ZIB

Annual Report 2013

PREFACE

Der Jahresbericht 2013 des Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB) bricht mit einigen Traditionen. From now on it is in English and called Annual Report. Introducing a new layout with various graphical components we try to render our report visually more appealing. A few broad surveys about current hot and important topics of our work, targeting a general audience, hopefully make the report more concise and readable. The readers interested in the specifics of ZIB's activities can find the technical details of our research and services, of our projects, and the statistics about publications, lectures, etc. on ZIB's Web pages (the starting point is www.zib.de).

With the official introduction of Christof Schütte as Vice President on January 10, ZIB had an excellent start into the year 2013. Christof immediately began to implement his ideas and to open new perspectives for ZIB.

The untimely death of Michael Hofmeister and the resignation of Thomas Liebling required a renewal of ZIB's Scientific Advisory Board. We are happy to welcome the new members Rainer Burkard (TU Graz, Austria), Michael Dellnitz (U Paderborn), Anna Schreieck (BASF, Ludwigshafen) and Kerstin Waas (Deutsche Bahn, Frankfurt) who, together with the continuing members Alfred Louis (U Saarbrücken), Jörg-Rüdiger Sack (Carlton U, Ottawa, Canada), Ludger D. Sax (Open Grid Europe, Essen), Horst D. Simon (Lawrence Berkely National Laboratory, Berkely, CA, USA), Reinhard Uppenkamp (Berlin Chemie, Berlin), will review our work and advise us on our future activities. At its June meeting, the Scientific Advisory Board elected J.-R. Sack as its new Chair. ZIB thanks his predecessor, A. Louis, for his long-term service as Chair of the Board.

The Board's evaluation of ZIB in June summarizes its findings as follows: "Wie in den Vorjahren sind an das ZIB sehr viele wissenschaftliche Auszeichnungen vergeben worden. Hervorzuheben ist dabei, dass regelmäßig junge wissenschaftliche Mitarbeiterinnen und Mitarbeiter für ihre Arbeit ausgezeichnet werden, dass also im ZIB erfolgreiche Nachwuchsarbeit betrieben wird. Zusammenfassend sieht der Wissenschaftliche Beirat das ZIB mit seiner eingeleiteten Neuausrichtung auf einem zukunftsweisenden Weg, der die hohe Forschungsqualität, das internationale Ansehen sowie die Rolle als Motor für die Berliner mathematiknahe angewandte Forschung bestens sichert. Der Leuchtturm ZIB wird auch in Zukunft große Strahlkraft haben und dies zum Wohle Berlins einsetzen."

Supercomputing at ZIB started in 1984 with the legendary Cray 1M. It entered its eighth generation in September 2013 when the new Cray XC30 was put into service. The two tightly linked systems, installed at ZIB and at the Leibniz Universität in Hannover and jointly called HLRN-III, will comprise in its final configuration 85,000 cores, 2,5 petaflops peak performance, 222 terabyte memory and 7,2 petabyte disk space for top researchers in application areas such as earth, material and life sciences, chemistry, physics and engineering who heavily employ advanced methods of scientific computing. ZIB does not only offer hardware, it continues to serve its external customers via HPC software consultants who, since 1984, help to make best use of this high performance computing environment. The HLRN-III is financed by the North-German Supercomputing Alliance (HLRN), comprising Berlin, Brandenburg, Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen and Schleswig-Holstein. The operation costs have become, mainly due to very high energy prices, a major burden – shared since 2013 by all HLRN members.

One of the new features of this Annual Report is that, instead of touching all aspects of our work, we provide only a few reports, partly crossing department boundaries, which combine brief surveys of very active topics with outlines of highlights of our own work – keeping global developments in mind.

The departments Numerical Analysis and Modeling and Visualization and Data Analysis feature two cross-department reports: "Solutions for Individualized Medicine" illustrates how ZIB supports modern medicine by the development of algorithms and demonstrators facilitating computer-assisted, model-guided therapy planning, and that therapeutic interventions often start with groundbreaking research in computing or data science. In "Design of Functional Molecules" the focus is on new and very efficient modeling techniques and simulation approaches that help to design molecules with desired functions, as, e.g., new pain relief drugs, but at the same time allow for the estimation of toxic risks and other adverse side effects.

The department Optimization indicates in its report "Mathematics in Traffic and Transport" that public transport, air and railway traffic, and logistics can be viewed as systems of gigantic optimization problems. Since many years ZIB has been contributing to the progress in traffic optimization by advancing the mathematical theory of large-scale transportation problems and by developing state-of-the-art algorithms to solve them – employing modern information technology. Today's challenges arise from urbanization, aging societies and the scarcity of resources.

"Shepherding the Bits – Challenge Accepted" is the title of the report by the department Scientific Information which describes ZIB's contributions to preserving the cultural heritage, in particular in its digital form. This ranges from the administration and advancement of the Cooperative Library Network Berlin-Brandenburg (KOBV), through our museum project that develops and maintains software for the indexing and administration of museum collections (GOS), to the project digiS which supports digitization efforts of cultural assets in Berlin.

The two departments of the computer science division of ZIB were very active in preparing the transition to the new HLRN-III system, see above. Of course, research continued as well. Highlighting its focus on methods for large-scale data analytics, the department Distributed Algorithms and Supercomputing describes in its report "Similarity Search and Insect Classification" the difficulties computers have in understanding the similarity of objects. Once meaningful similarity metrics and the like are designed, the methodology's use can range from taxonomy in biology to data handling in file systems.

More time than expected was spent preparing the establishment of the Einstein Center Mathematics that (in cooperation with FU, HU, TU and WIAS) will take up operation in June 2014. Even more effort went into the layout of the legal framework and the project design of the Research Campus MODAL that will go full speed (after a hopefully successful evaluation) in the fall of 2014. By winning a number of new projects and the third-party financing associated with these, ZIB was able to continue its research and service activities at its standard high level. Without the exceptionally motivated members in research, service and administration this would be impossible.

Finally, ZIB is grateful for having many supporters and friends. Thanks to all those who provide us with encouragement, assistance and advice.

Berlin, April 2014
Martin Grötschel
ZIB President

CONTENTS

EXECUT SUMMA

DISCRETE MATHEMATICS

All branches of optimization are currently witnessing a strong growth of the size of the industry problems to be addressed and of the range of requirements (linear, integer, nonlinear, stochastic, robust, multi-objective, ...) to be satisfied. The department Optimization reacts to this, e.g. within its working group Linear and Nonlinear Integer Optimization, by advancing the theory and algorithmic methodology (parallelization, utilization of supercomputers, etc.) for solving large-scale problems of this type. General progress would not exist here without close collaboration with our working groups on Energy, Telecommunications, Traffic and Logistics (the area covered in detail in this annual report) where we continue our application-oriented research in seventeen projects, mostly with close ties to industries.

The department Scientific Information focusses on planning and developing information systems and services for the sciences and specializes, in particular, on libraries (KOBV), museums (GOS), and support for digitization of cultural object (digiS).The article "Shepherding the bits" describes our efforts in this area in 2013.

TIVE
ARY

NUMERICAL MATHEMATICS

In 2013 the departments Numerical Modelling and Analysis and Visualization and Data Analysis have continued to advance their competences in mathematical modelling, simulation, optimization, and (visual) data analysis. At present the two departments are active in more than 60 third-party funded projects with topics ranging from basic research in mathematics via application-oriented research to demonstrators jointly developed with cooperation partners. Application areas include medicine, molecular science, neuro- and systems biology, nano-optics, and materials science.

During the last year the two departments have started to identify the focus areas in which application-oriented research will be concentrated in coming years. The departments' joint strengths will be combined with additional in-house expertise in (big) data analysis and high performance computing to maximize the impact of mathematics-centered research in these focus areas. Two of these focus areas, "Solutions for individualized medicine" and "Design of functional molecules", are featured in this annual report.

COMPUTER SCIENCE

The two departments Distributed Algorithms and Supercomputing and IT Service spent much effort in the preparation of the HLRN-III installation. The Cray XC30 was up and running already two weeks after the first truck unloaded its precious freight on the ZIB premises. To complete the setup, all data needed to be copied to the new file system, the job queues had to be configured and new software for project and job accounting had to be developed. Despite its complexity, the HLRN-III was opened to the users within a month after its physical installation.

Another highlight in 2013 was the collaboration with Intel, which selected ZIB as a Research Center for Many-core High-Performance Computing – one of only five centers worldwide. First research results have already been published and found their way into the consultancy of HLRN users.

As a fundamental building block for big data analysis we intensified our research on parallel file systems and distributed key-value stores. Supported by venture capital, three of our Post-Docs founded the XtreemFS Storage Solutions GmbH, which soon doubled in size and changed its name to Quobyte, now with headquarters in Berlin and in Boston.

1

1 Zuse Institute Berlin, main campus Dahlem

ADMINISTRATIVE BODIES

The bodies of ZIB are the President and the Board of Directors (Verwaltungsrat).

President of ZIB is:
PROF. DR. DR. H.C. MULT. MARTIN GRÖTSCHEL

Vice President is:
PROF. DR. CHRISTOF SCHÜTTE
(since January 10, 2013).

The Board of Directors was composed in 2013 as follows:

PROF. DR. PETER FRENSCH
Vice President, Humboldt-Universität zu Berlin, (Chairman)

PROF. DR. UWE PAUL THAMSEN
Vice President, Technische Universität Berlin, (Vice Chairman)

PROF. DR. BRIGITTA SCHÜTT
Vice President, Freie Universität Berlin

BERND LIETZAU
Senatsverwaltung für Wirtschaft, Technologie und Forschung

N.N.
Senatsverwaltung für Bildung, Jugend und Wissenschaft

PROF. DR. MANFRED HENNECKE
Bundesanstalt für Materialforschung und -prüfung (BAM)

PROF. DR. ANKE KAYSSER-PYZALLA
Helmholtz-Zentrum Berlin für Materialien und Energie (HZB)

PROF. DR. NIKOLAUS RAJEWSKY
Max-Delbrück-Centrum für Molekulare Medizin (MDC)

The Board of Directors met on May 15, 2013 and November 20, 2013.

SCIENTIFIC ADVISORY BOARD

The Scientific Advisory Board advises ZIB on scientific and technical issues, supports ZIB's work and facilitates ZIB's cooperation and partnership with universities, research institutions and industry.

The Board of Directors appointed the following members of the Scientific Advisory Board:

PROF. DR. ALFRED K. LOUIS
Universität des Saarlandes, Saarbrücken, (Chair until June 10, 2013).

PROF. DR. JÖRG-RÜDIGER SACK
Carleton University, Ottawa, Canada, (Chair since June 11, 2013)

PROF. EM. DR. RAINER E. BURKARD
Technische Universität Graz, Austria (since May 15, 2013)

PROF. DR. MICHAEL DELLNITZ
Universität Paderborn (since May 15, 2013)

PROF. DR. THOMAS LIEBLING
Ecole Polytechnique Fédérale de Lausanne, Switzerland (until May 15, 2013)

DIPL.-MATH. LUDGER D. SAX
Open Grid Europe GmbH, Essen

DR. ANNA SCHREIECK
BASF SE, Ludwigshafen (since May 15, 2013)

DR. HORST D. SIMON
NERSC, Lawrence Berkeley National Laboratory, USA

DR. REINHARD UPPENKAMP
Berlin Chemie AG, Berlin

DR. KERSTIN WAAS
Deutsche Bahn AG, Frankfurt am Main (since May 15, 2013)

The Scientific Advisory Board met on June 10 and 11, 2013 at ZIB.

OR

SCIENTIFIC ADVISORY BOARD

Chairman Prof. Dr. Jörg-Rüdiger Sack, Ottawa
Prof. Dr. Rainer E. Burkard, Graz
Prof. Dr. Michael Dellnitz, Paderborn
Prof. Dr. Alfred K. Louis, Saarbrücken
Ludger D. Sax, Essen
Dr. Anna Schreieck, Ludwigshafen
Dr. Horst D. Simon, Berkeley
Dr. Reinhard Uppenkamp, Berlin
Dr. Kerstin Waas, Frankfurt am Main

THE STATUTES

The Statutes, adopted by the Board of Directors at its meeting on 30 June 2005, define the functions and procedures of ZIB's bodies, determine ZIB's research and development mission and its service tasks and decide upon the composition of the Scientific Advisory Board and its role.

GANIZATION

BOARD OF DIRECTORS

Chairman: Prof. Dr. Peter Frensch, Humboldt-Universität zu Berlin (HUB)

PRESIDENT

Prof. Dr. Dr. h.c. mult. Martin Grötschel

VICE PRESIDENT

Prof. Dr. Christof Schütte

ADMINISTRATION AND LIBRARY

Annerose Steinke

NUMERICAL MATHEMATICS

Prof. Dr. Christof Schütte

DISCRETE MATHEMATICS

Prof. Dr. Martin Grötschel

COMPUTER SCIENCE

Prof. Dr. Alexander Reinefeld

NUMERICAL MATHEMATICS – Schütte

- **NUMERICAL ANALYSIS AND MODELLING** – Weiser
  - COMPUTATIONAL MEDICINE – Weiser, Zachow
  - COMPUTATIONAL MOLECULAR DESIGN – Weber
  - COMPUTATIONAL NANO OPTICS – Schmidt
  - COMPUTATIONAL SYSTEMS BIOLOGY – Röblitz
- **VISUALIZATION AND DATA ANALYSIS** – Hege
  - VISUALIZATION ALGORITHMS – Hege
  - VISUALIZATION SYSTEMS – Prohaska
  - MEDICAL PLANNING – Zachow
  - MEDICAL BIOINFORMATICS – Conrad

DISCRETE MATHEMATICS – Grötschel

- **OPTIMIZATION** – Grötschel/Borndörfer
  - TRAFFIC AND LOGISTICS – Borndörfer
  - LINEAR AND NONLINEAR INTEGER PROGRAMMING – Koch
  - TELECOMMUNICATION – Werner
  - ENERGY – Hiller
- **SCIENTIFIC INFORMATION** – Koch
  - WEB TECHNOLOGY AND MULTIMEDIA – Dalitz
  - INFORMATION SYSTEMS – Koch
  - MUSEUM SOFTWARE – Saro
  - KOBV – Koch/Rusch
  - SERVICE CENTER DIGITALIZATION – Koch

ZIB is structured into four divisions, three scientific divisions and ZIB's administration. Each of the scientific divisions is composed of two departments which are further subdivided into research groups. For further information on the broad spectrum of thematic and application foci of the research groups please visit www.zib.de.

ZIB STRUCTURE

COMPUTER SCIENCE
Reinefeld

- **DISTRIBUTED ALGORITHMS AND SUPERCOMPUTING**
  Reinefeld/Schintke
  - **DISTRIBUTED DATA MANAGEMENT**
    Schintke
  - **SCALABLE ALGORITHMS**
    Schütt
  - **SUPERCOMPUTER ALGORITHMS AND CONSULTING**
    T. Steinke
  - **SUPERCOMPUTER SYSTEMS**
    Schimmel
  - **SOFTWARE SERVICE**
    Löbel
- **IT-SERVICE**
  Pyszkalski
  - **MASS STORAGE**
    Pyszkalski, Pham
  - **SERVER AND NETWORKS**
    Krickel, Düring
  - **BERLIN METROPOLITAN AREA NETWORK (BRAIN)**
    Pyszkalski

ADMINISTRATION AND LIBRARY
A. Steinke

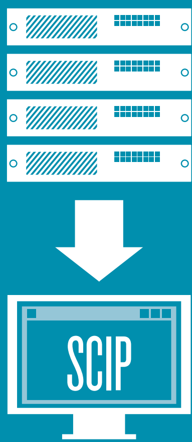
LEGEND

- SCIENTIFIC DIVISIONS AND DEPARTMENTS
- RESEARCH GROUPS
- RESEARCH SERVICE GROUPS

ZIB IN NUMBERS

PROMOTION OF YOUNG SCIENTISTS:

HABILITATIONS
DISSERTATIONS
DIPLOMA
MASTER



8.000

DOWNLOADS
OF MIP SOLVER SCIP

SEMINARS
GIVEN BY ZIB SCIENTIST AT UNIVERSITIES

LECTURES
GIVEN BY ZIB SCIENTIST AT UNIVERSITIES

1.200 VISITORS

LONG NIGHT OF THE SCIENCES

5 PROFESSORSHIPS OFFERED TO ZIB RESEARCHERS

109

PEER-REVIEWED PUBLICATIONS IN INTERNATIONAL SCIENTIFIC JOURNALS

233 SCIENTIFIC TALKS

41 DISTINGUISHED

192 INVITED

AMIRA AVIZO

8.000 LICENSES OF AMIRA AND AVIZO INSTALLED WORLDWIDE IN 2013

47 OUTREACH EVENTS FOR SCHOOL CLASSES AND GENERAL PUBLIC

4.026.607 €
PUBLIC PROJECTS REVENUES

1.864.118 €
REVENUES FROM RESEARCH AND DEVELOPMENT (INDUSTRY)

ZIB ACQUISITION OF THIRD-PARTY FUNDS

THIRD-PARTY FUNDING INCREASED BY MORE THAN 20%!

15 AWARDS

5.000 INTERNATIONAL GUESTS AT ZIB IN 2013

100 CONFERENCES AND WORKSHOPS AT ZIB

OCTOBER 25
ZIB AND INTEL ESTABLISH THE RESEARCH CENTER FOR MANY-CORE-HIGH-PERFORMANCE-COMPUTING AT ZIB

PERMANENT DATA STORAGE
18,000 SLOTS FOR TAPES IN 3 SL 8,500 LIBRARIES
9 PETABYTE LIBRARY STORAGE CAPACITY

SEPTEMBER 1
START OF OPERATION OF HLRN-III

NOVEMBER 18
HLRN-III NO. 120/121 IN THE TOP 500 LIST OF SUPER COMPUTERS IN THE WORLD

FOR FURTHER INFORMATION PLEASE CONTACT:
Annerose Steinke
a.steinke@zib.de
+49 30 84185-100

MODAL

The Forschungscampus ("research campus") Modal is a platform for public-private partnership for innovation built by ZIB and Freie Universität Berlin together with 12 participating companies. Modal is an acronym derived from "Mathematical Optimization and Data Analysis Laboratories". The Forschungscampus is located at ZIB where partner researchers from the public and private sectors are directly working together in joint laboratories. Modal aims at sustainable innovations in process optimization through mathematical research on modelling, simulation, and mixed discrete-continuous-stochastic optimization problems that originate from real-world application problems beyond reach of today's computing and optimization technologies. Present main applications are railway and gas transport optimization, and information-based medicine.

Modal resulted from a national competition in which the German Ministry for Research and Education (BMBF) awarded the title "Forschungscampus" to ten public-private partnerships out of almost a hundred candidates. The Forschungscampus funding initiative supports long-term public-private partnerships which bring their partners together in one place. Each selected Forschungscampus will receive funding of 1 to 2 million Euros per year by BMBF for a maximum of 15 years. The funding coming from the private consortium partners must exceed the BMBF funding, though.

FUTURE PRESENT PAST

ZUSE INSTITUTE BERLIN

The Konrad-Zuse-Zentrum für Informationstechnik Berlin (official name of ZIB) was established on July 17, 1984. In the terminology of the German public law system, ZIB is a "rechtsfähige Anstalt des öffentlichen Rechts mit kaufmännischer Buchführung". ZIB receives its basic funding from the State of Berlin, covers most of its research activities through project funding from various sources, and operates in close interdisciplinary cooperation with the three large Berlin universities and many research institutions in the field of information technology, particularly in applied mathematics and practical computer science.

Image courtesy: Horst Zuse

1 Konrad Zuse – Patron of ZIB

Konrad Zuse, known as the inventor of the world's first freely programmable computer, is the patron of the Zuse Institute Berlin (ZIB) . He was born in Berlin in 1910. Zuse had a great talent for technology and art and as a child he made first intelligent inventions. Inspired by the movie "Metropolis" he designed a road system for his visionary future city; this road map forms the current logo of ZIB. After having passed high school, he studied engineering and architecture at the antecessor of today's TU Berlin.

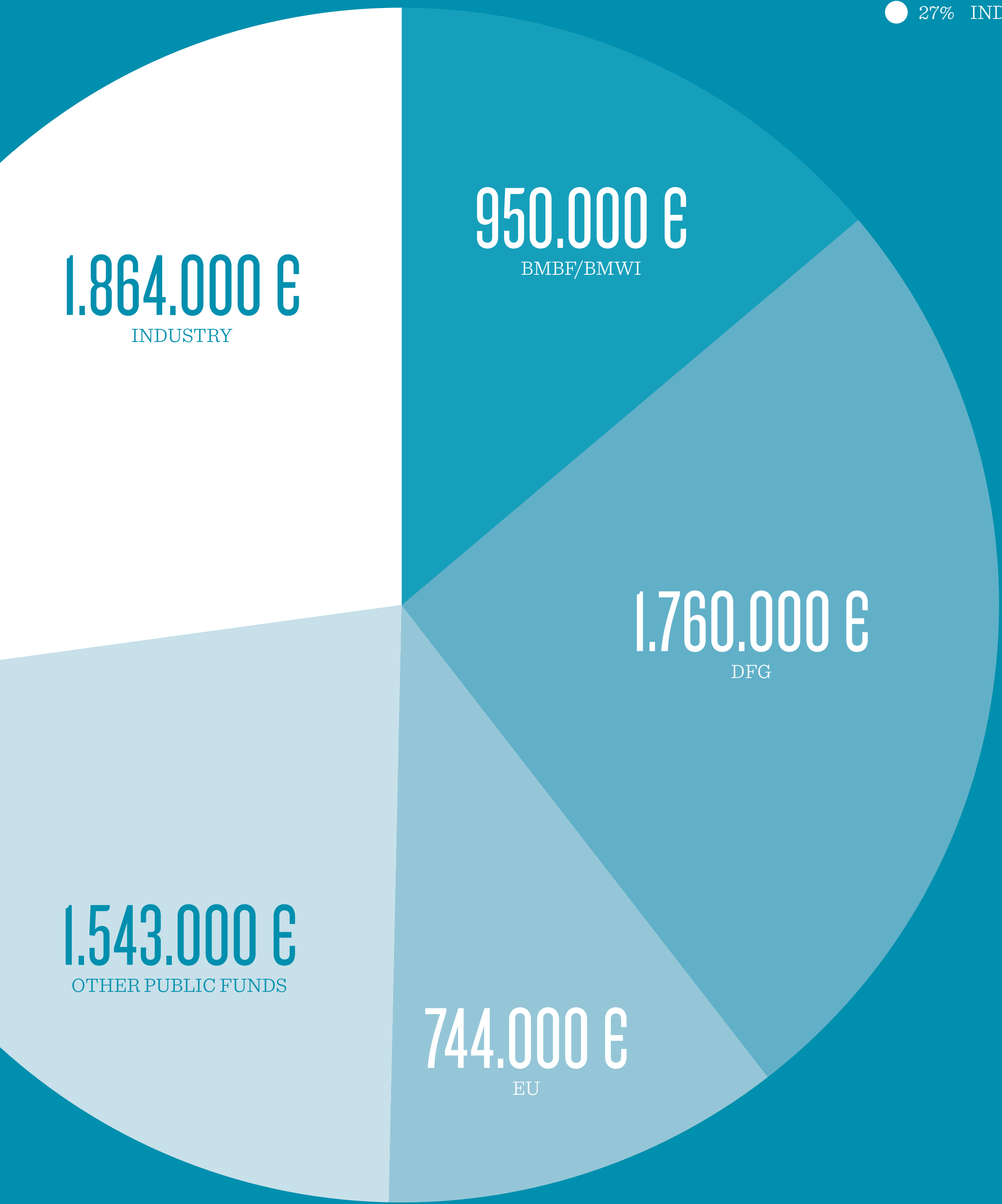
In 1941, Zuse developed the world's first free programmable computer (Z 3) which is today regarded as the first functioning computer in the world. After the end of World War II, he built up his own Company, the Zuse KG . After his retirement he devoted himself increasingly to its second passion, which was painting. Using the pseudonym "Kuno See", he created abstract paintings, two of which he donated to ZIB. He still experienced the preliminary planning for the new Zuse Institute Berlin.

Konrad Zuse, who himself never earned a PhD, received numerous honorary doctorates and awards. He died at the age of 85 in Hünfeld near Fulda.

ECONOMIC SITUATION IN 2013

ZIB THIRD PARTY FUNDS BY SOURCE

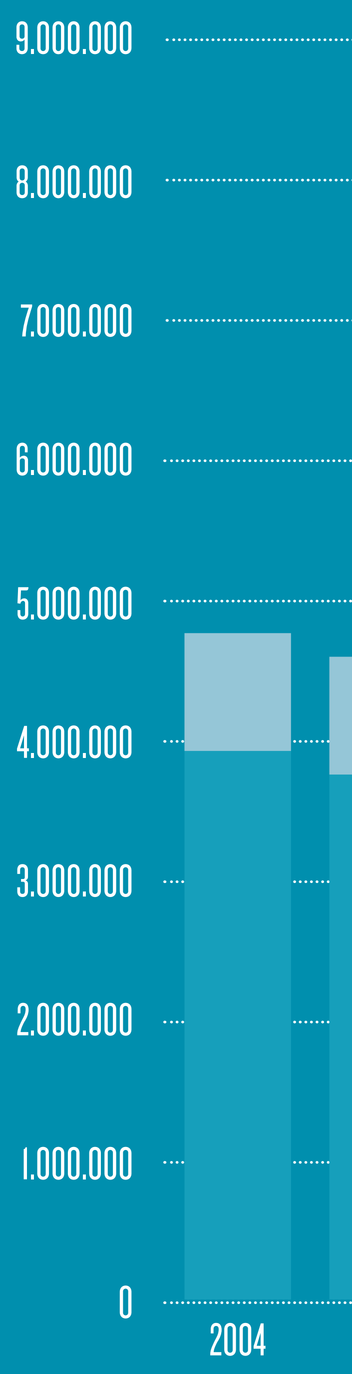
- 14% BMBF/BMWI
- 26% DFG
- 11% EU
- 22% OTHER PUBLIC FUNDS
- 27% INDUSTRY



The Zuse Institute Berlin (ZIB) finances its scientific work via three main sources: the basic financial stock of the Federal State of Berlin, and third-party funds of public sponsors as well as from industrial cooperation contracts.

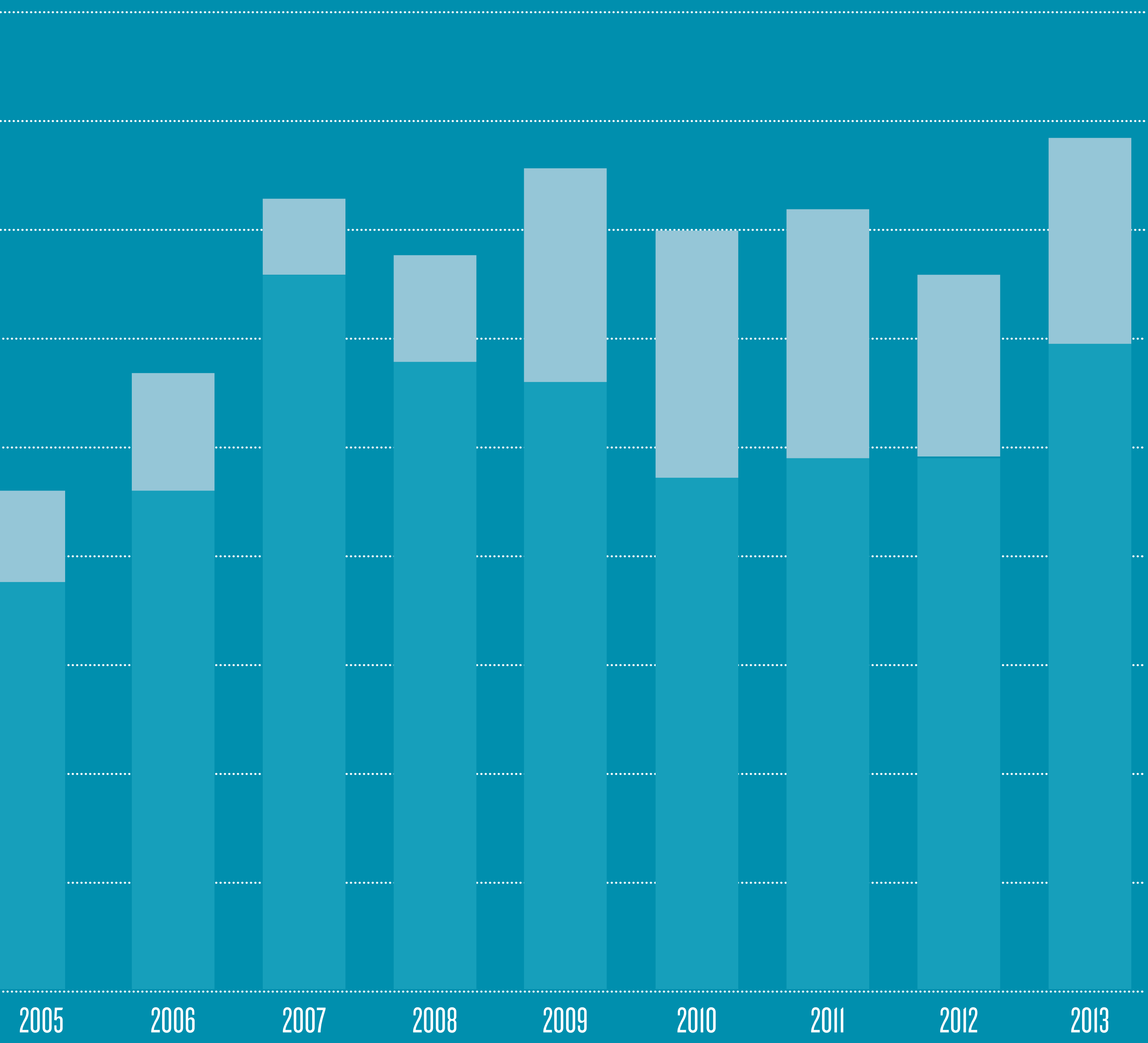
In 2013, ZIB raised third-party funding for a large number of projects. While the basic financial stock has stayed approximately the same, project-related public third-party funds have raised by more than 25% compared to last year. Despite the economically difficult times, ZIB was once again able to acquire almost 30% of its total third-party funding through research cooperation contracts with large and medium-sized companies. Based on these successful achievements ZIB strengthened its position as an attractive place for innovative mathematicians and computer scientists such that the number of employees at ZIB stayed almost constant despite the fact that many highly-qualified scientists were offered careers in academia or industrial institutions research elsewhere.

Big investments were taken: The new HLRN-III high-performance computer at ZIB started operation in September 2013, and the institute's long-term data storage was further expanded. As a consequence, a large number for new research cooperations with Berlin-based, national and international institutions in this area could be initiated. With these developments ZIB strengthened its position as a competence center for computing and data science.



ZIB THIRD PARTY FUNDS IN EURO

INDUSTRY
PUBLIC FUNDS



NUMBER OF EMPLOYEES

218 persons were employed at ZIB in the year 2013, thereof 151 positions were financed by third-party-funds

	01.01.13			01.01.14		
	TOTAL	*permanent*	*non-permanent*	TOTAL	*permanent*	*non-permanent*
MANAGEMENT	4	4	0	4	4	0
SCIENTISTS	102	19	83	99	18	81
SERVICE PERSONNEL	45	37	8	48	35	13
KOBV-HEADQUARTERS	16	7	9	13	7	6
STUDENTS	51	0	51	49	0	49
TOTAL	218	67	151	213	64	149

COMPUTING IN TECHNOLOGY GMBH (CIT)
1992 | www.cit-wulkow.de
Mathematical modelling and development of numerical software for technical chemistry

RISK-CONSULTING PROF. DR. WEYER GMBH
1994 | www.risk-consulting.de
Database marketing for insurance companies

INTRANETZ GMBH
1996 | www.intranetz.de
Software development for logistics, database publishing, eGovernment

AKTUARDATA GMBH
1998 | www.aktuardata.de
Development and distribution of risk evaluation systems in health insurance

VISAGE IMAGING GMBH
(originating from spin-off Visual Concepts GmbH)
1999 | www.visageimaging.com
Visualization and data analysis (Amira etc), especially medical visualization

ATESIO GMBH
2000 | www.atesio.de
Development of software and consulting for planning, configuration, and optimization of telecommunication networks

BIT-SIDE GMBH
2000 | www.bit-side.com
Telecommunication applications, and visualization

DRES. LÖBEL, BORNDÖRFER & WEIDER GBR
2000 | www.lbw-berlin.de
Optimization and conculting in public transport

LENNÉ3D GMBH
2005 | www.lenne3d.com
3d landscape visualization, software development and service

JCMWAVE GMBH
2006 | www.jcmwave.com
Simulation software for optical components

ONSCALE SOLUTIONS GMBH
2006 | www.onscale.de
Software development, consulting and service for parallel and distributed storage and compute systems

MOLCONCEPT GMBH
2008 | www.molconcept.com
Mathematical concepts for molecular design, software development for computational drug design

LAUBWERK GMBH
2009 | www.laubwerk.com
Construction of digital plant models

1000SHAPES GMBH
2010 | www.1000shapes.com
Statistical shape analysis

TASK – BERTHOLD GLEIXNER HEINZ KOCH GBR
2010
Distribution, service and consulting for ZIB's optimization suite

QUOBYTE INC.
2013 | www.quobyte.com
Quobyte develops carriergrade storage software that runs on off-the-shelf hardware

SPIN OFFS

SOLUTIONS FOR INDIVIDUALIZED MEDICINE

Enabling technologies for complex therapies

FOR FURTHER INFORMATION PLEASE CONTACT:
Dr. Martin Weiser | weiser@zib.de | +49 30 84185-170

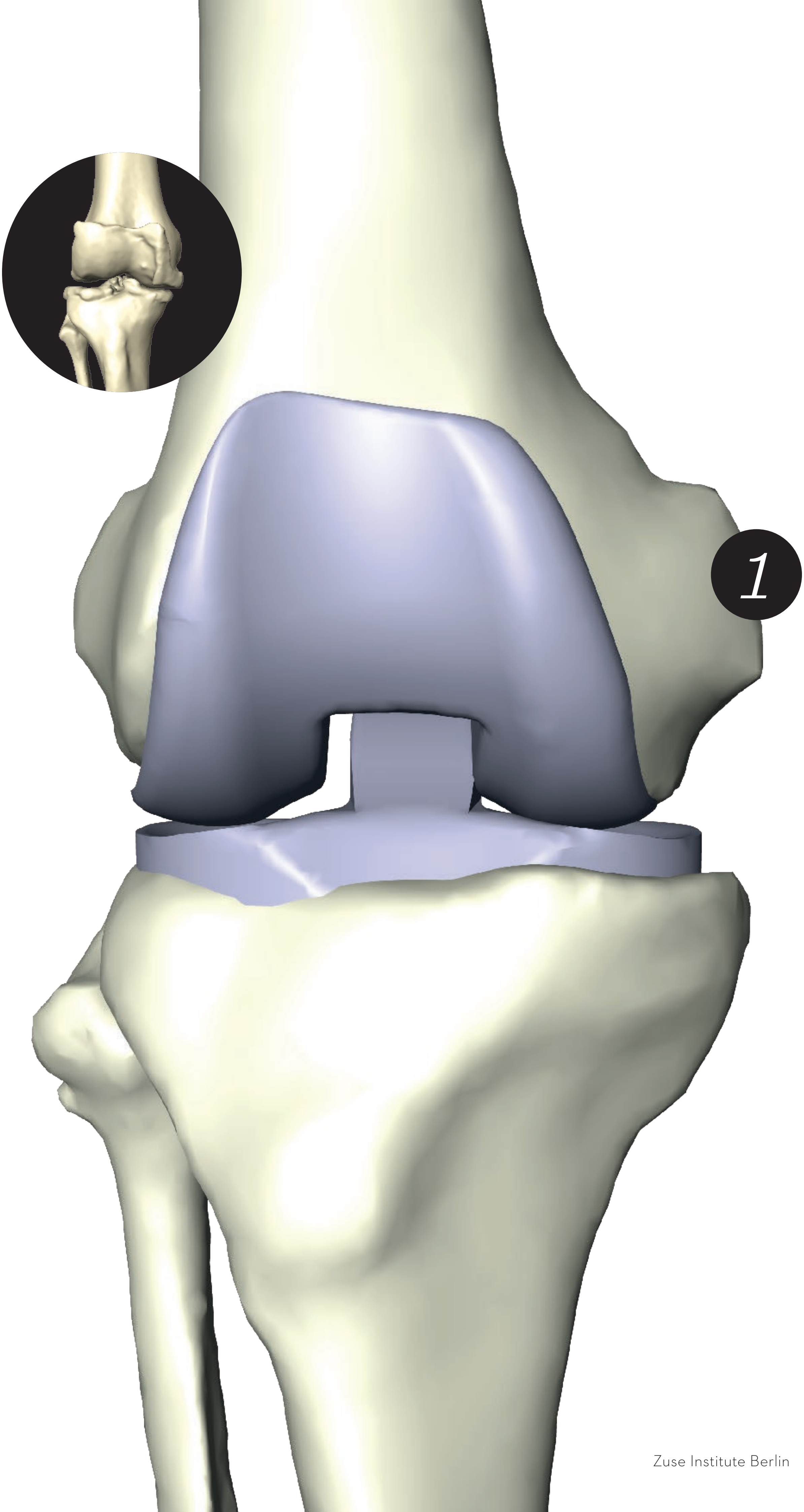
In today's medicine therapeutic interventions are getting increasingly complex. Functional and physiological data need to be monitored, structures of risk are to be considered, and a lot of technology is involved. In order to master the complexity and to deliver successful treatments, such therapies need to be carefully planned, taking individual anatomy and pyhsiological processes into account. We develop algorithms and demonstrators supporting computer-assisted, model-guided therapy planning.

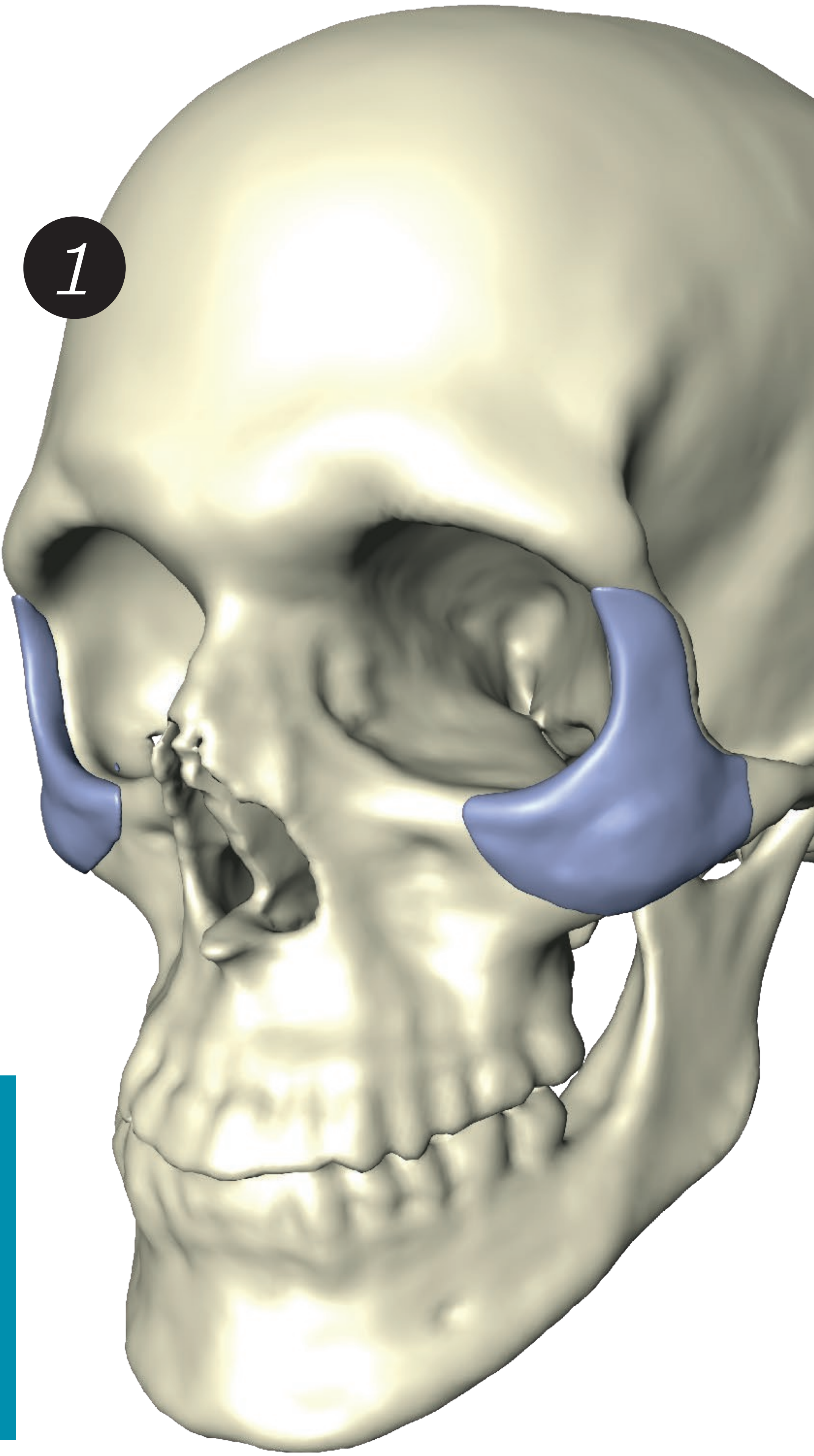
❶ Individual anatomy reconstruction of the knee bones allows 3D planning of implant positioning. Mechanical simulations taking cartilage, patella, tendons, and ligaments into account improve the understanding of degenerative diseases and the functionally appropriate positioning of implants.

❷ Knee implant wear testing is a time consuming process. Efficient simulation of wear can speed up the design process significantly.

ORTHOPEDIC SURGERY

Orthopedic surgery is concerned with the treatment of pathologies in shape or function of the musculoskeletal system that may be caused by trauma, degenerative or congenital diseases. Therapy planning in orthopedics comprises the identification of pathological situations as well as the design of therapeutic concepts under consideration of functional rehabilitation. Based on patient specific anatomical models, for instance, pose and shape of anatomical structures can be analysed and both can be compared to nonpathological findings. Using kinematic models also functional aspects can be analyzed, e.g., the range of motion of joints [1]. Stresses within anatomical structures can be simulated under particular loading conditions, and parts of the anatomical structures can be virtually removed and replaced by implants in order to analyze the performance of the planned configuration [2]. With the help of computer-assisted and model-guided therapy planning, different strategies can be preoperatively assessed and the most promising approach can be chosen for implementation. We are developing techniques for 3D shape analysis and establishing 3D shape databases representing various anatomical structures that can be stratified according to gender, age, body weight, and so forth. Within the stratified groups the range of variation of properties like shape or mineral density can be analyzed. An individual patient's anatomy can then be matched against the shape database in order to assess its pathological state and to determine how to surgically correct skeletal structures, mechanical axis or joint contact regions. In addition, optimally fitting implants can be selected from a set of different types or manufacturers, or even new implant types can be designed that either fit an individual patient or a particular population, whose anatomical average shape can be derived from the database (individual vs. mass customization).


1

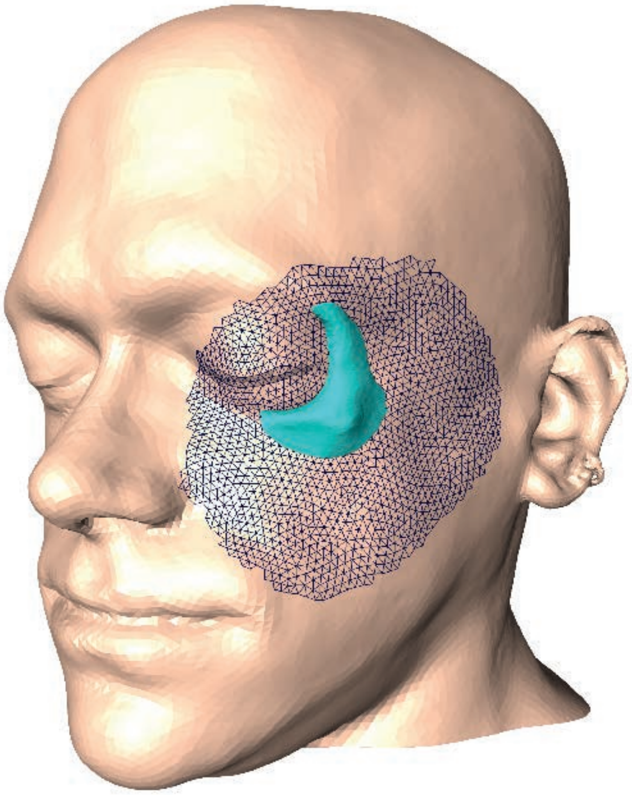


IN A NUTSHELL

- Modern interventions are increasingly complex.
- Many medical areas are affected: osteotomic surgery, facial surgery, implant design, cardiac diseases, and more.
- Patient-specific anatomical models and simulations thereon are necessary for improved reliability of treatment decisions.

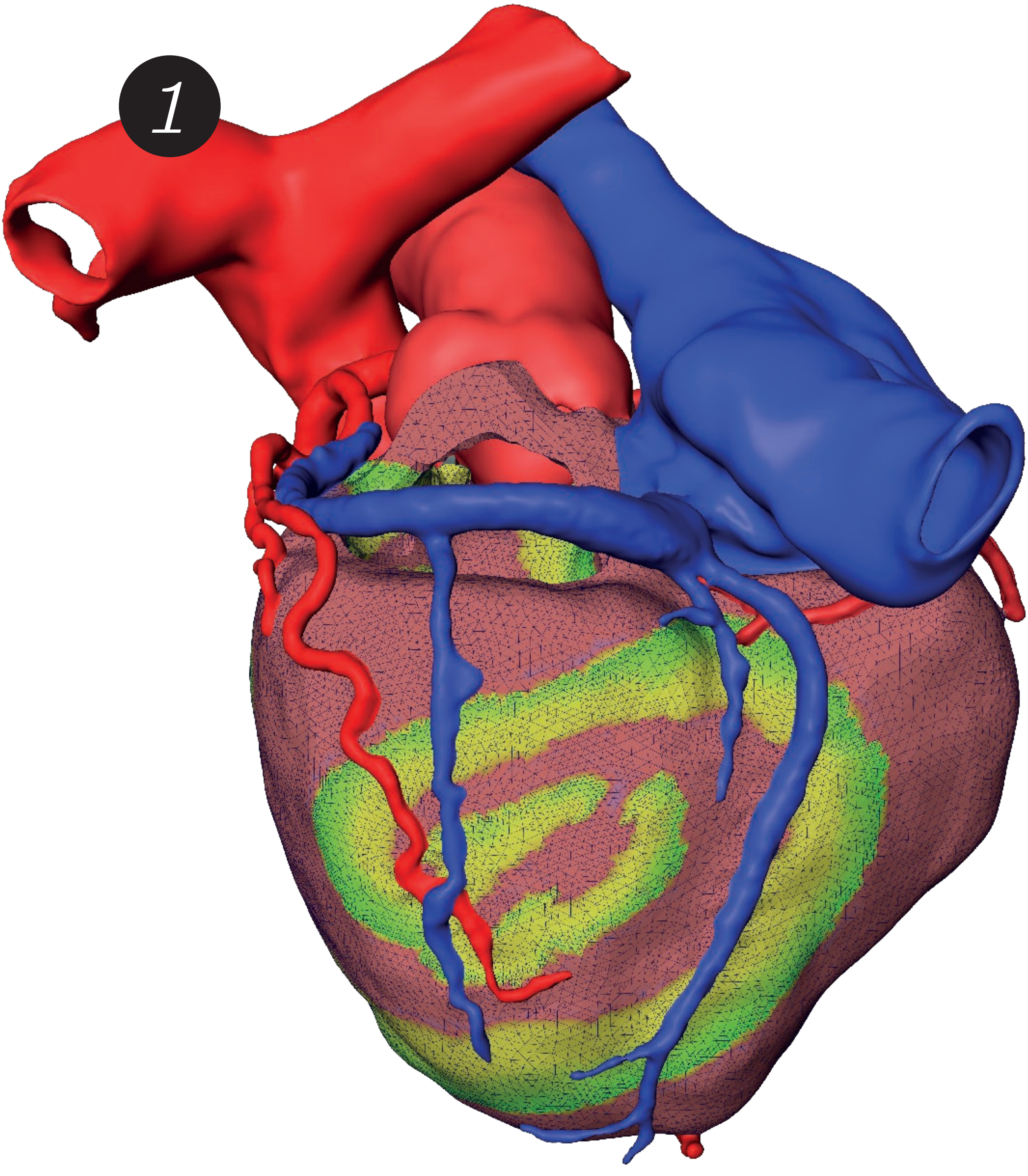
FACIAL SURGERY

Preoperative planning of plastic and reconstructive surgery of the head, not only involves functional but also aesthetic aspects. This is especially the case for congenital malformations, where a face or the shape of the head will be heavily altered without having an individual reference. Either rearrangements of bony structures need to be planned, or implants need to be carefully designed to model a new face or head. In both cases the facial appearance needs to be taken into account in order to achieve the best possible aesthetic result under consideration of functional rehabilitation [3]. We are working on software tools to perform interactive 3D planning of bone cuts and rearrangements based on digital anatomy as well as to design implant shapes. In addition, we work on models and algorithms to realistically simulate mechanical behaviour of biological tissues. We also develop methods to establish a database of 3D face models that can be used to derive an objective for facial reconstruction. Furthermore, we develop methods to determine the shape of implants for an individual patient with regard to a desired facial appearance. That way surgeons can directly shape a face, automatically receiving a blueprint for implants that once inserted lead to the planned aesthetic outcome.



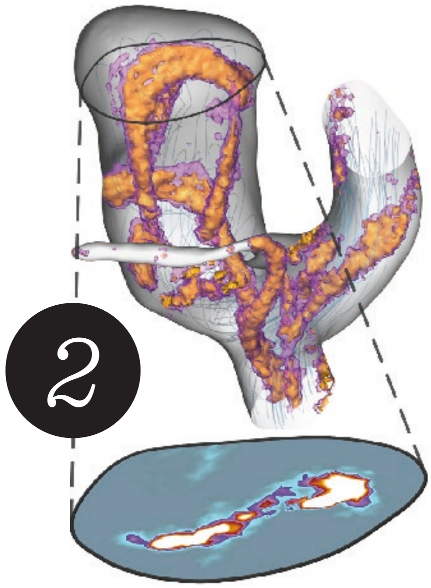
1 Augmentation implants replace traumatized bone or fill voids left by malformed bones. An important aspect is to restore a socially accepted facial shape, for which the appropriate implant shape has to be determined.

CARDIOVASCULAR DISEASES



1 Changes in transmembrane voltage induce calcium release and myocard contraction. Simulation of the excitation fosters understanding of fibrillation onset and stability of spiral wave patterns, improving individual therapy planning.

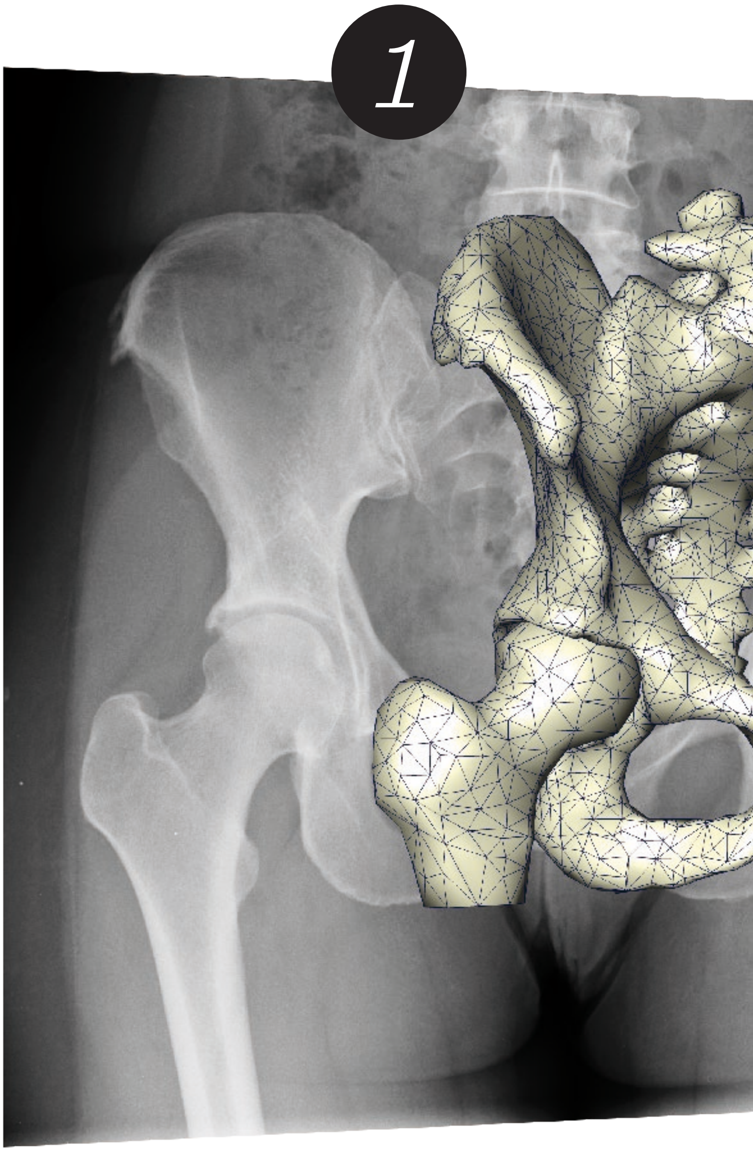
2 Probability density of vortex cores in an aneurysm.



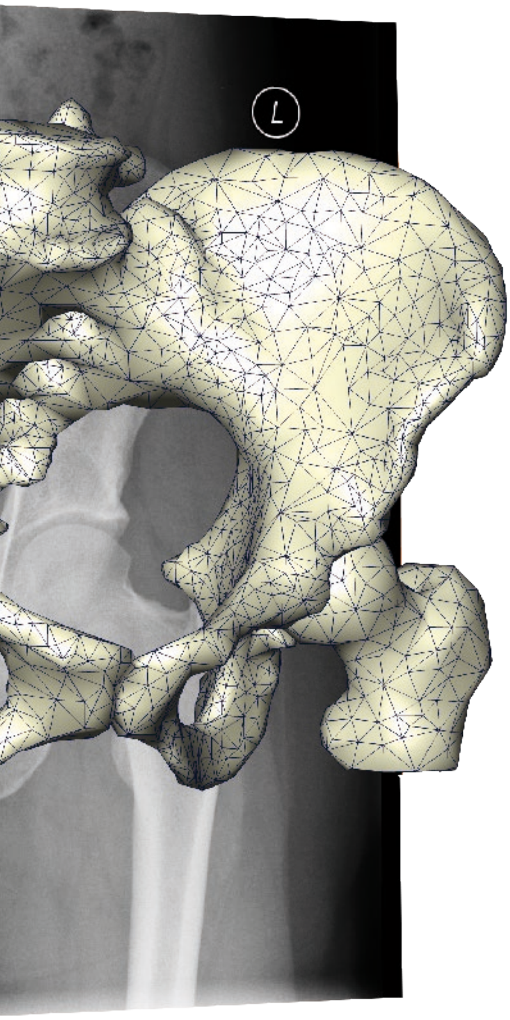
Treatment decisions for cardiovasclar diseases, e.g., in valve replacement, resynchronization therapy, or ablation therapy, should be based on the benefit of the patient. Large inter-individual variations of the outcome make the prediction of the benefit difficult, e.g., some patients develop a ventricular hypertrophy after valve replacement, and some do not. Some aneurysms remain the same size over years, and some grow quickly with the danger of rupture. A reliable patient-specific simulation of the relevant physiological processes is often limited by the sheer complexity of the cardiovascular system, involving different scales from mitochondrial metabolism to systemic circulation and effects from ion transport in myocytes to fluid dynamics. Therefore, individual therapy planning is still far from clinical routine. We are working on efficient simulation methods for the electrical excitation of the heart and the induced mechanical contraction. This includes not only adaptive, error controlled solvers for partial differential equations, but also data compression algorithms for the efficient storage of simulation results for optimization purposes. E.g., implantable cardioverter-defibrillators can detect the onset of life-threatening ventricular fibrillation and terminate this condition by applying a sufficiently strong electrical impulse to the myocard. “Pushing this big reset button of the heart is painful and shocking for the patient”, explains PhD student Sebastian Götschel. “One goal in designing the pulse shape is to minimize its amplitude while maintaining the effectivity”.

Estimating the rupture risk of aneurysms is important to decide whether to perform surgery or not. This risk depends on the flow patterns of blood in the aneurysm and the resulting wall shear stresses. The stochastic nature of these quantities makes a reliable interpretation difficult. Uncertainty visualization methods can help to understand the risk factors better.

MEDICAL IMAGE AND GEOMETRY PROCESSING



Computer-assisted therapy planning in medicine is heavily based on medical imaging technology. Therefore, image data are not only to be visualized but often also to be converted into individual 3D patient models, sub-divided into various anatomical structures of interest [4]. In order to faithfully extract anatomical 3D models from medical image data, we are developing algorithms for image enhancement, segmentation, and geometry reconstruction. Converting segmented image data into 3D surfaces or volumetric grids makes them accessible to computational methods for mathematical analysis and simulation. In addition, interactive visualization methods are developed for presenting image and simulation data to the domain experts in an intelligible manner and software prototypes to our clinical research collaborators.



1 *Comparing projections of shape models with actual X-ray images allows to estimate individual 3D geometry from few X-ray images, reducing the radiation dose compared to CT scans.*

IMAGE PROCESSING

We are developing efficient and robust algorithms for medical image segmentation deploying so called shape priors. Such shape priors represent averaged 3D shapes of many similar anatomical structures, implicitly describing the full range of variation in shape that typically exist in any population. For fully automated segmentation of arbitrary medical image data, like computed tomography, magnetic resonance imaging, or ultrasound, shape priors are enriched with modality specific image information. Within an optimization scheme these shape and appearance models are plausibly deformed until shape and image features possibly coincide [5]. Our algorithms have been sucessfully applied for fully automatic segmentation of a broad range of different anatomical structures from different imaging modalities. In addition, highly parallelized image filtering, registration, and segmentation algorithms are being developed – deploying modern hardware architectures to handle an ever increasing amount of medical image data in reasonable time [6] or to solve complex 3D reconstruction problems. Moritz Ehlke, PhD student in the Berlin School of Regenerative Therapies explains his work [7]: "Reconstructing 3D anatomies from few projective X-ray images is only possible with very good knowledge of the arising shapes, as it is encoded in shape priors. Still, the necessary simulation of X-ray projections requires the computing power of modern GPUs."

MEDICAL GEOMETRY PROCESSING

Though medical image data can be directly visualized for diagnostic purposes, interactive, model-guided planning and simulation often requires three-dimensional geometric models representing individual patients' anatomy. Therefore, we are developing algorithms for the conversion of segmented image data into geometrical representations like surface meshes or volumetric grids, making them accessible to computational methods for interactive therapeutic modifications and numerical simulations employing, for instance, advanced finite-element techniques.

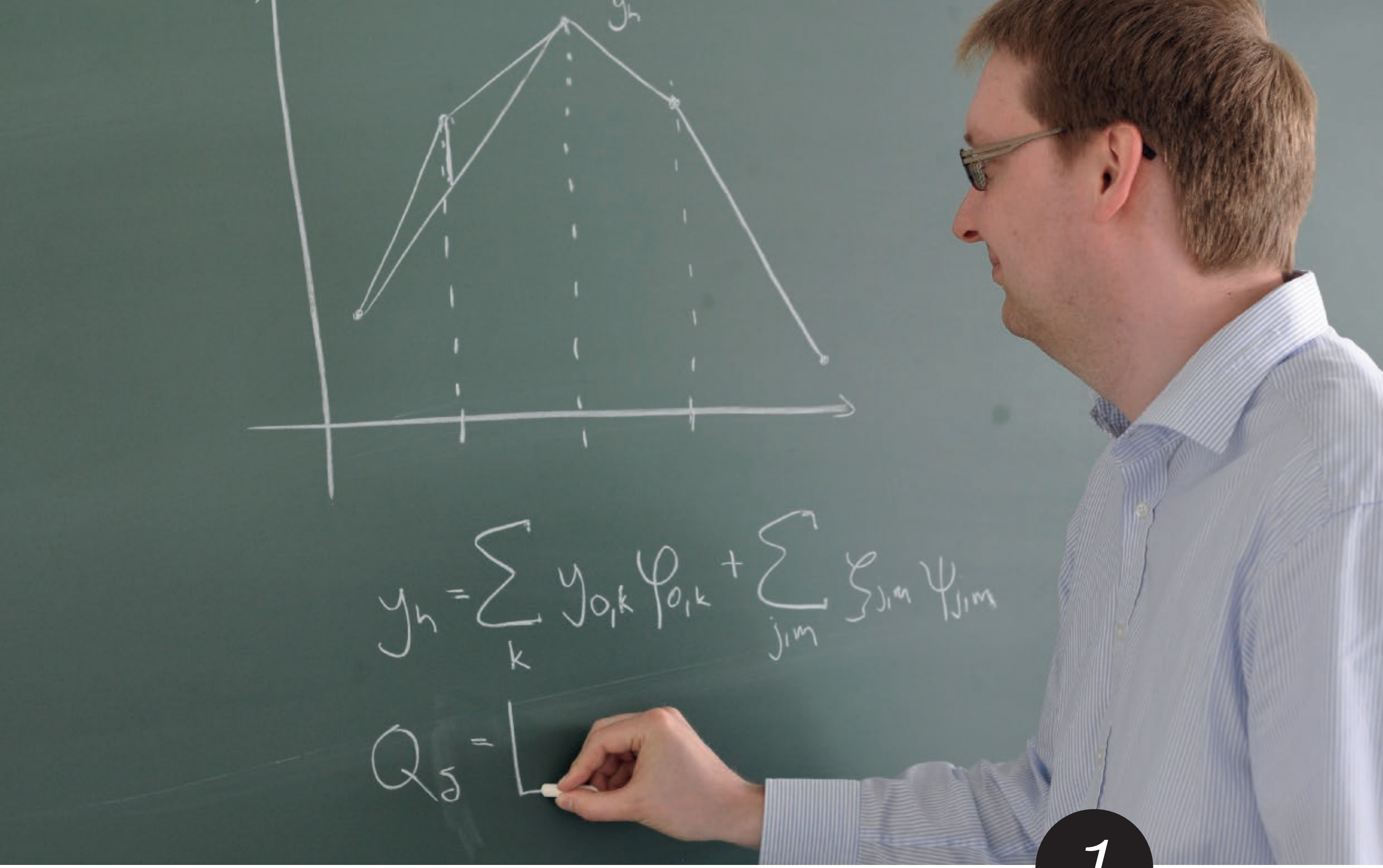
We are also developing methods to establish and to improve statistical shape models. One important step in building shape models is the registration of complex shapes onto an average. Among others, we are investigating large deformation diffeomorphic mapping methods that use a deformation of a spatial embedding for the registration of shapes. Decoupling shape representation and space deformation allows a better adaptation of the discretizations to the solution. Statistical shape models are the basis for the generation of new anatomical atlases that not only show averaged shapes of anatomical structures but also the variation in shape of such structures. The latter is important for morphological analysis in general, discrimination of pathologies, classification, and even the human-centered design of implants or prostheses. To understand complex relationships between shape and other factors or categorial information, we are developing methods to visually explore and to analyze shape models.

NUMERICAL METHODS FOR BIOMEDICAL PDES

The physical and physiological processes that are relevant in complex interventions, such as electrical excitation, heat transport, or mechanical behaviour of soft tissues, bones and implants, are usually described by partial differential equations (PDEs). In order to be able to predict the impact of treatments and to optimize their parameters, we develop finite element methods and optimization algorithms. One focus of our research are adaptive, error-controlled methods combining reliability and efficiency.

TIME INTEGRATION

Dynamic problems such as cardiac excitation or implant wear simulation require a time stepping scheme that integrates well with adaptive spatial finite element discretization. Recently we turned to spectral deferred correction (SDC) methods, which are simply structured iterative methods for solving implicit collocation systems. Interleaving SDC iterations and mesh refinement, as well as amortizing the overhead of adaptivity over longer time steps possible with higher-order methods, bear the promise of improving overall efficiency. The deteriorating convergence rate of SDC methods on nonequidistant time grids can be accelerated significantly by a special construction of the SDC sweeps [8].



TRAJECTORY COMPRESSION

In time-dependent 3D optimization problems, such as optimal control of implantable cardioverter-defibrillators or shape optimization of knee joint prosthesis for minimal wear, adjoint gradient computation for fast minimization requires the repeated storage of a full 4D simulation for adjoint gradient computation. This can impose a significant difficulty in terms of both, storage capacity and storage bandwidth. In a Matheon project we have developed lossy compression of finite element solutions on adaptively refined meshes, reducing the memory demand as well as the memory bandwidth by a factor of more than 20 without impeding the convergence of the optimization algorithm [9, 10]. “It’s similar to MPEG video compression, just that it works on unstructured grids instead of on pixels, and respects the error transport through the adjoint equation rather than visual artifact perception”, says PhD student S. Götschel.

OPTIMAL CONTROL

Complex therapies often exhibit a moderate to large number of parameters to fix, e.g., the shape of augmentation implants or the positioning of joint implants. Choosing parameters for an optimal outcome can be rather challenging. We work on optimization algorithms suggesting optimized therapy parameters. Sometimes this involves basic research: “Very little is known about the structure of nonlinear elasticity solutions”, PhD student Lars Lubkoll says, “so we are glad we could prove the existence of minimizers” [11].

1 S. Götschel works on trajectory compression for finite element computations.

SOFTWARE PROTOTYPES FOR THERAPY PLANNING

1

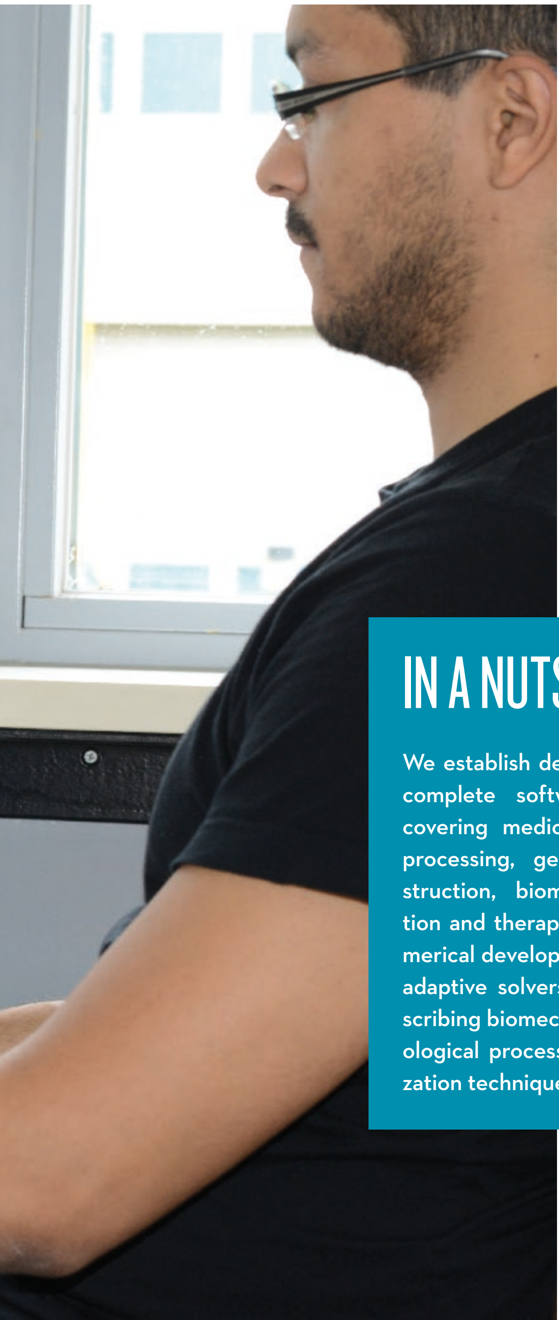
Algorithmic developments have to be implemented as a prototype software to be evaluated in clinical research. For the construction of such demonstrators we make use of the Amira and Kaskade 7 software developed at ZIB.

VISUALIZATION FRAMEWORK AMIRA

Amira is a visualization and data analysis software that was designed and developed at ZIB. It is now commercially developed and distributed under the product names Amira and Avizo by FEI Visualization Sciences Group. ZIB Amira is a variant of Amira that ZIB uses for ongoing research with cooperation partners. It contains well-known visualization and data analysis methods and results of ongoing research in various fields, such as image analysis, geometry processing, flow analysis, and molecular visualization. Prototypes of interactive therapy planning tools are implemented as special purpose modules in Amira.

FINITE ELEMENT TOOLBOX KASKADE 7

The workhorse for PDE problems is our finite element toolbox Kaskade 7 based on the DUNE libraries. Its C++ template structure allows a flexible, type-safe, and efficient mix-and-match approach to arbitrarily discretized systems of PDEs. The toolbox has been used to solve problems from various medical areas including cardiac excitation, hyperthermia cancer treatment planning, soft tissue deformation, and implant design. Recent improvements comprise spectral deferred correction methods for time integration, discontinuous Galerkin methods, and better memory access patterns on NUMA machines.

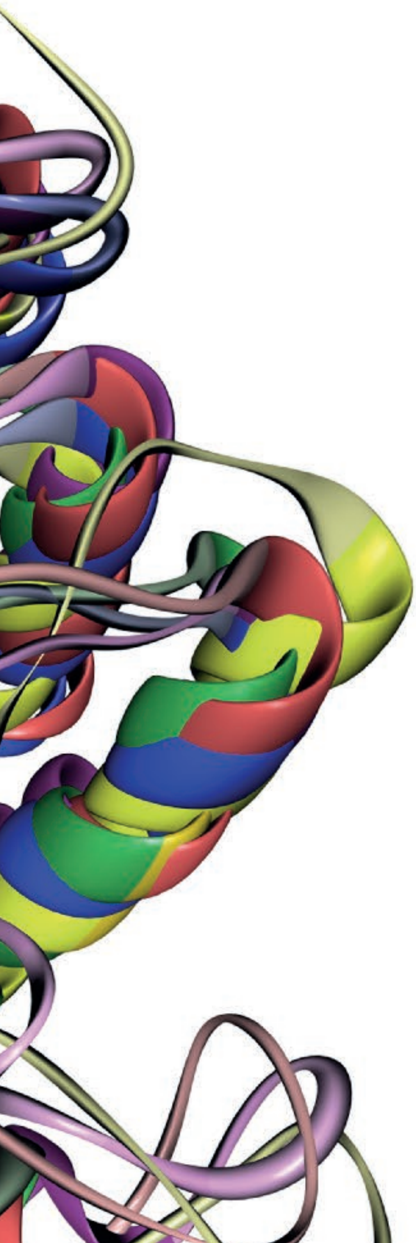


IN A NUTSHELL

We establish demonstrators of complete software pipelines covering medical image data processing, geometry reconstruction, biomedical simulation and therapy planning. Numerical developments focus on adaptive solvers for PDEs describing biomechanics or physiological processes and optimization techniques.

1 Software prototype for interactive patient-specific planning of bone anchored hearing aids. Research cooperation with 1000shapes GmbH and Unfallkrankenhaus Berlin-Marzahn (UKB).

1 *Antibiotic drug molecule interacting with different, superimposed bacterial target proteins*



DESIGN OF FUNCTIONAL MOLECULES

Mathematics, visualization and informatics join forces in molecular research

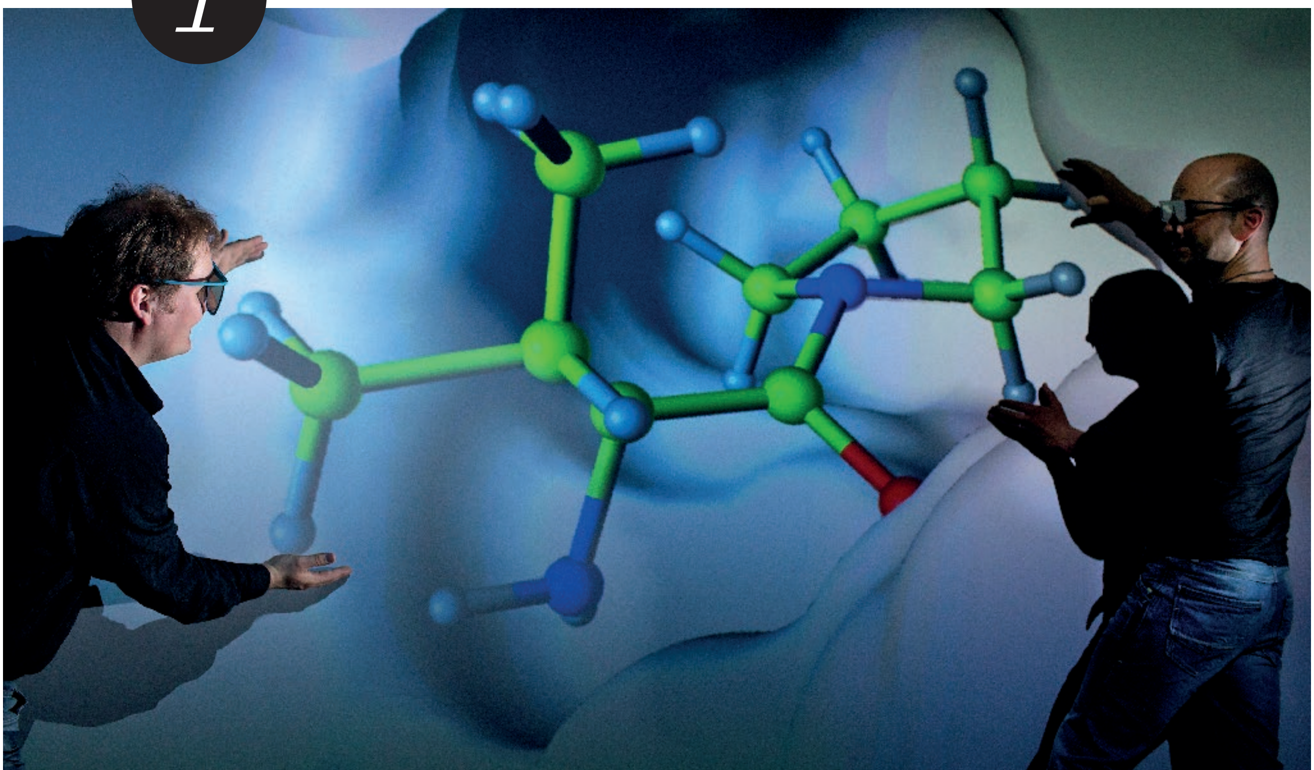
Small molecules like drugs have effects on humans. Some of these effects may be desired in a medical treatment. For a sustainable design of functional molecules, however, it is necessary to analyze undesired side effects, too. Only very efficient modelling and simulation approaches can tackle this problem.

FOR FURTHER INFORMATION PLEASE CONTACT:
Dr. Marcus Weber | weber@zib.de | +49 30 84185-189

CONTROLLING RISK EMERGING FROM DRUGS IN THE WATER CYCLE

After passing through our body remnants of drugs get into the water cycle. The largest part of these drug wastes are degraded in water purification processes. However, traces of the drugs themselves and in particular their transformation products can still be found in our drinking water and constitute a significant threat to our health. Systematical research on the transformation process has started only a few years ago. Just recently first ideas of how to save the water cycle from potential risks have been presented.

1



❶ *ZIB provides 3D facilities for the design of functional molecules*

❷ *A simple test in laboratory shows that a transformation product of SMZ has an antibiotic effect. This transformation product can be found in the water cycle and is less degradable compared to SMZ.*

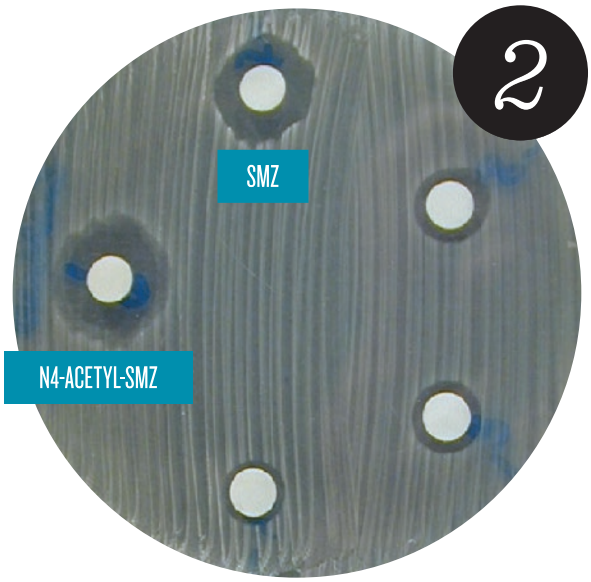
ANTIBIOTICS RESISTANCE

Since 2011, Vedat Durmaz is working at ZIB for the research project "TransRisk" funded by the Federal Ministry of Education and Research. TransRisk deals with the characterization, communication and minimization of risk emerging from new pollutants and transformation products in the water cycle. Together with Harald Mückter at LMU Munich, about 30 transformation products have been identified for the antibiotics sulfamethoxazole SMZ (Fig. 2). Transformation products may also have an antibiotic effect and, thus, can cause the increase of antibiotic resistant bacteria

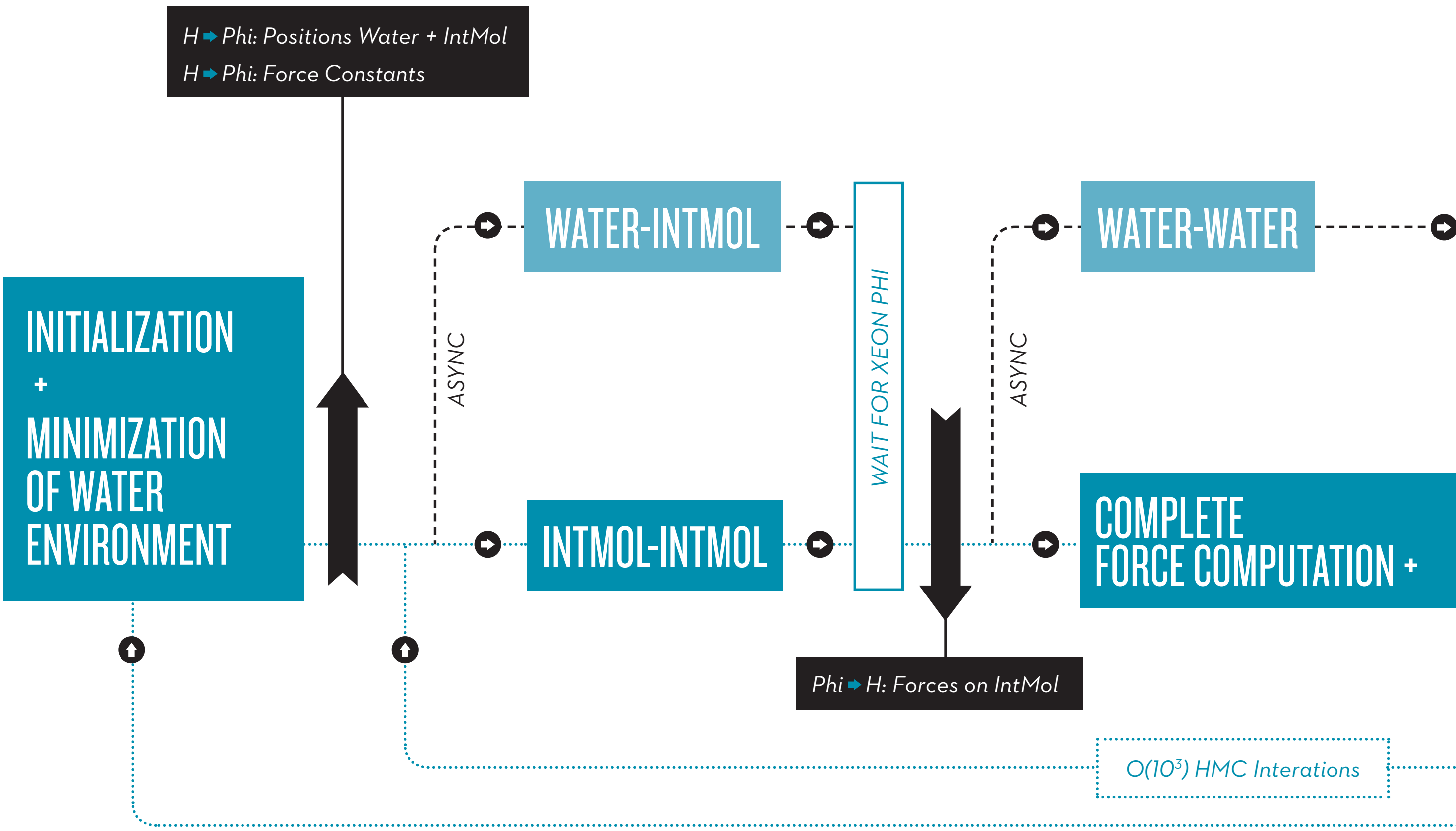
"With our computational strategy it is possible to prioritize the transformation products for further toxicological investigations", Vedat Durmaz says. The aim is, e.g., to compare the 30 transformation products with SMZ and to estimate which one has similar effects on the essential bacterial target protein. Marthe Solleder, a student of Vedat Durmaz, translates the molecular structures of the transformation products into physical models in the computer. On the basis virtual screening of all possible transformation products via molecular simulations are possible [12, 13]. By this strategy, previously unknown transformation products have already been identified that exhibit significantly strong antibiotic effects but are hardly degradable by water cleaning technology.

ESTROGENE EFFECTS

The virtual screening concept has already been successful for the estimation of estrogenic effects caused by transformation products. In this case, the target protein is the human estrogene receptor. In a cooperation with Christian Piechotta from the Federal Institute for Materials Research, the relative binding affinities measured in the laboratory were compared with the results from molecular simulations. "We found a very high correlation between these two quantities", Marcus Weber, leader of the ZIB working group, says [12].



TWEAKING MOLECULAR SIMULATIONS FOR MANY-CORE PROCESSORS



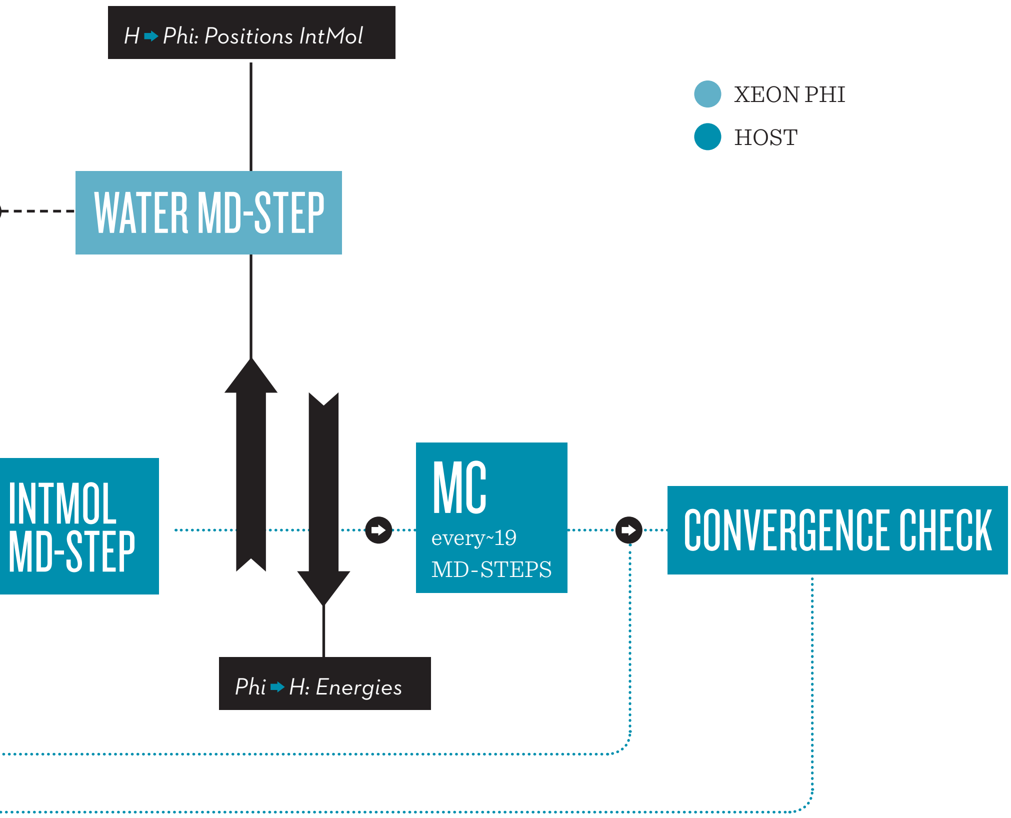
As described above, the simulation of drug-like molecules and the prediction of their interaction properties are of great interest. Unfortunately, for these small-sized molecular species, traditional computational approaches fail to benefit from many-core processors simply due to their insufficient problem size. Even if one considers molecules in its aqueous environment, when solvation effects have to be taken into account, all-atom simulations including sufficiently many water molecules tend to be still too small-sized to exploit the computational capabilities.

CONCURRENT KERNEL EXECUTION

One possible solution for modern many-core processors is to utilize the feature of Concurrent Kernel Execution, meaning that multiple computational tasks are offloaded to the many-core device so that a higher resource utilization and significant speed-ups can be achieved. Thomas Steinke, the head of the supercomputing group at ZIB, explains a problem with this approach: "Although the Concurrent Kernel Execution feature is offered on the Nvidia Fermi GPU architecture we identified weaknesses in its implementation." Together with cooperation partners his group developed an alternative approach to allow for a higher computational throughput for molecular simulations [1]. The software GLAT [2] is based on a new idea of how to compute the interactions of surrounding water molecules on many-core processors (see Fig. below).

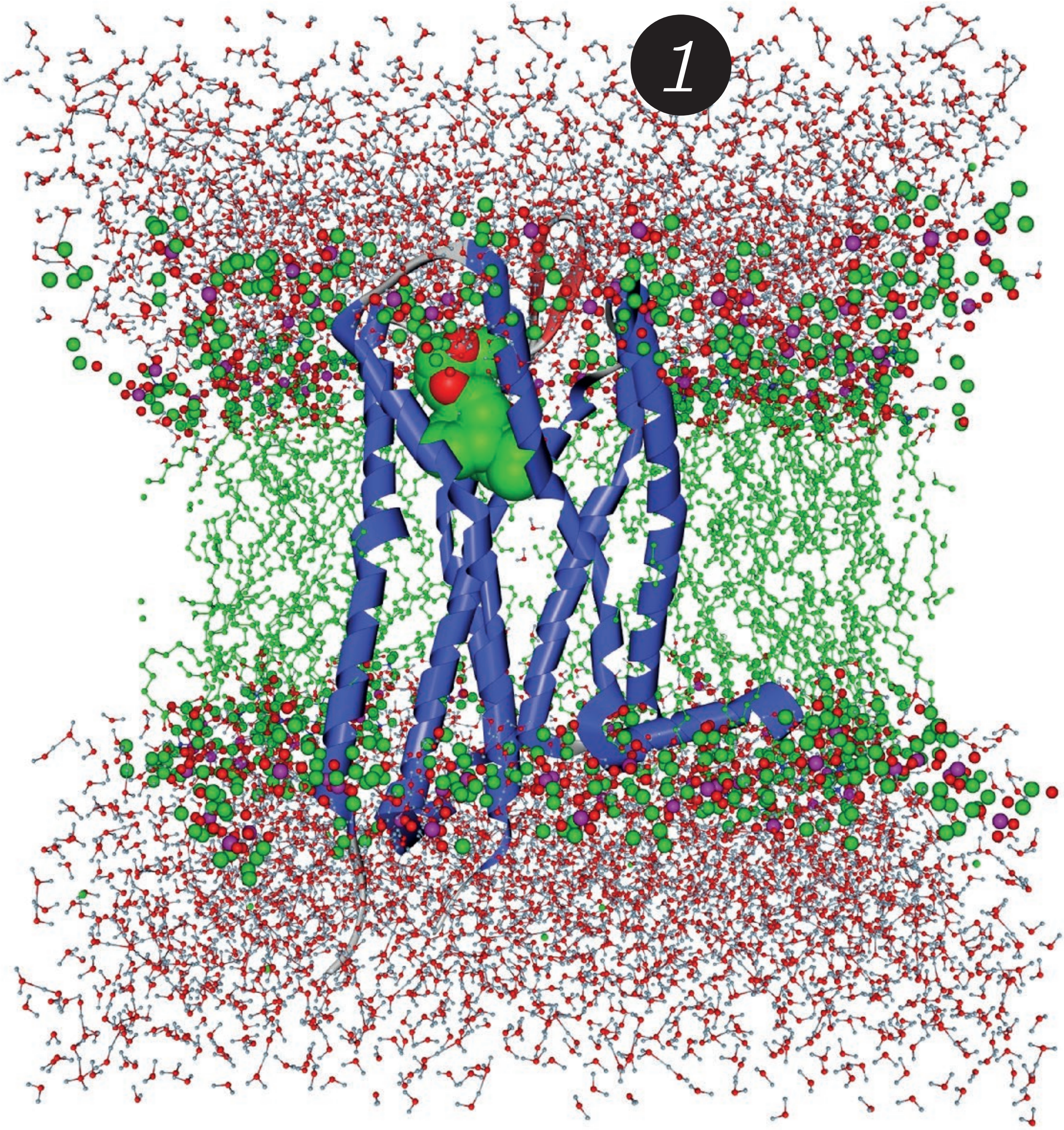
HIGH SPEED-UP FACTORS

"We continued our architectural analysis on today's many-core platforms NVidia Kepler GPU and Intel Xeon Phi", Thomas Steinke says. His group demonstrated improvements of vendor architectures and the advantages of their software approach [14, 15]. "With two Xeon Phi processors we can handle twice the number of simulations with almost no increase in the overall execution time", he explains the advantages of the new approach which helps researchers in the CMD group to approach more advanced problems.



EXPLANATION BOX

Heterogeneous computer architectures with many-core processors like general-purpose graphics processors (GPU) or Xeon Phi provide data processing capabilities which are an order of magnitude larger than today's standard multi-core CPUs. This computational potential can be profitably unlocked for molecular simulations if the implemented algorithms are highly parallel and the problem size suitable large. Otherwise, these resources cannot be fully utilized and, thus, speedups are rather unlikely.



COMPUTATIONAL DRUG DESIGN

Small molecules not only lead to undesirable effects. Our new molecular simulation methods are also able to investigate new drug candidates for clinical research. The CMD working group at ZIB has successfully invented a new pain relief drug concept.

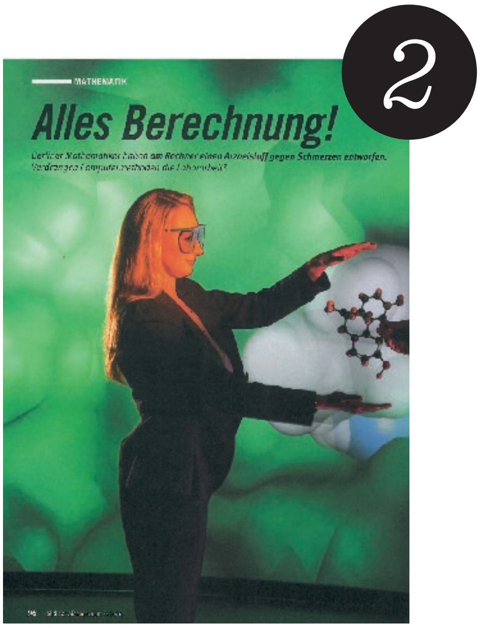
DESIGN OF PAIN RELIEF DRUGS

Opioid agonists, like the well-known morphine, are essential analgesics in the treatment of moderate up to severe pain. Prof. Stein from Charité Berlin explains the problem: "Systemically applied conventional opioid agonists activate opioid receptors (Fig. 1), resulting in analgesia. The use of currently available opioids, however, is limited by major adverse side effects such as the central or intestinal effects [16, 17]." The aim of a BMBF cooperation project is to design opioids acting selectively outside the brain and gut to avoid such side effects.

Olga Scharkoi (Fig. 2) from the Computational Molecular Design working group describes the idea: "We want to activate peripheric opioid receptors only in inflamed tissue. This tissue has a lower pH-value than healthy tissue." By the use of conformation dynamics methods, three opioid candidates have been proposed, which are presently undergoing successful in-vitro and in-vivo experiments at Charité [18, 19].

MECHANISMS OF OPIOID RECEPTORS

Besides the design of new drugs, the low pH-value in inflamed tissue is supposed to have further effects on the opioid receptor and the activation process. Martina Klimm is a fellow of the SALSA graduate school. She points out: "Because the opioid receptor system is rather complex, classical molecular simulation methods are not leading to sufficient results. One has to use our new methods to obtain statistically appropriate amounts of data from molecular simulation." A lot is known about the rhodopsin receptor which is a very similar G-protein-coupled receptor. The available simulation data for rhodopsin can be reconsidered for the opioid receptor. Martina Klimm describes her work: "The challenge is to derive the global behavior of the opioid receptor from the available local samplings of the rhodopsin receptor. For this we invented direct reweighting strategies [20]." In order to analyze the huge statistical data in the end, coarse graining methods are needed [21]. Together with spectroscopic experiments performed at Technical University Berlin, within the group of Prof. Peter Hildebrandt, the results will help to understand the effects of analgesics at different pH-values.



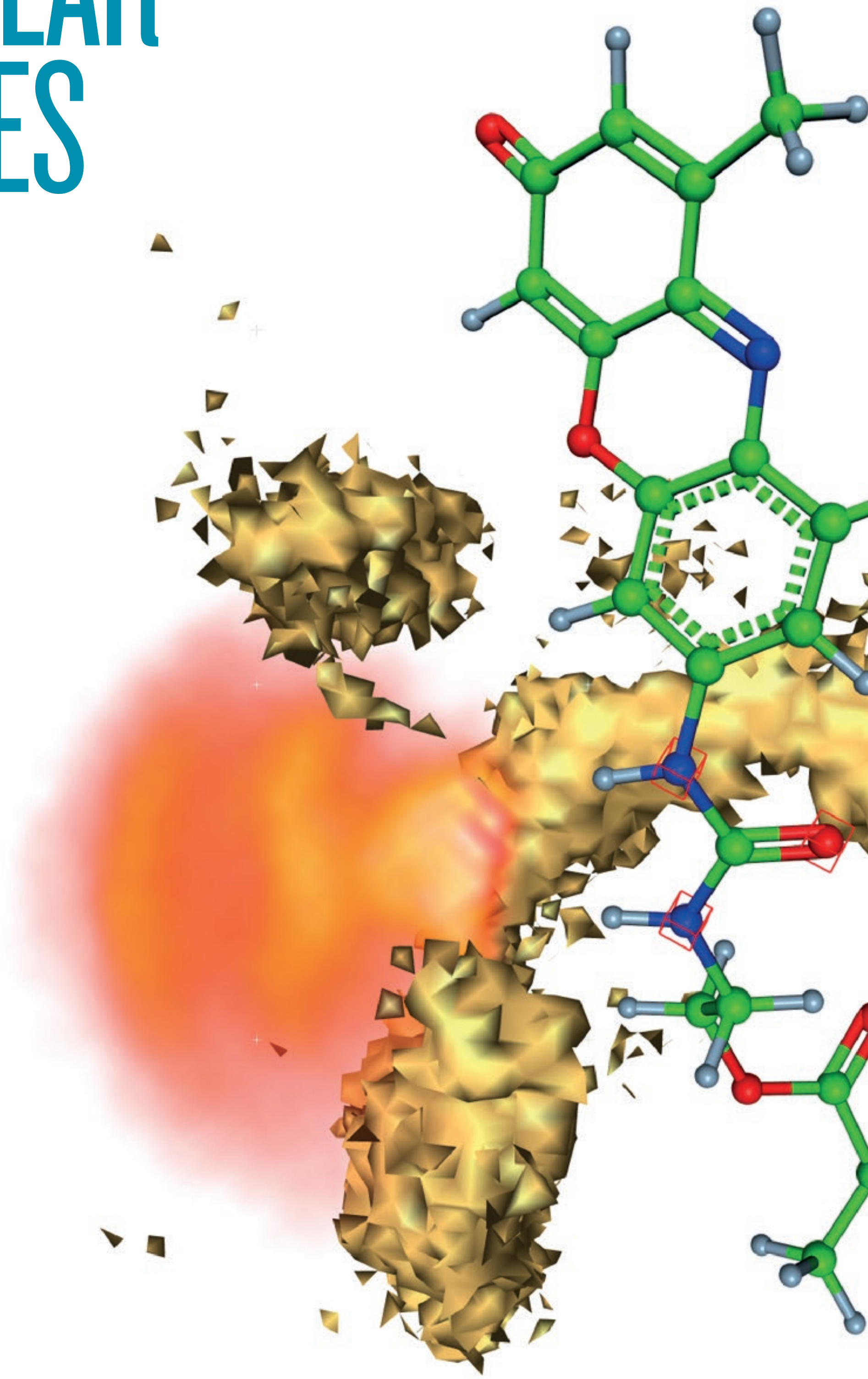
1 Opioid and μ-opioid-receptor within its aqueous-membrane surroundings

2 Publication in »Bild der Wissenschaft« concerning the opiod design.

CAVITY DETECTION IN MOLECULAR STRUCTURES

Opioid receptors are located in the cellular membrane. Membrane proteins are often used for the transport of signals or for pumping ions by forming cavities.

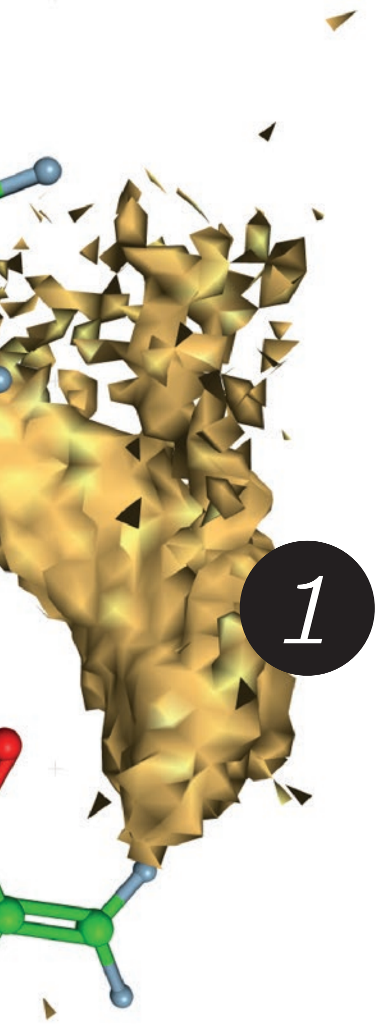
The visual reconstructions of the geometrical shape of cavities are of high interest. During the last decades many algorithms have been presented to achieve this. However, the geometrical accuracy of most of these algorithms is often insufficient. Additionally, many techniques are restricted to compute only a small subset of the internal cavities.



1 The molecular structure visible in ball-and -stick representation is the template-detecting part inside a MIP. Isosurfaces show the most probable position of polymer atoms which form a specific binding cavity for the template molecule. The flexibility of the template molecules is indicated with a »cloud«.

GENERALIZED VORONOI DIAGRAMS

To overcome these limitations, the visualization group of Hans Christian Hege developed a new approach based on the generalized Voronoi Diagram for spheres that represent the atoms [22, 23]. Norbert Lindow is a coworker in this group: “We have implemented an efficient algorithm to compute the Voronoi (see Fig. 1, page 46) diagram and to extract and visualize all cavities in a molecule at once.”



1

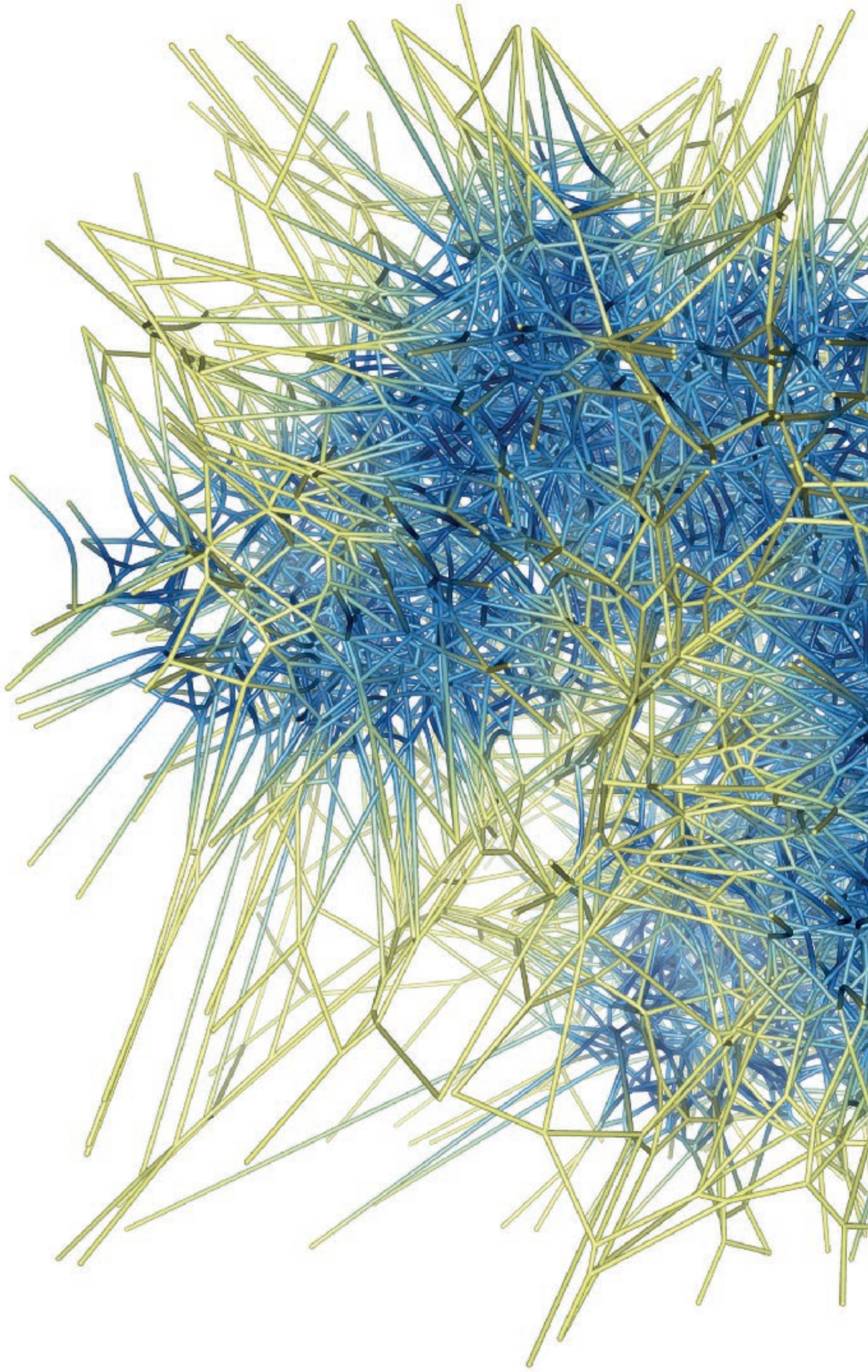
MOLECULAR IMPRINTING – AN EMERGING TECHNOLOGY

The mechanisms which are responsible for the formation of binding cavities are very important for the studies of molecularly imprinted polymers (MIPs) [24]. “This understanding will facilitate the rational development and the actual preparation of MIPs, in particular, of MIPs with intrinsic signal generation”, says Knut Rurack the cooperation partner from the Federal Insitute of Materials Reseach and Testing: “Since several decades, the concept of molecular imprinting is successfully explored as a synthetic alternative to immunochemical recognition methods [25].” The molecule to be detected by the MIPs is denoted as template molecule. Vedat Durmaz from the CMD working group creates physical models for each proposed polymer mixture: “Each polymer mixture is simulated on the computer twice, once with template molecule in it and once without.” From this computational data the cavity forming quality is derived in order to find the best MIP with the most specific binding of the template molecule.

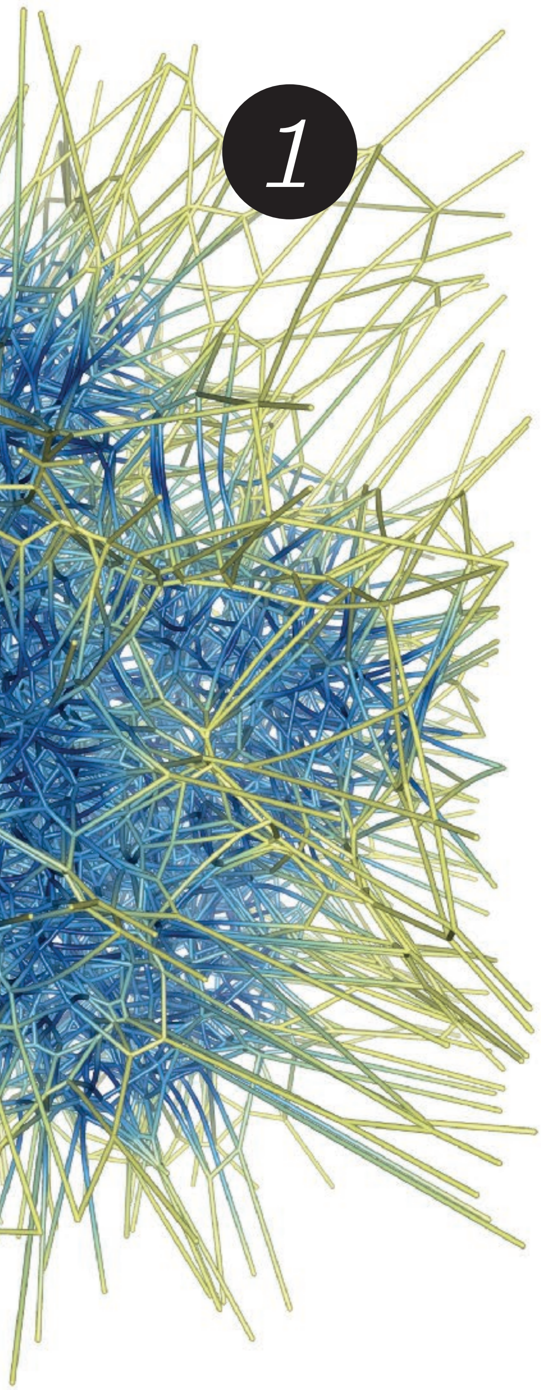
MOLECULAR CAVITY DYNAMICS

The internal cavities of proteins are dynamic structures and their dynamics may be associated with conformational changes that are required for the functioning of the proteins. In order to study the dynamics of internal protein cavities, appropriate tools are required that allow rapid identification of the cavities as well as assessment of their time-dependent structures. In a cooperaion with Ana-Nicoleta Bondar from FU Berlin [26, 27], the visualization group of Hans Christian-Hege applied their above mentioned static cavity detection algorithm to dynamic data from molecular simulations. “Initially, the cavities are computed for each time step separately”, explains Norbert Lindow: “The new feature allows the user to interactively select a set of cavities in an arbitrary time step and trace them over time.” By this procedure the geometrical overlap of cavities of consecutive time steps is analyzed. Another coworker in the visualization group is Daniel Baum: “Our techniques for visualizing the cavity dynamics in a single image helps to analyze the behavior of the cavity structure. Apart from the shape dynamics one can also analyze topological events like splits and merges by inspecting abstract graph representations.”

ADVANCED MATHEMATICAL METHODS FOR MOLECULAR SIMULATION



1 *Voronoi diagram of possible molecular pathways through a protein structure.*



REBINDING EFFECTS

Classically ligand receptor interaction is described using a key-lock analogy: the ligand (the key) seeks for a suitable binding site at the receptor (the lock). "However, this analogy is no longer justified, when considering a real docking process in a thermodynamic environment", Konstantin Fackeldey from the CMD working group argues: "Here, the ligand and the receptor are flexible and constantly in motion. Moreover, a receptor often has more than one binding site." The receptors and the ligands can be presented in a multivalent manner, allowing for a simultaneous binding on multiple sites. This kind of binding is investigated together with the Collaborative Research Center CRC 765 located at FU Berlin. In the joint project it has been figured out that the spatial arrangement of multivalent ligands and receptors [28, 29] lead to faster association and slower dissociation kinetics if the ligands approach the receptors.

One class of efficient simulation methods to model such a binding process are a special class of conformations dynamics algorithms, the so-called Markov State Models, that aim at a partition of the state space into subsets such that the dynamical behavior of the system can be described in terms of transition probabilities which lead to a much better understanding of the processes.

NEW MATHEMATICAL APPROACHES

All conformation dynamics approaches like Markov State Modelling require a Galerkin discretization of the underlying transfer operator. Despite considerable progress in the theory of transfer operators and metastability in the last decade, fundamental problems are still open. Adam Nielsen, working at ZIB as a graduate fellow at the Berlin Mathematical School, the joint graduate school of Berlin's mathematics institutes, has solved one of these problems: "We were able to show for which classes of ansatz spaces a set-based Galerkin discretization of the generalized transfer operator converges strongly in the L2 norm" [30].

Besides such rather fundamental results the research at ZIB still aims at novel algorithmic breakthroughs. One of the main algorithmic challenges in MD still is efficient generation of rare event statistics: Most of the available approach to sampling rare event statistics suffers from very large variance and the need to sample very long MD trajectories [31, 32]. Very recently the CMD group and working groups at FU Berlin have joined forces to develop a novel approach that allows for the construction of low variance estimators while keeping required trajectories short. The key idea is to use non-equilibrium driving protocols to enforce the rare event under consideration, thus pushing it to shorter time scales. Novel mathematical results show that there is an optimal driving protocol that allows for reproduction of the exact equilibrium rare event statistics via the non-equilibrium sampling with zero variance. Wei Zhang, a postdoc at the CMD group, has even been able to show that suboptimal low variance estimators can be computed using low-dimensional reaction coordinates of the system. "These surprising results are the basis of the development of a new class of rare event simulation algorithms", Christof Schütte explains the future strategy of ZIB in this field.

FOR FURTHER INFORMATION PLEASE CONTACT:
Prof. Dr. Ralf Borndörfer | borndoerfer@zib.de | +49 30 84185-243

MATHEM
IN T
AND TRANS

Traffic in Hanoi in 2013

MATICS RAFFIC SPORT

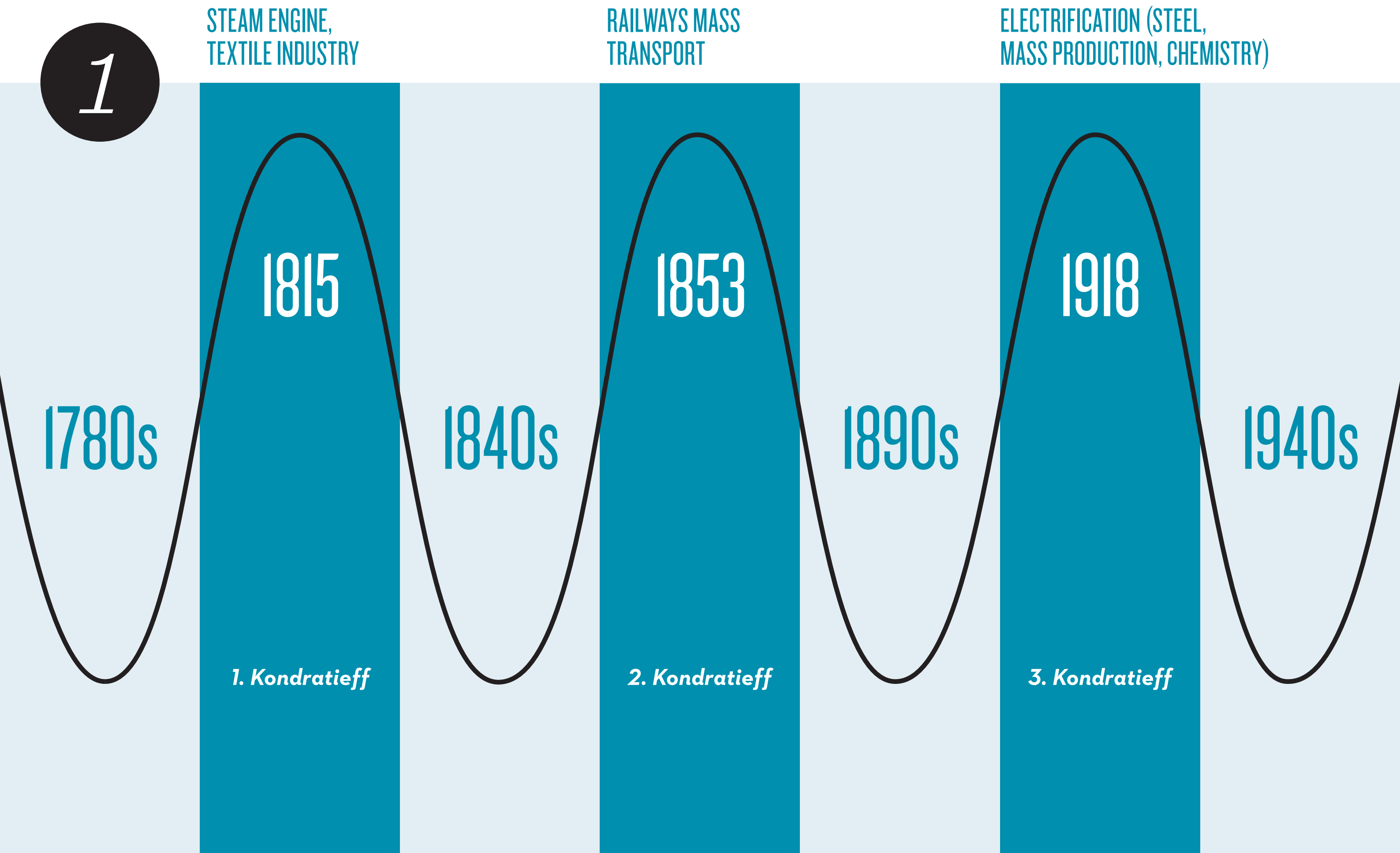
Will transport become again a revolutionary key technology?

Public transit, railways and air transport are all about timetables, schedules and plans. What else is this than a gigantic optimization problem? Its solution "cries" for optimization methods. And these are now available: Modern IT provides the necessary data and control, and mathematics the "artificial intelligence" that we need to solve the transportation problems of today. Traffic and transport can play a leading role in establishing mathematics as a key technology of the 21st century in industry and society and, once again, become a driving force of innovation.

WHAT IS THE NEXT KEY INNOVATION?

Traffic and transport have always been key industries in the history of mankind. Progress often came from revolutionary innovations in vehicle technologies such as domesticated animals, the wheel, sailing ships, railways, subways, cars, and aircraft, as well as from the buildup of corresponding infrastructure networks and suitable organizational structures with which the transportation problems of their time could be solved, see Fig. 1.

»A model that might have taken a year to solve 10 years ago, might now solve in less than 10 seconds.«



TODAY'S TRANSPORTATION CHALLENGES

These arise from urbanization, ageing societies, and a general scarcity of resources, including environmental ones. More than half of the world population already lives in cities, more will, and everybody has witnessed chaotic traffic situations in the megacities of Asia, America, or Africa. In Europe, people also move into the cities, depopulating the countryside, with mostly elderly people staying. In general, costs for traffic and transport are rising, and people complain about congestion and pollution.

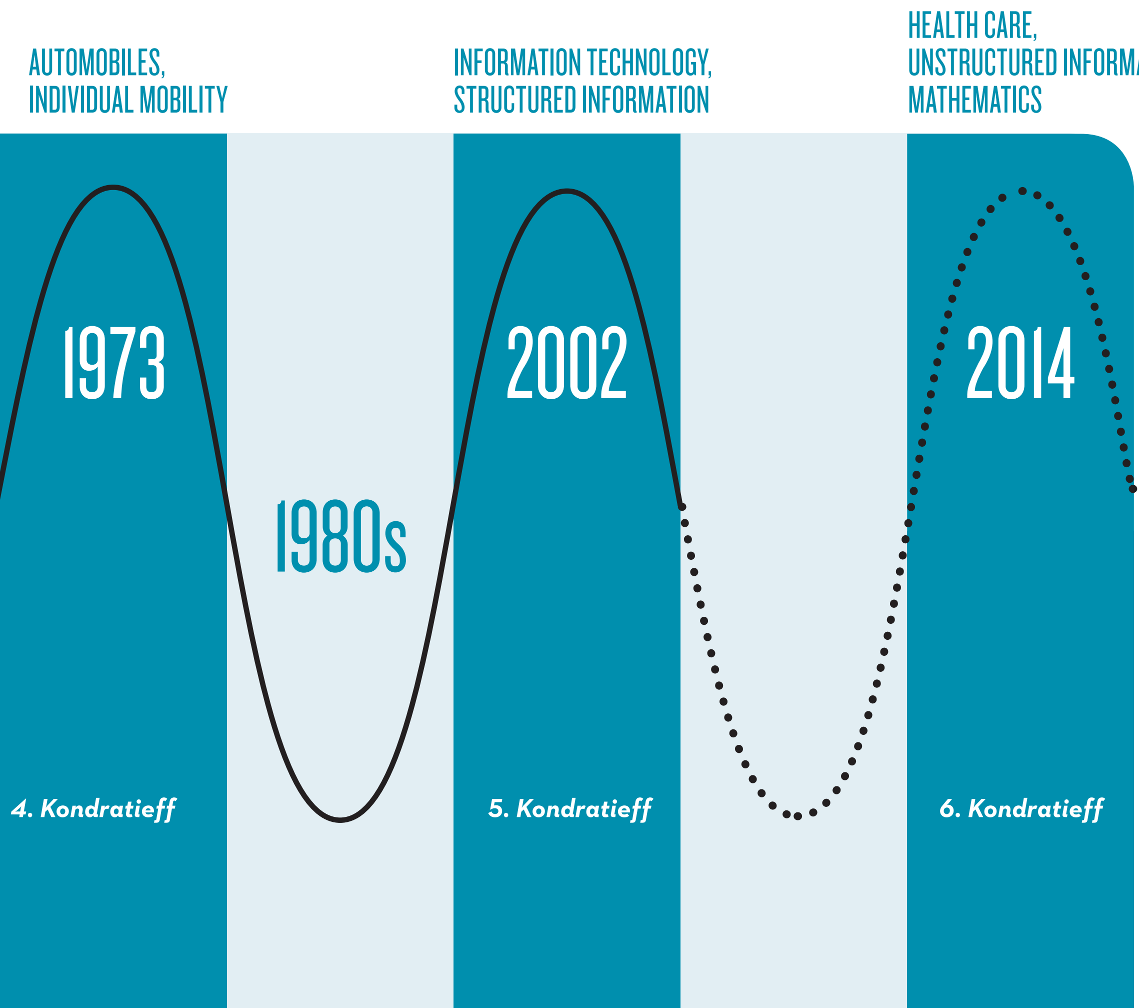
WHAT CAN WE DO?

It would be great if a technological breakthrough like a teleporter would come to our rescue, but this is unlikely – and we also do not need it! We witness significant progress on improving the efficiency of the existing vehicle technologies by downsizing engines, new fuels, better aero- und hydrodynamics, and by using modern information technology for trip planning and route guidance, online information, and internet ordering. In fact: The existing means of transportation could be extremely efficient, if their system advantages would be put to use and the disadvantages could be avoided.

All we need is to make efficient use of these technologies, and the key to this is mathematics and information technology.

WHY MATHEMATICS CAN IMPROVE TRAFFIC AND TRANSPORT

This is actually an old dream, at least 100 years old. Nobel laureates like Leonid Kantorovich and Tjalling Koopmans had it, and operations research was the discipline that tried to put it into practice in the 1960s. We know it didn't work that time. But things have changed. Modern IT can provide all needed data and control. Computing power has increased dramatically. And what is not widely known: mathematical algorithms have improved at least as much as CPUs. Progress in computing speed and algorithms multiplies, making it now possible, for the first time in history, to solve large scale transportation problems to proven optimality. Traffic and transport are particularly suited to be optimized. They can spearhead revolutionary applications of mathematical optimization methods in industry and society [33].



1 Several (Kondratieff) waves of innovations here triggered by progress in traffic and transport.

MATHEMATICS IN PUBLIC TRANSIT

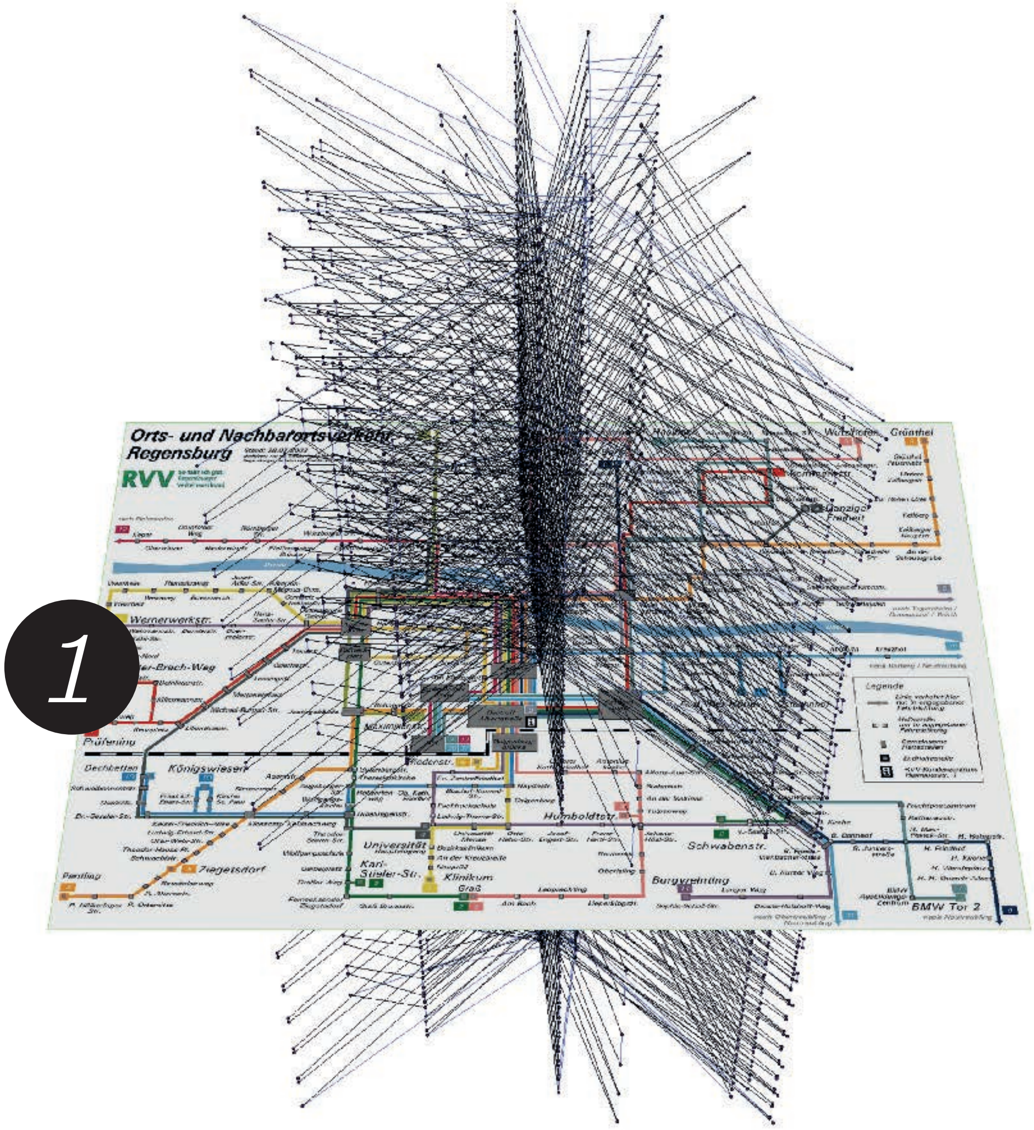
Optimization and public transit fit well together. The rotations of buses and subways, and the schedules of their crews are nowadays commonly computed by optimization algorithms. Vehicle and crew scheduling lead to classical combinatorial optimization problems on network flows and path covers that are nowadays well understood. A solid theory and powerful algorithms have led to the development of mathematical vehicle and crew "optimizers". These are on a mature level and available within the leading software systems for public transit companies [34].

SUCCESS STORIES IN VEHICLE SCHEDULING

Vehicle scheduling is about the "flow" of multiple types or "commodities" of vehicles through a graph of timetabled and deadhead trips (see Fig. 1). In Berlin, e.g., more than 1,000 buses of 50+ types serve more than 28,000 timetabled trips a day, connected by 100 million possible deadhead trips. Algorithmic techniques such as Lagrangean pricing, developed at ZIB, can identify the relevant parts of such gigantic graphs, and solve such problems to proven fleet optimality. Our BVG results still mark a world record in this area. The system-wide view allowed for shifts of vehicles between the morning peaks in the Eastern and Western parts of the city, resulting in substantial savings.

COMBINING ALGORITHMS FOR INTEGRATED PLANNING

Similar, but more complex methods, are used to schedule crews, producing a process chain, in which vehicles are scheduled first and crews second. Of course, doing both steps at the same time would be even better, and this type of integrated scheduling is a very active research topic. ZIB has pioneered this area by developing a bundle method, that allows the integrated optimization of vehicle and crew schedules for regional scenarios. This approach is based on a Lagrangean relaxation of the problem, which assumes that the two parts of the problem can also be addressed independently. This is sometimes a problem, for example in staff rostering, where it is hard to come up with a crew roster for a month if one has no idea about the daily duties. We have recently shown that Bender's decomposition approaches can work in such a setting. Integrated optimizers based on Lagrange and Bender's approaches will be the next level of optimization in public transit optimization.



WHY NOT OPTIMIZE THE BIG DECISIONS?

All of these methods only deal with implementing a given service at minimum cost. The real potentials, however, lie in the optimization of the public transport system itself. The network and the line plan, the timetable, and the fares can be optimized in order to find a best compromise between quality and costs. These problems deal with passenger behavior and are mathematically difficult. ZIB has developed a theory of path connectivity in order to solve line planning problems with a dynamic passenger routing, including a novel method to deal with direct connections and transfers. This model has been used to compute the line plan that is in operation at ViP Potsdam, a real breakthrough in this area.

1 The typical star-shaped structure of an urban vehicle scheduling graph (only timetabled trips, time goes upwards).

RAILWAY OPTIMIZATION FINALLY WORKS

Railways are at the origins of traffic optimization, because trains must operate according to precisely planned schedules. Early works such as Tolstoí's augmenting cycles, Ford and Fulkerson's theory of network flows, and Charnes and Miller's set partitioning approach were all invented for railway applications. These are, however, difficult: trains must keep safety distances, vehicles must be combined into trains, etc., i.e., railways require integrated planning approaches with an eye to many details.

1

Image courtesy: www.siemens.com/presse

WHY RAILWAY OPTIMIZATION IS SO DIFFICULT

Vehicle rotation planning (vehicle scheduling) must include train composition, and cannot be done on the level of individual vehicles, as in public transit. Extending bus rotation models by additional constraints did not work. In a project with DB Fernverkehr AG, we came up with a model that expresses transitions of vehicles in a train and the formation of trains from vehicles by configuration variables, that can be viewed as hyperarcs in a scheduling hypergraph, and consequently, vehicle rotations correspond to hyperflows. We have started to develop an algorithmic theory of hypergraphs that allows to solve, for the first time, large scale railway vehicle rotation planning problems for the entire ICE fleet of Deutsche Bahn on a strategic planning level. Now that the basic approach works, we can start to develop adaptive discretization techniques in order to address more detailed dated problems closer to the day of operation.

TRACK ALLOCATION

Once the rotations have been determined, train paths have to be embedded into the available network infrastructure. To this purpose, trains have to be assigned to particular tracks and platforms, and at precise times, such that minimum headways and other safety and technical constraints are satisfied. This so-called track allocation problem is also useful to determine infrastructure capacities in investment planning, or in a marketing process that allocates capacities according to the highest bid (this is actually prescribed in the German railway legislation). The track allocation problem is notoriously difficult. ZIB has developed a novel approach that is based on the consideration of train configurations on railway track segments using a bundle approach. With this technique, it was possible to solve, for the first time, instances involving several hundred infrastructure elements and several hundred trains. We used this method in project with the Swiss railways SBB in order to analyze the capacity of the Simplon corridor through the Alps under different regimes of operation [35].

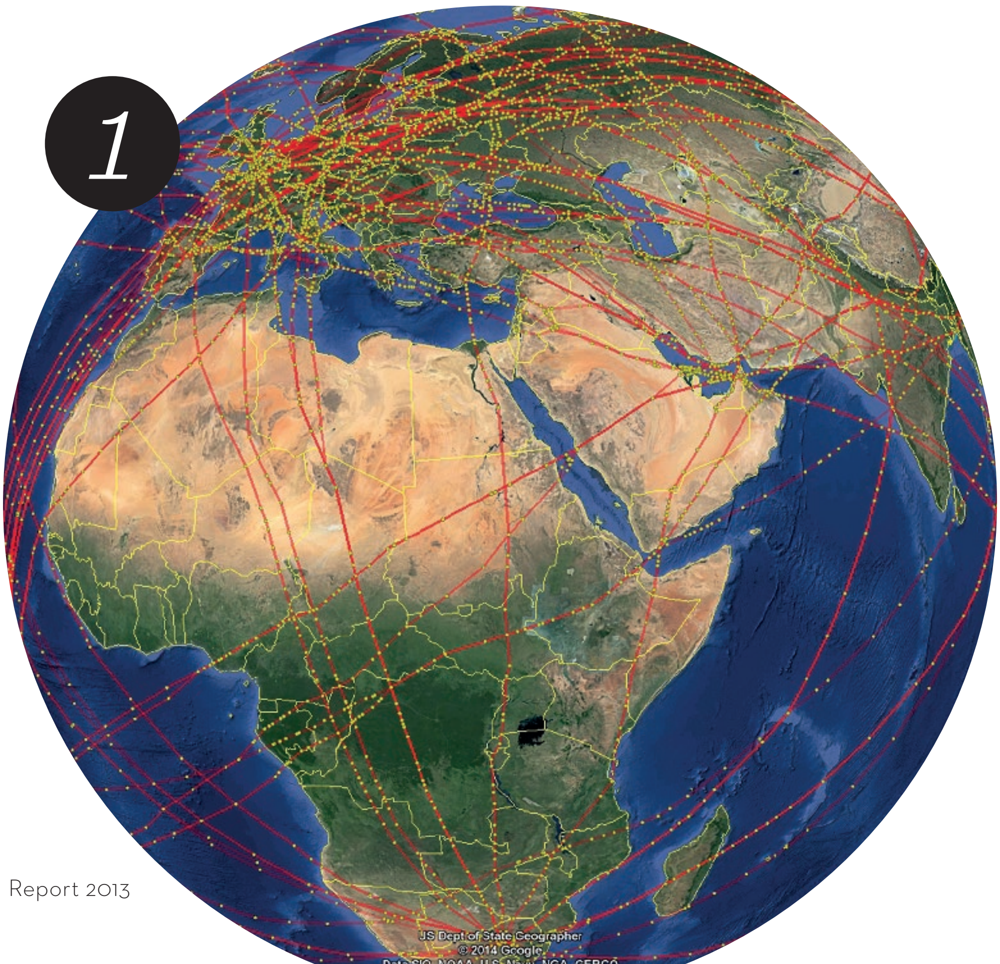
Railway optimization has finally started to work. Enormous potentials wait to be uncovered. These can substantially increase the overall attractivity of the railway system.

1 ICE high speed trains can consist of several railcars in different orders and orientations.

AIR TRAFFIC OPTIMIZATION FLIES

The airline industry started to boom in the eighties. Networks were reorganized in a hub-and-spoke fashion, new low cost carriers emerged, and prices fell, sometimes to unbelievably low levels. Many of these changes are closely linked to progress in mathematical optimization. Ticket prices are the result of sophisticated demand forecasts and revenue maximization methods that try to segment the market and skim the customers' willingness to pay. Efficiency increases are due to fleet, aircraft rotation, and crew scheduling, they were first applied here at an industrial level. Today, all planning processes in the airline industry are heavily supported by mathematical optimization methods.

1



FAILURE SAFE PLANNING

In fact, efficiency has been pushed to a level that has eliminated many traditional buffers. In such a system, disruptions and delays can spread easily and it can happen that a plan that looked perfect on paper turns out to be very bad under only slightly unfavorable conditions. The consequence is to optimize not only under ideal conditions, but to anticipate possible failure situations. There are various concepts how to do that, ranging from stochastic optimization, which assumes complete knowledge of the probabilities of all possible failures, via a safeguarding against a worst case in robust optimization to online optimization without knowledge about the future. Failure safe planning is of course much more difficult from an algorithmic point of view, and it is not only a purely mathematical question to find the right concept.

MAKING AIRCRAFT ROTATIONS ROBUST

Computing robust rotations that can absorb delays up to a certain point by judiciously allocated buffer times is one question of this type. We have proposed a stochastic optimization concept that is based on propagating delays along individual aircraft rotations by an efficient computation of delay convolutions. It turned out that delays depend on relatively few factors, such that the associated probabilistic delay model can be used for forecasts. In simulations, reductions in expected propagated delays of up to 10% could be demonstrated [36].

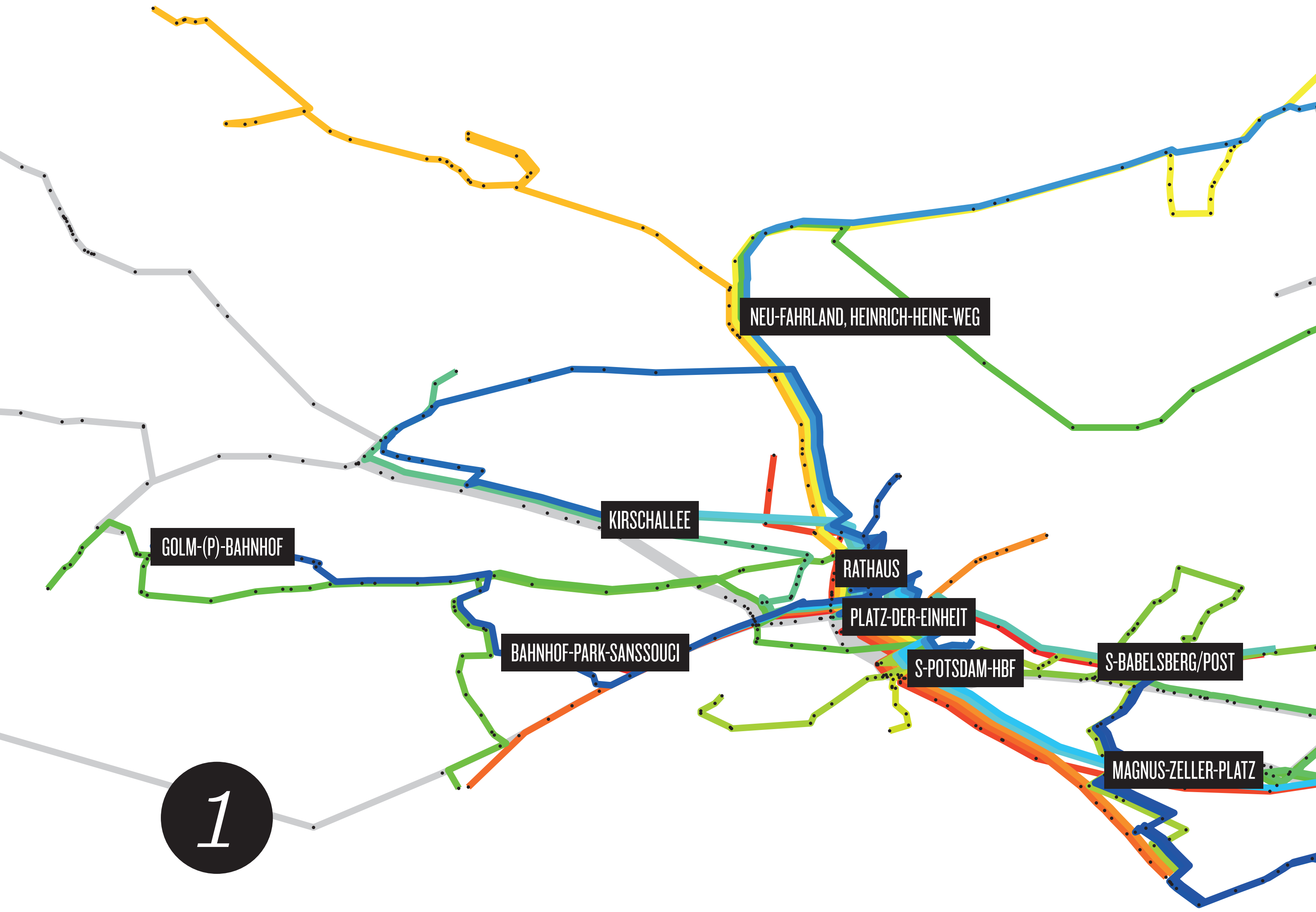
FUEL EFFICIENT AIRCRAFT TRAJECTORIES

Air traffic has nowadays also reached a density where the airspace becomes crowded. Air traffic control reacts by enforcing more and more traffic restrictions. Their complexity has reached a level at which traditional optimization methods for aircraft trajectories are at their limits, and the increasing ranges of modern aircrafts don't make the task easier, see Fig. 1 . New methods are needed in order to compute fuel efficient trajectories subject to changing weather conditions and complex air traffic restrictions. We investigate this problem in a new project with Lufthansa Systems Berlin, the manufacturer of the market leading LIDO software system for aircraft trajectory optimization.

1 Air traffic network.

INDIVIDUAL TRAFFIC

Car traffic is much more "anarchistic" than public transit. Rail and air traffic require optimization by a central powerful agency. Car traffic is more like a game, in which players pursue their interests and finally arrive at some kind of equilibrium solution. In fact, if we get stuck in a traffic jam every morning, we will try alternative routes, until we cannot improve our travel times any more. If we want to optimize car traffic, we must therefore take user behavior into account. This is, of course, difficult to predict. On the other hand, it is even more difficult to come up with a good system by trial and error, or by relying on some kind of evolution.



1 The public transport line system in Potsdam has been optimized for a best possible compromise between cost and quality of service.

INSPECTORS AGAINST FARE EVADERS

Highway fees are an example of an instrument that can be used to control individual traffic, in this case, truck traffic. In Germany, all trucks with a weight of at least 3.5 tons have to pay a highway fee depending mainly on the distance travelled and the vehicle weight. The toll is enforced by automatic gantry bridges and by highway inspectors, which perform mobile controls. In this case, the game is how to allocate the mobile controls in such a way as to establish a maximal level of control. In a project with the Bundesamt für Güterverkehr, which carries out these controls, we have developed a game theoretic approach to this problem, that predicts the behavior of fare evaders and plays a stochastic maximum control strategy against them. Moreover, the strategy can even be published without losing its efficiency. Such a strategy is way more effective than controlling, e.g., by the level of usage [37].

OPTIMIZING TRANSPORTATION SYSTEMS

Mathematical optimization has just started to unveil its enormous potentials in traffic and transport optimization. These are needed in order to solve the gigantic transportation problems of the present and future. Mathematics currently contributes most to operational planning problems to allocate and schedule vehicles and crews. This is important, but has a much smaller leverage than the big decisions of system design, see Fig. 1. It is not a good idea to pour the concrete first and think about an optimal operation of the resulting system afterwards. More effort must be invested into good planning, and mathematics is the key technology that makes the difference. This certainly also needs substantial mathematical progress. The start has been made, and the perspective is there: Mathematics can revolutionize traffic and transport.

SHEPHERDING THE BITS

Gelbe Box mit Datenträgern, Tim Hasler, cc-by-nc

Challenge accepted!

Our society has become digital. The Internet is omnipresent. We read, listen, and view content through data streams. We save mails, photos, videos, music, personal and work-related data in the virtually unlimited Internet cloud. We utilize search engines to acquire new knowledge. We quantify ourselves by constantly providing information to the Net and to our "friends and followers". What we do, where we are right now, how many steps we have made and the calories we have consumed is often shared with the world automatically . Those who do not also embrace a digital lifestyle may soon be sidelined.

FOR FURTHER INFORMATION PLEASE CONTACT:
Prof. Dr. Thorsten Koch | koch@zib.de | +49 30 84185-213

PRESERVING OUR DIGITAL HERITAGE

Even supposedly archetypical analogue areas such as those venerable institutions preserving our cultural heritage or humanities research, have arrived in the digital age. Thus, archives digitize files and estates in order to save them from decay. These collections archive increasingly more digitally born files. Museums present digital representations of their analogue collections in portals, such as the German Digital Library or the Europeana, entire art forms such as video installations, are developed purely digitally. Libraries buy electronic books, journals and databases and administrate licenses instead of shelf metres. But, also science acts digitally. Increasingly, not only publications are considered to be the result of often long years of scientific work, but also the underlying research data. Besides classical research data such as results of measurements or observations, due to the increasing digitization of humanities, data records that supposedly do not fall into the area of research also become the subject of investigation. For instance, social behavioural patterns may even be extracted from series of "Twitter feeds" that appear superficial. Another research area in humanities is concerned with investigations about how far modern, to a certain extent equalizing communication channels may even influence the development of societies.

An example for this is the "Arab Spring", which gained significant momentum by the uncensored coverage in so-called "Social Networks". Due to their temporal relevance, these data collections are, comparable to atmospheric measurements in Antarctica, unrepeatable, irretrievable and therefore a treasure for future generations of researchers, which should be preserved carefully. With the digitization of society, culture, research and science, a digital cultural heritage is growing day by day, which demands sustainable preservation.

Digital cultural assets – just like analogue objects – must also be administrated, preserved, maintained and be made available on a permanent basis by institutions of cultural heritage. The new digital collections pose enormous challenges on cultural heritage and research institutions. The long-term and trustworthy safeguarding of the availability of digital information comprises the entire life cycle of digital objects which includes data management as well as rule-based enrichment and processing of data, the digitization and the persistent long-term availability of information. All this can not be mastered by small and medium institutions alone.

In Berlin, it has been decided to opt for a cooperative way to digitize cultural assets and ensure their long-term availability. In 2012, the Cultural Affairs Department of the Federal State of Berlin initiated an interdisciplinary digitization support programme that was hitherto unprecedented in the whole of Germany. With this initiative, digitization projects in various institutions of cultural heritage in Berlin are supported financially, and a central consulting and coordination office of the Federal State was established. Since 2012, ZIB has hosted this Service Center for Digitization Berlin, in short digiS. The tasks of digiS start with a sustainable definition of digitization. digiS networks the individual projects of the support programme, provides consulting services for the implementation of digitization and aims at enabling institutions to play an active role in their digital practice. The projects supported represent a cross-section of the cultural landscape in Berlin. In total, 17 projects in 11 institutions have been supported by 2013, mainly in the area of museums. A survey of the projects can be found on the digiS Website: www.servicestelle-digitalisierung.de. All projects within the Berlin support programme aim at the preservation of the stock just as well as at the creation of new entries to the collections. Another goal is to pave the way into the large national and European cultural portals such as the German Digital Library and the Europeana.

COMMUNITY AND CAPACITY BUILDING

DIGIS, KOBV, MUSEUM PROJECT AT ZIB

If you understand digitization merely as photographing or scanning of an object or the coding of an audio recording, you are aiming too short. The process starts much earlier with a comprehensive description of the formal aspects and the content of the objects. This is the only way to capture contexts and keep an overview over collections containing millions of objects. Inventory-taking or indexing, and thus the capture of metadata, precede and accompany every digitization project. In institutions of cultural heritage, this process is largely standardized. However, there is no obligation to apply these standards homogeneously. This is where the longest-standing project at ZIB has its scope of functions. The museum project develops and maintains comprehensive software for the indexing and administration of museum collections and stocks (GOS). Furthermore, the working group is doing research in the area of data modelling and data exchange. While the digital wave has demanded new structures for museums only recently, libraries, to stay with the metaphor, have already become experienced surfers. Electronic search in catalogues, defined metadata schemes for indexing, administration of electronic journals, ebooks and multimedia access portals are concepts with high penetration in the library world. The Cooperative Library Network Berlin-Brandenburg (KOBV), which has been located at ZIB since 1997, supports its more than 80 member libraries by rendering development work and services in the area of indexing, presentation and search systems for libraries. Since the Library Congress 2013 in Leipzig at the latest, German library networks are also discussing the administration and long-term preservation of research data. Since 2011 at KOBV, research data management beyond classical biographical description of library catalogues has been investigated in the project EWIG; starting in 2014, a distributed Humanities Data Center (HDC) will be developed.

Berliner Kul

turleben der 60er, Fotos: Harry Croner, © Stiftung Stadtmuseum Berlin

'BERLINER APPELL 2013' – LONG-TERM AVAILABILITY AS A COOPERATIVE TASK

The joint central research task for all projects within the Scientific Information Department (SIS), is the actual safeguarding of long-term availability of data. By the support programme of the Federal State of Berlin, Service Center for Digitization Berlin (digiS) has been given the mandate to ensure the long-term availability of digital reproductions. In September 2013, ZIB signed the so-called "Berliner Appell", which calls for joint reflection about the future of our digital society and the preservation of our digital cultural heritage. Since 2013, ZIB has also been a regular member in the Nestor Network of Expertise, thereby actively taking part in the shaping of national structures in this area.

LONG-TERM PRESERVATION STRATEGIES AT ZIB

Since more and more of the output of art and science is born digital, the access to it becomes more and more fragile. For rendering contemporary digital cultural manifestations, a complex infrastructure is needed: Starting with the description of the data formats, via the software needed to interpret them up to the hardware running this software. If you are lucky to find the remnants of an ancient culture written in stone, all you need is an understanding of the language. The rendering process is no problem, since eyesight and sunlight are more or less stable for the next millennia. But imagine someone finding an eighty-year old 5.1/4" floppy disk. Given that the magnetization is still intact; do you have a reading device for it? Are you able to determine the file format? Do you happen to have a version of the word processor or worse the database application used to create it? The answer to all that questions is most probably: no.

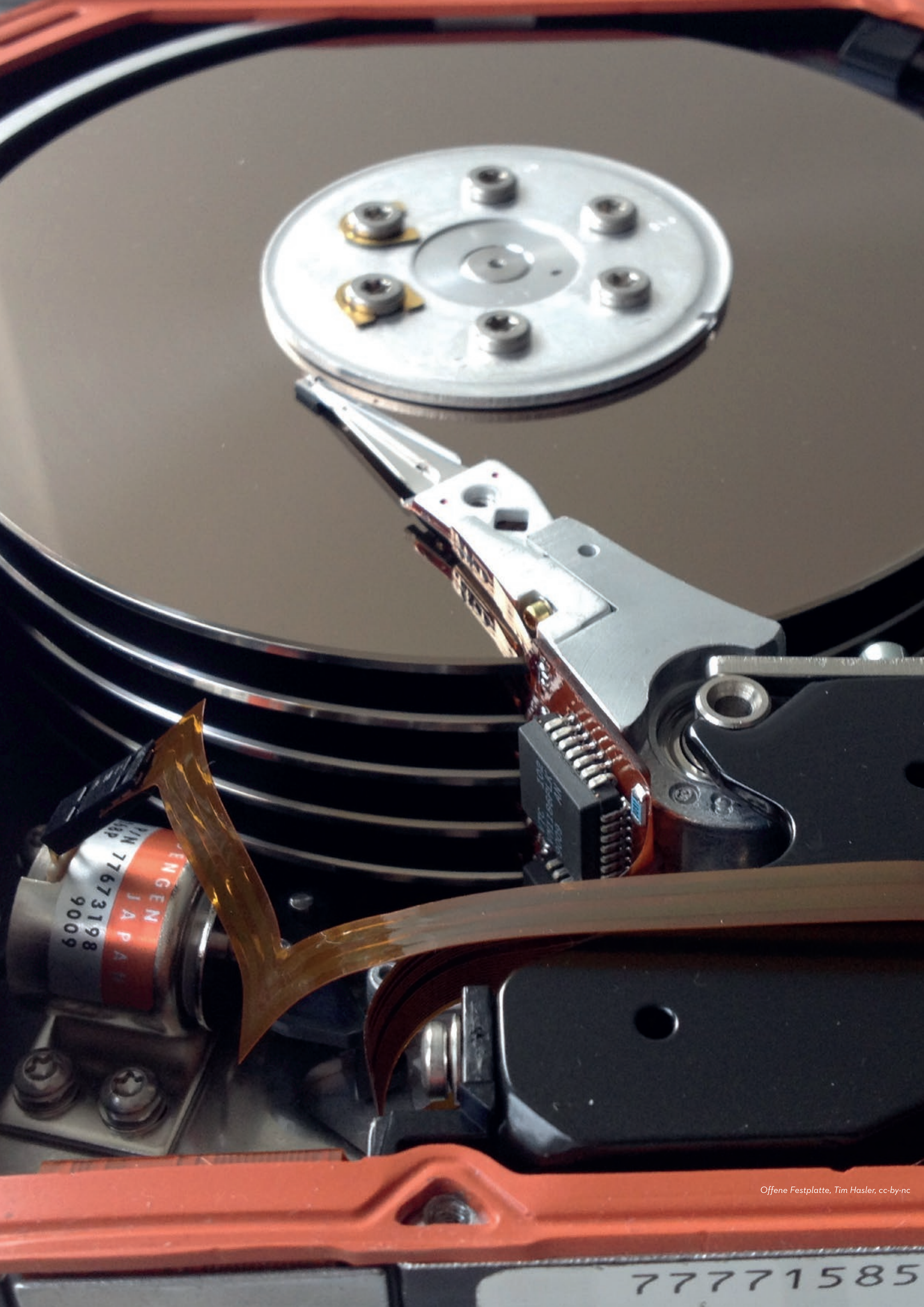
For digital objects to stay usable in the future, you actually do have to preserve them as intellectual entities in their complex context, but change them, if needed, on a technological level in order to guarantee their long-term accessibility. Digital preservation is an actively maintained process in the presence ensuring the access to digital objects in the future. There are still lots of –analogue as well as digital– recordings, photographs and measurements that are stored without proper care. In geosciences there is even an annual data rescue award, "to improve preservation and access of research data, particularly of dark data, and share the varied ways that these data are being processed, stored, and used."

To perform a proper digital preservation, you have to be able to understand the objects you are curating. This is only possible in the future if some criteria are met: One is the open availability of the format specifications, the other is the lack of any patent or other right restraining the free use of the standard. These two characteristics are not as widespread as we would like. Some file formats have been patented by the inventors. So, if you want to deal with that format in your software, you have to pay a license fee. Even worse, you could be forced to work with proprietary parts of vendor-provided software. Open Standards is a key issue in long-term preservation. That mindset becomes more common with the open access and open knowledge movement.

The Wikimedia Foundation for instance describes its mission to: "empower and engage people around the world to collect and develop educational content under a free license or in the public domain, and to disseminate it effectively and globally".

Another actor in the field, the Open Knowledge Foundation, states its vision as: "We believe that a vibrant open knowledge commons will empower citizens and enable fair and sustainable societies." What is Open Knowledge? 'Open knowledge' is any content, information or data that people are free to use, reuse and redistribute – without any legal, technological or social restriction.

In that line of thought belongs also the Free and Open Source Software (FOSS) Movement, which spawned not only some of the software that is today's backbone of the Internet like the Apache Web Server, but also a vast amount of tools used in the long-term preservation area. Some of these tools are used in a test bed installation of an OAIS compliant archive at ZIB. ZIB itself has more than 15 years of experience in bitstream preservation, which is nowadays more often dubbed 'passive preservation' compared to the actively maintained process described in the OAIS model.



Offene Festplatte, Tim Hasler, cc-by-nc

OPEN SOURCE IN LONG-TERM PRESERVATION

Building on the extremely robust storage infrastructure at ZIB, a long-term archiving system with different service levels to preserve technical availability as well as availability in terms of content also for heterogeneous data, is to be developed. The individual service levels are directly dependent on the requirements of the various partners (institutions of cultural heritage and research institutions). In this way, the previous approach of a "Swiss bank deposit box" will be complemented successively by standards-compliant services for long-term archiving.

First tests with the open-source system Archivematica have convinced particularly because the business logics of the OAIS model (see above) can be mapped by means of extremely transparent micro services that can be modelled on a granular level. In this way, diversity of data as well as partner-specific requirements for individual long-term availability / curation levels can be taken into account. In its functionalities and services, Archivematica does not orient itself towards the implementation of a complex overall system, but institute-specific preservation policies and the tasks and goals embedded there. Jointly with the Yale University Library, MoMA, the University of British Columbia and the Rockefeller Archive Center, ZIB is engaged in the further development of Archivematica.

DIGITAL HUMANITIES AND HUMANITIES DATA CENTER

The term Digital Humanities originally just referred to computer-assisted methods in the humanities. The increased opening of digital data sources, the world-wide networking of these resources, and the possibility to process huge amounts of data (keyword: Big Data) have widened the concept of Digital Humanities in the recent years.

Digital research data infrastructures have been present for more than 50 years, e.g., in the geosciences with their World Data System. So far, most of the humanities neither have research data infrastructures nor the resources to run them on their own, although there is a growing need for quality-assured long-term archiving, re-use of and online access to digital research data, too. Some digital research collections for the humanities are already maintained by institutions such as DNB, DDB, Europeana or IANUS. They, however, cover only a fraction of the relevant research data.

The purpose of the Humanities Data Center granted in 2013 (and funded by the Niedersächsisches Ministerium für Wissenschaft und Kultur for the period 2014 -2016) is to assume responsibility for data hitherto not available in sufficient digital quality for research purposes and for research data from branches of the humanities currently having no corresponding long-term repository. The partners of ZIB in this project are: Berlin-Brandenburgische Akademie der Wissenschaften, Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen, Akademie der Wissenschaften zu Göttingen, Max- Planck-Institut zur Erforschung multireligiöser und multiethnischer Gesellschaften Göttingen, Niedersächsische Staats- und Universitätsbibliothek Göttingen.

Bandroboter, Jürgen Keiper, cc-by-nc

Zersetzter Film in Dose, Jürgen Keiper, cc-by-nc

EWIG

EWIG stands for "Entwicklung von Workflowkomponenten für die Langzeitarchivierung von Forschungsdaten in den Geowissenschaften" (Development of workflow components for the long-term archiving of research data in geosciences). The project funded by the DFG aims at supporting the transfer of research data from different research environments into digital long-term archives. For this purpose, EWIG has identified three focal points. On the one hand, the development of rules for the handling of research data that should be generic in the beginning and ideally be accepted in the entire institution later, so called policies. On the other hand, exemplary workflows are developed, which enable the transfer of research data to an OAIS-compliant long-term archive (see ISO Standard 14721:2012) on the basis of Free and Open Source Software (FOSS) components. The third aspect addresses topic-specific qualification. In this context, 'archived' research data are analyzed in a manner of speaking as an archaeological item in geoscientific teaching. The student learns what type of metadata is relevant and required for the interpretation of preserved data. This is accompanied by a handout to teachers for teaching in the field of research data management.

Project partners: ZIB, Helmholtz-Zentrum Deutsches GeoForschungsZentrum Potsdam and Institut für Meteorologie der Freien Universität Berlin. Funding: DFG. Duration: 2011-2014.

ACCESS TO CONTENT

In order to enable comfortable access to saved data, the use of a repository middleware is required. The open-source software Fedora Commons and Islandora as potential Web front end are currently being evaluated with regard to search and access to the data of the long-term archive. Fedora is to be used as repository for the data storage of derivatives and/or the associated metadata and should enable the continued right and role-based access to these data as well as the master files in the long-term archive – also in the form of data harvesting via standardized and open protocols, such as the Protocol for Metadata Harvesting by the Open Archives Initiative (OAI-PMH).

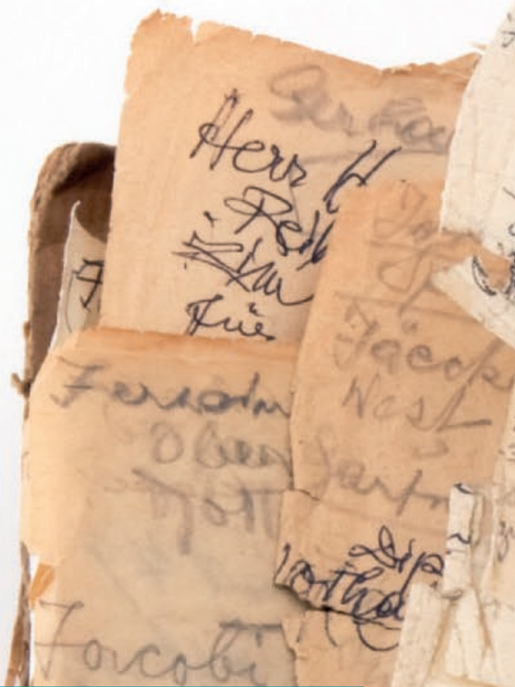
Thus, not only the institutions themselves, but also the broad public has transparent access to these data. This is a significantly more differentiated use of data, not only for science, but also, e.g., in the area of education or the pursuit of personal interests. Furthermore, contextualization and thus enrichment of data is made possible, thus creating additional added value [38]. For a permanent, unique identification of these research data ZIB has become a data center in the context of the DataCite DOI scheme. Based on a contract with the TIB Hannover, ZIB can mint its own DOIs. Due to the handle system underlying the DOI concept, the URL stored for a representation may change, the object, however, remains reliably citable and identifiable by means of the DOI - as long as the URL is adjusted, if a change has been made. Thus, a unique and permanent public referencing of these objects is possible for users

For the future it is planned that ZIB not only assures the long-term preservation of (research) data from cooperation partners of digiS and from projects within the support programme, but also for data from member libraries of the KOBV as well as associated research partners. Since, as described above, in the sense of the 'Open Science' paradigms the data records underlying publications are moving into the focus of research, archiving these data records is becoming indispensable.

Today ensuring the long-term availability of our digital culture and research is no longer on the political agenda of Berlin alone, but also throughout Germany. In Berlin, with the support of ZIB a start has been made towards dedicating this topic the social attention and publicity that is appropriate and necessary in view of its relevance.

The project-related support of individual projects defined in the support programme of the Federal State of Berlin in conjunction with the accompanying development of an organizational and technical infrastructure at ZIB, takes digitization / long-term preservation out of its niche that has become too narrow and turns it into a task that has to be dealt with strategically at ZIB and beyond.

With its existing infrastructure and services as well as due to its networking capabilities and outstanding cooperation among different projects and institutions (digiS, KOBV, museum project, ITS), ZIB offers excellent opportunities to tackle this not only technical but also social task. A start has been made – you learn on the way.



MUSUEM PROJECT

The project “Development of information technology tools for museums” has been in place since 1981 and is therefore older than ZIB itself. Its tasks include on the one hand the IT-technical support of inventory and documentation projects in museums and on the other hand the communication and distribution of internationally recognized standards within the German cultural heritage institutions.

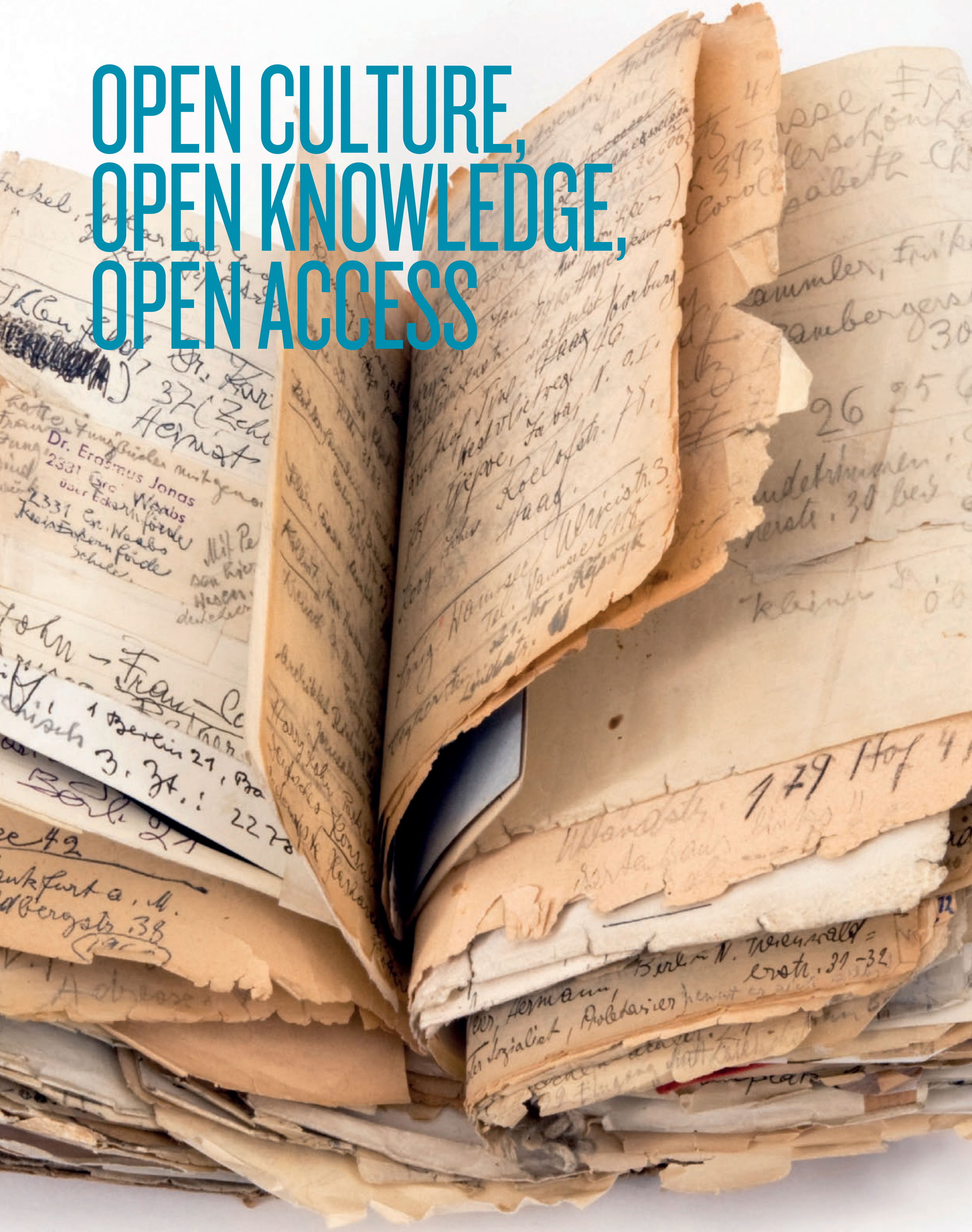
SOFTWARE

The basic software for all projects is the GOS database system, which was originally developed by the British Museum Documentation Association (MDA, in the meantime renamed to Collection Trust) especially for use in inventory and documentation in museums. Since 1981 it has been further developed by ZIB and is made available to interested museums.

Due to close contacts to decision-making bodies, e.g. the “Fachgruppe Dokumentation” in the Deutscher Museumsbund and the CIDOC (Comité international pour la documentation des International Council of Museums), ZIB is now well established in the field of museum documentation. The autumn meeting on museum documentation, which each year is jointly organized by the Fachgruppe Dokumentation and the Institut für Museumsforschung, has traditionally taken place at ZIB for more than 15 years now.

Members of the museum project participate in the “data exchange” working group, which was founded by the Fachgruppe Dokumentation. The working group is involved in the principles of structuring data required for the inventory of museum objects and examines (metadata) formats with respect to their usability for the needs of the different collections. The basis for this is the “Conceptual Reference Model”10 (CRM), which has been developed by the CIDOC. The working group is also active in maintaining the data harvesting format “LIDO” (http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido/), which has become the standard format for many museum portals.

OPEN CULTURE, OPEN KNOWLEDGE, OPEN ACCESS



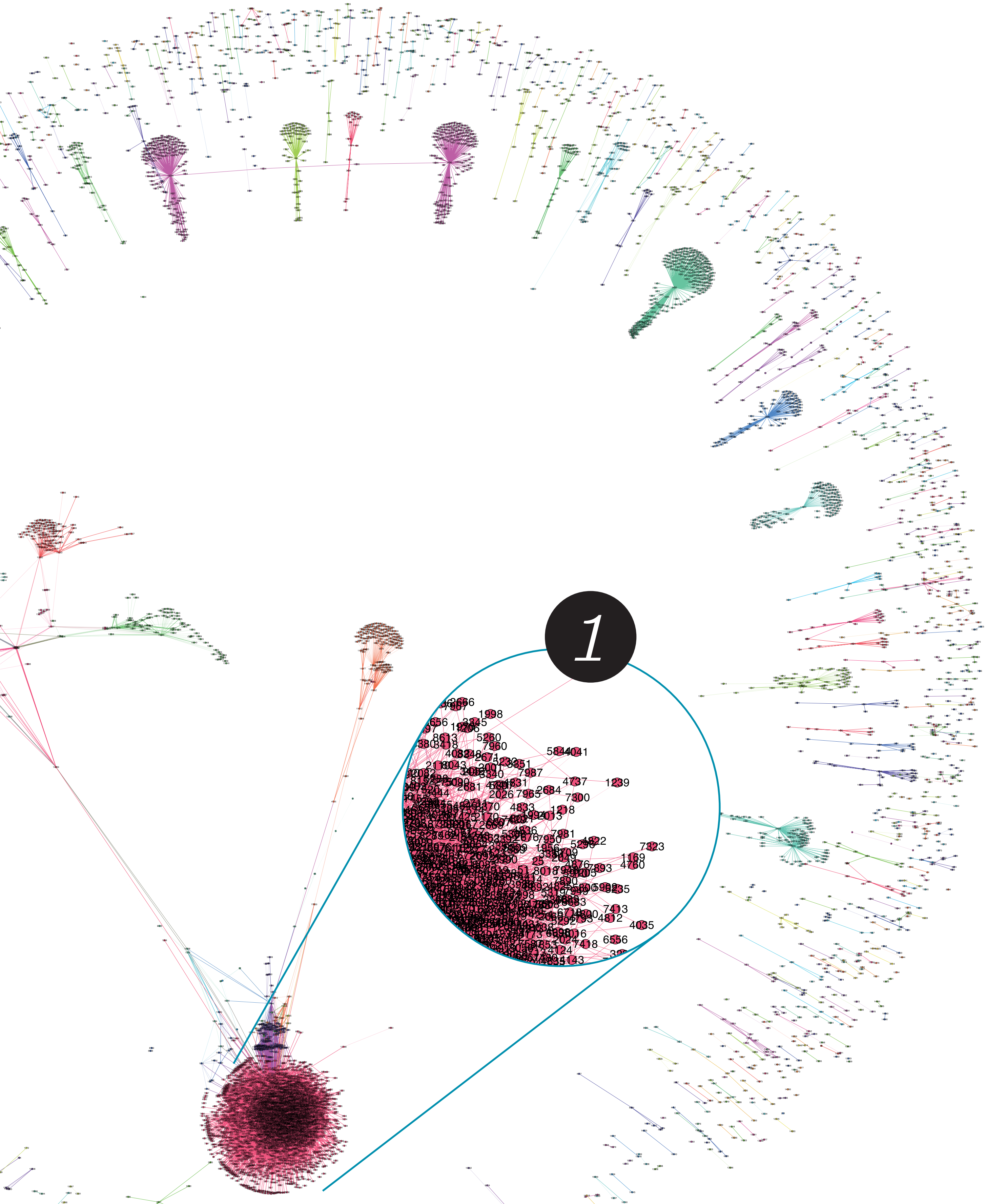
Adressbuch Hannah Höch, Foto: Kai Annett Becker, © Berlinische Galerie

FOR FURTHER INFORMATION PLEASE CONTACT:
Patrick Schäfer | patrick.schaefer@zib.de | +49 30 84185-168

SIMILARITY SEARCH AND INSECT CLASSIFICATION

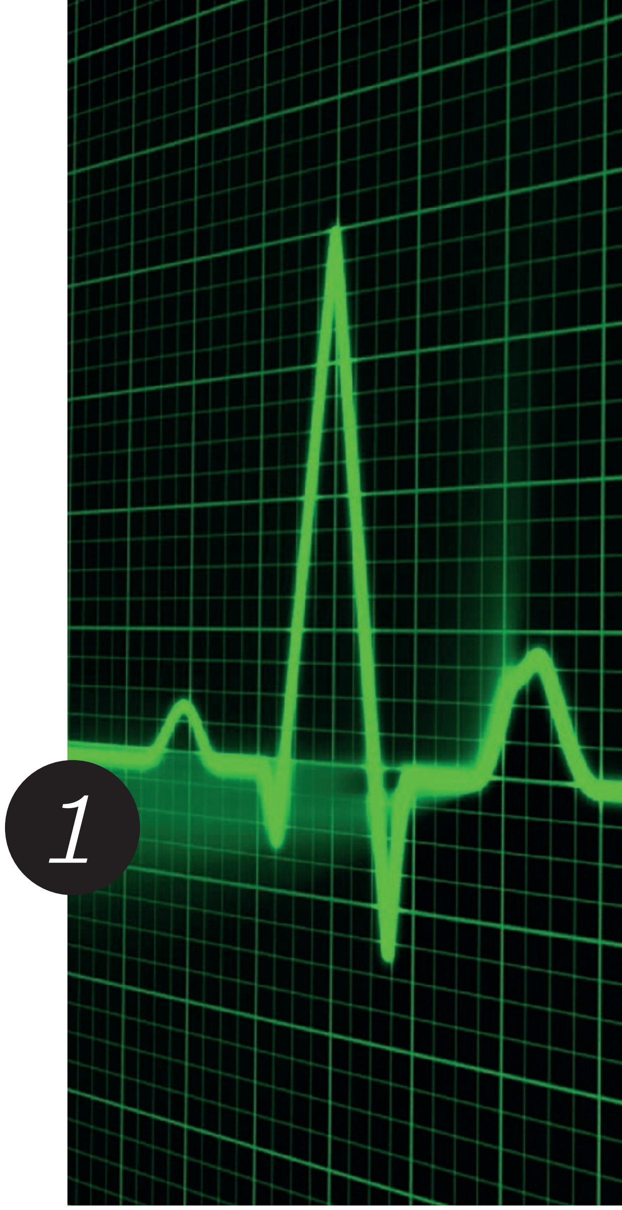
Emulating the human intuition regarding similarity

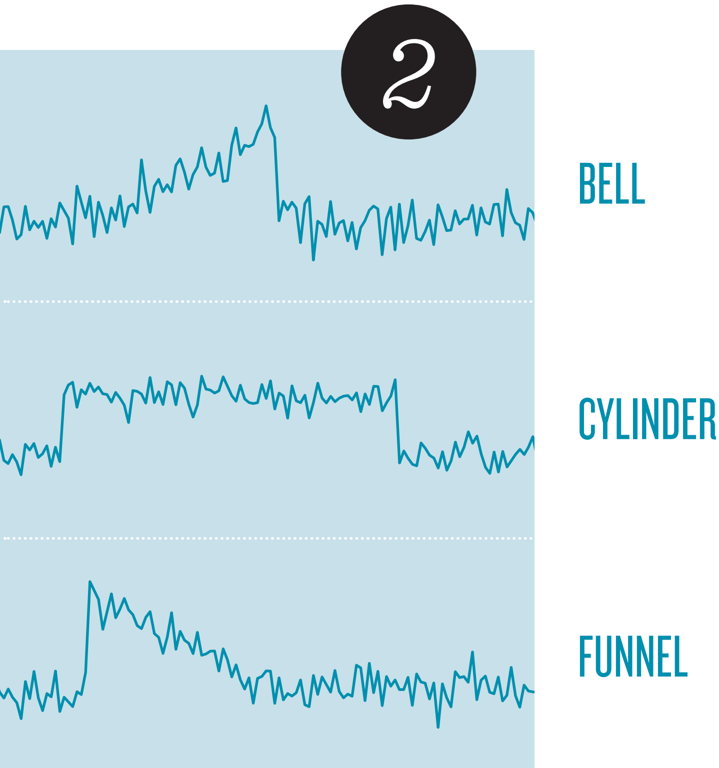
While humans have an intuitive understanding of the similarity of two objects, this task becomes complex for a computer. Apart from designing a meaningful similarity metric, scalability remains a main challenge. Use cases of similarity search range from time series indexing or classification to insect wing-beat detection.



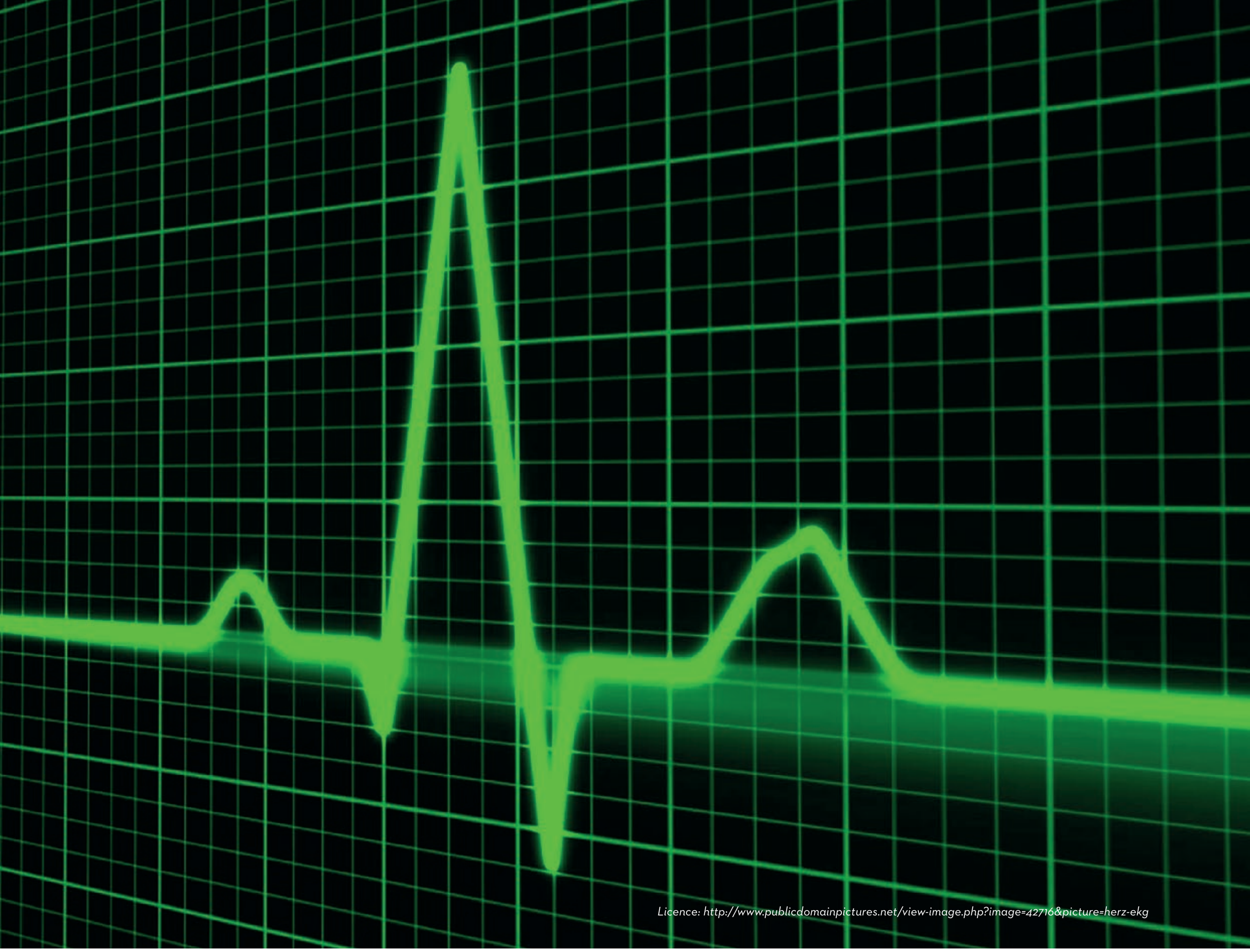
1 A time series can be represented by a point in high dimensional space. Searching close points leads to finding similar time series.

TIME SERIES SIMILARITY





Time series are recorded from sensors and other input sources over time. Application domains include personalized medicine, human walking motions, anthropology, security, historical documents, astronomy, or spectrographs. For data analysis, the input data must be cleaned by eliminating erroneous, extraneous, and repetitive substructures. The signals from Electrocardiography (ECG) recordings, for example, are highly redundant and they may contain noisy or extraneous data before or after the sensor was connected to the patient.



While a human has an intuitive understanding of the similarity of objects, this task becomes very complex for a computer. A similarity measure is commonly used to compare two time series which qualifies similar objects by a small number and dissimilar objects by a large number. The most simple similarity measures include the Euclidean Distance (ED) or Dynamic Time Warping (DTW). Despite the fact that DTW is several decades old, it has shown to be highly competitive and is up to date used as a benchmark. ED and DTW have some known short comings. The former fails to match time series if they contain reordered substructures, like heart beats in an ECG recorded in a different ordering. The latter provides a peak-to-peak and valley-to-valley alignment of two time series, which fails if there are a variable number of peaks and valleys, e.g. a different number of heart beats in an ECG recording.

A toy example contains three basic shapes: bells, cylinders and funnels, whereas each shape is distorted by noise, see Fig. 2. For the human eye it is easy to distinguish between these curve types. Given a large number of these shapes both ED and DTW fail to capture the dissimilarity of the bell and funnel curves correctly. This toy example gives an impression on the difficulties that may arise with respect to time series similarity. In general, several sources of invariance like amplitude/offset, warping, uniform scaling, occlusion, and complexity have been presented. The three shapes require invariance to phase (horizontal alignment), warping (local scaling), occlusion (noise) and amplitude/offset. State-of-the-art algorithms that provide subsets of these invariances are based on structure-based similarity. That is two time series are compared based on their higher level structures like reoccurring or characteristic patterns.

Interestingly, while structure-based similarity measures provide very good results on some datasets, they are not more accurate than DTW in general. We believe that one reason for this is that invariance to noise was paid little attention to and the similarity is often estimated based on raw and noisy time series data.

1 Electrocardiography (ECG).
2 Basic shapes of curves.

1 The query (top) is cut into windows, and an SFA representation (bottom) is calculated using an alphabet of size 26 (»a« to »z«).

FROM REAL VALUES TO WORDS

Our Symbolic Fourier Approximation (SFA) [39, 40, 41, 42] is a symbolic representation of time series. SFA is composed of two operations, namely low pass filtering and quantization. These aim at reducing noise from a time series.

SFA is a symbolic representation. That is, the SFA transformation of a real valued time series results in a lower dimensional character string, named SFA word, using an alphabet of symbols. The figure 1 top represents the wing-flapping of a mosquito. These were recorded by a sensor that measures the fluctuations of light received by a photoreceptor as an insect flies through a laser beam and partially occludes light with its flapping. The main source of misclassifications results almost completely from species that are very similar in size and morphology: aedes aegypti (yellow fever mosquito) male can be misclassified to culex tarsalis (mosquito) male but never as Apis mellifera (common bee), which is a much larger insect compared to mosquitoes. Figure 1 center illustrates the SFA words over the subsequences of a time series. Figure 1 bottom gives an impression of the inverse transformation of an SFA word which represents an area. In the context of insect classification SFA smoothens out anatomic differences of animals of the same species or other sources of misclassification like different entrance angles or speed of the insect when passing the photoreceptor.

In general SFA has two parameters: the size of the alphabet and the word length. These have the following effects:

➡ Low pass filtering: Rapidly changing sections of a time series are often associated with noise. This can be reduced by a low pass filter. The SFA word length determines the bandwidth of the low pass filter.

➡ String representation: SFA makes use of quantization and character strings. Thereby it allows for string matching algorithms to be applied. The size of the alphabet determines the degree of quantization, which has an additional noise reducing effect.

SFA is composed of two operations:

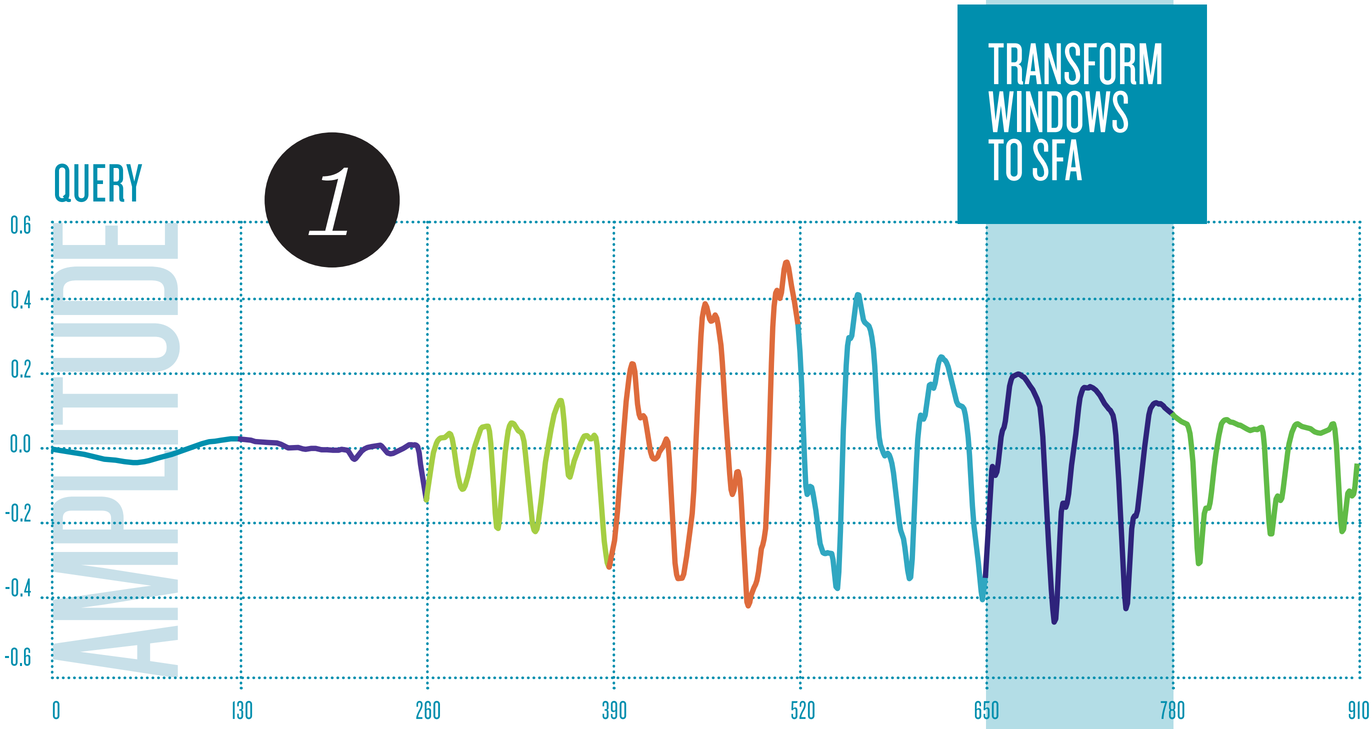
1. Approximation using the Discrete Fourier Transform (DFT).
2. Quantization using a technique called Multiple Coefficient Binning (MCB).

Approximation aims at representing a signal by a transformed signal of reduced length. Higher order Fourier coefficients represent rapid changes like dropouts or noise in a time series. This time series is low pass filtered by using only the first few Fourier coefficients. Quantization adds to the noise reduction by dividing the frequency domain into frequency ranges and mapping each real valued Fourier coefficient to its range. Our quantization scheme MCB is data adaptive and thereby minimizes the loss of information introduced by the quantization.

CASE STUDIES USING SFA

SFA was applied in the context of:

1. Time series similarity search [39],
2. Time series classification [40],
3. Insects wing-beat classification [41,42].



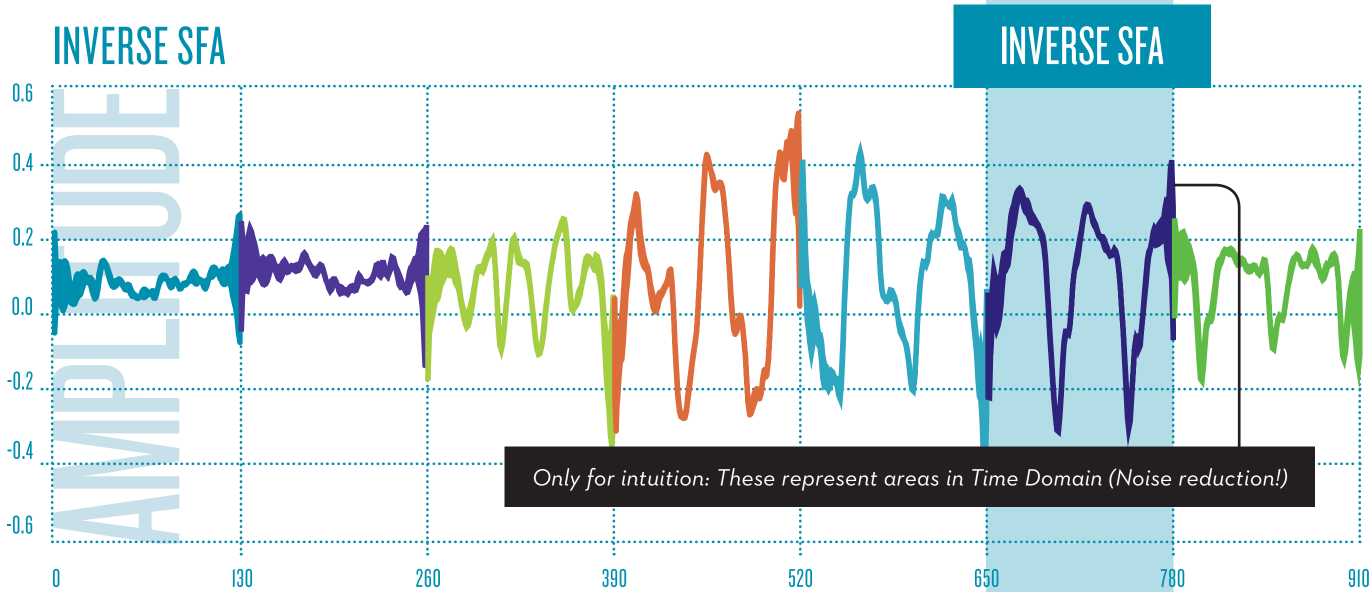
SFA WORDS

SFA

0–130	130–260	260–390	390–520	520–650	650–780	780–910
nmmlmnlmjm	nmllmnkmjm	nkmkmmjmjh	npmtkirlkg	nhngtqingk	nnknhkqmlj	nkmknpjnim
lpnplonnlm	lpnolnnnlm	kqmpknnokl	moppmnpnnn	jqmplqjokn	kqoqmlnnmn	ipnolopnmm
pnomniokpk	onomniokoj	noommiolol	sloloipiqj	oononkjkmk	pmpkoinjqj	pnnmmilknj
okommllkim	nkommmlkhm	nkpmmlklin	pltloinjkl	mokokmimfp	pgqknjljlk	mmqnmomkhm
jolpmnmnnn	ipkqmnlolo	iolqmokomo	lnpoplompk	fqirkpgpkr	lmopomonpk	iolqmolomo
qjmnqnjqog	pklopnirmh	pklopojrnh	tiqltkmnse	mmirnpgukj	siplsllore	pkloqnjqnh
lonnn	komnm	kpmom	olrjp	hriqj	nmqko	komnm

This is SFA: Words in Frequency Domain

0 130 260 390 520 650 780 910



TIME SERIES SIMILARITY SEARCH

A problem of time series similarity search is that the amount of data can be massive. For example, 1 hour of an ECG log sums up to 1 GB, or the sensors of a space shuttle record up to 200GB for each start. To avoid sequentially scanning the data for each query, an index structure must be used. A time series consisting of n measured values can be seen as a point in n-dimensional space, where the i-th measured value represents the i-th dimension. By indexing this n-dimensional point using a spatial index, the problem of finding similar time series is reduced to finding nearby points in n-dimensional space using a similarity metric as introduced before.

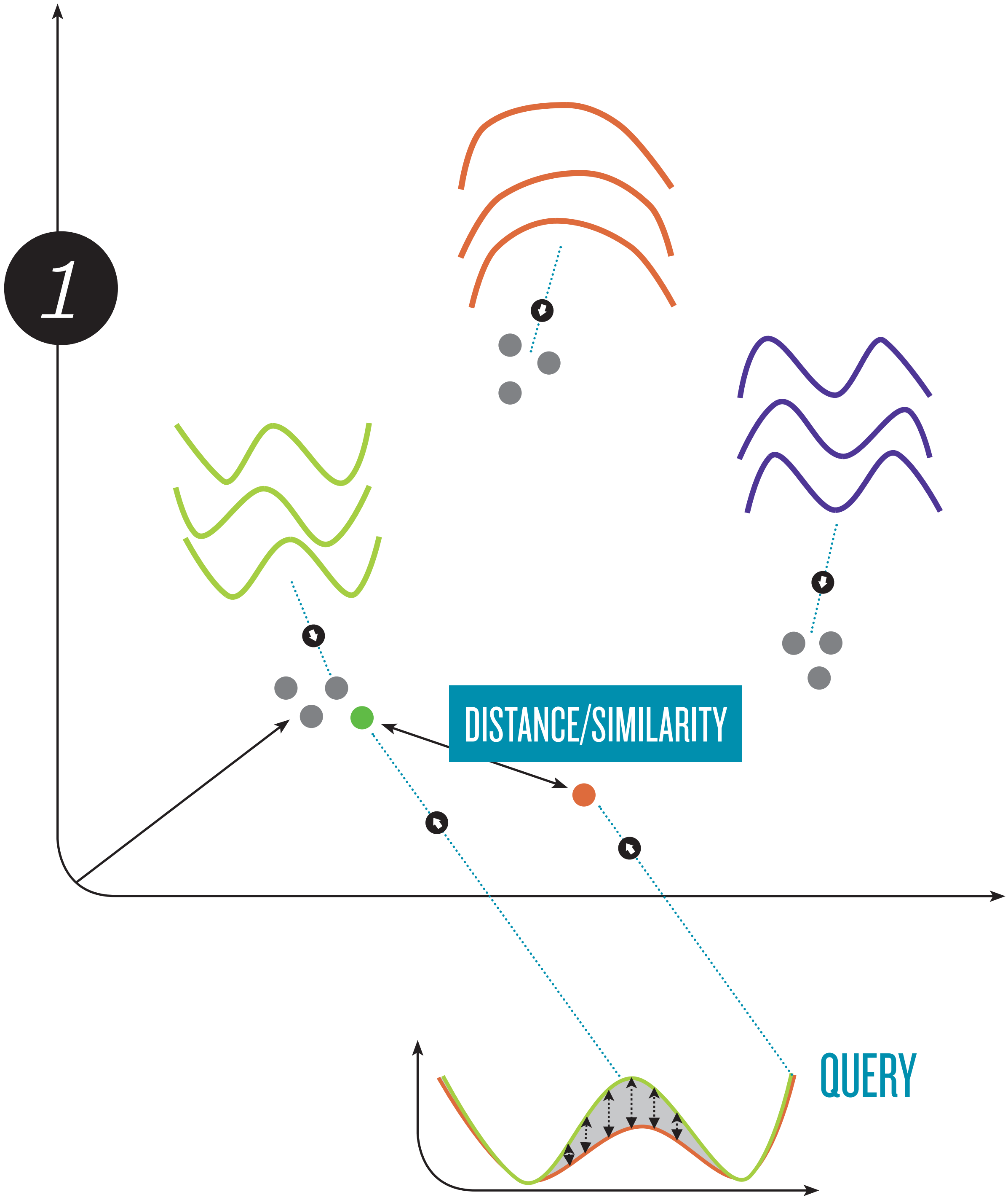
Indexing high dimensional data is a considerable challenge, as spatial index structures, like the R-Tree suffer from a phenomenon called the Curse of Dimensionality: with increasing dimensionality of the search space, the performance of similarity based queries on the index becomes worse than a linear scan of all data. Spatial index structures usually degenerate with 10-20 dimensions.

By using dimensionality reduction prior to indexing and a lower bounding distance measure, queries are guaranteed to return the exact same result in reduced dimensional search space as if they were executed in the original space. In order to reduce the dimensionality of a time series an approximation technique is applied, effectively reducing dimensionality of the time series by 10:1 to 50:1. By use of an approximation technique the Curse of Dimensionality is shifted to 10-20 indexable dimensions of the approximations. The problem with approximation is that information from the original time series is lost. A goal of approximation is to find representations of the original data, so that each representation is distinct. For example, similar time series will have the same representation after an approximation if the reduced dimensionality is too low.

AN INDEX STRUCTURE FOR TIME SERIES

The purpose of indexing is to partition the search space into equal-sized groups containing similar entries and to provide a fast lookup for these entries. The idea of indexing SFA is simple: grouping is done based on the first k characters of the SFA words. Increasing k leads to smaller sized groups. As the entries are not equally distributed in the search space, k is individually increased until all entries are equally distributed over each group. This means that SFA is best computed at a high word length, but only a few groups exploit all symbols of the words.

In essence, our SFA try is a prefix tree built over the prefixes of the SFA words of the time series. In the SFA trie, each node represents all time series sharing the same prefix of an SFA word. The novelty of the SFA trie is that it adapts to dense leaf nodes by increasing the length of the prefix, which equates to improving the quality of the approximation without altering all the other leaf nodes. This increases accuracy of the similarity search, i.e. less leaf nodes have to be searched for answering a similarity query.



SFA in combination with the SFA trie build an exact method for time series similarity search. They offer a very low memory footprint due to quantization applied on top of approximation. By utilizing our variable length symbolic representation, we were able to index and search terabyte sized time series datasets with commodity hardware.

1 The similarity of time series corresponds to closeness of points in n-dimensional space.

TIME SERIES CLASSIFICATION

Designing a time series classification algorithm is a challenging task. It is non-trivial to extract a statistical model from time series as these may be non-stationary, and show varying statistical properties with time. Data mining algorithms, on the other hand, degenerate due to the high dimensionality of the time series and noise. Existing techniques for time series classification can be categorized into shape-based and structure-based. Shape-based techniques use a similarity measure in combination with 1-nearest-neighbor search. These are very competitive on cleaned datasets but fail on long or noisy data. Structure-based techniques transform a time series into a different representation or extract characteristic features from the time series like reoccurring, characteristic patterns. This comes at a higher computational cost when compared to shape-based approaches.

We looked at case studies covering human walking motions, personalized medicine, anthropology, historical documents, mass spectrography and security. Let's have a closer look at walking motions of four different subjects. Each motion was categorized by the labels normal walk and abnormal walk. The data were captured by recording the z-axis accelerometer values of either the right or the left toe. The difficulties in this dataset result from variable-length gait cycles, gait styles and pace due to different subjects throughout different activities including stops and turns.

1

1 Human walking motions are recorded by an accelerometer.

Approaches based on similarity measures like the Euclidean Distance or Dynamic Time Warping fail to identify the abnormal walking styles. Our Bag-of-SFA-Symbols (BOSS) model combines the extraction of substructures with the tolerance to extraneous and erroneous data using the noise reducing SFA representation of the time series. The similarity of two time series is then estimated based on the differences in the set of noise reduced patterns. As a result the separation of the normal walking motions from the abnormal walking motions is much clearer with just one walking motion being misclassified. As opposed to rivaling methods our approach offers multiple advantages: (a) it is robust to noise, (b) it allows for fast data analytics by use of string matching algorithms, (c) it is more accurate than state of the art time series classifiers, and (d) it provides invariance to the phase shifts, offsets and amplitudes and occlusions by comparing two time series based on their higher-level substructures.

ON CLASSIFYING INSECTS FROM THEIR WING-BEAT

Producers set up traps in the field that lure and capture pests, in order to detect and count them. Manual inspection of traps is a procedure that is both costly and error prone. Inspection must be carried out manually as a repeated process, sometimes in areas that are not easily accessible. These traps can also, and often do, accidentally capture other species of insects rendering the counting difficult especially for species that are morphologically similar. Pest surveillance, adaptable for other flying insect pests, can provide pest information at regional and national scales, giving authorities a powerful tool to understand at a higher level the impacts and risks imposed by the presence of these pests.

A novel approach to wing-flapping recording devices has been announced recently. The core idea behind these new sensors is to embed a photoreceptor in an insect trap to record the fluctuations of light as an insect flies through the laser beam and partially occludes light with its flapping. The sonification of the baseband signal produced by the light fluctuations represents normal sound that is subsequently recorded using audio recorders.

We put new work in context by providing algorithmic details associated with two different datasets that have been made publicly available by the University of California Riverside. The recorded insects are:

1. AEDES AEGYPTI
   yellow fever mosquito
2. ANOPHELES GAMBIAE
   mosquito
3. APIS MELLIFERA
   western honey bee
4. COTINIS MUTABILIS
   figeater beetle
5. CULEX QUINQUEFASCIATUS
   southern house mosquito
6. CULEX TARSALIS
   mosquito
7. CULEX STIGMATOSOMA
   mosquito
8. DROSOPHILA MELANOGASTER
   common fruit fly
9. FUNGUS GNAT
   short-lived flies
10. MUSCA DOMESTICA
    common housefly
11. PSYCHODIDAE DIPTERAL
    moth fly

❶ *A photoreceptor records the fluctuations of light as an insect flies through a laser beam and partially occludes light with its flapping.*

The main source of misclassifications results almost completely from species that are very similar in size and morphology, e.g. aedes aegypti male can be misclassified to culex tarsalis male but never as apis mellifera, which is a much larger insect compared to mosquitoes. Note also that mosquitoes are dimorphic with females being larger than males. Therefore, acoustic discrimination of sex is not as difficult as it might seem on first impression. In our experiments the most accurate approaches never confuse a culex quinquefasciatus male with a culex quinquefasciatus female.

An audio file can be reinterpreted as a time series. This connects bio-acoustics to time series analysis by the transformation of the process of finding similar audio files to finding similar time series. We built feature vectors based on SFA. Our approach reduces random noise and accounts for differences in insect movement by the use of quantization on top of the Fourier spectrum. SFA proved to be one of the top scoring approaches applied in the evaluation.


CRAY

SUPER COMPUTING AT ZIB

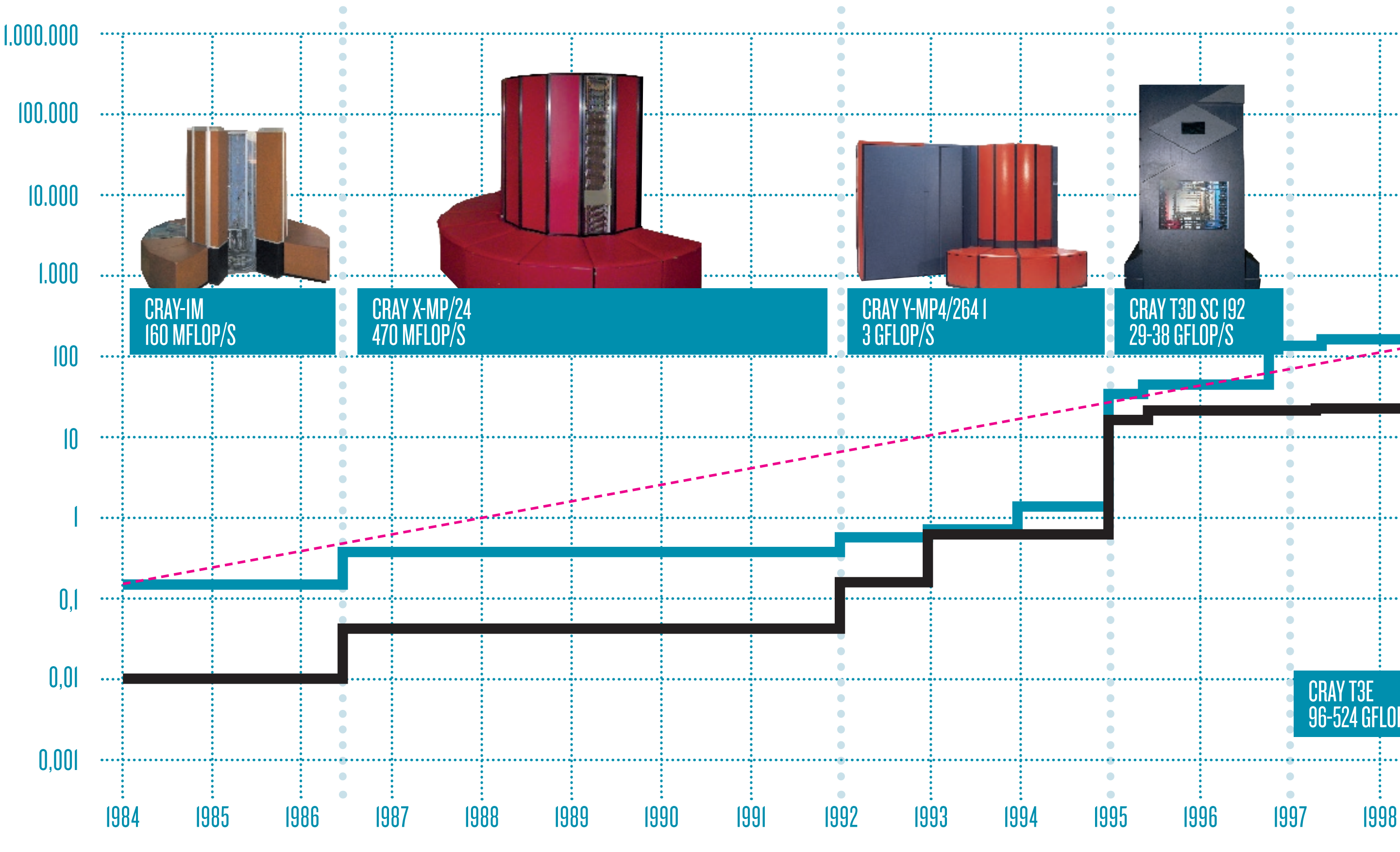
Solving complex problems

Our basic principle “Fast Algorithms – Fast Computers”, coined almost thirty years ago is still up-to-date. Today, large-scale computer simulations and optimization runs on the supercomputer resources at ZIB help researchers to solve complex problems in science, engineering, environment, and society.

FOR FURTHER INFORMATION PLEASE CONTACT:
Dr. Thomas Steinke | steinke@zib.de | +49 30 84185-144

SUPER-COMPUTERS AT ZIB: PAST TO PRESENT

Supercomputing at ZIB started in 1984 with the installation of the legendary Cray 1M system in Berlin. But it is not the hardware resources alone which make supercomputing feasible for various disciplines in sciences, engineering and other areas. The idea of HPC software consultants was born at ZIB and quickly taken up by other centers as well. Today, the HPC consultants at ZIB pursue their own domain-specific research in fields such as chemistry, engineering, earth sciences or bioinformatics while providing consultancy to external users of the supercomputers. Our basic principle "Fast Algorithms – Fast Computers", which was coined almost thirty years ago holds until today.

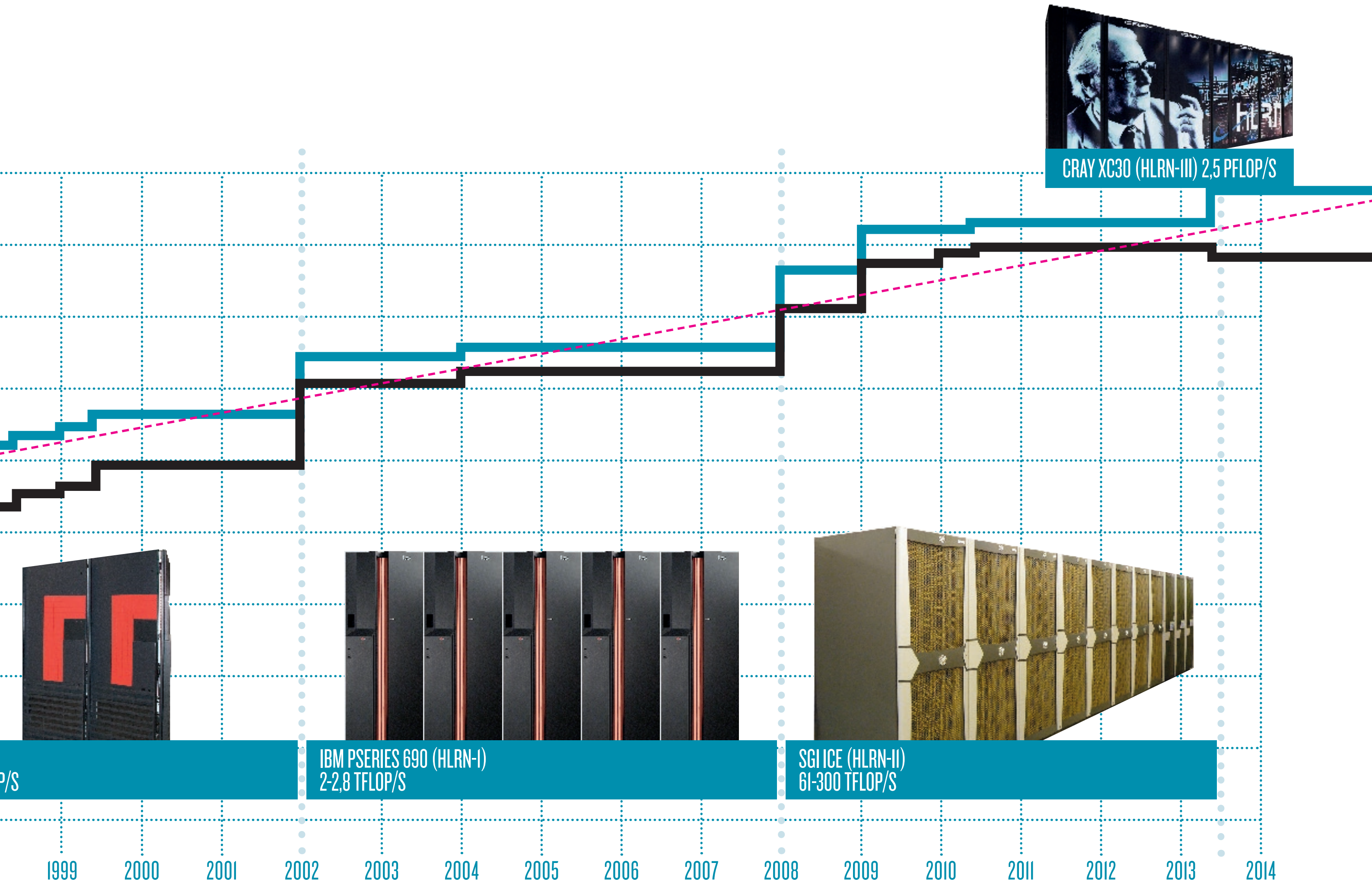


SUPERCOMPUTER INSTALLATIONS AT ZIB

The story of supercomputer installations at ZIB started in 1984. All systems up to the Cray T3E were financed solely by the State of Berlin, while the systems after 2002 were jointly financed by the HLRN member states and operated at the two sites ZIB/Berlin and LUIS/Hanover. The systems' peak performance (blue curve) and memory capacity (black curve) increased as compared to Moore's law (dashed line) shown in the Fig. below. Over the years, ZIB regularly made a system listed in the TOP500 list—despite the fact, that we did not focus on peak performance but rather on having well-balanced systems with powerful interconnects and sufficiently large main memories. Moreover, the supercomputers operated at ZIB do not comprise single monolithic systems, but rather a number of system components with complementary architectures for massively parallel processing (MPP) and shared-memory processing (SMP). The driving force on the selection of our system architectures has, of course, always been the user demand. Apart from the high investment costs, the sharp increase in the operation costs became of major burden for HPC service providers. Today, Germany is probably among the worst of all European countries with respect to energy prices. From 2013 this financial burden could no longer be carried by the two HLRN sites in Berlin and Hanover. Consequently the HLRN board decided to jointly finance the costs for system maintenance and energy, starting with the HLRN-III in 2013.

EVOLUTION OF SUPERCOMPUTER INSTALLATIONS AT ZIB

- Rpeak (GFlop/s)
- Memory (GBytes)
- Moores Law



THE HLRN-III PROCUREMENT: AN OPTIMIZATION PROBLEM

With the long and time-consuming application process for HPC resources in Germany, the HLRN consortium started the planning process for the HLRN-III in 2009/2010, shortly after the installation of the predecessor system. In spring 2011, the HLRN-III-proposal was submitted to the WR and DFG. After their positive decision, the HLRN alliance mandated ZIB to coordinate the Europe-wide procurement. All relevant HPC vendors were asked to participate. After several negotiation rounds Cray's bid was formally accepted and the competing bidders were informed. The contract was signed in December 2012 and the system was delivered in August 2013. From then everything progressed very fast: Cray needed only a little more than one week to install the XC30 from scratch at our premises. More time-consuming was the software installation on the various support servers (login, batch, data mover, firewalls, monitoring, accounting, etc.). Because of the complex dependencies of these servers, this difficult part of the system installation is most often underestimated. This is especially true for the HLRN system with its unique "single-system view" that provides a single user login, project management, and accounting for a system that is operated at two sites in Berlin and Hanover.

HOW TO SELECT THE RIGHT SUPERCOMPUTER?

This section is a short excursion on how we selected our HLRN-III supercomputer. Finding the perfect system that matches the requirements of their users, administrators and – last but not least – the capabilities of the providers, is far from an easy task. And yes – we found an almost "perfect" system architecture ...
We received proposals from many well-known HPC vendors on our European procurement (RFP). The selection process was primarily guided by the performance results of the HLRN-III Benchmark Suite. This benchmark suite aims to rate the overall compute and I/O performance of the two-site HLRN-III configuration over the operating time with two installation phases. Thus, the idea is to value the performance of different HLRN-III configurations during the negotiation stage at the end by a single score. The HLRN-III benchmark comprises the application benchmark as its major part and is complemented by an I/O benchmark section and the HPCC benchmark to gain insight into low-level characteristics of the system architecture. Each of these benchmark sections are differently weighted with the major impact by results of the application benchmark.

APPLICATION BENCHMARK

The HLRN-III has to serve a broad range of HPC applications across various scientific disciplines in the North-German science community. Therefore, and not for the first time, our RFP and subsequent decision-making process was based on a balanced set of important applications running on HLRN resources. The balance here refers to cover a representative spectrum of traditional and foreseeable major applications across the HLRN users community. The application benchmark is composed of eight program packages: BQCD, CP2K, PALM suite, OpenFOAM, NEMO, FESOM, and FRESCO. Workloads for the benchmark suite are defined by the HLRN benchmark team in close cooperation with major stakeholders from the different scientific fields. Reference benchmark results were obtained on the HLRN-II configuration with representative input data sets and possibly utilizing large portions of the HLRN-II resources.

PERFORMANCE RATING METHODOLOGY

Our goal was to maximize the work that can be carried out on the computer system over its total operation time. With the HLRN-III Benchmark Rating Procedure (described below) this overall serviceable work is calculated.

Our rating procedure is based on the fundamental assumption that the work which can be carried out by the system is equally split across the various user communities – the HLRN policy implies no bias towards a specific science and engineering field. Each of the stakeholders is ideally represented by one of the application benchmark. The achievable sustained performance of each of the application benchmarks determines the amount of work per community within the system's operating time. Together with the assumption described above the geometric mean of all application benchmarks gives an average performance of a system configuration – the key performance indicator – which then can be translated to the target work value. To allow for flexibility in terms of up-to-date technology the vendors were allowed to offer a second installation phase with a technology-update or complete hardware exchange. In that case, the vendor has an additional degree of freedom in choosing an optimal point in time for the second installation phase. Again, the optimization target was to maximize the overall work which can be delivered by the system – now a time-weighted sum of each key performance indicator per installation phase.

For each of the application benchmarks a minimum performance was defined in the RFP making sure that the new system shows no performance degradation for a specific application case. Thus, the minimum performance serves as a strong penalty in the evaluation. The application performance is measured on a fully loaded system with as many application instances as possible. Application performance is either determined inside the application itself (if simple operation counts can be captured), or application wall-times are converted to performance numbers by predefined operations counts.

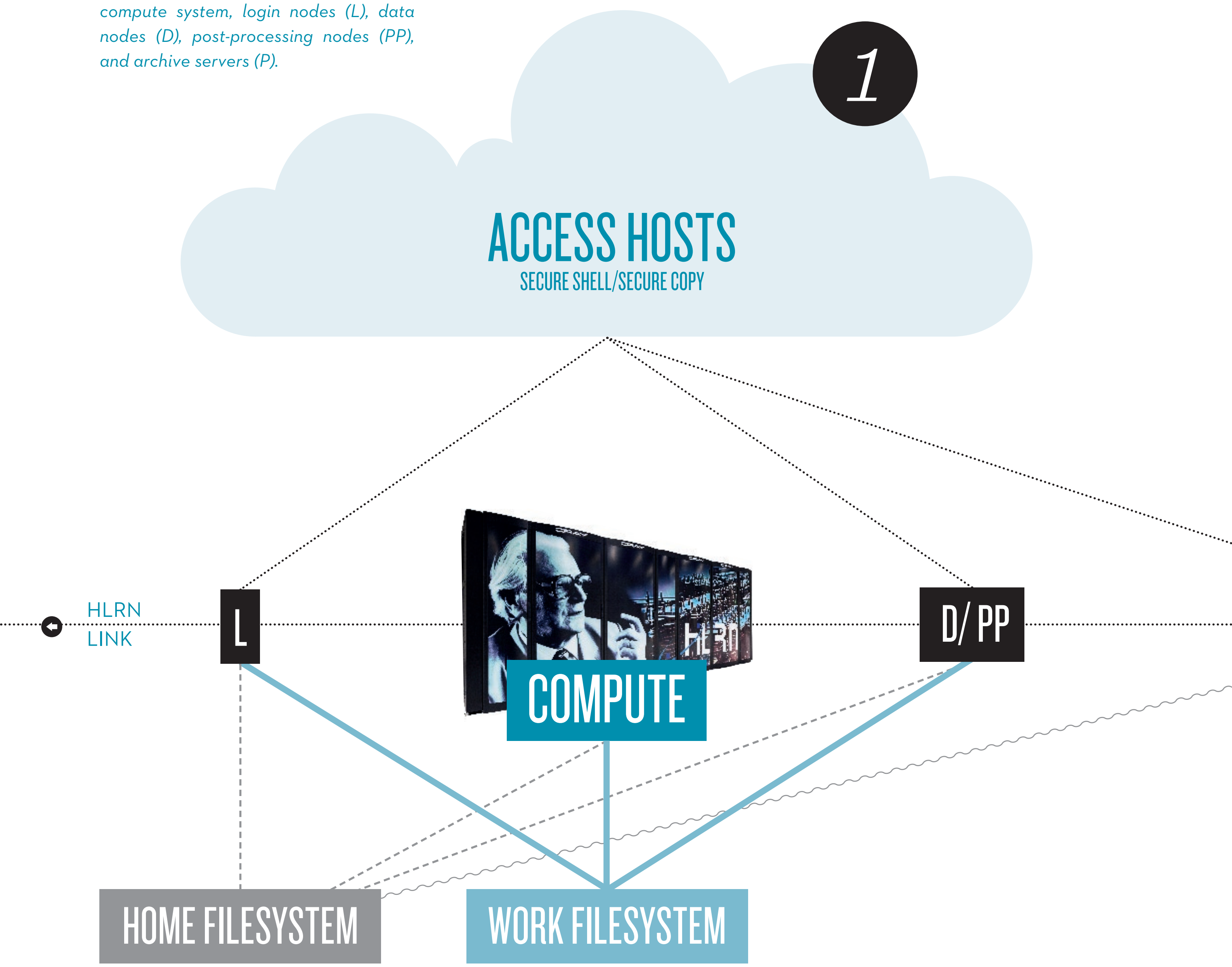
We did not want to leave the reader of this section without disclosing at least one benchmark number, that is the HPL performance. The two Cray XC30 systems of installation phase I are rated with a Linpack performance (Rmax) of 295.7 TFlop/s. With that, the two MPP components of HLRN-III were ranked at position 120 and 121, respectively, in the November 2013 issue of the TOP500 list.

I/O BENCHMARKS

The I/O benchmark suite measures bandwidth (IOR) and metadata performance (MDTEST) of the global parallel file system. In particular, the I/O benchmark suite is used (i) to measure the I/O bandwidth to the global file systems HOME and WORK, and (ii) to evaluate the metadata performance of the parallel file system, i. e., the achievable file creation, remove, and stat operations per time.

1 One of the compute cabinets of the Cray XC30 "Konrad" at ZIB (phase 1).

1 HLRN System Architecture per site: The two HLRN-III complexes with the compute system, login nodes (L), data nodes (D), post-processing nodes (PP), and archive servers (P).



HLRN APPLICATION SOFTWARE

ENGINEERING:
Abaqus | Ansys/CFX/Fluent | OpenFOAM | Star-CD | StarCCM+

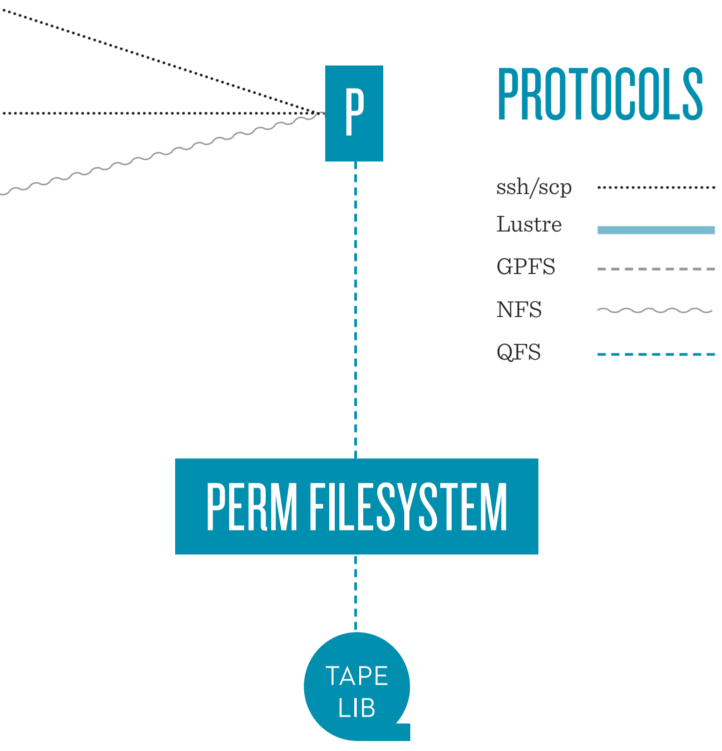
CHEMISTRY:
Abinit | CP2K | CPMD | Crystal | Desmond | Gamess-US | Gaussian | Gromacs | LAMMPS | NAMD | NWchem | Turbomole | VASP

NUMERICS:
Cray SciLib | Intel MKL | FFTW | PETSc | Scalapack/Blacs | SLEPc | SuiteSparse | Trilinos | Cray TPSL | Numpy | Scipy

DATA:
netCDF | pnetCDF | HDF5 | CGNS | NCL | Splash | CDO Operators | NCO Tools

OVERALL SYSTEM ARCHITECTURE

The two HLRN-III complexes, »Konrad« in Berlin and »Gottfried« in Hanover, are connected by a dedicated 10 Gb/s fiber link. The current installation (phase 1) comprises at each site a four cabinet Cray XC30 system of 744 nodes with 24 Intel Ivybridge cores each, 46.5 TByte distributed memory and 1.4 PByte on-line disk storage in a Lustre filesystem.



VIZ:
Ferret | *Grads* | *Splash* | *Visit* | *NcView*

KEY HARDWARE PROPERTIES OF THE CRAY XC30 SYSTEM (PHASE I)

Each of the two Cray XC30 compute system provides a peak-performance of 343 TFlop/s. The 744 compute nodes within one installation are connected via the Cray Aries network which provides scalability across the entire system. Within a node, 24 IvyBridge cores clocked at 2.4 GHz have access to 64 GByte memory per node. The compute and service nodes have access to global storage resources, i. e. a common HOME and a parallel Lustre filesystem. The HOME filesystem is implemented as network attached storage (NAS) with a total capacity of 1.4 PByte. The Lustre filesystem provides high-speed access to temporary, large data sets with a total capacity of 2.8 PBytes.

SCIENTIFIC DOMAINS ON HLRN-III

WORKLOADS AND THEIR RESOURCE REQUIREMENTS

The projects at HLRN cover a wide variety of scientific areas and research topics with different computational demands. The pie-chart on the right illustrates the CPU time usage on the HLRN-II system in the year 2013 for the different scientific areas.

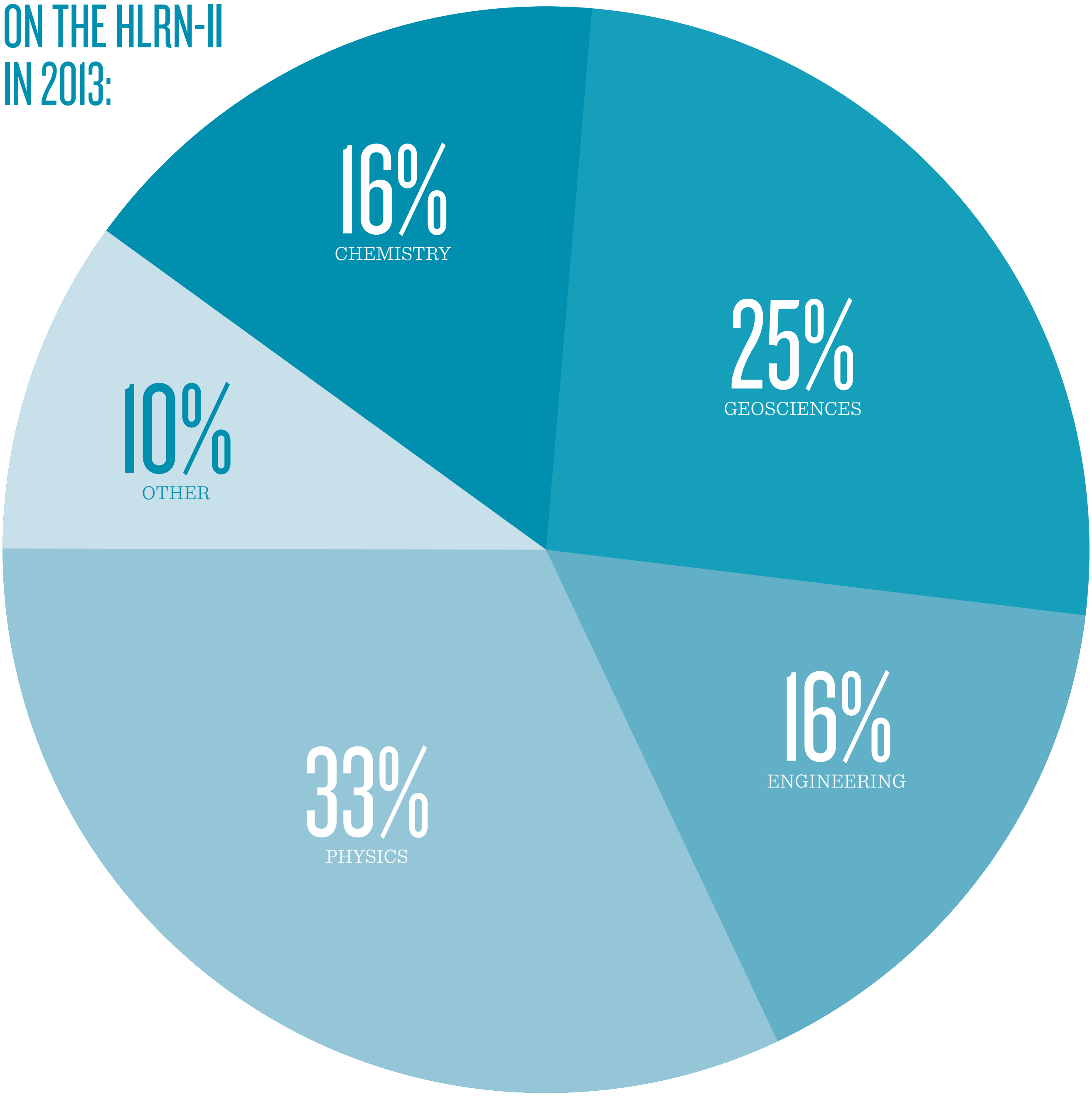
The job sizes are dictated by the problem size in terms of memory requirements as well as the CPU time in regards to the time to solution. First the problem has to fit into memory. If parallel efficiency and scalability of the code and the problem allows, the number of used nodes can be increased to obtain a shorter time to solution. The workload is a broad mix of all kinds of job types. Among the projects we observe two extremes. The first type of projects submit hundreds or thousands of independent jobs ("job clouds") that themselves are either moderately or massively parallel computations. In principle the workload presented by these projects is able to adapt to the system utilization. Projects in particle physics are prominent representatives for this kind of work flow.

Projects of the second type conduct only a small number of simulations that require a long total runtime until the desired solution is obtained. Here each computation should be performed in one long run, requiring several days, weeks or even months of wall time. System stability considerations as well as fair-share aspects in the overall system usage have lead us to the policy of limiting the maximum job run time to 24 hours. For the affected projects this requires user initiated checkpoint and restart capabilities in their codes. These projects may require special treatment by the batch system (e.g. special queues, reservations). Climate development or long-term statistics for unsteady fluid-dynamic flow are typical examples for this type of work flow. The requirements of all other projects lies in between the described extremes.

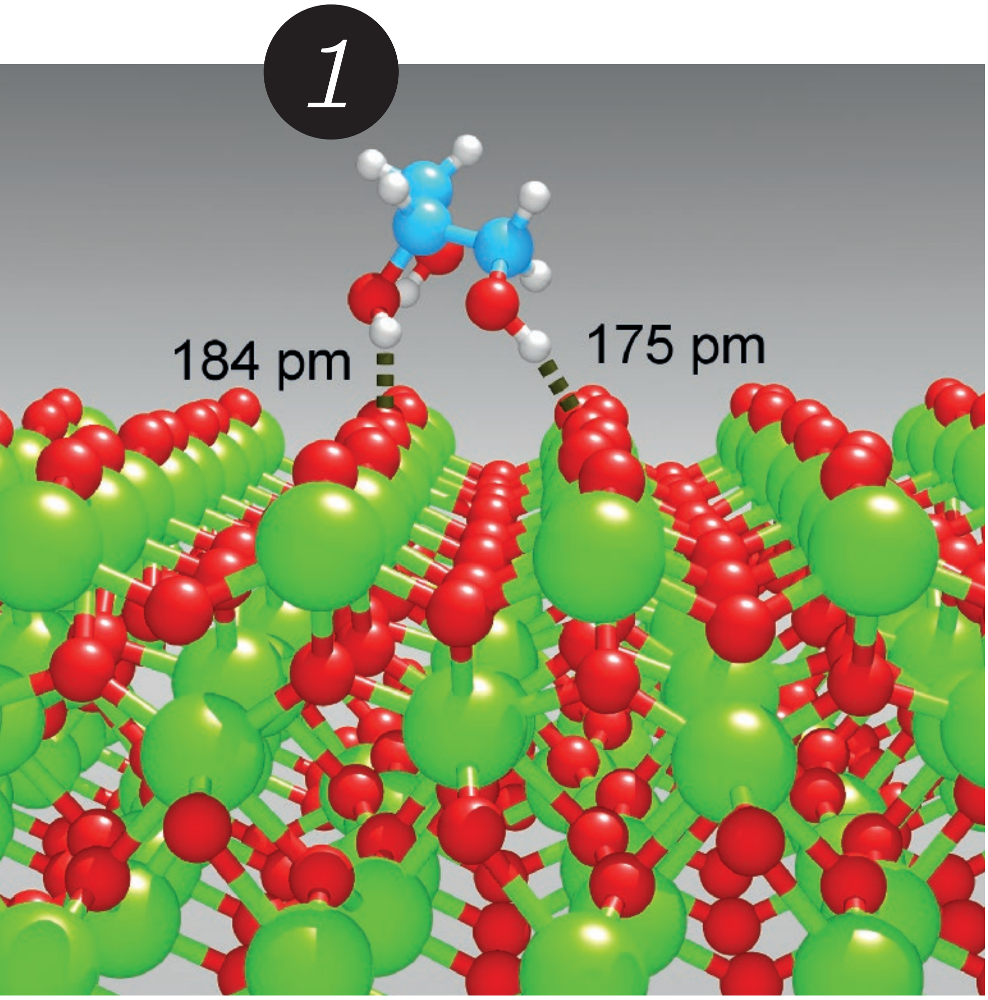
Another aspect influencing the work flow of many projects is I/O. Here the requirements vary from very small data to be read or written to massive data transfer throughout the course of the computations. I/O can be concentrated to one or very few files on one side, while other projects use several files per MPI task, leading to I/O from/to several thousands to tens of thousands of files per job, written in each simulated time step. The overall amount of data transferred in one job can be in the order of several Terabytes.

Especially those projects with big result data sets face the problem of demanding pre- and post-processing procedures. This part of their work flow can often be more time consuming than the computations themselves. Thus these projects need to carefully balance their data production with the capabilities of pre- and post-processing their data.

CPU TIME USAGE ON THE HLRN-II IN 2013:



RESEARCH FIELDS UTILIZING THE HLRN-III



CHEMISTRY AND MATERIAL SCIENCES

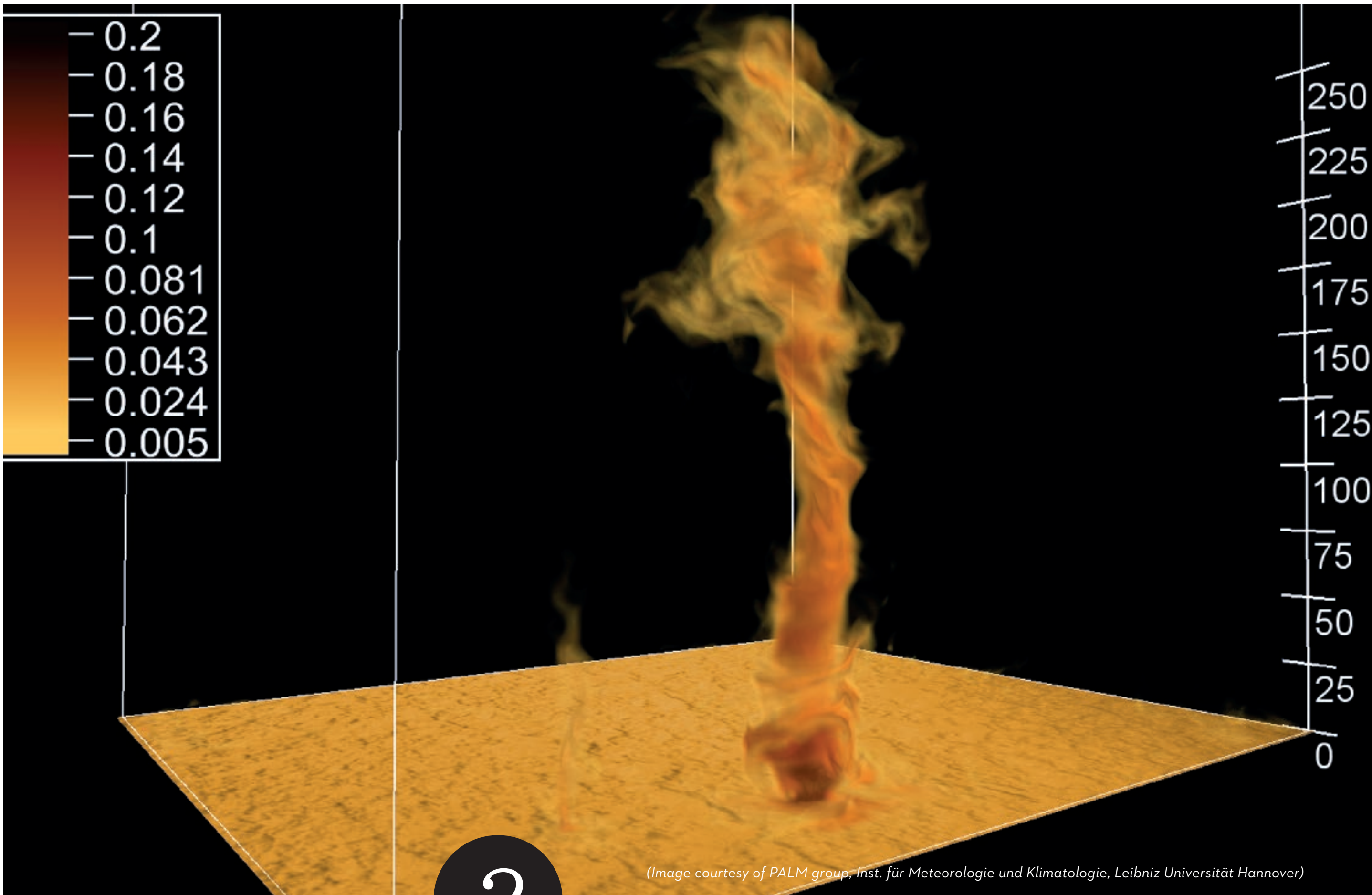
Numerical simulations of molecular complexes range from high-level ab initio electronic structure calculations on systems of only few atoms in size to coarse-grained dynamics simulations of macromolecular systems. The scientific projects employ either empirical interatomic potential functions or density functional theory (DFT) for the description of their model systems. DFT is the workhorse to perform electronic structure calculations on new materials for hydrogen storage, photoelectrolysis, molecular electronics, and spintronics. It is also employed at HLRN to describe properties of transition metal oxide aggregates, nanostructured semiconductors, functionalized oxide surfaces and supported graphene, as well as zeolite and metal-organic framework catalysts.

EARTH SCIENCES

The earth sciences domain plays an important role in the HLRN alliance. Projects are conducted in atmospheric sciences, meteorology, oceanography, and climate research. The PALM software – a parallelized large-eddy simulation model for atmospheric and oceanic flows – is developed at the Institute of Meteorology and Climatology (Leibniz University Hanover) and is especially designed for massively parallel computer architectures. The scalability of PALM has been tested on up to 32,000 cores and has been proven on the Cray XC30 at HLRN-III by simulations with 4320^3 grid points that require approx. 13 TByte memory.

❶ Hydrogen bond distances of a glycerol molecule adsorbed at the surface of a zirconia crystal.

❷ Dust devil simulated with PALM. The displayed area of 200x200 m² represents only a small part of the total simulated domain.



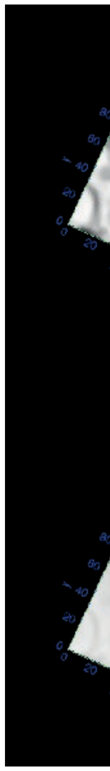
2

3 Large eddy simulation of passive scalar mixing at Reynolds number 10000 and Schmidt number 1000 along the centerline of a mixing device. The area of interest is 2x1x1mm resolved with approximately 200 mio. cells using OpenFOAM on 2000 cores.

4 Visualization of the results for continuum 3D radiation transfer for a stellar convection model for pure absorption (top panel) and strong scattering (bottom panel). A total of about 1.5 billion intensities are calculated for each iteration with the PHOENIX code.

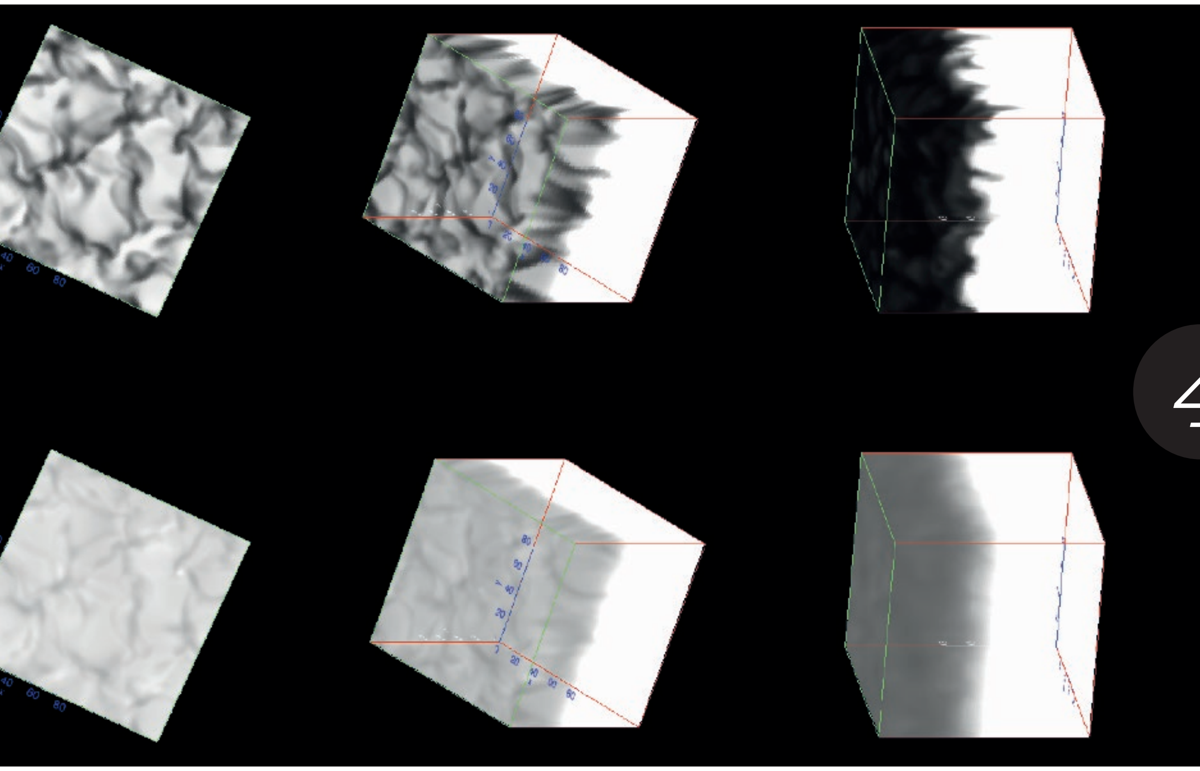
PHYSICS

Scientific questions in elementary particle physics, astrophysics and material science are solved on the HLRN-III resources. The PHOENIX program – developed at the Atrophysics Inst. (University Hamburg) – is designed to model the structures and spectra of a wide range of astrophysical objects, from extrasolar planets to brown dwarfs and all classes of stars, extending to novae and supernovae. The main results from the calculations are synthetic spectra and images, these can be directly compared to observed spectra. PHOENIX/3D simulations show strong and weak scaling (94% efficiency) on up to at 131,072 processes. With that, 3D simulations which consider non-equilibrium thermodynamics produce a minimum of 2.6 TB spectra data, saving the complete imaging information would deliver about 11 PB raw data.



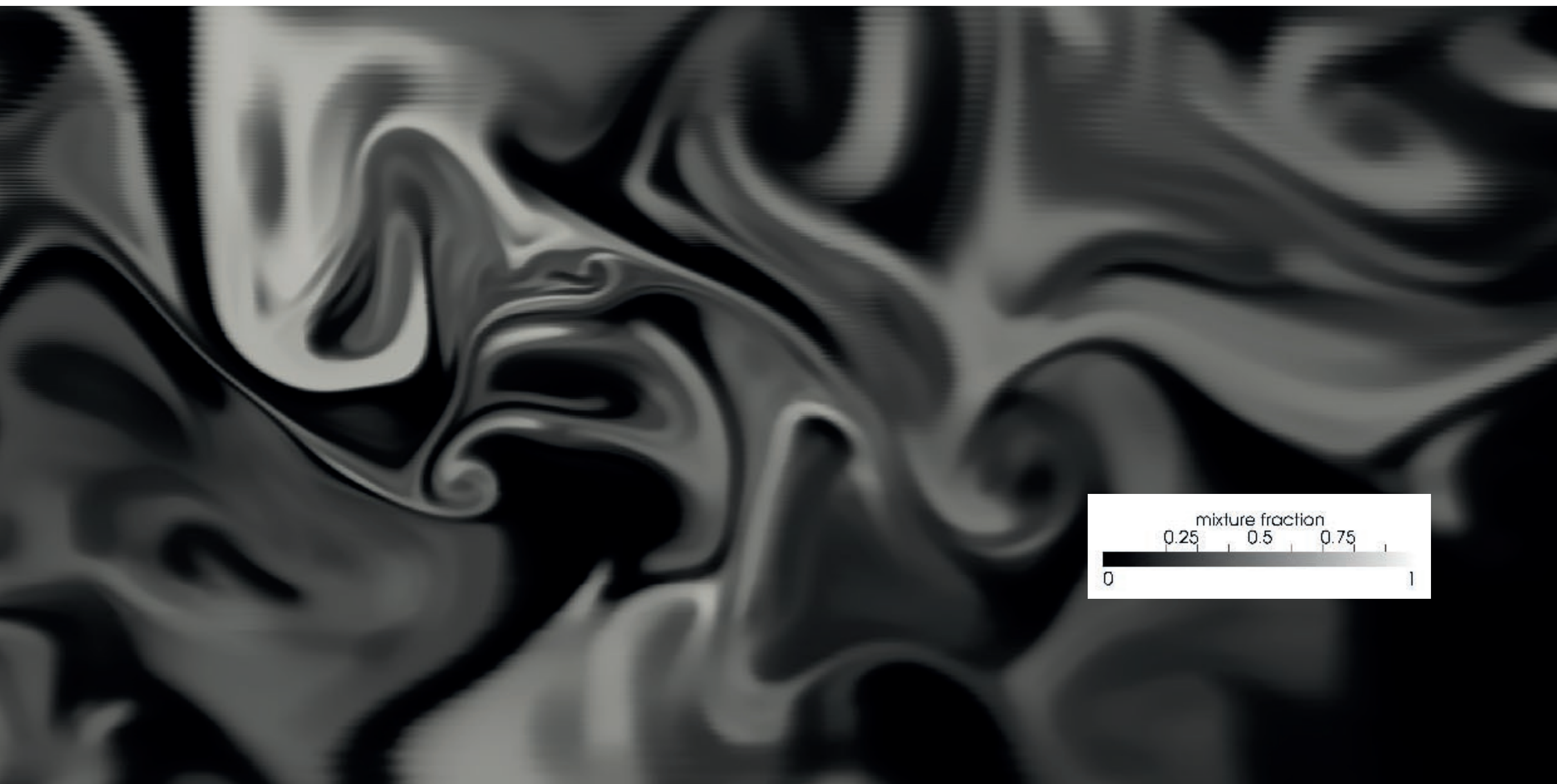
ENGINEERING

Projects in engineering sciences come from the fields of fluid dynamics, combustion, aerodynamics, acoustics and mechanics with a focus on ship hydromechanics and marine technologies. Other research topics are flow optimization, aerodynamics and acoustics of planes, wings, cars, trucks, and trains, also in gas turbine engines, fans, and on propellers for higher flow efficiency. This area is also important in ship hull and propeller design with respect to drag and cavitation reduction and stability. Environmental aspects are important in research on fuel efficiency, pollutants and noise reduction in combustion processes.



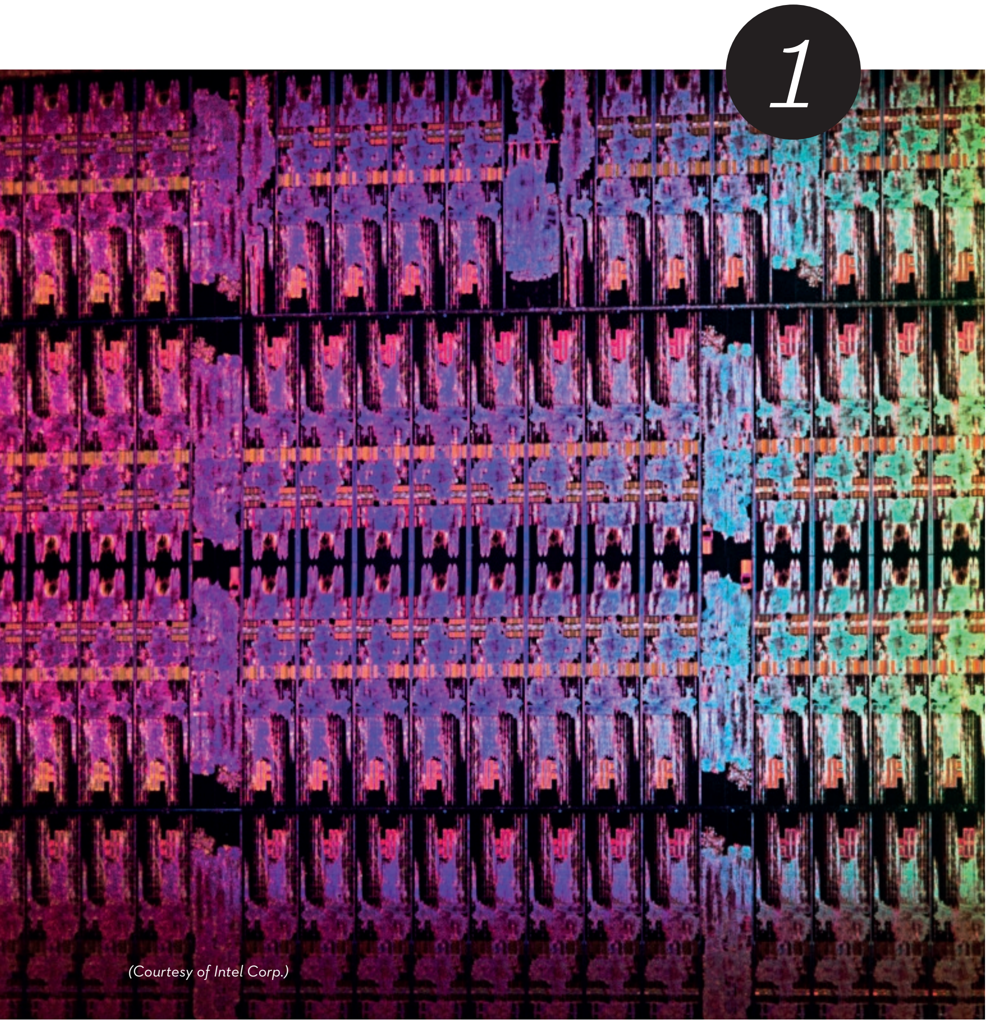
Images courtesy: P. Hausschildt, Astrophysics Institute, University Hamburg

Images courtesy of Institute for Modeling and Numerical Simulation (LEMOS), Faculty of Mechanical Engineering and Marine Technology, University of Rostock.



THE MANY-CORE CHALLENGE

For the last fifty years, the number of transistors on processor chips doubled every 18 to 24 months. This observation became a "law", named after the Intel co-founder Gordon Moore, who first described the trend in his famous 1965 article.



(Courtesy of Intel Corp.)

Folks wisdom simplified Moore's law to the doubling of single processor performance every 18 months. Interestingly, this "law" was a perfect orientation line for many decades. But why should we expect this trend to continue in the future? Physical and architectural restriction cannot be denied. Today, the number of transistors on a chip still increases at a fast pace, but the transistors' switching capabilities are now used to put more processor cores onto a chip. Today, multicore chips with 2 to 12 cores are standard computing devices and many-core-chips with 60 and more cores are also available. We believe that the trend to many-core CPUs will become mainstream in high-performance computers. Future supercomputers will comprise heterogeneous compute and storage architectures to leverage compute and storage resources without sacrificing the power envelope.

»As we look to reaching exascale high performance computing by the end of the decade, we have to ensure that the hardware doesn't get there and leave the applications behind. This means not only taking today's applications and optimizing them on parallel architectures, but also co-designing future application software and the future versions of these architectures to achieve optimal performance at all levels from the computing node to the entire system.« *Raj Hazra, VP of Intel's Architecture Group*

PREPARING FOR THE MANY-CORE FUTURE

With respect to the envisaged growing importance of many-core computing in HPC, the Parallel and Distributed Computing group at ZIB evaluated disruptive processing technologies, including field programmable gate arrays (FPGAs), the ClearSpeed processor, and the Cell Broadband Engine (Cell BE) [43, 44], as well as general-purpose graphics processors (GPUs) from Nvidia [45, 46] and recently the Intel MIC architecture [47].

In 2013, our research in the area of many-core computing was recognized by Intel, who established one of only five worldwide Intel Parallel Computing Centers at ZIB. Because of our deep expertise in the efficient use of many-core architectures for HPC codes, Intel and ZIB jointly established a "Research Center for Many-core High-Performance Computing" located at ZIB. This center fosters the uptake of current and next generation Intel many- and multi-core technology in high performance computing and big data analytics. Our activities are focused on enhancing selected workloads with impact on the HPC community to improve their performance and scalability on many-core processor technologies and platform architectures.

The selected applications cover a wide range of scientific disciplines including materials science and nanotechnology, atmosphere and ocean flow dynamics, astrophysics, drug design, particle physics and big data analytics. Novel programming models and algorithms will be evaluated for the parallelization of the workloads on many-core processors. The workload optimization for many-core processors is supported by research activities associated with many-core architectures at ZIB. Furthermore, the parallelization work is complemented by dissemination and education activities within the North-German Supercomputing Alliance (HLRN) to diminish the barriers involved with the introduction of upcoming highly parallel processor and platform technologies.

1 *The Intel Xeon Phi die with its 61 cores*

ZIB PUBLICATIONS

PEER REVIEWED

Karsten Andrae, Vedat Durmaz, Konstantin Fackeldey, Olga Scharkoi, Marcus Weber (2013). Medizin aus dem Computer. *Der Anaesthesist*, 62(7):561-557.

C. Arellano, S. Mingaleev, I. Koltchanov, A. Richter, J. Pomplun, S. Burger, F. Schmidt (2013). *Efficient design of photonic integrated circuits (PICs) by combining device- and circuit-level simulation tools. Proc. SPIE*, 8627:862711.

Cornelia Auer, Jens Kasten, Andrea Kratz, Eugene Zhang, Ingrid Hotz (2013). Automatic, Tensor-Guided Illustrative Vector Field Visualization. In *Proc. IEEE Pacific Visualization Symposium 2013*, 265-275.

C. Becker, J. Xavier, V. Preidel, P. Wyss, T. Sontheimer, B. Rech, J. Probst, C. Hülsen, B. Löchel, A. Erko, S. Burger, F. Schmidt, F. Back, E. Rudigier-Voigt (2013). Nanophotonic light trapping in polycrystalline silicon thin-film solar cells using periodically nanoimprint-structured glass substrates. In *Next Generation (Nano) Photonic and Cell Technologies for Solar Energy Conversion IV, volume 8824 of Proc. SPIE*, page 88240D.

C. Becker, D. Amkreutz, T. Sontheimer, V. Preidel, D. Lockau, J. Haschke, L. Jogschies, C. Klimm, J. J. Merkel, P. Plocica, S. Steffens, B. Rech (2013). Polycrystalline silicon thin-film solar cells: Status and perspectives. *Solar Energy Materials and Solar Cells*, 119:112-123.

Michael Berlin (2013). Consistent and Partition-Tolerant File Replication in the Cloud File System Xtreem FS. *Praxis der Informationsverarbeitung und Kommunikation*, 36(1):39.

Timo Berthold, Domenico Salvagnin (2013). Cloud branching. In Carla Gomes, Meinolf Sellmann, editors, *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, volume 7874 of *Lecture Notes in Computer Science*, pages 28-43.

Timo Berthold (2013). Measuring the impact of primal heuristics. *Operations Research Letters*, 41(6):611-614.

Timo Berthold, Ambros M. Gleixner (2013). Undercover Branching. *In Experimental Algorithms, 12th International Symposium, SEA* 2013, *Rome, Italy, June 5-7, 2013, Proceedings*, volume 7933 of *Lecture Notes in Computer Science*, pages 212-223.

Andreas Bley, Fabio D'Andreagiovanni, Daniel Karch (2013). Scheduling technology migration in WDM Networks. In *Photonic Networks*, 14. 2013 ITG *Symposium. Proceedings*, 1-5.

Andreas Bley, Fabio D'Andreagiovanni, Daniel Karch (2013). WDM Fiber Replacement Scheduling. In *Proceedings of INOC 2013*, volume 41 of *Electronic Notes in* Discrete Mathematics, pages 189-196.

M. Bock, J. Skibina, D. Fischer, M. Bretschneider, R. Wedell, R. Grunwald, S. Burger, V. Beloglazov, G. Steinmeyer (2013). Nanostructured fibers for sub-10 fs optical pulse delivery. *Laser & Photonics Reviews*, 7:566-570.

Heide Boeth, Georg N. Duda, Markus O. Heller, Rainald M. Ehrig, Ralf Doyscher, Tobias Jung, Philippe Moewis, Sven Scheffler, William R. Taylor (2013). Anterior cruciate ligament-deficient patients with passive knee joint laxity have a decreased range of anterior-posterior motion during active movements. *The American Journal of Sports Medicine*, 41(5):1051-7.

Ralf Borndörfer, Julia Buwaya, Guillaume Sagnol, Elmar Swarat (2013). Optimizing Toll Enforcement in Transportation Networks: a Game-Theoretic Approach. In *Proceedings of INOC'2013*, volume 41 of *Electronic Notes in Discrete Mathematics*, pages 253-260.

Ralf Borndörfer, Heide Hoppmann, Marika Karbstein (2013). A Configuration Model for the Line Planning Problem. In Daniele Frigioni, Sebastian Stiller, editors, *ATMOS 2013 – 13th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems*, 33:68-79.

Ralf Borndörfer, Nam Dung Hoang, Marika Karbstein, Thorsten Koch, Alexander Martin (2013). How many Steiner terminals can you connect in 20 years?. In Michael Jünger, Gerhard Reinelt, editors, *Facets of Combinatorial Optimization; Festschrift for Martin Grötschel*, 215-244, Springer.

Ralf Borndörfer, Marika Karbstein (2013). The Steiner connectivity problem. *Mathematical Programming A*, 142(1):133-167.

Andreas Brandt, Manfred Brandt (2013). Workload and busy period for M/GI/1 with a general impatience mechanism. *Queueing Systems*, 75(2-4):189-209.

Achim Brose, Armin Fügenschuh, Pia Gausemeier, Ingmar Vierhaus, Günther Seliger (2013). A System Dynamic Enhancement for the Scenario Technique. In *Proc. 11th Global Conference on Sustainable Manufacturing*, 561-566.

S. Burger, L. Zschiedrich, J. Pomplun, M. Blome, F. Schmidt (2013). Advanced finite-element methods for design and analysis of nano-optical structures: *Applications. Proc.* SPIE, 8642:864205.

S. Burger, L. Zschiedrich, J. Pomplun, F. Schmidt, B. Bodermann (2013). Fast simulation method for parameter reconstruction in optical metrology. *Proc. SPIE*, 8681:868119.

S. Burger, L. Zschiedrich, J. Pomplun, F. Schmidt (2013). Finite-element based electromagnetic field simulations: Benchmark results for isolated structures. In *Photomask Technology*, 8880:88801Z.

Christina Büsing, Fabio D'Andreagiovanni, Annie Raymond (2013). Robust optimization under multiband uncertainty. In *CTW 2013*, 35-38.

Christina Büsing, Fabio D'Andreagiovanni, Annie Raymond (2013). 0-1 Multiband Robust Optimization. In *Operations Research Proceedings 2013, Selected Papers of the International Annual Conference of the German Operations Research Society (GOR)*.

Fabio D'Andreagiovanni, Carlo Mannino, Antonio Sassano (2013). GUB Covers and Power-Indexed Formulations for Wireless Network Design. *Management Science*, 59(1):142-156.

Peter Dely, Fabio D'Andreagiovanni, Andreas Kassler (2013). Fair Optimization of Mesh-Connected WLAN Hotspots. *Wireless Communications and Mobile Computing*.

Vincent J. Dercksen, Hans-Christian Hege, Marcel Oberlaender (2013). The Filament Editor: An Interactive Software Environment for Visualization, Proof-Editing and Analysis of 3D Neuron Morphology. *NeuroInformatics*, 1-15. (epub ahead of print)

V. Durmaz, S. Schmidt, P. Sabri, C. Piechotta, M. Weber (2013). A hands-off linear interaction energy approach to binding mode and affinity estimation of estrogens. . *Chem. Inf. Model*, 53(10), 2681-2688

Robert Döbbelin, Thorsten Schütt, Alexander Reinefeld (2013). Building Large Compressed PDBs for the Sliding Tile Puzzle. In *Workshop on Computer Games*.

Moritz Ehlke, Heiko Ramm, Hans Lamecker, Hans-Christian Hege, Stefan Zachow (2013). Fast Generation of Virtual X-ray Images for Reconstruction of 3D Anatomy. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2673-2682.

M. Enciso, Ch. Schütte, L. Delle Site (2013). pH-dependent Response of Coiled Coils: A Coarse-Grained Molecular Simulation Study. *Molecular Physics*, 111(22-23):3363-3371.

T. Endres, S. Burger, M. Wurm, B. Bodermann (2013). Numerical investigations of the influence of different commonly applied approximations in scatterometry. *Proc. SPIE*, 8789:878904.

Marianna Euler, Norbert Euler, Thomas Wolf (2013). The Two-Component Camassa-Holm Equations CH(2,1) and CH(2,2): First-Order Integrating Factors and Conservation Laws. *Journal of Nonlinear Mathematical Physics*, 19(1):13-22.

Armin Fügenschuh, Björn Geißler, Ralf Gollmer, Christine Hayn, René Henrion, Benjamin Hiller, Jesco Humpola, Thorsten Koch, Thomas Lehmann, Alexander Martin, Radoslava Mirkov, Antonio Morsi, Werner Römisch, Jessica Rövekamp, Lars Schewe, Martin Schmidt, Rüdiger Schultz, Robert Schwarz, Jonas Schweiger, Claudia Stangl, Marc C. Steinbach, Bernhard M. Willert (2013). Mathematical optimization for challenging network planning problems in unbundled liberalized gas markets. *Energy Systems*, 1-25.

Armin Fügenschuh, Roel van Veldhuizen, Ingmar Vierhaus (2013). Production Planning for Non-Cooperating Companies with Nonlinear Optimization. In *11th Global Conference on Sustainable Manufacturing : Proceedings*, 536-541.

Armin Fügenschuh, Ingmar Vierhaus (2013). System Dynamic Optimization in the Sustainability Assessment of a World-Model. In *11th Global Conference on Sustainable Manufacturing : Proceedings*, 530-535.

Francis Galloway, Max Kahnt, Heiko Ramm, P. Worsley, Stefan Zachow, P. Nair, Mark Taylor (2013). A large scale finite element study of a cementless osseointegrated tibial tray. *Journal of Biomechanics*, 46(11):1900-1906.

ZIB PUBLICATIONS

Gerald Gamrath (2013). Improving strong branching by propagation. Carla Gomes, Meinolf Sellmann, editors, *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, volume 7874 of *Lecture Notes in Computer Science*, pages 347-354.

Ambros M. Gleixner, Stefan Weltge (2013). Learning and Propagating Lagrangian Variable Bounds for Mixed-Integer Nonlinear Programming. In *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems, 10th International Conference, CPAIOR 2013, Yorktown Heights, NY, USA, May 18-22, 2013*, volume 7874 of *Lecture Notes in Computer Science*, pages 355-361.

Richard Grunzke, Sebastian Breuers, Sandra Gesing, Sonja Herres-Pawlis, Martin Kruse, Dirk Blunk, Luis de la Garza, Lars Packschies, Patrick Schäfer, Charlotta Schärfe, Tobias Schlemmer, Thomas Steinke, Bernd Schuller, Ralph Müller-Pfefferkorn, René Jäkel, Wolfgang Nagel, Malcolm Atkinson, Jens Krüger (2013). Standards-based metadata management for molecular simulations. *Concurrency and Computation: Practice and Experience.*

Martin Grötschel, Christian Raack, Axel Werner (2013). Towards optimizing the deployment of optical access networks. *EURO Journal on Computational Optimization.* (epub ahead of print)

Sebastian Götschel, Chamakuri Nagaiah, Karl Kunisch, Martin Weiser (2013). Lossy Compression in Optimal Control of Cardiac Defibrillation. *J. Sci. Comput.*, 1-25.

Sebastian Götschel, Martin Weiser, Christiane Maierhofer, Regina Richter, Mathias Röllig (2013). Fast Defect Shape Reconstruction Based on the Travel Time in Pulse Thermography. *Nondestructive Testing of Materials and Structures*, 6:83-89.

Tobias Günther, Alexander Kuhn, Benjamin Kutz, Holger Theisel (2013). Mass-Dependent Integral Curves in Unsteady Vector Fields. *Computer Graphics Forum*, 32:211-220.

Andreas Günther, Hans Lamecker, Martin Weiser (2013). Flexible Shape Matching with Finite Element Based LDDMM. *International Journal of Computer Vision*, 105(2):128-143.

F. Haack, K. Fackeldey, S. Röblitz, O. Scharkoi, M. Weber, B. Schmidt (2013). Adaptive spectral clustering with application to tripeptide conformation analysis. *The Journal of Chemical Physics*, 139:110-194.

M. Hammerschmidt, D. Lockau, S. Burger, Frank Schmidt, C. Schwanke, S. Kirner, S. Calnan, B. Stannowski, B. Rech (2013). FEM-based optical modeling of silicon thin-film tandem solar cells with randomly textured interfaces in 3D. *Proc. SPIE*, 8620:86201H.

P. Hansen, S. Burger (2013). Investigation of microstructured fiber geometries by scatterometry. *Proc. SPIE*, 8789:87890R.

Steven Harrod, Thomas Schlechte (2013). A Direct Comparison of Physical Block Occupancy Versus Timed Block Occupancy in Train Timetabling Formulations. *Transportation Research Part E: Logistics and Transportation Review*, 54:50-66.

C. Hartmann, R. Banisch, M. Sarich, T. Badowski, Ch. Schütte (2013). Characterization of Rare Events in Molecular Dynamics. *Entropy (Special Issue)*, 16 (1):350-376.

Olga Heismann, Ralf Borndörfer (2013). The Random Hypergraph Assignment Problem. In *Proceedings of the 16th International Multiconference INFORMATION SOCIETY – IS 2013*, 599-602.

Thomas Hildebrandt, Leonid Goubergrits, Werner Heppt, Stefan Bessler, Stefan Zachow (2013). Evaluation of the Intranasal Flow Field through Computational Fluid Dynamics (CFD). *Journal of Facial and Plastic Surgery*, 29(2):93-98.

Achim Hildenbrandt, Gerhard Reinelt, Olga Heismann (2013). Integer Programming Models for the Target Visitation Problem. In *Proceedings of the 16th International Multiconference INFORMATION SOCIETY* – IS 2013, 569-572.

Benjamin Hiller, Thorsten Klug, Andreas Tuchscherer (2013). An Exact Reoptimization Algorithm for the Scheduling of Elevator Groups. *Flexible Services and Manufacturing Journal*, 1-24.

Ulrike Homberg, Daniel Baum, Alexander Wiebel, Steffen Prohaska, Hans-Christian Hege (2013). Definition, Extraction, and Validation of Pore Structures in Porous Materials. In *Topological Methods in Data Analysis and Visualization III.* (accepted for publication)

Dagmar Kainmüller, Hans Lamecker, Markus Heller, Britta Weber, Hans-Christian Hege, Stefan Zachow (2013). Omnidirectional Displacements for Deformable Surfaces. *Medical Image Analysis*, 17(4):429-441.

G. Kewes, A. W. Schell, R. Henze, R. S. Schonfeld, S. Burger, K. Busch, O. Benson (2013). Design and numerical optimization of an easy-to-fabricate photon-to-plasmon coupler for quantum plasmonics. *Applied Physics Letters*, 102:051104.

Stefan Koch, Alexander Wiebel, Jens Kasten, Mario Hlawitschka (2013). Visualizing Linear Neighborhoods in Non-Linear Vector Fields. In *Proceedings of the IEEE Pacific Visualization Symposium*, 249-256.

Andrea Kratz, Daniel Baum, Ingrid Hotz (2013). Anisotropic Sampling of Planar and Two-Manifold Domains for Texture Generation and Glyph Distribution. T*ransactions on Visualization and Computer Graphics (TVCG)*, 19:1782-1794.

Andrea Kratz, Cornelia Auer, Markus Stommel, Ingrid Hotz (2013). Visualization and Analysis of Second-Order Tensors: Moving Beyond the Symmetric Positive-Definite Case. *Computer Graphics Forum – State of the Art Reports*, 1:49-74.

Alexander Kuhn, Norbert Lindow, Tobias Günther, Alexander Wiebel, Holger Theisel, Hans-Christian Hege (2013). Trajectory Density Projection for Vector Field Visualization. In *EuroVis 2013, short papers. M. Hlawitschka, Tino Weinkauf (eds.)*, 31-35.

Julián Lamas-Rodríguez, Dora Blanco Heras, Francisco Argüello, Dagmar Kainmueller, Stefan Zachow, Montserrat Bóo (2013). GPU-accelerated level-set segmentation. *Journal of Real-Time Image Processing*, 1-15. (epub ahead of print)

Markus Leitner, Ivana Ljubić, Markus Sinnl, Axel Werner (2013). On the Two-Architecture Connected Facility Location Problem. *Electronic Notes in Discrete Mathematics*, 41:359-366.

H. Lie, K. Fackeldey, M. Weber (2013). A Square Root Approximation of the Transition Rates for a Markov State Model. *SIAM. J. Matrix Anal. Appl.*, 34(2):738-756.

Norbert Lindow, Daniel Baum, Ana-Nicoleta Bondar, Hans-Christian Hege (2013). Exploration of cavity dynamics in biomolecular systems. *BMC Bioinformatics*, 14.

D. Lockau, T. Sontheimer, C. Becker, E. Rudigier-Voigt, F. Schmidt, B. Rech (2013). Nanophotonic light trapping in 3-dimensional thin-film silicon architectures. *Opt. Express*, 21:A42.

Daniel Lockau, Tobias Sontheimer, Veit Preidel, Christiane Becker, Florian Ruske, Frank Schmidt, Bernd Rech (2013). Optical properties and Limits of a Large-Area Periodic Nanophotonic Light Trapping Design for Polycrystalline Silicon Thin Film Solar Cells. *MRS Proceedings*, 1493:59-64.

Daniel Lockau, Tobias Sontheimer, Veit Preidel, Florian Ruske, Martin Hammerschmidt, Christiane Becker, Frank Schmidt, Bernd Rech (2013). Advanced microhole arrays for light trapping in thin film silicon solar cells. *Solar Energy Materials and Solar Cells*. (epub ahead of print)

Lars Lubkoll, Anton Schiela, Martin Weiser (2013). An optimal control problem in polyconvex hyperelasticity. *SIAM J. Control Opt.*. (accepted for publication)

B. Maes, J. Petráček, S. Burger, P. Kwiecien, J. Luksch, I. Richter (2013). Simulations of high-Q optical nanocavities with a gradual 1D bandgap. *Opt. Express*, 21:6794-6806.

B. A. Nikitin, M. Wulkow, Ch. Schütte (2013). Modeling of Free Radical Styrene/Divinylbenzene Copolymerization with the Numerical Fractionation Technique. *Macromolecular Theory and Simulation*, 22(9):475-489.

Alexander Odesskii, Thomas Wolf (2013). Compatible quadratic Poisson brackets related to a family of elliptic curves. *Journal of Geometry and Physics*, 63(0):107-117.

F. Orejas, A. Boronat, U. Golas, N. Mylonakis (2013). Checking Bisimilarity for Attributed Graph Transformation. In E. Pfenning, editor, *Foundations of Software Science and Computation Structures. Proceedings of FOSSACS 2013*, volume 7794 of *LNCS*, pages 113-128.

F. Paulus, M. E. R. Weiss, D. Steinhilber, A. Nikitin, Ch. Schütte, R. Haag (2013). Anionic Ring-Opening Polymerization Simulations for Hyperbranched Polyglycerols with Defined Molecular Weights. *Macromolecules*, 46(21):8458-8466.

Michael Poss, Christian Raack (2013). Affine recourse for the robust network design problem: Between static and dynamic routing. *Networks – Special Issue: INOC 2011 Conference, June 13–16, 2011, Hamburg, Germany*, 61(2):180-198.

ZIB PUBLICATIONS

Veit Preidel, Daniel Lockau, Tobias Sontheimer, Franziska Back, Eveline Rudigier-Voigt, Bernd Rech, Christiane Becker (2013). Analysis of absorption enhancement in tailor-made periodic polycrystalline silicon microarrays. *Physica status solidi (RRL)-Rapid Research Letters*, 7(12):1045-1049.

Kai Pöthkow, Hans-Christian Hege (2013). Nonparametric Models for Uncertainty Visualization. *Computer Graphics Forum*, 32(3):131-140.

Kai Pöthkow, Christoph Petz, Hans-Christian Hege (2013). Approximate Level-Crossing Probabilities for Interactive Visualization of Uncertain Isocontours. *International Journal for Uncertainty Quantification*, 3(2):101-117.

Christian Raack, Annie Raymond, Thomas Schlechte, Axel Werner (2013). Standings in sports competitions using integer programming. *Journal of Quantitative Analysis in Sports*. (epub ahead of print)

Heiko Ramm, Oscar Salvador Victoria Morillo, Ingo Todt, Hartmut Schirmacher, Arneborg Ernst, Stefan Zachow, Hans Lamecker (2013). Visual Support for Positioning Hearing Implants. In Wolfgang Freysinger, editor, *Proceedings of the 12th annual meeting of the CURAC society*, 116-120. (epub ahead of print)

B. Rath, K. Yousef, D. Katzenstein, R. Shafer, Ch. Schütte, M. von Kleist, T. Merigan (2012). HIV-1 Evolution in Response to Triple Reverse Transcriptase Inhibitor Induced Selective

F. Ries, M. Ditze, A. Piater, E. Singer, V. Fäßler, S. Dreßler, T. Soddemann (2013). Tool-Supported Integration of Hardware Acceleration in Automotive CFD-Simulations. In *NAFEMS 2013 Proceedings*.

Susanna Röblitz, Claudia Stötzel, Peter Deuflhard, Hannah M. Jones, David-Olivier Azulay, Piet van der Graaf, Steven Martin (2013). A mathematical model of the human menstrual cycle for the administration of GnRH analogues. *Journal of Theoretical Biology*, 321:8-27.

Susanna Röblitz, Marcus Weber (2013). Fuzzy spectral clustering by PCCA+: application to Markov state models and data classification. *Advances in Data Analysis and Classification*, 7(2):147-179.

Stefan Saevarsson, Gulshan Sharma, Heiko Ramm, Robert Lieck, Carol Hutchison, Jason Werle, Sigrun Montgomery, Carolina Romeo, Stefan Zachow, Carolyn Anglin (2013). Kinematic Differences Between Gender Specific And Traditional Knee Implants. *The Journal of Arthroplasty*, 28(9):1543-1550.

Guillaume Sagnol (2013). On the semidefinite representation of real functions applied to symmetric matrices. *Linear Algebra and its Applications*, 439(10):2829-2843.

Guillaume Sagnol (2013). Approximation of a maximum-submodular-coverage problem involving spectral functions, with application to experimental designs. *Discrete Applied Mathematics*, 161(1-2):258-276.

M. Sarich, N. Djurdjevac, S. Bruckner, T. Conrad, Ch. Schütte (2013). Modularity revisited. *Journal of Computational Dynamics*. (accepted for publication)

M. Sarich, R. Banisch, C. Hartmann, Ch. Schütte (2013). Markov State Models for Rare Events in Molecular Dynamics. *Entropy (Special Issue)*, 16(1):258-286.

O. Scharkoi, K. Fackeldey, I. Merkulow, K. Andrae, M. Weber, I. Nehls (2013). Conformational Analysis of Alternariol on the Quantum Level. *J. Mol Mod Vol, Molecular Simulation*, 19:2567-2572.

P. Schäfer, S. Dreßler (2013). Shooting Audio Recordings of Insects with SFA. In *AmiBio Workshop, Bonn, Germany*.

T. Schütt, R. Döbbelin, A. Reinefeld (2013). Forward Perimeter Search with Controlled Use of Memory. In *International Joint Conference on Artificial Intelligence, IJCAI-13, Bejing*.

Christof Schütte, M. Sarich (2013). Metastability and Markov State Models in Molecular Dynamics: Modeling, Analysis, Algorithmic Approaches. *Courant Lecture Notes*, 24.

M. Shan, K. E. Carlson, A. Bujotzek, A. Wellner, R. Gust, M. Weber, J.A. Katzenellenbogen, R. Haag (2013). Nonsteroidal Bivalent Estrogen Ligands – An Application of the Bivalent Concept to the Estrogen Receptor. *ACS Chem. Biol.*, 707 8(4)-715.

Yuji Shinano, Nobuo Inui, Youzou Fukagawa, Noburu Takakura (2013). Optimizing Movement Sequences for Step-and-Scan Lithography Equipment. *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, 7(4):608-618.

V. Soltwisch, S. Burger, F. Scholze (2013). Scatterometry sensitivity analysis for conical diffraction versus in-plane diffraction geometry with respect to the side wall angle. *Proc. SPIE*, 8789:878905.

Tobias Sontheimer, Veit Preidel, Daniel Lockau, Franziska Back, Eveline Rudigier-Voigt, Bernd Löchel, Alexei Erko, Frank Schmidt, Alexander Schnegg, Klaus Lips, Christiane Becker, Bernd Rech (2013). Correlation between structural and opto-electronic characteristics of crystalline Si microhole arrays for photonic light management. *Journal of Applied Physics*, 114(17):173513.

J. Stender, M. Berlin, A. Reinefeld (2013). XtreemFS – a File System for the Cloud. In D. Kyriazis, A. Voulodimos, S. Gogouvitis, T. Varvarigou, editors, *Data Intensive Storage Services for Cloud Environments*, IGI Global.

R. Tyagi, S. Malhotra, A. Thünemann, A. Sedighi, M. Weber, A. Schäfer, R. Haag (2013). Investigations of Host-Guest Interactions with Shape-persistent Nonionic Dendritic Micelles. *J. Phys. Chem. C*, 117(23):12307-12317.

J. Tyminski, J. Pomplun, L. Zschiedrich, D. Flagello, T. Matsuyama (2013). Topographic mask modeling with reduced basis finite element method. *Proc. SPIE*, 8683:86831C.

H. Wang, C. Hartmann, Ch. Schütte, L. Site (2013). Grand-canonical-like molecular-dynamics simulations by using an adaptive-resolution technique. *Phys. Rev. X*, 3:011018.

H. Wang, C. Hartmann, Ch. Schütte (2013). Linear response theory and optimal control for a molecular system under nonequilibrium conditions. *Molecular Physics*, 111:3555-3564.

Marcus Weber (2013). Adaptive Spectral Clustering in Molecular Simulation. In: Studies in Classification, Data Analysis, and Knowledge Organization. *XIV: Classification and Data Mining, A.Giusti, G. Ritter, M. Vichi (Eds.), Springer Series*, 147-157.

Martin Weiser (2013). On goal-oriented adaptivity for elliptic optimal control problems. *Opt. Meth. Softw.*, 28(13):969-992.

Alexander Wiebel, Philipp Preis, Frans Vos, Hans-Christian Hege (2013). 3D Strokes on Visible Structures in Direct Volume Rendering. In Mario Hlawitschka, Tino Weinkauf, editors, *EuroVis – Short Papers*, 91-95.

Anna Zakrzewska, Fabio D'Andreagiovanni, Sarah Ruepp, Michael S. Berger (2013). Biobjective Optimization of Radio Access Technology Selection and Resource Allocation in Heterogeneous Wireless Networks. In *Proceedings of the 11th International Symposium on Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt 2013) – RAWNET/WNC3'13*, 652-658.

Max Zinser, Stefan Zachow, Hermann Sailer (2013). Bimaxillary "rotation advancement" procedures in patients with obstructive sleep apnea: A 3-dimensional airway analysis of morphological changes. *International Journal of Oral & Maxillofacial Surgery*, 42(5):569-578.

Valentin Zobel, Jan Reininghaus, Ingrid Hotz (2013). Visualization of Two-Dimensional Symmetric Tensor Fields Using the Heat Kernel Signature. In *Topological Methods in Data Analysis and Visualization III*. (accepted for publication)

L. Zschiedrich, H. Greiner, S. Burger, F. Schmidt (2013). Numerical analysis of nanostructures for enhanced light extraction from OLEDs. *Proc. SPIE*, 8641:86410B.

ZIB PUBLICATIONS

BOOKS

Katja Biermann, Martin Grötschel, Brigitte Lutz-Westphal, editors (2013). Besser als Mathe – Moderne angewandte Mathematik aus dem Matheon zum Mitmachen. Springer Spektrum.

Christof Schütte, M. Sarich (2013). Metastability and Markov State Models in Molecular Dynamics: Modeling, Analysis, Algorithmic Approaches. *Courant Lecture Notes*, 24.

HABILITATIONS

Thorsten Koch (2013). Progress in computational integer programming. Technische Universität Berlin.

DISSERTATIONS

Stephan Aiche (2013). Inferring Proteolytic Processes from Mass Spectrometry Time Series Data. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4968> Freie Universität Berlin.

Alexander Bujotzek (2013). Molecular Simulation of Multivalent Ligand-Receptor Systems. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4965> Freie Universität Berlin.

Pooja Gupta (2013). Integrating high-density marker information into the genetic evaluation of the honey bee. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4966> Freie Universität Berlin.

Dagmar Kainmüller (2013). Deformable Meshes for Accurate Automatic Segmentation of Medical Image Data. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4452> Universität zu Lübeck.

Marika Karbstein (2013). Line Planning and Connectivity. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4874> Technische Universität Berlin.

Andrea Kratz (2013). Three-Dimensional Second-Order Tensor Fields: Exploratory Visualization and Anisotropic Sampling. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4245> Freie Universität Berlin.

D. Lockau (2013). Optical modeling of thin film silicon solar cells with random and periodic light management textures. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4642> Technische Universität Berlin.

Dany Pascal Moualeu-Ngangue (2013). A Mathematical Tuberculosis Model in Cameroon. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4303> Freie Universität Berlin.

Jan Stender (2013). Snapshots in Large-Scale Distributed File Systems. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/3380> Humboldt-Universität zu Berlin.

Stefanie Winkelmann (2013). Markov Decision Processes with Information Costs. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4967> Freie Universität Berlin.

Kaveh Yousef (2013). Stress responses in Escherichia coli and HIV as model systems of adaptation to the environment. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/4969> Freie Universität Berlin.

Kathrin Sobe (2013). Consistency and Fault Tolerance of Distributed Storage Systems. <http://opus4.kobv.de/opus4-zib/frontdoor/index/index/docId/3379> Humboldt-Universität zu Berlin.

OTHER PUBLICATIONS

Rita Albrecht, Klaus Ceynowa, Andreas Degkwitz, Jiri Kende, Thorsten Koch, Gabriele Messmer, Uwe Risch, Beate Rusch, Robert Scheuerl, Michael Voß (2013). Cloudbasierte Infrastruktur für Bibliotheksdaten – auf dem Weg zu einer Neuordnung der deutschen Verbundlandschaft. *Bibliothek Forschung und Praxis*, 37(3):279-287.

Thomas Dierkes, Susanna Röblitz, Moritz Wade, Peter Deuflhard (2013). Parameter identification in large kinetic networks with BioPARKIN. *CoRR – Computing Research Repository*, abs/1303.4928.

Harry Enke, Norman Fiedler, Thomas Fischer, Timo Gnadt, Erik Ketzan, Jens Ludwig, Torsten Rathmann, Gabriel Stöckle, Florian Schintke (2013). Leitfaden zum Forschungsdaten-Management. In Harry Enke, Jens Ludwig, editors, *Leitfaden zum Forschungsdaten-Management*, Verlag Werner Hülsbusch, Glückstadt.

Carl Martin Grewe, Hans Lamecker, Stefan Zachow (2013). Landmark-based Statistical Shape Analysis. In Michael Hermanussen, editor, *Auxology – Studying Human Growth and Development url*, 199-201.

Martin Grötschel (2013). Vertrauen in die/in der Wissenschaft: Einführung in das Thema. *DEBATTE*, 9-16.

Martin Grötschel (2013). Vertrauen in die/in der Wissenschaft: Einführung. *DEBATTE*, 63-69.

Anja Müller, Beate Rusch, Kilian Amrhein, Marco Klindt (2013). Unfertige Dialoge – Das Berliner Förderprogramm Digitalisierung. *Bibliotheksdienst*, 47(12):931-942.

Olaf Paetsch, Daniel Baum, David Breler, Karsten Ehrig, Dietmar Meinel, Steffen Prohaska (2013). 3-D-Visualisierung und statistische Analyse von Rissen in mit Computer-Tomographie untersuchten Betonproben. In *Tagungsband der DGZfP Jahrestagung 2013*.

René Scheumann, Ingmar Vierhaus, Ya-Ju Chang, Armin Fügenschuh, Matthias Finkbeiner (2013). Identification of trade-offs for sustainable manufacturing of a Bamboo Bike by System Dynamics. In *Proceedings of the 27. Conference on Environmental Informatics – Informatics for Environmental Protection, Sustainable Development and Risk Management*, 523-531.

F. Schintke (2013). XtreemFS & Scalaris. *Science & Technology*, 54-55.

ZIB REPORTS

Martin Ballerstein, Dennis Michaels, Stefan Vigerske (2013). Linear Underestimators for bivariate functions with a fixed convexity behavior. ZIB-Report 13-02.

Thomas Bauschert, Christina Büsing, Fabio D'Andreagiovanni, Arie M. C. A. Koster, Manuel Kutschka, Uwe Steglich (2013). Network Planning under Demand Uncertainty with Robust Optimization. ZIB-Report 13-59.

Timo Berthold, Domenico Salvagnin (2012). Cloud branching. ZIB-Report 13-01.

Timo Berthold, Gregor Hendel (2013). Shift-And-Propagate. ZIB-Report 13-03.

Timo Berthold, Ambros M. Gleixner, Stefan Heinz, Stefan Vigerske (2013). Analyzing the computational impact of MIQCP solver components. ZIB-Report 13-08.

Timo Berthold (2013). Measuring the impact of primal heuristics. ZIB-Report 13-17.

Timo Berthold, Ambros M. Gleixner (2013). Undercover Branching. ZIB-Report 13-14.

Timo Berthold (2013). Primal MINLP Heuristics in a nutshell. ZIB-Report 13-42.

Marco Blanco, Thomas Schlechte (2013). Analysis of Micro-Macro Transformations of Railway Networks. ZIB-Report 13-65.

ZIB PUBLICATIONS

Andreas Bley, Fabio D'Andreagiovanni, Daniel Karch (2013). Scheduling technology migration in WDM Networks. ZIB-Report 13-62.

Ralf Borndörfer, Heide Hoppmann, Marika Karbstein (2013). A Configuration Model for the Line Planning Problem. ZIB-Report 13-40.

Ralf Borndörfer, Marika Karbstein (2013). A Primal-Dual Approximation Algorithm for the Steiner Connectivity Problem. ZIB-Report 13-54.

Ralf Borndörfer, Nam-Dung Hoang, Marika Karbstein, Thorsten Koch, Alexander Martin (2013). How Many Steiner Terminals Can You Connect in 20 Years?. ZIB-Report 13-57.

Ralf Borndörfer, Julika Mehrgardt, Markus Reuther, Thomas Schlechte, Kerstin Waas (2013). Re-optimization of Rolling Stock Rotations. ZIB-Report 13-60.

Ralf Borndörfer, Guillaume Sagnol, Thomas Schlechte, Elmar Swarat (2013). Optimal Toll Enforcement – an Integration of Vehicle Routing and Duty Rostering. ZIB-Report 13-79.

Achim Brose, Armin Fügenschuh, Pia Gausemeier, Ingmar Vierhaus, Günther Seliger (2013). A System Dynamic Enhancement for the Scenario Technique. ZIB-Report 13-24.

Alexander Bujotzek, Ole Schütt, Adam Nielsen, Konstantin Fackeldey, Marcus Weber (2013). Efficient Conformational Analysis by Partition-of-Unity Coupling. ZIB-Report 13-58.

Christina Büsing, Fabio D'Andreagiovanni (2013). A new theoretical framework for Robust Optimization under multi-band uncertainty. ZIB-Report 13-61.

Christina Büsing, Fabio D'Andreagiovanni (2014). Robust Optimization under Multi-band Uncertainty - Part I: Theory. ZIB-Report 13-10.

Christina Büsing, Fabio D'Andreagiovanni, Annie Raymond (2014). 0-1 Multiband Robust Optimization. ZIB-Report 13-77.

Fabio D'Andreagiovanni, Annie Raymond (2013). Multiband Robust Optimization and its Adoption in Harvest Scheduling. ZIB-Report 13-76.

Fabio D'Andreagiovanni, Jonatan Krolikowski, Jonad Pulaj (2014). A hybrid primal heuristic for Robust Multiperiod Network Design. ZIB-Report 13-78.

Vincent J. Dercksen, Hans-Christian Hege, Marcel Oberlaender (2013). The Filament Editor: An Interactive Software Environment for Visualization, Proof-Editing and Analysis of 3D Neuron Morphology. ZIB-Report 13-75.

Sebastian Dreßler, Thomas Steinke (2013). An Automated Approach for Estimating the Memory Footprint of Non-Linear Data Objects. ZIB-Report 13-46.

Robert Döbbelin, Thorsten Schütt, Alexander Reinefeld (2013). Building Large Compressed PDBs for the Sliding Tile Puzzle. ZIB-Report 13-21.

Moritz Ehlke, Heiko Ramm, Hans Lamecker, Hans-Christian Hege, Stefan Zachow (2013). Fast Generation of Virtual X-ray Images from Deformable Tetrahedral Meshes. ZIB-Report 13-38.

Hildegard Franck, Stefan Lohrum, Jannis Polojannis (2014). Automatisierte Fernleihe im KOBV – Handbuch für Bibliotheken – 4. Auflage. ZIB-Report 13-35.

Armin Fügenschuh, Björn Geißler, Ralf Gollmer, Christine Hayn, Rene Henrion, Benjamin Hiller, Jesco Humpola, Thorsten Koch, Thomas Lehmann, Alexander Martin, Radoslava Mirkov, Antonio Morsi, Werner Römisch, Jessica Rövekamp, Lars Schewe, Martin Schmidt, Rüdiger Schultz, Robert Schwarz, Jonas Schweiger, Claudia Stangl, Marc C. Steinbach, Bernhard M. Willert (2013). Mathematical Optimization for Challenging Network Planning Problems in Unbundled Liberalized Gas Markets. ZIB-Report 13-13.

Armin Fügenschuh, Ingmar Vierhaus (2013). System Dynamic Optimization in the Sustainability Assessment of a World-Model. ZIB-Report 13-23.

Armin Fügenschuh, Roel van Veldhuizen, Ingmar Vierhaus (2013). Production Planning for Non-Cooperating Companies with Nonlinear Optimization. ZIB-Report 13-25.

Armin Fügenschuh, Ingmar Vierhaus (2013). A Global Approach to the Optimal Control of System Dynamics Models. ZIB-Report 13-28.

Armin Fügenschuh, Jesco Humpola (2013). A Unified View on Relaxations for a Nonlinear Network Flow Problem. ZIB-Report 13-31.

Armin Fügenschuh, Stefan N. Grösser, Ingmar Vierhaus (2013). A Global Approach to the Control of an Industry Structure System Dynamics Model. ZIB-Report 13-67.

Gerald Gamrath, Thorsten Koch, Alexander Martin, Matthias Miltenberger, Dieter Weninger (2013). Progress in Presolving for Mixed Integer Programming. ZIB-Report 13-48.

Gerald Gamrath (2013). Improving strong branching by domain propagation. ZIB-Report 13-47.

Ambros M. Gleixner, Stefan Weltge (2013). Learning and Propagating Lagrangian Variable Bounds for Mixed-Integer Nonlinear Programming. ZIB-Report 13-04.

Martin Grötschel, Christian Raack, Axel Werner (2013). Towards optimizing the deployment of optical access networks. ZIB-Report 13-11.

Sebastian Götschel, Chamakuri Nagaiah, Karl Kunisch, Martin Weiser (2013). Lossy Compression in Optimal Control of Cardiac Defibrillation. ZIB-Report 13-26.

Sebastian Götschel, Martin Weiser (2013). Lossy Compression for PDE-constrained Optimization: Adaptive Error Control. ZIB-Report 13-27.

Sebastian Götschel, Christoph von Tycowicz, Konrad Polthier, Martin Weiser (2013). Reducing Memory Requirements in Scientific Computing and Optimal Control. ZIB-Report 13-64.

Steven Harrod, Thomas Schlechte (2013). A Direct Comparison of Physical Block Occupancy Versus Timed Block Occupancy in Train Timetabling Formulations. ZIB-Report 13-18.

Carsten Hartmann, Ralf Banisch, Marco Sarich, Thomas Badowski, Christof Schütte (2013). Characterization of Rare Events in Molecular Dynamics. ZIB-Report 13-51.

Tim Hasler, Wolfgang Peters-Kottig (2013). Vorschrift oder Thunfisch? – Zur Langzeitverfügbarkeit von Forschungsdaten. ZIB-Report 13-70.

Stefan Heinz, Wen-Yang Ku, Christopher Beck (2013). Recent improvements using constraint integer programming for resource allocation and scheduling. ZIB-Report 13-05.

Olga Heismann, Achim Hildenbrandt, Francesco Silvestri, Gerhard Reinelt, Ralf Borndörfer (2013). HUHFA: A Framework for Facet Classification. ZIB-Report 13-45.

Benjamin Hiller, Torsten Klug, Jakob Witzig (2013). Reoptimization in branch-and-bound algorithms with an application to elevator control. ZIB-Report 13-15.

Ulrike Homberg, Daniel Baum, Alexander Wiebel, Steffen Prohaska, Hans-Christian Hege (2013). Definition, Extraction, and Validation of Pore Structures in Porous Materials. ZIB-Report 13-56.

Jesco Humpola, Armin Fügenschuh (2013). A New Class of Valid Inequalities for Nonlinear Network Design Problems. ZIB-Report 13-06.

Jesco Humpola, Armin Fügenschuh, Thomas Lehmann (2013). A Primal Heuristic for MINLP based on Dual Information. ZIB-Report 13-49.

Marco Klindt, Daniel Baum, Steffen Prohaska, Hans-Christian Hege (2013). iCon.text – a customizable iPad app for kiosk applications in museum exhibitions. ZIB-Report 13-07.

Marco Klindt, Kilian Amrhein, Anja Müller, Wolfgang Peters-Kottig (2013). Dealing with all the data – Participating in workflows to transform digital data about cultural heritage objects within a digital long-term preservation infrastructure. ZIB-Report 13-69.

Torsten Klug, Ralf Borndörfer, Armin Fügenschuh, Thilo Schang, Thomas Schlechte, Hanno Schülldorf (2013). The Freight Train Routing Problem. ZIB-Report 13-36.

Dorian Krause, Konstantin Fackeldey, Rolf Krause (2013). A parallel multiscale simulation toolbox for coupling molecular dynamics and finite elements. ZIB-Report 13-50.

Alexander Kuhn, Wito Engelke, Christian Rössl, Markus Hadwiger, Holger Theisel (2013). Time Line Cell Tracking for the Approximation of Lagrangian Coherent Structures with Subgrid Accuracy. ZIB-Report 13-71.

Felix Lehmann (2013). Inexaktheit in Newton-Lagrange-Verfahren für Optimierungsprobleme mit Elliptischen PDGL-Nebenbedingungen. ZIB-Report 13-37.

Markus Leitner, Ivana Ljubic, Markus Sinnl, Axel Werner (2013). On the Two-Architecture Connected Facility Location Problem. ZIB-Report 13-29.

ZIB PUBLICATIONS

Han Cheng Lie, Konstantin Fackeldey, Marcus Weber (2013). A square root approximation of transition rates for a Markov State Model. ZIB-Report 13-43.

Kristin Meisel, Susanna Röblitz (2013). Mathematische Modellierung der Knochenremodellierung. ZIB-Report 13-41.

Dany Pascal Moualeu, Susanna Röblitz, Rainald Ehrig, Peter Deuflhard (2013). Parameter Identification for a Tuberculosis Model in Cameroon. ZIB-Report 13-72.

Dany Pascal Moualeu, Martin Weiser, Rainald Ehrig, Peter Deuflhard (2013). Optimal control for a tuberculosis model with undetected cases in Cameroon. ZIB-Report 13-73.

Anja Müller, Beate Rusch, Kilian Amrhein, Marco Klindt (2013). Unfertige Dialoge – Das Berliner Förderprogramm Digitalisierung. ZIB-Report 13-68.

Adam Nielsen, Konstantin Fackeldey, Marcus Weber (2013). On a Generalized Transfer Operator. ZIB-Report 13-74.

Julia Plöntzke, Mascha Berg, Amiya Olany, Sabine Leonhard-Marek, Kerstin Elisabeth Müller, Susanna Röblitz (2013). Modeling potassium balance in dairy cows . ZIB-Report 13-09.

Wolfgang Pyszkalski (2013). Übersicht über die Datenhaltung im ZIB und die Möglichkeiten einer Nutzung durch Projekte. ZIB-Report 13-16.

Christian Raack, Annie Raymond, Axel Werner, Thomas Schlechte (2013). Integer Programming and Sports Rankings. ZIB-Report 13-19.

Heiko Ramm, Oscar Salvador Morillo Victoria, Ingo Todt, Hartmut Schirmacher, Arneborg Ernst, Stefan Zachow, Hans Lamecker (2013). Visual Support for Positioning Hearing Implants. ZIB-Report 13-53.

Guillaume Sagnol, Radoslav Harman (2013). Computing exact D-optimal designs by mixed integer second order cone programming. ZIB-Report 13-39.

Marco Sarich, Ralf Banisch, Carsten Hartmann, Christof Schütte (2013). Markov State Models for Rare Events in Molecular Dynamics. ZIB-Report 13-52.

René Scheumann, Ingmar Vierhaus, Ya-Ju Chang, Armin Fügenschuh, Matthias Finkbeiner (2013). Identification of trade-offs for sustainable manufacturing of a Bamboo Bike by System Dynamics. ZIB-Report 13-32.

Yuji Shinano, Tobias Achterberg, Timo Berthold, Stefan Heinz, Thorsten Koch, Stefan Vigerske, Michael Winkler (2013) SCIP. ZIB-Report 13-22.

Yuji Shinano, Stefan Heinz, Stefan Vigerske, Michael Winkler (2013). FiberSCIP – A shared memory parallelization of SCIP. ZIB-Report 13-55.

Yuji Shinano, Tobias Achterberg, Timo Berthold, Stefan Heinz, Thorsten Koch, Michael Winkler (2013). Solving hard MIPLIB2003 problems with ParaSCIP on Supercomputers: An update. ZIB-Report 13-66.

Han Wang, Carsten Hartmann, Schütte Christof (2013). Linear response theory and optimal control for a molecular system under nonequilibrium conditions. ZIB-Report 13-33.

Marcus Weber, Konstantin Fackeldey (2013). Computing the Minimal Rebinding Effect Included in a Given Kinetics. ZIB-Report 13-12.

Martin Weiser (2013). Faster SDC convergence on non-equidistant grids by DIRK sweeps. ZIB-Report 13-30.

Florian Wende, Thomas Steinke (2013). Swendsen-Wang Multi-Cluster Algorithm for the 2D/3D Ising Model on Xeon Phi and GPU. ZIB-Report 13-44.

Alexander Wiebel, Philipp Preis, Frans M. Vos, Hans-Christian Hege (2013). Computation and Application of 3D Strokes on Visible Structures in Direct Volume Rendering. ZIB-Report 13-20.

Stefanie Winkelmann, Christof Schütte, Max von Kleist (2013). Markov Control Processes with Rare State Observation: Theory and Application to Treatment Scheduling in HIV-1. ZIB-Report 13-34.

Anna Zakrzewska, Fabio D'Andreagiovanni, Sarah Ruepp, Michael S. Berger (2013). Biobjective Optimization of Radio Access Technology Selection and Resource Allocation in Heterogeneous Wireless Networks. ZIB-Report 13-63.

REFERENCES

SOLUTIONS FOR INDIVIDUALIZED MEDICINE

[1] S. Saevarsson, G. Sharma, H. Ramm, R. Lieck, C. Hutchison, J. Werle, S. Montgomery, C. Romeo, S. Zachow, C. Anglin (2013). Kinematic differences between gender specific and traditional knee implants. The Journal of Arthroplasty, 28(9):1543 – 1550.

[2] F. Galloway, M. Kahnt, H. Ramm, P. Worsley, S. Zachow, P. Nair, M. Taylor (2013). A large scale finite element study of a cementless osseointegrated tibial tray. Journal of Biomechanics, 46(11):1900 – 1906.

[3] M. Zinser, S. Zachow, H. Sailer (2013). Bimaxillary "rotation advancement" procedur es in patients with obstructive sleep apnea: A 3- dimensional airway analysis of morphological changes. International Journal of Oral & Maxillofacial Surgery, 42(5):569 – 578.

[4] S. Zachow, M. Zilske, H.-C. Hege (2007). 3d reconstruction of individual anatomy from medical image data: Segmentation and geometry processing. ZIB Report 07-41, Zuse Institute Berlin, Takustr.7, 14195 Berlin.

[5] D. Kainmüller, H. Lamecker, M. Heller, B. Weber, H.- C. Hege, S. Zachow (2013). Omnidirectional displacements for deformable surfaces. 17(4):429 – 441.

[6] J. Lamas-Rodríguez, D. Blanco Heras, F. Argüello, D. Kainmueller, S. Zachow, M. Boo (2013).Gpu-acceleratedlevel-setsegmentation. Journal of Real-Time Image Processing, pages 1 – 15. epub ahead of print.

[7] M. Ehlke, H. Ramm, H. Lamecker, H.-C. Hege, S. Zachow (2013). Fast generation of virtual x-ray images for reconstruction of 3d anatomy. Visualization and Computer Graphics, IEEE Transactions on (Volume: 19, Issue: 12) pp. 2673 - 2682, 19(12).

[8] M. Weiser (2013). Faster SDC convergence on non-equidistant grids with DIRK sweeps. ZIB Report 13-30.

[9] S. Götschel, C. von Tycowicz, K. Polthier, M. Weiser (2013). Reducing memory requirements in scientic computing and optimal control. ZIB Report 13-64.

[10] S. Götschel, M. Weiser (2013). Lossy compression for PDE-constrained optimization: Adaptive error control. ZIB Report 13-27.

[11] L. Lubkoll, A. Schiela, M. Weiser (2012). An optimal control problem in polyconvex hyperelasticity. ZIB-Report 12-08.

DESIGN OF FUNCTIONAL MOLECULES

[12] V. Durmaz, S. Schmidt, P. Sabri, C. Piechotta, M. Weber (2013). A hands-off linear interaction energy approach to binding mode and affinity estimation of estrogens. J. Chem. Inf. Model, 53(10):2681–2688.

[13] V. Durmaz, M. Weber, R. Becker (2012). How to Simulate Affinities for Host-Guest Systems Lacking Binding Mode Information: Application in the Liquid Chromatographic Separation of Hexabromocyclododecane Stereoisomers. Journal of Molecular Modeling, 18(6):2399-2408.

[14] F. Wende, F. Cordes, T. Steinke (2012).On Improving the Performance of Multi-threaded CUDA Applications with Concurrent Kernel Execution by Kernel Reordering, in Proceedings of the 2012 Symposium on Application Accelerators in High Performance Computing, SAAHPC '12, (Washington, DC, USA), pp. 74-83, IEEE Computer Society, 2012.

[15] F. Cordes (2013). GLAT – Global Local Adaptive Thermodynamics." GETLIG&TAR.

[16] E. Freye (2010): Opioide in der Medizin, Springer, 8. Auflage.

[17] C. Zöllner, C. Stein (2007): Opioids, Handb. Exp. Pharmacol. (177):31-63.

[18] C. Stein, M. Weber, O. Scharkoi, P. Deuflhard. Method and system for identifying compounds that bind and preferably activate a target opioid receptor in a pH-dependent manner. EP261327

REFERENCES

[19] C. Stein, C. Zöllner, M. Weber, O. Scharkoi. Fentanyl derivatives as pH-dependent opioid receptor agonists. EP2559685

[20] M. Klimm, A. Bujotzek, M. Weber (2011). Direct Reweighting Strategies in Conformation Dynamics. MATCH Communications in Mathematical and in Computer Chemistry, 65(2):333-346.

[21] K. Fackeldey, M. Klimm, M. Weber (2012). A Coarse Graining Method for the Dimension Reduction of the State Space of Biomolecules. Journal of Mathematical Chemistry, 50(9):2623-2635.

[22] N. Lindow, D. Baum, S. Prohaska, H.-C. Hege (2010). Accelerated Visualization of Dynamic Molecular Surfaces. Comput. Graph. Forum, Vol. 29, pp. 943-952.

[23] N. Lindow, D. Baum, H.-C. Hege (2011). Voronoi-based Molecular Path Detection and Visualization. IEEE Trans. Vis. Comput. Graph., Vol.17, pp. 2025-2034.

[24] B. Sellergren (2001). Molecularly imprinted polymers. Man made mimics of antibodies and their applications in analytical chemistry. Elsevier, Amsterdam.

[25] R. Wagner, W. Wan, M. Biyikal, E. Benito-Peña, M.C. Moreno-Bondi, I. Lazraq, K. Rurack, B. Sellergren (2013). J. Org. Chem. 78:1377-1389,

[26] N. Lindow, D. Baum, A.-N. Bondar, H.-C. Hege (2012). Dynamic Channels in Biomolecular Systems: Path Analysis and Visualization. Proceedings of IEEE Symposium on Biological Data Visualization (Biovis '12), pp 99-106, 2012

[27] N. Lindow, D. Baum, A.-N. Bondar. H.-C. Hege (2013). Exploration of Cavity Dynamics in Biomolecular Dystems. BMC Bioinformatics, 14 (Suppl 19): S5.

[28] M. Weber, K. Fackeldey (2013). Computing the Minimal Rebinding Effect Included in a Given Kinetics, accepted for publication in SIAM Mult Mod.

[29] A. Bujotzek, O. Schütt, K. Fackeldey, A. Nielsen, M. Weber (2013). Efficient Conformational Analysis by Partition-of-Unity Coupling. Math Chem accepted (2013), http://dx.doi.org/10.1007/s10910-013-0265-1, also as ZIB-Report 13-58.

[30] A. Nielsen, K. Fackeldey, M.Weber (2013). On a generalized transfer operator, ZIB-Report 13-74.

[31] Ch. Schütte, A. Nielsen, M. Weber (2013). Markov State Models and Molecular Alchemy, submitted to Mol. Phys.

[32] W. Zhang, C. Hartmann, M. Weber, Ch. Schütte (2013). Importance sampling in path space for diffusion processes, submitted to JCP.

MATHEMATICS IN TRAFFIC AND TRANSPORT

[33] M. Grötschel , R. Borndörfer, (2014). Mathematik im Verkehr. Heureka '14. FGSV Köln.

[34] R. Borndörfer, M. Grötschel, U. Jäger, (2010). Planning Problems in Public Transit. In Grötschel, Lucas, Mehrmann, (EDS). Production Factor Mathematics. Acatech/Springer, p. 95-122.

[35] R. Borndörfer, B. Erol, T. Graffagnino, T. Schlechte, E. Swarat (2010). Optimizing the Simplon Railway Corridor. Annals of OR.

[36] R. Borndörfer, I. Dovica, I. Nowak, T. Schickinger (2010). Robust tail assignment. ZIB Report 10-08.

[37] R. Borndörfer, G. Sagnol, E. Swarat (2011). An IP Approach to Toll Enforcement Optimization on German Motorways. OR 2011 Proceedings.

SHEPHERDING THE BITS

[38] A. Müller, B. Kusch, K. Amrhein and M. Klindt (2013). Unfertige Dialoge - Das Berliner Förderprogramm Digitalisierung Bibliotheksdienst 2013; 47(12):931-942.

SIMILARITY SEARCH AND INSECT CLASSIFICATION

[39] P. Schäfer, M. Högqvist (2012). SFA: a symbolic Fourier approximation and index for similarity search in high dimensional datasets. EDBT Proceedings of the 15th International Conference on Extending Database Technology, pp. 516-527.

[40] P. Schäfer (2014). Towards Time Series Classification without Human Preprocessing, Machine Learning and Data Mining in Pattern Recognition: 10th International Conference, MLDM 2014, St. Peterburg, Russia.

[41] P. Schäfer, S. Dreßler (2013). Shooting Audio Recordings of Insects with SFA, AmiBio Workshop, Bonn, Germany.

[42] I. Potamitis, P. Schäfer (2014). On Classyfing Insects from their Wing-Beat: New Results, Symposium on "Ecology and Accoustics: Emergent Properties from Community to Landscape", Paris, France.

SUPERCOMPUTING AT ZIB

[43] P. May, G. W. Klau, M. Bauer, T. Steinke (2006). Accelerated microRNAPrecursor Detection Using the Smith-Waterman Algorithm on FPGAs. In Werner Dubitzky, Assaf Schuster, Peter M. A. Sloot, Michael Schroeder, and Mathilde Romberg, editors, GCCB, volume 4360 of Lecture Notes in Computer Science, pages 19–32. Springer.

[44] T. Steinke, A. Reinefeld, T. Schütt (2006). Experiences with High-Level Programming of FPGAs on Cray XD1. In Cray Users Group (CUG 2006).

[45] T. Steinke, K. Peter, S. Borchert (2010). Efficiency Considerations of Cauchy Reed-Solomon Implementations on Accelerator and Multi-Core Platforms. In Symposium on Application Accelerators in High Performance Computing (SAAHPC).

[46] F. Wende, F. Cordes, T. Steinke (2012). On Improving the Performance of Multithreaded CUDA Applications with Concurrent Kernel Execution by Kernel Reordering. Symposium on Application Accelerators in High-Performance Computing (SAAHPC), 0:74–83.

[47] F. Wende, T. Steinke (2013). Swendsen-Wang Multi-cluster Algorithm for the 2D/3D Ising Model on Xeon Phi and GPU. In Proceedings of SC13: International Conference forHigh Performance Computing, Networking, Storage and Analysis, SC '13, pages 83:1–83:12, New York, NY, USA. ACM.

IMPRINT

Published by
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustraße 7 | 14195 Berlin | Deutschland

Phone +49-30-841 85-0
Telefax +49-30-841 85-125
Contact bibliothek@zib.de

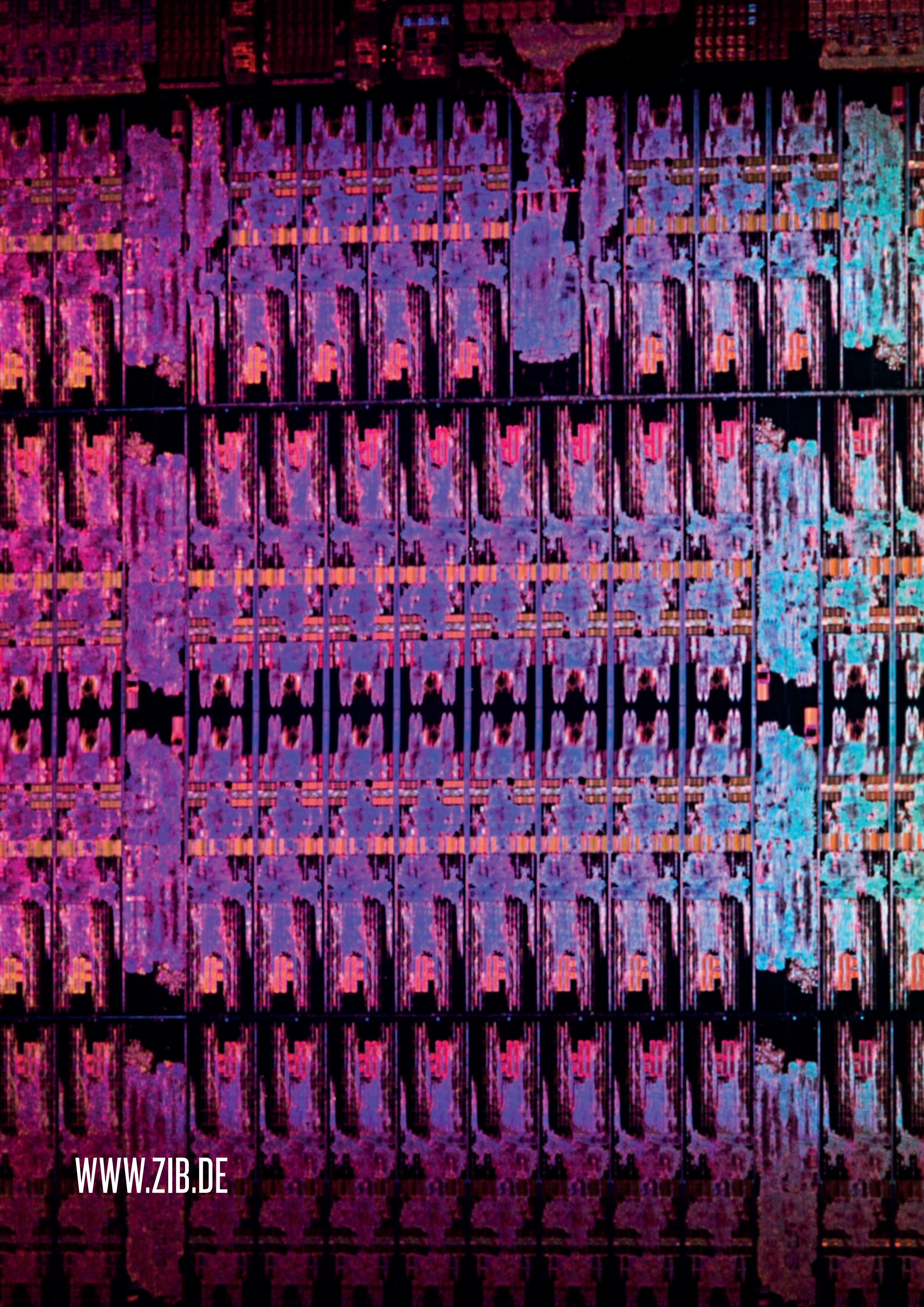
www.zib.de

Layout
SCHIEBE PREIL BAYER
werbung und kommunikation - Berlin

ISSN
0934-5892

NOTES

NOTES


WWW.ZIB.DE