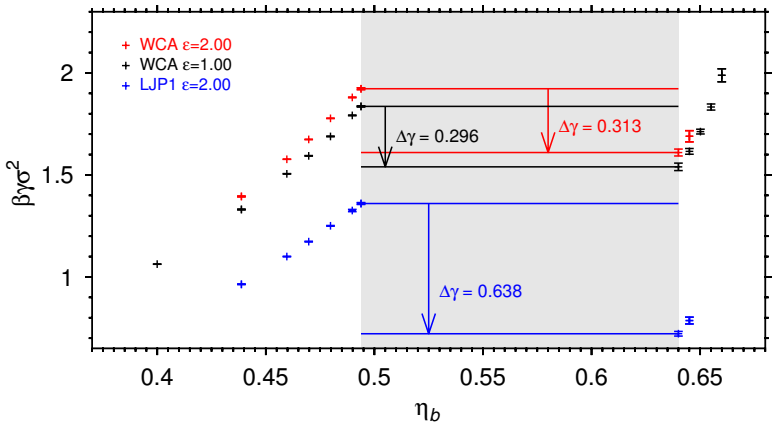


Figure 6. Wall tensions of the AO model plotted vs. bulk colloid packing fraction η_b, as obtained from the ensemble mixing method. Two choices of WCA potentials and one choice of the LJP1 potential are included, and the resulting differences $\Delta\gamma = \gamma_{wf} - \gamma_{wc}$ are indicated. The shaded area shows the liquid-crystal coexistence region.

density (and L_z must be compatible with an integer number of (111) planes stacked on top of each other in the standard face-centred cubic arrangement). Fig. 6 shows our final results for the wall tensions of the AO model.

4 Conclusions

In this investigation, a comparative exploration of methods to obtain wall-liquid and wall-crystal interface tensions was carried out, and their accuracy and efficiency was established.

The newly developed ensemble mixing method was found to be most useful, and could be applied to estimate $\Delta\gamma = \gamma_{wf} - \gamma_{wc}$ both for the hard sphere and AO models (using different choices of wall potentials). In future work, these results will be used to analyze wall-attached crystalline nuclei in the two-phase coexistence region (such as Fig. 4), to provide an understanding of heterogeneous nucleation phenomena.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grants No BI 314/19 (D. D.), TR6/C4 (D. W.), TR6/A5 (A. W.). We are grateful to NIC Jülich for a generous grant of computer time at the JUROPA supercomputer of the JSC. We are grateful to J. Horbach, M. Oettel, M. H. Yamani, and T. Zykova-Timan for a fruitful collaboration (Refs. 6,11).

References

1. P. N. Pusey, in *Liquids, Freezing, and Glass Transition* (J. P. Hansen, D. Levesque, and J. Zinn-Justin, eds.) North-Holland, Amsterdam (1991) p. 763.
2. H. Löwen, J. Phys.: Condens. Matter **20**, 404201, 2008.
3. S. Asakura and F. Oosawa, J. Chem. Phys. **22**, 1255, 1954.
4. W. C. K. Poon, J. Phys.: Condens. Matter **14**, R 859, 2002.
5. M. Dijkstra, R. van Roij, and R. Evans, Phys. Rev. E **59**, 5744, 1999.
6. T. Zykova-Timan, J. Horbach, and K. Binder, J. Chem. Phys. **133**, 014705, 2010.
7. D. Turnbull, J. Appl. Phys. **21**, 1022, 1950.
8. D. Winter, P. Virnau, and K. Binder, Phys. Rev. Lett. **103**, 225703, 2009.
9. D. Bonn, J. Eggers, J. Indekeu, J. Meunier, and E. Ratley, Rev. Mod. Phys. **81**, 939, 2009.
10. J. D. Weeks, D. Chandler, and H. C. Andersen, J. Chem. Phys. **54**, 5237, 1971.
11. D. Deb, A. Winkler, M. H. Yamani, M. Oettel, P. Virnau, and K. Binder, J. Chem. Phys. **134**, 214706, 2011.
12. D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulation in Statistical Physics, 3rd ed.* (Cambridge Univ. Press, Cambridge, 2009).
13. J. S. Rowlinson and B. Widom, *Molecular Theory of Capillarity* (Clarendon, Oxford, 1982).
14. E. De Miguel and G. Jackson, Mol. Phys. **104**, 3717, 2006.
15. M. Heni and H. Löwen, Phys. Rev. E **60**, 7057, 1999.
16. B. B Laird and R. L. Davidchack, J. Phys. Chem. C **111**, 15902, 2007, J. Chem. Phys. **132**, 204101, 2010.
17. T. Schilling and F. Schmid, J. Chem. Phys. **131**, 231102, 2009.
18. D. Frenkel and A. J. C. Ladd, J. Chem. Phys. **81**, 3188, 1984.
19. J. P. Hansen, and I. R. MacDonald, *Theory of Simple Liquids, 3rd ed*. (Academic, New York, 2006.
20. M. Schrader, P. Virnau, and K. Binder, Phys. Rev. E **79**, 061104, 2009 and references therein.

Coupled Lattice Boltzmann and Molecular Dynamics Simulations on Massively Parallel Computers

Jens Harting[1,2], **Stefan Frijters**[1], **Florian Janoschek**[1,2], **and Florian Günther**[1]

[1] Department of Applied Physics, Eindhoven University of Technology
Den Dolech 2, 5600MB Eindhoven, The Netherlands
E-mail: {f.a.r.janoschek, s.c.j.frijters, f.s.guenther}@tue.nl

[2] Institute for Computational Physics, University of Stuttgart
Pfaffenwaldring 27, 70569 Stuttgart, Germany
E-mail: jens@icp.uni-stuttgart.de

Complex colloidal fluids play an important role in many branches of industry. Examples include particle-stabilized emulsions in cosmetics, food and oil industries. Their understanding requires a study which resolves their microscopic structure, while still attaining sufficiently large length scales. The lattice Boltzmann method, owing to its high degree of locality, allows for such studies, when parallelized on modern supercomputers. However, exploiting these possibilities on hundreds of thousands of cores is a non-trivial task. We report on our experiences when employing large fractions of the IBM Blue Gene/P system at the Jülich Supercomputing Centre for our simulations and summarize recent results on particle-stabilized emulsions.

1 Introduction

Stabilizing emulsions by employing colloidal particles is a very attractive tool in the food, cosmetics and medical industries. The underlying microscopic processes of this stabilization can be explained by assuming an oil-water mixture. Without any additives, these liquids will phase separate, but the mixture can be stabilized by adding small particles. If these particles diffuse to the interface they will adsorb there, reducing interfacial free energy and stabilizing the mixture. The particles in these mixtures block Ostwald ripening, which is one of the main processes leading to drop coarsening in emulsions. Thus, blocking this process allows for long-term stabilization of these emulsions. The stabilized drops can be used to produce new materials with complex hierarchical structure. Interestingly, many of the properties of such systems cannot be explained with theories derived for surfactant-stabilized systems. This is due to the larger size of the particles compared to the surfactant molecules, and their lack of amphiphilic properties. New theoretical models have been developed (and have been underlined experimentally), incorporating specific features of particle-stabilized systems that have no direct analogue in surfactant systems. Quantitatively, however, the description of these systems still leaves to be desired. A promising method to understand the dynamic properties of particle-stabilized multiphase flows can be found in computer simulations. However, a suitable simulation algorithm must be able to deal not only with simple fluid dynamics, but it is also required to simulate several fluid species and particle-particle and particle-fluid interactions. Some recent approaches trying to solve these problems utilize the lattice Boltzmann (LB) method for the description of the solvents[1]. The LB method can be seen as an alternative to conventional Navier-Stokes solvers and is well-established in the literature. A number of multiphase and multicomponent models that are comparably straightforward to implement exist, making it attractive

for our current application. In addition, boundary conditions have been developed to simulate suspended finite-size particles in flow. Microscopic studies such as these allow for a direct link to macroscopic experimental data and will lead to a sustantial improvement of our understanding of particle-laden multiphase flows. These simulations remain computationally very challenging because typical particles need to be at least 10 LB length units in diameter and typical droplet diameters should be larger by an order of magnitude. To reduce finite-size effects and to acquire a statistically relevant number of droplets, the side length of the volume of interest easily reaches 1000-2000 LB length units—touching the limit of what is currently possible on high-end supercomputers. Another constraint is given by the very small time step in the simulations. The hydrodynamic interactions between the particles need to be resolved on the same scale as the particle motion, pushing our time resolution down to the nanosecond scale. To still be able to reach steady states, typical simulation runs require millions of LB time steps. This means that 0.5 to 1 rack months on the Blue Gene/P system "JUGENE" in Jülich might be required for a single simulation.

2 Simulation Method and Implementation

We use the lattice Boltzmann (LB) method as the basis for our simulations[1]. In our simulations, we effect the fluid-fluid interactions through the Shan-Chen multicomponent model[2,3], while we use a variant of Ladd's method to couple the particles to our fluid[4,5]. Particle-particle interactions and particle dynamics are modelled by a molecular dynamics algorithm. The code has recently been expanded to allow for spherical particles as well as ellipsoidal ones. For a more detailed description of our simulation methods the reader is referred to Ref. 6, 7. The development of our simulation code LB3D started already in 1999 as a parallel LB solver. As the scientific focus at that time was on the behaviour of complex fluid mixtures under shear, the ability to model up to three fluid species, one of them with amphiphilic properties, using Shan and Chen's aforementioned approach was one the first features implemented. Already in 2004, LB3D was awarded a gold star rating by the Edinburgh Parallel Computing Centre for scaling almost linearly to 1024 processors and it also won several further prizes as part of the TeraGyroid project[8]. Further applications of the code encompass flow in porous media[9] and fluid interactions with rough and hydrophobic surfaces[10–13]. In the meantime, the application was ported to most supercomputing platforms available. A parallel molecular dynamics (MD) code was integrated into LB3D in 2008. After considerable changes and extensions both parts of the code now use the same 3-dimensional spatial decomposition scheme, the Message Passing Interface (MPI) for communication, and modern Fortran 95 language elements. The simultaneous availability of Lagrangian particles in the same code as the Eulerian LB fluid opened up new applications for LB3D, such as tracking the velocity field in micromixers using massless tracer particles[14], a coarse-grained model for red blood cells[15], or colloidal particles with variable wettability as described in this report[6]. Due to the strong locality of the LB equation and because the relevant interactions in colloidal systems are short-range, an efficient parallelization of such systems can be expected. In the past, LB3D indeed showed very good scalability. On a machine such as JUGENE, the IBM Blue Gene/P system at Jülich Supercomputing Centre (JSC), with 294 912 cores, however, it is crucial to minimize communication costs and serial parts of a code even further in order to maintain high efficiency when running on large portions of the machine. Initially, LB3D showed only low

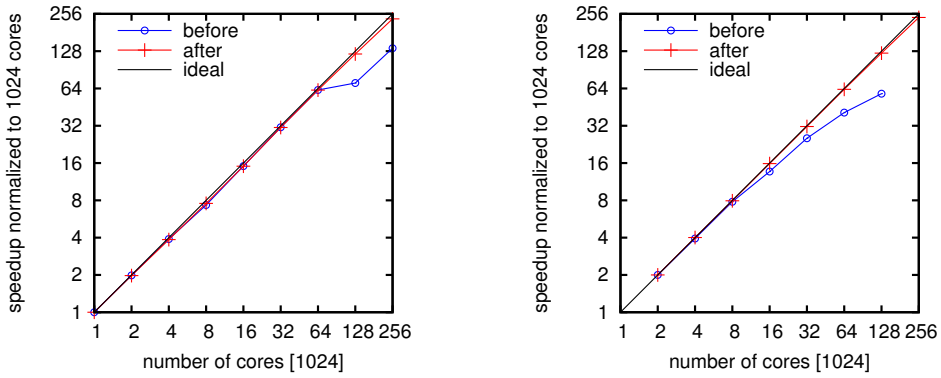


Figure 1. Strong scaling of LB3D on the Blue Gene/P before and after our optimizations. (a) relates to a system with only one fluid component so the effect of matching or mismatching topologies of network and domain decomposition can be examined better. (b) refers to a system with two fluid species and suspended particles as they are of interest in this paper. The absolute execution times for small core counts did not change significantly (taken from Ref. 7).

efficiency there in strong scaling beyond $65\,536$ cores. At the Jülich Blue Gene/P Extreme Scaling Workshop 2011 we could relate this to a mismatch of the network topology of the domain decomposition in the code and the network actually employed for point-to-point communication. The Blue Gene/P provides direct links only between direct neighbours in a three-dimensional torus, so a mismatch can cause severe performance losses. We investigated this issue for a system of $1\,024^2 \times 2\,048$ lattice sites carrying only one fluid species and no particles. Limiting ourselves to this simpler system compared to the systems of interest later on has the benefit that the reduced ratio of computation to communication costs makes possible bottlenecks more visible. We find that if we allow `MPI_Cart_create()` to reorder process ranks and manually choose a domain decomposition based on the known hardware topology, efficiency can be brought close to ideal for this system. A comparison of the speedup before and after this optimization is shown in Fig. 1(a). Systems containing colloids and two fluid species were known to slowly degrade in parallel efficiency when the number of cores was increased. This is demonstrated in Fig. 1(b) for a benchmark system of $1\,024^2 \times 2\,048$ lattice nodes and $4\,112\,895$ uniformly distributed particles with a radius of 5 lattice units (resulting in a volume concentration of about $20\,\%$). Since the degradation was not visible for a pure LB system (Fig. 1(a)), it initially was attributed to load imbalances and communication overhead related to the suspended particles. During the Scaling Workshop we identified a non-parallelized loop over all particles in one of the subroutines implementing the coupling of the colloidal particles and the two fluids as the actual reason. Due to the low computational cost per iteration compared to the overall coupling costs for colloids and fluids, at smaller numbers of particles or CPU cores this part of the code was not recognized as a possible bottleneck. A complete parallelization of the respective parts of the code produced a nearly ideal speedup up to $262\,144$ cores

also for this system. Both speedup curves are depicted in Fig. 1(b). Actual production runs require checkpointing and the output of physical observables to disk, which is not accounted for in the benchmarks above. Pickering systems require output at most once per 100 to 1 000 LB time steps, so the performance of the output routines is of inferior importance. Using parallel HDF5 output allows us to store fluid density fields of $4.6\,\mathrm{GB}$ size for a system of $1\,024^2 \times 1\,152$ lattice sites and $454\,508$ particles within on average $29\,\mathrm{s}$ when using the whole system ($294\,912$ cores). This corresponds to the time required to simulate 100 LB steps, so even at maximum core count possible and maximum I/O frequency not more than $50\,\%$ of the time is spent on storing output. Particle configurations are typically smaller by two orders of magnitude than a density field. Therefore, this information is still written serially by the root process. Since strongly asymmetric point-to-point communication patterns are likely to produce buffer overflows and fail at high core counts, collective MPI operations are applied for data accumulation. However, the intuitive choice of `MPI_Gatherv()` is not optimal since on Blue Gene/P, the specially optimized implementation of `MPI_Allgatherv()` proved to be significantly faster at the cost of requiring one receive buffer per task. For `MPI_Allgatherv()` we measure an absolute time required to write one complete particle configuration of approximately $10\,\mathrm{s}$ and a speedup of 77 on $131\,072$ cores compared to `MPI_Gatherv()`. For even larger systems, however, `MPI_Allgatherv()` imposes a lower limit regarding the maximum number of particles, as receive buffers able to store the whole particle configuration need to be allocated four times per node (once per core) instead of only once on the node running the root process.

3 Simulation Results

We apply our method to study the deformation of a liquid droplet under steady shear. Here, the simulation setup is a typical Couette flow where the upper and lower boundary are moved by Lees-Edwards boundary conditions. A spherical droplet with radius R is initially placed at the centre of the system, which can be characterized by its capillary number $\mathrm{Ca} \equiv \frac{\eta\dot{\gamma}R}{\sigma}$. The surface tension σ can be related to the Shan-Chen parameter $g_{cc'}$, η is the dynamic viscosity and $\dot{\gamma}$ is the shear rate. The deformation of the droplet is measured by fitting an ellipse to a two-dimensional projection of the droplet, where L and B are the length and breadth of the ellipse, respectively: $D \equiv \frac{L-B}{L+B}$. As shown by Taylor and reconfirmed in our simulations, for two fluids of equal density and equal viscosity one obtains $D = \frac{35}{32}\mathrm{Ca}$ for $D \ll 1$[16,17]. While there have been numerous analytical studies, experiments and simulations on the deformation of simple liquid droplets within the last century, the case of particle covered droplets has not been covered in such detail and it is still not fully clear how the stability and deformability changes when particles have adsorbed to the liquid-liquid interface. In our simulations we can study these effects in detail by varying the droplet size, the number of particles, the particle size, their contact angle, as well as the shear rate and interfacial tension. Fig. 2 shows the deformation of a particle-covered droplet for various moderate capillary numbers. As can be observed from the plot, the deformation shows an almost parabolic behaviour and a non-uniform particle distribution can be observed at the interface (see inset figures). For even larger shear rates, the droplets break up. The capillary numbers have been computed based on the radius of the large original droplet and if one would rescale it with the radius of the small droplets

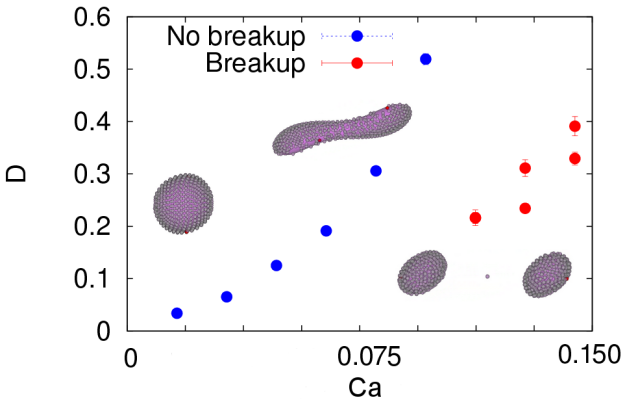


Figure 2. Droplets subjected to shear—but stabilized by particles—deform and might even break for high shear rates.

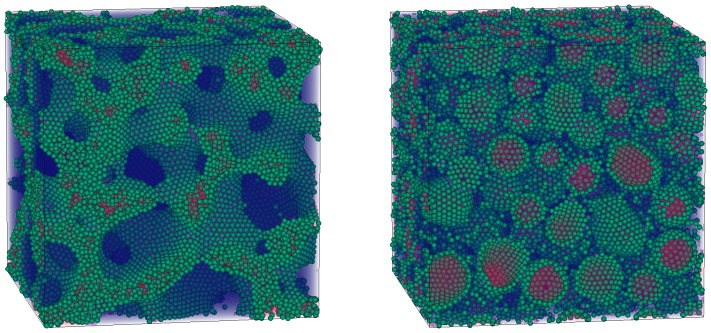
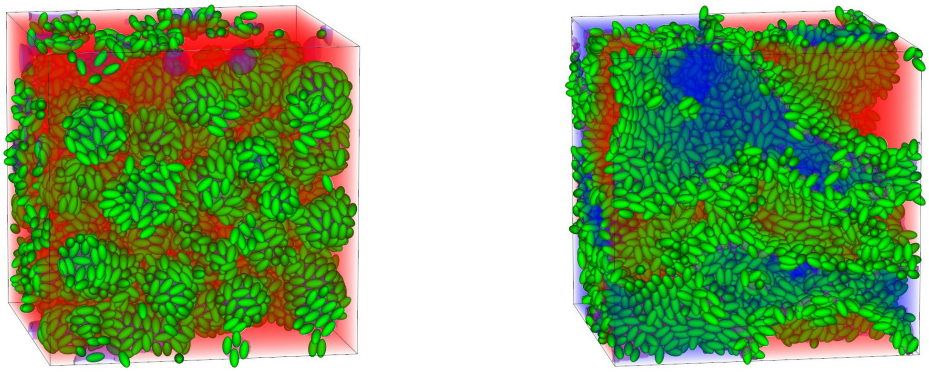


Figure 3. Snapshots of a 3D Bijel (a) and a Pickering emulsion (b): particles (green) self-assemble at the interface between two immiscible fluids (shaded in red and blue) and stabilize a bicontinuous interface (a) or droplets (b).

occurring after breakup, the curves would be shifted towards smaller Ca.

We studied in detail how particles travel towards the fluid-fluid interface and get jammed there generating fluid-bicontinuous gels (so-called "Bijels"[18]) as shown in Fig. 3a. Further, it was demonstrated that by modifying the lyophobic/lyophilic properties of the solved particles, the particles travel towards the fluid-fluid interface and accumulate at the surface of a droplet thus forming a "Pickering emulsion" (see Fig. 3b). Large scale parameter studies as well as investigations of the time dependent domain or droplet growth in Bijels and Pickering emulsions were performed. All simulations start from a random mixture of fluids and particles, modelling a temperature quench. At the beginning of the simulation, small scale droplets nucleate due to ballistic motion of particles and fluids. After the nucleation regime, the droplets grow due to spinodal decomposition. At some point the phase separation comes to a halt and droplets can only continue to grow due to coalescence (Ost-



(a) (b)

Figure 4. Two snapshots of emulsions stabilized by particles are shown for an aspect ratio $m = 2$, a volume concentration of $C \approx 0.2$ and a contact angle $\theta = 90$. (a) Pickering emulsion for a fluid radio of $5 : 2$. (b) Bijel for a fluid radio of $1 : 1$[7].

wald ripening), which is limited due to the particles that were captured at the fluid-fluid interface. Ostwald ripening is one of the main processes leading to drop coarsening in emulsions and foams; hence stopping it allows long-term stabilization. Time-dependent measurements like these are very hard or even impossible to perform experimentally, but the processes involved determine the final properties of the emulsion. Therefore, we expect to be able to contribute to a better understanding of how the final product depends on fluid composition, fluid properties, particle concentration, particle size, or the contact angle the fluids form with the particle surface. We have studied in detail the phase behaviour of our system and determined a phase diagram demonstrating for which values of the particle concentration, fluid decomposition or contact angle either a Bijel or a Pickering emulsion is formed[6]. It is our current main focus for these phases to understand the rheological properties of Bijels and Pickering emulsions in detail. For this, further large-scale sheared simulations will be performed in order to explore the complex non-Newtonian properties of these systems. Collecting the data from a number of such large-scale runs will allow us to quantitatively compare the size distribution of the stabilized droplets as well as the rheological properties of such systems with experimental data.

We also studied emulsions stabilized by non-spherical, anisotropic colloidal particles: ellipsoids with an aspect ratio of $m = 2$. Also for these ellipsoids, we observe two different phases: the Pickering emulsion (see Fig. 4(a)) and the Bijel (see Fig. 4(b)). Fig. 5 shows the transition from a Bijel to a Pickering emulsion for these ellipsoids with an aspect ratio of $m = 2$ and a particle volume concentration of $C \approx 0.2$. The two control parameters used for the study of the phase transition are the fluid ratio and the contact angle θ. The squares denote the configurations which lead to a Bijel whereas the circles denote a resulting Pickering emulsion. If the amount of the two fluids present in the simulation is equal or not too different (e.g. a ratio of $4 : 3$) we find a Bijel for all considered contact angles. However, if the fluid ratio is increased we find a Pickering emulsion. For intermediate fluid ratios the obtained phase depends on the chosen contact angle. For example for a ratio of

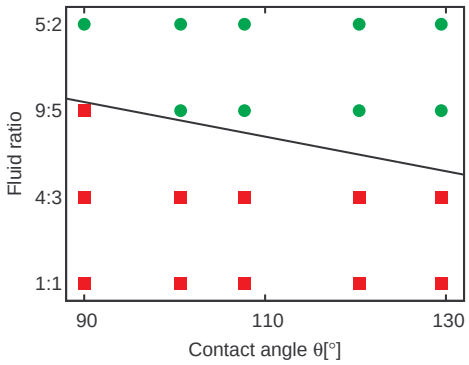


Figure 5. Phase diagram demonstrating the transition from a Bijel to a Pickering emulsion. The contact angle θ and the density ratio between the two fluids are varied. The squares denote the configurations which lead to a Bijel whereas the circles denote a Pickering emulsion[7].

$9:5$ we get a Bijel for a contact angle of $90°$ and a Pickering emulsion for all higher values of θ.

4 Conclusion and Outlook

Suspensions of colloids in multiphase flows can exhibit many surprising effects, which are of potential interest to both fundamental science and industrial applications.

Our investigation of the effect of colloids on the deformability of droplets has shown that when the fluid-fluid interface is almost saturated with colloids, the deformability of a droplet at a constant capillary number increases dramatically, and the breakup behaviour of those droplets at higher capillary number is also changed quantitatively. The results obtained for the particle-covered droplets will be compared to surfactant-covered droplets in similar systems, making a quantitative link between the two.

On slightly larger scale, we investigated the behaviour of the particle-stabilized Pickering emulsions and Bijels, and found that the arrival at either phase after a temperature quench depends on many properties of the system, such as fluid ratio, and size and contact angle of the colloids. Our simulation code has been extended to include ellipsoidal colloids, which will introduce another parameter that can influence the formation of a particular phase. In the future, Janus particles will also be considered, having different wetting properties on different parts of the surface of the particle.

Acknowledgements

Besides support from JSC and IBM experts present at the Blue Gene/P Extreme Scaling workshop 2011 we acknowledge computing resources that were granted via GSC grants `hss08` and `hss10` and by PRACE (project `pra036`).

References

1. S. Succi, *The Lattice Boltzmann Equation for Fluid Dynamics and Beyond*, Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2001.
2. X. Shan and H. Chen, *Lattice Boltzmann model for simulating flows with multiple phases and components*, Phys. Rev. E, **47**, no. 3, 1815–1819, Mar 1993.
3. H. Chen, B. M. Boghosian, P. V. Coveney, and M. Nekovee, *A ternary lattice Boltzmann model for amphiphilic fluids*, Proc. R. Soc. Lond. A, **456**, 2043, 2000.
4. C.K. Aidun, Y. Lu, and E.-J. Ding, *Direct analysis of particulate suspensions with inertia using the discrete Boltzmann equation*, J. Fluid Mech., **373**, 287, 1998.
5. A.J.C. Ladd and R. Verberg, *Lattice-Boltzmann simulations of particle-fluid suspensions*, J. Stat. Phys., **104**, no. 5/6, 1191–1251, 2001.
6. F. Jansen and J. Harting, *From bijels to Pickering emulsions: A lattice Boltzmann study*, Phys. Rev. E, **83**, no. 4, 046707, Apr 2011.
7. F. Günther, F. Janoschek, S. Frijters, and J. Harting, "Complex colloidal fluids: enabling a lattice Boltzmann solver for massively parallel simulations on the Blue Gene/P", submitted for publication, 2011.
8. J. Chin, P. V. Coveney, and J. Harting, *The TeraGyroid project – collaborative steering and visualization in an HPC grid for modelling complex fluids*, in: Proceedings of the UK e-Science All Hands Meeting, 2004.
9. A. Narváez, T. Zauner, F. Raischel, R. Hilfer, and J. Harting, *Quantitative analysis of numerical estimates for the permeability of porous media from lattice-Boltzmann simulations*, J. Stat. Mech: Theor. Exp., **2010**, P211026, 2010.
10. J. Harting, C. Kunert, and H.J. Herrmann, *Lattice Boltzmann simulations of apparent slip in hydrophobic microchannels*, Europhys. Lett., **75**, 328–334, 2006.
11. C. Kunert and J. Harting, *Roughness induced apparent boundary slip in microchannel flows*, Phys. Rev. Lett., **99**, 176001, 2007.
12. J. Hyväluoma and J. Harting, *Slip flow over structured surfaces with entrapped microbubbles*, Phys. Rev. Lett., **100**, 246001, 2008.
13. C. Kunert, J. Harting, and O.I. Vinogradova, *Random-roughness hydrodynamic boundary conditions*, Phys. Rev. Lett., **105**, 016001, 2010.
14. A. Sarkar, A. Narváez, and J. Harting, "Quantification of the degree of mixing in chaotic micromixers using finite time lyapunov exponents", submitted for publication; arXiv:1012.5549, 2011.
15. F. Janoschek, F. Toschi, and J. Harting, *Simplified particulate model for coarse-grained hemodynamics simulations*, Phys. Rev. E, **82**, 056710, 2010.
16. G.I. Taylor, *The Viscosity of a Fluid Containing Small Drops of Another Fluid*, Proc. R. Soc. Lond. A, **138**, no. 834, 41–48, 1932.
17. G.I. Taylor, *The Formation of Emulsions in Definable Fields of Flow*, Proc. R. Soc. Lond. A, **146**, no. 858, 501–523, 1934.
18. K. Stratford, R. Adhikari, I. Pagonabarraga, J.-C. Desplat, and M.E. Cates, *Colloidal Jamming at Interfaces: A Route to Fluid-Bicontinuous Gels*, Science, **309**, no. 4, 2198–2201, 2005.

Computational Studies of Insulating Magnetic Oxides

Marjana Ležaić, Phivos Mavropoulos, Konstantin Z. Rushchanskii, Ivetta Slipukhina, and Stefan Blügel

Peter Grünberg Institut and Institut for Advanced Simulation
Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {m.lezaic, ph.mavropoulos, k.rushchanskii, i.slipukhina, s.bluegel}@fz-juelich.de

We conduct computational, first principles studies of insulating magnetic oxides. These materials have a potentially crucial place in future electronic memory devices, for example as active barriers in magnetic tunnel junctions. While indisputably useful, insulating oxides with a net magnetization are very scarce. Therefore, significant effort is invested in understanding the origin of ferromagnetism in insulators. Computer experiments play an important role in these investigations, as well as in the design of new magnetic insulators with predetermined properties specific to the intended applications. We present here a study of the origin of ferromagnetic spin-coupling in N-doped MgO, as well as the computational design of a multiferroic material for an experimental search for the permanent electric dipole moment of the electron.

1 Introduction

In the past decades, we have been witnessing a rapid development of various mechanisms and techniques for storing, manipulating and securing information. This is an important component in industrial development: the world-wide hard disk revenue is about to reach 30 billion Euros and is still rising. Many devices used in everyday life are becoming smaller, faster and have multiple functionalities. Ever increasing demands of the market are a great challenge for the information storage industry. While different research groups are working on further improving the existing technologies, new ones are already being conceived and their potentials explored in search for the ones that could bring a breakthrough in the future. This is a challenging task which involves conducting and combining research in different areas and at different length scales. An important position in this race is taken by the computer experiments, where the behaviour of the new devices can be simulated at every step of their creation, from materials engineering at the atomic level, to the final circuits. The multifunctionality of the new devices puts stringent constraints on the materials that can be used: the higher the functionality, the more complicated the components. For example multifunctional oxides, that are intensively studied by many groups around the world, have a very high potential for future applications. However, this comes at a price: such materials usually have a very complex crystal structure, with unit cells consisting of many atoms and strongly interacting magnetic and structural properties. This complexity necessarily presents new challenges for crystal growers and device engineers, but the theoreticians are not carefree either: translated to the computing language, simulating properties of such multifunctional materials means a very high memory and CPU-time consumption. So we are in a sort of an infinite spiral, where we need ever stronger computers in order to create the newer, even stronger ones.

One of the device concepts that was, e.g, implemented in 2005 in the new type of drives produced by Seagate, relies on the so-called *tunneling magnetoresistance* (TMR) effect in *magnetic tunnel junctions* (MTJs). The working principle of MTJs is shown in

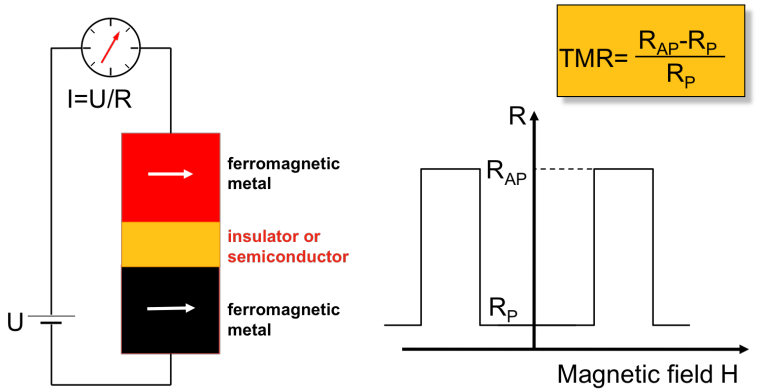


Figure 1. Working principle of a magnetic tunnel junction.

Fig. 1: a junction is made of two ferromagnetic electrodes separated by an insulating layer (*barrier*). When voltage is applied, a weak (*tunneling*) current runs through the circuit and the resistance it experiences is measured. This resistance can have two values: a high one, $\mathrm{R_{AP}}$, when the magnetization directions in the two ferromagnetic electrodes are antiparallel, or a low one, $\mathrm{R_P}$, when they are parallel. The junction is characterized by the so-called TMR ratio, TMR$= (\mathrm{R_{AP}} - \mathrm{R_P})/\mathrm{R_P}$. The mutual orientation of the magnetization directions in the electrodes can be controlled by an external magnetic field and can be used to store information (1bit= 2 resistive states, $\mathrm{R_{AP}}$ and $\mathrm{R_P}$). It is worth mentioning that such a tunneling current is purely a quantum effect which has no analogue in classical physics.

While there is still a substantial effort to increase the efficiency of MTJs, the next generation junctions, with a barrier material which also plays an active role (in contrast to the present passive barriers) are already investigated. Usually one considers a ferromagnetic, ferroelectric or, ideally, a *multiferroic* (magnetic and ferroelectric in the same phase) barrier which in combination with the ferromagnetic electrodes offers a possibility of multi-bit memory. Presently, the main obstacle on the road to applications of such junctions is the scarcity of multiferroic materials (especially the ones that retain their multiferroic properties at room temperature).

Here we present some of our computational studies of insulating magnetic oxides. In Sec. 2 we describe the computational method and in Sec. 3 we outline two example studies: doping as a way to induce magnetism in nonmagnetic oxides and an unusual application of a (computationally designed and consequently synthesized) low-temperature multiferroic, (Eu,Ba)$\mathrm{TiO_3}$. For our studies of room-temperature multiferroics with potential applications in memory devices, we refer the interested reader to Ref. 1.

2 Computational Method

We investigate the properties of materials using *first principles* calculations. This means that the calculations rely on the laws of quantum mechanics, without adjustable parameters. A crystal is a many-body system described by the Schrödinger equation. However, due to

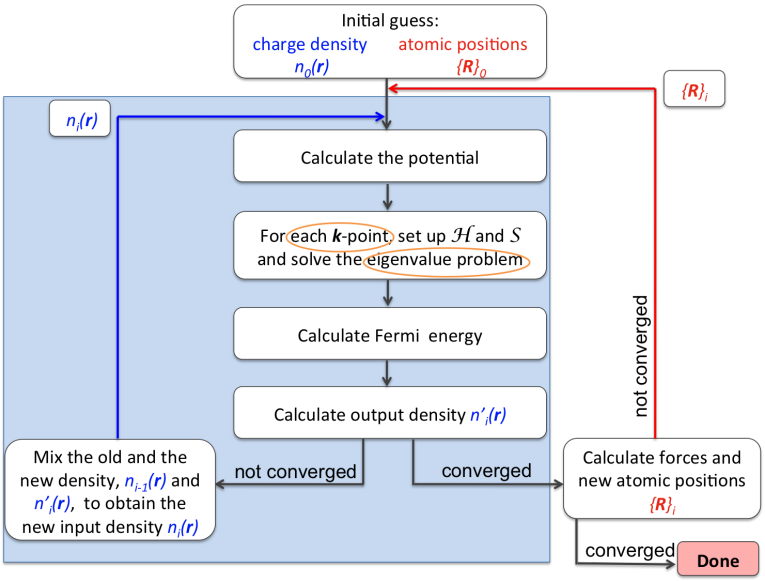


Figure 2. Self-consistency cycle of the electronic ground state (indicated in the blue rectangle) and a self-consistent determination of the atomic positions (indicated by the red arrow).

the very large number of electrons in such a system, this equation would be impossible to solve. Density functional theory solves this problem and describes the system by the Kohn-Sham equations, which are differential quantum-mechanical equations similar to the Schrödinger equation. In practice, this is done at the price if introducing an approximation to the electron-electron interaction.

The main property entering the equations (along with the fixed atomic positions) is the electron density n. A set of wavefunctions is determined from the equations that, in return, determine the output electron density n' and the process is repeated until the output density is (within a predefined tolerance) equal to the one entering the equations. This process is called *electronic self-consistency* (see Fig. 2). If in the specific method of calculation a basis set is used, the equations which are solved during the self-consistency cycle can be expressed as

$$\mathcal{H}[n]c = \varepsilon \mathcal{S}c, \tag{1}$$

i.e. as a standard eigenvalue problem. Here, the eigenvector c specifies the wavefunction in the given basis and ε is the corresponding eigenvalue. $\mathcal{H}$ is the Hamiltonian matrix and $\mathcal{S}$ the overlap matrix occurring in the case of non-orthogonal basis sets.

The dimensions of the matrices $\mathcal{H}$ and $\mathcal{S}$ are determined by the number of the basis functions. Due to the translational symmetry of periodic crystals, these matrices can be block-diagonalized. Each of the blocks is labeled by a vector called ***k**-point*. As the system is infinite, the blocks on the diagonal will be infinitely many. Their size will grow with the volume of the unit cell of the crystal, i.e. the more complex the system, the big-

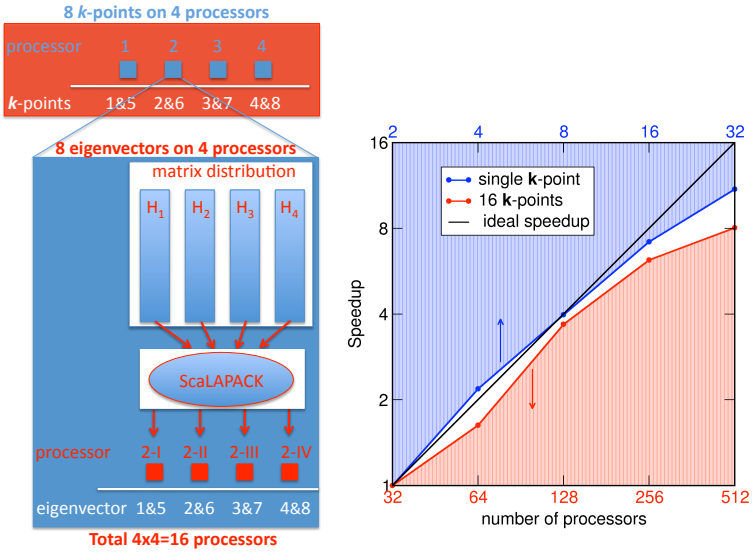


Figure 3. Left: An example of $\boldsymbol{k}$-point- and eigenproblem parallelization in FLEUR code. Right: the speedup achieved on JUROPA supercomputer by the parallelization, for a system with 82 atoms (matrix size 9500x9500). The blue line corresponds to the eigenproblem parallelization with 1 $\boldsymbol{k}$-point only and the red one to the same system with 16 $\boldsymbol{k}$-points and the use of both $\boldsymbol{k}$-point- and eigenproblem parallelization (data courtesy of Dr. G. Bihlmayer).

ger the matrices. In reality, only a finite number of these matrices is calculated and for the rest an interpolation in $\boldsymbol{k}$-space is done. For a system with a big unit cell, we need a smaller number of explicitly calculated matrices (smaller number of $\boldsymbol{k}$-points) for the interpolation. In contrast to this, for the interpolation of a system with a small unit cell, we need many calculated matrices (many $\boldsymbol{k}$-points). Consequently, we either have to solve few large eigenvalue problems, or many small ones. Therefore, depending on the size of the system one wants to calculate, an appropriate parallelization scheme can be applied. In the FLEUR code[2], which is developed at our institute, two schemes are implemented (see Fig. 3): for the smaller systems, where one solves many small eigenvalue problems, a simple distribution of these eigenproblems over many processors is a straightforward solution; for the large systems, there is the option of parallelization of the large eigenproblems, which is handled over ScaLAPACK.

It is worth mentioning that in case of magnetic systems with e.g. spin-orbit coupling or non-collinear magnetism the problem doubles in size. However, even collinear magnetism can in some antiferromagnetic systems enlarge the problem by increasing the unit cell size. Furthermore, various lattice distortions that are often found in oxides, charge order, or orbital order require unit cells that are sometimes much larger than the chemical formula unit. Finally, the coexistence of two (or more) orders in the same material like, e.g., magnetic and charge-order, demands a unit cell that can contain both (or more) order patterns. It then becomes clear why the multiferroic materials, or their junctions, are computationally extremely demanding.

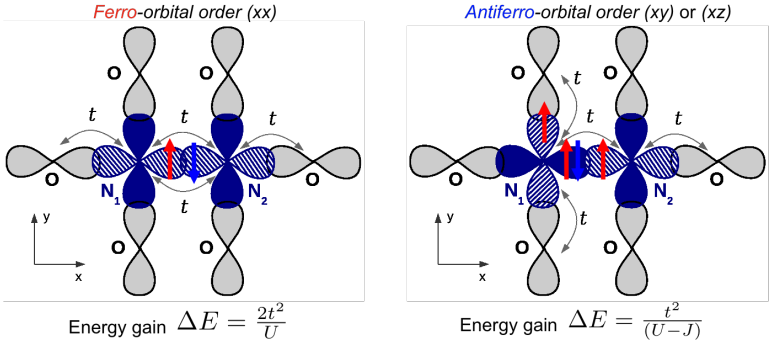


Figure 4. Ferro-orbital (left) and antiferro-orbital (right) arrangement of the half-filled Nitrogen p-orbitals in N-doped MgO. The arrows indicate the preferred spin arrangement: AFM for the ferro-orbital and FM for the antiferro-orbital arrangement. When the Coulomb intersite interaction is neglected, the configuration on the left is the ground state (as indicated by the corresponding energy gain). First principles calculations, however, identify the configuration on the right as the ground state.

3 Applications

3.1 How to Make a Nonmagnetic Oxide go Magnetic?

In the recent years, a new strategy in engineering ferromagnetic insulators is at work: instead of the traditional doping of nonmagnetic insulators with transition-metal ions, doping with sp-elements is explored. The observed magnetism is associated with the partially filled p-electronic states of the dopants (or of Oxygen, induced by the dopants) and is usually referred to as *sp-magnetism*. Among the studied insulators, MgO takes a special place as currently the most important barrier material for MTJs. Experimentally ferromagnetism in Nitrogen-doped MgO films was recently demonstrated[3].

For a nonmagnetic system to become magnetic due to doping, several conditions have to be fulfilled: (*i*) local magnetic moments have to be formed (usually, on dopants) (*ii*) the locally formed moments should interact with each other; note that the radius of such interaction is limited, and usually does not extend further than a couple of lattice constants in doped systems (*iii*) the magnetic moments should form a network spreading throughout the sample. In order to achieve ferromagnetism in the sample, the interaction of the local moments should be ferromagnetic, i.e. the atomic spins should be mutually parallel. The aim of our study is to determine what drives the ferromagnetic interaction of the spins located on N-dopants in MgO[4,5].

First, one should understand the origin of the local magnetic moment formed on N dopants. Pure MgO is a nonmagnetic band insulator. Due to its large electronegativity, Oxygen attracts two electrons from the Mg outer electronic shell in order to complete its own outer shell with eight electrons. Six of these electrons are situated in p-orbitals, two in each one of p_x, p_y and p_z (according to Pauli principle, each orbital can take maximum two electrons of opposite spins). In N-doped MgO, N substitutes one O atom in the crystal lattice. However, N has one electron less, which means that one of the three p-orbitals remains occupied with only one electron. The $2p$ wavefunctions are rather localized on the N atom and so the spin of one unpaired electron gives rise to the magnetic moment

localized on this atom.

Next, we study the interaction of two N-impurities positioned next to each other in the MgO lattice. There are two points to pay attention two: (*i*) the mutual orientation of the half-filled orbitals on the two N atoms and (*ii*) the mutual orientation of spins on the two N atoms - ferromagnetic (FM) or antiferromagnetic (AFM). We define a coordinate system with the p_x and p_y orbitals pointing toward neighbouring N (or O) atoms in the xy plane (Fig. 4), while the p_z orbital points out of plane, toward Mg. We omit Mg atoms from the discussion because they do not contribute to the magnetic interaction between the N-dopants.

The half-filled p-orbitals on two N atoms can be parallel to each other without any overlap (yy and zz configuration) or with a direct overlap (xx configuration); further in a T-configuration the half-occupied orbitals on the two Nitrogens are perpendicular to each other (xy and xz configuration). Since the magnetic interaction is carried by the electrons that can move between the two N-sites, we can omit from our considerations the configurations in which only fully occupied Nitrogen orbitals overlap with each other, and we are left with only two configurations: xx (Fig. 4 left) and the T-configuration (Fig. 4 right). In order to understand what kind of magnetic interaction one can expect in these two configurations, three factors should be taken into account in a minimal model: the kinetic energy of the electrons (usually expressed by the so-called *hopping, t*), the onsite Coulomb-repulsion between the electrons U and the Hund's exchange J. The latter takes care that the Hund's first rule is fulfilled. This rule states that if on one atom two orbitals of the same energy (like two of the p orbitals here) are occupied by one electron each, then these electrons will have parallel spins. J expresses the energy gain of such a configuration compared to that with the two electrons having antiparallel spins. Considering the FM and AFM spin-orientation in both of the two orbital configurations, one arrives to the conclusion that energetically the most favourable state should be the one with xx-orbital and AFM-spin configuration, i.e. one would expect no net magnetization in N-doped samples of MgO.

Experimentally, however, the samples were found to be ferromagnetic. We also performed first principles calculations looking for the answer to this puzzle. We used supercells of 64 atoms of host MgO, with a $2 \times 2 \times 2$ $\boldsymbol{k}$-point mesh and investigated all possible mutual orientations of orbitals and spins on the two N atoms (12 different spin-orbital configurations in total). The calculations yielded a ferromagnetic ground state with the T-orbital configuration, in contrast to our expectations relying on the minimal model. The only answer could be that the model was incomplete. Indeed, it turns out that the three factors mentioned above were not accounting for all the important interactions. Extending our model to take into account the intersite Coulomb interaction which favours the T-shaped configuration, we arrive at the conclusion in accordance to our first principles findings[5].

The result that the FM spin-orientation is in fact stabilized by the intersite Coulomb interaction is rather exciting, as it is probably a characteristic of the systems presenting sp-magnetism exclusively, due to a larger spatial extent and directionality of the p-wavefunctions compared to the d-ones. However, further studies are needed in order to establish the role played by this interaction in insulators doped with transition metals.

3.2 Search for a Permanent Electric Dipole Moment of the Electron

Although most of the research of multiferroic materials is concentrated on the search for the multiferroics that would be used in room-temperature applications, we recently inves-

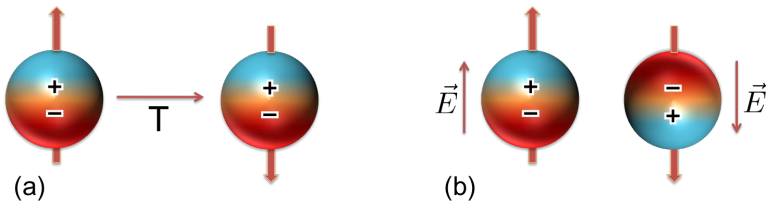


Figure 5. (a) Breaking of the time-inversion symmetry T due to the electric dipole moment of the electron. (b) The basic idea of the experiment searching for EDM: an electric field $\vec{E}$ controls the direction of the EDM and simultaneously the direction of the spin, leading to a net magnetization of the sample.

tigated the possibility of employing a multiferroic material in a solid-state based search for the permanent electric dipole moment of the electron[6], in an experiment conducted at liquid Helium temperature (4.2 K). Three intrinsic characteristics of the electron are so far established: mass, charge and the magnetic dipole moment (spin). If an electric dipole moment of the electron exists, it would be its fourth intrinsic characteristic and it would, for reasons of symmetry, be aligned with its spin.

It is important that the existence of the electric dipole moment (EDM) of the electron would violate the discrete symmetry of time reversal, T. The reason for this can be intuitively understood from the Fig. 5 (a), where the electronic spin is depicted with an arrow and the EDM with the spatially separated "+" and "-" (this picture should of course be taken only as a way to imagine the EDM, not as a reality). T will reverse the direction of spin, but it leaves the electric dipole moment intact. Thus, after the application of T we end up with a system different from the initial one, i.e. T is violated. According to the so-called CPT theorem, all the processes in nature should be invariant under a combined action of the three discrete symmetries, the charge conjugation C, the parity P and the time reversal T. Therefore, if the EDM violates the time reversal, it also violates the combined CP symmetry. The CP-symmetry violation is a necessary condition for explaining the observed matter-antimatter disbalance in the Universe.

Let us consider the properties of an ideal material for the EDM search. Firstly, EDMs of electrons add up to give an atomic EDM. The latter is maximal if all the electronic EDMs are mutually parallel, i.e. in a magnetic atom (because EDM is aligned with the electronic spin). Next, the atomic EDM is proportional to the cube of the atomic number. We therefore need a compound containing heavy magnetic atoms, e.g. one of the rare earth atoms. The basic idea of the EDM experiment is that, if an external electric field E is applied on the material, the EDMs of most of the magnetic atoms in the solid align with it taking their spins along and inducing in this way a net magnetization. Flipping the direction of the E-field will therefore flip the sample magnetization. This change in the magnetization can be experimentally measured. The induced magnetization is proportional to the density of magnetic atoms and to the strength of the applied electric field, and inversely proportional to the temperature. Therefore, the experiments are conducted at liquid Helium temperature (4.2 K) and, in order to enhance the externally applied electric field maximally, we use a ferroelectric. In order to avoid any linear magneto-electric coupling that could mimic the EDM signal, we use a paramagnet where the linear magneto-electric effect is forbidden by symmetry. We therefore need a ferroelectric material which is, at 4.2K, paramagnetic but

still has a sufficiently high density of the magnetic ions. These combined conditions are difficult to meet in a single material.

In designing the appropriate material, we note two perovskites, $EuTiO_3$ and $BaTiO_3$. The former is paraelectric and antiferromagnetic, with magnetic ordering temperature of 7 K - low, but not lower than 4.2 K, which is what we need for the experiment. The latter is a nonmagnetic ferroelectric. Some paraelectrics become ferroelectric when doped by larger cations that enlarge the lattice parameter of the compound. As Ba^{2+} is larger than Eu^{2+}, we make a hypothesis that by alloying $EuTiO_3$ at the Eu site by 50% of Ba atoms, this will result in a ferroelectric (Eu,Ba)TiO_3 compound. Moreover, the alloying should dilute the magnetic Eu sublattice, lowering in this way the magnetic ordering temperature. We investigated computationally the structural and magnetic properties of (Eu,Ba)TiO_3, concluding that the material should be an ideal candidate for the planned experiment: the calculations predict a material that is ferroelectric with a significant polarization value of P=23 $\mu C/cm^2$ and paramagnetic down to very low temperatures. The compound was subsequently synthesized and characterized by our collaborators; the measured properties confirmed our predictions-the compound is indeed a multiferroic, with a magnetic ordering temperature of 1.9 K. The first measurements searching for the EDM are currently conducted on this material.

Acknowledgements

We thank Prof. Dr. Nicola Spaldin for the longterm collaboration and guidance, Dr. Gustav Bihlmayer for many fruitful discussions and the Jülich Supercomputing Centre for the provided computing time. We gratefully acknowledge the financial support from the Young Investigators Group Program of the Helmholtz Association, Contract VH-NG-409.

References

1. M. Ležaić and N. A. Spaldin, *High-temperature multiferroicity and strong magnetocrystalline anisotropy in $3d-5d$ double perovskites*, Phys. Rev. B **83**, 024410, 2011.
2. `http://www.flapw.de`.
3. C.-H. Yang, *PhD thesis* (Stanford University, 2010).
4. Ph. Mavropoulos, M. Ležaić, and S. Blügel, *Ferromagnetism in nitrogen-doped MgO: Density-functional calculations* Phys. Rev. B **80**, 184403, 2009.
5. I. Slipukhina, Ph. Mavropoulos, S. Blügel, and M. Ležaić *Ferromagnetic spin coupling of $2p$-impurities in band insulators stabilized by intersite Coulomb interaction: Nitrogen-doped MgO*, Phys. Rev. Lett. **107**, 137203, 2011.
6. K. Z. Rushchanskii, S. Kamba, V. Goian, P. Vaněk, M. Savinov, J. Prokleška, D. Nuzhnyy, K. Knížek, F. Laufek, S. Eckel, S. K. Lamoreaux, A. O. Sushkov, M. Ležaić and N. A. Spaldin *A multiferroic material to search for the permanent electric dipole moment of the electron* Nature Mater. **9**, 649, 2010.

Quantum Spin Liquid in Correlated Fermions on a Graphene-Like Structure

Z. Y. Meng[1], T. C. Lang[2], S. Wessel[2], F. F. Assaad[3], and A. Muramatsu[1]

[1] Institut für Theoretische Physik III, Universität Stuttgart
Pfaffenwaldring 57, 70550 Stuttgart, Germany
E-mail: {meng, mu}@itp3.uni-stuttgart.de

[2] Institute for Theoretical Solid State Physics, RWTH Aachen University
Otto-Blumenthal-Strasse 26, 52056 Aachen
E-mail: {wessel, lang}@physik.rwth-aachen.de

[3] Institut für Theoretische Physik und Astrophysik, Universität Würzburg
Am Hubland, 97074 Würzburg, Germany
E-mail: assaad@physik.uni-wuerzburg.de

Among the newly discovered exotic quantum states of matter, like in the high temperature superconducting cuprates and pnictides, the fractional quantum Hall states, supersolidity, in topological insulators, or in frustrated magnets, quantum spin liquids occupy a prominent place. This fact is due, on the one hand, to their ubiquity in the theoretical description of high temperature superconductors and frustrated quantum magnets, and, on the other hand, to their elusiveness once models proposed to capture such phases are subjected to closer examination. Here, we discuss a quantum spin liquid phase, that was recently found by our collaboration, in electrons on a two dimensional honeycomb lattice, the structure of graphene, within an intermediately correlated regime at an average density of one electron per site. Our quantum Monte Carlo simulations showed that this unexpected spin liquid emerges between a state described by massless Dirac fermions and an antiferromagnetically ordered Mott insulator. Moreover, we found that this quantum-disordered state is a resonating valence-bond (RVB) liquid, akin to the one proposed for high temperature superconductors. This was the first unbiased determination of a RVB-liquid in an electronic system.

1 Introduction

In recent years states of quantum matter appeared[1] that go well beyond the generic behaviour, where, as temperature is lowered, a system passes from a disordered, high temperature phase to a symmetry broken ordered one. Of particular interest are spin liquids, where quantum fluctuations preclude ordering from the liquid state to a solid, i.e. to a periodic ordering of the magnetic moments, even at zero absolute temperature. Generally it is expected, that such strong quantum fluctuations arise due to competing interactions that frustrate the formation of an ordered state[2].

A system underlying strong quantum fluctuations is graphene, a carbon allotrope that was recently obtained experimentally by exfoliation of graphite[3]. By means of such a micromechanical cleavage, single layers of carbon atoms with a honeycomb structure (Fig. 1) are produced, that, remarkably, subsist as free-standing two-dimensional crystals. The honeycomb lattice is a bipartite one, i.e. it consists of two sublattices, where the nearest neighbours of the sites of one of them are sites belonging to the other sublattice. Hence, in such a lattice geometric frustration is absent, allowing e.g. antiferromagnetic order by placing magnetic moments pointing in one direction on one sublattice and pointing in the

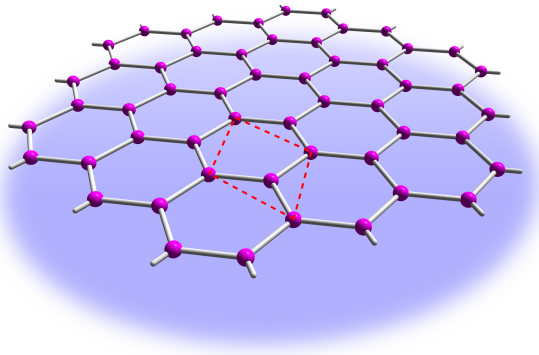


Figure 1. The honeycomb lattice with its unit cell, displayed by the dashed red line. The unit cell contains two atoms, each one belonging to a different sublattice. The coordination number is 3, the lowest in two dimensions.

opposite one on the other sublattice. However, three factors enhance the incidence of fluctuations. The first one is the low dimensionality, the second is the fact that the honeycomb lattice has the smallest coordination number in two dimensions. The third one is due to the particular band structure of graphene, where the electron (particle) states and the hole (antiparticle) states are symmetrically placed around the zero of energy, such that particle-hole (charge conjugation) symmetry is present. The zero of energy corresponds to the Fermi energy when the average density of electrons is unity per lattice site, i.e. for a half-filled band. Moreover, the low energy states around particular points in the two-dimensional Brillouin zone display a relativistic dispersion and can be readily described by Dirac's equation[3]. We therefore, will refer to those points as Dirac points in the following. The relativistic dispersion in the low energy sector at half-filling leads to a vanishing density of states at the Fermi energy. Therefore, a finite strength for interactions promoting a spontaneous symmetry breaking is necessary, enhancing thus the role of fluctuations.

In order to put fluctuations at work, many-body interactions should be introduced that may lead the system to a non-trivial phase-diagram, as detailed next.

2 Numerical Simulation of Correlated Electrons on the Honeycomb Lattice

The effects of electronic correlations in its most basic form can be studied with the Hubbard model, where only an on-site interaction, termed U, is present. The Hamiltonian is given by

$$H = -t \sum_{\substack{\langle i,j \rangle \\ \sigma}} \left(c^{\dagger}_{i\sigma} c_{j\sigma} + \text{h.c.} \right) + U \sum_{i} n_{i\uparrow} n_{i\downarrow} \, , \tag{1}$$

where $c^{\dagger}_{i\sigma}$ and $c_{i\sigma}$ respectively denote the creation and annihilation operators for fermions with spin $\sigma = \uparrow, \downarrow$ on lattice site i, and $n_{i\sigma} = c^{\dagger}_{i\sigma} c_{i\sigma}$ is the corresponding density operator.

$\langle i, j \rangle$ restricts the sum to nearest-neighbour sites, t denotes the nearest-neighbour hopping amplitude, and $U \geq 0$ the strength of the onsite repulsion.

Such a model is a paradigm for strongly correlated electrons, as in the case of high temperature superconductors[4]. For large values of the repulsive interaction U, and at half-filling, the ground state corresponds to a Mott insulator, i.e. an insulating state due to interactions in contrast to the metallic state of the noninteracting system. Furthermore, in this limit antiferromagnetic correlations dominate due to Pauli's exclusion principle and the necessity of gaining kinetic energy. Therefore, on a bipartite lattice, they will lead to an antiferromagnetically ordered ground state. However, as the interaction strength diminishes, a competition between the tendency to order and quantum fluctuations will set in, so that a detailed analysis of correlations is needed to characterize the possible phases.

An unbiased study of the Hubbard model in two dimensions is only possible by numerical means. Among the different methods, quantum Monte Carlo (QMC) simulations are the most appropriate ones since a careful extrapolation to the thermodynamic limit is necessary to determine whether a spontaneous symmetry breaking has taken place. We implemented a projective (temperature $T = 0$) determinantal QMC algorithm in the canonical ensemble that is free of the sign-problem at half-filling[5]. This algorithm allows the calculation of the expectation value of any physical observable in the ground state by performing an imaginary time evolution of a trial wave function that is required to be nonorthogonal to the ground state. The value Θ reached in the imaginary time evolution corresponds to a projection parameter[5]. For a spin-singlet trial wave function, we found $\Theta = 40/t$ to be sufficient to obtain converged ground state quantities within statistical uncertainty. In the presented simulations, we used a finite imaginary time step $\Delta\tau = 0.05/t$. We verified by extrapolating $\Delta\tau \to 0$ that this finite imaginary time step produces no artifacts. The phases described in the following were determined by a finite-size extrapolation to the thermodynamic limit with lattices of $N = 2L^2$ sites with periodic boundary conditions, and linear sizes L in terms of the unit cell containing two sites, with $L \leq 18$. L was taken as a multiple of 3 in order to be able to include the Dirac points in our Brillouin zones, such that the low energy physics is correctly represented.

3 Phase Diagram of the Hubbard Model on the Honycomb Lattice at Half-Filling

We first consider the single-particle excitation gap $\Delta_{sp}(\vec{k})$ that we extracted from the imaginary-time displaced Green's function (see Ref. 6 for details). $\Delta_{sp}(\vec{k})$ gives the minimal energy necessary to extract one fermion with momentum $\vec{k}$ from the system, and corresponds to the gap that can be observed in photoemission experiments. As shown in Fig. 2, $\Delta_{sp}(K) = 0$ for $U < U_c \approx 3.6t$. Here we denote by K the point in the Brillouin zone corresponding to the Dirac point. The vanishing gap corresponds to a metal, that in the present case is commonly called a semimetal (SM) due to the fact that the Fermi surface is reduced to a point, as a consequence of the Dirac dispersion. Beyond U_c, the system enters into an insulating phase due to interactions, and hence, as expected for large values of U, the system becomes a Mott insulator. The values of the gap are obtained via an extrapolation of the QMC data to the thermodynamic limit[6] with energies given in units of t.

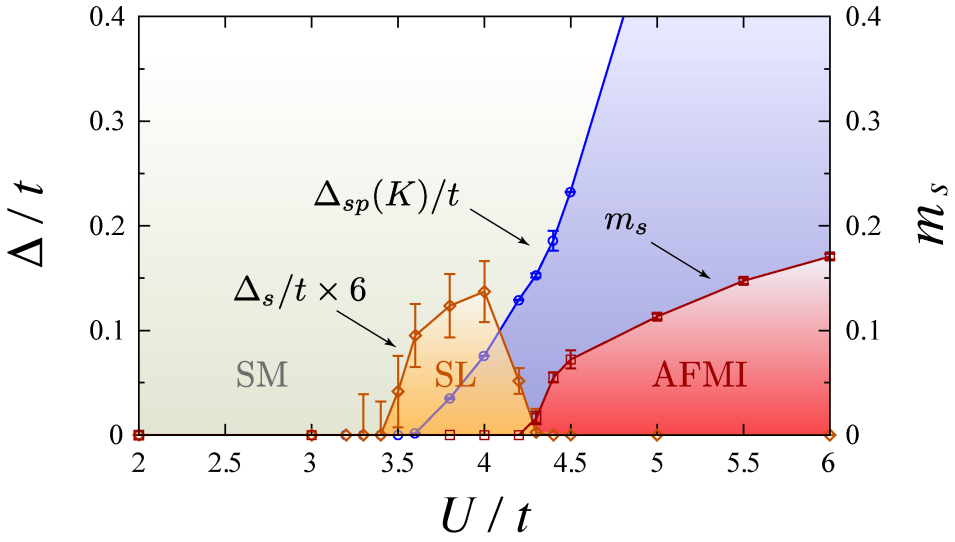


Figure 2. Phase diagram for the Hubbard model on the honeycomb lattice at half-filling The semimetal (SM) and the antiferromagnetic Mott insulator (AFMI) are separated by a gapped spin liquid (SL) phase in an intermediate coupling regime. $\Delta_{sp}(K)$ denotes the single-particle gap at one of the Dirac points (K), and Δ_s the spin gap. m_s denotes the staggered magnetization whose saturation value is $1/2$.

For values of U large enough, one expects long-range antiferromagnetic (AF) correlations due to the bipartite nature of the underlying lattice. We therefore measured the AF spin structure-factor S_{AF} that reveals long-range AF order if $\lim_{N\to\infty} S_{AF}/N > 0$. The results of a finite-size extrapolation are also presented in the phase diagram of Fig. 2. AF order appears beyond $U/t \approx 4.3$. Hence, contrary to the usual expectation for a bipartite lattice, AF long-range order sets in later than the insulating phase, leaving an extended window $3.6 \lesssim U/t \lesssim 4.3$, within which the system is neither a semimetal, nor an AF Mott Insulator.

Further details on the nature of this intermediate region are obtained by examining the spin excitation gap, extracted from the long-time behaviour of the imaginary-time displaced spin-spin correlation function[6]. We consider first the spin gap Δ_s in the staggered sector at $\vec{k} = 0$, which vanishes inside the AF phase due to the emergence of two Goldstone modes, as well as in the gapless metallic phase. Fig. 3 shows finite size estimates of Δ_s for different values of U/t, along with an extrapolation to the thermodynamic limit. A finite value of Δ_s persists within an intermediate parameter regime $3.5 \lesssim U/t \lesssim 4.3$, while it vanishes both within the metallic and the AF phase. We also calculated the uniform spin gap Δ_u by extrapolating the spin gap observed at the smallest finite $\vec{k}$-vector on each cluster to the thermodynamic limit. Δ_u is found to be even larger than Δ_s inside the intermediate region (e.g. $\Delta_u = 0.099(1)$ at $U/t = 4$), and vanishes in the metallic and the AF phase[6].

The observation of a finite spin gap rules out gapless phases such as triplet superconductivity as well as quantum spin Hall states. The remaining possibilities can be enumer-

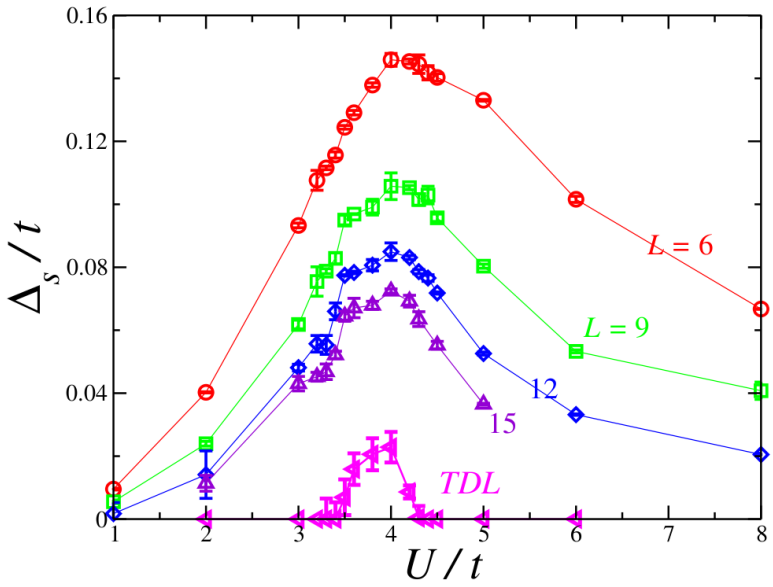


Figure 3. Spin gap in units of t as a function of U/t for various systems sizes. The lowest curve corresponds to the extrapolation to the thermodynamic limit (TDL).

ated by considering the coupling to order parameters that lead to the opening of a mass gap in Dirac fermions, and hence to account for the single-particle gap observed in the QMC data: (i) singlet superconductivity, (ii) a quantum Hall state (QHS), (iii) charge density wave (CDW) order, and (iv) a valence bond crystal (VBC).

In order to assess if superconductivity arises in the vicinity of the Mott transition, we used the method of flux quantisation which probes the superfluid density and is hence independent of the specific symmetry of the pair wave function[6]. Let Φ be a magnetic flux traversing the centre of a torus on which the electronic system lies and $E_0(\Phi/\Phi_0)$ the total ground state energy, Φ_0 being the flux quantum. A superconducting state of Cooper pairs is present, if in the thermodynamic limit the macroscopic energy difference $E_0(\Phi/\Phi_0) - E_0(\Phi/\Phi_0 = 1/2)$ is a function with period $1/2$. In contrast, a metallic (insulating) phase is characterised by an (exponential) vanishing of $E_0(\Phi/\Phi_0) - E_0(\Phi/\Phi_0 = 1/2)$ as a function of system size. The QMC data is consistent with the vanishing of this quantity in the thermodynamic limit. In addition, we measured singlet superconducting order parameters of (extended) s-, p-, and f-wave symmetry, which turn out to all vanish in the thermodynamic limit[6]. Hence, both flux quantisation as well as a direct measurement of pairing correlations in various symmetry sectors lead to no sign of superconductivity.

Both the CDW and QHS trigger a breaking of the sub-lattice symmetry and thereby open a mass gap at the mean-field level. A detailed analysis of the charge-charge correlation functions rules out a CDW. Furthermore, we have found no signature for the presence of (spin) currents in the ground state. This rules out the breaking of sublattice and time reversal symmetries as required for the QHS[6].

In order to detect the occurrence of a VBC, we probe for dimer-dimer correlations

between a dimer formed by nearest neighbour sites $\langle ij \rangle$ and a distant bond formed by sites $\langle kl \rangle$[6]. We have found no VBC, neither in the charge, nor in the spin sector. The left side of Fig. 4 shows the results of this measurement in the spin sector, i.e. the correlation between singlet dimers at $U/t = 4.0$. The striped bond is the one with respect to which correlations were determined. They are found to be short-ranged, and consistent with the dominance of a resonating valence bond (RVB) state within the hexagons of the honeycomb lattice. This can be seen by comparing the singlet-correlations with those of an isolated hexagon (right side of Fig. 4), the classical example of the resonance phenomenon in conjugated π-electrons[7]. Accordingly, we find no long-ranged order from the dimer-dimer structure factors in Fourier space.

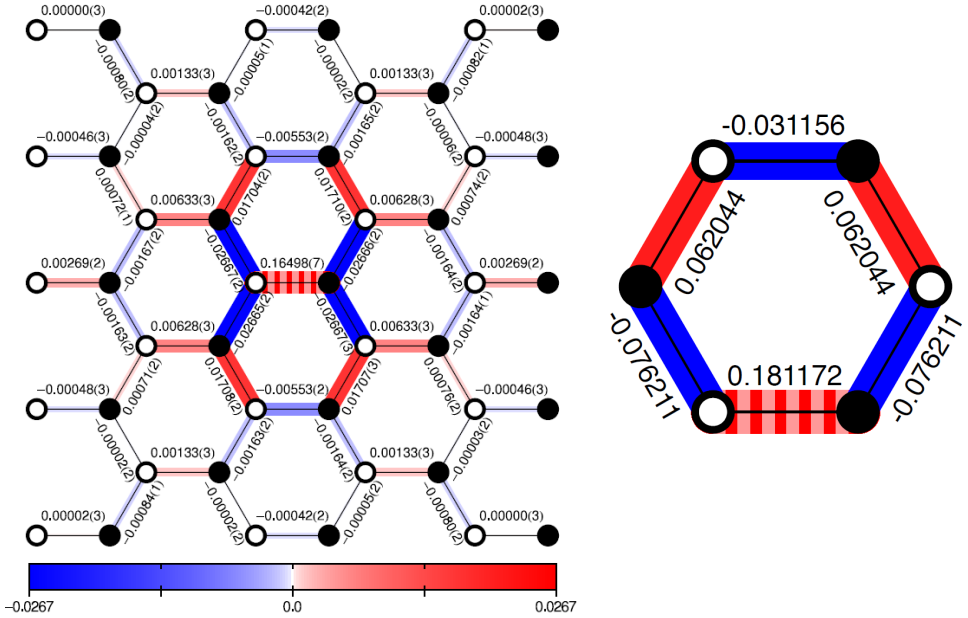


Figure 4. Real space plot of the spin dimer-dimer correlations. Left side: the dimer-dimer correlation function in the spin-channel for a $L = 6$ system at $U/t = 4$. Right side: the same correlation for the isolated Hubbard hexagon also at $U/t = 4$. The reference bonds are dressed with stripes. Numbers in parenthesis indicate the standard error of the last digit.

Our results reveal thus a genuinely exotic state of matter, where no spontaneous symmetry-breaking is observed, while a spin gap is present. It corresponds to a spin liquid RVB state in the intermediate coupling regime in the vicinity of the Mott transition.

4 Discussion

The presence of a spin liquid in the Hubbard model on the honeycomb lattice close to an antiferromagnetic Mott insulator was hitherto unexpected, due to the bipartite nature of the

honeycomb lattice, and hence, the absence of frustration. However, our results indicate that strong enough fluctuations, that develop close to the quantum critical point where AF order sets in, lead to such an exotic state of matter. It could be expected, that such fluctuations would promote some broken symmetry states like superconductivity. However, the vanishing density of states at the Fermi energy may be responsible for its absence, since in this case, a finite coupling strength is needed, at least in the BCS-frame.

Having an unexpected realisation of a short-range RVB state, it would be highly interesting to explore the consequences of doping, in a spirit rather close to the original scenario proposed by Anderson[4] and Kivelson *et al.*[8] for the cuprates. In particular, for the fully gapped short-range RVB state, the finite spin gap sets the energy scale of pairing in the superconducting state[8]. In this respect, the value obtained for the spin gap is rather promising. The largest value attained is $\Delta_s \sim 0.025t$ (Fig. 2), that for t of the order of 1 eV (in graphene is $t = 2.8$ eV[3]) corresponds to a temperature scale of around 290 K.

Although studies of doping are beyond the power of our quantum Monte Carlo approach due to the sign problem, they could open interesting perspectives e.g. in future experiments with ultra-cold atoms on a honeycomb optical lattice. Also solid-state realizations with honeycomb lattices based on group IV elements could be of interest. Recent estimates of the strength of the effective Coulomb interaction in graphene[9] set the value of the interaction $U/t \simeq 3.3$ and hence below the region where the spin liquid state sets in. However, it was shown recently both theoretically[10] and experimentally[11] that silicon can also form a monolayer structure with a honeycomb lattice, silicene, where the nearest neighbour distance is expected to be approximately 50% larger than in graphene[10], such that correlations effects are enhanced.

5 Outlook

From the theoretical point of view, one of the most interesting aspects of a fully gapped spin liquid is that such exotic states of matter are expected to possess unconventional order, of topological nature, where the ground state displays a degeneracy not due to spontaneous symmetry breaking, but due to states in different topological sectors[12]. Such topological degeneracy, originally demonstrated in quantum dimer models[13] that aim at modeling spin liquids, could until now be detected in few spin-systems[14]. Since the spin liquid discussed here takes place in an itinerant electron system, a hitherto unexplored class in this respect, it is of primordial interest to settle down the topological characteristics of such a state. The search for topological order constitutes a challenging enterprise since its unambiguous detection requires simulations of systems with linear sizes larger than the correlation length of magnetic excitations, that in our case extends to around 40 unit cells. However, such simulations are feasible with recently installed high performance computing resources, so that simulations on very large systems as required here, open the perspective of determining a topologically non-trivial state in a paradigmatic model, that is close to experimentally accessible systems[9–11].

Acknowledgements

F.F.A. acknowledges support by the DFG through AS120/4 and FG1162. A.M. acknowledges partial support by the DFG through SFB/TRR21. S. W. acknowledges support by

the DFG through SFB/TRR21 and WE3649. We thank NIC Jülich, HRL Stuttgart, the BW Grid and the LRZ München for the allocation of CPU time.

References

1. Nature Insight, *Exotic Matter*, Nature **464**, 175–216, 2010.
2. L. Balents, Nature **464**, 199–208, 2010.
3. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov and A. K. Geim, Rev. Mod. Phys. **81**, 109, 2009.
4. P. W. Anderson, Science **235**, 1196, 1987.
5. F. F. Assaad and H. G. Evertz, *Computational Many-Particle Physics*, Lecture Notes in Physics P**739** (Springer-Verlag, Berlin, 2008).
6. Z.-Y. Meng, T. C. Lang, S. Wessel, F. F. Assaad, and A. Muramatsu, Nature **464**, 847–851, 2010.
7. L. C. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: an Introduction to Modern Structural Chemistry* (Cornell University Press, 20th Edition, Ithaca, New York, USA, 1986).
8. S. A. Kivelson, D. S. Rokhsar, and J. P. Sethna, Phys. Rev. B **35**, 8865, 1987.
9. T. O. Wehling *et al.*, Phys. Rev. Lett. **106**, 236805, 2011.
10. S. Cahangirov, M. Topsakal, E. Aktürk, H. Şahin, and S. Ciraci, Phys. Rev. Lett. **102**, 236804, 2009.
11. B. Aufray *et al.*, Appl. Phys. Lett. **96**, 183102, 2010; P. De Padova *et al.*, Appl. Phys. Lett. **96**, 261905, 2010; B. Lalmi *et al.*, Appl. Phys. Lett. **97**, 223109, 2010.
12. X. G. Wen, Phys. Rev. B **44**, 2664, 1991.
13. D. S. Rokhsar and S. A. Kivelson, Phys. Rev. Lett. **61**, 2376–2379, 1988.
14. G. Misguich, C. Lhuillier, M. Mambrini, and P. Sindzingre, Eur. Phys. J. B **26**, 167–183, 2002.

KKRnano: Precise Density Functional Calculations for Thousands of Atoms

Rudolf Zeller[1], Alexander Thiess[2,3,1], Matthias Bolten[4], Peter H. Dederichs[3], and Stefan Blügel[1,3]

[1] Institute for Advanced Simulation, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: ru.zeller@fz-juelich.de

[2] German Research School for Simulation Sciences, 52425 Jülich, Germany
E-mail: a.thiess@fz-juelich.de

[3] Peter Grünberg Institut, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {ph.dederichs,s.bluegel}@fz-juelich.de

[4] Department of Mathematics and Science, University of Wuppertal, 42119 Wuppertal, Germany
E-mail: bolten@math.uni-wuppertal.de

KKRnano is a newly developed electronic structure code applicable to insulating, semiconducting and metallic materials. The code is massively parallelized for use of up to a hundred thousand processors. We describe its computational efficiency and the underlying physical and mathematical methodology and present first applications to a ferromagnetic semiconductor and a phase change material.

1 Introduction

Density functional theory is a formulation of quantum mechanics which provides an excellent balance of accuracy and computational efficiency for a wide range of problems in materials science. It is used in more than 10000 papers per year, however, for complex systems with internal interfaces, structural and chemical disorder, point and extended defects its applicability is severely limited by poor scaling if standard techniques are used to solve the density functional equations. Here the computational effort increases with the third power of the number of atoms. This means that even with a thousandfold increase of computing power – if exaflop supercomputers replace the present petaflop ones – only ten times more atoms can be treated which approximately increases the diameter of the systems by a factor two only.

To overcome this length-scale limitation, in recent years much effort has been spent worldwide to exploit a locality principle of quantum mechanics, the nearsightedness of electronic matter[1], with the aim to achieve a better, optimally linear scaling. Then, in the exaflop-era thousand time more atoms could be treated with system dimensions ten times larger than today. In the present contribution we will outline our own effort in this respect. Starting in 2008 we developed a computer program based on the full-potential Kohn-Korringa-Rostoker Green's function method[4] and applied it to study the electronic structure of systems with up to 8000 atoms. Since these systems extend over several nanometers, we consequently name our computer code KKRnano.

KKRnano was designed from the outset to utilize supercomputers consisting of a large number of multicore nodes as for the JUGENE and JUROPA machines available in Jülich.

The main goal was to have the ability to give the density functional answer to problems with many thousand atoms without any compromise on accuracy. This goal is reached in KKRnano with computing times which scale quadratically with system size. Optionally, linear scaling effort can also be achieved in KKRnano if a controllable and tolerable reduction of accuracy is traded for speed.

Below we will describe the method, the applied parallelization strategy and its efficiency, and results of applications to systems of technological interest, the semiconductor GaN doped with Gd, which shows room temperature ferromagnetism and colossal magnetic moments of about 4000 μ_{B} per Gd atom[2], and the phase change material $GeSb_2Te_4$ which shows indications for an Anderson localization[3].

2 Method

Contrary to most other density functional methods, which obtain the Kohn-Sham single-particle wavefunctions by expanding these wavefunctions into a set of basis functions and minimizing the total energy with respect to the expansion coefficients, the KKR Green's function method exploits the concept of reference systems and directly determines the Kohn-Sham Green's function by solving an integral equation. The integral extends over the space covered by the system and, for the solution, space is divided into non-overlapping cells around the atoms and in each cell the angular momentum representation of the free space Green's function (the one for vanishing potential) is used to separate the problem into the so called single and multiple scattering parts.

The single scattering part of the calculations requires the solution of integral equations where the integral extends only over the volume of a single cell. Since these integral equations for the cells are independent of each other, the computational effort for this part of the calculations naturally scales linearly with system size. The multiple scattering part, which describes the interrelation between the cells, requires to solve systems of linear algebraic equations of the form

$$G^{nn'}_{LL'}(\epsilon_i) = G^{r,nn'}_{LL'}(\epsilon_i) + \sum_{n''}\sum_{L''} G^{r,nn''}_{LL''}(\epsilon_i) \sum_{L'''} \Delta t^{n''}_{L''L'''}(\epsilon_i) G^{n''n'}_{L'''L'}(\epsilon_i)\,. \qquad (1)$$

Here $\Delta t^{n}_{L''L'''}$ is a difference of single-cell t matrices for the system and the reference system which are obtained in the single scattering part of the calculations. The Green's function matrix elements $G^{r,nn'}_{LL'}$ of the reference system are known and the Green's function matrix elements $G^{nn'}_{LL'}$ are to be calculated. The upper indices n label the cells around the atoms and the lower indices L the angular momentum components. The parameters ϵ_i denote mesh points used for the integration to obtain the electronic density, the fundamental quantity of density functional theory. The number of L's determines the angular resolution in the single scattering integral equations. Experience has shown that usually 16 L values are enough, which leads to matrices of dimension 16 times N in (1), where N denotes the number of atoms (cells). Here it is important to note that the radial resolution used in the single scattering integral equations does neither affect the structure of (1) nor the matrix dimension. It is also important to note that the solution of (1) represents the only part in the calculations with scaling proportional to N^3 and that standard methods of numerical algebra can be used to deal with (1).

The main difference in KKRnano compared to conventional KKR programs consists in the solution of (1). Whereas conventionally this equation with non-Hermitian matrices is solved by Gauss elimination, in KKRnano it is solved iteratively by the quasi-minimal residual (QMR) method of Freund and Nachtigal[6]. Since for real values of ϵ_i the Green's function matrix $G_{LL'}^{nn'}(\epsilon_i)$ can have singularities which prevent convergence of the QMR iterations, KKRnano is based on finite temperature density functional theory with complex contour integrations to obtain the density[7]. During the QMR iterations the calculations essentially consists of matrix multiplication with matrices of dimension N. If these matrices are dense, as they are for the standard free space reference system of the KKR method, nothing is gained by iterative solution. Fortunately, an advantageous reference system was proposed[8] with a reference Green's function matrix for which the number of nonzero elements is proportional to N if exponentially small elements are neglected. Then the overall computing effort increases only with N^2 instead of N^3. If the concept of nearsightedness of electronic matter[1] is exploited, the effort can be further reduced. The concept means that in systems without long range electric fields the electronic density at a point in space does not depend relevantly on potential changes far away. In KKRnano this fact is exploited by choosing a truncation region around each atom and neglecting matrix elements arising from the outside of the truncation region for the Green's function of the considered material. Then only sparse matrices must be multiplied and the scaling becomes linear.

3 Parallelization

The natural parallelization strategy in KKRnano is domain decomposition with a simple matching between cells and processors. Both, the single scattering part of the calculations and the iterative solution of (1) can be done separately for each cell. Communication between the processors is required to distribute the t matrices and the reference Green's function matrix before (1) is solved. Communication also occurs after the density is calculated because density information must be distributed to obtain the electrostatic potential needed in the next density functional self-consistency step.

We have also implemented three other levels of MPI parallelization within KKRnano, parallelization over the energy points ϵ_i, over the two spin directions in magnetic systems and over the 16 L components. The energy parallelization is dynamically load balanced during the self-consistency steps by putting the energy points into two or three groups to deal with the considerably different number of iterations at different energy points. The four MPI levels of parallelization can be used with up to 96 processors per atom. This was tested during the Blue Gene/P Extreme Scaling Workshop 2010, where we were able to run a system of 3072 atoms using from 3072 to 294912 processors. Increasing the number of processors from 147456 to 294912 gave a wall clock time speedup of 1.6 in the L parallelized part of KKRnano, whereas the time became larger for the complete program which was fastest with 147456 processors.

Since then we have modified and improved KKRnano in several aspects. We now work with better initial guesses for the QMR iterations, we use preconditioning (see next section), and instead of MPI parallelization over L we use OpenMP threads for low level parallelization in various parts of the program. This is more advantageous for routine applications because some matrix-vector operations can be replaced by matrix-matrix ones. Moreover, with OpenMP threads larger systems with up to 16000 atoms are possible be-

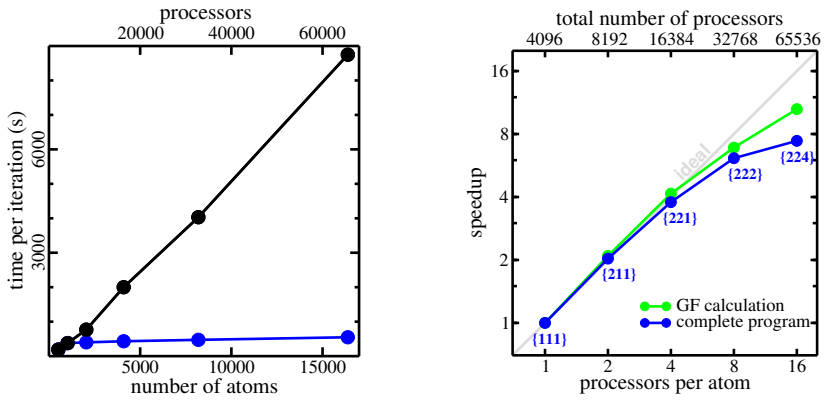


Figure 1. Wall clock time behaviour and parallel speedup for dilute NiPd systems. The left picture shows the wall clock time for one self-consistency step for one density functional self-consistency step as function of system size with no truncation of the Green's function (upper black curve) and for a truncation region of 959 atoms (lower blue curve). The right picture shows the speedup for a fixed system size of 4096 atoms for a full self-consistency step (lower blue curve) and for the Green's function calculation alone (upper green curve). The numbers in brackets specify the number of processors used per spin and energy parallelization and the number of OpenMP threads.

cause we can utilize the shared memory available on JUGENE where four CPU's on a node share 2 GBytes of memory.

A typical example for the parallel efficiency of KKRnano in routine applications is shown in Fig. 1. The left picture (weak scaling) shows the increase of wall clock time with and without truncation of the Green's function matrix for increasing system size using always four processors per atom. Without truncation the time increases almost linearly which means that the quadratically scaling total computational work is well distributed over the increasing number of processors. Moreover, with truncation (for the dependence of density functional total energies on the size of the truncation region, see next section) the calculations are considerably faster. With truncation the wall clock times increase only slightly with system size indicating that the desired linear scaling computational effort is almost achieved. The right picture (strong scaling for fixed system size) shows the speedup obtained by MPI parallelization over spin and energy points and by OpenMP threads. One sees that spin and energy parallelization perform very nicely and that additional speedup is gained by the OpenMP implementation which is an essential part of KKRnano for memory bound systems.

4 Precision

The precision achievable in KKRnano compared to standard KKR programs depends on the number of QMR iterations which one likes to spend and on the size of the truncation region. In our experience, the QMR iterations always converge to any desired precision but convergence may be slow so that many iterations are needed. Since it is well known that preconditioning often accelerates convergence, we tried to use conventional preconditioners like incomplete LU decomposition. They work in KKRnano, but unfortunately, they

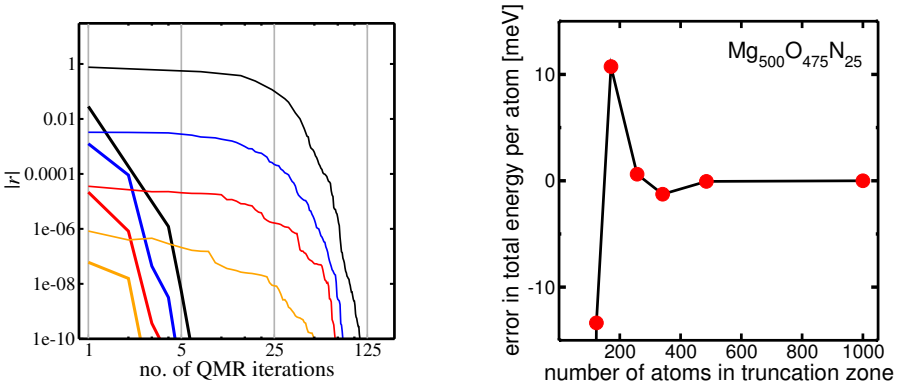


Figure 2. Precision achievable in KKRnano. The left picture shows the decrease of the residual norm as function of the number of QMR iterations for Si_{432} system with preconditioning switched on (thick curves) and off (thin curves). Four different sets of starting values are used for the QMR iterations rated by their initial residual norms $|r| = 1$, 10^{-2}, 10^{-4}, and 10^{-6} (black, blue, red and orange curves.) The right picture shows the density functional total energy as function of the number of atoms in the truncation region for a MgO system of 1000 atoms with 5% oxygen atoms replaced by nitrogen atoms.

cannot be computed with the high degree of parallelization used in KKRnano. To overcome this difficulty, we developed a new multi-level block-circulant preconditioner[9] that exploits the fact that the reference Green's function matrix only depends on the geometrical structure of the system and that only the t matrices carry the chemical information. We split the large unit cell into small subcells, take the geometrical and chemical average over the subcells and build a matrix from these averages. Its inverse is the preconditioner which we obtain by Fourier transformation, Fig. 2 (right picture) shows the typical reduction of the number of QMR iterations by roughly one order of magnitude which we obtain. Despite the overhead due to the construction of the preconditioner, the full calculation becomes three to ten times faster.

The effect of truncation on the total energy is shown in Fig. 2 (right picture) for MgO with 5% of the oxygen atoms replaced by nitrogen atoms. This semiconducting material is discussed as a room temperature ferromagnet with resistive switching properties useful for the next generation of nonvolatile memory. The total energy error rapidly decreases with increasing size of the truncation region and is small enough if the truncation region contains more than a few hundred atoms. Similar accuracy for metallic systems was found in a model study for periodic Cu and Pd supercells containing 131072 atoms[5].

5 Applications

5.1 Gd in GaN: the Role of Interstitial Defects

Starting with the discovery of ferromagnetic coupled colossal magnetic moments above room temperature in Gd doped GaN by Dhar et al.[2], the search for the origin of this effect is challenging the scientific community. While it seems to be clear that the striking magnetic properties are induced by extrinsic defects, answering the question which type of defect

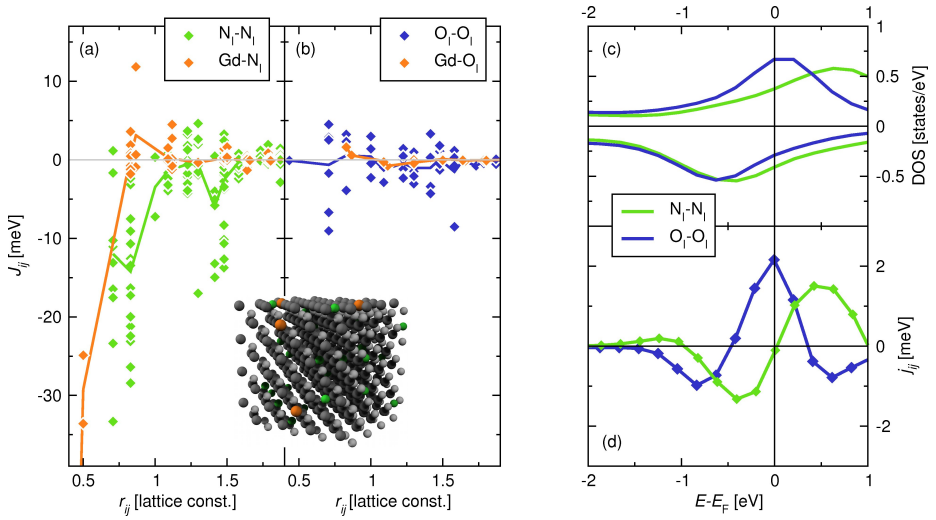


Figure 3. Magnetic interaction of nitrogen and oxygen interstitial defects, N_I and O_I in Gd doped GaN. (a) Magnetic exchange interactions J_{ij} obtained by the Liechtenstein formula between Gd and N_I (orange) and N_I and N_I sites (green) for $Ga_{248}N_{256}Gd_8N_{I,32}$. Average exchange interactions are shown as lines in the corresponding colour. (b) Same plot as in (a) but for the exchange interactions between Gd and O_I (orange) and O_I and O_I (blue) in $Ga_{253}N_{256}Gd_3O_{I,20}$. In (c) and (d) the averages of the LDOS for O_I and N_I and the energy resolved exchange interaction j_{ij} between O_I and O_I and between N_I and N_I are compared.

plays the crucial role is a highly demanding task. Firstly, a realistic description of the dilute limit of Gd doping can only be modelled accurately with supercells of several hundreds of atoms. Secondly, the co-doping of Gd and additional extrinsic effects increases the phase space of local defect configurations vastly, a large set of defects must be considered to obtain realistic and reliable results. On top of this, insight on the individual magnetic exchange interactions is crucial for an microscopic explanation of the magnetic ordering. With KKRnano we have a method at hand with which we can address these points. A particular advantage is the ability to calculate magnetic exchange interactions parameters, J_{ij}, making explicit use of the Green's function available by means of the Liechtenstein formula[10].

Since nitrogen as well as oxygen interstitial defects, N_I and O_I, respectively, have been proposed as possible candidates for ferromagnetic ordering above room temperature[11], we constructed zincblende GaN supercells with randomly distributed Gd on Ga sites and N_I or O_I on interstitial sites (see caption of Fig. 3 for more details) and find a sizable spin-polarization for all probed configurations of interstitials. As Fig. 3 (a) shows, the N_I's tend to couple antiferromagnetically amongst each others and to Gd. For O_I the picture changes (see Fig. 3 (b)), the calculated J_{ij}'s fluctuate around zero, leaving room for partly ferromagnetic and partly antiferromagnetic interactions. This qualitatively different behaviour of N_I's and O_I's can be understood considering the local density of states (LDOS) and the energy resolved magnetic exchange interactions shown in Fig. 3 (c) and (d) indicating that the balance of double and superexchange is significantly different for N_I's and O_I's. Since both defects do not show strong ferromagnetic couplings, we can conclude the important result that neither N_I's nor O_I's can support or induce a strong ferromagnetic coupling in

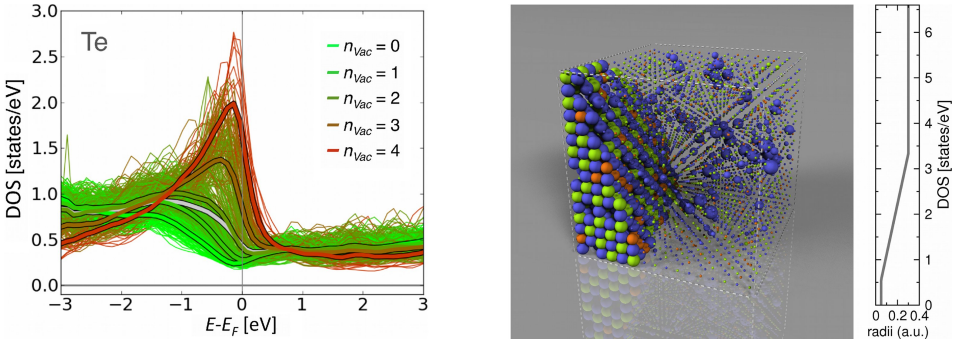


Figure 4. (Left) Local density of states (LDOS) of Te atoms in a $Ge_{125}Vac_{125}Sb_{250}Te_{500}$ supercell. Five different types of Te atoms are grouped according to their number of nearest neighbour sites occupied by vacancies. For each group the average LDOS is shown as thick line. The global average over the LDOS of all Te sites is represented by the grey line. (Right) Spatial distribution of local density of states for a $Ge_{500}Vac_{500}Sb_{1000}Te_{2000}$ supercell. In the left lower part the chemical information, in the upper right part the value of the LDOS is displayed. Here, large (small) radii of the spheres correspond to high (low) DOS values as specified in the right panel. For both parts of the plot Ge, Vac, Sb, and Te are shown in orange, transparent, green, and blue.

Gd doped GaN. Hence, the search for the microscopic origin of ferromagnetically coupled colossal magnetic moments will continue ... .

5.2 Disorder and Localization in Phase Change Materials

Our investigation of the phase change material $GeSb_2Te_4$ was motivated by recent experimental findings of Siegrist et al.[3], who reported pronounced localization in $GeSb_2Te_4$ and strong dependence of this localization on annealing. With KKRnano we are able to conduct extensive studies to identify the role of disorder on the electronic structure as well as the finite-size scaling of the states. In such calculations it is of high importance to consider sufficiently large supercells to minimize the influence of statistical fluctuations and include local atomic configurations of lower probability. In our calculations we use supercells of up to 4000 atoms and the most common structural model for the rocksalt-structure of $GeSb_2Te_4$, where one sublattice is completely occupied by Te atoms and the other one randomly by vacancies, Ge and Sb atoms (see Fig. 4 (right)). Focusing here on the question how the electronic states are affected by disorder, we reveal that the fundamental local properties of the sites can be well understood by taking into account only the chemical configuration on few neighbouring shells. For localization of states the distribution and ordering of vacancies turns out to be of particular importance. In the vicinity of regions with high vacancy density the Ge-Te and Sb-Te bonding states lie in a pseudo band-gap and are therefore highly localized. As a consequence especially Te sites in vicinity of such vacancy complexes possess very high density of states at the Fermi energy as shown in Fig. 4 (left) and (right). From the perspective of electronic transport, vacancy complexes can therefore be interpreted as strong scattering centres leading to enhanced resistivity. Thereby our calculations provide insight on the origin of localization of electronic states in $GeSb_2Te_4$. In addition, this understanding and the ability to extract the energetics of local configurations from KKRnano provide the key for further modeling and understanding of the temperature dependence of the localization.

6 Concluding Remarks

In this report we summarized our development of KKRnano, a massively parallel, full-potential all-electron program for arbitrarily shaped supercells. We illustrated that many thousands of atoms can be treated efficiently and that significant advances in understanding materials on the atomistic level are possible by exploiting this strength of KKRnano.

Acknowledgements

We thank Professor Walter Lambrecht for very fruitful discussions on the GaN(Gd) system. We also gratefully acknowledge financial support of the DAAD and computational resources at as well as support of the Jülich Supercomputing Centre.

References

1. E. Prodan and W. Kohn, *Nearsightedness of electronic matter*, Proc. Natl. Acad. Sci. USA **102**, 11635-11638, 2005.
2. S. Dhar, O. Brandt, M. Ramsteiner, V. F. Sapega, and K. H. Ploog, *Colossal Magnetic Moment of Gd in GaN*, Phys. Rev. Lett. **94**, 037205, 2005.
3. T. Siegrist, P. Jost, H. Volker, M. Woda, P. Merkelbach, C. Schlockermann, and M. Wuttig, *Disorder-induced localization in crystalline phase-change materials*, Nat. Mater. **10**, 202-208, 2011.
4. N. Papanikolaou, R. Zeller, and P. H. Dederichs, *Conceptual improvements of the KKR method*, J. Phys.: Condens. Matter **14**, 2799-2823, 2002.
5. R. Zeller, *Towards a linear-scaling algorithm for electronic structure calculations with the tight-binding Korringa-Kohn-Rostoker Green's function method*, J. Phys.: Condens. Matter **20**, 294215, 2008.
6. R. W. Freund and N. M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math. **60**, 315-330, 1991.
7. K. Wildberger, P. Lang, R. Zeller, and P. H. Dederichs, *Fermi-Dirac distribution in ab initio Green's-function calculations*, Phys. Rev. B **52**, 11502-11508, 1995.
8. R. Zeller, P. H. Dederichs, B. Újfalussy, L. Szunyogh, and P. Weinberger, *Theory and convergence properties of the screened Korringa-Kohn-Rostoker method*, Phys. Rev. B **52**, 8807-8812, 1995.
9. M. Bolten, I. Yavneh, A. Thiess, and R. Zeller, *Preconditioning systems arising from the KKR Green's function method using block-circulant matrices*, in press, Linear Algebra Appl., doi:10.1016/j.laa.2011.05.019.
10. A. I. Liechtenstein, M. I. Katsnelson, V. P. Antropov, and V. A. Gubanov, *Local spin density functional approach to the theory of exchange interactions in ferromagnetic metals and alloys*, J. Magn. Magn. Mater. **67**, 65-74, 1987.
11. C. Mitra and W. R. L. Lambrecht, *Magnetic exchange interactions in the gadolinium pnictides from first principles*, Phys. Rev. B **78**, 134421, 2008.

Computational Soft Matter Science

Kurt Kremer

Max Planck Institute for Polymer Research, 55021 Mainz, Germany
E-mail: kremer@mpip-mainz.mpg.de

Soft Matter is the generic term for all complex (organic) materials, which are characterized by a relatively low internal cohesive energy density and nanoscopic to microscopic length scales. They range from conventional plastic materials to almost all biological materials (for instance tissue, proteins etc.) and colloidal suspensions. Traditionally homogeneous systems have played the major role in research as well as in most applications. Over the years, however, nanoscopically structured materials more and more came into focus of science and molecular engineering. A material is nanoscopically structured if the ratio of the volume of the system to the overall internal or external surfaces and interfaces is clearly below a micrometer. This certainly holds for almost all biological systems but also for components of organic electronics or modern functional materials. Thus it is not surprising that properties of macromolecules close to surfaces are an important research topic, not only experimentally but also theoretically and especially in computational soft matter science.

Both contributions in this year's issue deal with problems of that kind. The first one by the group of W. Janke in Leipzig is dealing with properties of isolated macromolecules close to an adsorbing surface. There, a detailed comparison of a macromolecule which is already grafted to the surface with one chain end to the situation of a chain in free solution is presented. While problems like this have been matter of research for more than thirty years, only recently the whole complexity of the problem has been realized and investigated. Even on the level of the highly idealized models of the present study, the delicate interplay between strength of the adsorption energy, intrinsic chain stiffness, and interaction between the different monomers of the chain leads to a rather complex and rich "phase diagram". This is not to be confused with a true phase diagram since the limiting case of an infinite chain length might be theoretically of interest. However, for all experiments and all practical questions this limit typically is far away from the actual experimental set up. Depending on the specific situation there are significant differences between the end grafted polymer and the free floating polymer. Not only that the translational entropy of the overall chain plays an important role, also whether the condensation into an amorphous or ordered globule occurs already in the "bulk" or in the adsorbed phase plays an important role. Especially for the case, where the coil-globule transition occurs above the adsorption transition, end grafted and the free polymers display a qualitatively different properties near the adsorption threshold. Janke and co-workers study these problems with extensive parallel tempering simulations.

In contrast to that, in a second contribution W. Paul, K. Binder and co-workers study dynamical properties of 1,4-Polybutadiene melts close to graphite surfaces. By adapting and optimizing the codes they made efficient use of the JUGENE machine. For typical mid-size systems which then have to run for a very long times, this is a challenging task.

After porting their code they were able to perform extensive studies of the structure and dynamics of 1,4-Polybutadiene on the basis of the, so called united atom model close to (0,0,1) graphite surfaces. In addition to phenomena based on generic models, as in the above described study by Janke et al., the effect of specific chemistries is of equal relevance, especially when it comes to a comparison to experiment. Generally speaking, there are two ways to adjust/manipulate soft matter, and both ways allow to shift characteristic scales by many orders. One is to adjust generic properties, for instance changing the molecular weight of the investigated material, while the other one is to change local chemical characteristics and interactions. Because of that it is very important to include studies of specific chemistries in this kind of long time simulations as well. The results revealed not only a layering of monomers close to the surface but also a fairly distinct layer of centre of mass positions of the chains close to the surface. In a next step, a detailed analysis of time-dependent correlation functions and diffusion constants (perpendicular and parallel to the surface) as a function of the distance from the surface have been performed. These simulations provide a rather comprehensive study, which exceeds typical experimental approaches, because experiments typically cannot cover that many different dynamical processes within one study. In addition experiments so far report rather conflicting results. Here the simulations play an important role in shedding some light on these problems. The present model first was validated very carefully by bulk simulations and then used for a polymer melt close to a graphite surface. It was found that desorption of the chains from the wall as well as exchange between different layers is strongly slowed down compared to the typical bulk mobility at the same temperatures. This effect is dominated by the segment wall interaction and should not be confused with the effect of the confining geometry on the glass transition, even though the interplay of the two certainly will affect the approach to the glass transition temperature of the polymers close to the surface. This also points to an important future direction as the position dependent distance from the apparent glass transition temperature might reveal very interesting structural dynamical properties of the polymer film.

Again, these two studies show how important computer simulations can be to the context of soft matter science in general. On the one hand, analytic theory is only able to describe rather limited ideal situations, which are, however, very important in the context of our basic understanding and also as anchor points for simulation studies. On the other hand simulation studies cannot only provide an important bridge to actual experiments, but due to the power of modern software as well as hardware they can provide investigations and new ways of studying these systems in their own right. The present two studies are examples of this modern field of science.

Grafted versus Non-Grafted Polymer Adsorption

Monika Möddel[1], Michael Bachmann[2,3], and Wolfhard Janke[1]

[1] Institut für Theoretische Physik and Centre for Theoretical Sciences (NTZ), Universität Leipzig
Postfach 100 920, D-04009 Leipzig, Germany
E-mail: {Monika.Moeddel, Wolfhard.Janke}@itp.uni-leipzig.de

[2] Center for Simulational Physics, The University of Georgia, Athens, GA 30602, U.S.A.
E-mail: bachmann@smsyslab.org

[3] Institute for Complex Systems (ICS-2) and Institute for Advanced Simulation (IAS-2)
Forschungszentrum Jülich, D-52425 Jülich, Germany

The conformational behaviour of a semi-flexible self-interacting finite polymer near an attractive substrate was investigated with the parallel tempering Monte Carlo method. This method allows for a precise estimate of canonical equilibrium data over a wide range of surface attraction strengths and temperatures, which enables us to identify the different phases of the finite system via thermal fluctuations of canonical expectation values of energetic and structural quantities. Complementary microcanonical information is extracted from the density of states. The resulting phase diagrams for grafted and non-grafted polymers are discussed and compared.

1 Introduction

In a diluted polymer solution extended polymer conformations are favoured at high temperatures due to their higher conformational entropy compared to globular conformations. Reducing the temperature, globular conformations gain thermodynamic weight, and the polymer collapses in a cooperative rearrangement of the monomers. Those globular conformations are relatively compact but have only little internal structure. Hence, they are still entropy-dominated, and a further transition towards low-degenerate crystalline energetic states is expected and indeed observed: the freezing transition[1,2]. Although it has been shown that these two transitions fall together for sufficiently short-range interactions[3], they are clearly distinct in general.

The presence of an attractive substrate modifies this behaviour of the polymer and gives rise to theoretical and computational challenges. It not only adds the adsorption transition, but also induces several phases by the competition between monomer-monomer and monomer-surface attraction. Although the properties of some phases and subphases depend on the exact number of monomers, they are not less interesting, because with the advent of new sophisticated experimental techniques such small scales become accessible. Among such techniques at the nanometer scale is, e.g., atomic force microscopy (AFM), where it is possible to measure the contour length and the end-to-end distance of individual polymers[4] or to quantitatively investigate the peptide adhesion on semiconductor surfaces[5]. Another experimental tool with an extraordinary resolution in positioning and accuracy in force measurements are optical tweezers[6,7].

Numerous detailed numerical studies have been performed to elucidate the conformational behaviour of homopolymers and heteropolymers near substrates. Compared to experiments, computer simulations have the advantage that a wide range of control parameters can be scanned at will. Theoretical studies have, e.g., been performed analytically

using scaling theory[8,9], mean-field density functional theory[10], and series expansions[11] and numerically by employing off-lattice models such as a bead-spring model of a single polymer chain grafted to a weakly attractive surface[12,13], Monte Carlo studies of lattice homopolymers[8,12,14–17], or exact enumeration[18] to name only a few.

Our focus here is on the classification of thermodynamic phases and phase transitions for a range of surface attraction strengths and temperatures and we compare the results for end-grafted polymers with those of non-grafted polymers that can move freely within a simulation box. After we recently constructed such a phase diagram for the non-grafted case[19], and also analyzed the influence of the simulation box size[20], the question arises whether there is a qualitative difference between grafted and non-grafted adsorption. This gap of a systematic comparison of both cases using an otherwise identical coarse-grained model shall be filled here[21]. The reduction of degrees of freedom speeds up the simulation considerably, but still the problem is computationally very demanding, such that the well established parallel tempering Monte Carlo (MC) method with MPI is applied. This allows for an easy and efficient parallelization on the Jülich supercomputer JUROPA.

2 Model

The model used here is a simple bead-stick model of a linear polymer with three terms that contribute to the energy[19,20],

$$E = 4\sum_{i=1}^{N-2}\sum_{j=i+2}^{N}\left(r_{ij}^{-12} - r_{ij}^{-6}\right) + \frac{1}{4}\sum_{i=1}^{N-2}\left(1-\cos\vartheta_i\right) + \epsilon_s\sum_{i=1}^{N}\left(\frac{2}{15}z_i^{-9} - z_i^{-3}\right), \quad (1)$$

where the first two terms are the energy in bulk that consists of the standard 12-6 Lennard-Jones (LJ) potential and a weak bending energy. The distance between the monomers i and j is r_{ij} and $0 \leq \vartheta_i \leq \pi$ denotes the bending angle between the ith, $(i+1)$th, and $(i+2)$th monomer. The third term is the attractive surface potential, obtained by integration over the continuous half-space $z < 0$ (cf. Fig. 1), where every space element interacts with a single monomer by the 12-6 LJ expression[22]. Hence, the parameter ϵ_s weighs the monomer-surface and monomer-monomer interaction. We use generic units, in which $k_B = 1$.

The relative strength of the two interactions is continuously varied by considering ϵ_s as a control parameter. We simulate the polymer once grafted with one end to the substrate in the potential minimum and once freely moving in the space between the substrate and a hard wall a distance $L_z = 60$ away. To describe the system, we measure several canonical expectation values $\langle O \rangle$,

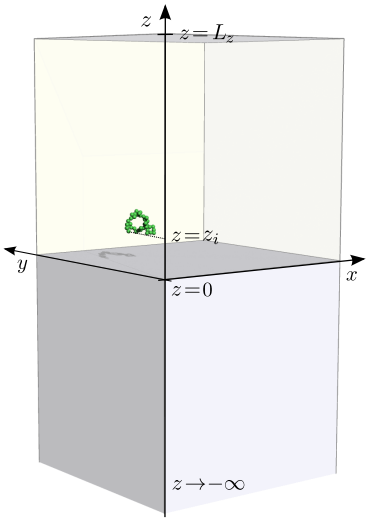


Figure 1. Graphical representation of the system of a single polymer close to an attractive substrate at $z = 0$. The hard wall at $z = L_z$ prevents a non-grafted polymer from escaping.

determine their temperature derivative and also keep a look at their microcanonical behaviour with energy $O(E)$ and the density of states $g(E)$. The most basic observable O is the energy E itself, whose temperature derivative is the specific heat $c_V(T)$. The squared radius of gyration $R^2_{\mathrm{gyr}} \equiv \sum_{i=1}^{N} (\vec{r}_i - \vec{r}_{\mathrm{cm}})^2$, with $\vec{r}_{\mathrm{cm}} = (x_{\mathrm{cm}}, y_{\mathrm{cm}}, z_{\mathrm{cm}}) = \sum_{i=1}^{N} \vec{r}_i/N$ being the centre-of-mass of the polymer, as a measure for the extension of the polymer as well as its tensor components parallel and perpendicular to the substrate, $R^2_{\parallel} = \sum_{i=1}^{N}[(x_i - x_{\mathrm{cm}})^2 + (y_i - y_{\mathrm{cm}})^2]$ and $R^2_{\perp} = \sum_{i=1}^{N} (z_i - z_{\mathrm{cm}})^2$, offer rich information since the substrate introduces a structural anisotropy into the system. One indicator for adsorption is the distance of the centre-of-mass of the polymer to the surface. Additionally, we analyze the mean number of monomers docked to the surface n_s. For the continuous substrate potential we define a monomer i to be docked if $z_i < z_c \equiv 1.5$.

3 Methods

To simulate this model system, the parallel tempering or replica exchange MC method[23,24] is applied. The basic idea is to run several Metropolis simulations in parallel at different inverse temperatures $\beta_1 < \ldots < \beta_n = 1/T_n$. Every once in a while two copies of the system exchange their current conformations with the Metropolis acceptance probability

$$A(\mu\nu \to \nu\mu) = \min\left(1, \frac{p_{\nu\mu}}{p_{\mu\nu}}\right) = \min\left(1, e^{\Delta\beta\Delta E}\right), \tag{2}$$

where $p_{\mu\nu} = e^{-\beta E_\mu} Z_\beta^{-1} e^{-\beta' E_\nu} Z_{\beta'}^{-1}$ is the joint probability of the systems at β and β' to be in state μ and ν, respectively. Z_β and $Z_{\beta'}$ are the associated canonical partition functions. This way conformations that were stuck in a valley of the energy landscape at low temperatures can escape at higher temperatures and eventually explore other regions of phase space. Such an approach combines two important advantages. It strongly decreases the autocorrelation time at low temperatures and is highly parallelizable.

In the case of a large temperature separation of the two replicas that attempt an exchange of their conformations, the acceptance probability gets small. Hence, for the performance of this method both, the number and the temperatures of the replicas are essential. In order that two replicas exchange their conformations, their energy histograms have to have a sufficient overlap as can be seen from Eq. 2. If one is interested in low-temperature properties, the range from those low temperatures to high enough temperatures to cross free energy barriers has to be covered. This sets a minimum of necessary replicas. Choosing too many, on the other hand, slows down the simulation effectively, since it takes longer on average for a replica to go from low to high temperatures and back again. There exist several attempts to optimize the choice of the β_i,[25] but usually one can get a reasonable performance when observing the histograms and ensuring the acceptance probability to be around 50%, which approximately requires an equidistribution in β. Here, 64-72 different replicas are used with 50 000 000 sweeps each, from which every 10th value was stored in a time series – the autocorrelation time in units of sweeps is of the order of thousands. For the canonical analysis, this statistics is very generous, but to obtain data with low statistical errors for the microcanonical entropy $S(E) = \ln g(E)$ and its derivatives, higher precision data are necessary. This was done for $\epsilon_s = 0, 0.1, \ldots, 5$.

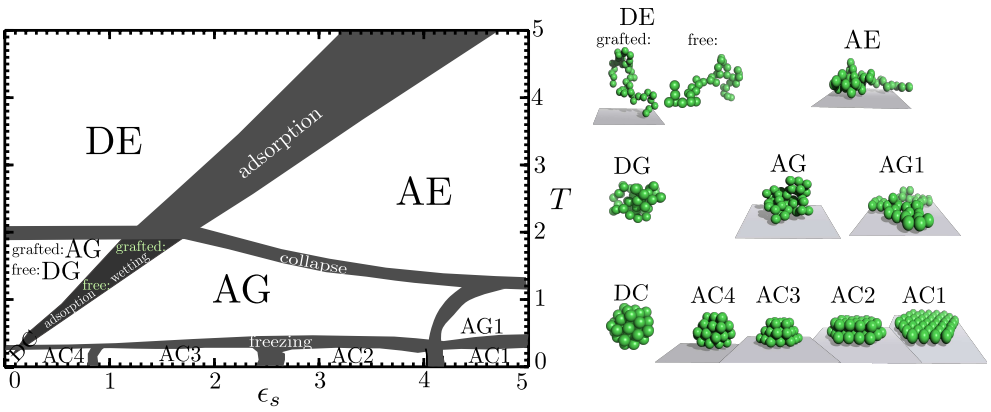


Figure 2. The pseudo-phase diagram parametrized by adsorption strength ϵ_s and temperature T for a 40mer. The purple transition regions have a broadness that reflects the variation of the corresponding peaks of the fluctuations of canonical expectation values we investigated. Phases with an "A/D" are adsorbed/desorbed. "E", "G" and "C" denote phases with increasing order: expanded, globular and compact/crystalline. The right panel shows representative conformations of the individual phases.

After having performed this simulation, one is confronted with the problem of how to combine all those canonical histograms $H_i(E)$, $i = 1, \ldots, n$, to get an optimal combination of the data available[26]. We use an error-weighted histogram reweighting method similar to the one in Ref. 27. All histograms $H_i(E)$ can be reweighted to yield an estimate of the density of states at inverse temperatures β_i, $g_i(E) \propto H_i(E)e^{\beta_i E}$, that is only of reasonable quality in an energy regime with sufficient statistics. Since absolute estimates of the partition function cannot be obtained in MC simulations, there is an unknown prefactor that is different for every β_i. To get rid of it, we work with the ratio $g_i(E + \Delta E)/g_i(E)$ and since $g(E)$ spans many orders of magnitude, it is advantageous to use the logarithm:

$$\begin{aligned} \Delta S_i(E) &= S_i(E + \Delta E) - S_i(E) = \ln\left[g_i(E + \Delta E)/g_i(E)\right] \\ &= \ln\left[H_i(E + \Delta E)\right] - \ln\left[H_i(E)\right] + \beta_i \Delta E. \end{aligned} \tag{3}$$

Now, an error-weighted average over all histograms,

$$\overline{\Delta S}(E) = \frac{\sum_i \Delta S_i(E) w_i(E)}{\sum_i w_i(E)}, \tag{4}$$

can be taken with $w_i(E) = 1/\sigma^2(\Delta S_i(E)) \simeq [H_i(E + \Delta E)H_i(E)]/[H_i(E + \Delta E) + H_i(E)]$. This gives an excellent estimate of $\beta(E) \approx \overline{\Delta S}(E)/\Delta E$ as long as all histograms overlap – which parallel tempering requires anyway. Up to a constant, the microcanonical entropy $S(E)$ is obtained by integration.

4 Results

All transitions are contained in the phase diagram for a 40mer in Fig. 2. It is constructed using the profile of several canonical fluctuations as shown for the specific heat in Fig. 3. A strong difference between the grafted and non-grafted case is observed at the adsorption transition, while the other transitions remain more or less unaffected.

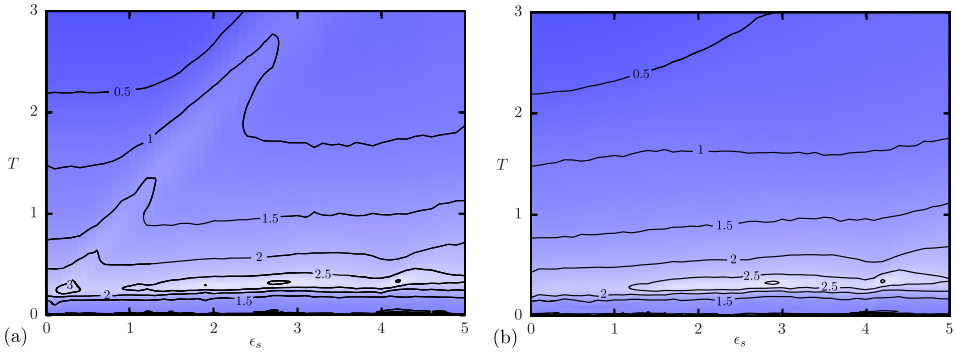


Figure 3. Specific-heat profile, $c_V(\epsilon_s, T)$, for (a) the non-grafted and (b) the grafted polymer.

Non-Grafted Polymer. The specific heat for the non-grafted polymer in Fig. 3(a) clearly reveals the freezing and adsorption transitions. The freezing leads to a pronounced peak near $T = 0.25$ almost independently of the surface attraction strengths. That this is indeed the freezing transition is confirmed by the very rigid crystalline structures found below this temperature and the rapidly decreasing density of states in the corresponding energy regime [cf. Fig. 4(a)].

To differentiate between the different crystalline structures, the radius of gyration, its tensor components parallel and perpendicular to the substrate, and the number of surface contacts were analyzed. This revealed that the crystalline phases arrange in a different number of layers to minimize the energy. For high surface attraction strengths, a single layer is favoured (AC1), and for decreasing ϵ_s the number of layers increases until for the 40mer a maximal number of 4 layers is reached (AC4), cf. Fig. 2. The fewer layers are involved in any layering transition, the more pronounced is that transition.

Raising the temperature above the freezing temperature, polymers form adsorbed and still rather compact conformations. This is the phase of adsorbed globular (AG) conformations that can be subdivided into drop-like globules for ϵ_s that are not strong enough to induce a single layer below the freezing transition and more pancake-like flat conformations (AG1) at temperatures above the AC1 phase. At even higher temperatures, two scenarios can be distinguished. In the first case, the polymer first desorbs from the sur-

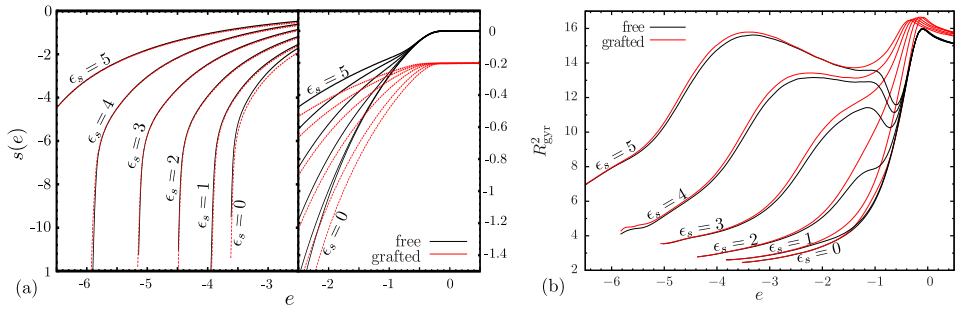


Figure 4. (a) Microcanonical entropy $s(e) = \ln g(e)/N$. Since $g(e)$ spans many orders of magnitude at low energies, the ordinate is divided into two different regimes. (b) Squared radius of gyration $R^2_{\rm gyr}(e)$ versus energy.

face [from AG to the desorbed globular (DG) bulk phase] and disentangles at even higher temperatures [from DG to the desorbed expanded bulk phase (DE)]. In the latter case, the polymer expands while it is still on the surface (from AG/AG1 to AE) and desorbs at higher temperatures (from AE to DE). The collapse transition in the adsorbed phase takes place at a lower temperature compared to the desorbed phase because the deformation at the substrate leads to an effective reduction of the number of contacts.

Fig. 4(a) reveals a convex regime at the adsorption transition in the microcanonical entropy $s(e)$ for extended chains. Hence, the adsorption transition of non-grafted chains is first-order like for finite chains. This convex intruder vanishes for longer chains[20] such that the continuous behaviour is regained in the thermodynamic limit.

Grafted Polymer. The grafting mainly has an influence on the adsorption transition. Fig. 3(b), e.g., reveals that it is strongly weakened for all ϵ_s, and Fig. 4 shows why this is so. The offset between the microcanonical entropy $s(e) = \ln g(e)/N$ of grafted and non-grafted polymers at high temperatures in Fig. 4(a) is proportional to the translational entropy difference of the two cases[20]. Due to the grafting, the translational entropy for desorbed chains is strongly reduced such that no convex intruder occurs for grafted chains and hence the adsorption of finite extended grafted polymers is continuous. Fig. 4(b) provides additional information about the conformational entropy. Directly at the adsorption transition, the overall radius of gyration gets reduced for non-grafted polymers compared to grafted ones if ϵ_s has a value large enough for the adsorption transition temperature to exceed the collapse transition. Conformations with such a reduced overall size – the same size that is on average adopted in bulk solution at this energy – are the transition states of the phase coexistence. Grafted chains are influenced by the surface in all cases, because they cannot escape. Hence, the conformational rearrangement of extended non-grafted polymers upon adsorption is not necessary and the adsorption is continuous.

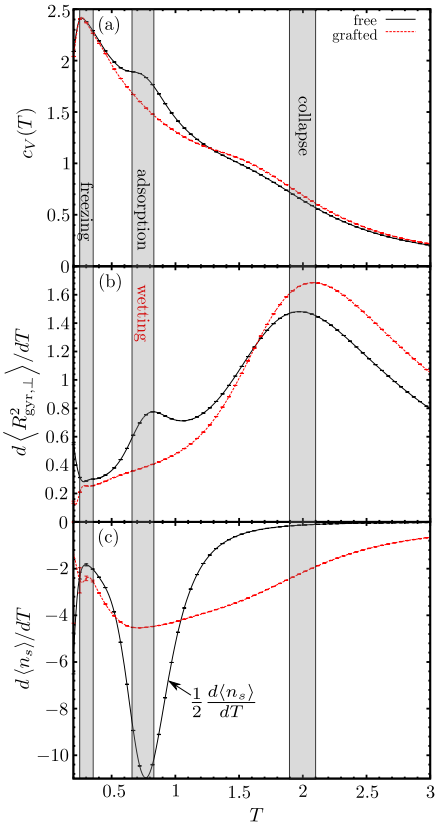


Figure 5. (a) Specific heat $c_V(T)$, (b) fluctuation of the radius of gyration component perpendicular to the substrate $d\left\langle R_{\mathrm{gyr},\perp}\right\rangle(T)/dT$, and (c) fluctuation of the number of monomers in contact with the substrate $d\langle n_s\rangle(T)/dT$ for weak surface attraction, $\epsilon_s = 0.7$, where the adsorption occurs at a lower temperature than the collapse.

The case of globular chains has to be discussed separately. While non-grafted globular chains adsorb continuously here, for grafted globular chains it even is nontrivial to identify an adsorption transition. A globular chain attached to a substrate always has several surface contacts such that a "desorbed globule" stops to be a well-defined description

here. One might, however, identify the transition from attached globules that only have a few contacts to docked conformations for stronger surface attraction strengths with the wetting transition. This roughly coincides with the position of the adsorption transition for the free chain between DG and AG in the phase diagram and is illustrated for $\epsilon_s = 0.7$ in Fig. 5. For the free polymer, at the adsorption transition a peak is visible in $c_V(T)$, $d\left\langle R^2_{\mathrm{gyr},\perp}\right\rangle/dT$ and $d\langle n_s\rangle/dT$. For the grafted polymer, the first two peaks disappear and with it the adsorption transition. Only a signal in the number of surface contacts is left. This change of surface contacts in an otherwise unchanged attached globule signals the wetting transition.

5 Conclusion

By using extensive parallel tempering simulations, we have analyzed and compared the whole phase diagram of a generic off-lattice model for grafted and non-grafted polymer chains for a range of temperatures and surface interaction strengths. The main differences were found at the adsorption transition. Here, the restriction of translational entropy due to grafting is much stronger above than below the transition. Additionally, grafting reduces the necessary rearrangement of segments to form substrate contacts and to adsorb such that grafted adsorption is always continuous whereas the adsorption of the free chain exhibits first-order like signatures for strong surface attraction and short chains. When for grafted chains, the adsorption temperature is below the coil-globule transition temperature, there are always several surface contacts present and the adsorption changes into the wetting transition. For free chains, a continuous adsorption transition exists here.

Acknowledgements

M. M. thanks the Forschungszentrum Jülich for hospitality during extended visits. This work is partially supported by the DFG under Grant No. JA 483/24-3, the Leipzig Graduate School of Excellence GSC 185 "BuildMoNa", the SFB/TRR 102, the Deutsch-Französische Hochschule (DFH-UFA) under Grant No. CDFA-02-07, and by the German-Israel Program "Umbrella" under Grant Nos. SIM6 and HPC_2. Support by supercomputer time grant hlz17 of NIC, Forschungszentrum Jülich is gratefully acknowledged.

References

1. W. Paul *et al.*, *Unexpectedly normal phase behavior of single homopolymer chains*, Phys. Rev. E **75**, 060801 (2007).
2. T. Vogel, M. Bachmann, and W. Janke, *Freezing and collapse of flexible polymers on regular lattices in three dimensions*, Phys. Rev. E **76**, 061803 (2007).
3. M. P. Taylor, W. Paul, and K. Binder, *All-or-none proteinlike folding transition of a flexible homopolymer chain*, Phys. Rev. E **79**, 050801 (2009).
4. F. Kuhner, M. Erdmann, and H. E. Gaub, *Scaling exponent and Kuhn length of pinned polymers by single molecule force spectroscopy*, Phys. Rev. Lett. **97**, 218301 (2006).
5. M. Bachmann *et al.*, *Microscopic mechanism of specific peptide adhesion to semiconductor substrates*, Angew. Chem. Int. Ed. **122**, 9721 (2010).

6. D. E. Smith *et al.*, *The bacteriophage Φ29 portal motor can package DNA against a large internal force*, Nature **413**, 748 (2001).
7. K. Kegler, M. Salomo, and F. Kremer, *Forces of interaction between DNA-grafted colloids: An optical tweezer measurement*, Phys. Rev. Lett. **98**, 058304 (2007).
8. E. Eisenriegler, K. Kremer, and K. Binder, *Adsorption of polymer chains at surfaces: Scaling and Monte Carlo analyses*, J. Chem. Phys. **77**, 6296 (1982).
9. Z. Usatenko, *Adsorption of long flexible polymer chains at planar surfaces: scaling analysis and critical exponents*, J. Stat. Mech. Theor. Exp. **2006**, P03009 (2006).
10. J. Forsman and C. E. Woodward, *Prewetting and layering in athermal polymer solutions*, Phys. Rev. Lett. **94**, 118301 (2005).
11. R. Rajesh *et al.*, *Adsorption and collapse transitions in a linear polymer chain near an attractive wall*, Phys. Rev. E **65**, 056124 (2002).
12. J. Luettmer-Strathmann *et al.*, *Transitions of tethered polymer chains: A simulation study with the bond fluctuation lattice model*, J. Chem. Phys. **128**, 064903 (2008).
13. S. Metzger *et al.*, *Surface excess in dilute polymer solutions and the adsorption transition versus wetting phenomena*, J. Chem. Phys. **118**, 8489 (2003).
14. J. Krawczyk *et al.*, *Layering transitions for adsorbing polymers in poor solvents*, Europhys. Lett. **70**, 726 (2005).
15. M. Bachmann and W. Janke, *Conformational transitions of nongrafted polymers near an absorbing substrate*, Phys. Rev. Lett. **95**, 058102 (2005); *Substrate adhesion of a nongrafted flexible polymer in a cavity*, Phys. Rev. E **73**, 041802 (2006).
16. T. Vrbová and S. G. Whittington, *Adsorption and collapse of self-avoiding walks in three dimensions: A Monte Carlo study*, J. Phys. A **31**, 3989 (1998).
17. R. Hegger and P. Grassberger, *Chain polymers near an adsorbing surface*, J. Phys. A **27**, 4069 (1994).
18. Y. Singh, D. Giri, and S. Kumar, *Crossover of a polymer chain from bulk to surface states*, J. Phys. A **34**, L67 (2001).
19. M. Möddel, M. Bachmann, and W. Janke, *Conformational mechanics of polymer adsorption transitions at attractive substrates*, J. Phys. Chem. B **113**, 3314 (2009).
20. M. Möddel, W. Janke, and M. Bachmann, *Systematic microcanonical analyses of polymer adsorption transitions*, Phys. Chem. Chem. Phys. **12**, 11548 (2010).
21. M. Möddel, W. Janke, and M. Bachmann, *Comparison of the adsorption transition for grafted and non-grafted polymers*, preprint (2011).
22. W. A. Steele, *The physical interaction of gases with crystalline solids*, Surf. Sc. **36**, 317 (1973).
23. K. Hukushima and K. Nemoto, *Exchange Monte Carlo method and application to spin glass simulations*, J. Phys. Soc. Japan **65**, 1604 (1996).
24. D. J. Earl and M. W. Deem, *Parallel tempering: Theory, applications, and new perspectives*, Phys. Chem. Chem. Phys. **7**, 3910 (2005).
25. E. Bittner, A. Nußbaumer, and W. Janke, *Make life simple: Unleash the full power of the parallel tempering algorithm*, Phys. Rev. Lett. **101**, 130603 (2008).
26. A. M. Ferrenberg and R. H. Swendsen, *Optimized Monte Carlo data analysis*, Phys. Rev. Lett. **63**, 1195 (1989).
27. M. K. Fenwick, *A direct multiple histogram reweighting method for optimal computation of the density of states*, J. Chem. Phys. **129**, 125106 (2008).

Structure and Dynamics at Polymer-Solid Interfaces: A Molecular Dynamics Simulation of 1,4-Polybutadiene at Graphite

Wolfgang Paul[1], Leonid Yelash[2], Peter Virnau[2], and Kurt Binder[2]

[1] Institute of Physics, Martin Luther University, 66099 Halle (Saale), Germany
E-mail: wolfgang.paul@physik.uni-halle.de

[2] Institute of Physics, Johannes Gutenberg University, 55099 Mainz, Germany
E-mail: yelash, virnau, kurt.binder@uni-mainz.de

We will discuss Molecular Dynamics simulation results on the structural and dynamic behaviour in a polymer melt in contact with a solid surface. For sake of concreteness and to enable quantitative comparisons with experiment, we will focus on a chemically realistic model of 1,4-polybutadiene at a graphite surface. Understanding the influence that the presence of a solid surface exerts on the behaviour of the polymer melt is interesting both technologically and from a general statistical physics perspective. Technologically, it provides an important step into the rational design of polymer based nano-composite systems. From a statistical physics perspective, the results are expected to shed light onto the controversially discussed question of a diverging length scale underlying the structural glass transition.

1 Introduction

Our aim in this contribution is to provide a general introduction to the problems and questions we are addressing in our research, to explain the computational strategies employed to address these questions and to present a selected set of results obtained so far and interpret them in the light of the questions posed at the outset. In section 2 we will explain the motivation for our work, followed by a discussion of the model studied and the method employed in section 3. Section 4 will then present some results on structural and dynamic properties obtained from the simulation and discuss their relevance for experiments, while section 5 will offer some conclusions and an outlook on ongoing work.

2 Motivation

The properties of nano-structured materials, e.g., nano-particles or fibers embedded into a polymer matrix, are to a large degree determined by interfacial effects[1]. Our work is devoted to improve our understanding of the structure and dynamics of the interphase region between a solid and a bulk polymer melt which are still insufficiently understood. Clearly, there has to exist a region of atomic size, the interface, where the chemical structure changes from the solid to the polymer. But beyond that, there exists a regime of unknown and varying extent, the so-called interphase, over which structure and dynamics in the polymer melt differ from their bulk behaviour. The term interphase has been coined to describe this third phase besides bulk solid and bulk polymer introduced in the phenomenological modeling of the properties of composite materials[2]. From such phenomenological modeling this interphase was found to sometimes have a thickness of the order of 100 nm and more[2].

But also looking on a molecular scale on the self-diffusion coefficient of entangled polystyrene chains close to a solid wall by secondary ion mass spectroscopy[3] it was found that the diffusion close to the wall was slowed down by two orders of magnitude and that the effect of the wall extended into the melt over a distance of $10R_g$ (where R_g is the radius of gyration of the chains which is typically of the order of a few nanometers). In contrast, recent NMR[4] and neutron scattering experiments[5] observed only a very small effect on the diffusion parallel to the walls even for highly confined polymer melts. These findings are also compatible with recent results from simulations of a coarse-grained polymer model[6]. Both experiments, however, observe a slowing down of the polymer dynamics on intermediate time and length scales, compatible with a layer of modified dynamics of thickness of the order of one or a few nanometers also found in simulations[7,8]. This length scale increases upon lowering the temperature towards the glass transition[9,10]. Connected with the increase of this structural length scale is an increase of the length scale over which dynamic properties in the melt are altered by the presence of a wall (or a free surface)[9–11]. This in turn changes the average glass transition temperature of, for example, a polymer film. Both from simulations[9–11] and most experiments, for a comprehensive review see[12], it has been found that the presence of free surfaces or at most weakly attractive surfaces decreases the glass transition temperature of a polymer film, however, some dielectric studies and calorimetric[13] reported no change in T_g for the same polymer-solid systems. The presence of a strongly attractive surface on the contrary tends to increase the glass transition temperature of the film[14].

So, while a significant amount of knowledge about the behaviour of polymers near solid surfaces has been accumulated, much of the simulation results on this topic where obtained from coarse-grained models only and most of the experimental results concern the average response of the confined polymer melt only. Furthermore the question about the exact form of the temperature dependence of the length scale over which the solid surface influences the dynamic response of the polymer melt is not answered yet. This problem has regained interest lately in connection with the question whether there exists a diverging length scale connected with the glass transition[15,16]. It is computationally and experimentally challenging to extract such a length scale from investigations of the glass transition in the bulk and while, e.g., single molecule experiments[17] and simulations of single molecule fluorescence in glass forming polymers[18] clearly reveal heterogeneous dynamics in the polymer matrix, it has not been feasible yet to extract a length scale of heterogeneity from such studies. For confined systems, however, it is possible in computer simulations to measure a spatially resolved relaxation time as a function of distance from a confining wall and extract the growing length scale from suitable analysis of such data[19], also helping to develop approaches how to extract such a length scale experimentally.

3 Model and Simulation Method

Here we will shortly describe the simulation model and refer to existing literature for more detail and then discuss the simulation method in the context of its realization on the JUGENE supercomputer.

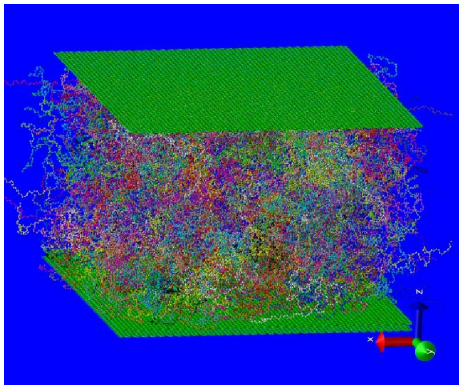


Figure 1. Snapshot configuration of a melt of 1,4-polybutadiene chains between 3 layers of graphite $(0, 0, 1)$. The lateral dimensions are 15 nm, the distance between the graphite walls is 10 nm.

3.1 Model

We employ a chemically realistic model of 1,4-polybutadiene developed in our group and shown to be able to quantitatively reproduce and predict experimental results for this polymer in the bulk (see Ref. 20 for a review). The model is given by a classical force field with intramolecular interactions (bond length, bond angle and torsion potential) and intermolecular interactions (van der Waals interaction parameterized in the form of a $12-6$ Lennard-Jones potential). There are no partial charges on the different atoms, i.e., 1,4-polybutadiene is a very apolar molecule. Carbon and hydrogen atoms are not represented as such, but carbon atoms and their attached hydrogen atoms are lumped into so-called united atoms having the combined mass of the constituting atoms. All interactions are parameterized from quantum chemistry for this representation[20].

The graphite surface is a $(0, 0, 1)$ surface defined by the equilibrium positions of the carbon atoms. The carbon-carbon interaction potential is of Lennard-Jones type and taken from the literature[21] with Lorentz-Berthelot combining rules determining the interaction between graphite carbon atoms and the united atoms of the polybutadiene chains. Three atomic layers are considered, and due to the use of periodic boundary conditions and the cutoff in the consideration of the Lennard-Jones interaction these represent two half spaces of graphite confining a slit filled with a polybutadiene melt as shown in Fig. 1.

3.2 Simulation Method

For our combination of apolar melt and apolar (and clean) solid surface, no chemically specific interactions like, e.g., hydrogen bonding can occur and also our simulations are at such high temperatures that it is not necessary to take quantum fluctuations into account. Therefore the simulation algorithm consists of a direct integration of Newton's equation of motion employing the forces derived from our chemically realistic model. As a further approximation we keep the carbon atoms in the graphite surface fixed.

Our simulations employ the GROMACS simulation package[22] which is one of the well-tested and highly optimized Molecular Dynamics simulation packages employed in

the community. At the beginning of our project it was not yet designed to be run on massively parallel machines like the JUGENE, but in close collaboration with the developers of GROMACS and IBM representatives the bottlenecks preventing the efficient implementation of GROMACS on the IBM Blue Gene were removed. For our systems of a typical size of 10^5 force centres we can employ one rack of the JUGENE machine with a reasonable efficiency (the speed-up for parallelization employing a domain decomposition scheme is roughly linear only to about half a rack for our system size).

Depending on the type of simulation method one is used to, this may not seem so impressive, so let us discuss some aspects of molecular dynamics simulations on parallel computers in general and the Blue Gene in particular. Molecular Dynamics simulations generate a time trajectory of a many particle system in phase space. The length of this trajectory has to be significantly longer than the largest time scales of the phenomena under study, whereas the single time step of the integrator has to be significantly smaller than the fastest time scale in the system. In our case, the integration time step is 1 fs whereas the simulated trajectory can be longer than 1μs, i.e, consist of more than 10^9 integration steps. System size is determined by the largest length scales in the problem, in our case the radius of gyration of the chains and the length scales over which the solid surfaces alter the behaviour of the polymer melt. These are of the order of a few nano-meter, which are well covered by the system size we are using. Increasing the system is, of course, always possible and would lead to a more parallel code, but would not provide additional physical insight. Consisting of relatively slow processors with relatively little memory, parallelism is by machine construction higher on the JUGENE than on, e.g, PC-type clusters. For the same reason, wall-clock time for the generation of long trajectories tends to be larger. With these machine constraints in mind, we geared our choice of system size and parallelization to optimally employ the potential of the machine.

4 Results

A first insight into the influence of the wall on the structure within the polymer melt can be obtained by looking at the segment density and chain centre of mass density as a function of distance to the wall, which are shown in Fig. 2. The layering in the segment density is a general phenomenon occurring even for dense liquids of hard spheres at a repulsive wall, i.e., it is of entropic origin. Nevertheless, it is much more pronounced for the case of a van der Waals attraction between wall atoms and segments which we study here. This attraction even leads to a well observable layering in the centre of mass density, a behaviour which only occurs at attractive surfaces like we are studying here. It was rarely heeded, yet, as we will see, it has pronounced influences on the dynamics of the polymer chains. Both layering effects extend over a distance of about 2 nm into the melt at $T = 353$ K which is about $2\,T_g$. Down to about $1.4\,T_g$ this range increases only very little. Concerning the influence of the solid surface on the dynamics in the adjacent polymer melt, let us focus on an observable which is measured in double quantum nuclear magnetic resonance (NMR) experiments. The following correlation function

$$C(t) = \langle P_2\left(\cos[\theta(t)]\right) P_2\left(\cos[\theta(0)]\right)\rangle \ , \tag{1}$$

where $\theta(t)$ is the angle between the orientation of the double bond in the polybutadiene monomer and a given magnetic field orientation, P_2 is the second Legendre polynomial

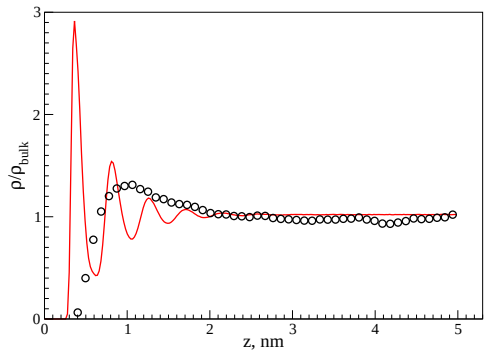


Figure 2. Density profiles of segments (red line) and chain centres (open circles) normalized by their bulk value as a function of distance to the solid surface. Half of the symmetric profiles is shown. The simulation temperature is $T = 353$ K.

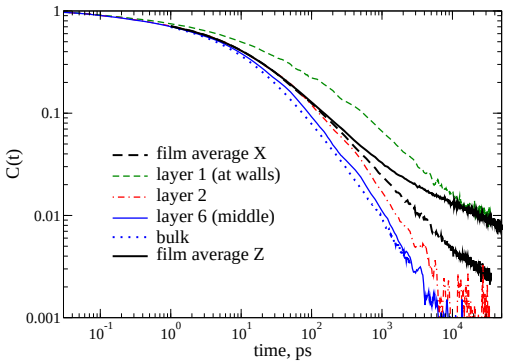


Figure 3. Correlation function $C(t)$ defined in Eq. 1 for three different layers within the film in comparison to the bulk behaviour (blue dashed line) and the film averaged behaviour (black line). All curves are for an orientation of the external magnetic field perpendicular to the walls (z-direction), except for the black dashed line which is the film averaged behaviour for an orientation parallel to the walls.

and the angular brackets indicate a thermal average is shown in Fig. 3. We have determined this correlation function in a layer resolved manner with a total of 11 layers between the two graphite surfaces at $z = 0$ and $z = D$. A given double bond contributes to the relaxation function for layer j if its centre is within this layer at $t = 0$. The relaxation behaviour close to the graphite walls is slowed down by more than an order of magnitude (green line) compared to the behaviour in the centre of the film (blue line) which agrees with the bulk relaxation behaviour (blue dashed line), i.e., the relaxation is heterogeneous

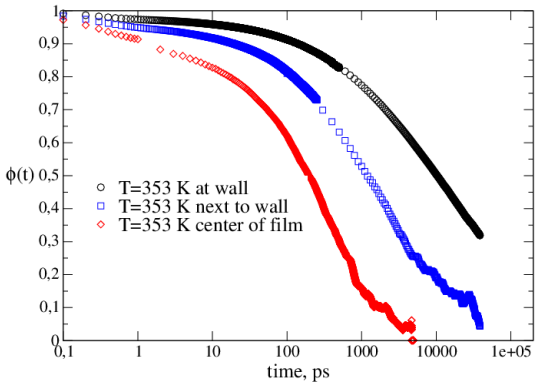


Figure 4. Correlation function $\phi(t)$ defined in Eq. 2 shown for chains located in different layers within the film.

across the film. The film averaged behaviour is still slowed down compared to the bulk relaxation and the degree of slowing down depends on the orientation of the external field. This proves that the dynamics is not only heterogeneous but anisotropic as well, agreeing with the different conclusions drawn from experiment susceptible to the dynamics normal to the surface[3] and parallel to it[4–6]. Looking at the averaged relaxation function for the field orientation normal to the surface, one can furthermore observe an additional slow process setting in at a time around 2 ns[23]. No such process exists in the bulk[20]. Clearly, this process is generated by the polymer-solid interactions which also generate the layering in the centre of mass density. The attractive interaction between the polymer segments and the wall atoms leads to a physisorption of the segments. To quantify the time scale of the desorption process, we define the following correlation function:

$$\phi(t) = \frac{\langle s(t)s(0)\rangle - \langle s\rangle^2}{\langle s(0)s(0)\rangle - \langle s\rangle^2} = \frac{\langle s(t)\rangle - \langle s\rangle^2}{1 - \langle s\rangle^2} \ , \tag{2}$$

where $s(t) = 1$ if the monomer is adsorbed at time t and $s(t) = 0$ otherwise, and where we have employed the constraint $s(0) = 1$ for the sampling of the autocorrelation function which is shown in Fig. 4. The adsorption autocorrelation shows a pronounced two-step decay, typical for correlation functions upon approach to the glass transition, however, at this high temperature all bulk relaxation functions show no such behaviour yet[20]. Most evidently for chains in the first layer, a time scale separation is occurring between a relaxation due to local segmental motions on a picosecond time scale and a desorption process of the complete chain on a nanosecond time scale. It is this chain desorption process which shows up as the additional relaxation process in the NMR observable displayed in Fig. 3. Since the amplitude of this process is very small for short chains, so far there exist only indications from NMR experiments that it might be observable. Of course, other techniques like dielectric spectroscopy or neutron scattering are susceptible to this process as well and it shows up as clearly predictable signatures in these experimental techniques[24]. Space limitations do not allow us to discuss this in more detail here.

5 Conclusions and Outlook

We have presented Molecular Dynamics simulations of 1,4-polybutadiene between graphite walls as a simple apolar model system for polymer-solid contacts. The polymer model employed was well validated in bulk simulations and was shown to quantitatively reproduce the relaxation processes accompanying the glass transition of this polymer. For the polymer adsorbed at graphite walls we established the existence of a layering effect in the centre of mass density, beyond the well-known and often studied segmental layering at the wall. This polymer effect has clearly observable consequences for the experimentally accessible relaxation functions of these confined polymer systems. The desorption of chains from the wall and their exchange between different layers, i.e., the motion perpendicular to the walls is strongly slowed down as compared to the bulk mobility at the same temperature. As this effect is caused by segment-wall interactions, it is controlled by different energetics than the glass transition in the bulk polymer melt and can be expected to influence the glass transition behaviour of these confined systems. First simulation studies of the effect of cooling the confined melt towards its bulk glass transition temperature have been performed and the grant of computer time allotted to our project in this year will enable us to continue these simulations down to approximately $1.2\,T_g$, a temperature regime where in the bulk the most interesting change in relaxation behaviour occurs. Thus we expect that we will be able to obtain some stimulating new results on the glass transition in confinement and on the existence of a growing length scale connected with the glass transition.

Acknowledgements

We are grateful for funding through the German science foundation within the SPP 1369. We are indebted to the Jülich Supercomputing Centre for a generous grant of computer time on the JUGENE which allowed us to perform this work.

References

1. S. G. Advani, *Processing and Properties of Nanocomposites* (World Scientific, Singapore, 2006).
2. L. T. Drzal, *The interphase in epoxy composites* in Advances in Polymer Sciences, Epoxy Resins and Composites II, Vol. 75, K. Durek, Ed., (Springer, Berlin, 1985).
3. X. Zheng, M. H. Rafailovich, J. Sokolov, Y. Strzhemechny, S. A. Schwarz. B. B. Sauer and M. Rubinstein, *Long-Range Effects on Polymer Diffusion Induced by a Bounding Interface*, Phys. Rev. Lett. **79**, 241 (1997).
4. S. Ok, M. Steinhardt, A. Serbescu, C. Franz, F. V. Chavez and K. Saalwächter, *Confinement Effects on Chain Dynamics and Local Chain Order in Entangled Polymer Melts*, Macromolecules **43**, 4429 (2010).
5. J. Martin, M. Krutyeva, M. Monkenbusch, A. Arbe, J. Allgeier, A. Radulescu, P. Falus, J. Maiz, C. Mijangos, J. Colmenero and D. Richter, *Direct Observation of Confined Single Chain Dynamics by Neutron Scattering*, Phys. Rev. Lett. **104**, 197801 (2010).

6. H. Meyer, T. Kreer, A. Cavallo, J. P. Wittmer and J. Baschnagel, *On the dynamics and disentanglement in thin and two-dimensional polymer films*, Eur. Phys. J. Special Topics **141**, 167 (2007).
7. K. Binder, A. Milchev, and J. Baschnagel, *Simulation Studies on the Dynamics of Polymers at Interfaces*, Ann. Rev. Mater. Sci. **26**, 107 (1996).
8. K. C. Daoulas and V. G. Mavrantzas, *Molecular Dynamics Simulation of a Polymer Melt/Solid Interface: Local Dynamics and Chain Mobility in a Thin Film of Polyethylene Melt Adsorbed on Graphite*, Macromolecules **38**, 5796 (2005).
9. K. Binder, J. Baschnagel, W. Paul, *Glass Transition of Polymer Melts: Tests of Theoretical Concepts by Computer Simulation*, Progress in Polymer Science **28**, 115 (2003).
10. J. Baschnagel, F. Varnik, *Computer simulations of supercooled polymer melts in the bulk and in confinement*, J. Phys.: Condens. Matter **17**, R851 (2005).
11. J.-L. Barrat, J. Baschnagel, A. Lyulin, *Molecular dynamics simulations of glassy polymers*, Soft Matter **6**, 3430 (2010).
12. M. Alcoutlabi and G. B. McKenna, *Effects of confinement on material behaviour at the nanometre size scale*, J. Phys.: Condens. Matter **17**, R461 (2005).
13. H. Huth, A. A. Minakow, A. Serghei, F. Kremer and C. Schick, *Differential AC-chip calorimeter for glass transition measurements in ultra thin polymeric films*, Eur. Phys. J. Spec. Top. **141**, 153 (2007).
14. G. D. Smith, D. Bedrov and O. Borodin, *Structural Relaxation and Dynamic Heterogeneity in a Polymer Melt at Attractive Surfaces*, Phys. Rev. Lett. **90**, 226103 (2003).
15. G. Biroli, J.-P. Bouchaud, K. Miyazaki and D. R. Reichmann, *Inhomogeneous Mode-Coupling Theory and Growing Dynamic Length in Supercooled Liquids*, Phys. Rev. Lett. **97**, 195701 (2006).
16. G. Szamel and E. Flenner, *Emergence of Long-Range Correlations and Rigidity at the Dynamic Glass Transition*, Phys. Rev. Lett. **107**, 105505 (2011).
17. R. A. L. Vallée, N. Tomczak, L. Kuipers, G. J. Vancso and N. F. Van Hulst, *Single Molecule Lifetime Fluctuations Reveal Segmental Dynamics in Polymers*, Phys. Rev. Lett. **91**, 038301 (2003).
18. R. A. L. Vallée, M. van der Auweraer, W. Paul, and K. Binder *Fluorescence lifetime of a single molecule as an observable of metabasin dynamics in fluids near the glass transition*, Phys. Rev. Lett. **97**, 217801 (2006).
19. P. Scheidler, W. Kob, K. Binder and G. Parisi, *Growing length scales in a supercooled liquid close to an interface*, Philos. Mag. B **82**, 283 (2002).
20. W. Paul and G. D. Smith *Structure and dynamics of amorphous polymers: computer simulations compared to experiment and theory*, Rep. Prog. Phys. **67**, 1117 (2004).
21. W. A. Steele, *The physical interaction of gases with crystalline solids: I. Gas-solid energies and properties of isolated adsorbed atoms*, Surf. Sci. **36**, 317 (1973).
22. `http://www.gromacs.org/`
23. L. Yelash, P. Virnau, K. Binder and W. Paul, *Slow process in confined polymer melts: Layer exchange dynamics at a polymer solid interface*, Phys. Rev. E **82** 050801 R (2010).
24. L. Yelash, P. Virnau, K. Binder and W. Paul, *Three-step decay of time correlations at polymer-solid interfaces*, submitted to Phys. Rev. Lett.

Earth and Environment

Ulrich Hansen

Institut für Geophysik, Westfälische Wilhelms-Universität Münster
Corrensstraße 24, 48149 Mnster, Germany
E-mail: hansen@earth.uni-muenster.de

Complexity seems inevitably to be connected to most of the problems from the section Earth and Environment. The range of spatial and temporal scales, as involved in these problems is immense and it seems almost impossible that experiments, conducted in laboratories are able to encompass the relevant scales. "Computer models", "numerical simulations", or, as I prefer to name them "numerical experiments" have become the most powerful research tools allowing for a continuous unraveling of interrelations which determine the structure and the dynamics of our living environment. Habitability of a planet, in today's view, results from a delicate interplay of processes in the atmosphere, the oceans and the planetary interior. Global, and local scales interact and are both of crucial importance. Properties of the Earth's subsurface, immediately beneath our feet need to be understood, in order to appropriately mange out groundwater resources. On global scale the Earth's magnetic field protects us from solar energetic particles and is today also viewed as a necessary ingredient for habitability. In order to understand the generation and maintenance of this phenomenon, not the subsurface region of planet, but the deep interior must be examined. Density driven convection currents in the molten interior of our planet at a depth of more than 3000 km are the machinery behind the Earth's magnetic field. Convection currents form a very important transport mechanism in all parts of the Earth's. Besides their importance for the Magnetic field, heat, as stored in the Earth is converted into mechanical energy in the Earth's mantle , providing the driving force for plate tectonics and, ultimately, for most tectonic activity on Earth. Boundary layers are known to play an essential role in all convective phenomena. In these boundary layers magnitudes like temperature, velocity or chemical composition vary strongly over relatively small distances. Of particular importance for the dynamics of the atmosphere is the so called "Planetary Boundary Layer".

The dynamics of the Planetary Boundary Layer is the topic of the contribution of Mellardo and Ansorge. This part of the atmosphere, which is directly influenced by the presence of surface (land or sea), has a large impact of the general behaviour of the atmosphere. Turbulent convection is a key mechanism determining the dynamics of the planetary boundary layer. In its simplest form it can be boiled down to investigating the problem of vigorous convection over a heated surface. However even fundamental aspects of this phenomenon are only poorly understood. While convection in the last years has become the paradigmatic example system to study various aspects of turbulence, many question do remain open. In their contribution "Direct Numerical Simulation of Turbulent Mixing in the Planetary Boundary Layer", Mellado and and Ansorge describe their efforts aiming at an understanding of the role of boundary conditions on the turbulent mixing by means of Direct Numerical Simulation. They investigate the case of a thermally driven and mechanically driven flow. Especially the effect of stratification on the latter one is examined. The

authors are able to identify three different regimes, depending on the strength of stratification. A "state of the art" DNS Approach allows them to gain insight into the interaction of turbulence and waves. Studies of this type were virtually impossible until the power of todays supercomputers became available. In order to study interactions of turbulent flow and wave motions (like gravity waves) extended computational domains over sufficiently long time periods are necessary.

The contribution "Simulation of Processes in Terrestrial Systems", by Vanderborght et al., describes a variety of phenomena focusing on subsurface flow problems. Clearly, less fundamental and more applied aspects are in the focus of this contribution. Topics range from influence of water uptake by plant roots on the subsurface water transport to characterization of the subsurface structure. Similarly diversified as the tackled problems are the applied methods. From simulations (forward modeling), over data assimilation to inversion ranges the spectrum of applied methods. Ground penetrating radar has proven its usefulness in several settings. It seems especially powerful in determining the thickness of ice shields in the polar regions of the Earth and also on mountain glaciers. In the contribution by Vanderborght et al., data from GPR measurements are used to unravel the structure of the subsurface ground. Not only a piece of information is used (for the example the speed of the waves), but instead a full waveform inversion, taking into account all information that influences the waveforming has been employed in this study.

The third contribution in the Earth and Environment section focuses on Atmospheric Processes ("Modeling of Atmospheric Processes", by König et al.), concentrating on practical as well as on fundamental aspects of atmospheric dynamics. The microphysics of clouds is today considered to have a profound influence on not only the short term weather development, but also on long term climatic changes. Aerosol particles, i.e. their concentration, their size distribution and their chemical properties are significant ingredients which need to be taken into account in order to get a proper view on the microphysics of clouds. There is however a further aspect, associated with aerosol particles. The question of air quality and particularly the distribution of fine dust depends largely on the behaviour of aerosols. The present study focuses on two topics: the flow in an urban setting is investigated by employing large eddy simulation techniques and further the dust distribution in Central Europe is studied. The authors conclude that their study is rather at the beginning, than at the end.

And since it seems a fairly general statement that in science new answers generate new question, this is possibly be true for all contributions. It also means that we can look forward, towards more exciting research in the following years.

Direct Numerical Simulation of Turbulent Mixing in the Planetary Boundary Layer

Juan Pedro Mellado and Cedrick Ansorge

Max Planck Institute for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany
E-mail: {juan-pedro.mellado, cedrick.ansorge}@zmaw.de

We present two problems of on-going research on turbulent mixing in the planetary boundary layer using direct numerical simulation. The first one addresses the topic of boundary effects in a buoyancy-driven system and describes the temporal evolution and structure of a free convective boundary layer over a heated plate. The stronger mixing in the near-wall region caused by free-slip conditions, compared to no-slip conditions, is studied. The second problem focuses on the Ekman layer, a shear-driven wall-bounded flow, and the effect of stable stratification on it. The three regimes found in the atmosphere are observed in the simulations, namely, weakly, intermediately and strongly stratified. The last two regimes provide a framework to investigate in depth the interaction between internal gravity waves and turbulence.

1 Introduction

The planetary boundary layer (PBL) is that region of the atmosphere directly influenced by the land or sea surface. Atmospheric turbulence is a key element of the PBL because it controls the transfer of mass (constituents like dry air, water vapor or carbon dioxide), momentum and energy between the surface and the upper, free atmosphere. Concomitantly, it still qualifies as *unsolved* problem and constitutes one major area of research[1,2]. In this context, basic, apparently simple turbulent processes are still lacking a fundamental understanding and a detailed characterization. One example is transients and changes of regimes, relevant because the external forcing on the PBL varies continuously in time – for instance, during the day. Another important case is the interaction between turbulence and gravity waves, the latter ubiquitous in the atmosphere because of buoyancy variations. The role that these basic processes play in small-scale regions of the earth system needs to be better understood, not only from a fundamental point of view but also to improve current models in climate research and numerical weather prediction. Our research group studies some of these problems and this paper presents two of them.

As a tool, we use direct numerical simulation (DNS). Significant contributions to the understanding of turbulence and turbulent mixing have been made based on DNS during the last decades[3]. DNS consists in solving as exactly as possible the partial differential equations governing the evolution of the mass, momentum and energy of a fluid or a fluids mixture – the Navier-Stokes equations. In a turbulent state, we need to account for the very broad range of spatio-temporal scales embracing almost all of the non-linearly interacting, three-dimensional modes. Hence, high performance computing is required.

Research on the PBL has often been based on observations, mixed layer theory and large eddy simulations. However, as explained before, unresolved small-scale mixing in certain regions of the system frequently prevents any definite conclusion. DNS is best suited for the study of small-scale mixing because it does not invoke any turbulence closure, and Reynolds number similarity (scale separation) counteracts the limitation to rel-

atively small domains imposed by current computational capabilities[4]. For instance, the potential of DNS in the investigation of these particular regions of the PBL has been shown recently for the case of the stratocumulus top[5,6].

The two problems presented here have been chosen because they concern some aspects of two important limiting cases of nature: the convectively-driven and the mechanically-driven regimes. The former is considered in Sec. 3 in terms of unsteady convection over a heated plate. The latter, discussed in terms of stratification effects in the Ekman layer, is described in Sec. 4. Before that, we briefly introduce the governing equations and the numerical algorithm.

2 Formulation

The system is described in terms of the velocity vector $\mathbf{v}(\mathbf{x},t)$ and the buoyancy $b(\mathbf{x},t)$ acting in the vertical direction Oz. The small magnitude of the velocities found in the PBL compared to the speed of sound and the small size of the domains considered in these studies compared to the scale height justifies an incompressible approach. Density variations smaller than 5% permits the Boussinesq limit. The governing equations are

$$\begin{aligned} &\frac{\partial \mathbf{v}}{\partial t} + \nabla\cdot(\mathbf{v}\otimes\mathbf{v}) + 2\Omega\mathbf{k}\times(\mathbf{v}-G\mathbf{i}) = -\nabla p + \nu\nabla^2\mathbf{v} + b\mathbf{k} \\ &\qquad \nabla\cdot\mathbf{v} = 0 \\ &\frac{\partial b}{\partial t} + \nabla\cdot(\mathbf{v}b) = \kappa\nabla^2 b \end{aligned} \tag{1}$$

The kinematic viscosity is ν, κ is a scalar diffusivity, p is a modified kinematic pressure, Ω is the angular velocity of the non-inertial system, G the corresponding geostrophic wind, and $\mathbf{i}$ and $\mathbf{k}$ are the unit vectors along Ox and Oz, respectively. The Prandtl number $Pr = \nu/\kappa$ is 1 in all the cases discussed here.

The geometry consists of a horizontal layer of fluid and the system is statistically homogeneous inside horizontal planes xOy. Periodicity is used in these two directions, which is acceptable as long as the domain size is large compared to the auto-correlation lengths of the turbulent fields. The top boundary is placed far enough from the region of interest, and no-penetration free-slip conditions are used there, along with Neumann ones for the scalar. Additional constraints are imposed depending on the problem and are discussed in Sec. 3 and Sec. 4.

High order spectral-like compact finite-differences are used to discretize (1) on a structured grid, along with a low-storage fourth-order Runge-Kutta scheme to advance in time. The discrete solenoidal constraint is satisfied down to machine accuracy using a Fourier decomposition along the horizontal directions and a factorization of the resulting set of equations along the vertical coordinate[7].

In terms of implementation, the code is written in Fortran90 and uses hybrid parallelization: MPI for the domain decomposition in Ox and Oy, and OpenMP within each node for the main loops. The linear systems are solved very efficiently using LU decompositions, and the code employs the level 1 BLAS routines from the IBM ESSL library and profits from SIMD vectorization. The library MPI_IO is used to read and write data in a parallel mode.

3 Unsteady Convection over a Heated Plate

It is a matter of common experience that a heated plate promotes motion inside the fluid lying on top of it[4]. This is a classical problem, but there are still several motivations for its study. One of them is to better understand the effect on the turbulence of having free- or no-slip conditions on the velocity field at the surface. One direct application is for instance the air-water interface: turbulent motion inside the atmosphere experiences close to no-slip conditions at the bottom (see Fig. 1), and free slip characterizes better the interface from the point of view of the sea surface turbulent layer. A more subtle case occurs when a relatively strong stable stratification at a density interface acts as a rigid plate, restricting the vertical motion, whereas the in-plane velocity components experience conditions somehow between free- and no-slip. One example is the inversion at the stratocumulus top in the sub-tropical, maritime PBL referred to in the introduction[5]. Previous work on this topic is scarce and we wanted to gain more insight into it and set a reference for future work on surface effects in the convective boundary layer. It is also noted that, to the authors' knowledge, this is the first DNS study of unsteady free turbulent convection over a heated plate and complements data that dates back to Townsend[8].

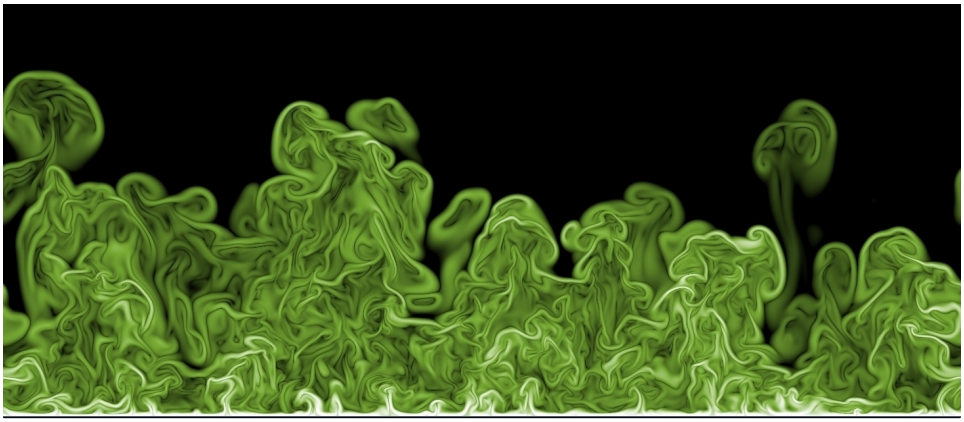


Figure 1. Logarithm of the magnitude of the gradient of the temperature field inside a vertical plane in the convective boundary layer over a heated plate. The magnitude increases according to the sequence black-green-white. The grid size is $3072 \times 3072 \times 1536$. Rayleigh numbers are between 10^8 and 10^9.

A second motivation is that free convection is usually studied in the context of Rayleigh-Bénard convection[9]. This configuration is statistically steady. Geometrically, it is characterized by the vertical separation of the two plates and, in cases of a closed container, the aspect ratio. Our work complements this research with an unbounded configuration. The basic difference is that the removal of the upper plate eliminates the length scale externally imposed on the system (the artificial effect here of the computational box is monitored and maintained within the desired limits), and the dependence of the statistics on time is introduced instead. This modification changes the large-scale structure and thus the well-known interaction of this large-scale organization of the flow and the near-wall region.

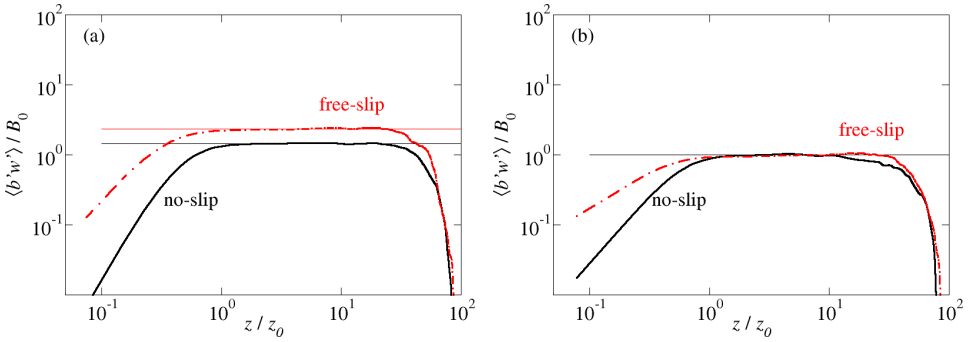


Figure 2. Vertical profiles of turbulent buoyancy flux, panel (a) showing cases of constant buoyancy at the wall, and panel (b) plotting cases of constant buoyancy flux. Colours denote different boundary conditions on the velocity field. Thin line indicates the inertial limit.

The set of equations we solve are (1) with $\Omega = 0$, i.e. no rotation is retained. The boundary conditions at the plate, which defines the origin $z = 0$, can be either constant buoyancy, b_0 given, or constant buoyancy flux, B_0 given. Dimensional analysis shows that the turbulent regime established after the initial transient does not depend on any non-dimensional parameter (recall that $Pr = 1$). One simulation per case is therefore enough.

Fig. 1 depicts the flow at a particular time after the transient for one of the cases. A turbulent boundary layer is observed, with the characteristic lamellar structure of the scalar field. As a function of time, the boundary layer simply thickens. Turbulence brings continuously relatively cold fluid from above (black colour) near the hot plate below, and a dynamic balance is eventually achieved in which some mean properties at the wall become stationary.

The vertical structure can be better quantified in terms of the mean buoyancy flux

$$B = \langle b'w' \rangle - \kappa \partial \langle b \rangle / \partial z \ , \tag{2}$$

where angle brackets indicate horizontal averages and an apostrophe denotes a turbulent fluctuation. The turbulent contribution $\langle b'w' \rangle$ is shown in Fig. 2. It is zero at the wall because of the no-penetration condition $w = 0$, and increases rapidly to a constant value over a relatively thin region of order z_0. Concomitantly, the molecular part $-\kappa \partial \langle b \rangle / \partial z$ decreases to become negligible for $z > z_0$ (not included in that same figure for clarity). The thickness z_0 of that diffusive sublayer is given by $10(\kappa^2/b_0)^{1/3}$, in the case of a constant buoyancy b_0 imposed at the wall, or by $(10^3 \kappa^3 / B_0)^{1/4}$ if the flux B_0 is fixed instead. The sum of both molecular and turbulent fluxes, B, exhibits a constant level all the way down to the wall at $z = 0$, conforming the so-called inner region. Eventually, the buoyancy transport falls to zero far enough from the wall, as we transverse the region of external intermittency and move deeper into the upper homogeneous layer. It is noted that this region of external intermittency is the thickest of the three layers, as observed in Fig. 1, and is one of the causes for the need of large-scale simulations.

The major difference between having no-slip or free-slip boundary conditions on the velocity is that free-slip conditions lead to significantly stronger mixing. This is so for both constant buoyancy and constant buoyancy flux cases. Consequently, some characteristics

of the near-wall structure of the flow are modified, since more kinetic energy is involved in the local dynamics. A second result from these simulations is that some terms of the budget equations of the Reynolds stresses (not shown) vary by an order of magnitude within the diffusive layer of order z_0 and there exists a strong anisotropy, which suggests more elaborated parametrizations than those currently used in low-order models.

4 Stably Stratified Boundary Layer

The Ekman layer typifies the mechanically-driven PBL[10]. Its defining characteristic is that the frame of reference rotates with an angular velocity $\mathbf{\Omega}$, and non-inertial forces are relevant. In such a system, one basic homogeneous solution is a horizontal wind $\mathbf{G}$ satisfying the geostrophic balance $2\mathbf{\Omega} \times \mathbf{G} + \nabla p = 0$. The Ekman layer develops when this flow encounters a no-slip condition at the bottom surface. Choosing the direction Ox along the horizontal wind and Oz along the angular velocity, the problem is described by the system of equations (1). Far above the surface, $\mathbf{u} \to (G, 0, 0)$; at the bottom $z = 0$, the boundary condition $\mathbf{u} = \mathbf{0}$ applies. In the neutral case, the buoyancy term b in the momentum equation is not retained and the scalar is passive. Then, a Reynolds number $Re = G/\sqrt{\Omega\nu}$ is the only non-dimensional parameter characterizing the system. The laminar solution is unstable for the Re values of interest and, when perturbed, a turbulent regime is established – we focus on this turbulent regime.

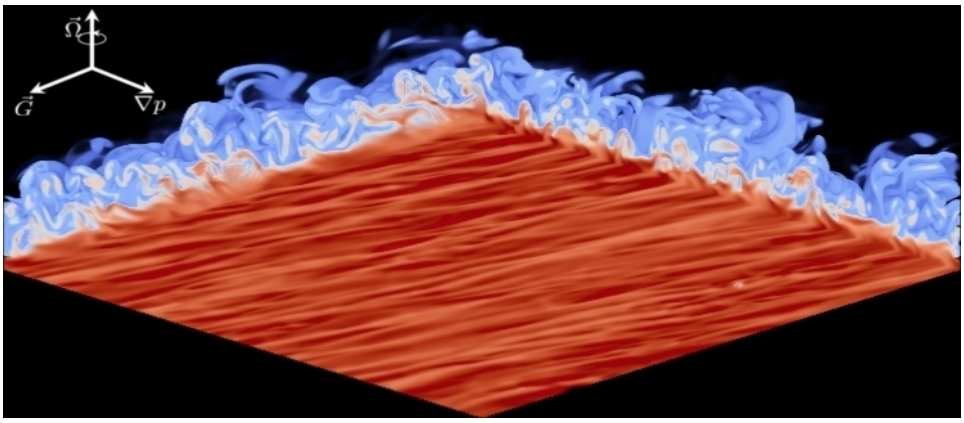


Figure 3. Neutrally stratified Ekman layer. Magnitude of the enstrophy increasing in the sequence red-white-blue, showing the turbulent boundary layer (blue) and the streaks pattern next to the wall (red). Only a subdomain $1/4 \times 1/4 \times 1$ of the total grid $2048 \times 2048 \times 192$ is shown, for clarity. The system is observed from below to reveal the streaks next to the wall.

Fig. 3 depicts the flow. The system is statistically steady and the mean thickness of the boundary layer observed there in blue colour is about $0.7\, u_*/(2\Omega)$, where u_* is the friction velocity defined in terms of the shear stress at the wall. The vertical structure is similar to that of other wall-bounded flows: viscous sublayer, logarithmic layer and outer region. However, the mean pressure gradient from the outer geostrophic balance turns the mean

flow to the left of the geostrophic wind $\mathbf{G}$ as we approach the surface; this is the so-called Ekman spiral. This is manifest in Fig. 3, where the streaks in the near-wall region (the elongated patterns of dark and light red colour) form an angle of about 20° with the Ox direction.

The velocity spiral in the Ekman layer complicates the problem significantly with respect to the convectional turbulent boundary layer. In fact, previous work on the topic is very scarce when compared to other wall-bounded flows[11]. This is the first reason to further investigate this problem, even in the neutrally stratified limit. Relatively small grid sizes have been used to validate the general setup following Coleman et al.[12]. However, the reference cases, like those shown in Fig. 3 and later in Fig. 4, have been defined using a significantly larger horizontal extent. This allows a better statistical convergence, and assures at the same time that large-scale structures can develop and de-correlate appropriately, the latter being crucial for the stably stratified system described below. Details of the structure and corresponding spectra are being currently analyzed and results will be submitted for publication.

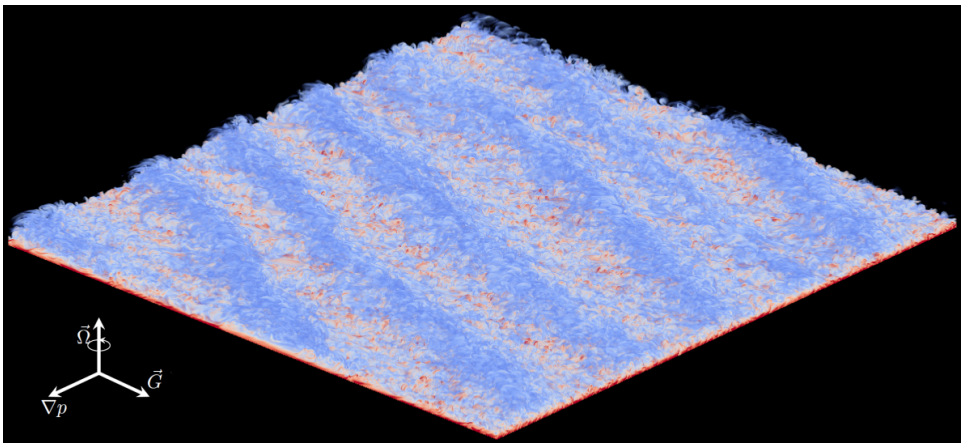


Figure 4. Stably stratified Ekman layer, observed from above to reveal the waves in the outer layer. Same legend as in Fig. 3, except that the complete domain is used here.

The second and main motivation for this investigation of the Ekman layer is the study of the interaction between internal gravity waves and turbulence. As explained in the introduction, this interaction is very relevant in general in geophysical fluid mechanics and the understanding thereof remains limited, in particular in the PBL[13–15]. Stable stratification is introduced by retaining the buoyancy term b in the momentum equation of the system (1). A buoyancy difference b_0 between the surface and the upper, free atmosphere, is added to the problem with an appropriate Dirichlet boundary condition at the lower wall. This translates into a second non-dimensional parameter, a Richardson number $Ri = b_0\sqrt{\nu/\Omega}/G^2$, in addition to the Reynolds number Re. This stratification represents the effect in nature of, for instance, the night-time radiative cooling of the surface or simply the horizontal advection over ice. If detailed work on the neutrally stratified case was said before to be scarce, detailed analysis based on DNS of the stably stratified Ekman layer practically does not

exist[16]. Some regimes of the turbulent flow are statistically steady and show certain similarities to channel flow, which is then often used as a surrogate[17,18]. However, there is one fundamental difference: the Ekman layer develops an outer region, where free atmosphere air is entrained into the turbulent zone. We retain here this characteristic.

A set of simulations has been carried out taking the neutral case as initial condition and varying the value of the Richardson number Ri. The three regimes commonly found in the atmosphere are clearly distinguished in these simulations as the parameter Ri is increased: weakly, intermediately and strongly stratified[14,15]. In the weak case the buoyancy field behaves practically as a passive scalar, with only a small modification in the outer layer. Previous results are reproduced, like the Monin-Obukhov scaling commonly used in turbulence closure models, at least for the moderate Reynolds numbers achieved in this work. In the strongly stratified case, turbulence decays monotonically and the flow tends to relaminarization, though turbulent bursts have been conjectured in the past; further study of this regime has been deferred for future work. Current research focuses on the intermediate regime, where the interaction between waves and turbulence characterizes the dynamics of the system. As explained before, DNS allows a detailed, controlled study of this problem without the spurious interference of closure models, and it constitutes a relative new approach. A three-dimensional rendering depicting that state is shown in Fig. 4. It is observed there that relatively large-scale waves form at a definite angle with respect to the direction Ox, an angle different from that of the streaks near the wall seen before in Fig. 3. The outer region changes therefore considerably with respect to the neutral case and this global external intermittency suggests the use of conditional statistics. On top of this intricate spatial structure, a temporal oscillation emerges in some of the statistics, like for instance the turbulent buoyancy flux within the outer layer. This oscillation cannot be easily explained in terms of the external parameters imposed on the system, like the rotation frequency 2Ω, and indicate an interaction between the inner and the outer region.

5 Concluding Remarks and Outlook

Some relevant aspects of turbulent mixing in the planetary boundary layer have been discussed. Both limiting regimes, the convectively-driven and the mechanically-driven, have been considered. In the first case, emphasis has been made on the role of boundary conditions (free- and no-slip) in the unsteady regime. A vertical structure in terms of an inner and outer layer appears. The outer layer is characterized by large-scale motions and external intermittency. The inner layer is formed by a thin diffusive layer dominated by molecular transport and an inertial layer where the turbulent buoyancy flux is constant and the relevant length scale is the distance to the wall. A stronger turbulence intensity develops in the case of free slip, which affects the local dynamics. In the second case, the Ekman layer exhibits a similar vertical structure. The crucial difference with other wall-bounded turbulent flows is the Ekman spiral, which translates into a different organization of the outer layer. Stable stratification further adds internal gravity waves. It has been shown that the different regimes can be successfully reproduced in a controlled way, and the work on one of these regimes has been initiated. Current capabilities in high performance computing enables now this detailed investigation on the interaction of waves and turbulence, impossible a few years ago due to the need of large domains and long simulations.

Acknowledgements

Support from the Max Planck Society through its Max Planck Research Groups program is gratefully acknowledged. The calculations were carried out at the Jülich Supercomputing Centre.

References

1. J. C. Wyngaard, *Atmospheric turbulence*, Annu. Rev. Fluid Mech., **24**, 205–233, 1992.
2. B. Stevens, *Atmospheric moist convection*, Annu. Rev. Earth Planet. Sci., **33**, 605–643, 2005.
3. P. Moin and K. Mahesh, *Direct numerical simulation: A tool in turbulence research*, Annu. Rev. Fluid Mech., **30**, 539–78, 1998.
4. D. J. Tritton, *Physical fluid dynamics*, Oxford Science Publications, 1988.
5. J. P. Mellado, *The evaporatively-driven cloud-top mixing layer*, J. Fluid Mech., **660**, 1–32, 2010.
6. B. Stevens, *Cloud-top entrainment instability?*, J. Fluid Mech., **660**, 1–4, 2010.
7. J.P. Mellado and C. Ansorge, *Factorization of the solenoidal constraint using implicit finite-difference schemes in colocated grids*, ZAMM - Zeitschrift für angewandte Mathematik und Mechanik (submitted), 2011.
8. A. A. Townsend, *Temperature fluctuations over a heated horizontal surface*, J. Fluid Mech., **5**, 209–241, 1959.
9. E. D. Siggia, *High Rayleigh number convection*, Annu. Rev. Fluid Mech., **26**, 137–168, 1994.
10. A.S. Monin, *The atmospheric boundary layer*, Annu. Rev. Fluid Mech., **2**, 225–250, 1970.
11. P. R. Spalart, G. N. Coleman, and R. Johnstone, *Retraction: Direct numerical simulation of the Ekman layer: A step in Reynolds number, and cautios support for a log law with a shifted origin*, Phys. Fluids, **21**, no. 109901, 1–3, 2009.
12. G. N. Coleman, J. H. Ferziger, and P. R. Spalart, *A numerical study of the turbulent Ekman layer*, J. Fluid Mech., **213**, 313–348, 1990.
13. H. J. S. Fernando, *Turbulent mixing in stratified fluids*, Annu. Rev. Fluid Mech., **23**, 455–493, 1991.
14. L. Mahrt, *Stratified atmospheric boundary layers*, Boundary-Layer Meteorol., **90**, 375–396, 1999.
15. T. Mauritsen and G. Svensson, *Observations of stably stratified shear-driven atmospheric turbulent at low and high Richardson numbers*, J. Atmos. Sci., **64**, 645–655, 2007.
16. G. N. Coleman, J. H. Ferziger, and P. R. Spalart, *Direct simulation of the stably stratified turbulent Ekman layer*, J. Fluid Mech., **244**, 677–712, 1992.
17. S. S. Zilitinkechich, *Comments on the numerical simulation of homogeneous stably-stratified turbulence*, Boundary-Layer Meteorol., **136**, 161–164, 2010.
18. O. Flores and J.J. Riley, *Analysis of turbulence collapse in the stably stratified surface layer using direct numerical simulation*, Boundary-Layer Meteorol., **139**, 241–259, 2011.

Simulation of Processes in Terrestrial Systems

Jan Vanderborght[1], Olaf Ippisch[2], Markus Blatt[2], Stefan Kollet[3], Shams Rahman[3], Taiseer Aljazzar[3], Reed Maxwell[4], Natalie Schröder[5], Mathieu Javaux[1,6], Harrie-Jan Hendricks-Franssen[1], Wolfgang Kurtz[1], Jan van der Kruk[1], Max Oberröhrmann[1], Anja Klotzsche[1], and Harry Vereecken[1]

[1] Agrosphere Institute, IBG-3, Forschungzentrum Jülich (Germany)
E-mail: j.vanderborght@fz-juelich.de

[2] Interdisciplinary Center for Scientic Computing, Heidelberg University (Germany)

[3] Meteorological Institute, Bonn University (Germany)

[4] Integrated GroundWater Modeling Center, Colorado School of Mines, Golden (USA)

[5] Jülich Supercomputing Centre, JSC, Forschungzentrum Jülich (Germany)

[6] Earth and Life Institute/Environmental Sciences, Université catholique de Louvain (Belgium)

We give an overview of simulations of processes in terrestrial systems that have been carried out recently using (or with potential use of) supercomputing infrastructure at the Jülich Supercomputing Centre. The focus of this contribution is more on the processes that are simulated than on the technical aspects of the simulations on the supercomputing systems. The simulation examples span a range of scales from a single root to a river catchment. Approaches to improve the parameterisation of simulation models making use of data assimilation techniques and high resolution geophysical methods are also presented.

1 Introduction

In this contribution we will give an overview of simulation studies in terrestrial systems that we have carried out recently. The general aim of these simulation studies is to obtain a better understanding of terrestrial systems. This understanding is required to develop and parameterize models that can predict the impact of changes and feedbacks in these systems due to external forcing. Concrete examples of processes that we investigate are water fluxes in soils and aquifers, water and heat transfers between the land surface and the atmosphere, transport of dissolved substances in soils and aquifers, assimilation of carbon by vegetation and carbon cycles in soils. These processes are closely linked to crop and food production and the management of soil and groundwater resources. Climate change and world population growth combined with changing consumption patterns are the main causes for changes that are expected in terrestrial systems. According to the World Economic Forums Global Risk Survey, perceived risks and financial impacts of global change and food and water security rank among the highest. This sets the relevance of improving the predictive power of models that simulate the impact of changes on processes in terrestrial systems.

A characteristic of terrestrial systems is that they are hierarchically structured. At small scales, the systems behaviour is determined by for instance the structure of the pore space or the structure of the root network of a plant. The distribution of geological layers in a geological basin or the distribution of soil types and vegetation in a landscape are important

for the systems behaviour at a larger scale. A central problem in terrestrial systems modelling is how to transfer the impact of small scale structures on processes at a larger scale. This transfer may involve a change in model structure and or a change in model parameters. Simulation studies that resolve the effect of smaller scale structures on larger scale processes are an ideal tool to realize this transfer of model structure and model parameters over different scales. A first set of studies that will be presented in this contribution are examples of this scale transfer. In a first set of studies, uptake of water and dissolved substances by plant roots, which is a key component in the terrestrial water cycle, is modelled at a spatial resolution that considers single plant roots. In a second set of studies, the effect of spatial variations in soil properties that exist at cm and m scales on water flow processes is modelled within a soil block of 10x10x10m. Simulations at this scale serve as a basis to parameterize grid cells of larger scale models. In a third set of studies, simulations of coupled atmosphere-soil-groundwater processes are carried out at the scale of a stream catchment. Using these kinds of simulations, the relation and feedbacks between topography, soil and vegetation distributions, groundwater depths on the one hand and the meteorology and hydrology in the region on the other can be evaluated. Recognition of these feedbacks is important to evaluate the effect of climate and land-use changes on soil and groundwater resources. For these three sets of simulation studies, supercomputing facilities are needed to describe systems with a sufficiently large extent and with a sufficiently fine spatial resolution.

Another approach to determine model parameters and improve the prediction quality of larger scale models is to assimilate observations in model predictions. The effectiveness of data-assimilation techniques is often hampered by the lack of computing power so that supercomputing facilities are needed to demonstrate their use.

Finally, given the importance of subsurface structures for the parameterisation of simulation models, the structure of the subsurface must be characterized with sufficient detail in order to ascertain the predictive capacity of simulation studies. In a last part, we will give an example of high resolution geophysical characterisation of the subsurface, which can be obtained by detailed, therefore computationally more expensive, simulation of the measured signals.

2 Detailed Simulation of Root Water and Nutrient Uptake

Uptake of water and nutrients by plant roots in a soil volume are treated as sink terms in macroscale models that do not consider flow and transport processes to single roots. To express the relation between root water uptake, root presence and water and nutrient availability in a soil volume, empirical relations between root length density, water potential and nutrient concentration on the one hand and water and nutrient uptake on the other hand are used. However, a large uncertainty about the functional forms and parameterisations of these relations exists. This leads to large uncertainties in the prediction of water transfer between the land surface and the atmosphere, in the prediction of nutrient uptake by plants, and in the prediction of leaching of nutrients and agrochemicals through the root zone. In order to improve the predictions of these processes, a better mechanistic description of water and nutrient uptake is required. Therefore, we developed a coupled soil-root model in which flow and transport processes in the soil towards to soil-root interface are modelled together with the water flow in the root system and the nutrient uptake by the root

tissue[1] In order to be able to simulation of large root systems, grid refinement methods were developed to adapt the simulation to the root architecture[2]. In an ongoing research project, the effect of root water uptake, the structure of the root system, and nutrient uptake mechanisms on solute transport is investigated. Fig. 1 shows simulated root systems and the simulated breakthrough of solutes that are not taken up by plants or are excluded. These breakthrough curves are used to parameterize simplified 1-D transport models and root nutrient uptake models. The results of these simulation studies demonstrate an interaction between water and nutrient uptake processes on the one hand and flow and transport processes on the other hand. The consequence of this interaction is that parameters of simplified models that parameterize one specific process, e.g. transport parameters, depend also on other processes that are simultaneously occurring such as nutrient uptake. The simulation studies also demonstrate that the nutrient uptake parameters that are used to describe nutrient uptake in the simplified model are different from the local scale nutrient uptake parameters. The consequence of this is that nutrient uptake parameters that are derived from experiments carried out in nutrient solutions should not be used to predict nutrient uptake in soils.

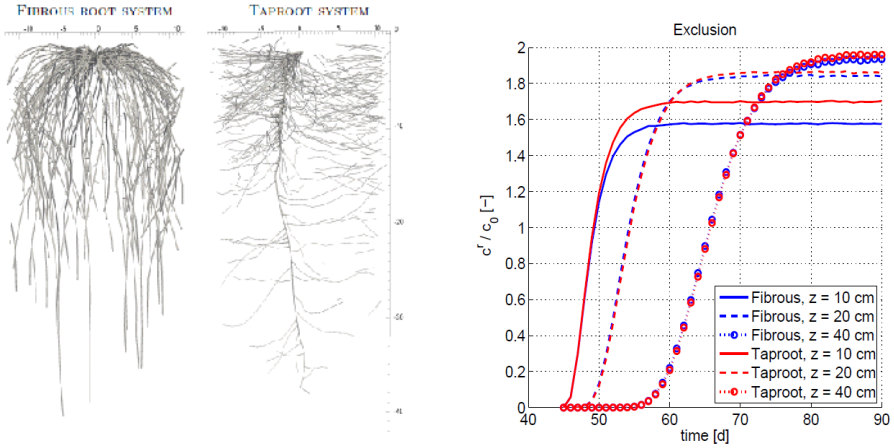


Figure 1. Simulated root structures: fibrous and taproot system (left), and simulated breakthrough of a solute that is not taken up or excluded by plants at different depths for the fibrous and taproot systems (right).

3 Simulation of the Effect of Multiscale Soil Heterogeneity on Flow Processes at the Soil Pedon Scale

In this set of simulations the effect of soil heterogeneity that may exists at different scales within a soil pedon is investigated. A soil pedon of 10x10x5 m soil block is considered. The soil structure within this block is represented by structures at two different scales. At the meter scale, soil horizons or layers with different soil hydraulic properties are considered. In the top soil, soil layers extend over the entire horizontal extent of the block whereas

in the bottom part of the profile, the horizontal scale of a layer is limited to the meter scale. An undulating boundary is considered between the top and bottom soil layer to investigate the effect of non-flat soil layer boundaries on lateral water fluxes and water redistribution. At the cm scale, stochastic variability that is represented using stochastic random fields is considered. In order to solve the unsaturated flow equation, i.e. the Richards equation, a cm-scale resolution was used. Fig. 2 shows simulated water potential and water contents in the 3-D soil block after 10 days of water infiltration. The simulation results illustrate the effect of small scale variability on soil water content and water fluxes (results not shown here) in the upper soil layers and the effect of the larger scale structure, i.e. the undulating boundary layer between top and sub soil, on pressure head and water content fields in the subsoil. These results indicate that in order to come to effective models and model parameters that do not resolve spatial variations at these scale, both small and large scale variability in hydraulic properties must be considered. Grid cells with a surface area that is similar to the total surface area of the simulation block are used in larger scale simulation of hydrological processes at the catchment scale. The simulation results of this study can thus be used to develop parameterizations for the entire soil block in which the lateral variations of soil hydraulic properties, water contents, potentials and fluxes, are averaged out.

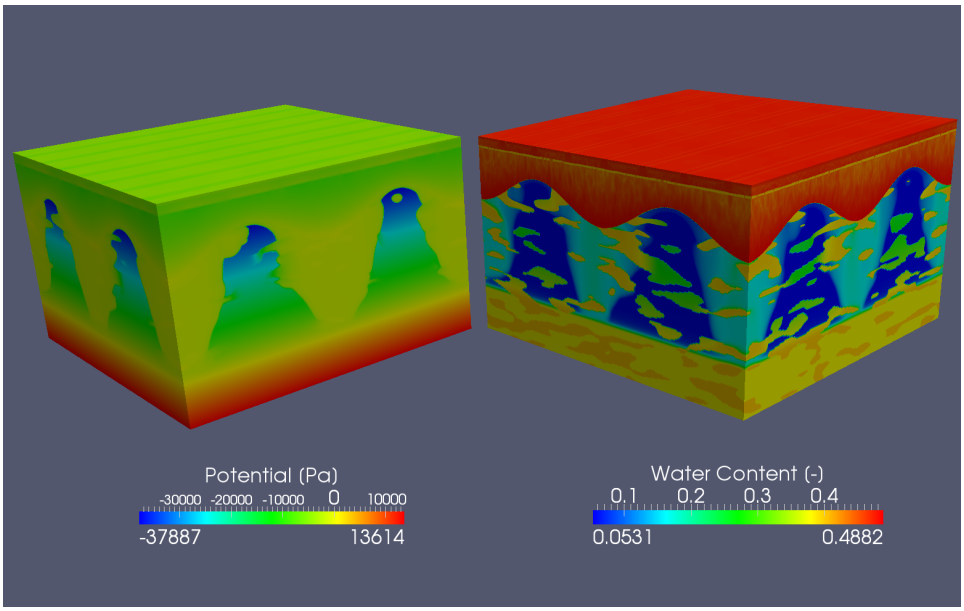


Figure 2. Simulated water potential fields (left) and water contents (right) after 10 days of water infiltration in the heterogeneous soil block.

In order to solve the Richards equation, the $\mu\phi$ code was used[3]. When using the parallel algebraic multigrid solver of DUNE-ISTL, a good scalability of the code on the JUGENE up to 287496 processes could be demonstrated[4].

4 Hydrodynamics at the Catchment Scale

Simulations at the hydrological catchment scale are used to investigate relations between topography, vegetation, land use and climate on the one hand and water (river discharge and groundwater recharge) and energy fluxes on the other and possible feedbacks between them. The outcome of such simulations provides insight in the effect of land-use and climate change on the hydrological fluxes and the hydrodynamic equilibrium in a region. In order to reproduce these interactions, flow processes should be resolved at a scale that represents the spatial structure of the land-surface (vegetation and topography) and the spatial structure of the subsurface (soil types and geology). Therefore high-resolution simulations at the catchment scale are required to resolve the variances in the different state variables and fluxes across a number of different space and time scales. These simulations are feasible only in the new generation supercomputing environment provided by the Jülich Supercomputing Centre, which has been shown in a previous scaling study and expanded by Ref. 5.

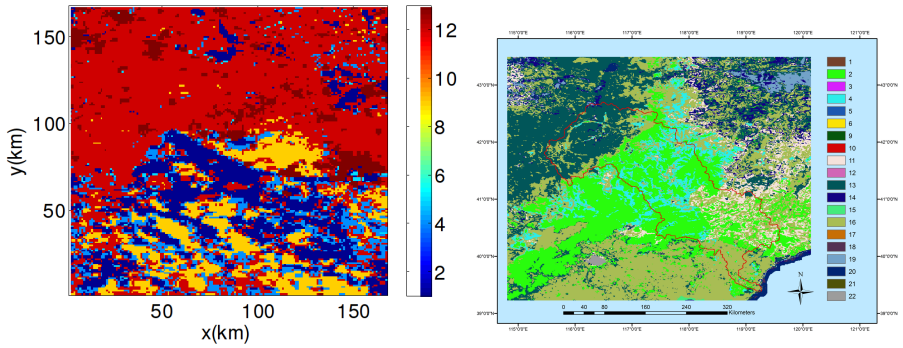
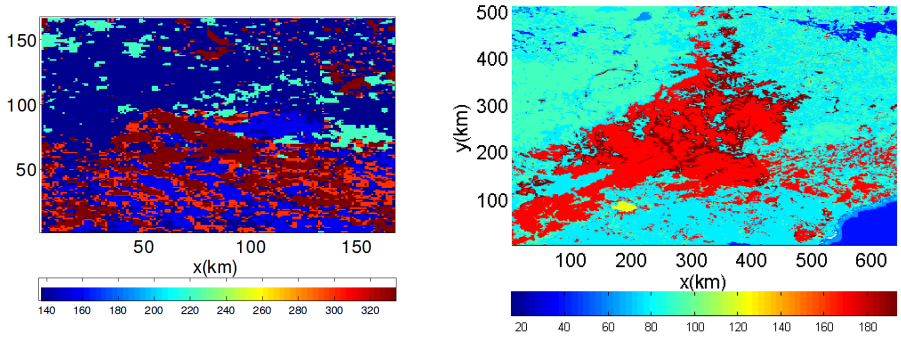


Figure 3. Vegetation cover in the Rur (left, colour code: 2 = evergreen needle leaf forest; 4 = evergreen broad leaf forest; 6 = deciduous needle leaf forest; 8 = deciduous broad leaf forest; 10 = mixed forest; 12 = closed shrub lands)) and Luahne catchment (right, colour code: 1=broadleaved, evergreen tree cover; 2=broadleaved, deciduous, closed tree cover; 3=broadleaved, deciduous, open tree cover; 4=needle-leaved, evergreen tree cover; 5=needle-leaved, deciduous tree cover; 6=mixed leaf type tree cover; 9=other natural vegetation; 10=burnt tree cover; 11=closed-open, evergreen shrub cover; 12=closed-open, deciduous shrub cover;13=closed-open herbaceous cover; 14=sparse shrub cover; 15=regularly flooded shrub and/or herbaceous cover; 16=cultivated/managed areas; 17=cropland; 18=cropland/shrub/grass cover; 19=bare areas; 20=water; 21=snow/ice; 22 = artificial surfaces).

Groundwater flow and transport processes, land surface-atmosphere interactions and surface flow are simulated using the PARFLOW/CLM simulation platform[6]. As an example to demonstrate the capabilities of the simulation platform, simulation results are shown for two different catchments: the Rur catchment in Germany and the Luanhe catchment in China. The model domains considered for the Rur and Luahne catchments were, respectively, 168km x 168km and 640km x 512km and in both cases, a spatial resolution of 1x1km in the horizontal direction was used. This corresponded to problems with 107 to 108 unknowns that had to be solved. Fig. 3 and Fig. 4 show respectively the vegetation cover and the simulated latent heat fluxes in the two model domains. These figures illus-



(a) Rur

(b) right Luahne

Figure 4. Simulated latent heat fluxes (Wm^{-2}) in the regions.

trate the strong correlation that exists between vegetation cover and latent heat flux which implies a strong correlation between land-use and hydrological fluxes. These simulations were carried out using a spatially uniform atmospheric forcing. In future simulations, the effect of spatially varying atmospheric forcing and the spatial coupling between atmospheric conditions and land surface properties (vegetation, topography, groundwater depth and soil wetness) will be included.

5 Improved Characterization of Stream-Aquifer Interactions, Streambed Properties and Aquifer Properties by Assimilation of Thermal Data

Groundwater models are an important tool for predicting the availability and quality of groundwater resources. A proper prediction of groundwater states with models relies on an adequate characterization of aquifer properties. Unfortunately, these aquifer properties show a strong spatial variability in nature and measurements are usually scarce due to the high financial demand associated with hydrogeological exploration techniques. As a consequence, predictions made with groundwater models are often associated with a high degree of uncertainty which is caused by the lack of knowledge on aquifer properties. Different studies[7,8] have shown that sequential data assimilation techniques like the Ensemble Kalman Filter[9] can be used to improve the predictive capability of groundwater models. This is done by using measured hydraulic heads of the groundwater to update model states and parameters in a Monte Carlo framework. Furthermore, the transport of heat and solutes within the aquifer is strongly influenced by its hydraulic conductivities and thus the spatial distribution and temporal evolution of temperature and solute concentration data carry important information about aquifer properties. For heat transport the highest temperature contrasts usually occur at the transition zone to surface water (e.g. rivers, artificial recharge) because groundwater temperatures are rather constant whereas the temperature of surface water usually follows a seasonal cycle. Due to the exchange fluxes between these two compartments a temperature signal propagates into the aquifer which can be used for a better characterisation of aquifer properties[10]. The general aim of our study is to

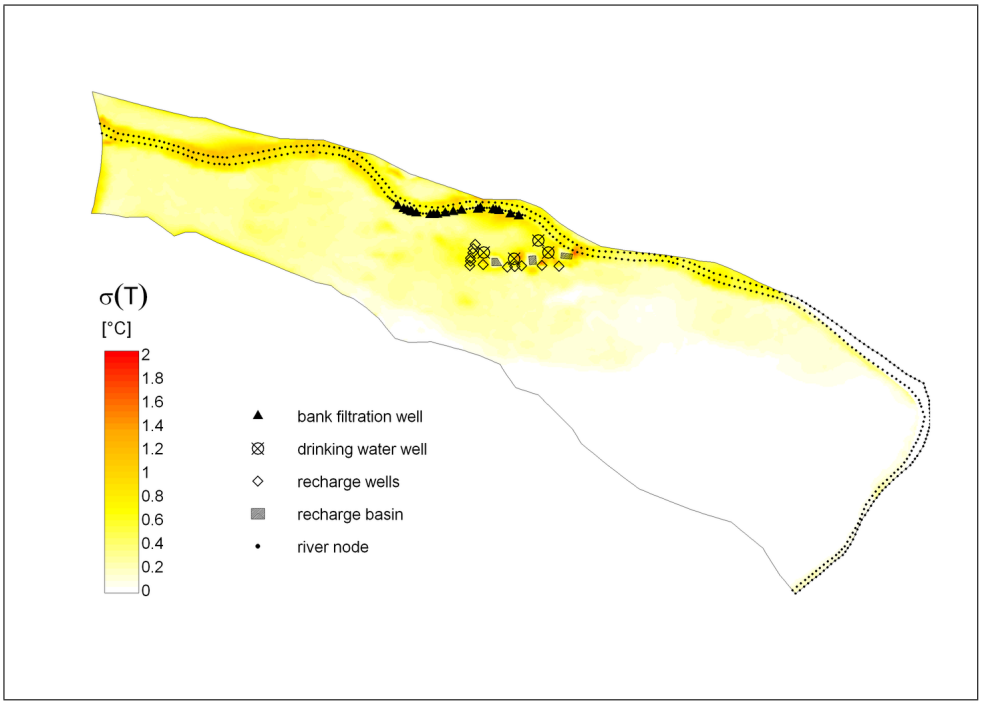


Figure 5. Spatial distribution of temperature standard deviations after 31 days (model layer 3) calculated from an ensemble 128 realisations of different hydraulic conductivity fields of the aquifer and river bed conductivities.

use data assimilation techniques to exploit the information given in temperature data close to streams to improve the characterisation of aquifer and stream bed properties.

For that purpose, we performed coupled flow and heat transport simulations with a 3D model of the Limmat aquifer in Zurich. Within the study area groundwater is strongly influenced by river-aquifer exchange fluxes and management activities (bank filtration, artificial recharge) both inducing temperature contrasts in the aquifer. Measurements of hydraulic heads (groundwater temperatures) for 87 (9) observation points as well as forcing data for the model were available on a daily basis.

In a first step we performed unconditional ensemble simulations for 128 different realisations of hydraulic conductivity fields of the aquifer and river bed conductivities for a period of 31 days. The spatial distribution of the variability of temperature data (expressed as standard deviation σ(T)) within the ensemble at the end of this period is shown in Fig. 5. It can be seen that the highest variability occurs close to the river and at the places where most of the groundwater management (pumping, artificial recharge) is performed. This variability of temperatures is caused by the variability of groundwater velocity and river-aquifer exchange fluxes which in turn are caused by the uncertainty within the ensemble regarding the hydraulic conductivities of the aquifer and the river bed. The generally higher variability in the western part of the model domain compared to the one in the eastern part is associated with the fact that the groundwater table is rather close to the surface in the western part what leads to a more intensive propagation of the temperature signal of pre-

cipitation into the aquifer. In a further step we performed simulations in which we used the measured hydraulic head and temperature data to update only model states (hydraulic heads and temperatures) with the Ensemble Kalman Filter (EnKF). Results indicate that in the first time steps hydraulic head and temperature predictions can be improved with EnKF. Afterwards simulated and measured temperatures show rather constant discrepancies what is probably caused by the biased parameterization of aquifer properties within the ensemble. In a next step we will perform ensemble simulations in which also the model parameters are updated with temperature data with EnKF in order to improve predictions of groundwater temperatures as well as the model parameters.

6 Characterisation of Subsurface Structure Using Full-Waveform Inversion of Crosshole Ground Penetrating Radar Data

Ground penetrating radar (GPR) is an important tool for a wide range of geological, hydrogeological and engineering investigations. Conventional processing techniques mainly use ray-based methods that may be either inappropriate or only provide limited resolution. Higher resolution images can be derived by using more sophisticated processing techniques that explicitly take into account the electromagnetic wave propagation characteristics. One of the most promising developments that return an unprecedented resolution is full-waveform inversion, which uses the full information content present in the measured traces. Recent publications[11,12] show the high potential of this method and return very similar values for the medium properties obtained by independent borehole measurements. Recently, borehole GPR measurements were carried out in Krauthausen and inverted on the JUROPA supercomputer using the full-waveform. Fig. 6 shows high resolution permittivity and conductivity images that are strongly correlated with the porosity present within the aquifer. In addition, a new checkerboard analysis was introduced that indicate the well-resolved regions of the inversion results (grey line in Fig. 6).

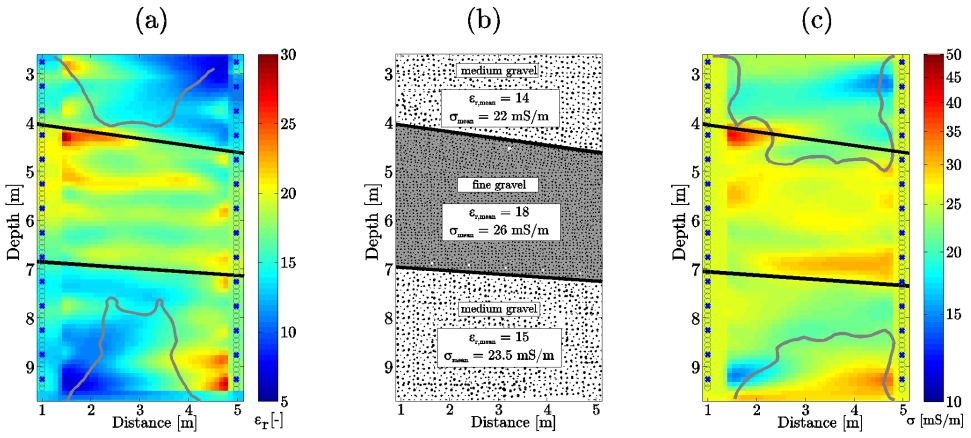


Figure 6. (a) Permittivity and (c) conductivity results obtained from GPR full-waveform inversion. The Crosses and circles on the left and right indicate source and receiver locations. (b) generalized stratigraphy that corresponds well with independent borehole information.

Acknowledgements

The calculations were carried out at the Jülich Supercomputing Centre.

References

1. Javaux, M., et al, *Use of a three-dimensional detailed modeling approach for predicting root water uptake*. Vadose Zone J., 2008. 7(3): p.1079-1088. DOI: 10.2136/vzj2007.0115.
2. Schröder, T., et al., *A grid refinement approach for a three-dimensional soil-root water transfer model*. Water Resour. Res., 2009. 45: p. W10412. DOI: 10.1029/2009wr007873.
3. Ippisch, O., H.J. Vogel, and P. Bastian, *Validity limits for the van Genuchten-Mualem model and implications for parameter estimation and numerical simulation*. Adv. Water Resour., 2006. 29(12): p. 1780-1789. DOI: 10.1016/j.advwatres.2005.12.011.
4. Ippisch, O. and M. Blatt, *Scalability Test of MuPhi and the Parallel Algebraic Multigrid solver of DUNE-ISTL, in Jülich Blue Gene/P Extreme Scaling Workshop 2011*, B. Mohr and W. Frings, Editors. 2011, Jülich Supercomputing Centre: Jülich. p. 21-26.
5. Kollet, S.J., et al., *Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources*. Water Resour. Res., 2010. 46: p. W04201. DOI: 10.1029/2009wr008730.
6. Kollet, S.J. and R.M. Maxwell, *Demonstrating fractal scaling of baseflow residence time distributions using a fully-coupled groundwater and land surface model*. Geophys. Res. Lett., 2008. 35(7): p. L07402. DOI: 10.1029/2008gl033215.
7. Chen, Y. and D.X. Zhang, *Data assimilation for transient flow in geologic formations via ensemble Kalman filter*. Adv. Water Resour., 2006. 29(8): p. 1107-1122. DOI: 10.1016/j.advwatres.2005.09.007.
8. Hendricks-Franssen, H.J. and W. Kinzelbach, *Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem*. Water Resour. Res., 2008. 44(9): p. W09408. DOI: 10.1029/2007wr006505.
9. Evensen, G., *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics*. J. Geophys. Res.-Oceans, 1994. 99(C5): p. 10143-10162. DOI: 10.1029/94jc00572.
10. Constantz, J., *Heat as a tracer to determine streambed water exchanges*. Water Resour. Res., 2008. 44: p. W00d10. DOI: 10.1029/2008wr006996.
11. Klotzsche, A., et al., *Full-waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland*. Near Surf. Geophys., 2010. 8(6): p. 635-649. DOI: 10.3997/1873-0604.2010054.
12. Meles, G.A., et al., *A New Vector Waveform Inversion Algorithm for Simultaneous Updating of Conductivity and Permittivity Parameters From Combination Crosshole/Borehole-to-Surface GPR Data*. IEEE Trans. Geosci. Remote Sensing, 2010. 48(9): p. 3391-3407. DOI: 10.1109/tgrs.2010.2046670.

Modelling of Atmospheric Processes

Marcel König, Ralf Wolke, Oswald Knoth, and Eberhard Renner

Leibniz Institute for Tropospheric Research
Permoserstrasse 15, 04318 Leipzig, Germany
E-mail: {marcelk, wolke, knoth, renner}@tropos.de

Modelling of atmospheric trace gases and particulate matter is of major importance for air quality studies as well as climate considerations. In this context, the modelling activities are aimed to problems in a widespread area of atmospheric research. They range from health effects to the impact of long-range transported smoke of forest fires on regional climate. Two selected applications from different objects of research are presented in the paper. The first study enables a detailed quantification of flow measurement in an urban like area and its numerical Simulation with LES. Because of an evaluating point, we try to hold our simulations right near the meteorological data set generated by these experimental cases of MUST (Mock Urban Setting Test).

In the second project, the contribution of secondarily formed inorganic particulate matter on the fine dust load in Central Europe is quantified. Both applications require a detailed process description as well as a high spatial resolution. Such simulations are very expensive in terms of computing time and demand the use of powerful parallel computers.

1 Introduction

The investigation of the interactions between aerosols, gases and clouds is very important for the understanding of tropospheric processes. The size distribution, number concentration and chemical composition of aerosol particles decisively influence microphysics of clouds (e.g., number and size of drops, precipitation formation). This has not only an impact on the operational weather forecast. By the change of the radiation properties this feeds also back on the atmospheric dynamics again. Therefore, it has an essential impact on climate changes. Moreover, investigations of the formation, propagation and modification of particles have great environmental importance as shown by the present fine dust discussion. Microphysical and dynamical processes proceed simultaneously, but very heterogeneous in space and time. For examining these processes and effects, the modelling department of the IfT has developed the multiscale model system COSMO-MUSCAT (Wolke et al.[1,2]). It can be applied to process studies as well as the operational forecast of pollutants in local and regional areas (Heinold et al.[3,4], Hinneburg et al.[5], Renner and Wolke[6]). The studies are performed from European to urban scale.

The large eddy simulation (LES) technique has been shown by many recent studies to yield results superior to the RANS standard, but at the cost of significant computational demands (CPU time, memory and storage). LES is becoming a more widely used simulation tool for the time-accurate prediction of unsteady flows at high Reynolds numbers. The effects of turbulent flow around buildings are usually studied with RANS models based on the Reynolds averaged Navier-Stokes equations. They do not explicitly resolve turbulence and yield only average flow information. Since they cannot produce instantaneous, time-dependent flow characteristics, one important limit is that peak values can only be estimated more or less accurately but not be predicted directly.

In the last years, more and more focus is directed on modelling of microscale effects in urban areas, which are very important for the assessment of possible health damages caused by air pollution. In recent years, considerable attention has focused on the development of models to predict the dispersion of toxic materials in the urban environment (Gifford and Hanna[7]). The systematic investigations initiated by Baechlin et al.[8] are notable and have contributed significantly to our current understanding of the behaviour of a passive neutrally buoyant plume passing through a large array of obstacles. The urban impact on meteorological flow fields is very important for the accuracy of flow simulation models. Otherwise the prediction of contaminant transport at urban scales has received a large amount of attention during the last decade.

2 Microscale Modelling of Urban Areas

2.1 The Model ASAM

The All Scale Atmospheric Model ASAM was developed at the IfT Leipzig to simulate flow fields in the atmosphere (Knoth and Hinneburg[9]). This code solves applications with scales ranging from the globe up to building resolved simulations to get the urban impact on the flow (ASAMWiki[10]). This flow simulation code works on a Cartesian Grid, with a Cut-Cell approach near buildings (Knoth and Hinneburg[9]). This fully parallelized code is a developing research code, which contains a lot of different chooseable options, for example numerical methods, physical processes and numbers of variables.

For a realistic description of the urban impact on the flow field in urban areas, we focus on SGS models, their implementation and problems, which appear near solid walls. Otherwise the additional produced turbulent energy, the development of wake diffusion, drag forces on building walls and changes in the radiation budget through sunlit and shaded surfaces in resolved street canyons are being developed for the code.

The atmospheric flow model ASAM use the compressible Euler equation in conservative form.

$$\begin{aligned}
\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) &= 0 \\
\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla(\rho \mathbf{v} \cdot \mathbf{v}) &= -\nabla p - \rho \mathbf{g} - 2\Omega \times (\rho \mathbf{v}) \\
\frac{\partial \rho \theta}{\partial t} + \nabla(\rho \mathbf{v} \theta) &= Q_\theta \\
p &= \rho R \theta (p/p_0)^\kappa \quad \text{or} \\
p &= \left(\frac{R \rho \theta}{p_0^\kappa} \right)^{1/(1-\kappa)}
\end{aligned} \tag{1}$$

The orography is incorporated in the model through a special grid system, where the orography is represented by cut cells in a Cartesian grid. The time integration is accomplished by a linear implicit method of Rosenbrock type. Because the method is fully implicit, the approach is able to employ time steps that result in Courant-Friedrichs-Lewy (CFL) numbers greater than one for advection, gravity, and sound waves; however, the dynamical time scale of the problem will be respected for accuracy by a dynamic time step procedure.

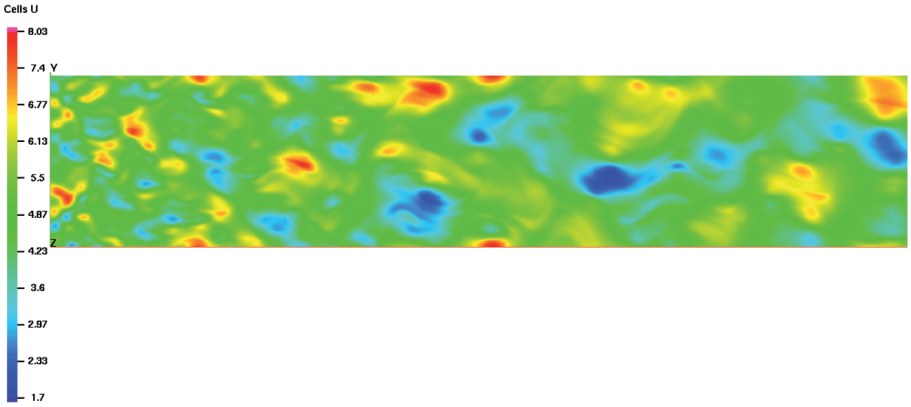


Figure 1. Screen shot of 2D-u-velocity component after 500 s.

Spatially inhomogeneous turbulent flows calculated with Large Eddy Simulation Code require in the majority turbulent inflow boundary conditions. In practice you can generate synthetic inflow data from statistical properties of flow which is known from experimental data or empirical correlation. We also investigate the applicability of Taylor's hypothesis in space-time correlation of fluctuations in velocity and vorticity.

In a first stage of our project, we test our inflow conditions and the resulting turbulent wind-field in an obstacle free space (see Fig.1). We used a method of generating inflow turbulence of Batten et al.[11] to provide a three-dimensional, unsteady velocity field at the inflow boundary layer. The velocity field is created using the sum of sines and cosines with random phase and amplitudes. We suppose at the inlet boundary layer ($x_1 = 0$) a velocity component in x_1 direction as follow:

$$v_1 = \bar{v}_1 + \sum_{n=1}^{N} E_1^n \cdot sin[k_1^n \hat{x}_2 + \omega_1^n t + \Delta\phi_1^n] \tag{2}$$

where

$$\hat{x}_2 = 2\pi x_2 / L_{Inlet}$$

$\hat{x}_2$ is a spatial coordinate normalized by the length-scale of turbulence, specially the length of inflow boundary. k_1^n and w_1^n are wave numbers and frequencies in x_1 direction. To get different turbulent eddies at the Inlet Boundary Layer, we use a sum over n, where n gives the number of different eddy sizes. For the whole spectrum of turbulence the wave number must range from the upper edge to the lower edge. These borders must be fit to the geometry and resolution of the simulation.

The amplitude E_1^n gives for every single mode n the value of energy in this special mode. You get this amplitude from a given energy spectrum with their associated fre-

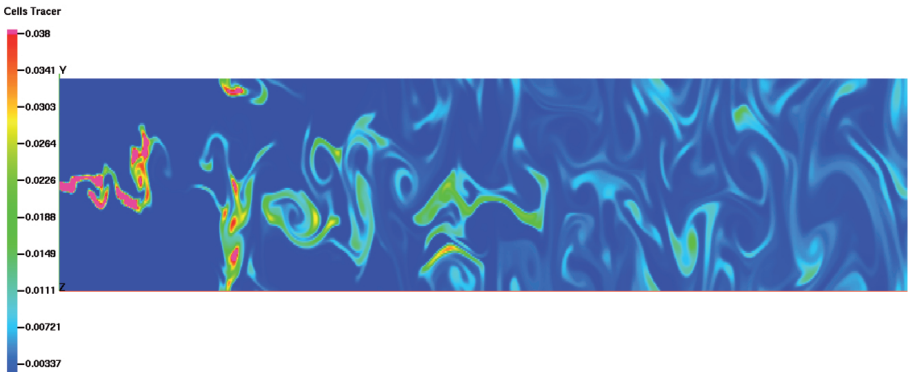


Figure 2. Screen shot of a Tracer gas plume which is released at the middle of the inlet Boundary.

quency. To describe this $E_1^n(\omega_1^n)$ amplitude from the target spectrum, Lee et al.[12] investigate

$$E(\omega) = \omega^4 exp[-2(\frac{\omega}{\omega_0})^2] \tag{3}$$

where ω_0 is the frequency with peak energy. In Fig. 2 we show our synthetic generated turbulent flow field symbolized as concentration field of a release point at the middle of the inlet boundary. After a short adjustment zone, we get a well mixed tracer gas flow field.

2.2 Results of Model Simulations

To realize the effect of different buildings to nearby buildings, it is important to resolve most of eddies and vortexes in the down wind zone of buildings. A more scientific understanding of the physics behind the interaction of different buildings represents our main goal. Additional activities are aimed at the behaviour of eddies in the flow field near solid walls. To achieve these goals we have to validate our code with measurements.

MUST (Mock Urban Setting Test) Experiment. This experiment is carried out in the great basin desert (Utah/USA) to investigate dispersion over an array composed by 12 by 10 Containers (Yee and Biltoft[13]). This configuration was chosen to represent an urban environment. They use these containers to realize the urban geometry of an urban like city (see Fig. 3). In order to have more controlled variables and conditions in the desert, we use this nearly urban-scale geometry to validate our model because of the existence of a considerably meteorological data set.

For our simulation of the MUST Experiment we use the method of generating Inflow Turbulence described above (see Sec. 2). The case, we choose, is a stable condition night time period. In a height of 10 m, we have a wind speed of about 4 m/s in front of the container domain. In Fig. 4 the wind speed at half the container height is shown. There are a lot of inhomogeneities in the streets. Based on this assumption, it is difficult to predict Tracer gas plumes in real cities. In our urban like area the plume spreads although nearly stable Boundary condition with only small turbulence at the Inflow Boundary were

Figure 3. Container array of MUST Experiment.

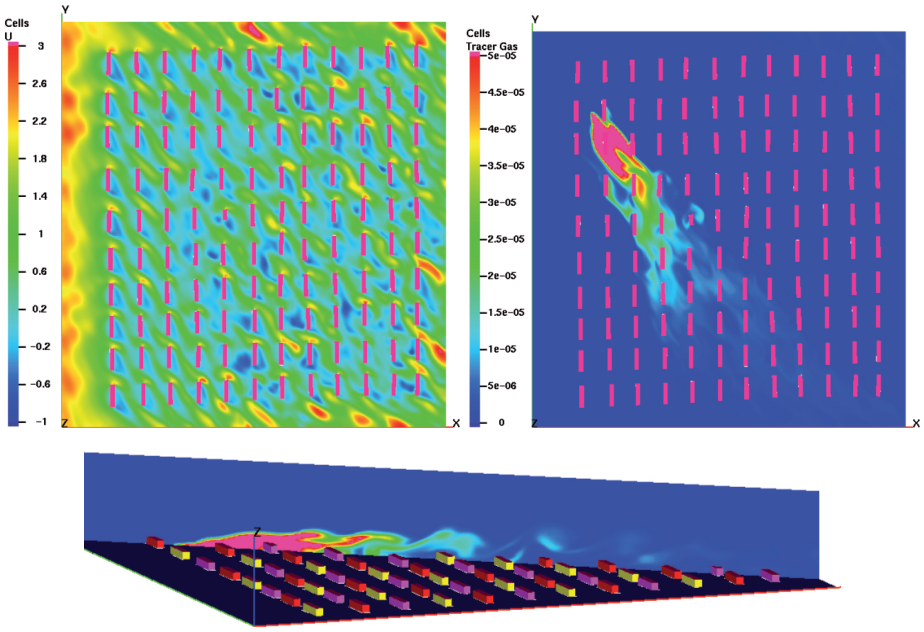


Figure 4. Top left: u - velocity component; Top right; Tracer gas plume at 1m above ground in the nightly Boundary Layer during the container array. Bottom; Vertical plane in wind flow direction.

established. In Fig. 4 the vertical dimension of the plume is coloured. Although there is a stable stratification, the tracer gas is pulled up with vertical winds in front of several containers.

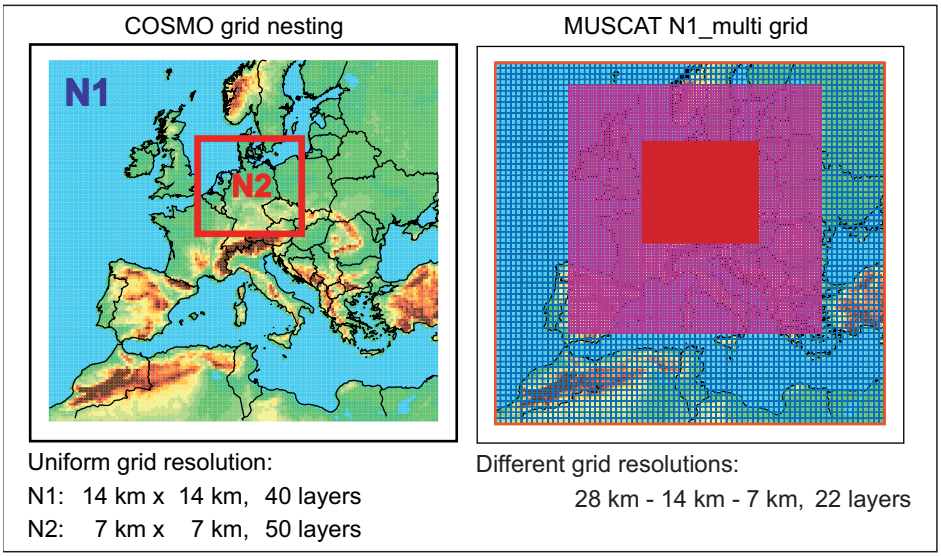


Figure 5. Nesting of COSMO and multiscale MUSCAT grid for the N1 domain.

3 Investigation of the Particle Composition in Central Europe

The investigations are focused on quantifying the large scale and local influence on aerosol concentrations and its composition in Germany by using COSMO-MUSCAT simulations. Several evaluation studies and investigations about urban-regional scale interactions were performed in this context. A one-way nesting strategy for the COSMO-MUSCAT grids is used (Fig. 5). In each of the regions N1 and N2, the meteorological code COSMO runs on a uniform grid. The online-coupled MUSCAT uses different resolutions (coarse near the boundaries and finer in the inner regions). This corresponds to a two-way nesting for the chemistry-transport part. The meteorological forcing of COSMO is performed by reanalysed data provided by the global meteorological model GME. Biogenic emissions and radiation activity are calculated by MUSCAT, whereas information on the cloud cover, temperature, and other meteorological parameters are taken from the coupled meteorological model.

Exemplarily for an analysis of the fraction of secondary particulate matter in 2005, the annual mean concentration of particulate matter (PM10 and PM2.5), ammonium nitrate and ammonium sulphate are shown in (Fig. 6). The averaged PM concentrations reach the maximum in the Ruhr and the Rhine-Main area around Frankfurt. These elevated concentrations are mainly caused by local emissions (traffic, heating). Ammonium sulphate has a similar patter as PM. Additionally, increased concentrations are found in Bavaria, which may be caused by higher biogenic emissions. The ammonium sulphate concentrations are

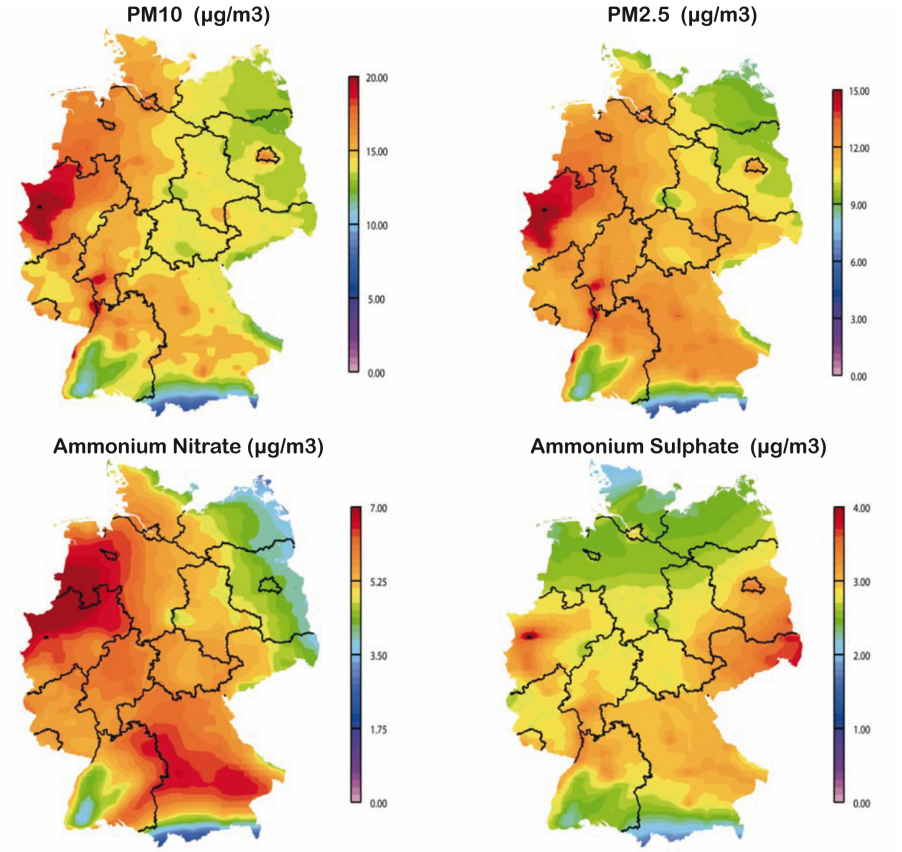


Figure 6. Annul mean concentrations of PM10 (mass of particles $< 10\ \mu m$), PM2.5 (mass of particles $< 2.5\ \mu m$), ammonium nitrate and ammonium sulphate for Germany in 2005..

very high in south-east Germany and have a strong gradient from east to west. This seems to be mainly generated by long-range transport of SO2 from Czech Republic.

4 Concluding Remarks

The LES simulations of urban areas are just at the beginning, so there are only poor statistical results available at the moment. But it is a computational challenge to to get closer to the measured data. Therefore we we want to continue our work also in collaboration with supercomputing infrastructure.

Acknowledgements

We thank the John von Neumann Institute for Computing (NIC) for supporting this work. Furthermore, we thank the Federal Environment Agency of Germany for providing monitoring station data and the DWD Offenbach for good cooperation.

References

1. R. Wolke, O. Knoth, O. Hellmuth, W. Schrder, and E. Renner, *The parallel model system LM-MUSCAT for chemistry-transport simulations: Coupling scheme, parallelization and applications*, vol. Volume 13, pp. 363–369, North-Holland, 2004.
2. Jrgen Helmert Detlef Hinneburg Matthias Lieber Eberhard Renner Wolfram Schrder Ina Tegen G. Mnster D. Wolf M. Kremer Ralf Wolke, Bernd Heinold, *Modelling of Atmospheric Chemistry Transport Processes*, NIC Symposium 2006, 2006.
3. B. Heinold, I. Tegen, K. Schepanski, M. Tesche, M. Esselborn, V. Freudenthaler, S. Gross, K. Kandler, P. Knippertz, D. Muller, A. Schladitz, C. Toledano, B. Weinzierl, A. Ansmann, D. Althausen, T. Muller, A. Petzold, and A. Wiedensohler, *Regional modelling of Saharan dust and biomass-burning smoke Part I: Model description and evaluation*, Tellus Series B-Chemical and Physical Meteorology, **63**, no. 4, 781–799, 2011, Sp. Iss. SI 811NA Times Cited:1 Cited References Count:63.
4. B. Heinold, I. Tegen, S. Bauer, and M. Wendisch, *Regional modelling of Saharan dust and biomass-burning smoke Part 2: Direct radiative forcing and atmospheric dynamic response*, Tellus Series B-Chemical and Physical Meteorology, **63**, no. 4, 800–813, 2011, Sp. Iss. SI 811NA Times Cited:1 Cited References Count:45.
5. E. Renner, D. Hinneburg, and R. Wolke, *Formation of secondary inorganic aerosols by power plant emissions exhausted through cooling towers in Saxony*, Environmental Science and Pollution Research, **16**, no. 1, 25–35, 2009, 392OV Times Cited:3 Cited References Count:43.
6. E. Renner and R. Wolke, *Modelling the formation and atmospheric transport of secondary inorganic aerosols with special attention to regions with high ammonia emissions*, Atmospheric Environment, **44**, no. 15, 1904–1912, 2010, 594LR Times Cited:5 Cited References Count:31.
7. F. A. Gifford and S. R. Hanna, *Modeling Urban Air-Pollution*, Atmospheric Environment, **7**, no. 1, 131–136, 1973, O4482 Times Cited:70 Cited References Count:21.
8. W. Baechlin, W. Theurer, and E. J. Plate, *Dispersion of Gases Released near the Ground in Built up Areas - Experimental Results Compared to Simple Numerical Modeling*, Journal of Wind Engineering and Industrial Aerodynamics, **44**, no. 1-3, 2721–2732, 1992, Ka516 Times Cited:0 Cited References Count:12.
9. D. Hinneburg and O. Knoth, *Non-dissipative cloud transport in Eulerian grid models by the volume-of-fluid (VOF) method*, Atmospheric Environment, **39**, no. 23-24, 4321–4330, 2005, 958AB Times Cited:2 Cited References Count:22.
10. Oswald Knoth, "Asam-wiki", 2010.
11. Batten, Paul, Goldberg, Uriel, Chakravarthy, and Sukumar, *Interfacing statistical turbulence closures with large-eddy simulation : Boundary conditions for large eddy simulation*, vol. 42, American Institute of Aeronautics and Astronautics, Reston, VA, ETATS-UNIS, 2004, 3 AIAA journal.
12. Sangsan Lee, Sanjiva K. Lele, and Parviz Moin, *Simulation of spatially evolving turbulence and the applicability of Taylor's hypothesis in compressible flow*, Physics of Fluids A: Fluid Dynamics, **4**, no. 7, 1521–1530, 1992.
13. E. Yee and C. A. Biltoft, *Concentration fluctuation measurements in a plume dispersing through a regular array of obstacles*, Boundary-Layer Meteorology, **111**, no. 3, 363–415, 2004, 774ZA Times Cited:55 Cited References Count:37.

Computer Science and Numerical Mathematics

Michael Griebel

Institute for Numerical Simulation, University of Bonn, Wegelerstr. 6, 53115 Bonn, Germany
E-mail: griebel@ins.uni-bonn.de

Computer science aspects on modern large scale parallel hardware platforms are discussed in the following section. Here we put the emphasis on parallel multigrid algorithms for the solution of partial differential equations using hierarchical hybrid grids and on a problem from statistical machine translation, where corpora, e.g. equal sentences from two languages, are extracted from Wikipedia data in English or German.

In their contribution *Highly Parallel Geometric Multigrid Algorithm for Hierarchical Hybrid Grids*, Björn Gmeiner, Tobias Gradl, Harald Köstler and Ulrich Rüde deal with efficient multigrid algorithms, which, due to their reduced number of levels, promise advantages in parallelization on, e.g. the Blue Gene/P computer. To this end, the underlying problem is decomposed into a large number of subproblems which can be treated in parallel and a certain part which cannot efficiently be dealt with in parallel. The authors present their hierarchical hybrid grid framework which allows to discretize and solve elliptic partial differential equations using a well-designed mixture of structured and unstructured grids by geometric multigrid methods. This approach is highly parallel and runs on the JUGENE (Blue Gene/P) system. Strong and weak scaling results are presented with up to 292.912 cores, which demonstrate the efficiency of this method.

In their contribution *Comparable Corpora in Wikipedia Text for Machine Translation*, Jia Xu, Casey Kennington, Česlav Przywara and Lilian Wanzare describe an approach to derive corpora for machine translation which is based on the maximum entropy technique. To this end, baseline scores for a number of language pairs are first established using the Experience Management System of Moses SMT. The authors make use of the JUROPA cluster to connect the necessary programs for training and testing and to organize all of the corpora. Then they use a previously processed set of Wikipedia data in English and German as comparison to their baseline score for the German-English language pair. This way, the new BLEU score was slightly increased to the value of 19.40.

Highly Parallel Geometric Multigrid Algorithm for Hierarchical Hybrid Grids

Björn Gmeiner, Tobias Gradl, Harald Köstler, and Ulrich Rüde

Chair for System Simulation, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
91058 Erlangen, Germany
E-mail: {bjoern.gmeiner, tobias.gradl, harald.koestler, ruede}@informatik.uni-erlangen.de

Current supercomputers are approaching a million cores and their compute power and amount of memory enable us to solve linear systems with more than 10^{12} degrees of freedom. However, this forces us to partition our problem into a large number of sub-problems that can be treated in parallel. Parts of the algorithm that do not permit such high degrees of parallelism thus easily become a bottleneck. Additionally, the performance analysis and debugging of programs for such a high number of cores become challenging tasks in themselves. Our Hierarchical Hybrid Grids framework is capable of solving elliptic partial differential equations discretized with finite elements on a compromise between structured and unstructured grids by a geometric multigrid method. It is designed to run highly parallel and was adapted within this project to run on the JUGENE located in Jülich. We present scaling results and discuss the specifics of an efficient implementation of our software on Blue Gene/P systems.

1 Introduction

A variety of applications have a high demand on algorithms which are able to provide very high spatial resolutions finite element (FE). Large-scale example include seismic sea wave and earthquake simulations or weather predictions. In much smaller scales, direct numerical simulations provide insight to highly turbulent phenomena in fluid mechanics. Here, it is only possible to calculate a domain of some cubic centimeters at very high Reynold numbers. But there are also interesting issues at other scales, e.g. in acoustics: An important design goal for concert halls is to aquire excellent room acoustics. A direct approach to capture the acoustics in such buildings is to solve the pressure equation. If an average concert hall is simulated at a reasonable resolution, we already can fill the complete memory of today's largest supercomputers.

A major aim of our report is to show that it is still possible to design a relatively flat, but still efficient multigrid (MG) algorithm on current highly parallel supercomputers. Flat means in this context, that we have a multigrid algorithm with up to seven grid levels and thus a relatively large number of unknowns on the coarsest grid.

Our test machine is a Blue Gene/P cluster located in Jülich, which has 73 728 nodes. Each node is equipped with four compute cores. Besides other software approaches e.g.[1–3], our framework Hierarchical Hybrid Grids (HHG) implements, a highly parallel, multigrid variant for such machines.

We present the HHG approach on Blue Gene/P, which is different in terms of its architecture and degree of parallelism. The next section introduces the problems with the coarsest grids and ways to treat it. Furthermore, some parallel issues are discussed, which arose in our implementations.

Section three presents strong and weak scaling results. In the strong scaling, we try to push our MG to its limits on the underlying computer architecture. We achieved a

speedup of 50 for a scaling from N to 96*N cores. The weak scaling shows a good parallel efficiency up to 292 912 cores, while the overhead for calculations on the coarsest grid stays reasonable.

1.1 Multigrid Algorithm

Discretizing a second-order differential equation with finite elements leads to a system of equations. An important property of the system matrix is sparsity. The sparsity pattern reduces the number of operations *per iteration* over the domain for iterative solvers to a complexity of O(N) for N unknowns. However, an increasing number of iterations are required for solvers like Gauss-Seidel (GS) or Conjugated Gradient (CG) with growing problem size. MG is a strategy to change this behaviour, such that the number of operations to solve a system is linearly dependent on the number of unknowns. This allows us to solve large systems in reasonable time. In the following, we give a quick overview on the multigrid method. For an introduction, we refer the reader to Ref. 4.

An observation for local acting iterative smoothers, like GS or Jacobi solvers, is that they smooth high frequent errors very well. In contrast, low frequent errors are reduced extremely slow. In order to resolve this shortcoming, MG uses a second effect: Let us consider a function on a fine grid, which has low frequencies. When we transfer this function to a coarser grid, the frequencies of the function increase with respect to the mesh size. Thus, we have changed the frequency of a function by changing its discretization. In MG this idea is applied to low frequencies of the error by transferring the error to coarser grids.

For the linear case, we can use the *error equation* (1) to calculate the error e_k on the coarser grids, where N_k is the discretization matrix, r_k the residual, f_k the right hand side, and v_k the approximated solution. The subindex k denotes the level of the grid. For a two-grid case k is the fine grid and $k-1$ is the coarse grid. Afterwards the solution is corrected by the error on the fine grid (see Fig. 1).

$$N_k e_k = r_k = f_k - N_k v_k \tag{1}$$

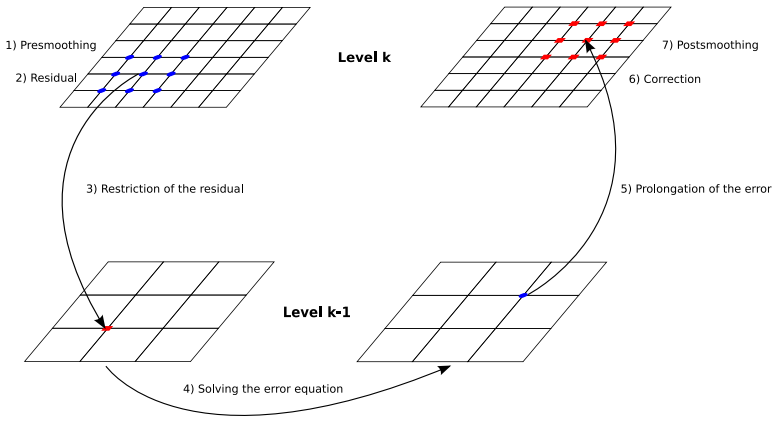


Figure 1. Multigrid two-grid cycle.

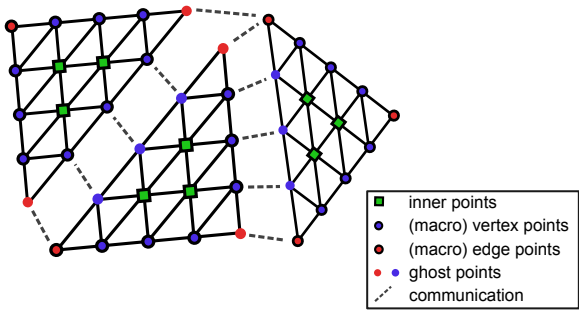


Figure 2. Space partitioning in HHG primitives.

1.2 Hierarchical Hybrid Grids

The HHG framework[5,6] uses a hybrid discretization strategy by combining structured and unstructured grids. A coarse input FE mesh is organized into the grid primitives vertices, edges, faces, and volumes. This grid is unstructured and thus provides geometric flexibility. The primitives are then refined in a structured way, resulting in semi-structured meshes (see Fig. 2). The generated structured regions are stored in a directly addressed way into the memory to allow an efficient execution. Moreover, the regularity of the resulting grids may be exploited in such a way that it is no longer necessary to explicitly assemble the global discretization matrix. In particular, given an appropriate input grid, the discretization matrix may be defined implicitly using stencils that are constant for each structured patch. Hence, HHG is designed to have low memory consumption as well as hardware efficient execution and a high degree of parallelism. This approach allows to solve elliptic partial differential equations with a very high resolution. HHG supports different point- and line-wise relaxation schemes for the smoothing procedure.

2 Grid Partitioning and the Coarsest Grids

From the theoretical point of view, MG has a linear complexity of $O(N)$ with respect to the number of unknowns for sparse systems. For serial runs, this observation can also be made in practical use. In parallel settings we are able to achieve this behaviour up to a certain degree that seems to be limited by the increasing amount of communication cost for the coarsest levels. Estimates for the decreasing volume to surface ratio for multigrid hierarchy are given in[7]. As an example, let us assume we want to utilize $300,000$ cores, having one process per core. With static grid partitioning the coarsest grid would consist of $300,000$ elements. At the latest at this stage there are two possibilities to treat this grid: Proceed with the construction of new coarser grid levels by collocating elements on fewer number of cores (agglomeration) or stop at this stage and apply any iterative or direct solver[8].

In our MG algorithm, we decided to go for the second approach. This strategy is also known as flat multigrid, truncated cycle, or U-cycle. For a regular tetrahedral grid it is too drastic to end up with one element per process, like in our example. This would lead to up to four local unknowns per process and at least 44 ghost points. HHG refines each input element twice to generate the coarsest multigrid level, so that the minimal size per process

is 64 elements (up to 35 local unknowns), which provide a reasonable volume to surface ratio.

In the case of grid partitioning two rules have to be considered, when dealing with a very high degree of parallelism. First, when constructing the grids in a setup phase, often global acting algorithms have to be used, e.g. to find communication neighbours or to perform global numbering. Let us assume an operation between two input values costs 10 processor cycles and let n_P be the number of processors. Algorithms of complexity of $O(n_P)$ and $O(n_P log_2(n_P))$ would need $3 \cdot 10^6$ and $5.5 \cdot 10^7$ cycles for $n_P = 3 \cdot 10^5$, respectively. A modern processor can perform these operations in far below one second. However, while for an $O(n_P^2)$ class it is often still possible with $n_P = 10 \cdot 10^4$ to treat our example in 10^9 cycles (around one second), $n_P = 3 \cdot 10^5$ processors would require 10^{11} cycles. This is in the range of a minute on current hardware. It can be acceptable, but our assumptions are quite optimistic. If there is more than one sub-grid per processor, an operation is more expensive, or thinking about the next generations of supercomputers a complexity of $O(n_P^2)$ has to be avoided by choosing an other algorithm or parallelization, if possible. In HHG we had to reduce the complexity for the mesh construction to $O(n_P log_2(n_P))$, since the setup times took much longer than the solving phase.

Second, we have to consider main memory requirements. Since we read in an unstructured input mesh to allow geometric flexibility, basic logical information of the mesh can easily grow up to more than hundreds of MB. This is already the case for a few million elements, which the coarsest grid might have when we use a flat multigrid algorithm. Thus one should consider to hold the coarse grid structure in main memory of each processor just in the setup phase, or to avoid that at all. Every instance of HHG only stores its own part of the coarsest grid to avoid memory problems.

Another issue is, if the problem solution on the coarsest grid of a truncated cycle can be approximated by a non-optimal solver in reasonable time. In this report, CG is used as a solver for the coarsest grid. To give an estimation of the necessary number of CG steps, we assume the following:

- The required number of CG steps is proportional to the diameter of the domain.
- We assume that high frequency error are eliminated by smoothing.
- One CG step for the diameter one with one coarse grid point is sufficient.

Consequently, our simple estimation for the required number of CG-steps n_{CG} for l multigrid levels and d dimensions (here $d = 3$) is

$$n_{CG} \approx \frac{\sqrt[d]{N}}{2^{l-1}}. \tag{2}$$

However, in pratice this is only a rough estimation, multiplied by a constant ($c \approx 1$) because of the third assumption.

3 Scaling on JUGENE

Next we present weak and strong scaling results of our multigrid algorithm. Hereby push HHG to the limits of current CPU parallelism. All our calculations in this report were performed on the super computer JUGENE in Jülich consisting of 73 728 compute nodes

or sockets. Each socket is equipped with IBM's Blue Gene/P (BGP) quad-core processor. The following results are measured by using VN (virtual node) mode on the JUGENE. This means each node executes 4 tasks, sharing 2 GB of main memory.

$$\begin{aligned} \Delta u &= f \text{ in } \Omega\,, \\ u &= 0 \text{ on } \partial\Omega \end{aligned} \tag{3}$$

As a test problem we use Poisson's equation (3) with a right hand side f. It is a hard test problem from a performance and scalability point of view, since it has a low computation to communication ratio compared to more complex PDEs.

3.1 Strong Scaling

To have smaller runtimes per time step, it is interesting not to utilize the full main memory. Fixing the problem size, but increasing the number of processors is known as strong scaling. For computational steering it can be useful to reduce the runtime of one time step to the order of 0.05 seconds to achieve real-time behaviour. In an implicit time stepping, one time step per frame would be feasible. In this section, we want to evaluate the potential of a multigrid algorithm to reach this goal.

In our experimental setting, we solved $2.14 \cdot 10^9$ unknowns using 512 to $49\,152$ cores. For the strong scaling experiment, we use the following multigrid components:

- V(3,3) row-wise red-black GS smoothing,
- 60 steps of a CG solver to approximate the coarsest grid,
- Five multigrid levels,
- Direct coarse grid approximation,
- Linear interpolation.

It can be shown by local fourier analysis[9], that three pre- and post-smoothing steps are necessary to achieve the "classical" multigrid convergence of 0.1 for tetrahedral grids. Fig. 3 shows that HHG achieves a good strong scaling behaviour over a wide range of cores.

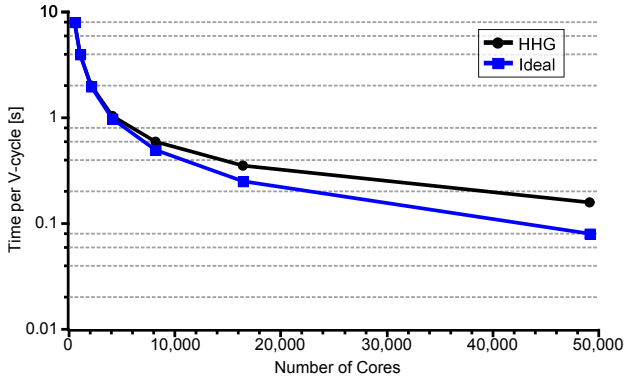


Figure 3. Strong scaling behaviour of HHG on PowerPC 450 cores of a Blue Gene/P located at Jülich. This test case was performed starting from 512 cores and solving a system of $2.14 \cdot 10^9$ unknowns.

Increasing the number of cores by a factor of 96 we are able to reduce the initial time from 7.95 seconds per V-cycle to 0.16 seconds. A perfect strong scaling would result in 0.08 seconds per V-cycle. Thus we can clearly observe that the communication overhead impacts the performance. For the largest run, we have quite small memory arrays of around 383 KB for each variable (right hand side, unknowns, residual) including ghost points. In this data volume, the ghost points require 69 KB. Thus, the volume to surface ratio is quite small. Furthermore, latency of the messages has a larger impact, since more messages per time have to be sent. A significant part of the time (about 38%) is spent for the approximation on the coarsest grid.

With a similar setup, but using V(1,1) cycles and 40 CG steps on the coarsest grid, one cycle takes 0.07 seconds. For an implicit real-time application, additional time is required for a visualization pipeline (post-processing) and the time stepping itself. However, it should be possible to tune the solving further, by e.g. optimizing the coarsest grid solver and the other multigrid components, having less unknowns per core, or utilizing stronger processors. So, highly parallel real-time simulations with a multigrid algorithm seem to be challenging but feasible on current hardware.

3.2 Weak Scaling

This section discusses some effects of solving large linear systems by MG with up to $294\,912$ compute cores. The setup of the weak scaling experiments corresponds to the strong scaling from the previous section except for:

- the number of CG steps depends on the size of the coarsest grid,
- we apply a constant numbers of MG levels:
  - Six for up to $262\,144$ cores and seven for $294\,912$ cores,
- we have 12 structured regions per core for six MG levels and
  we have 1 structured region per core for seven MG levels.

The parallel efficiency is reflected in Tab. 1 by the time per V-cycle. Utilizing the whole machine, HHG achieved a parallel efficiency of 69.2%. In the limit HHG solved up to 10^{12} degrees of freedom (DoF).

When fixing the number of levels for V-cycles, the number of required CG steps on the coarsest grid grows from 15 to 180 CG steps. However, for the largest runs the time on the coarsest mesh is around 12.5% of the total time for one V-cycle only. A comparison of the work done on the coarsest grid between two different MG cycles is given in Fig. 4. Our prediction is calculated by Eq. 2. In our setup an F-cycle is very similar to the prediction. An F-cycles is similar to a full-multigrid cycle and requires a better coarse grid approximation than a V-cycle, i.e. for a V-cycle approximately only half of the CG-steps are necessary. Moreover, we have to keep in mind that the number of CG-steps between the restarts is different. Restart means that only the approximated solution of the previous step is available, but no additional information like previous search directions. In our case, there are five times more CG-restarts for the F-cycle than for the V-cycle. However, the measured number of CG-steps is in the same order of magnitude as predicted. One reason for deviations from the prediction in the measurements is the shape of the domain. At many points in Fig. 4, the domain in one or two directions is twice as large as in the others. Furthermore, we do not consider a spherically shaped domain.

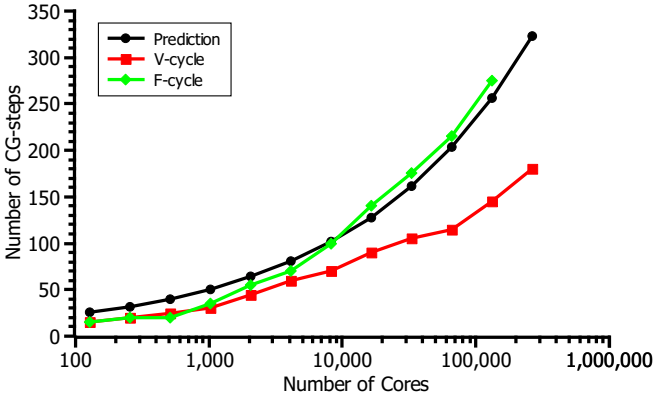


Figure 4. Required number of CG-steps on the coarsest grid per cycle with increasing numbers of compute cores.

Cores	**Struct. Regions**	**DoF** $(\cdot 10^6)$	**CG**	**Time (s)**
128	1 536	535	15	5.64
256	3 072	1 071	20	5.66
512	6 144	2 142	25	5.69
1024	12 288	4 287	30	5.71
2048	24 576	8 577	45	5.75
4096	49 152	17 159	60	5.92
8192	98 304	34 326	70	5.86
16384	196 608	68 669	90	5.91
32768	393 216	137 355	105	6.17
65536	786 432	274 744	115	6.41
131072	1 572 864	549 555	145	6.42
262144	3 145 728	1 099 176	180	6.52
294912	294 912	824 365	110	3.80

Table 1. Weak scaling behaviour of HHG on PowerPC 450 cores of a Blue Gene/P at Jülich.

The full machine run uses seven instead of six levels. In our semi-structured approach, a structured region cannot be shared by multiple processors. Thus, we are not able to utilize full main memory. The additional level reduces the number of CG iterations. The performance in terms of solved unknowns per second increases by 14% due to larger inner loops for the finest grids. Here, the overall achieved performance using the full machine is 59.8 TFLOP/s in the solving phase.

4 Conclusions and Future Work

We presented scaling results for geometric multigrid within the HHG software on the JUGENE supercomputer located at Jülich. We addressed scalability problems and communi-

cation overhead created by the coarsest grid in the multigrid hierarchy. A careful implementation results in excellent scalability results, i.e. the coarse grids do not seriously effect the overall parallel performance. To this end we explored and analyzed the weak scaling of numerical experiments with up to 10^{12} unknowns.

Next we plan to compare our results to another geometric multigrid framework on different architectures. In this context hybrid parallelization and GPU acceleration could be interesting issues. Apart from that we are extending HHG to treat the Poisson problem occurring in vortex particle direct numerical simulations.

Acknowledgements

The work was supported by the Kompetenznetzwerk für Technisch-Wissenschaftliches Hoch- und Höchstleistungsrechnen in Bayern (KONWIHR) and the International Doctorate Program (IDK) within the Elite Network of Bavaria.

We would like to thank Jutta Docter and Bernd Mohr, Jülich Supercomputing Centre (JSC) for their excellent support during the largest runs.

We are grateful to the Jülich Supercomputing Centre for providing the computational resources on JUGENE.

References

1. H.J. Bungartz, M. Mehl, T. Neckel, and T. Weinzierl, *The PDE framework Peano applied to fluid dynamics: an efficient implementation of a parallel multiscale fluid dynamics solver on octree-like adaptive Cartesian grids*, Computational Mechanics, **46**, no. 1, 103–114, 2010.
2. R. Falgout, V. Henson, J. Jones, and U. Yang, *Boomer AMG: A Parallel Implementation of Algebraic Multigrid*, Tech. Rep. UCRL-MI-133583, Lawrence Livermore National Laboratory, 1999.
3. R.S. Sampath and G. Biros, *A Parallel Geometric Multigrid Method for Finite Elements on Octree Meshes*, SIAM Journal on Scientific Computing, **32**, 1361, 2010.
4. W. Briggs, V. Henson, and S. McCormick, *A Multigrid Tutorial*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, USA, 2nd edition, 2000.
5. B. Bergen, T. Gradl, F. Hülsemann, and U. Rüde, *A Massively Parallel Multigrid Method for Finite Elements*, Computing in Science and Engineering, **8**, no. 6, 56–62, 2006.
6. T. Gradl, C. Freundl, H. Köstler, and U. Rüde, *Scalable Multigrid*, in: High Performance Computing in Science and Engineering. Garching/Munich 2007, pp. 475–483, 2008.
7. F. Hülsemann, M. Kowarschik, M. Mohr, and U. Rüde, "Parallel geometric multigrid", in: Numerical Solution of Partial Differential Equations on Parallel Computers, pp. 165–208. Springer-Verlag, Berlin, Heidelberg, New York, 2005.
8. D. Xie and L.R. Scott, *The parallel u-cycle multigrid method*, in: Proceedings of the 8th Copper Mountain Conference on Multigrid Methods. Citeseer, 1996.
9. R. Wienands and W. Joppich, *Practical Fourier Analysis for Multigrid Methods*, vol. 5 of *Numerical Insights*, Chapmann and Hall/CRC Press, Boca Raton, Florida, USA, 2005.

Comparable Corpora in Wikipedia Text for Machine Translation

Jia Xu[1], **Casey Kennington**[2], **Česlav Przywara**[2], **and Lilian Wanzare**[2]

[1] Institute for Interdisciplinary Information Sciences, Tsinghua University
Beijing, 100084, P. R. China

[2] Department of Computational Linguistics, Saarland University, D-66123, Saarbrücken, Germany
E-mail: {xujjia, bakuzen, ceslav.przywara, liliwanzie}@gmail.com

Machine translation systems need training data in order to function. Those tranining data are known as parallel corpora, which is a set of sentences in two languages where the sentences are translations of each other. At this point in time, there are several useful and sizable parallel corpora available, but there is still a need to gain more. We use a known method to extract parallel sentences from a comparable corpora, and use the extracted sentences as parallel corpus to help in MT accuracy. We describe the methodology for doing this, and how we used the Moses MT Experiment Management System to run our experiments.

1 Comparable Corpora

1.1 Overview

Statistical Machine Translation (MT) systems are designed to find patterns of how a source language is translated into a target language. In order to do this, large amounts of what are called parallel corpora are required for training the systems. These parallel corpora are sentences in the two languages, where each line in each corpus corresponds to its translation in the other language. There are several well-known corpora like these available, such as Europarl[1], JRC[2], and DGT[3]. These are very useful corpora for the languages that are supported. In general, the more training data like these that can be used, the better and more accurate the translations tend to be.

Though there are useful parallel corpora available, there are some limitations. First, not all languages are supported. Second, these training corpora only cover a small domain of possible translations. Natural human languages are dynamic and speakers of those languages create novel sentences continually, but if we train an MT system on corpora within a certain domain and context, it might be difficult to later discern how something should properly be translated. For example, Europarl contains proceedings of European Parliament and both JRC and DGT are translations of law into the European languages. Now, how should the MT system translate, for example, a text about sports or technology? Thus we need to increase the reach of domains that an MT system has. To combat these two problems, we need parallel corpora for each language and each domain. However, coming up with these translations it gets expensive and is very time-consuming. There is, however, an easier way.

The Internet has a vast amount of text in many languages. There are websites such as news sites and wiki sites that might contain similar information on a subject in different languages. Even though a news article in two different languages may not be translations of each other, they will most likely contain very similar information. In the same way, wiki

sites such as Wikipedia have articles and pages about millions of topics, many of them are in a number of different languages. Again, they may not necessarily be translations of each other, but they might contain similar information. Now imagine obtaining many thousands of these pages into a corpora in two languages. These corpora, though not quite ready to be used as MT training data, are called comparable corpora, and with some processing could be made to be useful to an MT system.

1.2 Previous Work

As stated earlier, this idea is not new. In face, several approaches have surfaced which have all had some success to some degree.

- Approach 1: WER, TER, TERp - Sadaf Abdul-Rauf and Holger Schwenk (2009)
- Approach 2: Maximum Entropy - Munteanu and Marcu (2004)
- Approach 3: Phrase Extraction - Munteanu and Marcu (2006)
- Approach 4: Information Networks - Heng Ji (2009)

1.3 Our Work

We take the approach given by the 2004 paper by Munteanu and Marcu. First, we spent time establishing baseline scores for a number of language pairs. In order to accomplish this, we used the freely available Moses SMT system, particularly the Experiment Management System (EMS) to encapsulate and organize our experiments. We used the previously mentioned corpora (Europarl, DGT, JRC) to train and a News-Commentary corpora to test. Our baseline scores are the well-known MT metric BLEU[4]. We made use of the JUROPA[5] cluster, which included installing and connecting together all of the necessary programs such as GIZA++[6], Moses[7], SRILM[8], and organizing all of the corpora. We then used a previously processed set of Wikipedia data in English and German to use as a comparable corpora, which we processed and compared to our baseline score for the German-English language pair.

2 Obtaining Baseline Scores

2.1 Moses EMS

Moses is a freely-available statistical machine translation system, and it comes with a lot of programs to help you train, tune, and evaluate your MT system. It comes with an experiment harness, known as the Experiment Management System that is a series of perl script files which read and parse a configuration file, then do all of the training, tuning, and evaluation for you. The said configuration file needs to have access to a working directory, and know where the training, tuning, and evaluation data files are. It also needs to know where all the executables are for the supporting programs that are to be run, such as SRILM, but for the most part everything it needs is relative to the moses installation directory.

One nice feature of the EMS system is that it keeps track of each step of the process, so it is easy to find and fix problems that might arise. Better still, when you find and fix a problem, you can invoke the EMS system again and it will pick up where it left off, not needing to redo everything from the start. It only redoes necessary steps. It can also detect changes in the configuration file, so if you change it and add, for example, another set of training corpora, it will know to re-train and any subsequent step that might need to be redone.

The EMS system further gives visual output of the process. Before you invoke the script to execute, you can run it in a test mode where it checks to make sure the files that your configuration file points to actually exist, and that your settings should work. It will then display an image file which is a graph, or one might call it a flowchart, of the process. Parts that need to be completed are green. Parts that are completed will turn blue. Parts that have errors and cause the process to stop will show in red, thus making it easy for you to know where to look in order to fix the problem. During execution, it will also display the image and if you leave it open, you will see it updating the colour scheme as it goes along the process. If you close the image, the script will still run, and at any time you can open the image to see what step the script is on. Each step of the EMS is kept, and log files are well organized. Every training file, or output of any supporting program are kept. This is quite useful if, for example, you want to use your trained files or language model file for another experiment. It can potentially take up quite a bit of space, but that is something that MT generally requires.

2.2 Data Preprocessing

Because of the number of language pairs we dealt with, we had to use several different corpora for training and baselines evaluation. Usually each of them comes in its own format (both charset encoding and internal structure), so a bit of preprocessing was required. For each part of the EMS pipeline Moses requires the input data to be provided in two files: one containing source language sentences and the other target language sentences. These files must be in a plain text, sentence per line format with parallel sentences placed on corresponding lines.

We should mention that Moses EMS can utilize multiple corpora for the given language pair in a single experiment. We took advantage of that whenever it was possible.

As the last step, all the data has been converted to UTF-8 encoding. Sometimes this required also replacement of XML character entities by corresponding UTF-8 characters - for example ü has been replaced by ü etc.

2.3 Baseline Scores

When evaluating the baseline scores we followed an approach taken by Koehn at al. as described in *462 Machine Translation Systems for Europe*[9] and we set up Moses system with the same settings: maximum sentence length 80 words, bi-directional msd reordering model and 5-gram language model. On the other hand we didn't experiment with exactly the same language pairs, thus we employed also several different datasets (corpora) in addition to JRC.

Following datasets have been provided by other parties and used without modification:

- *acquis* - dataset extracted from part of JRC corpus (as described in aforementioned paper by Koehn et al.), provided as a test set by Euromatrix project[10]. Domain: EU legislation.
- *eparl* - dataset based on a recent release (version 5) of the Europarl corpus, provided as a training data for WMT 2010 Translation Task[11]. Domain: proceedings of European Parliament.
- *nc* - dataset based on News Commentary corpus - provided as a training data for WMT 2010 Translation Task.
- *racai* - dataset of 3000 German-Romanian sentence pairs covering three domains: legal, medical and software related (1000 sentence pairs from each domain).
- *tilde* - dataset provided by the Tilde company to evaluate the translation quality.
- *wmt-dev* - dataset provided as a part of development data for WMT 2010 Translation Task (2008 test set).
- *wmt-test* - dataset provided as a part of development data for WMT 2010 Translation Task (2009 test set).

Following datasets have been extracted from freely available parallel corpora for purpose of the evaluation task:

- *dgt* - dataset extracted from entire DGT corpus. Domain: EU legislation.
- *dgt-f* - dgt dataset without the sentences found in acquis dataset (a match of sentence in any language resulted in whole sentence pair being taken off the dataset).
- *ijs-elan* - dataset extracted from IJS-ELAN corpus[12]. IJS-ELAN corpus contains 1 million words from 15 parallel Slovene-English / English-Slovene texts. Domain: not specified.
- *jrc* - dataset extracted from entire JRC corpus. Document pairs with the same number of sentences have been implicitly treated as 1:1 aligned. Other document pairs have been processed using sentence alignments provided with the corpus and only sentence pairs with 1:1 alignment have been put into dataset. Domain: EU legislation.
- *jrc-f* - dataset extracted from JRC corpus in the similar manner as in case of jrc dataset, but with two additional restrictions: (1) Sentence pairs from the documents used for composition of acquis dataset have been omitted. (2) Any other sentences found in acquis dataset have been also omitted.
- *setimes* - dataset based on parallel corpus of the Balkan languages, generated from Setimes news articles[13].

Tab. 1 presents baseline scores for all the language pairs evaluated in our task.

Source language	Target language	Training dataset	Tuning dataset	Testing dataset	BLEU score
Croatian	English	setimes	-	setimes[600]	26.0
English	Croatian	setimes	-	setimes[600]	30.2
English	Estonian	dgt, jrc	tilde[1000]	tilde[520]	10.1
English	Greek	setimes	-	setimes[600]	24.8
English	Latvian	dgt, jrc	tilde[1000]	tilde[520]	10.2
English	Lithuanian	dgt, jrc	tilde[1000]	tilde[520]	11.5
English	Romanian	setimes	-	setimes[600]	40.8
English	Slovenian	dgt, jrc	ijs-elan[1500]	ijs-elan[1000]	11.4
German	English	eparl, nc	wmt-dev	wmt-test	19.4
German	Romanian	dgt, jrc	racai[1500]	racai[1500]	18.1
Greek	Romanian	setimes	-	setimes[600]	35.4
Latvian	Lithuanian	dgt, jrc	tilde[1000]	tilde[520]	8.7
Lithuanian	Romanian	dgt-f, jrc-f	-	acquis	35.7
Romanian	English	setimes	-	setimes[600]	31.2
Romanian	German	dgt, jrc	-	racai[1500]	14.9
Romanian	Greek	setimes	-	setimes[600]	22.1
Slovenian	English	dgt, jrc	ijs-elan[2000]	ijs-elan[1000]	12.4

Table 1. Baseline scores for all evaluated language pairs.

3 Maximum Entropy

3.1 Munteanu and Marcu 2004

In the Munteanu and Marcu 2004 paper, they trained a maximum entropy classifier with some small parallel corpora and non-parallel corpora. Their training data included some in-domain traning data (news text), but mostly out-of-domain data (UN Proceedings). They also used a specific amount of non-parallel training data to add to the maximum entropy classifier. They did this so the classifier would be able to learn about truly parallel sentences and non-parallel sentences and be able to distinguish what it takes for two sentences to indeed be parallel.

In order to train a maximum entropy classifier, they needed to determine what features could be used to best distinguish if two sentences are parallel or not. Of course, one very important thing is to know if an individual word in a sentence translates into a word in the target sentence, so a dictionary will be necessary. They also made use of the IBM Model 1 which gave a word-level alignment which became a very strong indicator if two sentences are translations of each other. If two sentences have words that don't correspond via a dictionary, they obviously won't be parallel. Further, they took the sentence lengths, longest continuous span of translated words, longest unconnected string, and used those features for training with the before mentioned training data.

With a trained model, they then took large amounts of news text that they previously obtained. Now, which sentences should we compare? If we take a single sentence in the source language and then test it against every sentence in the target language, the processing could take a very, very long time. So, they took news articles that were within 5-10

days of each other, taking the assumption that news stories which discuss the same event will appear on websites in a similar tempora timeframe. This greatly reduced the number of comparisons, but left it open enough to find sentences that would indeed be translations of each other. They kept the sentences that had a 0.7 probability of being translations of each other and appended those to a list for each language, in the end making a parallel corpus, in their case, in Arabic and English.

3.2 Our Project

We did something similar to the 2004 Munteanu and Marcu paper. However, we didn't use the IBM Model 1. Instead, we found a custom dictionary and did our own alignment search. We used a freely available maximum entropy classifier[14] written in C++. It was quite easy to adapt to our needs. We used features such as number of translated words, longest contiguous span, longest unconnected substring, sentence lengths, and the difference between the source and target sentences.

Also unlike the paper, we used Wikipedia text. The set of all articles in German and English were previously processed for matching titles and sentence segmenting. We further processed the resulting text to make sentences into smaller segments by splitting on specific punctuation such as commas and parenthesis. This works because the sentences aren't aligned parallel, but are rather compared by taking a page in English and German where the title is the same and checking all sentences with all other sentences, where a sentence is on its own line in the file. Shorter segments are easier for an MT system to process and learn from, so if possible we keep the sentences short.

We train on 10,000 sentences from Europarl and another 10,000 sentences from the news-commentary corpora for the parallel training. For the non-parallel training, we use the same corpora, we only shift the source language by 1 sentence, so a sentence is compared to the sentence that is 1 line below its corresponding sentence. These sentences are non-parallel and useful to help the maximum entropy classifier distinguish between translations and non-translations. One thing to note here, and a similar issue arose in the 2004 Munteanu and Marcu paper, was the amount of training sentences for parallel vs. non-parallel. To be accurate, most sentences in existence in two languages are not translations of each other, but if we trained, for example, using 10,000 sentences of parallel text, and then trained each sentence on the 9,999 other non-parallel sentences, we would end up with a model that has seen so many non-parallel sentences that it would just assume all sentences are non-parallel and we wouldn't find any possible translations at all. Allowing the number of parallel training sentences and non-parallel traning sentences to be approximately the same allows the features to do most of the distinguishing.

As explained before, we took our trained model and used it to distinguish sentences in the two languages given Wikipedia articles on the same topic. We checked every sentence in the article in the source language against every sentence in the target language. Our threshold of probability for sentences to be translations of each other was set to 0.99, because during our tests it was considering some sentences to be translations that obviously weren't when it was set to 0.7 as in the paper.

The number of resulting sentences after processing was 56,000. On inspection, a look at some of the sentence translations looked reasonable. German and English had their own respective sentence files that were added to the Eurparl/News-Commentary Ger-

man/English baseline, which was 19.38, and retrained. The improvement was only slight, the new BLEU score increased to 19.40.

4 Conclusion

In our case, there wasn't much improvement but there are many improvements we can make to our system, including adding the IBM-1 model information as a feature. However, there are many languages and language pairs that need improvement in their machine translation ability, and making use of comparable corpora is a practical approach to getting the training data needed to improve machine translation for many language pairs.

5 Acknowledgements

This work was partly supported by the FZ Jülich with an access to the supercomputer JUROPA through project HSB07.

References

1. Various, *European Parliament Proceedings*, statmt.org.
   `http://www.statmt.org/europarl/`
2. Various, *JRC-Acquis Multilingual Parallel Corpus*, langtech.
   `http://langtech.jrc.it/JRC-Acquis.html`
3. Various, *The DGT Multilingual Translation Memory of the Acquis Communautaire*, langtech.
   `http://langtech.jrc.it/DGT-TM.html`
4. K. Papineni, S. Roukos, T. Ward and W. J. Zhu, *BLEU: a method for automatic evaluation of machine translation*, ACL 2002, pp. 311-138.
5. Various, *JUROPA - Jülich Research on Petaflop Architectures*, fz-juelich.de.
   `http://www.fz-juelich.de/jsc/juropa/`
6. F. J. Och and H. Ney, *A Systematic Comparison of Various Statistical Alignment Models*, Computational Linguistics, volume 29, number 1, pp. 19-51 March 2003.
   `http://www.fjoch.com/GIZA++.html`
7. Various, *Moses - statistical machine translation system*, statmt.org.
   `http://www.statmt.org/moses/`
8. A. Stolcke, *SRILM - An Extensible Language Modeling Toolkit*, Proceedings of International Conference on Spoken Language Processing, Denver, Colorado, September 2002.
   `http://www.speech.sri.com/projects/srilm/`
9. P. Koehn, A. Birch and R. Steinberger, *462 Machine Translation Systems for Europe*, Proceedings of International Conference on Spoken Language Processing, Denver, Colorado, September 2002.
10. Various, *Euromatrix Test Sets List*, statmt.org.
    `http://matrix.statmt.org/test_sets/list`
11. Various, *Shared Task: Machine Translation for European Languages*, ACL 2010 Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR, statmt.org.
    `http://www.statmt.org/wmt10/translation-task.html`

12. Erjavec T., *The ELAN Slovene-English Aligned Corpus*, Proceedings of the Machine Translation Summit VII, pp. 349-357. Singapore, September 1999.
http://nl.ijs.si/elan/
13. Tyers F. M. and Alperen M. S. *South-East European Times: A parallel corpus of the Balkan languages*, Proceedings of the Workshop on Exploitation of multilingual resources and tools for Central and (South) Eastern European Languages, LREC 2010.
http://www.setimes.com
14. Various, *A simple C++ library for maximum entropy classification*, University of Tokyo, Tsuji Labratory.
http://www-tsujii.is.s.u-tokyo.ac.jp/~tsuruoka/maxent/

Hydrodynamics and Turbulence

Dietmar Kröner

Department of Applied Mathematics, University of Freiburg
Hermann-Herder-Str. 10, D-79104 Freiburg, Germany
E-mail: dietmar.kroener@mathematik.uni-freiburg.de

In this section different application concerning incompressible and compressible flows and reactive flows, jets at high Mach number, aeroacustics, premixed preheated Bunsen flames, capillarity effects for liquid-gas systems, particulate flows, influence of sphere packings on hydrodynamic dispersion and two-phase flow of dilute suspensions of solid particles are considered. The underlying mathematical models are based on the compressible Navier-Stokes equations, low Mach number cases and probability density functions. Turbulence models also for turbulent combustions are included. Finite volume, AUSM and lattice Boltzmann methods are used for the discretization. The efficiency and the speedup for applications up to 10.000 mill. gridpoints and up to 16000 core on JUGENE are impressive.

A detailed numerical investigation of a "confined jet flame" and a "jet in cross flow" on high performance computers have been considered in "Numerical Investigation of Reacting and Non-Reacting Flows in Gas Turbine Related Configurations" by M. Di Domenico, E. Ivanova. The mathematical model is based on a low Mach number formulation for the transport equation and the discretization is given by a finite volume scheme on unstructured grids. The turbulence modeling has been done with three different methods: unsteady Reynold's-averaged Navier-Stokes equations, scale-adaptive and large Eddy simulations. The turbulence combustion coupling strategies uses Eddy dissipation concepts and probability density functions.

The results of the paper about "Walberla: Simulations of Complex Flows on Supercomputers" by J. Götz, St. Donath, C. Feichtinger, K. Iglberger, H. Köstler, U. Rüde has also been accepted for the supercomputing conference in 2010. They concern the simulations of capillarity effects (including wetting effects for liquid-gas systems) related to fuel cells in order to prove the underlying physics and to improve manufacturing processes. Furthermore they consider simulations of particulate flows, i. e. the motion of immersed non spherical particles by a dynamically coupled 3D-lattice Boltzmann and rigid body dynamics code. They got an excellent efficiency and 4.59 % peak GFLOP/S performance. This work was supported by the Kompetenznetzwerk für Technisches-Wissenschaftliches Hoch- und Höchstleistungsrechnen in Bayern und durch das Bundesministerium für Bildung und Forschung.

In the paper "Impact of the Heterogeneity of Random Sphere Packings on Hydrodynamic Dispersion" the authors S. Khirevich und U. Tallarek study the influence of sphere packings with different porosity and micro structures on advection diffusion transport, in particular on the hydrodynamic dispersion coefficients. It is shown that in addition to the porosity of the packing and to the average flow rate of the carrier fluid a third parameter, related to the packing micro structure is important for the hydrodynamic dispersion coefficient. The concept is performed on a lattice with up to 600x600x4.300 nodes and high

memory and CPU requirements are necessary.

G. Geiser, M. Meinke, W. Schröder consider in the paper "A Hybrid Aeroacoustic Prediction Method for Non-Reactive and Reactive Flows" flow induced noise. The main feature of this paper consists in the separation of the generating fluid field close to the acoustic source and the far field sound. Two important applications are treated: coaxial multi-spacious helium/airjets at high Mach number and lean premixed preheated Bunsen flame for low Mach number. The simulation is based on the acoustic perturbation equations and the large Eddy simulation model. Discretization of the Euler terms have been done by finite volume method based on AUSM schemes. The computations are performed on a grid up to 16 mill. grid points. They obtain an excellent speed up for up to 4.096 cores and a good speed up beyond.

The main focus of the contributions of M. Uhlmann and T. Doychev about "Finite-Size Particles in Homogeneous Turbulence" lies on two-phase flow of dilute suspensions of solid particles. In particular they study the influence of the particle to the turbulence, the influence of turbulence to the average particle velocity and the spatial structures of the dispersed face. The mathematical model consists of the incompressible Navier Stokes equations in the spaces between the moving particles, which are resolved with 15 grid points per diameter and up to 16.384 cores on JUGENE. For the computational grid they use up to 10.000 mill. grid points. The efficiency is decreasing for more than 100 cores.

Numerical Investigation of Reacting and Non-Reacting Flows in Gas Turbine Related Configurations

M. Di Domenico and E. Ivanova

DLR - German Aerospace Center, Institute of Combustion Technology, 70569 Stuttgart, Germany
E-mail: massimiliano.didomenico@dlr.de

Detailed analysis of relevant turbulent, reacting flows in typical and new gas turbine configuration by means of High Performance Computing (HPC) techniques are presented and discussed in this work. Both coaxial jet and jet in cross flow configurations are investigated under atmospheric and high-pressure conditions. The impact of turbulence and combustion models are outlined and major advantages and drawbacks are discussed. Several turbulence-chemistry interaction approaches like assumed PDF and Eddy Dissipation Concept approaches are introduced and used under different conditions. Detailed chemistry effects on flame stabilization mechanism and stability are shown. Autoignition and mixing phenomena in the practically relevant configurations are studied in detail by means of the unsteady scale-resolving simulations. The comparison with the experimental data are presented in order to provide a validation basis.

1 Introduction

The numerical investigation of turbulent reacting flows in gas turbine related configurations is nowadays of great interest. More stringent emission rules on pollutants, in particular NO_X, are issued and new combustor geometries must be designed to meet them. Therefore, these concepts need to be investigated and further developed[1]. A reliable way to reduce NO_X emissions and to increase combustion efficiency is represented by the fuel-lean and/or premixed combustion. However, the combination of such concepts issues problems related to an unstable combustion process, i.e. both low and high frequency oscillations may arise. In very lean cases local extinction phenomena accompanied by re-ignition or at worst lean blow out are observed. As several time- and space-scales are involved in these processes, such phenomena are far from being of easy modeling.

The use of H_2-rich fuels in lean premixed combustion systems leads to special challenges related to a lower autoignition temperature and flashback margins compared to natural gas[2]. This is especially true in the case of the reheat, or sequential, combustion systems[3,4]. This combustion concept is characterized by two fuel injection and energy conversion steps in two separate combustion chambers with an expansion step in a high pressure turbine stage in between. The exhaust gas is then further expanded in a low pressure turbine. The mixing zone of the reheat combustor is characterized by high temperatures (above 1000 K). These conditions lead to extremely short ignition delay times.

In order to shorten and make the design of new combustion chambers more efficient, Computational Fluid Dynamics (CFD) joined to (and in some cases in substitution to) experimental measurements should be used. However, a predictive use of the CFD analysis is possible only if new models are implemented and tested with help of both academic test cases and industrial relevant configurations, i.e. complex geometries, high-pressure operating conditions and highly turbulent flows. It should be kept in mind that the degree of complexity of a numerical simulation drastically increases with the number of phenomena

involved (reacting flows) and details included (three-dimensional flows). The application of simulation techniques to complex geometries typical of gas turbine model combustors also represents a difficult task because of the different timescales involved. For example, a reactive simulation with a simple one-step combustion model (one reaction, five species) would require computational resources which exceed of a factor two those needed by an aerodynamic calculation performed on the same grid, for which only velocity, density and energy fields are solved. Additionally, in order to properly take into account the highly unsteady dynamics of swirled flows, very small time steps are required. It turns out that the use of HPC facilities becomes the only practicable way to carry out such ambitious simulations.

2 Numerical Approach

2.1 Code Description

At the German Aerospace Center, Institute of Combustion Technology the in-house THETA code has been developed for the simulation of complex problems[5–8] as those described in the Introduction. The code is programmed in ISO C (C89), thus is fully portable on any machine which provides a compiler adherent to this standard. A domain-based parallelization realized by means of the MPI paradigm allows to run the code on multi-processors, multi-nodes architecture in a seamless fashion.

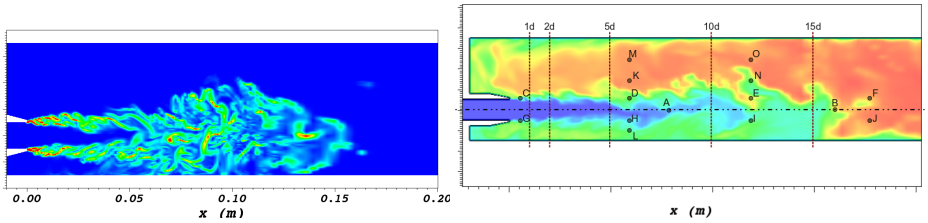
The finite-volume formulation on unstructured meshes allows an easy investigation of complex geometries typical of gas-turbine combustor. The low Mach number formulation of the transport equations and the application of *ad hoc* point-coupling techniques for the species transport equations yield an efficient and scalable algorithm.

In the THETA code flow solver a variety of the popular turbulence modeling concepts is implemented. In the present work turbulence is modelled according to three different approaches: unsteady Reynolds-Averaged Navier-Stokes (URANS), Scale-Adaptive Simulations (SAS)[9], and Large-Eddy Simulations (LES). Several turbulence-combustion coupling strategies are available in the THETA-Code, from a simple Eddy-Dissipation Concept (EDC)[10] to the more evolved assumed probability density function (PDF)[11]. As the code implements a general, detailed, multi-step combustion model, the timescales associated with different reactions and species can be included in a straightforward and seamless manner. A transport equation for the static enthalpy provides the thermodynamic field necessary to determine the temperature distribution and the density field is determined according to the law of perfect gases.

3 Simulations

3.1 Confined Jet Flame

As the experimental test rig was already presented in Ref. 12, 13, only few details are given here. The burner consisted of a single jet nozzle with an inner diameter d, and a combustion chamber with a rectangular cross section with edge lengths of $a \times b = 5\,d \times 4\,d$ and an overall height of $h = 60\,d$. The tip of the jet raised 2 d above the burner base plate. The nozzle was positioned at 3.5 d distance from the combustion chamber wall



(a) Gradient of the progress variable.

(b) Instantaneous temperature field and simulation instrumentation details.

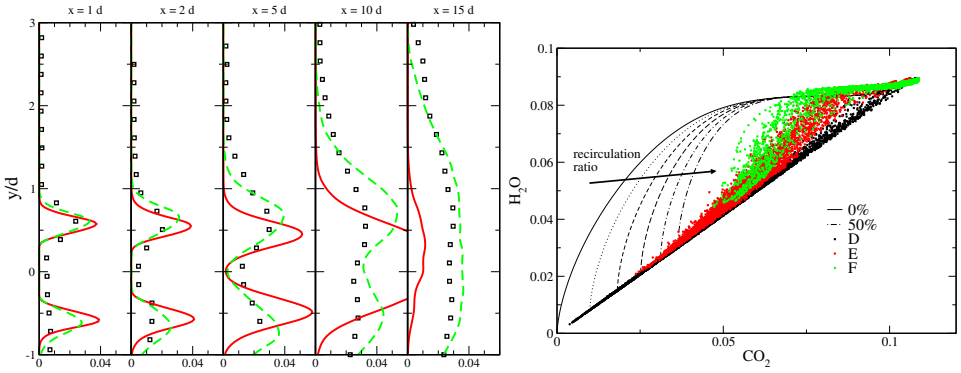
Figure 1. Two dimensional representation of the jet flame results.

referring to the longer side, and centred at 2 d distance in perpendicular direction. The off-centre arrangement was selected to obtain a pronounced recirculation on one side of the jet flow, thereby shaping a flow field in analogy to the inner recirculation zone of a FLOX®combustor. The walls of the combustion chamber were quartz glass plates with reasonable thickness to minimize heat exchange with the surroundings. As a result, the confined jet flames had mainly contact with hot glass walls, providing a very good optical access in addition.The domain included in the numerical simulations is represented by the inlet pipe (a length of 10 diameters are included) and the complete squared section combustion chamber. A fully hexaedral grid of about 5.6 million grid elements is used with more points put in the shear layer regions. The methane combustion is described by the reduced GRI mechanism DRM-19[14] including 19 reacting species and 84 reactions.

All CFD simulations have been run on the JUROPA cluster (located in Jülich, Germany) consisting of node having pairs of quad-core Xeon X5570 processors with Infiniband QDR connection. Starting from the unsteady EDC solution presented below, the finite-rate chemistry simulation takes around three weeks on 8 nodes to get converged statistics.

The instantaneous field of gradient of the progress variable on the chamber symmetry plane as obtained from the detailed chemistry model with the assumed PDF approach is given in Fig. 1(a). It clearly shows the potential core due to the incoming fresh mixture and the instabilities arising in the region where a strong mixing with the hot combustion products takes place. The intensity of these instabilities increases along the chamber and yields the formation of isolated pockets of fresh mixture surrounded by hot gases. From the plot it is also observed that the mixing process does not take place along a single flame front but in local sheets discontinuously distributed. The layer orientation also changes according to the region considered: while near the pipe exit highest gradients are orthogonal to the $x-$axis, downstream the layer orientation becomes more chaotic with highest gradients parallel to this direction.

Rms profiles of progress variable are shown in Fig. 2(a). Predictions given by the EDC combustion model are given in the same figure for comparison. Large discrepancies between the prediction given by the EDC model and the experimental data can be observed at locations higher than 1 d. While the measurements at $x = 10\,d$ still show a bimodal distribution, the simulation adopting the simplified EDC model predicts a single peak with values largely above the experimental levels. On the other hand, at $x = 15\,d$ the damping



(a) Rms of progress variable. (b) D-F.

Figure 2. Confined flame post processed data.

of progress variable fluctuations is largely over predicted.

A detailed analysis of the flame regime at different location of the combustion chamber is presented by comparing data extracted at the locations shown in Fig. 1(b) with laminar free flame calculations performed with the open-source software Cantera[15] adopting the GRI 3.0 mechanism[16]. As it is known that hot recirculating gases play an important role in the investigated configuration, these calculations are performed with a variable amount of hot combustion products that are adiabatically mixed to the inlet fresh mixture. Recirculation ratios ranging from 0% up to 50% (step 10%) are simulated and resulting trends are plotted on the same diagram.

A comparison of samples extracted from three points placed in the upper shear layer can be seen in Fig. 2(b). Samples belonging to the nearest location are mainly on the mixing line, although some reacting spots can be detected at high CO_2 concentrations (i.e. high temperature). At $x = 12\ d$ all thermodynamic conditions range from mixing-only to burning mixture are observed. It is worth to note that samples move away from the mixing line as CO_2 increases, unlike what it is observed for point D. Thus, at this stage the mixing process between hot and cold flow particles promotes auto-ignition with no inert samples available. At monitor point F a degree of recirculation which in some cases approaches 50% is assessed.

3.2 Jet in Cross Flow

The experimental investigations on the autoignition and mixing of H_2-rich fuels in a jet in cross-flow configuration under the practically relevant conditions were conducted in our institute, thus a comprehensive experimental validation database was available. For the current simulations the mixing section of the research reheat combustor described in detail in[17,18] is the object of interest. The mixing section is completely reproduced in the chosen computational domain as shown in Fig. 3.

The hot flue gas at the inlet of the mixing section comes from the hot gas generator[17,18]. The simulations are conducted under the same conditions as in experiment. These conditions are chosen to resemble the real temperatures, species concentrations, inlet velocities

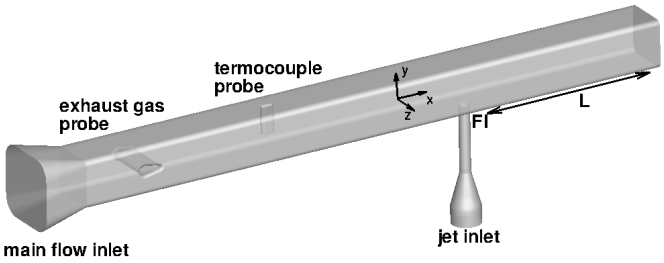


Figure 3. Computational domain - jet in cross-flow.

and pressures of the typical reheat combustors. The fuel is injected from the lower wall of the mixing section in a jet in cross-flow configuration as shown in Fig. 3. A more detailed description of the experimental and computational conditions can be found in[17–19].

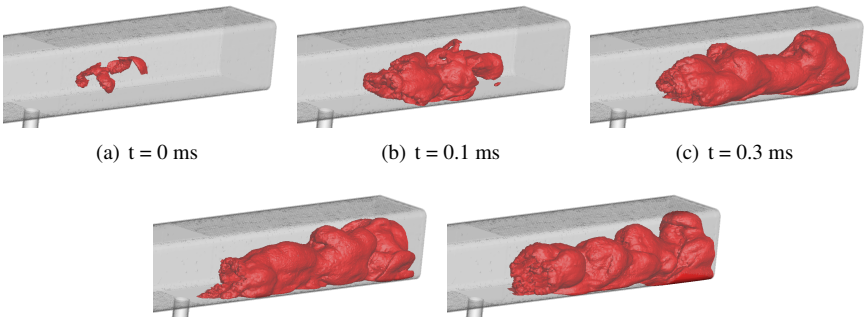
All calculations were performed on polyhedral unstructured grids with hexahedra-layers on the walls. The grids are refined in the important flow regions (jet shear layers) based on the results of preliminary calculations using the grid adaptation procedure available in the DLR THETA code. The basic computational grid consisted about 1M nodes, finer grids were used for the grid dependence studies.

For the scale-resolving unsteady turbulence simulations the SST-SAS model is used[9]. For the combustion modeling in the reacting case simulations the finite-rate combustion model was used in combination with the assumed joint PDF approach. The employed reaction mechanism is the 9-species 21-steps scheme of O' Conaire et al.[20]. More details can be found in[19].

The results of the autoignition simulations are presented in Fig. 4. Isosurfaces of the OH mass fraction Y_{OH} were chosen to be the flame front indicator. The time point of the data snapshot of Fig. 4(a) was taken shortly after occurrence of the ignition kernels. This time point is chosen to be a reference time point. The development of the flame can be very clearly seen in Fig. 4. The initial ignition occurs in the upper shear layer of the jet at the end of the first third of the distance L between the jet injection location and the end of the mixing section. Then the flame starts to propagate downstream along the jet shear layer. The position of the first occurrence of the ignition kernels as well as the flame front propagation time intervals correspond here very well with the experimental data of Ref. 17. At the later time points (between 1 and 2 ms) also the flame propagation upstream of the initial ignition location can be observed near to the lower wall of the mixing section. The same behaviour has also been observed in the experiments[17]. However in the experiments this phenomenon could not be described in detail because of the limited optical access in the near-wall region. A possible reason of this behaviour can be the propagation of the flame along the complex vortical structures formed in the jet in cross-flow configuration[21] and their interaction with the wall boundary layer. Here a more detailed study of the flame-vortex interaction is needed.

It should be mentioned that the steady-state RANS simulations are not able to predict the upstream near-wall flame front of the stable flame which establishes after ignition. This is clearly seen in Fig. 5. This figure compares the stable flame isosurface of

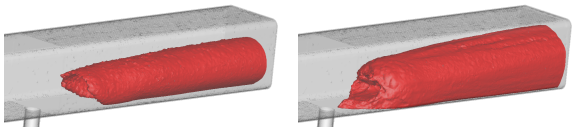
$Y_{OH} = 0.0001$ obtained in the SAS computations in the time range 2-8 ms with the isosurface of $Y_{OH} = 0.0001$ obtained in the steady-state RANS simulations. As can be seen, the form of these isosurfaces close to the lower wall of the mixing section is different. A better ability of the SAS to predict the upstream flame propagation can be explained by the good resolution of the jet in cross-flow vortical structures in such simulations[5].



(a) t = 0 ms (b) t = 0.1 ms (c) t = 0.3 ms

(d) t = 1 ms (e) t = 2 ms

Figure 4. Isosurfaces of $Y_{OH} = 0.0001$ at different time points of the SAS simulation.



(a) RANS (b) SAS

Figure 5. Stable flame isosurfaces of $Y_{OH} = 0.0001$.

4 Conclusions

A detailed numerical investigation of several flow configurations with the employment of HPC resources has been presented. The theoretical background and possible applications of CFD for the simulation of turbulent, reacting flows have been shown. The comparison with experimental data and the analysis of time-dependent data have demonstrated that this technology is able to provide valuable information on the flame stabilization mechanism. It was also shown that the application of the unsteady, scale-resolving turbulence modeling methods in conjunction with the detailed chemistry description allow to simulate such a complex phenomenon as autoignition of the H_2-rich fuels under realistic combustor conditions. Comparison of the obtained results with experimental observations made in[17] also allow the statement that the simulation represents the real flame propagation processes fairly well. This makes the simulation results very useful to complement the results

of the experimental studies. In such a "numerical experiment", the important combustion phenomena can be studied in zones not accessible by the measurement techniques thereby increasing the role of the numerical simulations in the detailed understanding of the physical processes in new combustion devices.

Acknowledgements

The authors would like to acknowledge A. Fiolitakis for the interesting discussions on laminar premixed flames. The Jülich Supercomputing Centre is also acknowledged for the computational time provided under the NIC project HSS09. Financial support has been provided by the Land Baden-Württemberg under the contract BW L 43GV within the scope of the KW21 project. The authors are grateful to the ALSTOM Power Generation AG for the encouragement of the autoignition study. We also would like to thank Julia Fleck, Michael Stöhr and Rajesh Sadanandan (Institute of Combustion Technology, DLR) for their support and discussions regarding the experimental autoignition data.

References

1. W. Lazik, Th. Doerr, S. Bake, R.v.d. Bank, and L. Rackwitz, *Development of a Lean-Burn Low-NOx Combustion Technology at Rolls-Royce Deutschland*, ASME paper GT2008-51115, 2008.
2. T. Lieuwen, V. McDonnel, E. Petersen, and D. Santavicca, *Fuel flexibility influences on premixed combustor blowout, flashback, autoignition and stability*, J. Eng. Gas Turbines Power, no. 130, 011506 1–10, 2008.
3. F. Joos, P. Brunner, B. Schulte-Werning, K. Syed, and A. Eroglu, *Development of the sequential combustion system for the ABB GT24/GT26 gas turbine family*, 1996, ASME Paper 1996-GT-315.
4. F. Güthe, J. Hellat, and P. Flohr, *The reheat concept: the proven pathway to ultralow emissions and high efficiency and flexibility*, J. Eng. Gas Turbines Power, no. 131, 021503 1–7, 2009.
5. E. Ivanova, M. Di Domenico, B. Noll, and M. Aigner, *Unsteady Simulations of Flow Field and Scalar Mixing in Transverse Jets*, ASME paper GT2009-59147, 2009.
6. M. Di Domenico, P. Gerlinger, and M. Aigner, *Development and validation of a new soot formation model for gas-turbine applications*, Combustion and Flame, 2009, in preparation.
7. F. Rebosio, M. Di Domenico, B. Noll, A. Meyer, H. Büchner, and M. Aigner, *Numerical Simulation of Flow and Combustion of a Laminar Multiple Port Burner Using Detailed Chemistry*, in: Proceedings of the European Combustion Meeting, 2007.
8. M. Di Domenico, P. Kutne, C. Naumann, J. Herzler, and M. Aigner, *Numerical and Experimental Investigation of a Semi-technical Scale Burner Employing Model Synthetic Fuels*, ASME paper GT2009-59308, 2009.
9. Y. Egorov and F. Menter, *Adwances in Hybrid RANS-LES Modelling*, chapter Development and application of SST-SAS turbulence model in the DESIDER project, pp. 261–270, Springer Verlag Beriln Heidelberg, 2008.

10. B.F. Magnussen and B.H. Hjertager, *On mathematical modelling of turbulent combustion with special emphasis on soot formation and combustion.*, in: Proceedings of the 16th International Symposium on Combustion, pp. 719–729, The Combustion Institute. 1976.
11. P. Gerlinger, B. Noll, and M. Aigner, *Assumed PDF Modeling and PDF Structure Investigation Using Finite-Rate Chemistry*, Progress in Computational Fluid Dynamics, **5**, no. 6, 334–344, 2005.
12. M. Di Domenico, C. Beck, O. Lammel, K. Werner, and B. Noll, *Experimental and Numerical Investigation of Turbulent, Lean, High-Strained, Confined, Jet Flames*, AIAA paper, 2011, submitted.
13. O. Lammel, M. Stöhr, P. Kutne, C. Dem, W. Meier, and M. Aigner, *Experimental Analysis of Confined Jet Flames by Laser Measurement Techniques*, ASME Conference Proceedings, **2011**, no. 45111, 2011.
14. M. Frenklach, H. Wang, M. Goldenberg, G.P. Smith, D.M. Golden, C.T. Bowman, R.K. Hanson, W.C. Gardiner, and V. Lissianski, *GRI-Mech—An Optimized Detailed Chemical Reaction Mechanism for Methane Combustion*, Tech. Rep. Report No. GRI-95/0058, Gas Research Institute Topical Report, November 1995.
15. D. Goodwin, "Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes", 2009.
16. G.P. Smith, D.M. Golden, M. Frenklach, N.W. Moriarty, B. Eiteneer, M. Goldenberg, C.T. Bowman, R.K. Hanson, S. Song, Jr Gardiner, W.C., V.V. Lissianski, and Z. Qin, "Gri mech 3.0", 1999, `http://www.me.berkeley.edu/gri_mech/`.
17. J. M. Fleck, P. Griebel, A. M. Steinberg, M. Stöhr, and M. Aigner, *Experimental investigation of a generic, fuel flexible reheat combustor of a gas turbine relevant operating conditions*, in: Proc. ASME Turbo Expo 2010: Power for Land, Sea and Air, Glasgow, UK, June 2010, GT2010-22722.
18. J. Fleck, P. Griebel, A. M. Steinberg, M. Stöhr, M. Aigner, and A. Ciani, *Autoignition limits of hydrogen at relevant reheat combustor operating conditions*, in: Proc. ASME Turbo Expo 2011: Power for Land, Sea and Air, Vancouver, Canada, June 2011, GT2011-46195.
19. E. Ivanova, B. Noll, P. Griebel, M. Aigner, and K. Syed, *Numerical Simulations of Turbulent Mixing and Autoignition of Hydrogen Fuel at Reheat Combustor Operating Conditions*, in: Proc. ASME Turbo Expo 2011: Power for Land, Sea and Air, Vancouver, Canada, June 2011, GT2011-46264.
20. M. O'Conaire, H.J. Curran, J.M. Simmie, R.W. Pitz, and C.G. Westbrook, *A comprehensive modeling study of hydrogen oxidation*, International Journal of Chemical Kinetics, no. 11, 602–622, 2004.
21. T.F. Fric and A. Roshko, *Vortical structure in the wake of a transverse jet*, J. Fluid Mech., **279**, 1–47, 1994.

Walberla: Simulations of Complex Flows on Supercomputers

Jan Götz[1], Stefan Donath[1], Christian Feichtinger[1], Klaus Iglberger[2], Harald Köstler[1], and Ulrich Rüde[1,2]

[1] Chair for System Simulation
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 91058 Erlangen, Germany
E-mail: {jan.goetz, stefan.donath, christian.feichtinger,ruede}@informatik.uni-erlangen.de

[2] Zentralinstitut für Scientific Computing (ZISC)
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 91058 Erlangen, Germany
E-mail: klaus.iglberger@zisc.uni-erlangen.de

In this report we will summarize our activities on the NIC systems JUGENE and JUROPA in the past two years. We will present simulation results for two highly relevant applications: the simulation of capillary effects, used to study the physical processes within a fuel cell, and simulations of fully resolved particulate flows, which are used to analyze sedimentation, fluidization and segregation processes. We will explain the important physics behind our simulations and show simulation results with up to the 294.912 cores of the full JUGENE system.

1 Introduction

The simulation of capillary effects as well as the simulation of particulate flows are crucial for the modeling of many natural phenomena and for the optimization of related industrial applications. Capillary effects are the origin of many interesting flow phenomena in micro scales such as porous media and nano tubes. They play an essential role in geological science, chemical applications and technical processes. Concrete examples are printable integrated circuits and liquid water management in the membrane of a fuel cell. The simulation of these effects is extremely difficult because most methods cannot cope with the strong forces arising in liquid–gas systems with high density ratios.

Particulate flows can be encountered in many natural processes. Important examples are sedimentation, erosion, and fluidization processes. Many of the currently established simulation methods for particulate flows, as for instance molecular dynamics or particle hydrodynamics, do not resolve the particles in the flow, but treat them as point masses without explicitly accounting for individual frictional collisions. In our approach, we fully resolve the particles in the flow, which allows a detailed analysis of the underlying phenomena.

The simulation of these applications, capillary flows and particulate flows, are very demanding in terms of compute resources. Therefore, there is a need for large and powerful computing systems and, moreover, for software frameworks that integrate the complex physics for these different applications while still sustaining high performance for various architectures. Focused on these challenging requirements is our software framework WaLBerla, which is the result of more than 30 man years of research and which is among the world-leading lattice Boltzmann CFD simulation tools. It was tested and validated on a variety of different supercomputers. Amongst them were heterogeneous GPU clusters[1], Intel

Xeon clusters like JUROPA, and also Blue Gene/P clusters like JUGENE. On all these systems WaLBerla was especially optimized and tuned for performance. In combination with our software framework *pe* for the simulation of large-scale particle scenarios, WaLBerla provides unique simulation capabilities for both capillary and particulate flows. The *pe* framework itself features a high number of numerical models in a modular and highly optimized fashion to satisfy both exactness and optimal compute performance.

In this report we will summarize the results achieved on the NIC systems for both capillary flows and particulate flows. First, the algorithms for the simulations are introduced in Sections 1.1 and 1.2. Then, Sec. 2.1 will present the scientific findings for capillary flows, whereas Sec. 2.2 will show the results for particulate flows. Besides the scientific outcomes of the simulation runs for particulate flows, also a performance study on large-scale parallel computers has been performed and is discussed within this report. Sec. 3 will conclude the report.

1.1 Free Surface Flows Simulation

The simulation of free surface flows employs a 3D lattice Boltzmann method (LBM) specifically designed for a liquid–gas system with high density ratios. It is assumed that the motion of the gas phase has a negligible influence on the liquid phase and thus, a free surface boundary condition is applied[2] at the explicitly tracked interface. Since the method is a Volume of Fluid approach, the surface information has to be reconstructed. Within this reconstruction, the local mean curvature can be computed which is taken into account by the surface tension calculation[3,4]. Capillary effects are introduced by capillary forces at the triple line, where the liquid–gas interface hits a solid wall. These capillary forces are incorporated into the free surface simulation by manipulating the surface curvature at the wall, such that the resulting forces represent the capillary forces[5]. With this model it is possible to simulate wetting effects for liquid–gas systems of high density ratios, such as water and air.

1.2 Particulate Flows

In our approach, a 3D lattice Boltzmann (LBM) fluid simulation[6,7] and a rigid body dynamics[8] simulation are dynamically coupled in order to fully resolve the motion of immersed, non-spherical particles. Both programs are fully parallelized and coupled via an explicit exchange of momenta from the fluid onto the objects[9,10], and by modeling moving boundaries on the surface of the objects to transfer momenta back to the fluid[11]. Further details on the rigid body dynamics solver and the LBM solver are described in Ref. 12 and Ref. 13, more details on the coupling can be found in Ref. 14.

2 Scientific Findings

2.1 Results: Free Surfaces Flows

The capillary flow model has been validated by two scenarios, the sessile microdrops and capillary rise and depression. Parts of the simulation scenarios have been computed on computers of the NIC.

	Microdrops			Capillary		
	Water	Oil	Mercury	Water	Oil	SF1.00
30°	2.5%	16.2%	5.6%	5%		0%
40°	9.7%	11.6%	1.6%	7%		3%
50°	4.5%	0.3%	4.1%	12%		6%
60°	4.2%	1.9%	7.8%	18%		3%
70°	2.8%	0.1%	2.7%	16%		0%
80°	0.5%	0.6%	0.1%	9%	8%	1%
90°	2.0%	1.9%	1.9%	0%	0%	0%
100°	2.9%	3.0%	2.8%	11%	31%	14%
110°	3.6%	4.1%	3.5%	11%		16%
120°	3.6%	4.8%	4.0%	11%		13%
130°	4.9%	4.4%	4.3%	8%		4%
140°	3.3%	4.7%	4.7%	2%		2%
150°	2.7%	3.2%	3.5%	1%		1%

Table 1. Relative errors of the wetting model with different liquids in the microdrop scenario and the capillary scenario for contact angles $30° < \vartheta < 150°$.

In the microdrop scenario, a small drop with radius below the capillary length is plunged onto a surface with a specific equilibrium contact angle ϑ. Since in the microscale the drop exhibits the shape of a spherical cap, the height of this cap can be analytically computed from the drop's volume and the contact angle. For evaluation of correctness, this height is compared to the equilibrium height exhibited by the simulation. Similarly, in the case of capillary rise and depression, where a tube is plunged into a liquid reservoir and the liquid is drawn upwards or pressed downwards in the tube due to its contact angle, the final equilibrium height can be analytically derived from the balance of the wetting and gravitational forces.

Tab. 1 shows, that the relative errors are below 20% for most of the regarded contact angles between $30°$ and $150°$. The validation has been carried out with different liquids that feature different properties in surface tension, viscosity and density. Fig. 1 exemplarily shows the time series of a microdrop falling on a surface with $\vartheta = 90°$ and a capillary rise scenario. Further details on the results can be obtained in a Ph. D. Thesis[4] and in reports about similar projects run on other cluster machines. Due to the shortage in the granted computing time, only a small subset of the results presented here have been computed on the computers of NIC.

The validated contact angle model has been applied to the simulation of the liquid water in the porous medium of a fuel cell. On the cathode side of the proton-exchange membrane, the protons, the electrons and the oxygen become water, which can be liquid, depending on the temperature. This liquid water blocks the pores of the porous medium

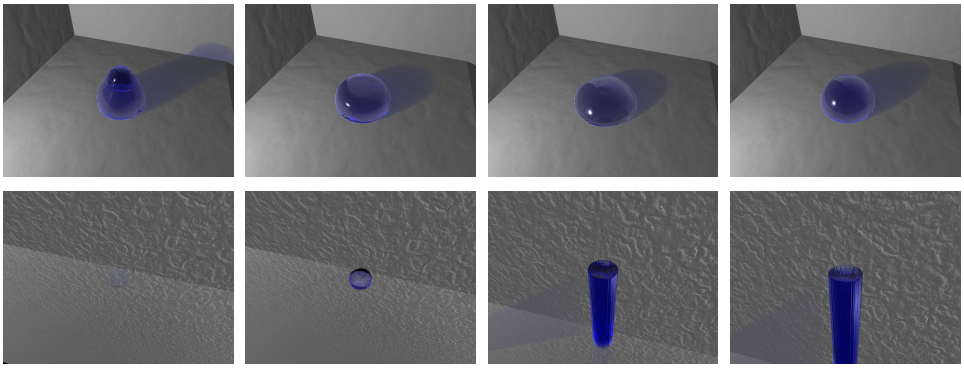


Figure 1. Simulation of aqueous microdrop plunging on a surface with $\vartheta = 90°$ (top), and capillary rise in an invisible tube with $\vartheta = 60°$ (bottom).

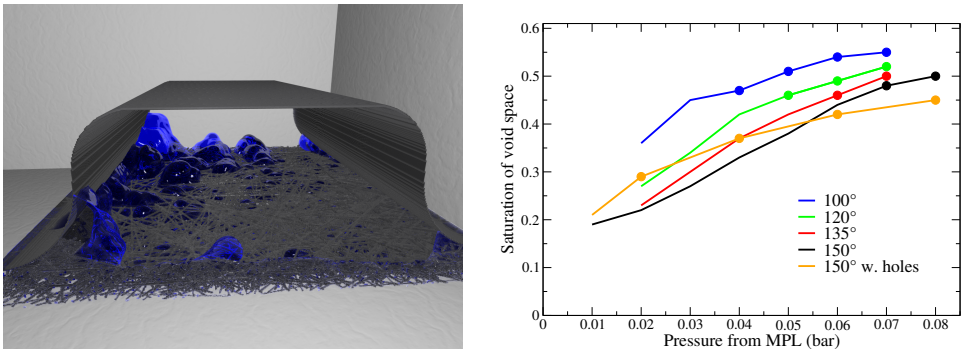


Figure 2. Simulation of liquid water in fibrous structure and gas channel of a fuel cell (left) and influence of contact angle on the saturation (right).

adjacent to the reaction zone. It is to be forced towards the gas channel, where it can be blown out of the fuel cell by the gas stream. This is why the porous medium is coated such that it is hydrophobic. However, liquid water is the source for degradation of this coating if it rests too long within the medium. Thus, an efficient liquid water management influences both the electrical performance and the duration of a fuel cell.

The simulations have been carried out within a European project called DECODE[15]. Fig. 2 shows a small excerpt of the results: First, it is evident that drop formation in the gas channel occurs at the channel wall rather than in the field, which is due to the hydrophilic coating of the walls. Second, the formation of drops depends on the contact angle and the pressure applied from below: Every simulation marked with a bullet point in the graph of Fig. 2 exhibits a drop formation on the top of the porous material. In the other cases, the capillary pressure exerted by the hydrophobicity of the material hinders the water from being evacuated. An alternation of geometry leads to significant improvement: If lasers are used to burn holes into the fibrous structure, the water is drawn out much more efficiently.

2.2 Performance Results: Particulate Flows

In this chapter we present performance and scalability results for particulate flows computed with up to 294 912 processor cores on JUGENE. To compile the program, the GNU C and C++ compiler in version 4.1.2 is used. In order to compare the performance values, we present the results in terms of *million lattice updates per second* (MLUPS). We evaluate two different scenarios, one labeled case A with sparsely packed particles, which appears in particulate flows and one case B with densely packed particles representative for sedimentation or segregation processes.

First, we evaluate the node performance of the coupled fluid particle interaction simulations, which will obviously depend on the domain size per core and the number of objects. The results are shown in Tab. 2, where the GFlop/s and the memory bandwidth values are again measured using the automatically available performance counters[16].

For the smaller domain size of 40^3 lattice cells per core, the performance is in general lower than for domain size of 80^3 lattice cells due to a larger communication to computation ratio. When the number of rigid bodies is increased (case B), the computations for mapping the objects to the grid, the force evaluation, the movement and the collisions of the objects are higher compared to simulations with smaller number of objects (case A), resulting in a lower performance in terms of MLUPS for case B. Thus, the highest MLUPS value is obtained for a domain size of 80^3 lattice cells and case A. For the same domain size and the same test case, also the highest measured values for GFlop/s and main memory bandwidth are achieved.

Test case	Lattice domain size per core	MLUPS	% of peak GFlop/s performance	% of peak memory bandwidth
A	40^3	3.0	4.12	14.61
	80^3	3.24	**4.59**	**22.32**
B	40^3	1.9	3.0	11.18
	80^3	1.92	3.15	15.63

Table 2. Node performance of coupled fluid structure interaction simulations.

In a second step the multi node performance of coupled fluid structure interaction simulations is evaluated. The weak scaling is performed with both scenarios and two different domain sizes. The experiment is started at 64 cores with a decomposition of $4 \times 4 \times 4$, when 8 processes communicate in all neighbouring directions. Note that placement of MPI processes needs to be done according to the requested torus shape and thus the underlying network shape on the system to maintain a high performance.

The weak scaling up to 294 912 cores of case A and case B is presented in Fig. 3. For case A, the reference values for calculating the efficiency are 44.07 MLUPS for a domain size of 40^3 lattice cells per core and 47.31 MLUPS for a domain size of 80^3 lattice cells per core for measurements with 64 cores. The lower efficiency for the smaller domain size results from a higher communication to computation ratio. The efficiency for case B is based on measurements on 64 cores, which result in 23.06 MLUPS and 25,14 MLUPS for domain sizes of 40^3 and 80^3 lattice cells per core, respectively. Over the whole range of cores, the efficiency remains over 98% for a domain sizes of 80^3 lattice cells per core. Again, the efficiency of the simulation with smaller domain size is lower, compared to the

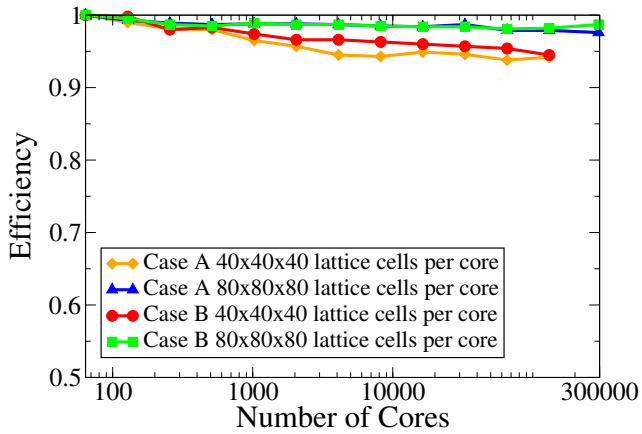


Figure 3. Weak scaling from 64 to 294 912 compute cores for sparsely packed (case A) and densely packed particles (case B).

simulation with larger domain size due to the communication to computation ratio. On 294 912 compute cores with case B, 264 331 905 rigid objects are simulated in a fluid grid with 150 994 944 000 lattice cells. To the best of our knowledge, this is currently the largest simulation of fully resolved objects in the flow.

Fig. 4 illustrates a simulation of a real world scenario. Objects with density values of $0.8\,kg/dm^3$ and $1.2\,kg/dm^3$ are immersed in water with density $1\,kg/dm^3$ and a gravitation field. Thereby light objects will rise to the top, while heavy objects will fall to the ground.

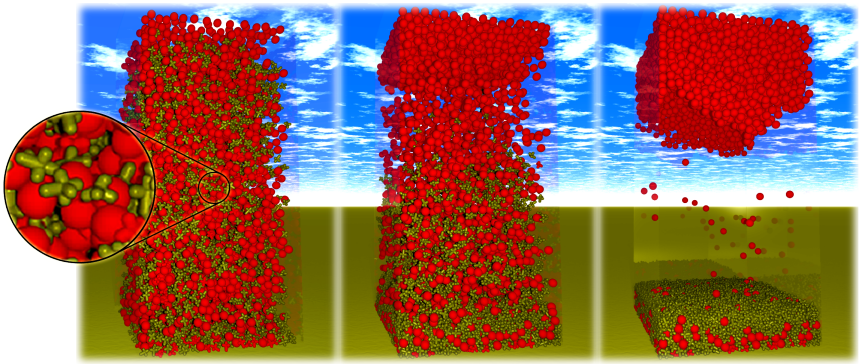


Figure 4. Segregation simulation of 12 013 objects with two different, non-spherical shapes in different time steps simulated on 2 048 cores. Density values of $0.8\,kg/dm^3$ and $1.2\,kg/dm^3$ are used for the objects in water with density $1\,kg/dm^3$ and a gravitation field. Light particles are rising to the top of the box, while heavy particle fall to the bottom.

3 Conclusion

In this report we have demonstrated two massively parallel simulation scenarios of scientifically important applications. We have shown the simulation of capillary effects that we use for a detailed analysis of the physical processes within a fuel cell. To a considerable amount the simulation runs were conducted on NIC systems and could contribute to the understanding of the underlying physics and with that to the improvement of the manufacturing process. Additionally we presented fully resolved particulate flow results of simulations with an unprecedented level of detail. Due to architecture-specific optimizations and the development of scalable algorithms, especially in the context of rigid body dynamics simulations, the scaling behaviour of these simulations is excellent up to the maximum number of cores on the JUGENE system. These results were also accepted for publication at the supercomputing conference 2010.

Acknowledgements

The work was supported by the Kompetenznetzwerk für Technisch-Wissenschaftliches Hoch- und Höchstleistungsrechnen in Bayern (KONWIHR) under project waLBerlaMC, the European Union (EU) project DECODE (grant 213295) and the Bundesministerium für Bildung und Forschung under the SKALB project no. 01IH08003A.

We would like to thank for their support: Bernd Mohr, Wolfgang Frings and Lukas Arnold, Jülich Supercomputing Centre (JSC).

We are grateful to the Jülich Supercomputing Centre for providing the computational resources on JUROPA and JUGENE.

References

1. C. Feichtinger, J. Habich, H. Köstler, G. Hager, U. Rüde, and G. Wellein, *A flexible Patch-based lattice Boltzmann parallelization approach for heterogeneous GPU-CPU clusters*, Parallel Computing, **37**, no. 9, 536 – 549, 2011.
2. C. Körner, M. Thies, and R. F. Singer, *Modeling of metal foaming with Lattice Boltzmann automata*, Advanced Engineering Materials, **4**, 765–769, 2002.
3. C. Körner, M. Thies, T. Hofmann, N. Thürey, and U. Rüde, *Lattice Boltzmann Model for Free Surface Flow for Modeling Foaming*, Journal of Statistical Physics, **121**, no. 1–2, 179–196, 2005.
4. T. Pohl, *High performance simulation of free surface flows using the lattice Boltzmann method*, PhD thesis, Jul 2008.
5. S. Donath, K. Mecke, S. Rabha, V. Buwa, and U. Rüde, *Verification of surface tension in the parallel free surface lattice Boltzmann method in waLBerla*, Computers & Fluids, **45**, no. 1, 177–186, 2011.
6. S. Succi, *The Lattice Boltzmann Equation - For Fluid Dynamics and Beyond*, Clarendon Press, 2001.
7. S. Chen and G. D. Doolen, *Lattice Boltzmann method for fluid flows*, Annu. Rev. Fluid Mech., **30**, 329–364, 1998.
8. M. Anitescu, *Optimization-based simulation of nonsmooth rigid multibody dynamics*, Math. Program., **105**, no. 1, 113–143, 2006.

9. A. J. C. Ladd, *Numerical simulations of particulate suspensions via a discretized Boltzmann equation. Part 1. Theoretical foundation.*, J. Fluid Mech., **271**, 285–309, 1994.
10. A. J. C. Ladd, *Numerical simulations of particulate suspensions via a discretized Boltzmann equation. Part 2. Numerical results.*, J. Fluid Mech., **271**, 311–339, 1994.
11. D. Yu, R. Mei, L.-S. Luo, and W. Shyy, *Viscous flow computations with the method of lattice Boltzmann equation*, Prog. Aerosp. Sci., **39**, no. 5, 329–367, 2003.
12. K. Iglberger and U. Rüde, *Massively Parallel Rigid Body Dynamics Simulation*, Comp. Sci. - Res. Dev., **23**, no. 3, 159, 2009.
13. C. Feichtinger, J. Götz, S. Donath, K. Iglberger, and U. Rüde, *Concepts of waLBerla Prototype 0.1*, Tech. Rep. 07–10, 2007.
14. K. Iglberger, N. Thürey, and U. Rüde, *Simulation of moving particles in 3D with the Lattice Boltzmann method*, Comput. Math. Appl., **55**, no. 7, 1461–1468, 2008.
15. DECODE, "Understanding of degradation mechanisms to improve components and design of PEFC", `http://www.decode-project.eu`, 2008–2011, CORDIS project 213295, funded by the European Commission within the Seventh Framework Programme.
16. G. Lakner, I-H. Chung, G. Cong, S. Fadden, N. Goracke D. Klepacki, J. Lien, C. Pospiech, S. R. Seelam, and H.-F. Wen, *IBM System Blue Gene Solution: Performance Analysis Tools*, pp. 29–35, 2008, IBM order number REDP-4256-01.

A Hybrid Aeroacoustic Prediction Method for Non-Reactive and Reactive Flows

Georg Geiser, Matthias Meinke, and Wolfgang Schröder

Institute of Aerodynamics, RWTH Aachen University, Wüllnerstraße 5a, 52062 Aachen, Germany
E-mail: {g.geiser, m.meinke, office}@aia.rwth-aachen.de

The accurate prediction of flow-induced noise is a challenging task even on modern supercomputers. This is mainly due to the large discrepancy in amplitude and time and length scales of flow phenomena like turbulent eddies and acoustic waves. A hybrid prediction method for aeroacoustic flow problems is presented that allows to restrict the more time-consuming simulation of the sound generating flow field to the vicinity of the acoustic source while the farfield sound is computed in a second step. A set of governing equations that accounts for linear sound propagation including convection and refraction effects is presented. In a sense of an aeroacoustic analogy the corresponding source terms exciting those equations are closely connected to the sound-generating mechanisms characterizing the flow problem. Two applications are presented, a coaxial multi-species helium/air jet and a lean-premixed preheated Bunsen flame.

1 Introduction

The noise generated by flow phenomena is highly relevant for many real-life problems, e.g., the noise emitted from jet engines[1–3] and the noise generated by unsteady combustion processes[4–6] that can induce instabilities resulting in a lower efficiency, higher pollution or even failure of combustor devices. Despite intensive aeroacoustic research since the 1940s the knowledge on the corresponding noise generating mechanism is not complete, yet. This is not astonishing since such flows are a highly intricate problem which is characterized by a large scale of involved length and time scales, multiple shear layers, mixing processes, density and/or temperature gradients, inflow conditions, boundary conditions, and so forth.

Since for most realistic flow problems the acoustic length and time scales restrict the size of computable aeroacoustic problems even on today's most advanced supercomputers, a hybrid approach has been developed[7,8] to limit the flow simulation to the direct vicinity of the acoustic source region, whereas specially derived acoustic equations are used to simulate the acoustic radiation including convection and refraction effects. Furthermore, due to the low amplitude of the acoustic perturbations and the large transport distances to the farfield observer positions carefully selected numerical methods have to be applied to ensure that the results are not spoiled by spurious numerical acoustics.

The paper is organized as follows. First the governing equations and the used numerical methods are briefly discussed. Then, results for two different flow cases are presented, a coaxial helium/air mixture jet, which was selected to investigate the impact of multi-species mixture processes on the acoustic source terms, and preliminary results for a premixed preheated Bunsen-type flame at high turbulence levels that will help to gain more insight into the thermoacoustic noise generation process. The paper concludes with some remarks on the parallel efficiency of the involved algorithms and a final discussion.

2 Governing Equations and Numerical Methods

The acoustic field is computed using the acoustic perturbation equations[7] (APE) with source terms sampled from compressible large eddy simulations (LES). Results of the acoustic fields have been computed using the APE-4 formulation with revised acoustic sources[9]

$$\frac{\partial p^a}{\partial t} + \overline{c^2}\boldsymbol{\nabla}\cdot\left(\overline{\rho}\boldsymbol{v}^a + \overline{\boldsymbol{v}}\frac{p^a}{\overline{c^2}}\right) = \overline{c^2}q_c \tag{1}$$

$$\frac{\partial \boldsymbol{v}^a}{\partial t} + \boldsymbol{\nabla}\left(\overline{\boldsymbol{v}}\cdot\boldsymbol{v}^a\right) + \boldsymbol{\nabla}\left(\frac{p^a}{\overline{\rho}}\right) = \boldsymbol{q_m}\,, \tag{2}$$

where p^a and $\boldsymbol{v}^a$ are the acoustic perturbation pressure and velocity, respectively, and q_c and $\boldsymbol{q_m}$ are the source terms given by

$$q_c = -\boldsymbol{\nabla}\cdot\left(\rho'\boldsymbol{v}'\right)' - \frac{\partial \rho_e}{\partial t} - \boldsymbol{\nabla}\cdot\left(\overline{\boldsymbol{v}}\rho_e\right) \tag{3}$$

$$\boldsymbol{q_m} = -\underbrace{\left(\left(\boldsymbol{\omega}'\times\overline{\boldsymbol{v}}\right) + \left(\overline{\boldsymbol{\omega}}\times\boldsymbol{v}'\right) + \left(\boldsymbol{\omega}'\times\boldsymbol{v}'\right)'\right)}_{=(\boldsymbol{\omega}\times\boldsymbol{v})'} - (\boldsymbol{\nabla}k)' \tag{4}$$

$$+\underbrace{\left(\boldsymbol{\nabla}\left(\frac{p'}{\overline{\rho}}\right) - \left(\frac{\boldsymbol{\nabla}p}{\rho}\right)'\right)}_{=\Delta p} + \underbrace{\left(\frac{\boldsymbol{\nabla}\cdot\underline{\boldsymbol{\tau}}}{\rho}\right)'}_{\text{neglected}}$$

with the excess density $\rho_e = \rho' - \frac{p'}{\overline{c^2}}$ and the instantaneous turbulent kinetic energy $k = \frac{\boldsymbol{v}'\cdot\boldsymbol{v}'}{2}$. Overbars indicate time-averaged values that constitute the underlying base flow that causes convection and refraction of acoustic waves while primes indicate perturbation values. The perturbed Lamb vector $-(\boldsymbol{\omega}\times\boldsymbol{v})'$ has been split into linear and non-linear parts for the coaxial jet case presented. Due to the high Reynolds number of the flow problem the viscous source terms can be neglected. The original APE-4 source term formulation used a first-order formulation of Gibbs' fundamental law of thermodynamics

$$\overline{\left(\frac{\rho}{c_p}\right)}s' \approx \frac{p'}{\overline{c^2}} - \rho' = -\rho_e \tag{5}$$

that was found to be too restrictive for the correct evaluation of entropy related sources. The formulation in terms of the excess density is a rigorous reformulation of the governing Navier-Stokes equations, resulting in correct prediction of radiated acoustics. Furthermore, the original formulation did not account for the chemical potential of the gas mixture that varies for reacting flows.

The time integration is accomplished using an alternating 5-6 stage low-dispersion low-dissipation Runge-Kutta (LDDRK) scheme[10] and the space derivatives are determined explicitly using a 7-point 4th-order dispersion-relation preserving (DRP) scheme[11]. Numerical stability is obtained by explicit filtering of the dependent variables. The described method has been successfully applied to several aeroacoustic problems including airframe noise[12,13], combustion noise of incompressible diffusion flames[8,14], and jet noise[15,16].

The acoustic domains contain virtual microphones measuring the acoustic perturbation pressure p^a (cf. Fig. 1). The microphones are arranged on multiple circles that form

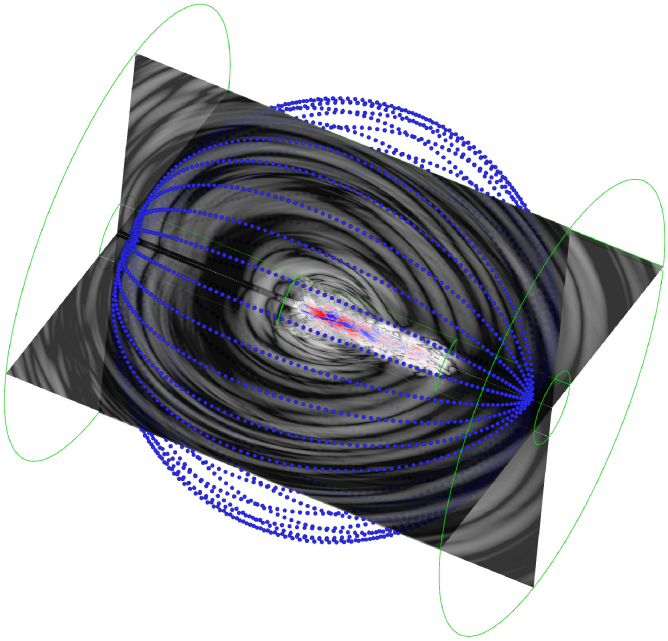


Figure 1. Acoustic domain equipped with virtual microphones given as blue spheres. Green lines indicate the edges of the sub-domains. The slices show an instantaneous snapshot of the acoustic perturbation pressure p^a.

spheres to acquire sufficient data for statistic averaging of the overall sound pressure level (OASPL) and power spectral density (PSD) results. Since the radiated acoustic field is azimuthally isotropic in a statistical sense this approach is justified. The spheres are centred at the main source region, i.e., the end of the jet's potential core and in the vicinity of the flame tip, respectively.

3 Computational Results

3.1 Coaxial Helium/Air Jet

The LES dataset involving injection of a helium/air mixture is computed by an implicit eddy-diffusivity approach according to Fureby and Grinstein[17]. The details of the general setup of the LES method are given by Meinke *et al.*[18] The LES is based on a finite-volume method, in which the Euler terms are spatially discretized by the AUSM scheme using upwind biased formulations and the MUSCL approach and a centred approximation for the pressure term. The approximation is second-order accurate. The non-Euler terms are approximated by a centred discretization. For the temporal integration an explicit five-step Runge-Kutta formulation is used. To eliminate spurious waves on the boundaries, a sponge layer is imposed[19]. The mean inflow condition of the LES is modelled to have a hyperbolic tangent profile. To ensure a turbulent inflow distribution, Bogey and Bailly's[20] artificial forcing based on several azimuthal modes is used. The capability of the LES code to compute non-reacting gas species flows is shown by Renze *et al.*[21]

A gas mixture of oxygen (O_2) and nitrogen (N_2) similar to air is used as coflow, while a jet with a 30% mass fraction of helium (He) in the primary stream ($0 \leq r \leq \frac{D}{2}$) is injected. The Reynolds number based on the diameter and the kinematic viscosity of the secondary stream, the Mach number of the primary and the secondary jet, and the density ratio of the primary and the ambient fluid are given in Tab. 1.

Re	Ma_p	Ma_s	s
40 000	1.0	0.9	0.37

Table 1. Parameters of the present simulations (Reynolds number $Re = \rho_s u_s D_s/\mu_s$, Mach number $Ma_p = u_p/a_\infty$, $Ma_s = u_s/a_\infty$, density ratio $s = \rho_p/\rho_\infty$). The subscript $_p$ denotes the primary and $_s$ the secondary jet stream. The index $_\infty$ indicates ambient values.

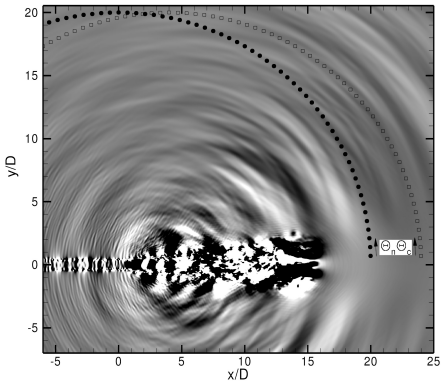


Figure 2. Virtual microphone positions for the nozzle exit (•, Θ_n) and the potential core (□, Θ_c) half circle.

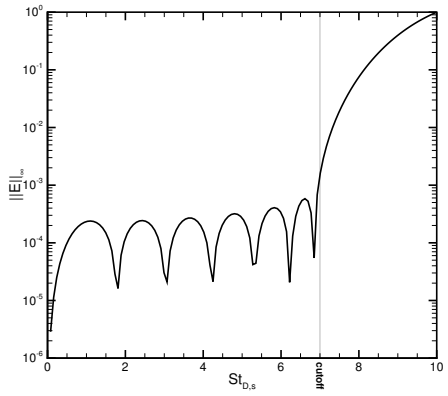


Figure 3. Infinity norm of the error induced by the interpolating FIR filter during source term reconstruction.

The positions of the virtual microphones displayed in Fig. 2 are centred at the breakdown of the jet's potential core (Θ_c). During the aeroacoustic simulation the source terms are reconstructed by a least-squares interpolating filter that guarantees a high accuracy up to $St_{D,s} = 7$ (cf. Fig. 3)

Figs. 4 and 5 display the overall sound pressure level (OASPL) for the single source components of the APE-4 equations and the result for the complete source formulation. The curves clearly indicate that the low radiation angles are dominated by the perturbed Lamb vector source $-(\boldsymbol{\omega} \times \boldsymbol{v})'$, while its OASPL significantly drops towards the sideline direction and finally rises again in the upstream radiation direction. For the upstream angles, there is a cancellation of the Lamb vector acoustics by the other source components. The perturbed Lamb vector source split into its linear components reveals the linear components $-\boldsymbol{\omega}' \times \overline{\boldsymbol{v}}$ and $-\overline{\boldsymbol{\omega}} \times \boldsymbol{v}'$ to almost solely contribute to the forward radiation, where the latter part is more dominant. At increasing polar angle the two linear Lamb vector

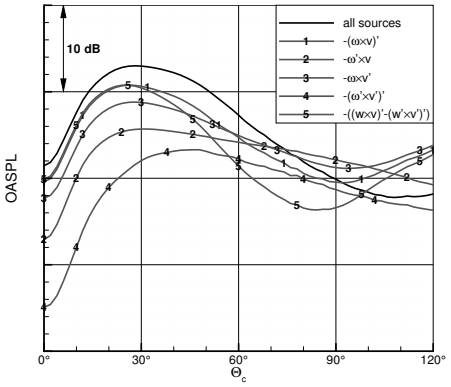


Figure 4. Averaged overall sound pressure level at standard atmospheric conditions obtained by the potential core virtual microphone circles for the coaxial helium/air jet. The Lamb vector source is split into linearized components.

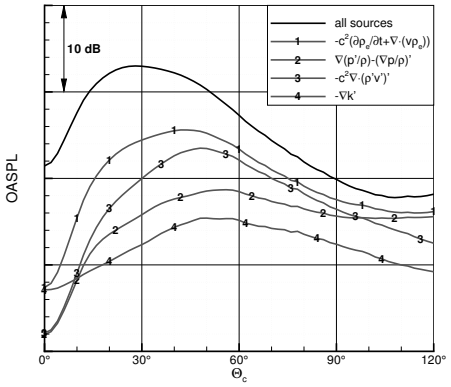


Figure 5. Averaged overall sound pressure level at standard atmospheric conditions obtained by the potential core virtual microphone circles for the coaxial helium/air jet. Several source terms listed in the insert are shown.

components show a strong cancellation (indicated by $-\left((\boldsymbol{\omega}\times\boldsymbol{v})'-(\boldsymbol{\omega}'\times\boldsymbol{v}')'\right)$) while the nonlinear part $-(\boldsymbol{\omega}'\times\boldsymbol{v}')'$ is most important. The excess density source in Fig. 5 radiates at a higher angle while contributing less energy to the overall peak radiation. For the sideline direction it becomes a dominant source. Here, also the remaining nonlinear sources are more relevant, while the value of the perturbed turbulent kinetic energy source $-(\boldsymbol{\nabla}k)'$ was found to be much smaller throughout the entire range.

3.2 Premixed Preheated Bunsen Flame

The combustion model used by the compressible large eddy simulation (LES) of the Bunsen flame by Zhang *et al.*[22] is an extension of the turbulent flame speed closure model (TFC)[23] with respect to LES and non-premixed combustion processes while taking detailed chemistry into account. The turbulent flame speed captures the interaction between the turbulent flow and the flame front and depends on thermophysical and turbulence parameters. It constitutes a source for the progress variable of the chemical reaction, since the TFC model states a direct relation to the chemical reaction rate. The simulation has been conducted using the open-source flow solver OpenFoam with an implicit second-order accurate finite-volume discretization. The computational domain for the LES contains 6.8 million cells. A turbulence generator[24] has been used to impose turbulent fluctuations based on the measured turbulence intensity and length scale onto the inlet velocity profile. Non-reflecting boundary conditions are enforced on the outflow boundaries to prevent spurious reflections of acoustic waves.

For the present study a lean-premixed preheated Bunsen-type flame[25] is computed. The burner is characterized by a rather homogeneous flow in the burner jet while the turbulence intensity and related volumetric heat release rates are at a high level. The parameters of the experimental setup and the flow problem are given in Tab. 2.

D	Re_D	Ma_u	Ma_∞	T_u	T_∞	λ	P
150 mm	41 000	0.046	0.069	673 K	300 K	1.75	260 kW

Table 2. Parameters of the present experiment (Reynolds number $Re = \rho_u u_u D/\mu_u$, Mach number $Ma_u = u_u/a_u$, $Ma_\infty = u_u/a_\infty$, temperature T_u, T_∞, air ratio λ, thermal power P). The subscript $_u$ denotes the unburnt gas and $_\infty$ the ambient fluid.

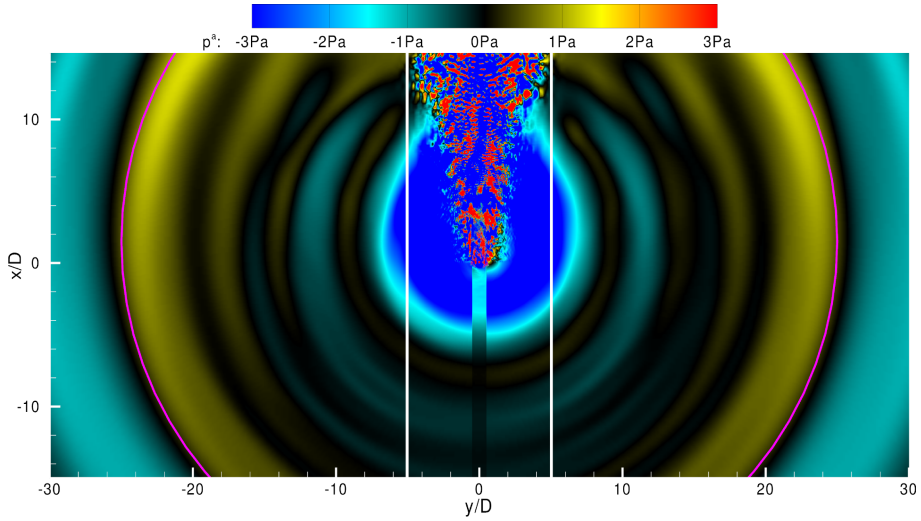


Figure 6. Instantaneous contour plot of the acoustic perturbation pressure p^a for the Bunsen flame. The white lines at $y/D \approx \pm 5$ display the extent of the source region while the pink circle indicates the position where the virtual microphones for the evaluation of the results have been placed.

Fig. 6 depicts an instantaneous contour plot of the acoustic perturbation pressure emitted by the investigated Bunsen flame. The major acoustic source region is located in the vicinity of the flame tip at $x \approx 1.5\,D$ while the flame brush generates no noticeable sound. This behaviour is expected, since the strong turbulent fluctuations induced by the breakdown of the potential core of the incoming jet flow results in heavy oscillations of the local reaction rate of the flame front of the premixed flame. This strong entropy source constitutes the dominant acoustic source for the present flow case and can also be interpreted as an unsteady heat release or volume expansion. The farfield sound shows a slight directional behaviour. The inhomogeneous speed of sound, caused by a strongly varying density and heat capacity ratio, in combination with the refraction due to velocity gradients induce a visible higher amplitude for the downstream facing direction. However, the refraction effects due to velocity gradients is probably quite small since the Mach number of the present case is very low.

The overall sound pressure level (OASPL) in Fig. 7 that was evaluated from the farfield data obtained by the virtual microphones located at $r = 25\,D$ from the flame tip supports the observations made for the instantaneous contour plot of the acoustic perturbation pressure p^a. A downstream-biased directivity is visible with a nearly linear slope (in semi-log

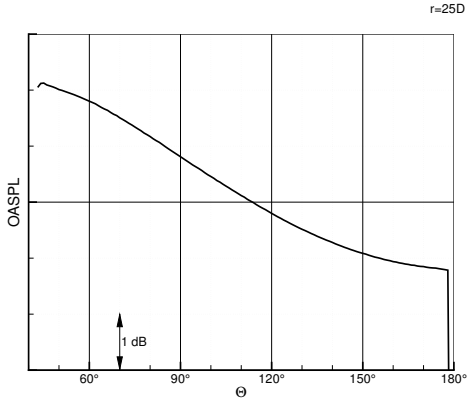


Figure 7. Overall sound pressure level (OASPL) for the Bunsen flame containing all APE-4 sources as measured by the farfield virtual microphones at a distance of 25 nozzle exit diameters from the flame tip.

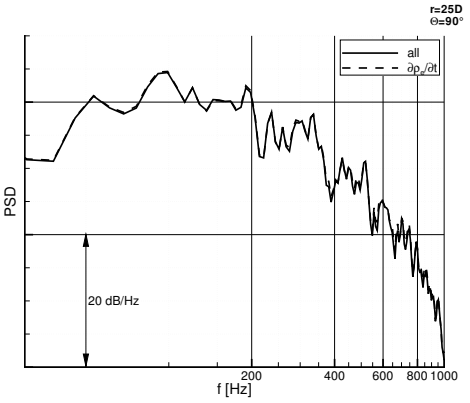


Figure 8. Power spectral density (PSD) estimate for the Bunsen flame as measured by the farfield virtual microphones at a distance of 25 nozzle exit diameters from the flame tip. The results for all acoustic sources and for the partial time derivative of the excess density are shown.

plotting) in the sideline direction ($\Theta = 90°$). The offset for the investigated polar angle range is approximately $3\,\mathrm{dB}$.

The power spectral density (PSD) estimate for the sideline direction ($\Theta = 90°$) at a distance of $r = 25\,D$ from the flame tip is shown in Fig. 8. It indicates a broadband spectral behaviour of the emitted acoustic field, peaking at $f \approx 90\,\mathrm{Hz}$ with a nearly linear decay towards the high frequencies in this double-log plot. Note that for the present flow case a strongly simplified source term formulation can be used. Results obtained solely by the partial time derivative of the excess density shown in the same plot are nearly identical with the result obtained by the entire source formulation. By Gibbs' fundamental law of thermodynamics this source is closely related to fluctuations of entropy and chemical potential of the flow thus representing major parts of the thermoacoustic source mechanism. Refraction and convection effects are mostly captured by the formulation of the APE operator such that the amplitude alteration by those effects is inherently contained in the simulation.

4 Parallel Efficiency

The aeroacoustic solver PIANO used for this research performs reasonably well on massively-parallel supercomputers like the IBM Blue Gene/P. This is indicated by a strong scaling for a 256^3 grid given in Fig. 9 and a weak scaling in Fig. 10 for 16^3 and 32^3 block sizes, where the latter one is the practically more relevant case. In practice the computations usually used at most 4096 cores (1 rack) and therefore are in a near-linear speedup range.

Parallel collective I/O is used to handle the large amount of data that is required for the source terms of the present approach. All processes share one global file so the file system overhead generated by many tiny files per process is avoided.

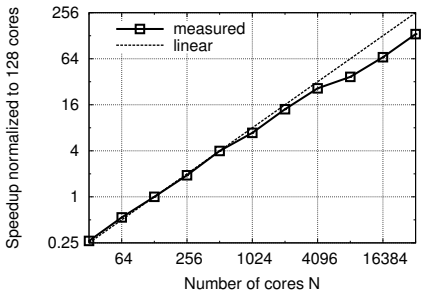


Figure 9. Parallel speedup in a strong scaling of the aeroacoustic solver PIANO on IBM Blue Gene/P. 128 cores have been used as the reference for scaling.

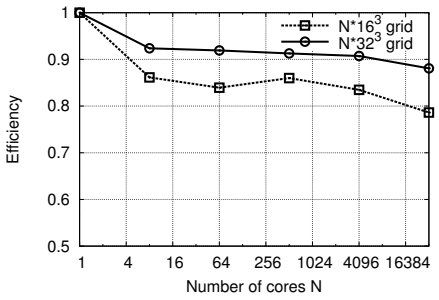


Figure 10. Parallel efficiency in a weak scaling of the aeroacoustic solver PIANO on IBM Blue Gene/P. The sequential code has been used as reference.

5 Conclusion

A hybrid aeroacoustic simulation method has been presented that allows to efficiently and accurately simulate the farfield acoustic characteristics of various flow-induced acoustic problems, including non-reactive and reactive flows. The capabilities of the method have been demonstrated for a non-reactive multi-species coaxial helium/air jet at a high Mach number and a lean-premixed preheated Bunsen flame at a low Mach number. Different source mechanisms for the two cases have been identified. The simulation method has been successfully adopted to run on massively-parallel supercomputers including parallel collective I/O.

Acknowledgements

The calculations were carried out at the Jülich Supercomputing Centre.

References

1. C. K. W. Tam, K. Viswanathan, K. K. Ahuja, and J. Panda, *The sources of jet noise: experimental evidence*, Journal of Fluid Mechanics, **615**, 253–292, 2008.
2. K. Viswanathan, *Aeroacoustics of hot jets*, J. Fluid Mech., **516**, 39–82, 2004.
3. J.B. Freund, *Noise sources in a low-Reynolds-number turbulent jet at Mach 0.9*, J. Fluid Mech., **438**, 277–305, 2001.
4. T. Lieuwen, *Modeling Premixed Combustion-Acoustic Wave Interactions: A Review*, Journal of Propulsion and Power, **19**, no. 5, 765–781, 2003, doi:10.2514/2.6193.
5. S. Candel, D. Durox, S. Ducruix, A.-L. Birbaud, N. Noiray, and T. Schuller, *Flame dynamics and combustion noise: progress and challenges*, International Journal of Aeroacoustics, **8**, no. 1 & 2, 1–56, 2009, doi:10.1260/147547209786234984.
6. A. Schwarz and J. Janicka, (Eds.), *Combustion Noise*, Fluid mechanics and its applications. Springer, 2009.
7. R. Ewert and W. Schröder, *Acoustic Perturbation Equations Based on Flow Decomposition via Source Filtering*, Journal of Computational Physics, **188**, 365–398, 2003.

8. T. P. Bui, W. Schröder, and M. Meinke, *Numerical analysis of the acoustic field of reacting flows via acoustic perturbation equations*, Computers & Fluids, **37**, 1157–1169, 2008.
9. G. Geiser, S. R. Koh, and W. Schröder, *Analysis of acoustic source terms of a coaxial helium/air jet*, in: 17th AIAA/CEAS Aeroacoustics Conference, 2011, AIAA-Paper 2011-2793.
10. F.Q. Hu, M.Y. Hussaini, and J.L. Manthey, *Low-Dissipation and Low-Dispersion Runge-Kutta Schemes for Computational Acoustics*, J. Comput. Phys., **124**, 177–197, 1996.
11. C. K. W. Tam and J. C. Webb, *Dispersion-relation-preserving finite difference schemes for computational acoustics*, J. Comput. Phys., **107**, 262–281, 1993.
12. R. Ewert and W. Schröder, *On the Simulation of Trailing Edge Noise with a Hybrid LES/APE Method*, Journal of Sound and Vibration, **270**, 509–524, 2004.
13. D. König, S.R. Koh, M. Meinke, and W. Schröder, *Two-step simulation of slat noise*, Computers & Fluids, **39**, no. 3, 512–524, 2010.
14. T. P. Bui, M. Ihme, W. Schröder, and M. Meinke, *Analysis of different sound source formulations to simulate combustion generated noise using a hybrid LES/APE-RF method*, International Journal of Aeroacoustics, **8**, 95–124, 2009.
15. E. R. Gröschel, W. Schröder, P. Renze, M. Meinke, and P. Comte, *Noise Prediction for a Turbulent Jet Using Different Hybrid Methods*, Computers & Fluids, **37**, no. 4, 414–426, 2008.
16. S. R. Koh, W. Schröder, and M. Meinke, *Turbulence and heat excited noise sources in single and coaxial jets*, Journal of Sound and Vibration, **329**, 786–803, 2010.
17. C. Fureby and F. Grinstein, *Monotonically integrated large eddy simulation of free shear flows*, AIAA Journal, **37**, 544–556, 1999.
18. M. Meinke, W. Schröder, E. Krause, and T. Rister, *A comparison of second- and sixth-order methods for large-eddy simulations*, Comput. Fluids, **31**, 695–718, 2002.
19. J. Freund, *A proposed inflow/outflow boundary condition for direct computation of aerodynamic sound*, AIAA Journal, **35**, no. 4, 740–742, 1997.
20. C. Bogey and C. Bailly, *Effects of inflow conditions and forcing on subsonic jet flows and noise*, AIAA Journal, **43**, no. 5, 1000–1007, 2005.
21. P. Renze, W. Schröder, and M. Meinke, *Large Eddy Simulation of Film Cooling Flows at Density Gradients*, International Journal of Heat & Fluid Flow, **29**, no. 1, 18–34, 2008.
22. F. Zhang, P. Habisreuther, M. Hettel, and H. Bockhorn, *Numerical Computation of Combustion Induced Noise Using Compressible LES and Hybrid CFD/CAA Methods*, Acta Acustica united with Acustica, **98**, 2012.
23. Hans-Peter Schmid, Peter Habisreuther, and Wolfgang Leuckel, *A Model for Calculating Heat Release in Premixed Turbulent Flames*, Combustion and Flame, **113**, no. 1-2, 79–91, 1998, doi:10.1016/S0010-2180(97)00193-4.
24. M. Klein, A. Sadiki, and J. Janicka, *A digital filter based generation of inflow data for spatially developing direct numerical or large eddy simulations*, Journal of Computational Physics, **186**, no. 2, 652–665, 2003, doi:10.1016/S0021-9991(03)00090-1.
25. M. Zajadatz, M. Hettel, and W. Leuckel, *Burning Velocity of High-Turbulence Natural Gas Flames for Gas Turbine Application*, in: International Gas Research Conference, pp. 793–803, 1998.

Impact of the Heterogeneity of Random Sphere Packings on Hydrodynamic Dispersion

Siarhei Khirevich and Ulrich Tallarek

Department of Chemistry, Philipps-Universität Marburg, 35032 Marburg, Germany
E-mail: {khirevic, tallarek}@staff.uni-marburg.de

We generate periodic random sphere packings of equal spheres and simulate advection–diffusion transport in the pore space to determine hydrodynamic dispersion coefficients. We employ the Jodrey–Tory algorithm to create sphere packings with various porosities and microstructures. We demonstrate that both porosity and packing microstructure may have a significant influence on the hydrodynamic dispersion coefficients. Further, we analyze pore space of each generated packing with Voronoi spatial tessellation and show a good correlation between a scalar geometrical descriptor, derived from the tessellation, and the corresponding hydrodynamic dispersion coefficients.

1 Introduction

Transport processes (such as diffusion and advection) in porous media occur in many diverse fields of science and engineering, and accurate prediction and optimization of the transport processes requires understanding, both qualitative and quantitative, of the underlying physical phenomena. In our study we are concerned with simulation of transport in computer-generated random sphere packings. Despite sphere packings are a simple example of porous media (compared to soil or rocks), they are a good model for the real-life packings of close-to-spherical objects, which are, in turn, experimentally available and of large practical importance.

In our study a transport coefficient of primary interest is the *hydrodynamic dispersion coefficient*, which can be understood as follows: a substance carried by the flow in the pore space of a sphere packing is spread around its mean position due to molecular diffusion and maldistribution of the flow field (assuming that substance is inert (i.e, nonreacting and nonadsorbing) and spheres are impermeable for the substance), and the rate of such spreading is quantified by the hydrodynamic dispersion coefficient. Depending on the application field, higher or lower values of the dispersion coefficient can be preferred: for example, in the field of chromatography, which deals with the separation of mixtures, transport of a mixture is typically performed in columns (or tubes) filled with spherical particles, and columns demonstrating higher dispersion coefficients should be avoided due to their poor separation efficiency.

Traditionally, the hydrodynamic dispersion coefficient in bulk (periodic along three spatial dimensions) sphere packings is assumed to be dependent on the porosity (void space fraction) of the packing and the average flow rate of a carrier fluid[1]. In the current work we demonstrate that an additional parameter, related to the packing microstructure, does have a significant impact on the hydrodynamic dispersion coefficient. For this purpose, we generate packings with various values of porosity and microstructure, and perform carefully conducted high-resolution simulations of hydrodynamic dispersion in the void space of the packings.

2 Numerical Simulation Approach

Our numerical approach includes the following consequent steps: i) generation of a random sphere packing, ii) simulation of fluid flow in the void space of the packing, and iii) simulation of advection–diffusion transport of small (compared to the sphere diameter) inert tracers in the packing void space. Further we provide a brief description of these steps.

2.1 Random Sphere Packing Generation

Random packings of monosized spheres were generated using Jodrey–Tory algorithm (JT)[2]. JT starts from a random distribution of N_{sp} sphere centres in a simulation box of volume V_{box}. Obviously, sphere overlap is typical in the initial configuration if the sphere packing to be generated is dense. The value of N_{sp} is calculated from the targeted (final) packing porosity $\varepsilon_{\mathrm{fin}}$, V_{box}, and the final sphere diameter d_{fin} as

$$\varepsilon_{\mathrm{fin}} = 1 - \frac{N_{\mathrm{sp}} \pi d_{\mathrm{fin}}^3}{6 V_{\mathrm{box}}}. \tag{1}$$

Each iteration of the algorithm includes i) search of two particle centres C_1 and C_2 with the minimum pair-wise distance d_{min}, where d_{min} defines the maximal sphere diameter at which no overlap occurs in the current packing configuration and, consequently, the corresponding packing porosity

$$\varepsilon_{\mathrm{min}} = 1 - \frac{N_{\mathrm{sp}} \pi d_{\mathrm{min}}^3}{6 V_{\mathrm{box}}}, \tag{2}$$

and ii) symmetrical spreading apart of these two particle centres along a line C_1C_2 up to a new distance, d_{max}, according to the following equation:

$$d_{\mathrm{max}} = d_{\mathrm{min}} \left(1 + \alpha \log_{10} \left(\frac{d_{\mathrm{fin}}}{d_{\mathrm{min}}} \right) \right), \tag{3}$$

where α is a constant. According to Eq. 3, as d_{min} asymptotically approaches d_{fin}, $\varepsilon_{\mathrm{min}}$ approaches $\varepsilon_{\mathrm{fin}}$. The algorithm exits when the condition

$$|\varepsilon_{\mathrm{min}} - \varepsilon_{\mathrm{fin}}| < 1.001\, \varepsilon_{\mathrm{fin}} \tag{4}$$

is satisfied. One should mention that a particular choice of functional dependence $d_{\mathrm{max}} = f(d_{\mathrm{min}})$ is optional assuming monotonically decreasing functions. However, the choice of a function may affect the minimal packing porosity, convergence possibility, and convergence rate.

As specified above, in this work we generate packings with different values of porosity (ε) and degree of heterogeneity (DoH). The value of ε is varied using different amounts of spheres placed in a simulation box of fixed volume (see Eq. 1), while the DoH is varied using i) initial distribution of sphere centres and ii) value of the displacement constant α. We generate packings with two initial distributions of sphere centres: R- and S-packings. *R-packings* have uniform random distribution of sphere centres in the simulation box, while initial distribution in *S-packings* is performed in two steps: i) the simulation box is divided into N_{sp} equal cubic cells, and ii) each sphere centre is placed in a random position into a cell. Both types of initial distribution result in a uniform random distribution of sphere centres within the simulation box. The constant α in Eq. 3 defines the magnitude

of each displacement of centres C_1 and C_2: with a small value of α the sphere centres tend to stay closer to their initial positions as far as possible during generation, and the final packing configuration reflects the randomness of the initial distribution. In our study packings generated with the JT algorithm are denoted as "TxA", where T is the type of initial distribution of particle centres (R or S) and A is the value of α. In case of $\alpha = 1$, the generated packings are referred to as just "T". Below the packings with different DoH's are mentioned as packings prepared with different generation protocols.

After the sphere packing is generated, it is discretized on a uniform three-dimensional lattice with a spatial resolution of n lattice nodes per sphere diameter. The resulting lattice is used to perform simulations of flow and hydrodynamic dispersion.

2.2 Simulation of Fluid Flow

The low-Reynolds number flow ($Re \ll 1$) of an incompressible fluid was simulated using the lattice-Boltzmann method (LBM)[3], which is a popular approach nowadays to simulate flow in complex geometries, like those occurring in random porous media. LBM is based on the kinetic theory of gases, and it recovers motion of a fluid using the lattice Boltzmann equation (LBE):

$$f_i(\vec{r} + \vec{e}_i, t + \Delta t) = f_i(\vec{r}, t) + \Omega_i(f_i(\vec{r}, t)), \tag{5}$$

which describes time evolution of the discrete one-particle distribution function f_i at lattice site $\vec{r}$ due to collision (denoted by the collision operator Ω_i) and propagation after the time step Δt along the lattice links $\vec{e}_i$. The local fluid density ($\rho(\vec{r}, t)$) and velocity ($\vec{v}(\vec{r}, t)$) are calculated from statistical moments of the distribution function as

$$\rho(\vec{r}, t) = \sum_i f_i(\vec{r}, t), \tag{6a}$$

$$\vec{v}(\vec{r}, t) = \frac{1}{\rho(\vec{r}, t)} \sum_i \vec{e}_i f_i(\vec{r}, t). \tag{6b}$$

In our simulations we use the BGK (Bhatnagar–Gross–Krook) collision operator, $\Omega_i(f_i(\vec{r}, t)) = (f_i(\vec{r}, t) - f_i^{\mathrm{eq}}(\rho, \vec{v}))/\tau$, where τ is the model relaxation time, and f_i^{eq} is the equilibrium distribution function of the following form:

$$f_i^{\mathrm{eq}}(\rho, \vec{v}) = w_i \rho \left(1 + \frac{3\vec{e}_i \cdot \vec{v}}{c_{\mathrm{s}}^2} + \frac{9(\vec{e}_i \cdot \vec{v})^2}{2c_{\mathrm{s}}^4} - \frac{3\vec{v} \cdot \vec{v}}{2c_{\mathrm{s}}^2}\right). \tag{7}$$

Here c_{s} is the speed of sound in the system, and weighting coefficients w_i depend on the chosen lattice type. In our simulations we employ a three-dimensional lattice with 19 lattice links: each lattice site has 6 "horizontal" links (connecting the site to its neighbours located at a distance of one lattice step Δr), 12 "diagonal" links (connecting the site with sites located on a distance of $\sqrt{2}\Delta r$), and one "rest" link which is a connection of the site to itself. Such a choice of the lattice links results in $w_i = 1/18$ for horizontal, $w_i = 1/36$ for diagonal, and $w_i = 1/3$ for the remaining links, respectively.

2.3 Simulation of Advection–Diffusion Transport

The last simulation step is concerned with the simulation of advection–diffusion mass transport (hydrodynamic dispersion) in stationary flow in the void space of the generated packings. For this purpose we employ a stochastic approach, the random walk particle tracking method (RWPT)[4]. The idea of the method is to distribute a large amount of point-like tracers in the void space of interest (which is the packing pore space in our study), and to perform iterative displacement of each tracer collecting statistical information over the whole tracer ensemble. In our simulations displacement of the i-th tracer (Δr_i) after time step Δt is calculated as a sum of advection and diffusion contributions:

$$\Delta r_i = \vec{v}(\vec{r}_i)\Delta t + \sqrt{2D_m \Delta t}\vec{N}, \tag{8}$$

where $\vec{v}(\vec{r})$ is the stationary flow field calculated on the previous step with LBM, D_m is the molecular diffusion coefficient of a tracer, and $\vec{N}$ is the random vector with components calculated using normally distributed random numbers with zero mean and unity variance. A time-dependent hydrodynamic dispersion coefficient ($D_z(t)$) characterizing the spreading rate of a tracer ensemble around its mean position is calculated as

$$D_z(t) = \frac{1}{2N_{\mathrm{tr}}\Delta t}\sum_{i=1}^{N_{\mathrm{tr}}}\left(\Delta r_{zi}(t) - \langle \Delta r_{zi}(t)\rangle\right)^2, \tag{9}$$

where N_{tr} is the amount of tracers, $\Delta r_{zi}(t)$ is the z-displacement of the i-th tracer after time t, and $\langle \ldots \rangle$ operator denotes averaging over the whole tracer ensemble.

2.4 Accuracy of the Transport Simulation

The discrete spatial lattice introduced for the simulation of flow and hydrodynamic dispersion, results in a "stair-step" approximation of the smooth boundary of the spheres. This approximation along with other causes (finite size of the periodic domain containing the packing, boundary conditions employed in LBM and RWPT simulations, finite time step used for the calculation of the displacement of a tracer) introduce numerical artifacts into the simulations. Quantification of these artifacts is a non-trivial task, and, due to the space limitation, here we provide only a brief discussion. For more detailed information the reader is referred to Ref. 5.

Periodic boundary conditions applied at the walls of the box containing the random sphere packing may introduce artifacts which will affect the value of the hydrodynamic dispersion coefficient. This situation is illustrated in Fig. 1a, where we present time-dependent hydrodynamic dispersion coefficients in packings with lateral dimensions of $10\ d_{\mathrm{p}} \times 10\ d_{\mathrm{p}}$ and the length between $6\ d_{\mathrm{p}}$ and $150\ d_{\mathrm{p}}$. The advective time scale employed to show the data in Fig. 1a characterizes the average distance in particle diameters travelled by the tracers. Hydrodynamic dispersion coefficients converge to the lower asymptotic values with an increase in the packing length. The time evolution of the coefficient becomes independent on the packing length for length value bigger than $75\ d_{\mathrm{p}}$. The demonstrated model behaviour is a numerical artifact related to the situation when certain tracers leaving and, due to periodic boundary conditions, entering the simulation domain from its opposite side re-experience particular flow stream lines persisting in a packing. The above artifact can be avoided if one uses "long enough" packings to achieve close-to-asymptotic behaviour

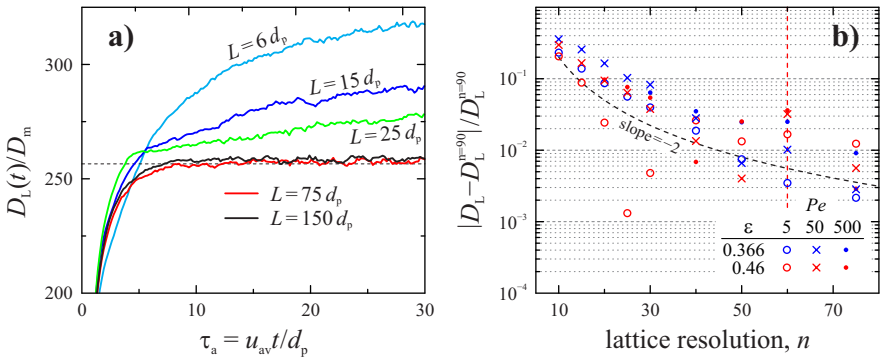


Figure 1. a) Evolution of the hydrodynamic coefficient as a function of advective time $\tau_{\mathrm{a}} = u_{\mathrm{av}} t/d_{\mathrm{p}}$, where u_{av} is the average flow velocity and d_{p} is the particle diameter. Simulations are performed at $Pe = u_{\mathrm{av}} d_{\mathrm{p}}/D_{\mathrm{m}} = 500$ in sphere packings with length of 6, 15, 25, 75, and 150 d_{p} and transverse dimensions of $10\,d_{\mathrm{p}} \times 10\,d_{\mathrm{p}}$, porosity of $\varepsilon = 0.366$, and discretized with spatial resolution of $n = 90$ lattice nodes per sphere diameter. b) Dependence of the relative error in close-to-asymptotic values of the hydrodynamic dispersion coefficient on the lattice resolution at $Pe = 5$, 50, and 500 in Rx0.001 packings with porosity values indicated in the figure legend.

of the dispersion coefficient *before* the average tracer displacement exceeds about half a packing length[6,7].

Fig. 1b shows a resolution study, where we present the dependence of the relative error in close-to-asymptotic values on the dispersion coefficient of two packings ($\varepsilon = 0.366$ and $\varepsilon = 0.46$) as a function of the lattice resolution (n). An error is calculated relative to the highest employed resolution, $n = 90$. As one can see, the relative error converges to its lattice independent value, and the difference between dispersion coefficients at $n = 75$ and $n = 90$ is about one percent. In the current study we use a lattice resolution of 60 nodes per sphere diameter, which, in combination with packing dimensions of $10 \times 10 \times 70\,d_{\mathrm{p}}^3$ as discussed above, results in a total lattice dimension of $600 \times 600 \times 4300$ nodes. These dimensions lead to the memory requirement of 110 GB to store in computer memory one velocity field during LBM simulations. In addition to the memory requirements, computational demands grow dramatically with the increase of discretization resolution: for example, computational time needed for LBM simulations grows as $O(n^5)$. Both high memory and computational time requirements are the result of the complexity of the pore space geometry of the sphere packings, and demand use of supercomputing facilities to perform simulations.

3 Results

Random sphere packings with periodic boundary conditions and dimensions of $10 \times 10 \times 70\,d_{\mathrm{p}}^3$ (d_{p} is the sphere diameter) were generated at porosity values of 0.366, 0.38, 0.40, 0.42, 0.44, 0.46, which cover the range between random close ($\varepsilon \approx 0.366$) and random loose ($\varepsilon \approx 0.46$) packing limits[8]. For each porosity, packings were created with four generation protocols (Rx0.001, R, S, Sx2; see Fig. 2, left). Each packing with certain porosity and protocol was generated using 10 different initial random distributions of sphere centres to address an eventual packing-to-packing variation in microstructure and resulting transport properties.

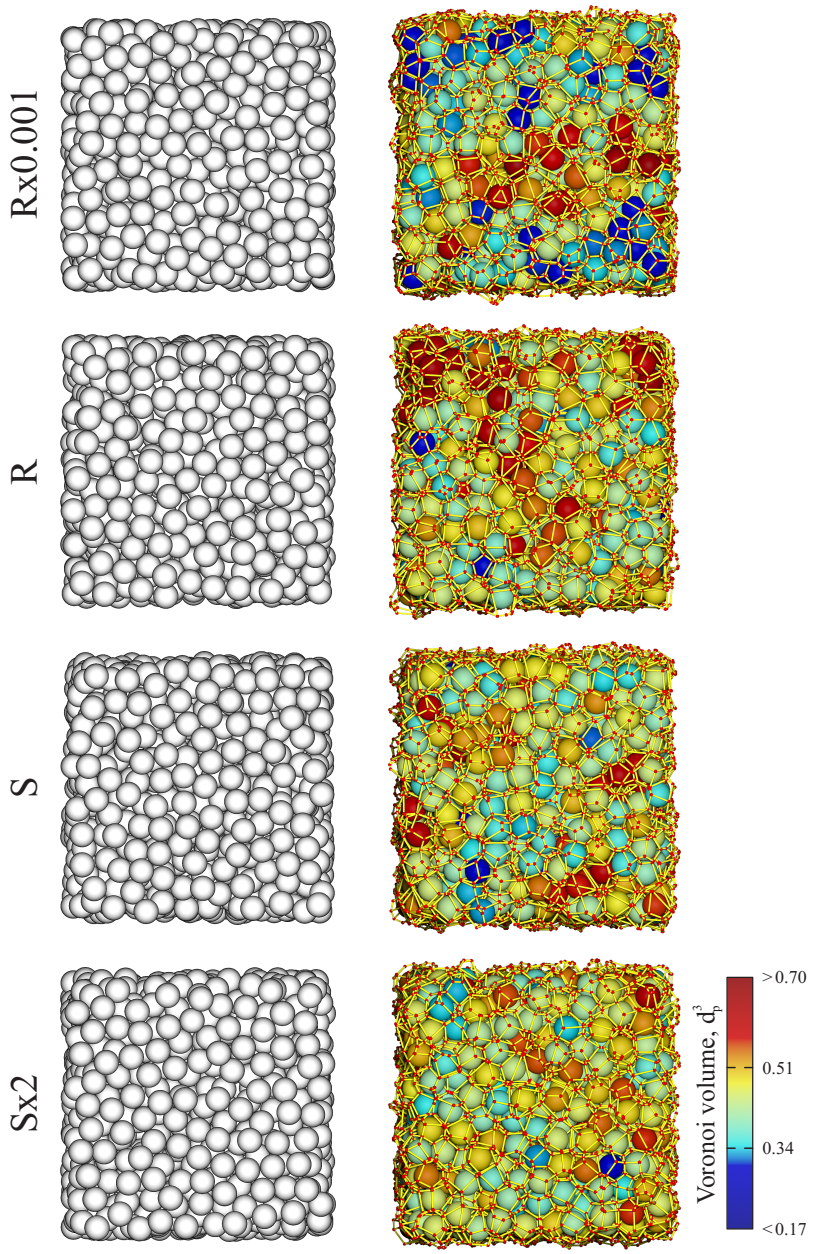


Figure 2. Left: front view onto the generated packings (spatial dimensions: $10 \times 10 \times 70\, d_\mathrm{p}^3$, where d_p is the sphere diameter). Presented packings have fixed porosity ($\varepsilon = 0.46$), and were prepared with different generation protocols (see main text). Figure on the left demonstrates no apparent difference for all the shown packings. Right: Voronoi tessellation of the packings presented on the left; yellow links and red nodes indicate polyhedra forming Voronoi cells; a Voronoi cell is the set of points closer to a given sphere than to any of its neighbours. Comparing the packings from top to bottom, one can see increase in uniformity of the Voronoi volumes distribution, indicated by the increasing amount of yellow and cyan colours.

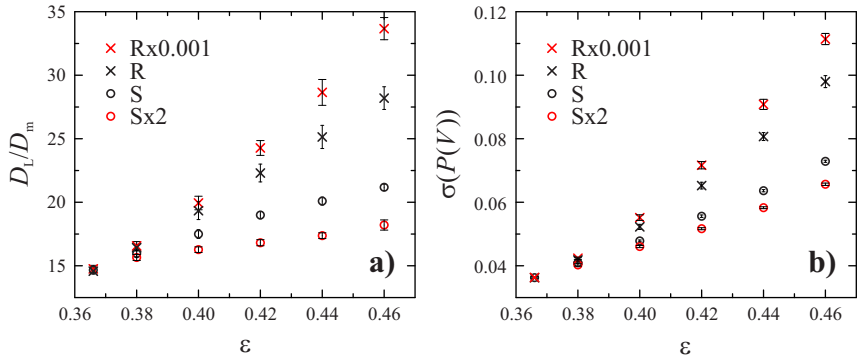


Figure 3. Normalized asymptotic dispersion coefficients $D_{\mathrm{L}}/D_{\mathrm{m}}$ (a) and standard deviation of the Voronoi volume distribution (b) for all generated packings as a function of packing protocol (Rx0.001, R, S, Sx2) and porosity. Dispersion coefficients were determined at $Pe = 50$. Error bars indicate 95% confidence intervals calculated using values of dispersion coefficient and standard deviation of 10 random realizations of each packing with a fixed porosity and packing generation protocol.

In Fig. 3a we show the asymptotic hydrodynamic dispersion coefficients for the generated packings, obtained by the simulation procedure described in Sec. 2. The dispersion coefficients are shown for the flow regime intermediate between diffusion- and advection-dominated (or, in other words, at intermediate Péclet number); a similar dependence of the dispersion coefficient on the porosity and protocol is observed as well as for the low ($Pe = 5$) and high ($Pe = 500$) flow rates[9]. Results in Fig. 3a show that i) all packings created with an individual protocol demonstrate a distinguished dispersion coefficient–porosity trend, and ii) variation in the dispersion coefficients (characterized by error bars in Fig. 3a) obtained for 10 different realizations of each packing is clearly smaller than the variation caused by the packing protocol.

We employ a sensitive analysis method, *Voronoi tessellation*, to analyze the microstructure of the generated packings. Applying the method to the monodisperse random sphere packings, it splits the packing pore space into irregular polyhedra (Voronoi cell) in a way where each polyhedron surrounds a given sphere and contains space points which are closer to the centre of this sphere than to any other. Each polyhedron can be characterized by its volume V (Voronoi volume), and analysis of the statistical distribution of the corresponding Voronoi volumes $P(V)$ is a straightforward way to quantify the distribution of the packing void space. Fig. 2 (right) shows the front view onto the packings of the same porosity ($\varepsilon = 0.46$) generated with the different protocols. The colours indicating the volume of a Voronoi cell of each sphere range mostly from dark blue to dark red for Rx0.001 packing, whereas they do from yellow to cyan for Sx2 packing. This difference in the colour ranges reflects the difference in maldistribution of the void space in the packings.

Distribution of the Voronoi volumes $P(V)$ can be quantified with its statistical moments, mean and standard deviation. The mean of $P(V)$ can be calculated analytically from the known porosity of a packing and it is independent on the packing microstructure (generation protocol). At the same time, Fig. 3b demonstrates that the standard deviation $\sigma(P(V))$ is strongly affected not only by porosity, but also by the packing generation protocol. More importantly, comparison between Figs. 3a and 3b, indicates a good agreement

of the dispersion coefficients with the standard deviation of the Voronoi volumes. This circumstance points out the importance of further investigation of the geometrical properties of sphere packings using spatial tessellations in order to establish quantitative geometry–transport correlations.

4 Conclusion

The numerical approach employed in this study allowed us i) to create random sphere packings with varied geometrical properties (porosity, degree of heterogeneity), ii) to perform accurate pore-scale simulations of the advection–diffusion transport (hydrodynamic dispersion), iii) to analyze pore space of the packings with Voronoi tessellation, and iv) to demonstrate a good agreement between a scalar geometrical measure derived from the Voronoi tessellation and transport (hydrodynamic dispersion) coefficient. Concluding, the presented approach provides a systematic way to the quantification of transport processes in (random) porous media, based only on their geometrical properties.

Acknowledgements

The calculations were carried out at the Jülich Supercomputing Centre.

References

1. U. M. Scheven, R. Harris, and M. L. Johns, *Intrinsic dispersivity of randomly packed monodisperse spheres*, Phys. Rev. Lett. **99**, 054502, 2007.
   DOI: 10.1103/PhysRevLett.99.054502
2. W. S. Jodrey and E. M. Tory, *Computer simulation of isotropic, homogeneous, dense random packing of equal spheres*, Powder Technol. **30**, 111–118, 1981.
   DOI: 10.1016/0032-5910(81)80003-4
3. S. Succi, *The lattice Boltzmann equation for fluid dynamics and beyond* (Oxford University Press, New York, 2001).
4. P. Salamon, D. Fernàndez-Garcia, and J. J. Gómez-Hernández, *A review and numerical assessment of the random walk particle tracking method*, J. Contam. Hydrol. **87**, 277–305, 2006.
   DOI: 10.1016/j.jconhyd.2006.05.005
5. S. Khirevich, *High-Performance Computing of Flow, Diffusion, and Hydrodynamic Dispersion in Random Sphere Packings*, (Ph.D. Thesis, Department of Chemistry, Philipps-Universität, Marburg, 2010).
6. C. P. Lowe and D. Frenkel, *Do hydrodynamic dispersion coefficients exist?* Phys. Rev. Lett. **77**, 4552–4555, 1996.
   DOI: 10.1103/PhysRevLett.77.4552
7. R. S. Maier, D. M. Kroll, R. S. Bernard, S. E. Howington, J. F. Peters, and H. T. Davis, *Pore-scale simulation of dispersion* Phys. Fluids **12**, 2065–2079, 2000.
   DOI: 10.1063/1.870452
8. S. Song, P. Wang, and H. A. Makse, *A phase diagram for jammed matter*, Nature **453**, 629–632, 2008.
   DOI: 10.1038/nature06981

9. S. Khirevich, A. Daneyko, A. Höltzel, A. Seidel-Morgenstern, and U. Tallarek, *Statistical analysis of packed beds, the origin of short-range disorder, and its impact on eddy dispersion*, J. Chromatogr. A **1217**, 4713–4722, 2010.
DOI: 10.1016/j.chroma.2010.05.019

Finite Size Particles in Homogeneous Turbulence

Markus Uhlmann and Todor Doychev

Institute for Hydromechanics, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany
E-mail: {markus.uhlmann, todor.doychev}@kit.edu

1 Introduction

Two-phase flows involving dilute suspensions of solid particles are of relevance to a large number of natural and technical processes. Examples include the motion of precipitation particles in the atmosphere, transport of solid-fluid mixtures in combustion devices as well as diverse chemical engineering applications. In many cases of practical importance, the flow is either per se (i.e. in the absence of particles) turbulent or does exhibit "turbulent" motion induced by the particles. The interaction between the turbulent carrier flow and solid particles can lead to a number of hydrodynamical coupling effects, which are most prominently manifested by (a) a peculiar spatial particle distribution (a.k.a. preferential particle concentration), (b) an increase or decrease in the particle settling velocity, (c) an enhancement or attenuation of flow turbulence. The vast majority of past studies has been devoted to the regime of small particles, i.e. those with a diameter less than the smallest length scale of the flow – often idealized as point particles. Conversely, particles with a size comparable to or larger than the smallest flow scales have received less attention, presumably because the lack of a scale-separation as well as the fact that the local flow around the particle happens at finite Reynolds numbers lead to considerable additional complexity. Not surprisingly, an adequate description of the dynamics of flows involving "large particles" is currently being considered as one of the major open questions in the multiphase flow community[1].

2 Numerical Method and Performance on JUGENE

The numerical method used to solve the incompressible Navier-Stokes equations in the presence of multiple moving immersed bodies is identical to the one of Ref. 2. Our finite-difference-based approach – employing an immersed-boundary technique for the representation of non-grid-conforming solid objects – is characterized by the following numerical/algorithmic characteristics. The semi-implicit predictor step of the fractional step procedure requires the solution of three Helmholtz equations. An approximate factorization technique is employed, which yields a sequence of one-dimensional implicit equations. The resulting matrix systems are solved by means of a parallel tridiagonal solver[3]. The projection step requires the solution of a Poisson equation for the pseudo-pressure. This is achieved through a multi-grid procedure[4], which typically needs 6-10 iterations for convergence. Parallelism is realized by standard three-dimensional domain decomposition. The overall operation count of the fluid solver is therefore $\mathcal{O}(N_1 N_2 N_3)$, where N_i is the number of grid points in the ith coordinate direction. For the evaluation of all central second-order finite difference operators, communication of a single ghost cell between neighbouring processes is sufficient. The representation of the solid phase essentially amounts to an

interpolation step and a subsequent spreading step in which data is transferred back and forth in between the Cartesian grid and Lagrangian marker points attached to the particle surfaces. These two operations can be realized in a relatively cost-efficient manner, since the interpolation functions have a limited spatial support. The operation count is therefore $\mathcal{O}(N_p\,(D/\Delta x)^2)$, where N_p is the number of particles. Particle-related work requires the knowledge of the surrounding flow field. Therefore, it is assigned to the process holding the respective sub-volume in which the particle currently resides. Whenever a particle crosses an inter-process boundary, control is passed over to the process holding the new sub-volume. Furthermore, whenever a particle overlaps sub-volume boundaries, two or more processors need to perform the computation of coupling forces jointly. A specific protocol for such communication steps of the dispersed phase has been developed. Again, communication only involves the 26 neighbouring processes. Balancing of fluid related work is automatically guaranteed by an equal distribution of the global grid. Particle-related work is difficult to balance exactly, since particles are free to move anywhere in the computational domain. However, our experience has shown that load balancing in statistically homogeneous flows is not an issue.

Fig. 1 presents weak scaling results for our code. As can be seen, the scaling is indeed good, considering that we solve an elliptic PDE. In particular, in the actual range of production runs with up to 64K processor cores, we achieve quite reasonable parallel efficiencies.

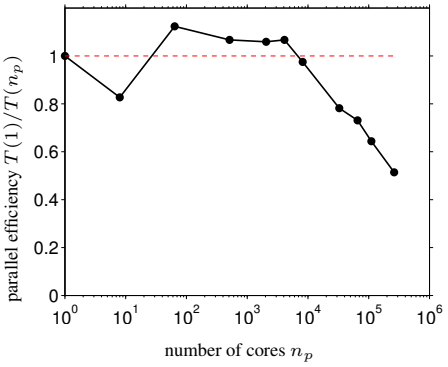


Figure 1. Weak-scaling of a full Navier-Stokes time step on JUGENE. In all cases each core is treating 128^3 grid cells. The largest number of cores used in this series is $n_p = 64^3 = 262144$ (64 racks).

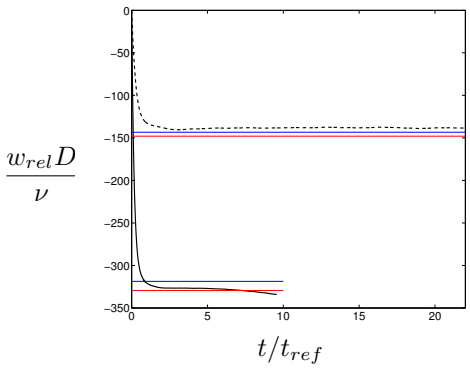


Figure 2. Average particle settling velocity (normalized with particle diameter and kinematic viscosity) as a function of time. The relative velocity is defined as $w_{rel} = \langle w^{(p)} \rangle - \langle w \rangle$. ——, present simulation (case A1); – – – –, previous simulation with identical density ratio and solid volume fraction, but smaller terminal Reynolds number and box size (case R). The terminal settling velocity for an isolated sphere in unbounded ambient fluid $w_\infty^{(p)}$ (cf. Ref. 5) is shown as ——; the data for an isolated sphere, corrected for collective hindrance[6] is shown as ——.

case	ϕ_s	ρ_p/ρ_f	$\frac{\sqrt{Dg_z}\,D}{\nu}$	box size	# of particles
A1	0.005	1.5	320.8	$137D \times 137D \times 546D$	97218
R	0.005	1.5	171.5	$34D \times 34D \times 273D$	3038

Table 1. Physical parameters describing the main simulation, here denoted as "case A1", as well as a smaller reference case ("case R") which we have simulated elsewhere in the past. Shown are the solid volume fraction ϕ_s, the solid-to-fluid density ratio ρ_p/ρ_f, as well as the gravitational to viscous velocity scale ratio $\sqrt{Dg_z}D/\nu$, the box size in multiples of the particle diameter, and the total number of particles.

3 Results

The physical problem involves the gravitational settling of approximately 10^5 spheres under dilute conditions (solid volume fraction $\phi_s = 5 \cdot 10^{-3}$), a system size which has not been accessible to high-fidelity numerical simulations in the past. The problem is fully characterized by three non-dimensional parameters, which (in addition to the volume fraction ϕ_s) can be taken as the density ratio and the gravitational to viscous velocity scale ratio. The corresponding values are given in Tab. 1. The table also shows that the box size of the current simulation ("case A1") is fourfold elongated in the vertical direction, amounting to nearly 550 particle diameters. We will see below that this size is indeed sufficient in order to guarantee a proper spatial decorrelation of velocity fluctuations. With the given solid volume fraction and the ratio of particle diameter to box size, a total number of 97218 particles is obtained. To our knowledge, no high-fidelity simulation with more than $\mathcal{O}(1000)$ resolved and fully mobile particles has been undertaken in the past.

In the numerical simulation we employ a (uniform and isotropic) grid with $2048 \times 2048 \times 8192$ nodes (total $3.4 \cdot 10^{10}$). Particles are resolved with 15 grid widths per diameter. The simulation was set up to run on 16384 processor cores on JUGENE. A snapshot of particle locations and vortical structures in the flow field can be seen in Fig. 3.

3.1 Particle Settling Velocity

Fig. 2 shows how the average settling velocity evolves with time. It should be noted that the simulation starts with all particles and the fluid at rest. Therefore, it requires an initial transient for particles to reach their approximate terminal velocity, after which the settling velocity exhibits only gradual changes. The smoothness of the curve in Fig. 2 is a manifestation of the large number of particles.

It can be seen from Fig. 2 that the present simulation yields settling velocities which are close to the values predicted by a balance between drag and buoyancy when using the standard drag law[5]. However, although the settling velocity seems to reach a nearly time-independent state by $t/t_{tref} \approx 5$, variation sets in at later times, when the average settling velocity starts to increase again. A similar behaviour was not observed in the reference case R (lower Reynolds number, smaller box), where the average settling velocity fluctuates little in time once the initial transient has been surpassed (cf. Fig. 2). The principal physical difference between the present case A1 and the reference case R is the type of flow induced in the particle wakes. As has been discussed by Ref. 7, the larger the extent of the wakes,

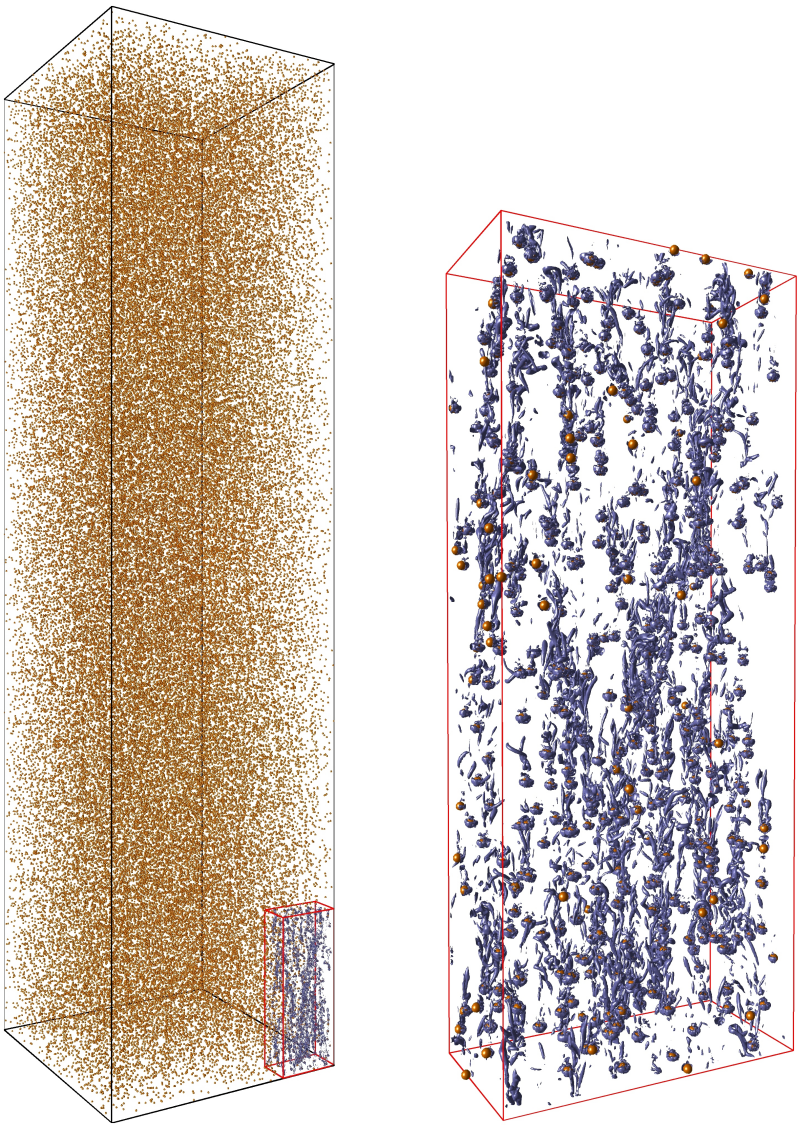


Figure 3. A dilute suspension of approximately 10^5 heavy spherical particles sedimenting in a viscous incompressible fluid in a periodic spatial domain. The image on the left shows the instantaneous spatial particle distribution; on the right corresponding coherent structures (vortices) are depicted in a small sub-volume. The terminal Reynolds number (for an isolated particle) is adjusted to a value of 354. Gravity acts from top to bottom.

the stronger the tendency of particles to accumulate. On the basis of the available previous data, it can be expected that significant particle clustering occurs in the present case A1. Consistently, the magnitude of the average settling velocity in Fig. 2 exhibits a visible trend to increase, as would be expected if a significant fraction of the overall number of particles is residing in other particles' wakes.

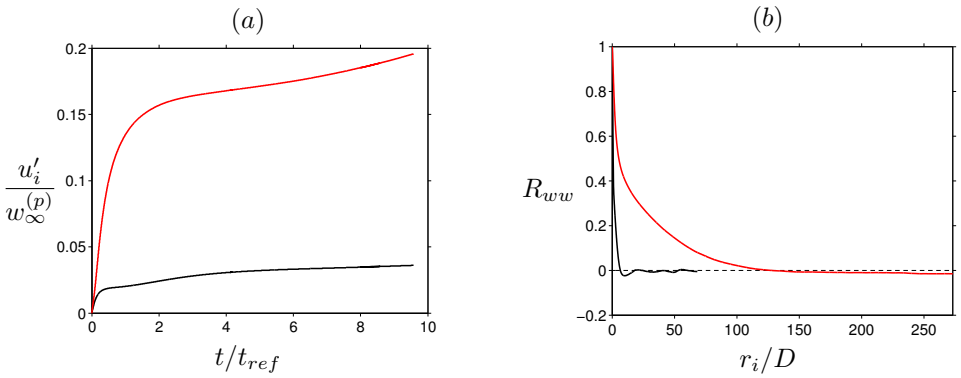


Figure 4. (a) R.m.s. fluid velocity fluctuations of the horizontal components (——) and the vertical (——) component. The terminal settling velocity $w_\infty^{(p)}$ given by the standard drag law[5] is used as the reference velocity. (b) Two-point correlation of the vertical component of the fluid-phase velocity, R_{ww}, as a function of vertical (——) and horizontal separations (——).

3.2 Turbulence Generation

As stated above, particles are released into ambient fluid. Therefore, any fluid motion is induced by the particle motion, which in turn is due to the gravitational potential. The relevant questions in this context are aimed at determining the intensity of self-induced fluid motion, as well as the structure of this particle-induced "turbulence". This analysis is different from (but related to) the study of the modulation of existing (background) turbulence due to the addition of particles, which will be studied in a future sub-project of the present study.

Fig. $4(a)$ shows the temporal evolution of the (box-averaged) r.m.s. values of fluid velocity fluctuations. It can be seen that the vertical component is highly dominant (i.e. the anisotropy b_{33} is close to the maximum value of $2/3$) as a consequence of the wake-induced character of the fluid motion. The quantity shown in Fig. $4(a)$ also directly represents the intensity of the fluctuations as commonly used with the purpose of parameterizing the influence of turbulence upon the particle motion. It can be seen that intensity values of the order of 0.2 (i.e. 20% of the terminal particle velocity) are reached. This level of intensity is comparable to the turbulence levels often considered in studies of the influence of background turbulence upon particle wakes[8,9] and its influence on hydrodynamic forces acting upon particles[10]. It can be concluded that, although the turbulent motion is self-induced and the particle volume fraction is dilute, appreciable average turbulence levels are generated in the present case.

Concerning the temporal evolution of the curves in Fig. $4(a)$, we observe that both components appear to level off around $t/t_{ref} = 5$, later starting to rise again (particularly the vertical component). This trend confirms the hypothesis that significant particle clustering is taking place towards the end of the present simulation interval, since residence of particles in fast-settling clusters is expected to increase the overall r.m.s. value of the fluid (and particle) velocities, particularly the vertical component.

A particular concern when using tri-periodic domains is the adequacy of the box size. One indicator of decorrelation is the two-point autocorrelation function. Fig. $4(b)$ shows

the spatial correlation of the vertical fluid velocity component as a function of horizontal and vertical separations. It can be observed that a complete spatial decorrelation in both directions is indeed warranted by the presently chosen box size. In particular, R_{ww} reaches a zero value at vertical separations of approximately 130 particle diameters. This length can be interpreted as a characteristic length of the wake structures. In the horizontal direction a small but significant negative correlation is observed for separations of approximately 10 particle diameters. Several local secondary extrema are recorded, each at a distance of approximately $10D$. These values can be taken as characteristic measures of average horizontal distances between individual particle wakes.

3.3 Spatial Structure of the Particle Phase

Several techniques are commonly used for the characterization of the spatial structure of the dispersed phase. Besides conventional box-counting[11] and genuine clustering detection algorithms[12], a technique based upon Voronoi tesselation has recently been proposed in the context of particulate flows[13]. Fig. 5(a) shows an example in two dimensions, where starting with random particle positions (the "sites") the space is covered with cells which have the property that each point in the cell is closer to the cell's site than to any other cell's site. As a consequence, the inverse of a Voronoi cell's volume is proportional to the local particle concentration.

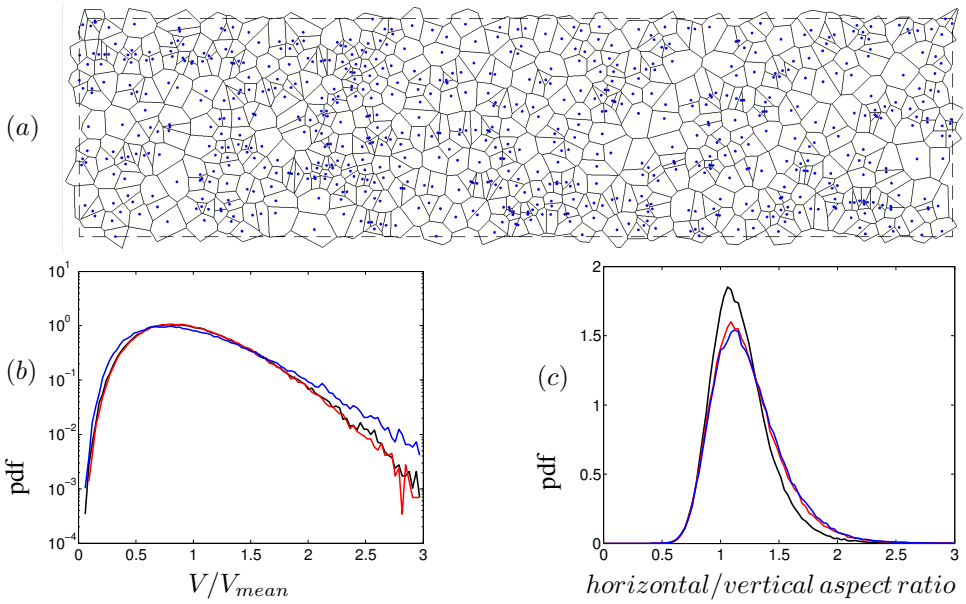


Figure 5. (a) Example of Voronoi tesselation of a bi-periodic plane, performed with respect to the locations indicated as blue dots. Note that the cells are continued periodically in order to guarantee a space-filling and non-redundant tesselation. The periodic boundaries are indicated by dashed lines. (b) Pdf of Voronoi cell volumes V at different times: ——, $t/t_{ref} = 0$ (random initial positions); ——, $t/t_{ref} = 5.0$; ——, $t/t_{ref} = 9.6$. (c) Pdf of the horizontal-to-vertical aspect ratio of Voronoi cells at the same times as in (b).

A graph of the pdf of Voronoi cell volumes based upon instantaneous particle positions at different times in our simulation is depicted in Fig. $5(b)$. Since the initial particle positions were chosen randomly, the initial pdf coincides with the theoretical Gamma distribution for random particle fields[14]. After an elapsed time of 5 reference time units the particle Voronoi cell volumes are still found to be approximately randomly distributed. At the latest time ($t/t_{ref} = 9.6$), a significant deviation from randomness is observed. The latest pdf indeed exhibits the behaviour reported in earlier studies[13], characterized by larger-than-random probabilities of finding very small and very large Voronoi cells while roughly maintaining the overall shape of the pdf. As a consequence, there exist two cross-over points (approximately at $V/V_{mean} = 0.6$ and 1.5) between the latest pdf and the reference pdf for random particle distribution. These cross-over points give us the possibility of defining an objective threshold cell-size for determining whether a particle belongs to a cluster region or not, as suggested by Ref. 13. As a consequence, the possibility of computing and analyzing conditional statistics exists, without the need to introduce ad hoc threshold values. This possibility will be further explored in the future.

Additional quantities of interest can be deduced from the Voronoi tesselation. In our case, where there exists one preferred spatial direction due to gravity, particle accumulation is not expected to take place in an isotropic manner. Rather do we expect particle clusters to exhibit a vertically elongated shape. This kind of non-isotropic cluster can sometimes be "invisible" in the light of a Voronoi analysis purely based upon cell volumes. Therefore, we have computed the aspect ratio of Voronoi cells dividing the largest horizontal extension by the largest vertical extension of each Voronoi cell. The corresponding pdfs are shown in Fig. $5(c)$, where it can be observed that both evolved particle fields (at $t/t_{ref} = 5$ and 9.6) are distinct from the random case: large aspect ratios (i.e. Voronoi cells which are "flat") are found to be clearly more probable than in the random case. This result is consistent with particles lining up in elongated, thin columns, where the vertical distances to the nearest neighbour are typically smaller than the corresponding horizontal distances. Therefore, considering the distribution of Voronoi cell aspect ratios has revealed a tendency to form elongated particle arrangements (significantly different from those of randomly positioned particles) at earlier times ($t/t_{ref} = 5$) than clustering diagnosed by pure Voronoi cell volume statistics ($t/t_{ref} = 9.6$).

4 Conclusion

Owing to continuous advances in computer hardware as well as efficient numerical algorithms it has now become feasible to faithfully simulate fluid-particle systems of considerable complexity (order of 10^5 spheres in dilute conditions, large box sizes). The data obtained through such numerical simulations will contribute to the elucidation of such fundamental questions as:

- How is the average particle settling velocity affected by (self-induced) turbulence?
- What are the characteristics of turbulence generated by the presence of particles?
- What is the spatial structure of the dispersed phase?

Integration over relatively long temporal intervals will be required in order to provide proper sampling of the large-scale dynamics. Since the time scales of particle "clusters"

are yet unknown, a certain amount of numerical exploration of the parameter space will be necessary in order to establish those requirements.

Acknowledgements

Support through a research grant from DFG (UH 242/1-1) is thankfully acknowledged. The calculations were carried out at the Jülich Supercomputing Centre.

References

1. S. Balachandar and J.K. Eaton, *Turbulent dispersed multiphase flow*, Ann. Rev. Fluid Mech., **42**, 111–133, 2010.
2. M. Uhlmann, *An immersed boundary method with direct forcing for the simulation of particulate flows*, J. Comput. Phys., **209**, no. 2, 448–476, 2005.
3. N. Mattor, T.J. Williams, and D.W. Hewett, *Algorithm for solving tridiagonal matrix problems in parallel*, Parallel Computing, **21**, 1769–1782, 1995.
4. B. Bunner, "Code package MGD: an MPI-parallel replacement for MUDPACK", URL: `http://www.mgnet.org/mgnet-codes-bunner.html`, 1998.
5. R. Clift, J.R. Grace, and M.E. Weber, *Bubbles, drops and particles*, Academic Press, 1978.
6. G.K. Batchelor, *Sedimentation in a dilute dispersion of spheres*, J. Fluid Mech., **52**, no. 2, 245–268, 1972.
7. T. Kajishima, *Numerical investigation of collective behaviour of gravitationally settling particles in a homogeneous field*, in: Proc. ICMF 2004 (5th Int. Conf. Multiphase Flow), Y. Matsumoto, K. Hishida, A. Tomiyama, K. Mishima, and S. Hosokawa, (Eds.), Yokohama, 2004.
8. P. Bagchi and S. Balachandar, *Response of the wake of an isolated particle to an isotropic turbulent flow*, J. Fluid Mech., **518**, 95–123, 2004.
9. D. Legendre, A. Merle, and J. Magnaudet, *Wake of a spherical bubble or a solid sphere set fixed in a turbulent environment*, Phys. Fluids, **18**, 048102, 2006.
10. P. Bagchi and S. Balachandar, *Effect of turbulence on the drag and lift of a particle*, Phys. Fluids, **15**, no. 11, 3496–3513, 2003.
11. J.R. Fessler, J.D. Kulick, and J.K. Eaton, *Preferential concentration of heavy particles in a turbulent channel flow*, Phys. Fluids, **6**, no. 11, 3742–3749, 1994.
12. J.A. Melheim, *Cluster integration method in Lagrangian particle dynamics*, Comput. Phys. Commun., **171**, no. 3, 155–161, 2005.
13. R. Monchaux, M. Bourgoin, and A. Cartellier, *Inertial particles clustering in turbulent flows: a Voronoi analysis*, in: ICMF 2010, S. Balachandar and J.S̃inclair Curtis, (Eds.), Proc. 7th Int. Conf. Multiphase Flow, CDROM, Tampa, USA, 2010.
14. J.-S. Ferenc and Z. Neda, *On the size distribution of Poisson Voronoi cells*, Physica A, **385**, 2007.

Computational Plasma Physics: New Insights into Laser-Matter-Interaction Processes

Karl-Heinz Spatschek

Institut für Theoretische Physik, Heinrich-Heine-Universität Düsseldorf
D-40225 Düsseldorf, Germany
E-mail: khs@tp1.uni-duesseldorf.de

Four areas, namely magnetic fusion, laser-matter-interaction, plasma astrophysics, and plasma technology are in the centre of actual plasma research. The first three necessitate computational support at highest level to achieve their goals. For magnetic fusion, special computers [e.g. the European "High Performance Computer For Fusion" (HPC-FF)] were installed to evaluate stability and confinement. Plasma astrophysics appears in this volume under a separate chapter. Thus, it is not astonishing that in the present chapter the emphasis is on laser-matter interaction. Short-pulse laser plasmas are inherently nonlinear, difficult to diagnose experimentally, and need extensive numerical simulations. The two projects being presented after this introductory remarks are typical examples of projects which lead into new insights into laser-matter-interaction processes via computational physics.

A new area of laser plasma interaction started approx. 40 years ago when laser fields became strong enough to directly ionize matter. With lasers operating at wave-lengths $0.25\,\mu m - 13.4\,\mu m$ one cannot expect the photoelectric effect to be effective on normal material since the photon energy $\hbar\omega$ is much smaller than the atomic potential barrier which the electron experiences in the vicinity of the nucleus (e.g. $13.6\,eV$ for the ground-state in hydrogen). With more powerful laser systems in the 60s and 70s of last century multiphoton ionisation became available. The rich physics being encountered already in the weakly nonlinear low-intensity regime is, e.g., summarized in the monograph by Kruer[1]. Before the CPA-area, the possibility of amplifying pulses directly was limited since intensities of some GW/cm^2 would damage the gain medium. Stretching the pulse in a controlled (reversible) manner by adding a chirp (temporal variation of the frequency), is the first stage in CPA (chirped pulse amplification). The chirped pulse is then amplified. For that usually one or more conventional laser amplifier stages can be used to increase the pulse energy by a factor $10^7 - 10^9$. Finally, a compressor performs the exact optical inverse of the strecher to deliver an amplified pulse of the same duration as the oscillator at the beginning. Optical Parametric Amplification (OPCPA) in a nonlinear crystal promises even higher intensities than conventional CPA.

To present an example, the "magic numbers" of the Jena TW laser are[2]: wavelength $800\,nm$, energy $0.8\,J$, pulse duration $80\,fs$, and focal spot size $5\,\mu m^2$, leading to an approximate intensity of $10^{20}\,W/cm^2$. In this regime, the generated laser plasma becomes relativistic. It is convenient to characterize the laser field by the dimensionless parameter $a_0 = eA_0/m_e c$, where A_0 is the peak vector potential of the laser pulse. Since a_0 corresponds to the normalized (by $m_e c$) maximum transverse momentum of an electron in the laser field, $a_0 \geq 1$ is often referred to as the entry into the relativistic optics. University facilities like the ARCTURUS at the University of Düsseldorf are able to produce $a_0 \sim 10$. A next big step will be reached at $a_0 \sim 1000$ when also ions will be accelerated to relativistic velocities. For higher intensities it will be possible to even study electron-positron

pair creation and other quantum-electrodynamic (QED) effects[3].

Laser-produced plasmas deliver unique conditions for studies of relativistic physics, sometimes called relativistic nonlinear optics. Laser-matter-interaction necessitates computational support at highest level to achieve its goals. Practically most of the new regimes in electron and ion acceleration, X-ray sources, etc. were proposed and supported by numerical modelling. Numerical approaches encompass single particle dynamics, particle-in-cell codes, kinetic, hydrodynamic (plasmadynamic), and magnetohydrodynamic (MHD) treatments, just to name the most popular ones.

The following two articles do represent results from high performance computing (HPC) in important areas of laser-matter interaction. The summary of Liseykina and Bauer[4] gives an overview over the state-of-art for formation and ionization dynamics in intense laser-droplet interaction. Particle-in-cell simulations with ionization included have shown that there exists a broad laser intensity regime for which wavelength-sized targets are not fully ionized. For higher-Z material this applies even to ultra-high intensities. Hence, for realistic simulations of laser plasma interaction the authors came to the conclusion that ionization processes have to be taken into account for reliable results. The three-dimensional fully kinetic PIC code UMKA was used to demonstrate that important conclusion.

The second contribution by Sperling et al.[5] gives a taste how Thomson scattering will evolve into an important tool for the diagnostics of dense strongly correlated plasmas. The Thomson scattering spectrum for inhomogeneous targets, as generated by irradiation of liquid hydrogen and helium droplets with optical laser light, is calculated by the authors. The method described in the project can be applied to various pump-probe scenarios by combining optical lasers, soft and hard X-ray sources. The simulations were performed with the JUROPA Supercomputer at the Jülich Supercomputing Centre (JSC).

References

1. W. Kruer, *The Physics of Laser-Plasma Interaction*, Addison-Wesley, 1988.
2. P. Gibbon, *Short pulse laser interactions with matter*, Imperial College Press, 2005.
3. K. H. Spatschek, *High Temperature Plasmas*, Wiley-VCH, 2012.
4. T. Liseykina and D. Bauer, *Plasma Formation and Ionization Dynamics in Intense Laser-droplet Interaction*, this volume.
5. P. Sperling, T. Liseykina, D. Bauer, and R. Redmer, *Thomson Scattering on Inhomogeneous Targets*, this volume.

Plasma Formation and Ionization Dynamics in Intense Laser-Droplet Interaction

Tatyana Liseykina and Dieter Bauer

Institut für Physik, Universität Rostock, 18057 Rostock, Germany
E-mail: {tatyana.liseykina,dieter.bauer}@uni-rostock.de

The progress in laser technology has led to light sources delivering pulses of femtosecond duration and focused intensities of up to 10^{22} W/cm^2. One important application of such laser light is the acceleration of particles. When intense laser light interacts with matter, the electric field of the laser easily rips off outer electrons, and a plasma is formed. Both electrons and ions can then be accelerated up to hundreds of MeV or even GeV[1]. The details of the acceleration mechanisms are still the subject of current research. One may distinguish between particle acceleration in underdense plasmas in which the laser pulse can propagate, and from overdense plasma. The electron acceleration in underdense plasmas is based on the generation of wake fields[2] or "bubbles"[3]. The mechanisms behind electron and ion acceleration from overdense and finite size targets such as thin foils, clusters, and droplets are multifaceted, complex, and still under debate[4]. Widely accepted scenarios of ion acceleration from thin foils are the "target normal sheath"[5] and — at higher laser intensities — the "radiation pressure acceleration"[6] or "light sail" regime[7].

While laser-based, "table-top" particle acceleration shows an unprecedented efficiency as far as particle energy per acceleration length is concerned, the beam quality is not yet sufficient for many applications. In fact, in order to be useful for, e.g., medical applications, the ion spectrum should be monochromatic, tunable, and reproducible. Researchers worldwide try to improve the beam quality by designing structured or multi-component targets[8]. Experimentalists in this field are guided by simulation results, mainly obtained using the "particle in cell" (PIC) method[9]. The plain PIC codes solve the Vlasov equation and thus describe collisionless plasmas only. Moreover, "standard" PIC codes start with a preformed plasma of a certain density profile and temperature, which means that ionization, both by the electric field and by collisions, is neglected. However, as the laser field propagation is determined by the plasma density but the plasma is generated by ionization due to the laser, the charge state and density distributions have to be calculated self-consistently. In fact, even the strongest present-day lasers cannot fully ionize heavier elements so that the assumption of a preformed plasma is often inadequate. Furthermore, in overdense targets the skin-effect prevents the laser from penetrating so that in general a complicated, space and time-dependent charge distribution develops. The goal of this paper is the study of the interplay between ionization and pulse propagation in intense laser-matter interaction using three-dimensional, relativistic PIC simulations with ionization included.

1 Introduction

The typical scenario of the intense laser-matter interaction will be that an overdense plasma is rapidly generated throughout an initially transparent target. Subsequently the laser is reflected and only within a thin skin layer at the front higher charge states may be produced. However, fast electrons will in general penetrate into the target and collisionally ionize the interior further. About 100 fs after the laser pulse is over the plasma is expected to have reached a thermal equilibrium. Present-day short-wavelength free-electron laser (FEL) sources (at, e.g., DESY in Hamburg or LCLS in Stanford) are used to probe via Thomson scattering[10] the inhomogeneous overdense plasmas generated in that way. More details

about the modeling of such FEL-experiments performed by the group of Prof. R. Redmer (University of Rostock) are given in the contribution "Thomson scattering on inhomogeneous targets" in this volume. Input for the radiation-hydrodynamics simulations in this context are the equilibrated charge and temperature distributions after the interaction with the long-wavelength, ionizing laser pulse. PIC simulations can deliver this input.

Moreover, we are interested in studying the possibility to generate more complex and structured inhomogeneous targets via tailored field configurations. It is expected, for instance, that in counter-propagating laser pulses the standing light wave pattern causes density modulations on the $\lambda/2$ length scale[11]. Furthermore, it is known that rapid field ionization may lead to scattering, frequency shifts, and trapping of electromagnetic radiation as well as ionization fronts, drift currents and related magnetic fields in the target[12]. Our goal is to explore these phenomena and their influence on the laser-to-particle coupling in full dimensionality.

Experimentalists in Rostock study the interaction of laser light with clusters embedded in Helium droplets[13]. If the laser intensity is too low for the ionization of He but high enough to ionize the embedded cluster, ionization avalanches from the centre to the droplet surface and shock-compression due to the Coulomb-explosion are expected. If, on the other hand, the intensity is high enough to directly ionize the He, the embedded cluster will be ionized mainly via collisions. However, for droplets with radii comparable to the wavelength, the scattered electromagnetic field is anticipated to show a complex spatio-temporal structure with field enhancements at the droplet surface and/or boundary. This, in turn, will influence the charge state distribution. It is interesting to investigate various scenarios in order to find out for which set-up the laser-matter coupling is particularly efficient so that high ion energies can be obtained.

2 Description of Methods and Algorithms

All simulations presented here were performed with the JUROPA Supercomputer at the Jülich Supercomputing Centre using the three-dimensional fully kinetic PIC code UMKA. UMKA employs state-of-the art, widely used numerical PIC algorithms such as the Yee Cartesian lattice for electromagnetic fields, the Boris pusher to integrate the Lorentz force and the current reconstruction scheme to satisfy the continuity equation[14]. UMKA results have been benchmarked with other codes, compared both with analytical modeling and experimental results and reported in several publications[15]. Generally speaking, a PIC code implements a particle-grid numerical method, where a collisionless plasma is sampled by a large number of computational particles and electromagnetic fields are discretized on a grid. The computational particles move across the grid cells and are accelerated by the electromagnetic fields defined at grid points, which are computed by reconstructing the source current density from the particle positions and velocities.

Tunneling and impact ionization have been implemented in the code. Ionization of an ion with charge state $\mathcal{Z}-1$ and ionization potential I due to the electric field $\mathcal{E}$ of both the laser and the mean field generated by all particles is implemented using the tunneling ionization rate formula[16]

$$w(\mathcal{E}) = \left(\frac{2\mathcal{E}_{ch}}{\mathcal{E}}\right)^{2n^*} \frac{k^2\hbar}{m}\frac{\mathcal{E}}{\mathcal{E}_{ch}} \exp\left(-\frac{2\mathcal{E}_{ch}}{3\mathcal{E}}\right) \tag{1}$$

with

$$k = \frac{\sqrt{2mI}}{\hbar},\ \mathcal{E}_{ch} = \frac{\hbar^2 k^3}{me},\ n^* = \mathcal{Z}\sqrt{\frac{I_H}{I}}. \tag{2}$$

Here m is the electron mass and I_H is the ionization potential of a hydrogen atom. The ionization probability for the i^{th} ion during time τ is then calculated from

$$P_i = \tau w(\mathcal{E}). \tag{3}$$

An ionization occurs if a uniformly distributed random number on the interval [0,1] is less than P_i. When an ionization event takes place a free electron is created at rest at the position of the ion. The energy needed for ionization has to be subtracted from the laser field, as only external fields can do work. To that end, after an ionization event has occurred, energy is taken out of the field via an "ionization current" $\vec{j}_{\rm ion}$ parallel to the electric field at the ion location. The value of the "ionization current" is such that $\vec{j}_{\rm ion} \cdot \vec{E}$ is the work spent on ionization per time step[17]. The energy conservation is taken care of during the whole process, i.e., if the remaining field energy after the ionization of particles in a cell is not sufficient for further ionization, this cell will not be considered anymore during the current time step.

In order to address the process of impact ionization we use the tabulated data of impact ionization cross-sections $\sigma_i(v_e)$[18]. The total collision ionization frequency for an electron with velocity v_e is then calculated from

$$\nu_{\rm tot}(v_e) = v_e \sum_{i=0}^{Z-1} \sigma_i(v_e) n_i, \tag{4}$$

and the total collision-ionization probability is

$$P_e = \tau \nu_{\rm tot}. \tag{5}$$

When an ionization event takes place the charge state of the ion under consideration is changed, and a free electron is created at rest at the position of the ion. The value of the impacting electron velocity is reduced in order to ensure energy conservation. Numerically, we use the Monte-Carlo null collision method[19] for the treatment of electron-impact ionization. Excited electronic states are not considered, neither are recombination processes.

The main building blocks of the code are illustrated in Fig. 1.

1. Assignment of the field grid data to the particle positions to obtain the accelerations and integration of the (relativistic) equations of motion for the particles in time;

2. Treatment of the ionization processes;

3. Assignment of the particle charges to the grid for the determination of the current density;

4. Integration of the Maxwell equations in space and time.

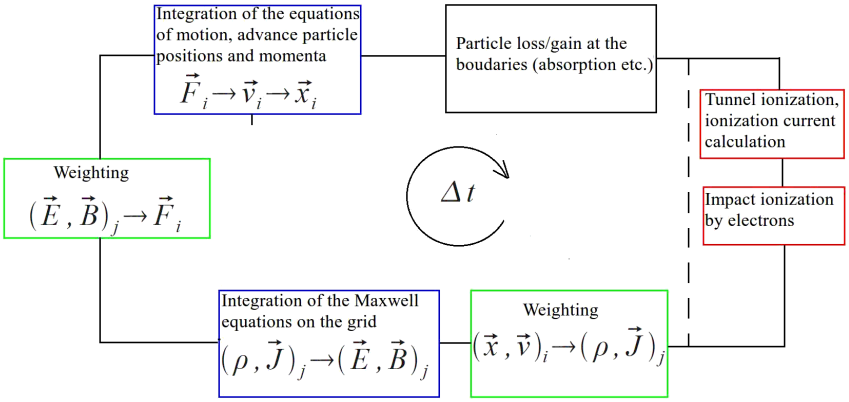


Figure 1. Extended PIC cycle.

From a computational point of view, items (1), (2) and (3), governing the particle-grid connection, are usually the most time-consuming, not only because of the large number of floating point operations, but also because addresses of particles and field data reside in different parts of the computer memory. Efficient parallel implementation of a PIC code is usually achieved by a domain decomposition technique, in such a way that each subdomain assigned to each processing element (PE) contains a comparable number of particles. Under these conditions it is evident that the work needed to exchange data among nearby subdomains is only a minor fraction of the total computational work.

In the UMKA code, parallelization is implemented by a decomposition of the (y, z)-discretized coordinates, perpendicular to the x-propagation direction of the laser pulse. This choice provides an initially balanced partition of the particle numbers for problems where the particle density is initially uniform along the (y, z)-plane. As particles move across different domains, a dynamical load balancing technique would be required to keep the data partition balanced. So far, a dynamical load balancing has not been implemented in UMKA. However, we have made some modification of the code using the HDF5 library for storing and managing data. With these modifications, the code allows us to choose a new, more balanced partitioning when restarting the simulation from a check-point. It is possible not only to change the partitioning while keeping the same number of PEs, but also to choose a different number of PEs.

3 Interaction of Intense Laser Pulse with Neutral He-Droplets

In this section we present typical results from PIC simulations of the interaction of an intense plane wave laser pulse with a neutral He droplet. For the PIC modeling of femtosecond ionization dynamics a high numerical resolution is crucial. The resolution must be sufficient to prevent field ionization from the field of a single macrocharge. We have used 100 cells per laser wavelength and 125 macroions per cell. The initial number of macroelectrons per cell is equal to 250 (the "naked" ion charge state times the number of macroions). Absorbing boundary conditions for the fields and for the particles were em-

ployed in propagation direction, periodic ones for the other directions. A linearly (in y direction) polarized 10-cycles $\sin^2$ -laser pulse enters the numerical box through the left boundary $x = 0$ and propagates into the $x > 0$ half-space. The dimensionless vector potential amplitude $a = |e\tilde{A}/mc| = |e\hat{E}/m\omega c|$ of the laser pulse ($\lambda = 2\pi c/\omega = 800$ nm) was $a = 0.5$, corresponding to a laser intensity $I \simeq 5.2 \times 10^{17}$ W/cm^2. The density of the 3.2μm diameter He-droplet was $\rho = 0.14$ g/cm^{-3}. If the droplet were completely preionized such a density would correspond to an electron density $n_{e0} = 22 \times n_{cr}$, where $n_{cr} = 1.8 \cdot 10^{21}$cm^{-3} is the critical density for $\lambda = 800$ nm wavelength light. In this simulation the impact ionization process was not considered, because for the given intensity its influence is much smaller than that of tunneling ionization.

Fig. 2 shows the distributions of the droplet's components (electrons, He^{1+} and He^{2+}) densities in two perpendicular planes $(\hat{\mathbf{k}}, \hat{\mathbf{E}})$ and $(\hat{\mathbf{k}}, \hat{\mathbf{B}})$ at times $t = 7$ and $t = 11$ cycles (corresponding to 18.5 and 32 fs accordingly). The distributions of the electric field components (longitudinal E_x and in the polarization direction of the laser electric field E_y) are shown in Fig. 3.

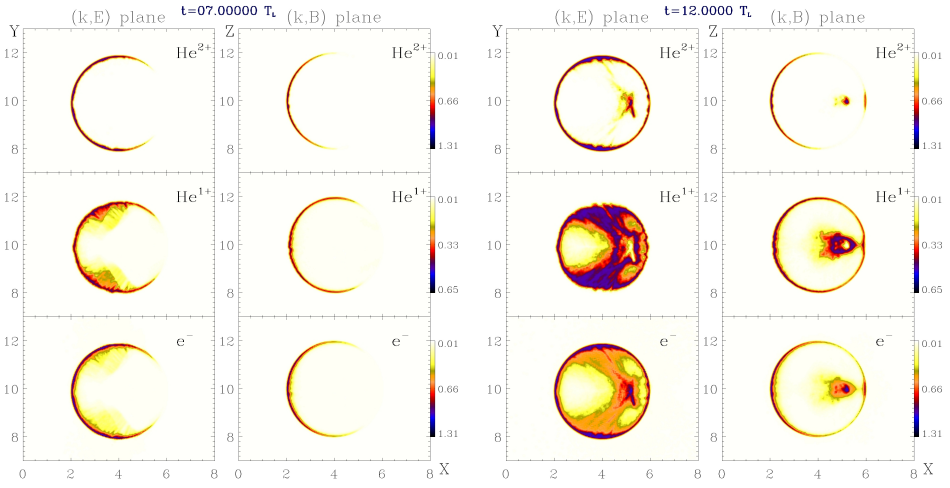


Figure 2. Densities of the droplet's components: electrons (bottom), He^{1+} (middle) and He^{2+} (top) in two perpendicular planes $(\hat{\mathbf{k}}, \hat{\mathbf{E}})$ (left) and $(\hat{\mathbf{k}}, \hat{\mathbf{B}})$ (right) at $t = 18.5$ and $t = 32$ fs. All lengths are measured in units of the laser wavelength λ.

In the beginning of the interaction the droplet is not ionized and the leading part of the laser pulse can penetrate through it. Later, ionization becomes more efficient and an overdense plasma is generated rapidly on the droplet surface. Subsequently, the laser pulse is reflected and only within a thin skin layer at the front higher charge states may be produced (see Fig. 2 upper frames for $t = 7$ cycles). Moreover, our simulations show that there is an angle of incidence at which oscillating electric fields penetrate into the droplet (see Fig. 3, $t = 7$ cycles), ionizing its interior (see Fig. 2). This leads to the formation of a highly asymmetric density distribution, concentrated mostly in the polarization plane, with the higher charge state not only within the skin layer on the surface but also in the droplet

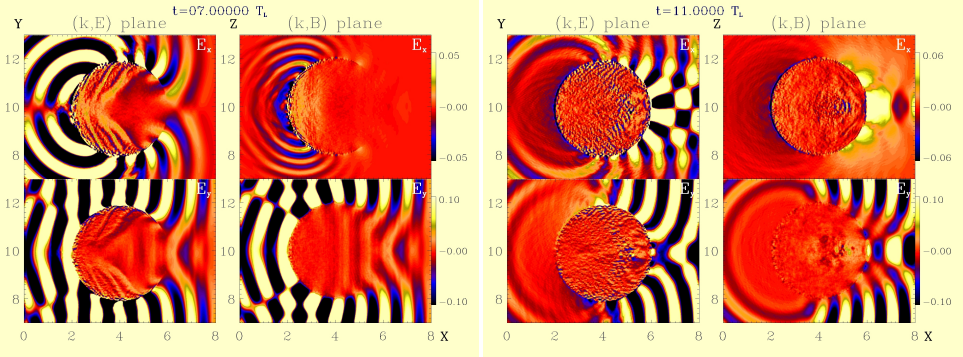


Figure 3. Electric fields (E_x-top, E_y-bottom) in two perpendicular planes ($\hat{\mathbf{k}}, \hat{\mathbf{E}}$) (left) and ($\hat{\mathbf{k}}, \hat{\mathbf{B}}$) (right) at $t = 18.5$ and $t = 29$ fs. All lenths are measured in units of the laser wavelength λ.

interior (see Fig. 2, upper frames at $t = 12$ cycles). Further investigations of this process are in progress.

4 Conclusion

The ionization dynamics during the interaction of intense laser pulses with neutral He-droplets has been studied by means of particle-in-cell simulations with ionization included. The numerical simulations have shown that there exists a broad laser intensity regime for which wavelength-sized targets are not fully ionized. For higher-Z material this applies even to ultra-high intensities. Hence, for realistic simulations of laser-plasma interaction ionization needs to be taken into account.

Acknowledgements

This work was supported by the DFG within the SFB 652 “Strong correlations and collective effects in radiation fields: Coulomb systems, clusters and particles”. PIC simulations were performed using the computing resources granted by the NIC/JSC of the Forschungszentrum Jülich under the project HRO01.

References

1. C. E. Clayton et al., Phys. Rev. Lett. 70, 37.(1993); Modena et al., Nature 337, 606 (1995); K. Nakajima et al., 1995, Phys. Rev. Lett. 74, 4428. (1995); D. Umstadter et al., Science 273, 472 (1996); D. Umstadter, J. K. Kim, E. Dodd, 1996, Phys. Rev. Lett. 76, 2073 (1996); R. Wagner et al., Phys. Rev. Lett. 78, 3125 (1997); S.-Y. Chen et al., Phys. Rev. Lett. 80, 2610 (1998); D. Gordon et al., Phys. Rev. Lett. 80, 2133 (1998); R. Bingham, J. T. Mendonca, P. K. Shukla, Plasma Phys. Controlled Fusion 46, R1 (2004); V. Malka et al., 2002, Science 298, 1596 (2002); C. G. R. Geddes, 2004, Nature 431, 538 (2004); J. Faure et al., Nature 431, 541 (2004); S. P.

D. Mangles et al., Nature 431, 535 (2004); S. Ya. Tochitsky et al., Phys. Rev. Lett. 92, 095004 (2004); T. Zh. Esirkepov et al., JETP Lett. 70, 82 (1999); K. Krushelnick et al. Phys. Rev. Lett. 83, 737 (1999); G. S. Sarkisov et al., Phys. Rev. E 59, 7042 (1999); S. V. Bulanov et al., JETP Lett. 71, 407 (2000); E. L. Clark, et al., Phys. Rev. Lett. 84, 670 (2000); Phys. Rev. Lett. 85, 1654 (2000); A. S. Maksimchuk et al., Phys. Rev. Lett. 84, 4108 (2000); D. Umstadter, J. Phys. D 36, R151 (2003); A. Maksimchuk et al., Plasma Phys. Rep. 30, 473 (2004); C. Joshi, V. Malka, New J. Phys. 12 045003 (2010).

2. T. Tajima, J.M. Dawson, PRL 43, 267 (1979).
3. A. Pukhov, J. Meyer-ter-Vehn, Appl. Phys. B 74, 355 (2002).
4. W. L. Kruer, K. Estabrook, Phys. Fluids 28, 430 (1985); F. Brunel, Phys. Rev. Lett. 59, 52 (1987); Phys. Fluids 31, 2714 (1988); P. Gibbon "Short Pulse Laser Interactions with Matter, an Introduction" (Imperial College Press, London 2005); C. H. Jungreuthmayer et al., Phys. Rev. Lett. 92, 133401 (2004); T.-Y. Bryan Yang et al., Phys. Plasmas 2, 3146 (1995); Y. Sentoku et al., Appl. Phys. B 74, 207 (2002); V. S. Rastunkov, V. P. Krainov, Laser Phys. 15, 262 (2005); D. F. Zaretsky et al., J. Phys. B 37, 4817 (2004); D. Bauer, P. Mulser, Phys. Plasmas 14, 023301 (2007).
5. M. Borghesi et al., Fusion Sci. Technol. 49 412 (2006); P. McKenna et al. Plasma Phys. Control. Fusion 49 B22331 (2007).
6. T. Esirkepov et al., Phys. Rev. Lett. 92 175003(2004); T. Esirkepov, M. Yamagiwa, T. Tajima, Phys. Rev. Lett. 96 105001 (2006).
7. A. Macchi, S. Veghini, F. Pegoraro, Phys. Rev. Lett. 103, 085003 (2009).
8. Yu Tong-Pu et al., Phys. Rev. Lett. 105, 065002 (2010).
9. C. K. Birdsall, A. B. Langdon, "Plasma physics via computer simulation"(Taylor & Francis, 2004); R.W Hockney, J.W Eastwood, "Computer Simulation Using Particles" (Taylor & Francis, 1989).
10. S.H. Glenzer, R. Redmer, Rev. Mod. Phys. 81, 1625 (2009).
11. E. Conejero Jarque, F. Cornolti, A. Macchi, J. Phys. B: At. Mol. Opt. Phys. 33 1 (2000).
12. A. Macchi et al., Phys. Rev. E 59, R36 (1999).
13. N.X. Truong et al., Phys. Rev. A 81, 013201 (2010).
14. J. Villasenor, O. Bunemann, Comput. Phys. Commun. 69, 306 (1992); T. Zh. Esirkepov, Comput. Phys. Commun. 135 (2), 144 (2001).
15. A. Macchi et al., Phys. Rev. Lett. 94, 165003 (2005); Y. Sentoku et al., Phys. Rev. E 62, 7271 (2000); T. V. Liseikina et al., Phys. Rev. E 60, 5991 (1999); A. Macchi et al., New J. Phys. 12, 045013 (2010); C. Benedetti et al., Nucl. Instr. Meth. A 606, 89 (2009); K. Quinn et al., Phys. Rev. Lett. 102, 194801 (2009); S. Kar et al., Phys. Rev. Lett. 100, 225004 (2008); L. Romagnani et al., Phys. Rev. Lett. 105, 175002 (2010); T.V. Liseykina, S. Pirner, D. Bauer, Phys. Rev. Lett. 104, 095002 (2010).
16. V.S. Popov, Phys. Usp. 47, 855 (2004).
17. P. Mulser, F. Cornolti, D. Bauer, Phys. Plasmas 5, 4466 (1998); S. C. Rae, K. Burnett, Phys. Rev. A 46, 20772083 (1992).
18. W. Lotz, ApJS, 1967; Zeitschrift fuer Physik 206, 205-211 (1967).
19. V. Vahedi, M. Surendra, Comput. Phys. Commun. 87, 199(1995).

Thomson Scattering on Inhomogeneous Targets

P. Sperling, T. Liseykina, D. Bauer, and R. Redmer

Institut für Physik, Universität Rostock, D-18051 Rostock, Germany
E-mail: philipp.sperling@uni-rostock.de

The introduction of free electron lasers enables new pump-probe experiments to characterize warm dense matter, i.e. systems at solid-like densities and temperatures of several eV. Such extreme conditions are relevant for the interior of giant planets and along the compression path of inertial confinement fusion capsules. Theoretical results for the pair distribution functions and the equation of state can now be checked using new experimental techniques. For instance, short-pulse optical lasers irradiate targets that are subsequently probed with brilliant X-ray radiation. The inhomogeneous plasma prepared by the optical laser is characterized with particle-in-cell simulations. The interaction of the X-ray probe radiation with the inhomogeneous plasma is also taken into account for different time delays between pump and probe via radiative hydrodynamic simulations.

For ongoing pump-probe experiments with liquid hydrogen and helium droplets, we calculate the respective scattering spectrum based on the Born-Mermin approximation for the dynamic structure factor considering the full density and temperature-dependent Thomson scattering cross section throughout the target. We can identify plasmon modes that are generated in different target regions and monitor their temporal evolution. Combining these results with the ion signal the temperature equilibration between electrons and ions can be studied as well. Therefore, such pump-probe experiments are promising tools to measure not only the important plasma parameters density and temperature but also to gain valuable information about their time-dependent profile through the target. The method described here can be applied to various pump-probe scenarios by combining optical lasers, soft and hard X-ray sources.

1 Introduction

X-ray Thomson scattering is a promising tool for the diagnostics of dense strongly correlated plasmas[1]. X-rays emitted from laser produced plasmas can probe the warm dense matter (WDM) region with temperatures of several eV and densities close to solid density[2] up to compressed matter well above solid density and electron temperatures T_e between 0.1 eV and several 10 eV[3]. These plasmas are opaque in the optical region since the frequency of light $\omega_0 = 2\pi c/\lambda_0$ is lower than the plasma frequency $\omega_{pe} = \sqrt{n_e e^2/(\epsilon_0 m_e)}$ of the free-electron subsystem, with the free-electron density n_e and the electron mass m_e. Therefore, probing plasmas with densities approaching solid or even higher densities requires efficient X-ray sources with high brightness which can be produced by energetic optical lasers. The study of dense plasmas is also possible with the implementation of free-electron lasers which provide brilliant radiation in the soft X-ray region as at the Free-electron LASer Hamburg (FLASH) at DESY, Hamburg[4,5] or in the X-ray region as at the Linac Coherent Light Source (LCLS) in Stanford, and at the future European XFEL in Hamburg.

Collective X-ray Thomson scattering experiments yield information on the density and temperature of dense plasmas[2]. Assuming a homogeneous density and temperature throughout the target, these parameters can be determined directly from the plasmon dispersion and the ratio of the plasmon amplitudes via the detailed balance relation[6]. For

longer laser wavelengths, i.e. optical lasers and wavelengths of about $40 - 50$ nm accessible with FLASH, the target is overdense and the absorption is limited to the skin depth, in contrast to excitation by short-wavelength FEL radiation where gradients are smaller[7,8]. The scattering signal represents an average of the local density- and temperature-dependent scattering cross sections weighted with the respective density and temperature profiles[8–10]. These profiles are generated by the pump pulse but also by the probe pulse if intense X-ray radiation pulses are used as in the case of the FLASH and LCLS facilities. Therefore, a realistic description of the light-matter interaction within a pump-probe experiment is of paramount importance to derive reliable information on the plasma parameters density and temperature, their profiles throughout the target, and their temporal evolution.

Here, we present the calculation of the Thomson scattering spectrum for inhomogeneous targets as generated by irradiation of liquid hydrogen and helium droplets with optical laser light. All simulations were performed with the JUROPA Supercomputer at the Jülich Supercomputing Centre (JSC). In Sec. 2 we outline simulations of the laser-plasma interaction for experiments performed with hydrogen at FLASH and calculate the corresponding Thomson scattering spectrum. In Sec. 3 we treat a similar setup with helium droplets which could be realized at FLASH and the future XFEL as well. Finally, we give conclusions.

2 Thomson Scattering on Inhomogeneous Hydrogen Droplets

In this section, a full pump-probe Thomson scattering experiment as performed at FLASH is calculated. Liquid hydrogen droplets with a diameter of 10 μm are pumped by an optical laser with a wavelength of $\lambda = 800$ nm, a pulse length of $t_{\mathrm{FWHM}} = 90$ fs, and an intensity of $I \approx 10^{15}$ W/cm^2. The pump pulse generates strong inhomogeneities in electron density and temperature in the droplet due to the short penetration depth of the optical laser. After a variable delay time of several picoseconds the droplet is probed by a soft X-ray FEL pulse with a wavelength of $\lambda = 13.5$ nm, a pulse length of $t_{\mathrm{FWHM}} = 200$ fs and an intensity of $I \approx 10^{13}$ W/cm^2. In contrast to the optical laser, the shorter soft X-ray probe radiation penetrates into the target and scatters from its total volume. In addition, the FEL radiation interacts with the droplet and yields further ionization which has to be considered. The spectrum of the scattered light is measured so that the plasma parameters can be extracted by comparing with the calculated data.

The simulation of such a pump-probe experiment includes three components. First, the interaction of the optical laser with the droplet (pump phase) was calculated using the 3d Particle-In-Cell (PIC) code VLPL3D[11] at the JUROPA supercomputer of JSC. Due to the short wavelength of the probe FEL radiation and the time delays of several picoseconds, PIC codes are too demanding for the present computing capacity even at the JSC. Therefore, we have used in a second step the 1d radiation-hydrodynamic code HELIOS[12] (i) to extend the simulations up to the nanosecond time scale, and (ii) to include also the interaction of the soft X-ray FEL radiation with the plasma. In a third step, the resulting density and temperature distribution throughout the droplet was used to calculate the full Thomson scattering spectrum for a scattering angle of $\theta_s = 90^\circ$ via the Born-Mermin-Approximation (BMA)[10]. Of special interest is the plasmon resonance from which the plasma density and temperature can be inferred[13].

We show the electron density distribution through half of the droplet and the corresponding Thomson scattering spectrum in Fig. 1. We can distinguish between two regions: a fully ionized surface layer generated by the optical laser (low penetration depth) and the interior of the droplet with a much lower ionization mainly due to the FEL radiation (high penetration depth). The Thomson scattering spectrum has an inelastic plasmon peak with an energy shift of $|\Delta E| \approx 8.5$ eV and an elastic scattering signal. The plasmon feature corresponds to an effective electron temperature of $T_e \approx 24$ eV and an electron density of $n_e = 5.3 \cdot 10^{22}$ cm^{-3} which can be related to the almost fully ionized surface layer of the droplet. The interior of the droplet with an electron density of $n_e \approx 5 \cdot 10^{20}$ cm^{-3} generates a plasmon with an energy shift of only $|\Delta E| = 0.83$ eV which is superimposed by the elastic scattering signal.

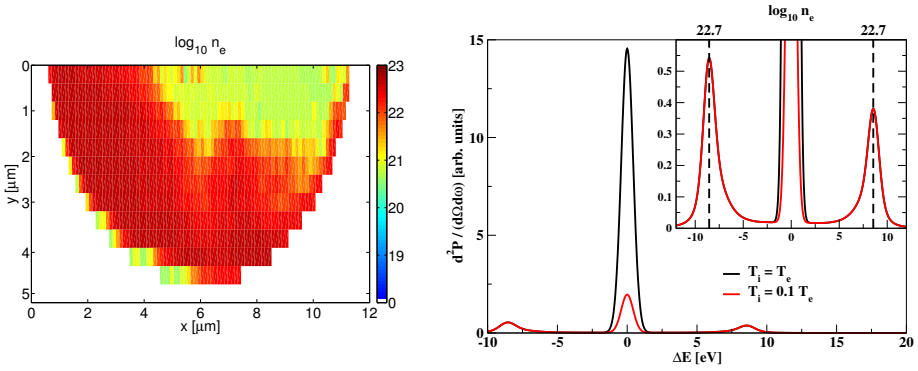


Figure 1. Electron density (left) and full Thomson scattering signal (right) for a pump-probe experiment with a delay time of 1 ps. The optical laser and the FLASH irradiate the droplet from the left. The front layer of the droplet is fully ionized and therefore opaque for the optical laser. The heating and ionization of the interior region is mainly caused by the FLASH radiation which can penetrate throughout the target due to its shorter wavelength. The resulting Thomson scattering spectrum is shown for two different ion temperatures which affect only the height of the elastic scattering signal.

3 Thomson Scattering on Inhomogeneous Helium Droplets

We propose a modified pump-probe experimental setup using helium instead of hydrogen. Helium is the prototypical two-electron system having two ionization energies ($E_1 = 24.587$ eV, $E_2 = 54.418$ eV) so that the more complex ultra-short time ionization dynamics may be studied via scattering experiments for the first time. To ionize the helium droplets of 6 μm diameter, we use a more intense optical laser with a wavelength of $\lambda = 800$ nm, a pulse length of $t_{\mathrm{FWHM}} = 30$ fs, and an intensity of $I \approx 10^{18}$ W/cm^2. This corresponds to a regime where the ionized electrons reach already relativistic velocities. Such a laser is currently not available at FLASH but plans for the high energy end station of the future XFEL in Hamburg include such experimental setups.

The interaction of the optical laser with the helium droplets is again calculated via a PIC code[14] for a duration up to 60 fs. More details about these kind of PIC simulations are given in the contribution “Plasma formation and ionization dynamics in intense

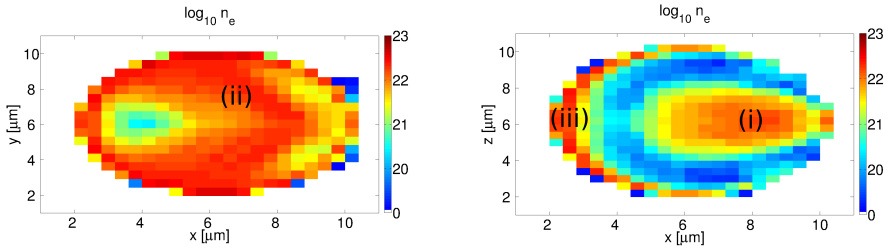


Figure 2. Electron density for two different droplet planes (in the laser-polarization plane and perpendicular to it) at a time of 0.06 ps. The optical laser irradiates the droplet from the left. We obtain a non-symmetric ionization which is caused by a focused penetration of the laser into the droplet. This is only possible with laser intensities close to the relativistic regime.

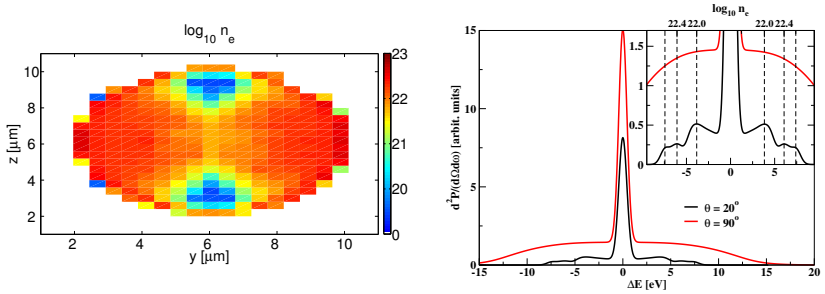


Figure 3. Electron density in the droplet for the plane perpendicular to the laser propagation (left) and the resulting Thomson scattering signal (right) at a time of 0.06 ps. The spectrum shows three different plasmons characterizing three different droplet regions, the ionized interior, the front surface layer and the laser focusing region inside the droplet.

laser-droplet interaction" by D. Bauer and T. Liseykina in this volume. The Thomson scattering spectrum is calculated via the BMA code. The electron density and the Thomson scattering spectrum for two different scattering angles and two different ion temperatures are shown in Fig. 2. A scattering angle of $\theta_s = 90^\circ$ represents non-collective scattering and the scattering angle $\theta_s = 20^\circ$ probes collective scattering. In addition, we observe three plasmon signals which characterize different regions of the droplet. We can derive the following plasma parameters effective electron density and temperature: (i) $n_e \approx 1.0 \times 10^{22}$ cm^{-3} and $T_e \approx 485.6$ eV for plasmons with an energy shift of $|\Delta E| \approx 3.8$ eV, (ii) $n_e \approx 2.7 \times 10^{22}$ cm^{-3} and $T_e \approx 780.5$ eV for plasmons with an energy shift of $|\Delta E| \approx 6.0$ eV, and (iii) $n_e \approx 4.0 \times 10^{22}$ cm^{-3} and $T_e = 850.3$ eV for plasmons with an energy shift of $|\Delta E| \approx 7.4$ eV. These plasma parameters correspond to the the interor (i), the focusing area (ii), and the surface (iii) of the droplet, see Fig. 3.

4 Conclusions

In this paper we have calculated the Thomson scattering signal off inhomogeneous hydrogen and helium droplets which are pumped with optical lasers and probed with FEL

radiation. For the pump phase we have applied the PIC codes VLPLD3D and another PIC code[14], on the JUROPA supercomputer of JSC. The probe phase and the evolution within the delay time is simulated using the radiation hydrodynamics code HELIOS. Subsequently we have derived the Thomson scattering signal via the BMA from the resulting density and temperature profiles. Important information on the density and temperature profiles throughout the targets and their temporal evolution can be gained from the Thomson scattering spectra. This will lead to new proposals for pump-probe experiments at advanced light sources with high peak brightness such as FLASH, LCLS, and the future European XFEL. Ultimate goals are a better understanding of matter under extreme conditions and of the ionization dynamics on ultra-short time scales.

Acknowledgements

This work was supported by the DFG within the SFB 652 "Strong correlations and collective effects in radiation fields: Coulomb systems, clusters, and particles" and the Federal Ministry of Education and Science (BMBF) under grant FSP 301-FLASH, project No. 05KS7HRA. PIC simulations were performed with the computing resources granted by the NIC at the Forschungszentrum Jülich under project HRO02 and HRO01.

References

1. S. H. Glenzer and R. Redmer, *X-ray Thomson scattering in high energy density plasmas*, Rev. Mod. Phys., **81**, 1625, 2009.
2. S. H. Glenzer, O. L. Landen, P. Neumayer, R. W. Lee, K. Widmann, S. W. Pollaine, R. J. Wallace, G. Gregori, A. Höll, T. Bornath, R. Thiele, V. Schwarz, W.-D. Kraeft, and R. Redmer, *Observations of Plasmons in Warm Dense Matter*, Phys. Rev. Lett., **98**, 065002, 2007.
3. A. L. Kritcher, P. Neumayer, J. Castor, T. Döppner, R. W. Falcone, O. L. Landen, H. J. Lee, R. W. Lee, E. C. Morse, A. Ng, S. Pollaine, D. Price, and S. H. Glenzer, *Ultrafast X-ray Thomson Scattering of Shock-Compressed Matter*, Science, **322**, 69, 2008.
4. S. Toleikis, R. R. Fäustlin, L. Cao, T. Döppner, S. Düsterer, E. Förster, C. Fortmann, S. H. Glenzer, S. Göde, G. Gregori, R. Irsig, T. Laarmann, H. J. Lee, B. Li, J. Mithen, K.-H. Meiwes-Broer, A. Przystawik, P. Radcliffe, R. Redmer, F. Tavella, R. Thiele, J. Tiggesbäumker, N. X. Truong, I. Uschmann, U. Zastrau, and Th. Tschentscher, *Soft X-ray scattering using FEL radiation for probing near-solid density plasmas at few electron volt temperatures*, High Energy Dens. Phys., **6**, 15, 2010.
5. R. R. Fäustlin, Th. Bornath, T. Döppner, S. Düsterer, E. Förster, C. Fortmann, S. H. Glenzer, S. Göde, G. Gregori, R. Irsig, T. Laarmann, H. J. Lee, B. Li, K.-H. Meiwes-Broer, J. Mithen, B. Nagler, A. Przystawik, H. Redlin, R. Redmer, H. Reinholz, G. Röpke, F. Tavella, R. Thiele, J. Tiggesbäumker, S. Toleikis, I. Uschmann, S. M. Vinko, T. Whitcher, U. Zastrau, B. Ziaja, and Th. Tschentscher, *Observation of ultrafast non-equilibrium collective dynamics in warm dense hydrogen*, Phys. Rev. Lett., **104**, 125002, 2010.

6. A. Höll, T. Bornath, L. Cao, T. Döppner, S. Düsterer, E. Förster, C. Fortmann, S. H. Glenzer, G. Gregori, T. Laarmann, K.-H. Meiwes-Broer, A. Przystawik, P. Radcliffe, R. Redmer, H. Reinholz, G. Röpke, R. Thiele, J. Tiggesbäumker, J. Toleikis, N. X. Truong, T. Tschentscher, I. Uschmann, and U. Zastrau, *Thomson scattering from near-solid density plasmas using soft x-ray free electron lasers*, High Energy Dens. Phys., **3**, 120, 2007.
7. B. Nagler, U. Zastrau, R. R. Fäustlin, S. M. Vinko, Th. Whitcher, A. J. Nelson, R. Sobierajsk, J. Krzywinski, J. Chalupsky, E. Abreu, S. Bajt, Th. Bornath, T. Burian, H. Chapman, J. Cihelka, T. Döppner, S. Düsterer, Th. Dzelzainis, M. Fajardo, E. Förster, C. Fortmann, E. Galtier, S. H. Glenzer, S. Göde, G. Gregori, V. Hajkov, P. Heimann, L. Juha, M. Jurek, F. Y. Khattak, A. R. Khorsand, D. Klinger, M. Kozlova, T. Laarmann, H. Ja Lee, R. W. Lee, K.-H. Meiwes-Broer, P. Mercere, W. J. Murphy, A. Przystawik, R. Redmer, H. Reinholz, D. Riley, G. Röpke, F. Rosmej, K. Saksl, R. Schott, R. Thiele, J. Tiggesbäumker, S. Toleikis, Th. Tschentscher, I. Uschmann, H. J. Vollmer, and J. S.Wark, *Turning solid aluminium transparent by intense soft X-ray photoionization*, Nat. Phys., **5**, 693, 2009.
8. P. Sperling, R. Thiele, B. Holst, C. Fortmann, S. H. Glenzer, S. Toleikis, Th. Tschentscher, and R. Redmer, *Two-color Thomson scattering at FLASH*, High Energy Dens. Phys., **7**, 145–149, 2011.
9. H. A. Baldis, J. Dunn, M. E. Foord, and W. Rozmus, *Thomson scattering diagnostic of solid density plasmas using x-ray lasers*, Rev. Sci. Instrum., **73**, 4223, 2002.
10. R. Thiele, P. Sperling, M. Chen, Th. Bornath, R. R. Fäustlin, C. Fortmann, S. H. Glenzer, W.-D. Kraeft, A. Pukhov, S. Toleikis, Th. Tschentscher, and R. Redmer, *Thomson scattering on inhomogeneous targets*, Phys. Rev. E, **82**, 056404, 2010.
11. A. Pukhov, *Three-dimensional electromagnetic relativistic particle-in-cell code VLPL (Virtual Laser Plasma Lab)*, J. Plasma Phys., **61**, no. 03, 425–433, 1999.
12. J. J. MacFarlane, I. E. Golovkin, and P. R. Woodruff, *HELIOS-CR-A 1-D radiation-magnetohydrodynamics code with inline atomic kinetics modeling*, J. Quant. Spectrosc. Radiat. Transfer, **99**, 381, 2006.
13. R. Thiele, T. Bornath, C. Fortmann, A. Höll, R. Redmer, H. Reinholz, G. Röpke, A. Wierling, S. H. Glenzer, and G. Gregori, *Plasmon resonance in warm dense matter*, Phys. Rev. E, **78**, 026411, 2008.
14. T. V. Liseykina, S. Pirner, and D. Bauer, *Relativistic Attosecond Electron Bunches from Laser-Illuminated Droplets*, Phys. Rev. Lett., **104**, 095002, 2010.

John von Neumann Institute for Computing
NIC Series

---

1. **Modern Methods and Algorithms of Quantum Chemistry - Proceedings**
   Winter School, 21 - 25 February 2000, Forschungszentrum Jülich, Germany
   edited by J. Grotendorst (2000), iv, 562 pages
   NIC Series Volume 1
   ISBN: 3-00-005618-1

2. **Modern Methods and Algorithms of Quantum Chemistry - Poster Presentations**
   Winter School, 21 - 25 February 2000, Forschungszentrum Jülich, Germany
   edited by J. Grotendorst (2000), xi, 77 pages
   NIC Series Volume 2
   ISBN: 3-00-005746-3

3. **Modern Methods and Algorithms of Quantum Chemistry - Proceedings, Second Edition**
   Winter School, 21 - 25 February 2000, Forschungszentrum Jülich, Germany
   edited by J. Grotendorst (2000), iv, 638 pages
   NIC Series Volume 3
   ISBN: 3-00-005834-6

4. **Nichtlineare Analyse raum-zeitlicher Aspekte der hirnelektrischen Aktivität von Epilepsiepatienten**
   by J. Arnhold (2000), ii, 118 pages
   NIC Series Volume 4
   ISBN: 3-00-006221-1

5. **Elektron-Elektron-Wechselwirkung in Halbleitern: Von hochkorrelierten kohärenten Anfangszuständen zu inkohärentem Transport**
   by R. Lövenich (2000), iv, 145 pages
   NIC Series Volume 5
   ISBN: 3-00-006329-3

6. **Erkennung von Nichtlinearitäten und wechselseitigen Abhängigkeiten in Zeitreihen**
   by A. Schmitz (2001), vi, 126 pages
   NIC Series Volume 6
   ISBN: 3-00-007871-1

7. **Multiparadigm Programming with Object-Oriented Languages - Proceedings**
   1st Workshop MPOOL, 18 June 2001, Budapest, Hungary
   edited by K. Davis, Y. Smaragdakis, J. Striegnitz (2001), iv, 160 pages
   NIC Series Volume 7
   ISBN: 3-00-007968-8

---

8. Europhysics Conference on Computational Physics - Book of Abstracts
Aachen, Germany, 5 - 8 September 2001
edited by F. Hossfeld, K. Binder (2001), 500 pages
NIC Series Volume 8
ISBN: 3-00-008236-0

9. NIC Symposium 2001 - Proceedings
5 - 6 December 2001, Jülich, Germany
edited by H. Rollnik, D. Wolf (2002), v, 514 pages
NIC Series Volume 9
ISBN: 3-00-009055-X

10. Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms - Lecture Notes
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, The Netherlands
edited by J. Grotendorst, D. Marx, A. Muramatsu (2002), v, 548 pages
NIC Series Volume 10
ISBN: 3-00-009057-6

11. Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms- Poster Presentations
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, The Netherlands
edited by J. Grotendorst, D. Marx, A. Muramatsu (2002), xiii, 85 pages
NIC Series Volume 11
ISBN: 3-00-009058-4

12. Strongly Disordered Quantum Spin Systems in Low Dimensions: Numerical Study of Spin Chains, Spin Ladders and Two-Dimensional Systems
by Y. Lin (2002), iv, 128 pages
NIC Series Volume 12
ISBN: 3-00-009056-8

13. Multiparadigm Programming with Object-Oriented Languages – Proceedings
Workshop MPOOL 2002, 11 June 2002, Malaga, Spain
edited by J. Striegnitz, K. Davis, Y. Smaragdakis (2002), iv, 133 pages
NIC Series Volume 13
ISBN: 3-00-009099-1

---

14. Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms - Audio-Visual Lecture Notes
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, The Netherlands
edited by J. Grotendorst, D. Marx, A. Muramatsu (2002)
NIC Series Volume 14
ISBN: 3-00-010000-8, DVD

15. Numerical Methods for Limit and Shakedown Analysis
Deterministic and Probabilistic Problems
edited by M. Staat, M. Heitzer (2003), xiii, 282 pages
NIC Series Volume 15
ISBN: 3-00-010001-6

16. Design and Evaluation of a Bandwidth Broker that Provides Network Quality of Service for Grid Applications
by V. Sander (2003), viii, 184 pages
NIC Series Volume 16
ISBN: 3-00-010002-4

17. Automatic Performance Analysis on Parallel Computers with SMP Nodes
by F. Wolf (2003), viii, 142 pages
NIC Series Volume 17
ISBN: 3-00-010003-2

18. Haptisches Rendern zum Einpassen von hochaufgelösten Molekülstrukturdaten in niedrigaufgelöste Elektronenmikroskopie-Dichteverteilungen
by S. Birmanns (2003), viii, 152 pages
NIC Series Volume 18
ISBN: 3-00-010004-0

19. Auswirkungen der Virtualisierung auf den IT-Betrieb
Proceedings, GI Conference, 4 - 5 November 2003, Forschungszentrum Jülich
edited by W. Gürich (2003), 110 pages
NIC Series Volume 19
ISBN: 3-00-009100-9

20. NIC Symposium 2004
Proceedings, 17 - 18 February 2004, Jülich, Germany
edited by D. Wolf, G. Münster, M. Kremer (2004), v, 463 pages
NIC Series Volume 20
ISBN: 3-00-012372-5

---

21. Measuring Synchronization in Model Systems and Electroencephalographic Time Series from Epilepsy Patients
by T. Kreutz (2004), ii, 120 pages
NIC Series Volume 21
ISBN: 3-00-012373-3

22. Computational Soft Matter: From Synthetic Polymers to Proteins - Poster Abstracts
Winter School, 29 February - 6 March 2004, Gustav-Stresemann-Institut Bonn
edited by N. Attig, K. Binder, H. Grubmüller, K. Kremer (2004), x, 103 pages
NIC Series Volume 22
ISBN: 3-00-012374-1

23. Computational Soft Matter: From Synthetic Polymers to Proteins - Lecture Notes
Winter School, 29 February - 6 March 2004, Gustav-Stresemann-Institut Bonn
edited by N. Attig, K. Binder, H. Grubmüller, K. Kremer (2004), iii, 427 pages
NIC Series Volume 23
ISBN: 3-00-012641-4

24. Synchronization and Interdependence Measures and their Applications to the Electroencephalogram of Epilepsy Patients and Clustering of Data
by A. Kraskov (2004), iii, 89 pages
NIC Series Volume 24
ISBN: 3-00-013619-3

25. High Performance Computing in Chemistry
Report of the Joint Research Project: High Performance Computing in Chemistry - HPC-Chem
edited by J. Grotendorst (2004), iv, 149 pages
NIC Series Volume 25
ISBN: 3-00-013618-5

26. Zerlegung von Signalen in unabhängige Komponenten: Ein informationstheoretischer Zugang
by H. Stögbauer (2005), iii, 90 pages
NIC Series Volume 26
ISBN: 3-00-013620-7

27. Multiparadigm Programming 2003
Joint Proceedings of the
3rd International Workshop on Multiparadigm Programming with Object-Oriented Languages (MPOOL'03)
and the
1st International Workshop on Declarative Programming in the Context of Object-Oriented Languages (PD-COOL'03)
edited by J. Striegnitz, K. Davis (2005), ii, 294 pages
NIC Series Volume 27
ISBN: 3-00-016005-1

28. Integration von Programmiersprachen durch strukturelle Typanalyse und partielle Auswertung
by J. Striegnitz (2005), 296 pages
NIC Series Volume 28
ISBN: 3-00-016006-X

29. OpenMolGRID - Open Computing Grid for Molecular Science and Engineering
Final Report
edited by M. Romberg (2005), iii, 73 pages
NIC Series Volume 29
ISBN: 3-00-016007-8

30. GALA Grünenthal Applied Life Science Analysis
edited by A. Kless and J. Grotendorst (2006), iv, 204 pages
NIC Series Volume 30
ISBN: 3-00-017349-8

31. Computational Nanoscience: Do It Yourself!
Lecture Notes
Winter School, 14 - 22 February 2006, Forschungszentrum Jülich, Germany
edited by J. Grotendorst, S. Blügel, D. Marx (2006), v, 527 pages
NIC Series Volume 31
ISBN: 3-00-017350-1

32. NIC Symposium 2006 - Proceedings
1 - 2 March 2006, Jülich, Germany
edited by G. Münster, D. Wolf, M. Kremer (2006), iv, 365 pages
NIC Series Volume 32
ISBN: 3-00-017351-X

33. Parallel Computing: Current & Future Issues of High-End Computing
Proceedings of the International Conference ParCo 2005
edited by G.R. Joubert, W.E. Nagel, F.J. Peters, O. Plata, P. Tirado, E. Zapata (2006), xiv, 930 pages
NIC Series Volume 33
ISBN: 3-00-017352-8

John von Neumann Institute for Computing
NIC Series

---

34. From Computational Biophysics to Systems Biology 2006 Proceedings
edited by U.H.E. Hansmann, J. Meinke, S. Mohanty, O. Zimmermann (2006), iv, 214 pages
NIC Series Volume 34
ISBN-10: 3-9810843-0-6, ISBN-13: 978-3-9810843-0-6

35. Dreistufig parallele Software zur Parameteroptimierung von Support-Vektor-Maschinen mit kostensensitiven Gütemaßen
by T. Eitrich (2007), xiv, 231 pages
NIC Series Volume 35
ISBN: 978-3-9810843-1-3

36. From Computational Biophysics to Systems Biology (CBSB07) Proceedings
edited by U.H.E. Hansmann, J. Meinke, S. Mohanty, O. Zimmermann (2007), vi, 312 pages
NIC Series Volume 36
ISBN: 978-3-9810843-2-0

37. Parallel Computing: Architectures, Algorithms and Applications - Book of Abstracts of the International Conference ParCo 2007
ParCo 2007 Conference, 4 - 7 September 2007
edited by G.R. Joubert, C. Bischof, F.J. Peters, T. Lippert, M. Bücker, P. Gibbon, B. Mohr (2007), xiv, 195 pages
NIC Series Volume 37
ISBN: 978-3-9810843-3-7

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
NIC Series Volume 38
ISBN: 978-3-9810843-4-4

39. NIC Symposium 2008 - Proceedings
20 - 21 February 2008, Jülich, Germany
edited by G. Münster, D. Wolf, M. Kremer (2008), iv, 358 pages
NIC Series Volume 39
ISBN: 978-3-9810843-5-1

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
NIC Series Volume 40
ISBN: 978-3-9810843-6-8

John von Neumann Institute for Computing
NIC Series

---

41. Multigrid methods for structured grids and their application in particle simulation
by M. Bolten (2008), viii, 132 pages
NIC Series Volume 41
ISBN: 978-3-9810843-7-5

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
NIC Series Volume 42
ISBN: 978-3-9810843-8-2

43. Towards the Confirmation of QCD on the Lattice
Improved Actions and Algorithms
by S. F. Krieg (2008), vi, 89 pages
NIC Series Volume 43
ISBN: 978-3-9810843-9-9

44. NIC Symposium 2010 – Proceedings
24 - 25 February 2010 | Jülich, Germany
edited by G. Münster, D. Wolf, M. Kremer (2012), v, 395 pages
NIC Series Volume 44
ISBN: 978-3-89336-757-3

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
NIC Series Volume 45
ISBN: 978-3-89336-758-0

The John von Neumann Institute for Computing (NIC) was established in 1998 by Forschungszentrum Jülich and Deutsches Elektronen-Synchrotron DESY to support the supercomputer-oriented simulation sciences. In 2006, GSI Helmholtzzentrum für Schwerionenforschung joined NIC as a contract partner.

The core task of NIC is the peer-reviewed allocation of supercomputing resources to computational science projects in Germany and Europe. The NIC partners also support supercomputer-aided research in science and engineering through a three-way strategy:

- Provision of supercomputing resources for projects in science research and industry.
- Supercomputer-oriented research and development by research groups in selected fields of physics and natural sciences.
- Education and training in all areas of supercomputing by symposia, workshops, summer schools, seminars, courses amd 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. This year's symposium stands out since it marks the 25th anniversary of the founding of the NIC's immediate predecessor, the Höchstleistungsrechenzentrum (HLRZ), in 1987.

The contributions for this sixth NIC Symposium are from projects that have been supported by the supercomputers JUROPA and JUGENE in Jülich. They cover selected topics in the fields of Astrophysics, Biophysics, Chemistry, Elementary Particle Physics, Materials Science, Condensed Matter, Computational Soft Matter Science, Earth and Environmental Research, Computer Science, Hydrodynamics and Turbulence, and Plasma Physics.

JÜLICH FORSCHUNGSZENTRUM

GSI

Deutsches Elektronen-Synchrotron DESY

NIC Series Volume 45

ISBN 978-3-89336-758-0