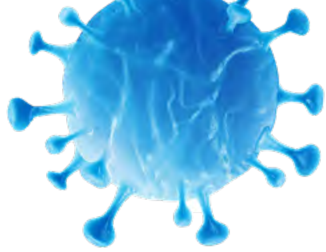
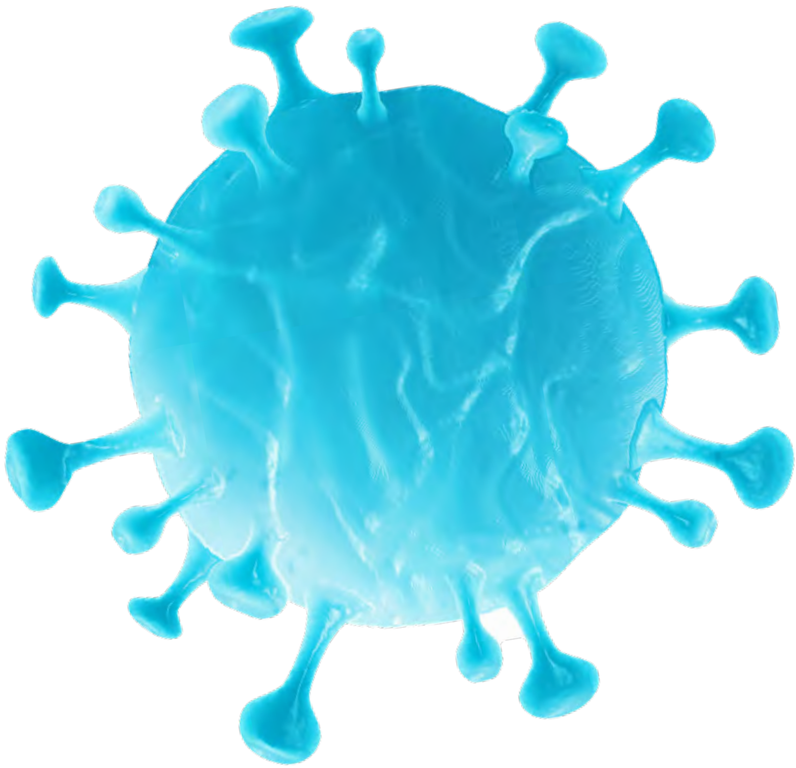


ANNUAL REPORT 2020



ZIB
ZUSE INSTITUTE BERLIN

CONTENTS

ZIB.DE


ZUSE INSTIT


TUTE BERLIN

EXECUTIVE SUMMARY

The corona crisis is massively changing the lives of individuals and their coexistence. This is true for everyday life as well as for the life at a research institution like ZIB. Many things have changed since the beginning of 2020. We learned how to work and do research together via virtual platforms and had to change the way we interact, coordinate, and more generally embrace digitalization. How fundamental and sustainable are these developments? Will research and scientific exchange be different after the corona crisis? At ZIB, cohesion was great and we believe that we kept the special team spirit of the institute, across all the different groups and people. However, at the same time everybody is eagerly waiting for the pandemic situation to improve such that we can meet each other in person again.

Research related to the corona crisis. Soon after the outbreak of the pandemic, several research projects related to corona started at ZIB, with model-based as well as data-based approaches. Our main project, entitled "Model-based investigation of school closures and other mitigation measures for COVID-19" brought together several research groups from TU Berlin, the Robert Koch Institute, and ZIB. The overarching goal of this joint project – funded by the German Ministry of Research and Education (BMBF) from 2020 until 2023 – is to investigate the effects of non-pharmaceutical interventions on the infection dynamics of COVID-19, ranging from school closures, over economic and societal lockdown to vaccination. The project also aims to improve our understanding of infection chains and the dynamics of spread within urban areas such as, e.g., Berlin, as well as in regional and national contexts. The results attracted attention of the federal government, among others, and were widely discussed and commented on in the media.

National alliance for High-Performance Computing. In 2020, the German federal and state governments decided to jointly initiate and found a National High-Performance Computing Alliance (NHR). The NHR alliance will create a nationally coordinated network of high-performance computing centers. Eight centers, one of which is ZIB, were selected through a national competition, evaluating both their scientific as well as HPC competency. The evaluation process involved both a review by the German Research Foundation (DFG) and an independent strategy committee appointed by the Germany's Joint Science Conference. The NHR center at ZIB will receive more than EUR 70 million of funding for the next ten years; see also the feature article on the NHR center.

Amendment to the ZIB Act. On the basis of an intensive debate with relevant parties lasting more than a year, an amendment to the ZIB Act was formulated, discussed in the responsible bodies and passed by the Parliament of the State of Berlin on December 2, 2020. The new law defines ZIB as a research institute for scientific computing, application-oriented mathematics, and high-performance computing with an interface to artificial intelligence. It also entangles ZIB more closely with the partners of the Berlin University Alliance. More details can be found in the article "New ZIB law passed by the parliament of Berlin."

Research Campus MODAL enters its second funding phase. After successful completion of its first funding phase from 2014 to 2020, the Research Campus MODAL was evaluated by an expert review panel and the grand jury of the German Ministry of Research and Education (BMBF). MODAL's proposal for a second phase was granted for the funding period 2020-2025. The grant includes funding of EUR 10 million by the BMBF matched by a considerably larger amount of the participating private companies. MODAL hosts more than 60 researchers from academia and industry under a single umbrella. More details, especially some success stories of public-private partnerships at ZIB, can be found in the feature article on the Research Campus MODAL.

More insights. In addition to the topics already mentioned, this annual report provides insights into a variety of other success stories and gives a general overview of ZIB's organization and key factors for its successful development. For example, the feature article "Research Campus MODAL: Success Stories" outlines what can be achieved regarding research transfer if private partners closely work together with institutions like ZIB. In "NHR@ZIB: Taking the Next Step in HPC," further insight is provided about the creation of the NHR center at ZIB and what it means for the future of high-performance computing at ZIB. The feature articles "Understanding and Modeling Complex Biological Systems" reports on ZIB's recent work on seamlessly integrating modeling approaches and machine learning for enabling new ways for extracting "dynamical laws" from the observation of complex systems on a macroscopic or microscopic level. This overview is complemented by the feature article "Simple Explanations, Sparsity, and Conditional Gradients" which explains why it is often important to find sparse solutions to optimization problems that appear at the core of machine learning methods and how this sparsity requirement can be realized algorithmically. Finally, "Wind under your Wings" reports on recent progress regarding discrete-continuous free flight planning related to the development of hybrid approaches combining discrete optimization algorithms with optimal control methods in order to compute globally optimal free flight trajectories with high accuracy in an efficient way.

Despite all the problems resulting from the COVID pandemic, 2020 was a very successful year for ZIB with key successes that will positively impact on ZIB's development for a long time. Many other very positive developments make us confident that our institute has a bright future. Put differently, ZIB continues to be a place for excellent research and first-rate scientific services and infrastructure.

ORGANIZATION

The Statutes

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

Administrative Bodies

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

President of ZIB
Prof. Dr. Christof Schütte

Vice President
Prof. Dr. Sebastian Pokutta

The Board of Directors was composed in 2020 as follows:

Prof. Dr. Peter Frensch
Vice President, Humboldt-Universität zu Berlin (Chairman)

Prof. Dr. Christian Thomsen
President, Technische Universität Berlin (Vice Chairman)

Prof. Dr. Günter Ziegler
President, Freie Universität Berlin

Thorsten Steinmann
Der Regierende Bürgermeister von Berlin, Senatskanzlei – Wissenschaft und Forschung

Dr. Jürgen Varnhorn
Senatsverwaltung für Wirtschaft, Energie und Betriebe

Prof. Dr. Manfred Hennecke
Bundesanstalt für Materialforschung und -prüfung (BAM)

Thomas Frederking
Helmholtz-Zentrum Berlin für Materialien und Energie (HZB)

Prof. Dr. Heike Graßmann
Max-Delbrück-Centrum für Molekulare Medizin (MDC)

The Board of Directors met on June 3, 2020, and November 16, 2020.

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 to the Scientific Advisory Board:

Prof. Dr. Jörg-Rüdiger Sack (Chairman)
Carleton University, Ottawa, Canada

Prof. Dr. Frauke Liers
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Prof. Dr. Michael Dellnitz
Universität Paderborn, Germany

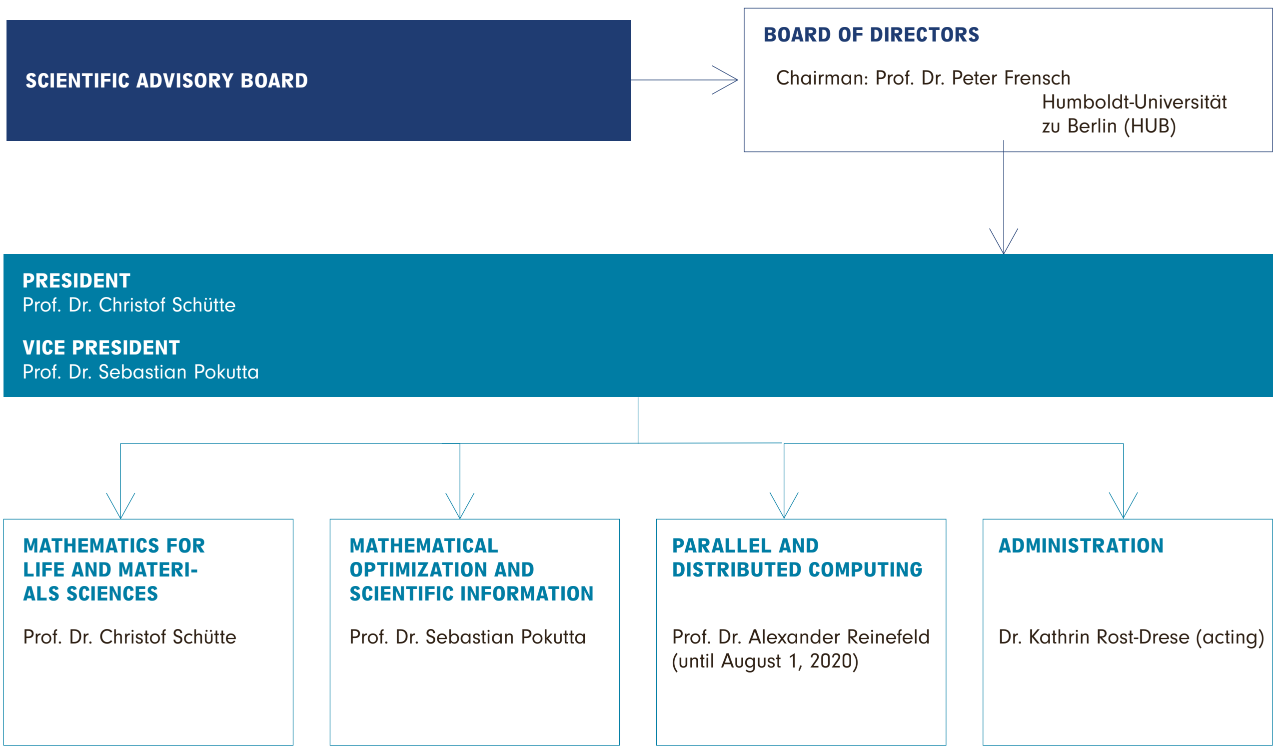
Prof. Dr. Rolf Krause
Université della svizzera italiana, Lugano, Switzerland

Ludger D. Sax
Grid Optimization Europe GmbH

Prof. Dr. Reinhard Schneider
Université du Luxembourg, Luxembourg

Prof. Dr. Dorothea Wagner
Karlsruher Institut für Technologie (KIT), Karlsruhe, Germany (Membership shall be held in abeyance during the continuing service as Chair of the German Science Council)

The Scientific Advisory Board met on July 2, 2020, online.



ZIB STRUCTURE

MATHEMATICS FOR LIFE AND MATERIALS SCIENCES
C. Schütte

- **MODELING AND SIMULATION OF COMPLEX PROCESSES** – M. Weiser
  - COMPUTATIONAL ANATOMY AND PHYSIOLOGY – M. Weiser, S. Zachow
  - COMPUTATIONAL MOLECULAR DESIGN – M. Weber
  - COMPUTATIONAL NANO OPTICS – S. Burger
  - COMPUTATIONAL SYSTEMS BIOLOGY – S. Winkelmann
  - UNCERTAINTY QUANTIFICATION – M. Weiser, T. Sullivan
  - COMPUTATIONAL HUMANITIES – N. Conrad
- **VISUAL AND DATA-CENTRIC COMPUTING** – T. Conrad
  - VISUAL DATA ANALYSIS – D. Daum, H. Hege
  - GEOMETRIC DATA ANALYSIS AND PROCESSING – C. von Tycowicz
  - COMPUTATIONAL DIAGNOSIS AND THERAPY PLANNING – S. Zachow
  - BIOINFORMATICS IN MEDICINE – T. Conrad
  - EXPLAINABLE AI FOR BIOLOGY – V. Sunkara

MATHEMATICAL ALGORITHMIC INTELLIGENCE
S. Pokutta

- **AI IN SOCIETY, SCIENCE, AND TECHNOLOGY (AIS^{2}T)** – S. Pokutta
  - MATHEMATICAL OPTIMIZATION METHODS – A. Gleixner
  - INTERACTIVE OPTIMIZATION AND LEARNING (IOL) – S. Pokutta
  - FAIRNESS, ACCOUNTABILITY, TRANSPARENCY, AND EXPLAINABILITY (FATE) – S. Pokutta
- **APPLIED ALGORITHMIC INTELLIGENCE METHODS (A^{2}IM)** – T. Koch
  - PRESCRIPTIVE SYSTEMS – J. Zittel
  - PREDICTIVE METHODS – M. Petkovic
  - MULTI ENERGY SYSTEMS – J. Zittel
  - SUSTAINABLE ENERGY PLANNING – C. Tawfik
- **NETWORK OPTIMIZATION** – R. Borndörfer
  - HEALTH LOGISTICS – G. Sagnol
  - INDIVIDUAL TRAFFIC (BAG) – E. Swarat
  - MOBILITYLAB – N. Lindner

RESEARCH SERVICE UNIT IT AND DATA SERVICES
C. Schäuble

- BRAIN BERLIN RESEARCH AREA INFORMATION NETWORK – C. Schäuble
- PROCESS MODELING AND MANAGEMENT – C. Schäuble, K. Rost-Drese
- DATA MANAGEMENT SERVICES – S. Prohaska

RESEARCH SERVICE UNIT DIGITAL DATA AND INFORMATION FOR SOCIETY, SCIENCE, AND CULTURE (D^{2}IS^{2}C)

- DIGIS: RESEARCH AND COMPETENCE CENTER DIGITALIZATION BERLIN – A. Müller
- KOBV LIBRARY NETWORK BERLIN-BRANDENBURG – B. Rusch, S. Lohrum

PARALLEL AND DISTRIBUTED COMPUTING
A. Reinefeld

- **DISTRIBUTED ALGORITHMS** – F. Schintke
  - **DISTRIBUTED DATA MANAGEMENT** – F. Schintke
  - **SCALABLE ALGORITHMS** – F. Schintke
  - **MASSIVELY PARALLEL DATA ANALYSIS** – F. Schintke
- **SUPERCOMPUTING** – T. Steinke
  - **HPC CONSULTING** – T. Steinke
  - **HPC SYSTEMS** – C. Schimmel
  - **ALGORITHMS FOR INNOVATIVE ARCHITECTURES** – T. Steinke
  - **DYNAMICS OF COMPLEX MATERIALS** – F. Höfling, T. Kramer

ADMINISTRATION
K. Rost-Drese (acting)

ZIB is structured into four divisions: three scientific divisions and ZIB's administration.

Each of the scientific divisions is composed of two departments that are further subdivided into research groups (darker bluish color) and research service groups (lighter bluish color).

B. Rusch, T. Koch

- **FRIEDRICH-ALTHOFF-KONSORTIUM** – U. Kaminsky
- **DIGITAL PRESERVATION** – W. Peters-Kottig
- **OPEN SCIENCE AND RESEARCH DATA (OSARD)** – W. Dalitz
- **ZIB LIBRARY** – K. Lachmann

LEGEND

- Scientific divisions and departments
- Research groups
- Research service groups
- Research service unit

RESEARCH TO FIGHT COVID-19

The COVID-19 virus has caused a worldwide pandemic with more than 120 million positive cases and more than 2.5 million deaths by March 2021. As long as effective medical treatment and vaccination are not available everywhere, non-pharmaceutical interventions such as social distancing, self-isolation, and quarantine as well as far-reaching shutdowns of economic activity and public life are the only available strategies to prevent the virus from spreading. Researchers at ZIB contributed to mathematical models and simulation tools that are capable of predicting the spread of the infection for different combinations of non-pharmaceutical interventions. Reports based on these tools supported governmental policy decisions on regional and federal level.

In 2020, in a worldwide effort, researchers started to construct a multitude of mathematical models for predicting the spread of the infection. At ZIB, we considered model-based and data-based approaches. Our main project entitled "Model-Based Investigation of School Closures and Other Mitigation Measures for COVID-19" (MODUS-COVID) brought together several research groups from the TU Berlin, Robert Koch Institute, and ZIB. The overall goal of this joint project (funded by the German Ministry of Research and Education (BMBF) from 2020 until 2023) is to investigate the effects of non-pharmaceutical interventions on the infection dynamics of SARS-CoV-2, initially with a special focus on schools closing. The project also aims to improve the understanding of infection chains and the dynamics of spread within urban areas as well as in a regional and nationwide context. In addition to examining schools closing, other measures and combinations of measures to curb the spread of SARS-CoV-2 are examined and model-based policy recommendations are formulated.

The main objective of MODUS-COVID is to build a micro-model in the form of an agent-based model (ABM), together with a pipeline that allows the simulation of a synthetic population with realistic movement patterns, subjection of the synthetic individuals to infection dynamics, and then testing the response of infection dynamics to different interventions to subsequently evaluate the effectiveness of these interventions. The main ABM allows the population of Berlin (approximately 5 million people, 3.6 of which are living inside the city limits) to be described moving about during the day, meeting other agents at work, in school, on public transport, at leisure activities, or when shopping. The mobility side of the model had been worked on for many years before the outbreak of SARS-CoV-2 and is based on cell-phone data [1].

The ABM assumes that an infection may occur when an infectious and a susceptible agent are in the same building or vehicle simultaneously. The transport model used is activity-based, meaning that it contains complete daily activity chains for every agent. These activities have types (for example home, work, leisure, shopping), durations, and locations. The activity chains are available for a typical weekday, a typical Saturday, and a typical Sunday. This means that for every simulated week, the weekday model is run five times and then the Saturday and Sunday models. In these simulations, every time two agents meet and one of them is contagious and the other susceptible, the probability of an infection is computed. This results in a large-scale spatiotemporal model, simulations of which require utilization of a high-performance computing infrastructure. All ABM simulations in the context of MODUS-COVID were performed on the supercomputer operated by ZIB.

ABM models provide microscale simulations of the spread of infection for different intervention strategies. However, their very high computational effort prevents the use of ABMs as the prediction core of a multi-objective optimization scheme that is required for identifying the best intervention strategies: interventions have to meet conflicting requirements where some objectives, such as the minimization of disease-related deaths or the impact on health systems, demand for stronger countermeasures, while others, such as social and economic costs, call for weaker countermeasures. Therefore, finding the optimal compromise of countermeasures requires the solution of a multi-objective optimization problem that is based on accurate prediction of future spread of infection for all combinations of countermeasures under consideration. A strategy for construction and solution of such a multi-objective optimization problem with real-world applicability was presented in [2]. The strategy is based on complementing the ABM by a surrogate macro-model that is much less computationally expensive and can therefore be used in the core of a numerical solver for the optimization problem. The resulting set of optimal compromises between countermeasures (Pareto front) can be computed, and then be discussed as a background for policy decisions.

The ABM developed in MODUS-COVID can also be used to model the spread of the B.1.1.7 mutation and other variants, for investigating the widespread use of rapid tests, and the effects of vaccination strategies. It allows quantitative predictions for differentiated strategy building blocks. The ABM-based tool for simulation-based investigation of infection dynamics developed as part of the project is available under an open-source license and will be available free of charge following the project. The biweekly MODUS-COVID reports that were addressed to the German Ministry of Research and Education (and have also been part of the reporting to the federal government) can be found on www.zib.de.

NEW ZIB LAW PASSED BY THE PARLIAMENT OF

On the basis of an intensive debate with relevant parties lasting more than a year, an amendment to the ZIB Act was formulated, discussed in the relevant bodies and passed by the Parliament of the State of Berlin on December 2, 2020. The new law links ZIB more closely with the partners of the Berlin University Alliance.

ZIB now is defined as a research institute of the State of Berlin with the purpose to promote science and research in the field of scientific computing, application-oriented mathematics and high-performance computing, including related development and research services.

The tasks of ZIB encompass the development of models and algorithms, in order to use computer simulations and optimization methods as well as data-driven procedures to solve problems from natural sciences, engineering, life sciences and medicine as well as the social sciences and humanities. Research is complemented by related development and research services, in particular regarding the operation of high performance research infrastructures. ZIB works in close cooperation with the universities of the state of Berlin and the Charité – Universitätsmedizin Berlin, Berlin's university hospital, in order to meet the research service requirements of the parties involved.

In order to create the right basis for this mission, the new law reconstitutes ZIB's Board of Directors with a different structure. From 2021 on, the Board of Directors is composed of the responsible member of the government of the State of Berlin, the heads of the partner institutions of the Berlin University Alliance (BUA) – the presidents of Berlin's three largest universities, FU, HU, and TU Berlin, the Chief Executive Officer of Charité – as well as one of the directors of the Helmholtz Zentrum Berlin. This composition reflects the importance of the link between ZIB and the BUA partners, and, by this, the importance of ZIB for the future of Berlin's research environment.

In addition, the management of ZIB will be expanded and strengthened. The new law creates the position of an additional vice-president. This position will be filled through a joint professorship with one of BUA's partner institutions, thus further strengthening the connection to BUA. Moreover, the head of ZIB's

BERLIN

administration will take over the role of the director of finance, so that the scientific management can focus more on strategic and scientific matters as well as on ZIB's cooperation with the international scientific community.

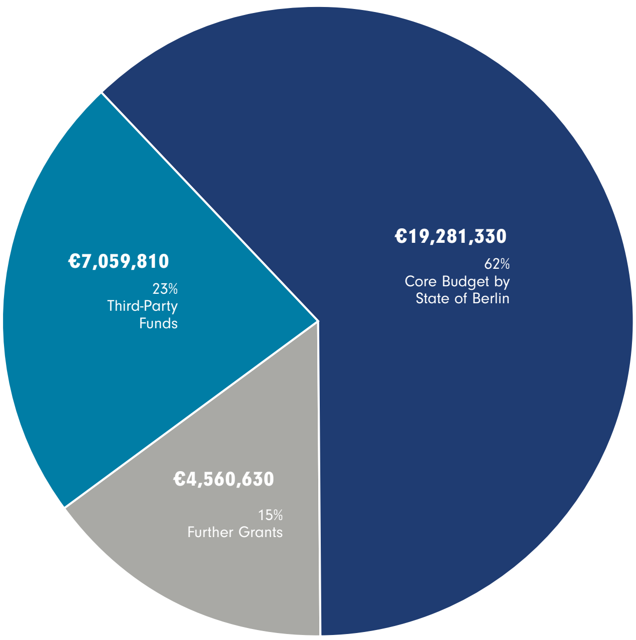
Furthermore, ZIB is given the right to independently enter contracts with scientific and technical personnel, which means a significant simplification of administrative procedures and a strengthening of strategic options regarding recruitment of personnel and fostering junior researchers. Last but not least, ZIB obtains more freedom regarding its economic situation. This leeway is necessary in order to be able to act in a more professional and targeted manner in the area of research-related services in the future. However, it also means that ZIB has to reform parts of its activities, e.g., regarding the expansion and digitization of its financial accounting.

In conclusion, the new law brings quite some changes for ZIB, its members and its environment. It will need some time to implement all these changes appropriately but the result will be an even stronger and more agile institute.

ECONOMIC SITUATION IN 2020

In 2020, the total income of ZIB comprised 30.9 million euros. The main part of this was made available by the Federal State of Berlin as the basic financial stock of ZIB (19.3 million euros) including investments and Berlin's part of the budget of HLRN at ZIB. The second largest part of the budget resulted from third-party funds (7.0 million euros) acquired by ZIB from public funding agencies (mainly DFG and BMBF) and via industrial research projects. This was complemented by a variety of further grants, such as the HLRN budget made available by other German states or the research service budget of KOBV, summing up to almost 4.6 million euros in total.

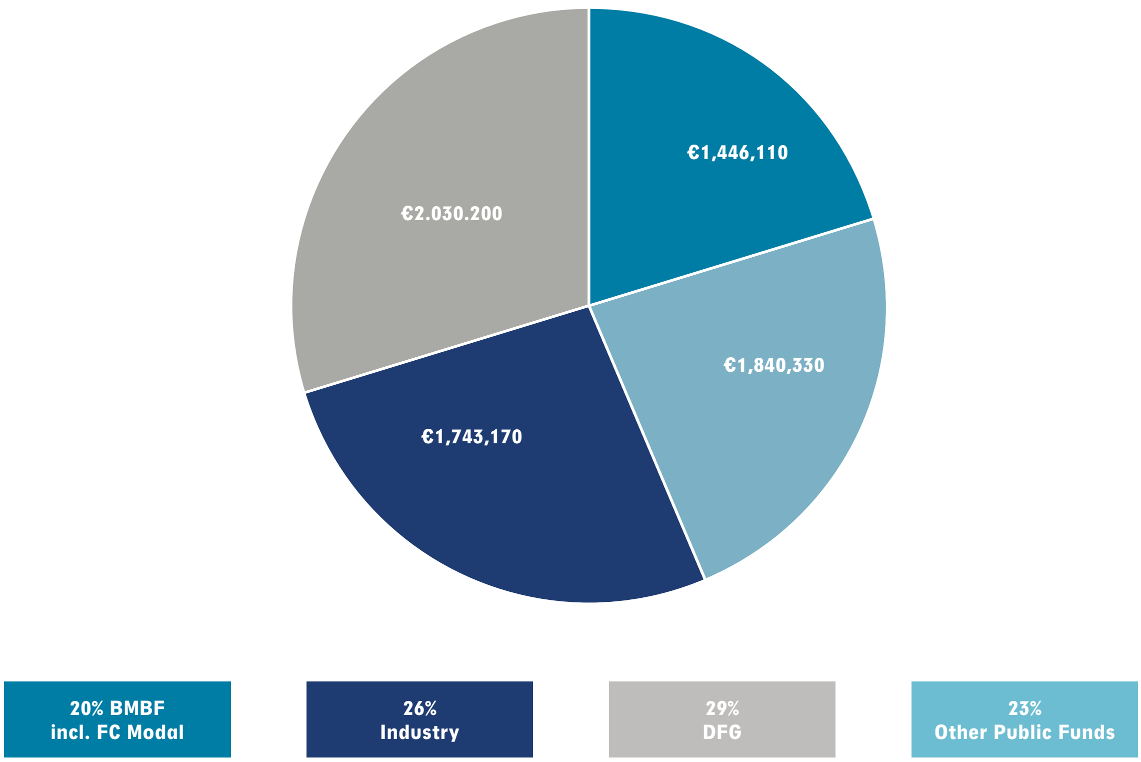
ZIB INCOME



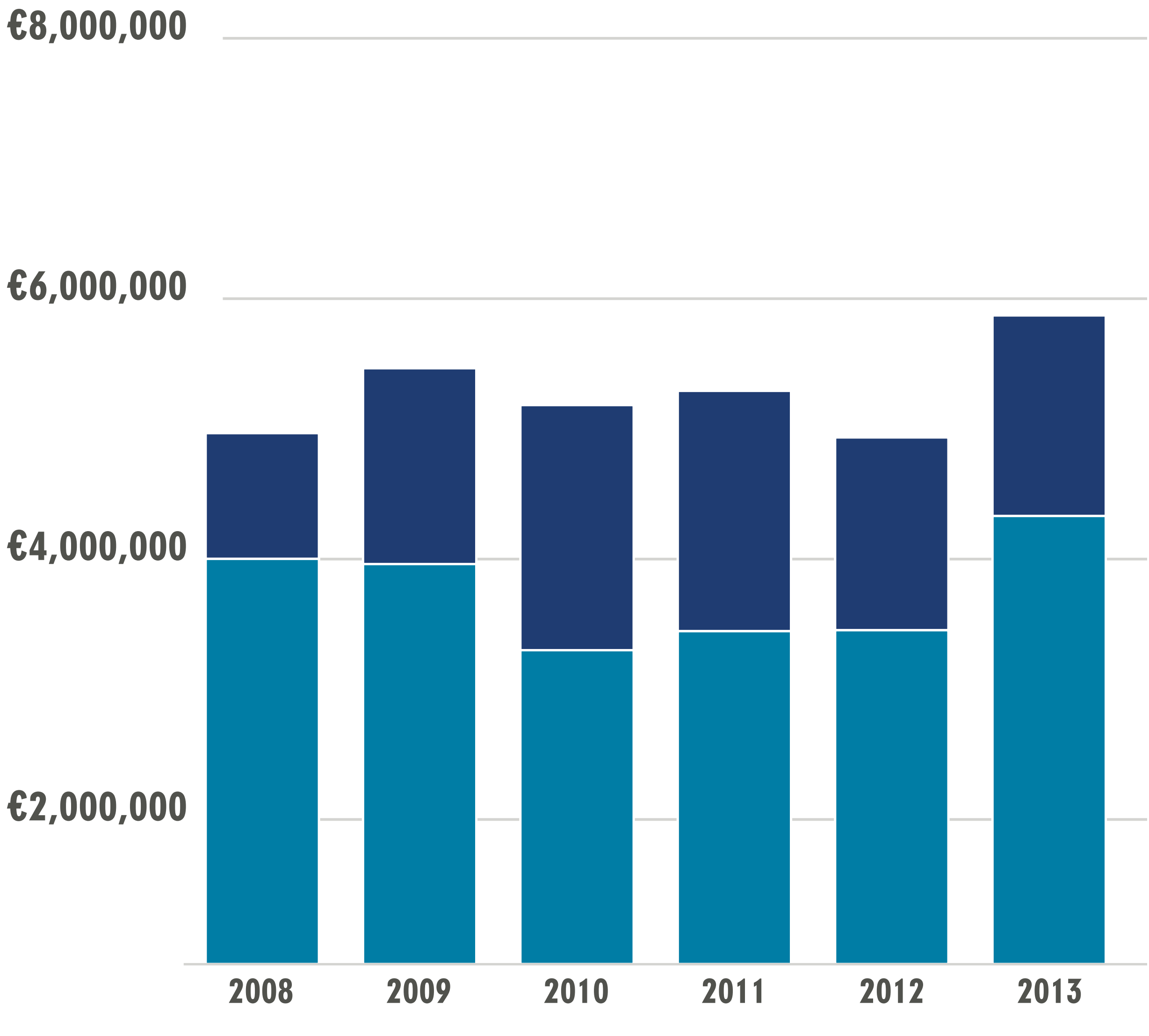
The Zuse Institute Berlin (ZIB) finances its scientific work via three main sources: the basic financial stock of the Federal State of Berlin and third-party funds from public sponsors and those of industrial cooperation contracts.

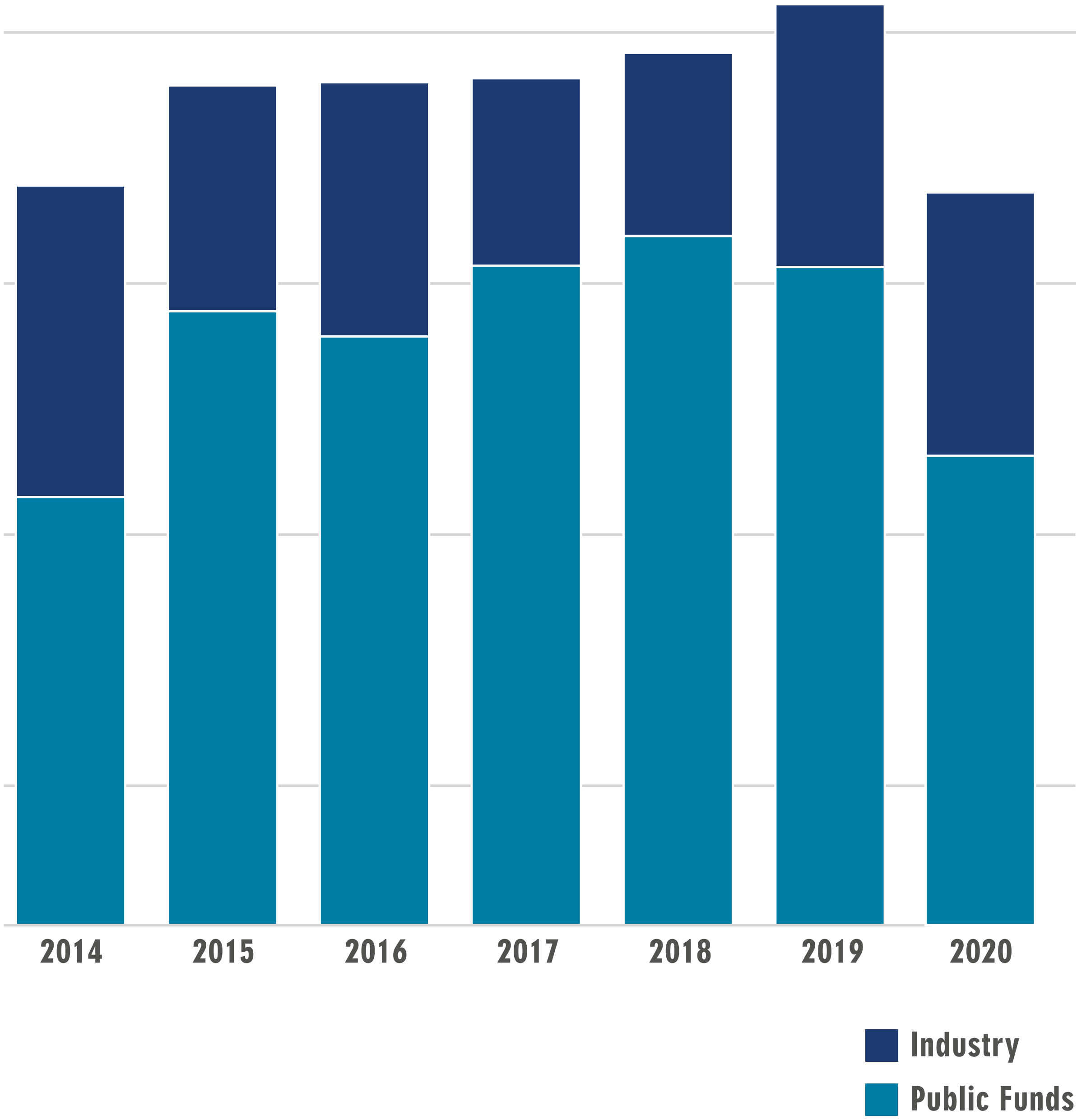
In 2020, ZIB raised third-party funding through a large number of projects. Project-related public third-party funds declined from 6.136 million euros in 2019 to 5.219 million euros in 2020, industrial third-party projects also declined from 2.450 million euros to 1.840 million euros. In total, 7.060 million euros in third-party funding marked a significant decrease, primarily due to the transition from the first to the second funding period of the Research Campus MODAL and the resulting funding gap, as well as reduced expenditures in the area of travel, guests, and conferences.

ZIB THIRD-PARTY FUNDS BY SOURCE



ZIB THIRD-PARTY FUNDS IN EUROS




2014
2015
2016
2017
2018
2019
2020
Industry
Public Funds

Spin-Offs

Computing in Technology GmbH (CIT)
1992 | www.cit-wulkow.de
Mathematical modeling 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, and e-government

Visage Imaging GmbH
(Originating from the ZIB spin-off Visual Concepts GmbH)
1999 | www.visageimaging.com
Advanced visualization solutions for diagnostic imaging

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

bit-side GmbH
2000
Telecommunication applications and visualization

Dres. Löbel, Borndörfer & Weider GbR / LBW Optimization GmbH
2000 | www.lbw-optimization.com
Optimization and consulting in public transport LBW Optimization GmbH was founded in 2017 and is a spin-off of LBW GbR

Lenné 3D GmbH
2005 | www.lenne3d.com
3-D landscape visualization, software development, and services

JCMwave GmbH
2005 | www.jcmwave.com
Simulation software for optical components

onScale solutions GmbH
2006 | www.onscale.de
Software development, consulting, and services for parallel and distributed storage and computing systems

Laubwerk GmbH
2009 | www.laubwerk.com
Construction of digital plant models

1000shapes GmbH
2010 | www.1000shapes.com
Statistical shape analysis

Quobyte Inc.
2013 I www.quobyte.com
Quobyte develops carrier-grade storage software that runs on off-the-shelf hardware

Keylight GmbH
2015 I www.keylight.de
Keylight develops scalable real-time Web services and intuitive apps. The focus is on proximity marketing, iBeacon, and Eddystone for interactive business models

Number of Employees

In 2020, 219 people were employed at ZIB; of these, 134 positions were financed by third-party funds. The number of employees grew in comparison to 2020, mainly because of increase in funding by the State of Berlin and as a consequence of the start of the second funding period of the Research Campus MODAL.

	1/1/2020			1/1/2021		
	Permanent	Temporary	**Total**	Permanent	Temporary	**Total**
MANAGEMENT *without temporary management	3	0	3*	2	0	2*
SCIENTISTS	21	78	99	28	86	114
SERVICE PERSONNEL	41	0	41	41	3	44
KOBV HEADQUARTERS	15	1	16	14	1	15
STUDENTS	0	40	40	0	44	44
Total	80	119	**199**	85	134	**219**

UNDERSTANDING AND MODELING COMPLEX BIOLOGICAL SYSTEMS

Marcus Weber | weber@zib.de | +49-30-8 41 85-189

PS

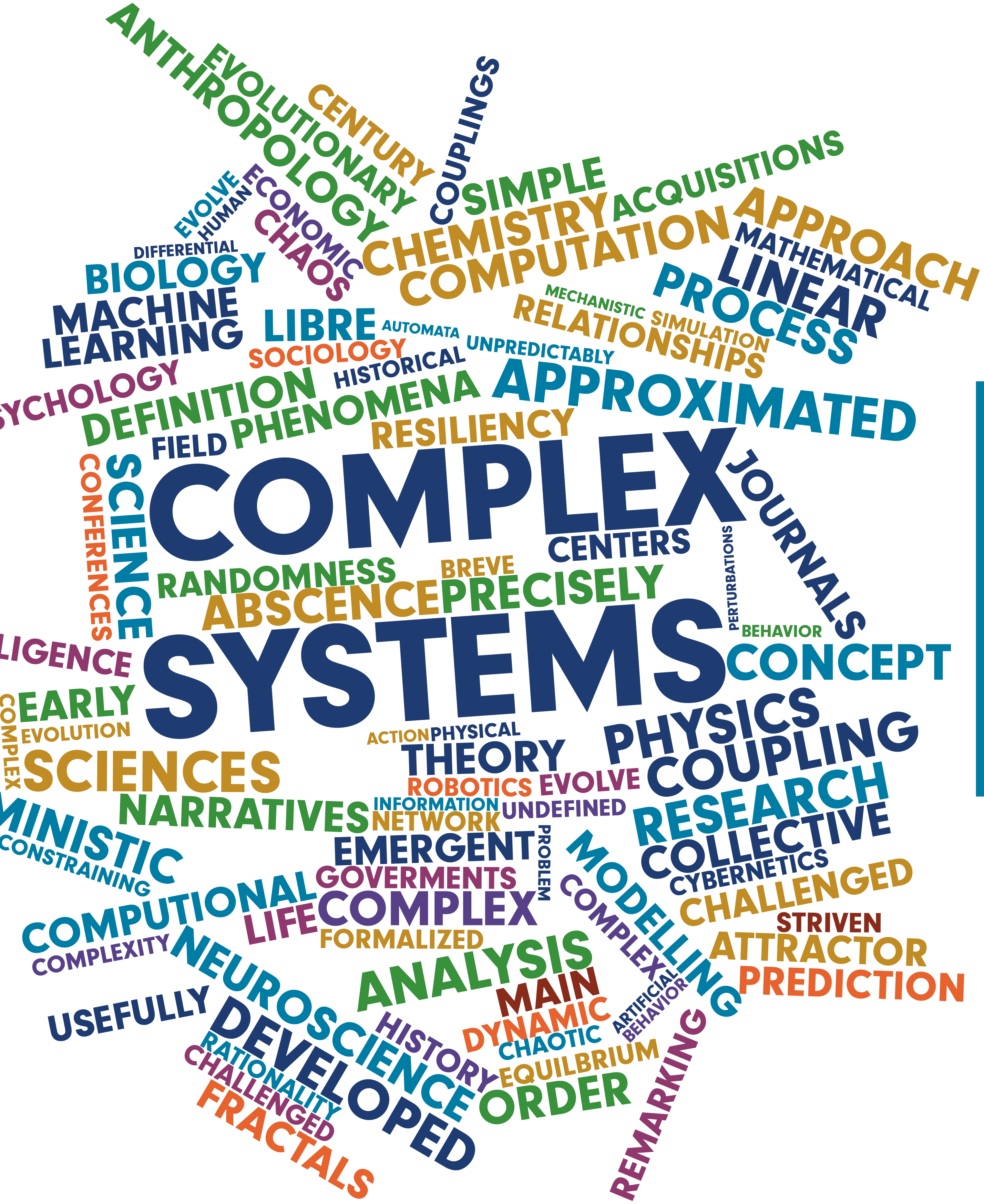
INTEL

DETER

Complex systems are characterized by their multiple scales. To better understand and ultimately allow the control of such complex systems, insights about their underlying laws are needed that can then be used to simulate those systems for validation and hypothesis testing. Researchers at ZIB develop new mathematical methods needed for these tasks. In particular, new approaches based on time-series analysis and machine learning have been researched that enable new ways for extracting "natural laws" from the observation of complex systems on a macroscopic or microscopic level.

Researchers at ZIB have contributed to many topics in this area, such as:

- **Data Analysis:** "Understanding Biological Data: How Machine Learning Can Help"
- **Simulation:** "Speeding up Simulations Using Artificial Intelligence"
- **Complexity Reduction:** "Reducing the Complexity: a Data-Driven Approach for Molecular Systems"
- **Modeling and Optimization:** "Learning Dynamical Laws from Noisy Data"


ANTHROPOLOGY
EVOLUTIONARY
CENTURY
COUPLINGS
SIMPLE
ACQUISITIONS
EVOLVE
HUMAN
ECONOMIC
CHAOS
CHEMISTRY
COMPUTATION
APPROACH
MATHEMATICAL
LINEAR
PROCESS
DIFFERENTIAL
BIOLOGY
MACHINE
LEARNING
LIBRE
AUTOMATA
MECHANISTIC
SIMULATION
RELATIONSHIPS
SOCIOLOGY
HISTORICAL
UNPREDICTABLY
PSYCHOLOGY
DEFINITION
PHENOMENA
APPROXIMATED
RESILIENCY
FIELD
CONFERENCES
SCIENCE
COMPLEX
JOURNALS
CENTERS
PERTURBATIONS
RANDOMNESS
BREVE
ABSCENCE
PRECISELY
BEHAVIOR
INTELLIGENCE
SYSTEMS
CONCEPT
EARLY
EVOLUTION
COMPLEX
SCIENCES
ACTION
PHYSICAL
THEORY
PHYSICS
COUPLING
ROBOTICS
EVOLVE
NARRATIVES
INFORMATION
NETWORK
UNDEFINED
RESEARCH
DETERMINISTIC
CONSTRAINING
EMERGENT
PROBLEM
MODELLING
COLLECTIVE
CYBERNETICS
COMPUTIONAL
GOVERMENTS
COMPLEX
CHALLENGED
LIFE
COMPLEX
STRIVEN
COMPLEXITY
FORMALIZED
ATTRACTOR
ANALYSIS
PREDICTION
NEUROSCIENCE
MAIN
ARTIFICIAL
BEHAVIOR
USEFULLY
DEVELOPED
HISTORY
DYNAMIC
CHAOTIC
RATIONALITY
CHALLENGED
EQUILBRIUM
REMARKING
ORDER
FRACTALS

Understanding Biological Data: How Machine Learning Can Help

Biomedical experiments often lead to large data collections. This could be images from high-resolution microscopy devices, genomic sequences from tumors, or a list of all protein concentrations that are contained in blood samples. Understanding the data and extracting useful information manually is not only tedious but even becomes impossible if the size of the data gets too large. Computer-based algorithms can help to automate and speed up this process. However, one of the main challenges is to create methods that recognize the useful information and ignore the rest. Machine learning-based approaches are a way to achieve this by training an algorithm until it delivers the desired output. Researches at ZIB have contributed to the design of such methods for several applications.

Within the area of *machine learning*, researchers are building and applying algorithms that can improve with experience. While this is quite a normal part of our everyday lives – given that we are humans – it is still surprising to see that a machine can also do the same thing, at least in some specialized areas, such as object and pattern recognition or prediction of future events. In some of these areas, many machines (or algorithms) already outperform humans not only in accuracy but also in terms of speed – owing to the ever-increasing power of new processor technologies.

The two main types of machine learning are supervised and unsupervised learning. Supervised learning is focused on figuring out the relationship between the given data and their respective labels, for example, whether an ECG measurement represents a medical condition known as bradycardia. In this case, the algorithm does not know a priori how to recognize bradycardia – instead, it will learn this from the data. Once this relationship has been learned, the algorithm can suggest (or predict) the label of an unknown instance and with this, might support a medical doctor recognizing heart diseases [Weimann2021]. The main promise of machine learning is: the more data has been used to train such an algorithm, the better the quality of the predicted labels.

The other main type of machine learning are unsupervised algorithms. Here, the data does not have labels. The goal is to find patterns in the data that can then be used to identify sub-groups (or clusters) in the data. In biomedical applications, such cluster-identification algorithms help researchers to better understand biological subtypes based on their genomics profile, which could lead to a better understanding of the underlying mechanisms [Rams2020].

Although more and more biological data is generated these days, still, in most situations, either not enough data or not enough labels are available for a successful training of machine learning algorithms. Most of the time, unfortunately, both options are true, and, on top of that, the available data contains errors and noise. In this situation, solutions are needed that somehow allow the pretraining of the machine learning algorithms on synthetically generated data first, where the quality can be controlled. However, this seems to be a paradox: if the system that should be learned is known, one does not need to learn it. But if the system is unknown, how can data about it be generated? Researchers at ZIB contributed to overcoming this challenge for applications in biomedical data analysis. The key insight is that it is beneficial to pretrain – or initialize – a machine learning algorithm with data that is similar – but not exactly the same – to the target data. This data can be synthetic, e.g., resulting from simulations using a well-established (but not perfect) model for the system at hand. This allows the algorithm to learn the general structure and is then followed by a fine-tuning step using the real data. This combination leads to better results compared to only training on the available real data and has been demonstrated to work very well for proteomics and heart activity (ECG) data [Iravani2019, Weimann2021].

One of the main goals while analyzing complex systems in general and biological systems in particular, is to understand the underlying rules or laws of the dynamic behavior over time. Opposed to the previously described situation, in this case the question is not about predicting labels of an unknown instance (supervised learning) or finding (sub-)groups in large datasets (unsupervised learning), but the idea is to get insights into the mechanisms that can be used to derive a mathematical formulation about a system. Once this can be formulated, simulations can be performed and deeper analysis and hypothesis testing can be carried out. ZIB researchers have presented new approaches that could successfully demonstrate the applicability of these ideas to real-world systems from the life sciences [Melnyk2020, Zhang2019].

Speeding up Simulations Using Artificial Intelligence

Simulations of biological systems play a crucial role in many biomedical applications, such as drug development. Here, for example, one is interested whether a drug is actually able to bind to a receptor long enough to really trigger the desired action, such as pain relief. Unfortunately, detailed and long enough simulations of these kinds of molecular processes are still infeasible these days. Researchers at ZIB have developed new approaches that can circumvent this problem by using methods from artificial intelligence.

Detailed simulations of a biological system – such as the interaction between a drug and a receptor – offer a deep understanding of the processes involved, for example, the time needed for the drug to really trigger the desired function or the on- and off-rates for the drug to bind to the receptor. However, in most cases these simulations are computationally too costly and cannot be performed even on a supercomputer. To overcome this problem, researchers at ZIB have developed alternative approaches that require only a fraction of the computation time and still deliver the required information.

The approach taken is based on so-called partial differential equations and committor functions. Their solution provides all the essential information about the macroscopic behavior of the studied stochastic dynamic microsystem. For each possible state of a microsystem, the key is that a committor function indicates how likely this state is to evolve to a certain target state B, before it returns to the initial state A. By this, for example, the on-rate with which a drug (starting in the unbound state A) binds to the binding pocket (state B) of the receptor can be computed from the committor function by means of transition path theory. The committor function is given by

$$L\chi = 0, \chi = 1 \text{ in the target state B,}$$
$$\chi = 0 \text{ in the initial state A,}$$

where L denotes the infinitesimal generator of the molecular dynamics (MD) of the receptor, drug, and their environment. In the simplest case where MD is modeled by an overdamped Langevin equation, the infinitesimal generator L takes the form of a partial differential operator, which turns the above equation into a partial differential equation (PDE). For more general situations, see [Donati2020]. Similar constructions can be used to compute the off-rates, for example, the frequency of the event that the drug leaves the binding pocket of the receptor, or ligand-protein dissociation rates [Ray2020], leading to linear equations of the form

$$L\chi = \varepsilon_1 \chi - \varepsilon_2 (1 - \chi),$$

the solution of which combines committor functions and mean holding times for appropriate boundary conditions, where ε_1 and ε_2 denote small scalar numbers related to the rates, see [Ernst2020].

Solving PDEs with Artificial Intelligence

The solution of these kinds of PDEs would allow the calculation of the desired macroscopic properties of the full microscopic system. Unfortunately, solving these equations can still be hard or even impossible due the extremely high dimension.

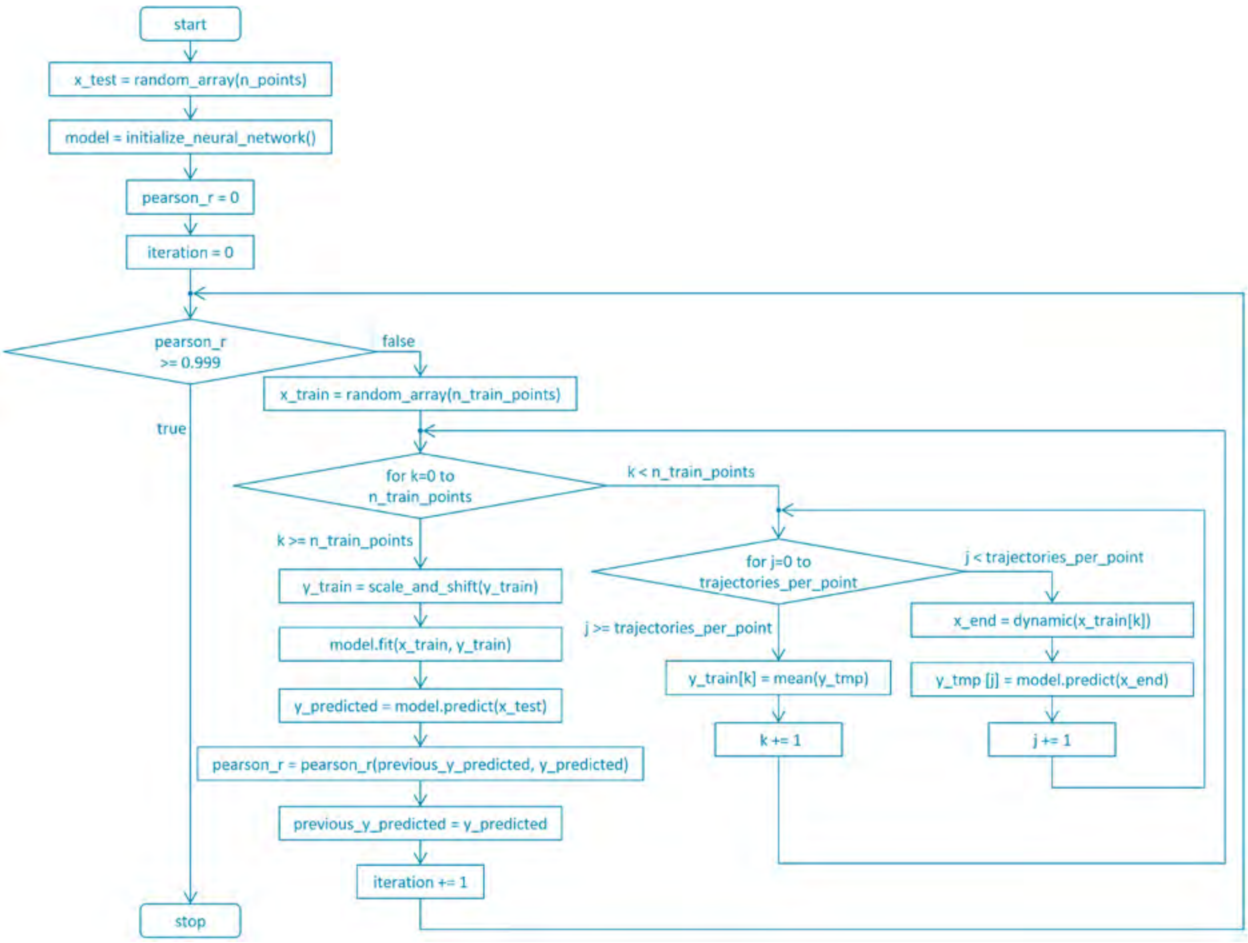


Figure 1: The flowchart of the algorithm ISOKANN, which combines adapted molecular simulations with deep learning.

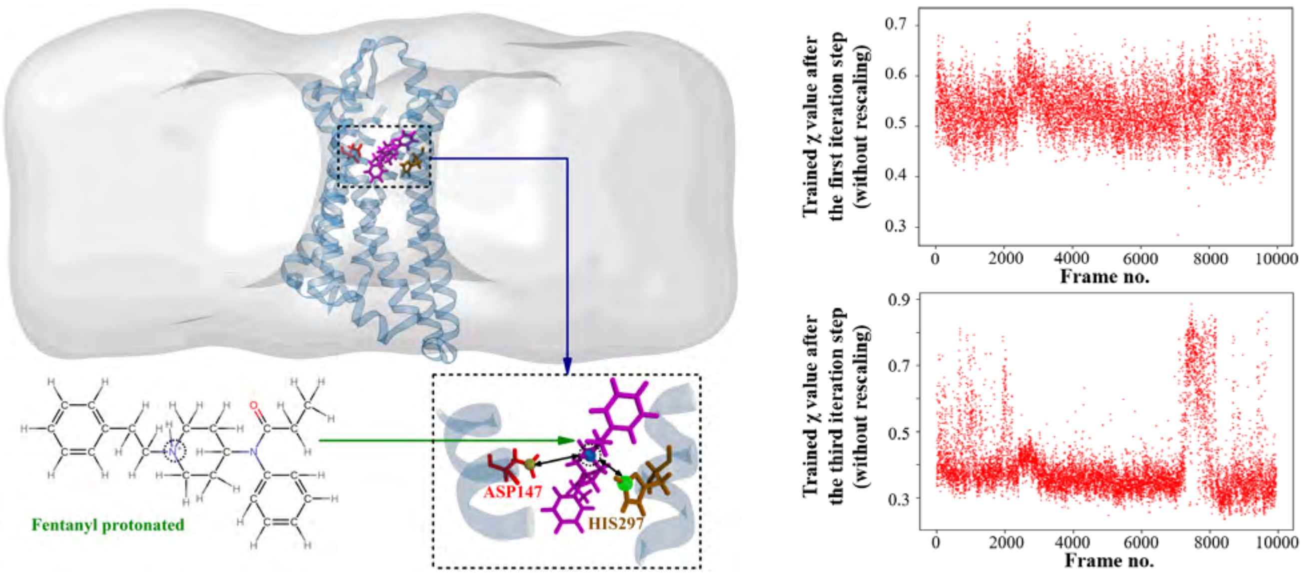


Figure 2: After just two iteration cycles of ISOKANN, the neural network has distinguished between two states of the microsystem and provides the relevant information about the macrosystem.

To tackle this, a new approach was found that turns the original problem of solving the PDE into a so-called functional fixed-point problem. This can be solved by a fixed-point iteration that turns out to be computationally feasible. The developed method, named ISOKANN [Rabben2020], is based on an artificial neural network that is trained to represent the iterated function (see Figure 1). ISOKANN proved to be highly efficient in cases that could not be solved by alternative means. The key ingredient is that the solution of the PDE can be regarded as a continuous classification problem: whether a microstate can be interpreted as "the drug is still at the receptor," or not (see Figure 2). This classification is done very efficiently by the neural network and allows application of adaptivity during the learning phase, where the set of training points can be adjusted in order to put an emphasis onto the "interesting" part of the molecular process (i.e., onto the transition region of the molecular system).

Reducing the Complexity: A Data-Driven Approach for Molecular Systems

Ab initio molecular dynamics allows simulations on spatial and temporal resolutions that are not accessible by traditional experimental methods. Often however, one is not primarily interested in the atomistic detail of a molecular system, but more in the chemically and biologically relevant processes that emerge on longer time and length scales. Gaining insight into these processes requires the careful construction of physically consistent reduced models from simulation data. Central here is the identification of good reaction coordinates (i.e., the “slow” degrees of freedom that form the parameter space of the reduced model).

A GEOMETRICAL VIEW OF REACTION COORDINATES

For given reaction coordinates (RCs), the corresponding reduced dynamical law can be constructed by truncation techniques such as the Mori-Zwanzig formalism. However, the actual consistency with the microscopic dynamics will depend heavily on the initial choice of the RCs. ZIB researchers were involved in formulating a new, general characterization of optimal timescale-preserving RCs [Bittracher2020a]. It revolves around the property of the system’s transition probabilities that with increasing lag time, their dependence on the quickly equilibrating, unimportant degrees of freedom vanish, so that eventually they essentially depend only on the relevant slow degrees of freedom.

Geometrically, this means that the system’s transition densities, considered as points in a certain metric function space, evolve toward a low-dimensional manifold. A parametrization of that manifold corresponds to the desired RCs.

This viewpoint allows a nonparametric representation of the RCs to be learned by applying unsupervised manifold learning methods, such as the diffusion-maps algorithm, to an empirical approximation of the system’s transition densities. This requires simulation data in the form of many short, independent trajectories, whose generation is favored by modern massively parallel computation hardware. Furthermore, the manifold learning step can be enhanced by embedding the data into a reproducing kernel Hilbert space prior to this, which regularizes the problem and minimizes the number of free parameters [Bittracher2020b]. The result is a robust and flexible machine learning algorithm that has been used to identify RCs, transition pathways, and metastable conformations of high-dimensional molecular systems such as peptides and proteins (Figure 3).

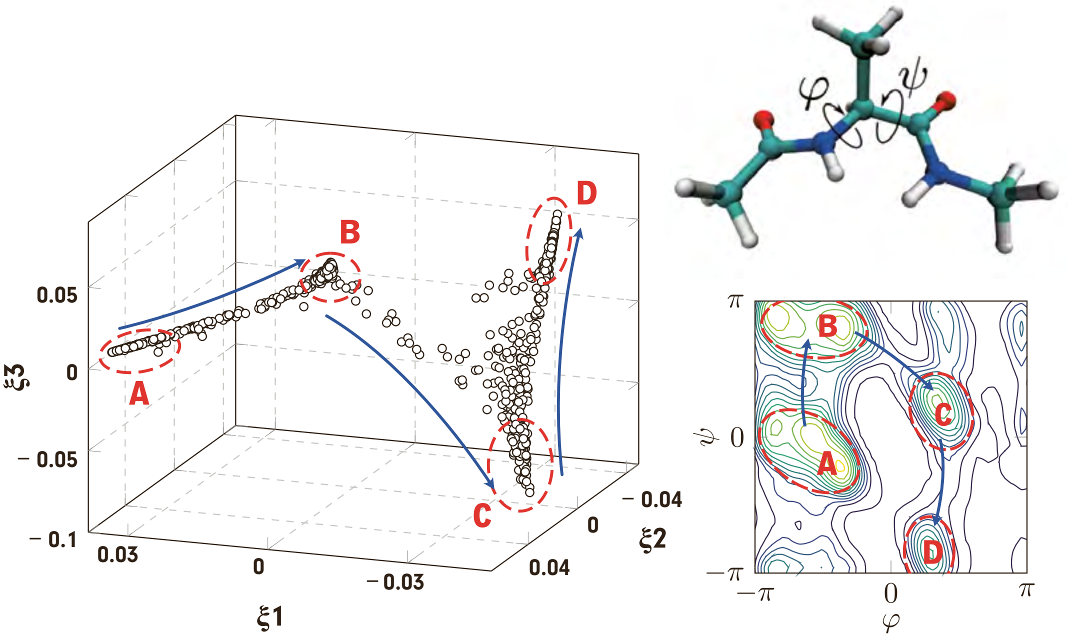


Figure 1: Three-dimensional reaction coordinates evaluated at sampled configurations of the Alanine dipeptide. The edges (arrows) in the low-dimensional manifold structure correspond to transition pathways between local minima of the potential energy surface.

MODEL REDUCTION OF SPARSELY SAMPLED SYSTEMS

Another important aspect of reducible systems is that, due to the existence of an underlying reduced model, the microscopic dynamics cannot become arbitrarily complex. This has the effect that the amount of dynamical data required to approximate the reduced dynamics up to a fixed level of accuracy does in fact not depend on the dimension of the full system, but only on the dimension of the reduced system. For the subclass of discrete Markov chains, ZIB researchers and colleagues were able to formulate a comprehensive theory behind this phenomenon, and explain why accurate reduced Markov jump models of high-dimensional metastable systems can be constructed from vastly sparse trajectory data [Bittracher2021]. This theory also has applications beyond molecular dynamics; for example, they demonstrated how an accurate large-scale model for urban taxi traffic can be constructed from only a small number randomly-chosen measurement points (Figure 4).

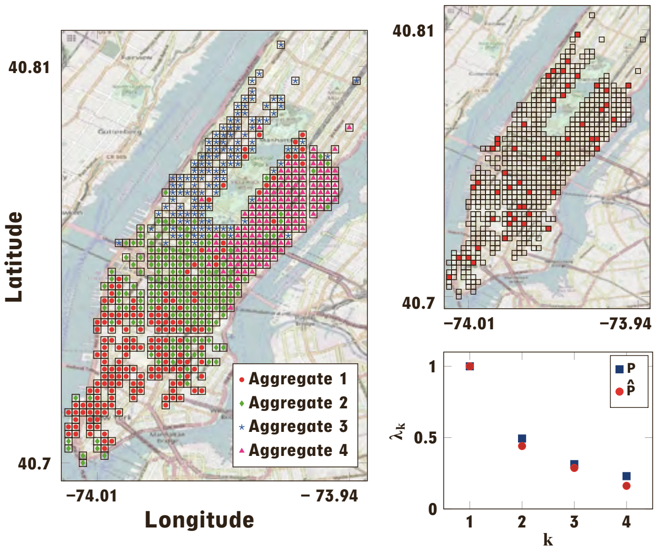


Figure 4: Four-state aggregated Markov model of the taxi traffic on Manhattan Island (left). The aggregates were identified based on traffic measurements at only a small subset of all microscopic states (top right, red boxes). The full and reduced model show excellent agreement (in the form of leading eigenvalues of the respective transition matrices, bottom right)

Learning Dynamical Laws from Noisy Data

Besides classification and prediction, a recent trend in machine learning is to infer the underlying physical laws from data that generated them, which makes use of data not only for concrete application purposes but also for advancing the understanding of complex processes. Since physical laws can usually be formulated as simple algebraic expressions, one way is to look for "simple" dynamical laws predicting the observed data well.

LEARNING DYNAMICAL LAWS

A tractable form of the symbolic regression approach is to build linear combinations of a few atoms selected from a large dictionary of potential contributions. This is particularly attractive for inferring dynamical laws (i.e., right-hand sides in ordinary differential equation systems such as (bio)chemical reactions) due to their uniform structure. Algorithms such as orthogonal matching pursuit (selecting promising terms sequentially) and formulations such as basis pursuit (regression with a sparsity-enhancing L^1 regularization, also known as least absolute shrinkage and selection operator, LASSO) have been proposed. One prominent example is sparse identification of nonlinear dynamics (SINDy), aiming expressly on learning dynamical laws [Brunton2016a].

Recent research has demonstrated that the idea of identifying dynamical laws via sparse solutions to regression and least-squares-like problems is impressively powerful in applications. Since 2016, several advanced variants of this idea have been introduced (e.g., extensions to high-dimensional systems like MANDy [Gelß2019]), combinations with the Mori-Zwanzig formalism, and Taken's embedding theorem leading to extensions of SINDy with memory terms like in sparse identification of nonlinear autoregressive models (SINAR) [Wulkow2020], with essential contributions by ZIB researchers.

In another branch of the literature, especially Koopman operator theory and extended dynamics mode decomposition (EDMD), approximations of dynamical systems by transfer operators have been combined with the idea behind SINDy [Brunton2016b, Klus2018]. Again, this leads to methods for the identification of dynamical laws via solutions to least-squares problems and allows generalizations toward control problems [Brunton2016b]. Recently, this kind of approach has been utilized for sparse identification by means of replacing the transfer operator by the underlying infinitesimal generator [Klus2020]. This technique, termed gEDMD, in particular shows remarkably nice properties in learning sparse representations of stochastic dynamics (SDEs) such as agent-based models, see Figure 5.

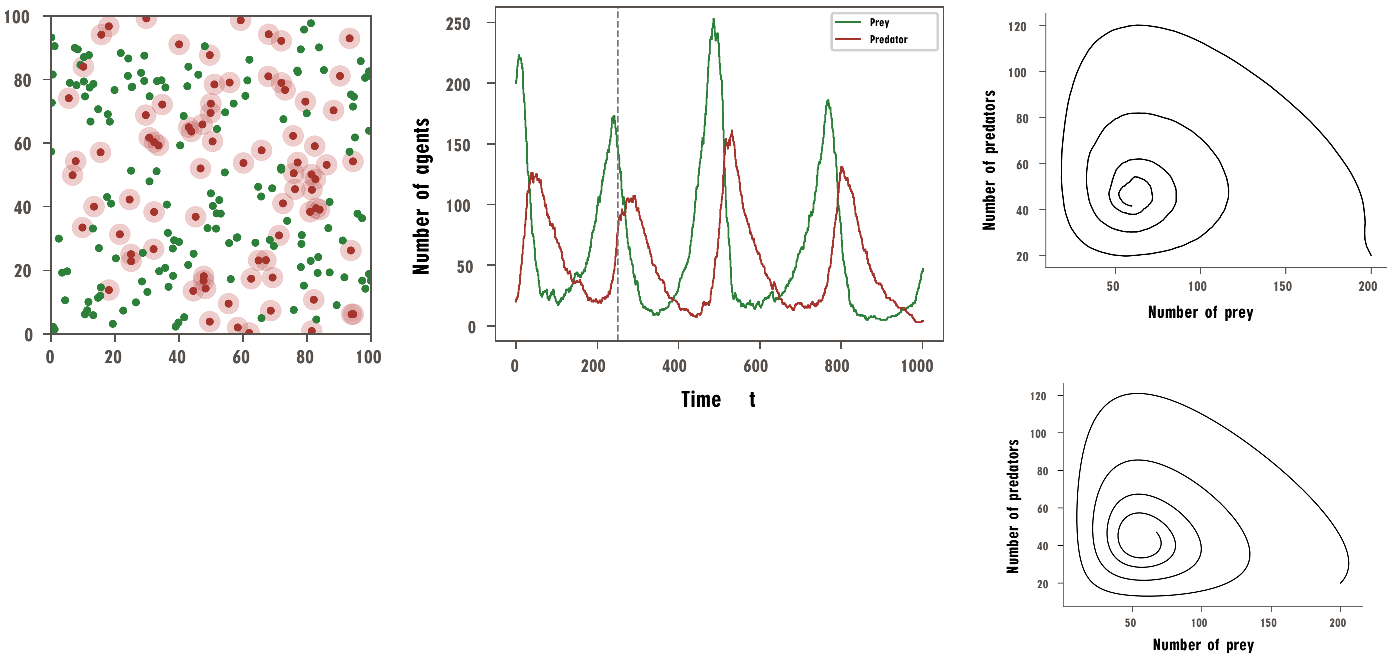


Figure 5: Learning the collective dynamics of predator-prey systems.
Left panel: Snapshot of the spatial state of the predator-prey system with 250 individuals at time t = 250. Predators (red dots) and prey (green) move randomly in space, if prey is within some radius (light red) of the predator, it is eaten with a certain probability and the predator reproduces; surviving prey always reproduce.
Middle panel: Evolution of the collective numbers of predators and prey in the system with time (sample trajectory). Time series of such sample trajectories are used as data for learning the laws of the collective dynamics via gEDMD.
Right panel: Phase diagrams of the learned dynamics (below) and phase diagram coming from a sample trajectory.
Data from Jan-Hendrik Niemann, Stefan Klus, and Christof Schütte (2020) Data-driven model reduction of agent-based systems using the Koopman generator, submitted to PLOS one.

All of these algorithms (see Figure 6 for a schematic representation) require an efficient and robust solver of the underlying sparse optimization problem, mostly one in a basis pursuit form. Until recently, the dominant trend in the literature was to apply off-the-shelf optimization algorithms like LASSO. In many cases, the inherent limitations prevented the successful application to realistic data.

	linear	**nonlinear, non-sparse**	**nonlinear, sparse**
Markovian	Dynamic Mode Decomposition (DMD)	Extended Dynamic Mode Decomposition (EDMD)	Sparse Identification of Nonlinear Dynamics (SINDy)
			Generator-based Extended Dynamic Mode Decomposition (gEDMD)
Non-Markovian	Autoregressive models and Hankel-DMD		Sparse Identification of Nonlinear Autoregressive Models (SINAR)

Figure 6: Methods for learning dynamical laws from data.

NOISY DATA

A challenge to be faced by SINDy-like or EDMD-based algorithms is inevitable noise in real-world data. Many available algorithms produce much less sparse results in the presence of noise – up to completely dense solutions. While this does not necessarily affect prediction accuracy, the reliability of insight in physical laws extracted from these solutions suffers tremendously. Thus, we set out to develop an algorithmic approach that is robust with respect to noise.

Conditional gradient (CG) methods are first-order sequential linear programming algorithms for convex optimization problems, with the compelling feature of only requiring the solution of linear programs as a basic building block, as opposed to projections onto the admissible set, which are in general much harder to solve. In the identification of nonlinear dynamical laws, this is the case in particular if additional equality constraints such as mass preservation need to be included.

Such CG methods have been applied to the basis pursuit formulation of learning nonlinear dynamical laws, leading to a method dubbed CINDy (conditional gradients-based identification of nonlinear dynamics) as an homage to SINDy [Carderera2021]. For more details, see the feature article "Simple Explanations, Sparsity, and Conditional Gradients."

WIND UNDER

Martin Weiser | weiser@zib.de | +49 30 841 85 170

The current system of routing flights on predefined airway networks does not harvest the benefits of exploiting tailwinds in an optimal way. Dropping the restriction to airways allows fuel to be saved and reduces CO_2 emissions – an important driver in view of growing numbers of air traffic.

Established graph-based routing algorithms face considerable computational challenges in this setting. At ZIB, we develop and analyze hybrid approaches combining discrete optimization algorithms with optimal control methods in order to compute globally optimal free-flight trajectories to high accuracy in an efficient way.

YOUR WINGS

Discrete-Continuous Free Flight Planning

The Future Sky

At the dawn of aviation, the sky was free, but navigation a problem. Apart from using a compass, the sun, moon, and stars, only landmarks and various sorts of beacons, vision and later radio provided orientation. Routes were set up between them, soon stacked into vertical flight levels, and monitored by air traffic control. In this way, today's airway network evolved – a virtual 3D graph of 100,000 "waypoints" and 600,000–700,000 "segments" on each of 20+ flight levels, on which all air traffic takes place (see Figure 1 for an example). A notable exception was the oceans, where free flight was always possible, and very well exploited to cut short or to take advantage of favorable wind conditions.

At a long-term exponential growth of 4.4% per year [Oxley, Jain 2015], congestion of the airway network was unavoidable, and started to become more and more of a problem for at least a decade. On the other hand, aircraft navigation equipment became more and more sophisticated, reducing the need for centralized guidance.

Can one simply abandon the airway network to fly free like a bird, with a tailwind propelling you forward? This would eliminate congestion and detours, and save time, fuel, and emissions – a five-fold-win situation. It therefore doesn't come as a surprise that the EU's Single European Sky program has exactly this vision, and there are similar initiatives all over the world.

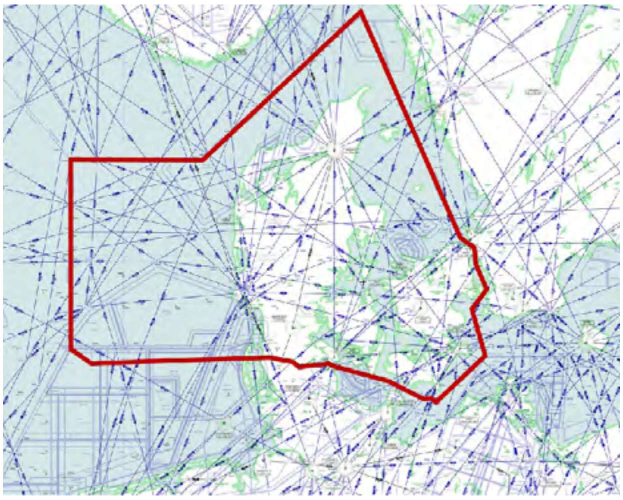


Figure 1: Danish airway network.

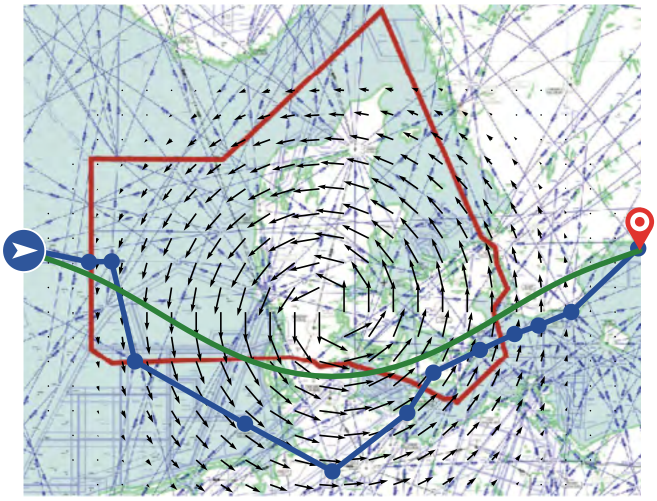


Figure 2: Restriction to the graph leads to suboptimal routes (blue) compared to the continuous optimum (green).

How much can one gain? Practical experience, recent studies [Wells et al. 2021], and our own numerical examples indicate potential savings in fuel and CO_2 emission of about 2%, and up to 10% in extreme cases. As an example, consider Figure 2: the optimal free-flight route (in green) is not only faster, but also smoother in operation, as the comparison of the heading angles of both trajectories in Figure 3 shows.

To harvest these benefits, flight-planning tools are needed that can find the quickest, most fuel efficient, etc. free-flight route taking into account the weather conditions – and a number of other side constraints such as traffic rules – in reasonable time.

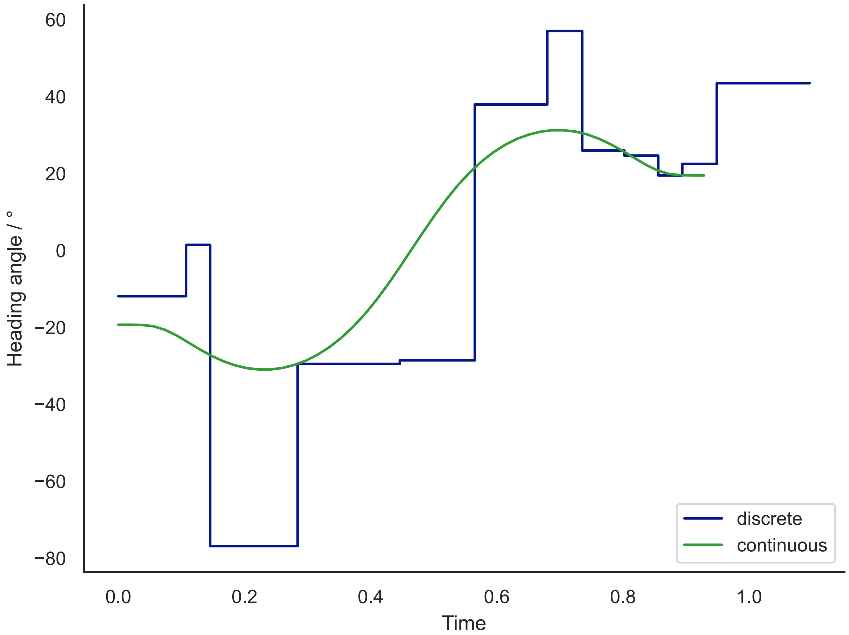


Figure 3: The heading angle shows that the discrete route tends to "zigzag" around the continuous optimum.

Discrete Flight Planning

The current state of the art in flight planning is marked by discrete dynamic programming algorithms that are based on superfast shortest-path algorithms. These methods exploit the characteristics of the problem to derive powerful lower bounds that effectively prune the search space, such that globally optimal solutions can be computed to any desired degree of accuracy, with respect to several objectives, and subject to complex side constraints, including route availability, terrain clearance, alternate airport rules, etc. For example, bounds from so-called super-optimal winds, developed at ZIB in collaboration with Lufthansa Systems [Blanco et al., 2016] can be used to significantly tighten traditional A* bounds, which in turn already do much better than a pure Dijkstra approach, as illustrated in Figure 4. A striking example that stresses the importance of the wind is shown in Figure 5. The search space is subdivided into two routes, a northern direct route, and a longer southern corridor, which can take advantage of strong jet-stream winds, and, despite a detour of 1,250 km, turns out to save 45 minutes of flight time, 5.5 tons of fuel, 18.8 tons of CO_2 and $1,100 of overflight fees.

For some time now, gradually expanding free-flight opportunities have been successfully approximated by graph densification (see Figure 6). However, this approach does not scale well, such that, at some point, the problem becomes intractable.

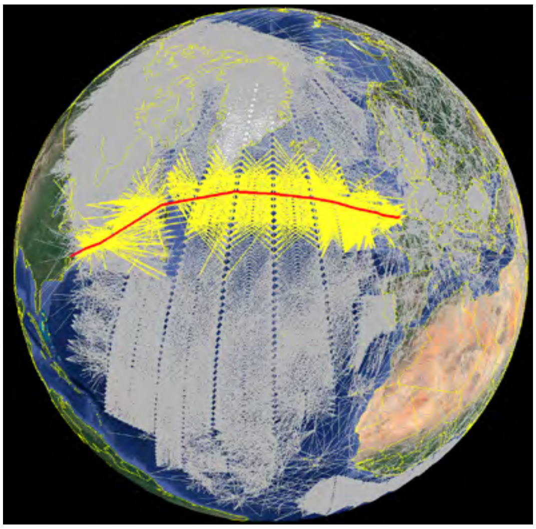


Figure 4: Search spaces of Dijkstra's algorithm (gray) and A* (yellow).

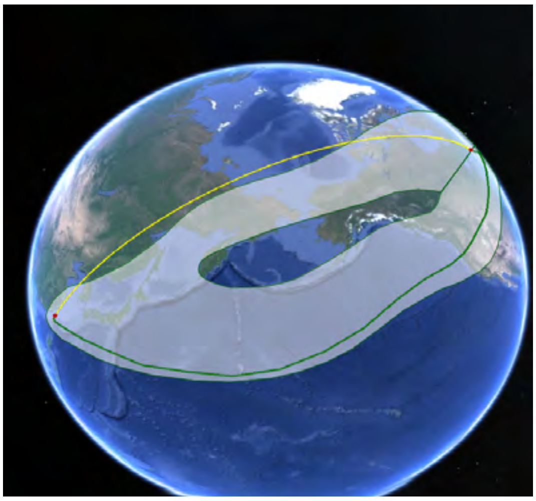


Figure 5: The search space is divided following two promising strategies. Direct connection (top) or a detour while making use of the strong jet stream (bottom).

Indeed, the run time of discrete shortest-path algorithms depends on the density of the underlying network. On sparse graphs, shortest paths can be computed extremely fast by one of several domain-specific superfast algorithms. These all have a worst-case time complexity of $O(|E| + |V| \log|V|)$, where $|E|$ and $|V|$ are the number of edges and vertices, respectively. The airway network, like a planar graph, even has a favorably low number of edges: $|E|=O(|V|)$, such that the overall runtime becomes a mere $O(|V| \log|V|)$, even without preprocessing. Naive free-flight networks, however, have a considerably larger number of vertices and $|E|=O(|V|^2)$ edges (see again Figure 6). This is an enormous difference, which at some point turns the computation of globally optimal shortest paths at high resolution and accuracy into an impossibility.

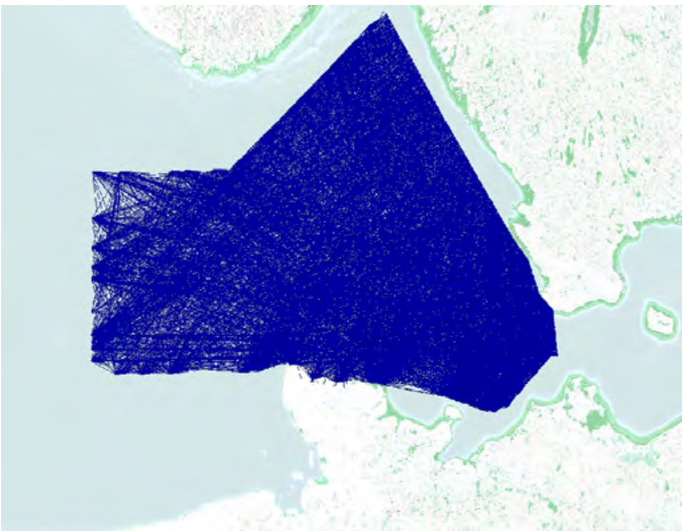


Figure 6: Free-flight approximation of the Danish airspace by using a denser graph.

Continuous-Flight Planning

Continuous methods of optimal control, which have been studied for a century, are an algorithmic alternative. Instead of covering the airspace by a graph, they only discretize the trajectory itself, exploiting necessary optimality conditions to move the waypoints continuously in space (see Figure 7). Established methods based on Pontryagin's maximum principle or direct collocation approximate the trajectories by higher-order polynomials, and thus translate the optimality conditions into large, sparse, and structured equation systems to be solved by Newton's method. This allows much coarser discretizations to achieve the same accuracy as graph-based approaches, and, combined with fast quadratic convergence, highly accurate free-flight trajectories to be computed in a much shorter time frame.

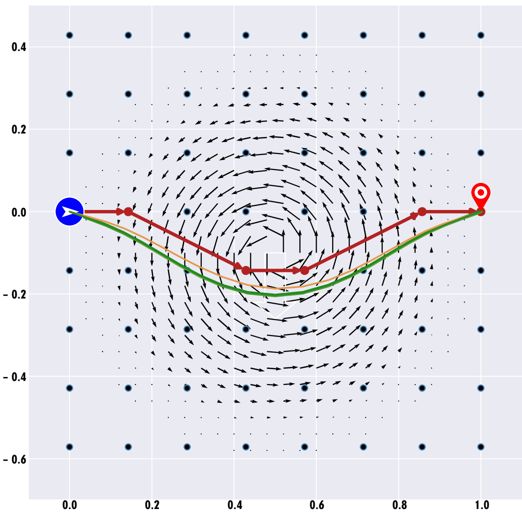


Figure 7: Newton iterates of a discretized trajectory converging quadratically toward the optimal solution.

The downside of that approach is that these methods only converge to some nearby local optimum, if at all, and thus provide no guarantee of global optimality (see Figure 8). Thus, in itself, optimal control approaches are insufficient for free-flight planning.

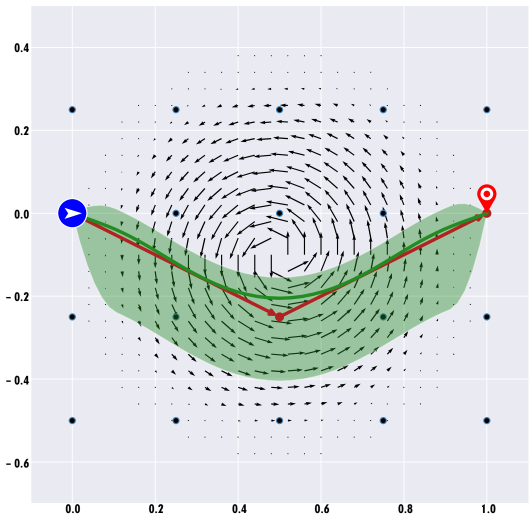
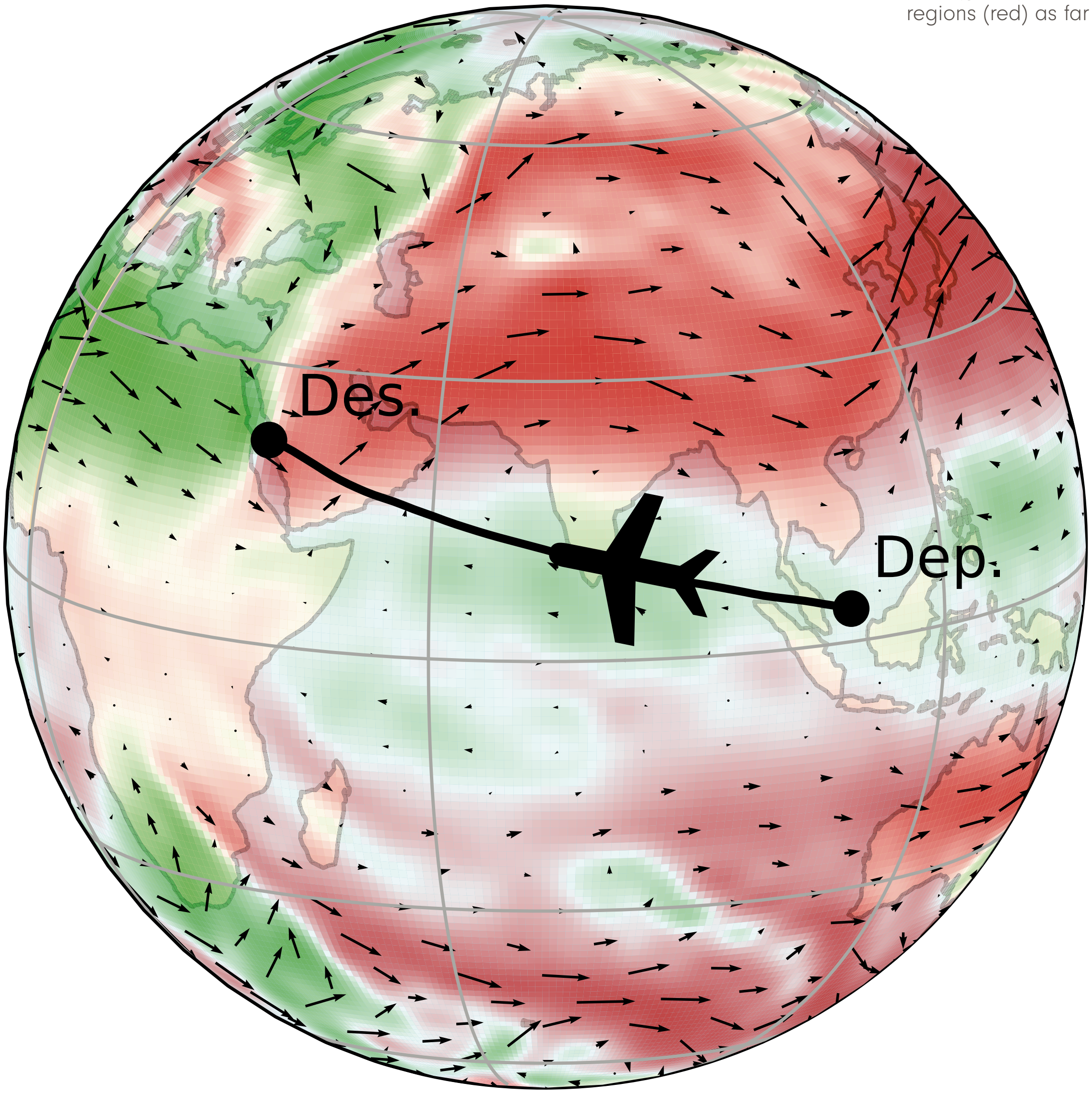


Figure 8: Empirical domain of convergence to the globally optimal trajectory. Starting Newton's method from an initial trajectory within the green region, it converges to the corresponding optimal solution.

Figure 9: The optimal trajectory from Brunei to Jeddah exploits tailwind (green) and avoids headwind regions (red) as far as possible.



Can discrete and continuous methods be merged into an approach that combines their strengths and eliminates their weaknesses, resulting in a method that provides, at the same time, global optimality and rapid convergence to an optimal free-flight trajectory at any desired accuracy? Exactly this is the aim of a new method that is currently developed at ZIB: we recently proposed a novel hybrid algorithm called DisCOptER [Borndörfer/Danecker/Weiser2021].

A Hybrid Algorithm

While discrete, graph-based approaches efficiently approximate global optima at low to medium accuracy, continuous optimal control methods excel at fast local convergence to high accuracy. The hybrid DisCOptER method exploits these properties in a two-stage approach. In a nutshell, it starts an optimal control algorithm from a suitable path provided by a discrete dynamic programming method.

First, an artificial, fully connected graph is created (cf. Figure 10, blue dots) that contains origin and destination vertices. Then, a shortest path on this graph is calculated (red path).

Finally, this shortest path is used as an initial guess for the following refinement stage of the algorithm. Here, an optimal control approach involving some variant of Newton's method is used to calculate a highly accurate solution (green trajectory). Provided that the first guess is within the radius of convergence of the actual continuous optimum (green area), Newton's method converges extremely fast.

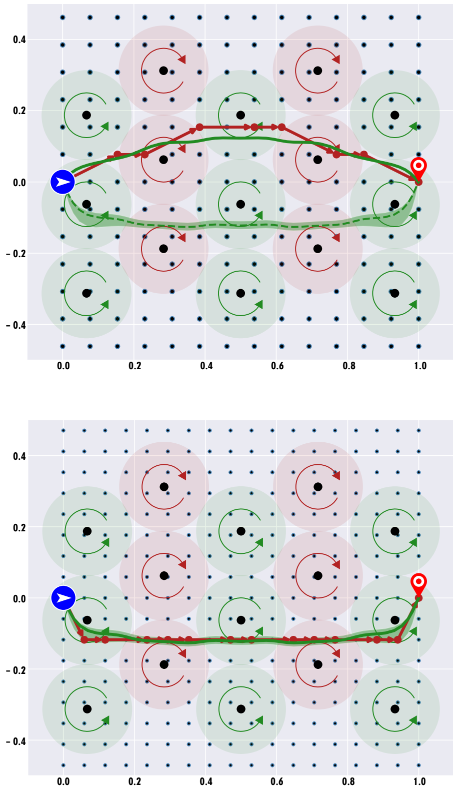


Figure 10: Continuous solution found if initialized with different discrete routes.

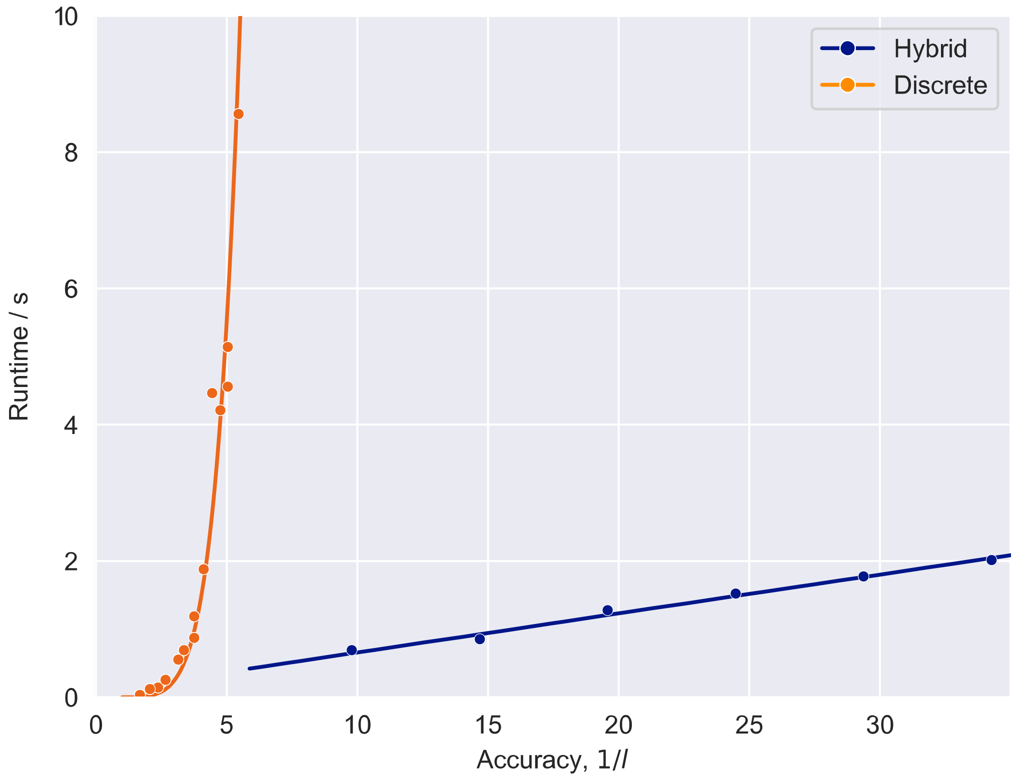


Figure 11: Time complexity of the purely discrete approach $(O(1/l^6))$ vs. the hybrid algorithm $(O(1/l))$.

For the second stage to converge reliably to the global optimum in the vicinity of the path provided by the first stage, this path must be sufficiently close to the global optimum, more precisely, within the domain of convergence of Newton's method. This can be ensured by making the first-stage graph sufficiently dense. The required graph resolution depends only on external factors like the wind conditions, but not on the desired accuracy, which only affects the trajectory discretization in the second stage.

It turns out that in practically relevant examples, the required graph density is rather low (again, see Figure 10). Consequently, the discrete shortest path can be found quickly, the problematic asymptotic runtime behavior of the discrete graph search is circumvented, and the overall time complexity is governed by the asymptotically efficient optimal control stage (cf. Figure 11, $1/l$ is a measure for the solution accuracy). But can we decide within the algorithm how dense the graph needs to be in a concrete case?

Error Estimates for Discrete Paths

For free-flight trajectory planning, the graph used either stand-alone for path planning or as the first stage in a hybrid algorithm needs to be sufficiently fine for the discrete optimal path to be accurate enough or to lie within the convergence domain of Newton's method, respectively. To ensure this, a priori error estimates for optimal flight paths in locally densely connected graphs have been derived [Borndörfer/Danecker/Weiser2021-b]. (h,l)-dense graphs are those where the h-balls around vertices cover the plane, and all vertices with a distance of $l+2h$ at most are connected by an edge (see Figure 12).

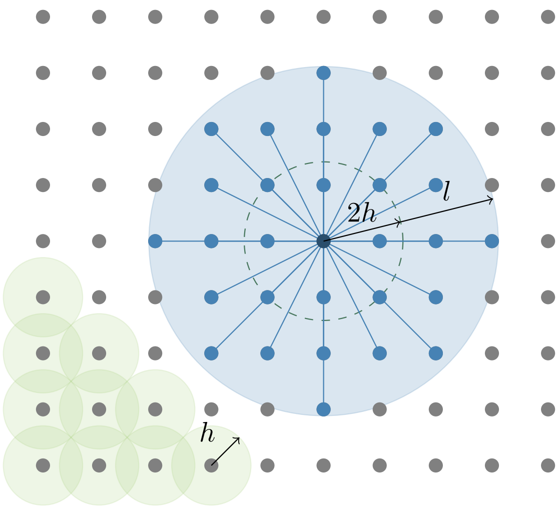
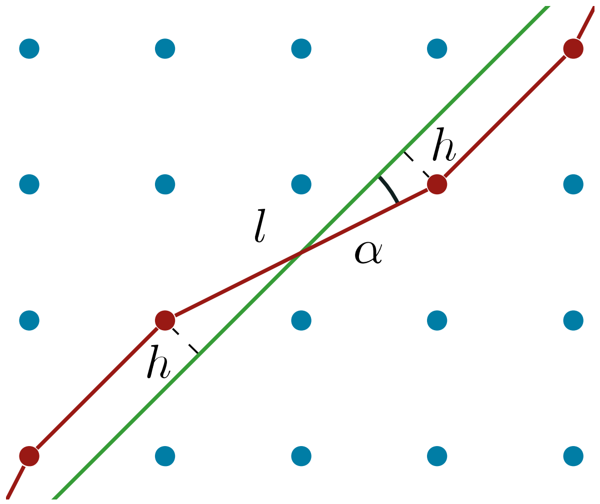


Figure 12: (h,l)-dense graph.

Following the standard paradigm of a priori error estimates for finite element discretizations [Deuflhard/Weiser, 2020], the distance of an optimal continuous path to its interpolate in the graph is bounded, in terms of the flight duration, by bounding the second derivative of the objective with respect to path deviations. The structure of (h,l)-dense graphs guarantees the existence of a reasonable interpolant. Since the optimal discrete path in the graph is not worse than the considered interpolant, this provides an error bound for the discrete solution in terms of vertex density h, local connectivity length l, and magnitudes of wind speed and its spatial derivatives.

The excess in flight duration is due to spatial displacement of the trajectory on one hand, and zigzagging due to limited angular resolution on the other hand. The latter is bounded by l/h, whereas the former is controlled by h (see Figure 13). Based on this distinction, the error estimate allows the selection of a theoretically optimal relation of vertex density h and connectivity radius l, which turns out to be $h = O(l^2)$. The resulting error bound guarantees a convergence of optimal discrete paths toward the continuous optimum in terms of flight duration at a rate of $O(l^2)$.



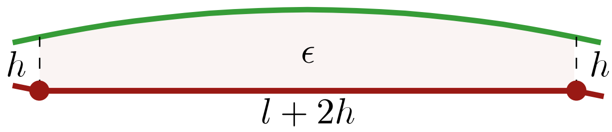


Figure 13: Angular error (l) and distance error (r) between a discrete (red) and a continuous route (green).

As usual, the a priori error bounds are reliable but far from sharp. Nevertheless, they capture the numerically observed convergence order accurately (see Figure 14). This provides the basis on which efficient a posteriori error estimators can be designed, which can actually guide the graph refinement to the desired accuracy and allows implementing efficient and reliable hybrid algorithms.

DisCOptER is a first example of a new type of discrete-continuous algorithms that combine global optimality with fast local convergence. The underlying principle is not limited to shortest path problems, but is general enough to be applicable to many other problems of similar nature.

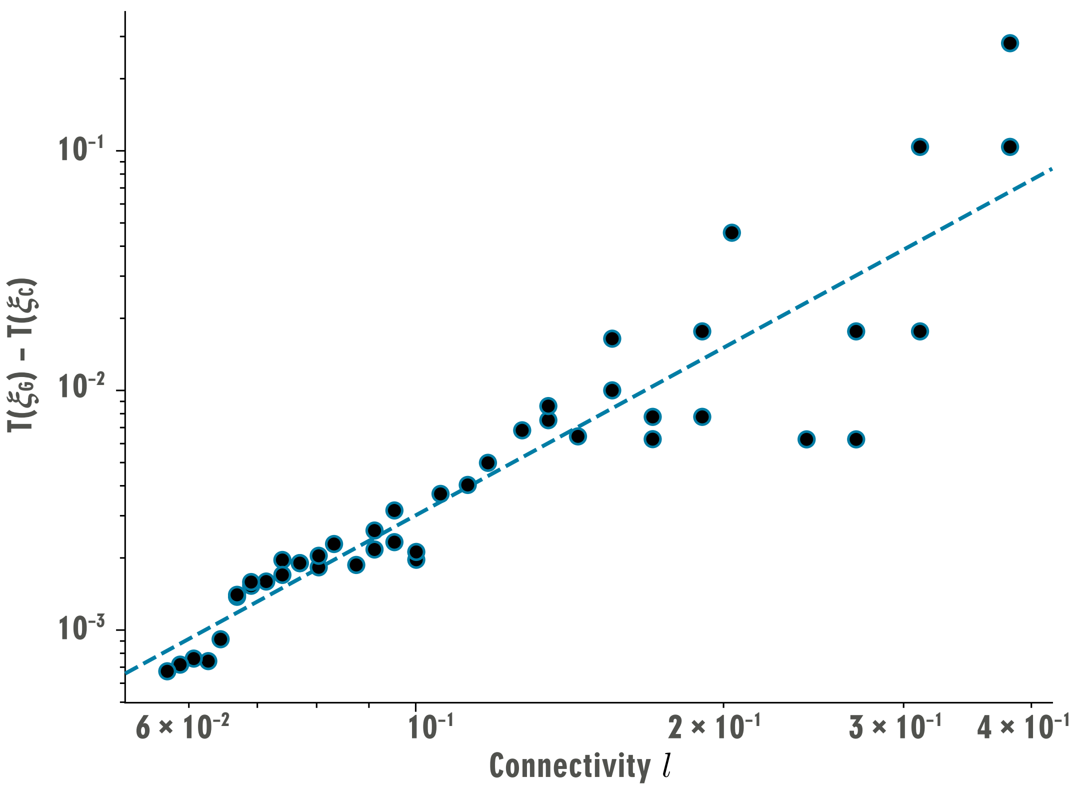
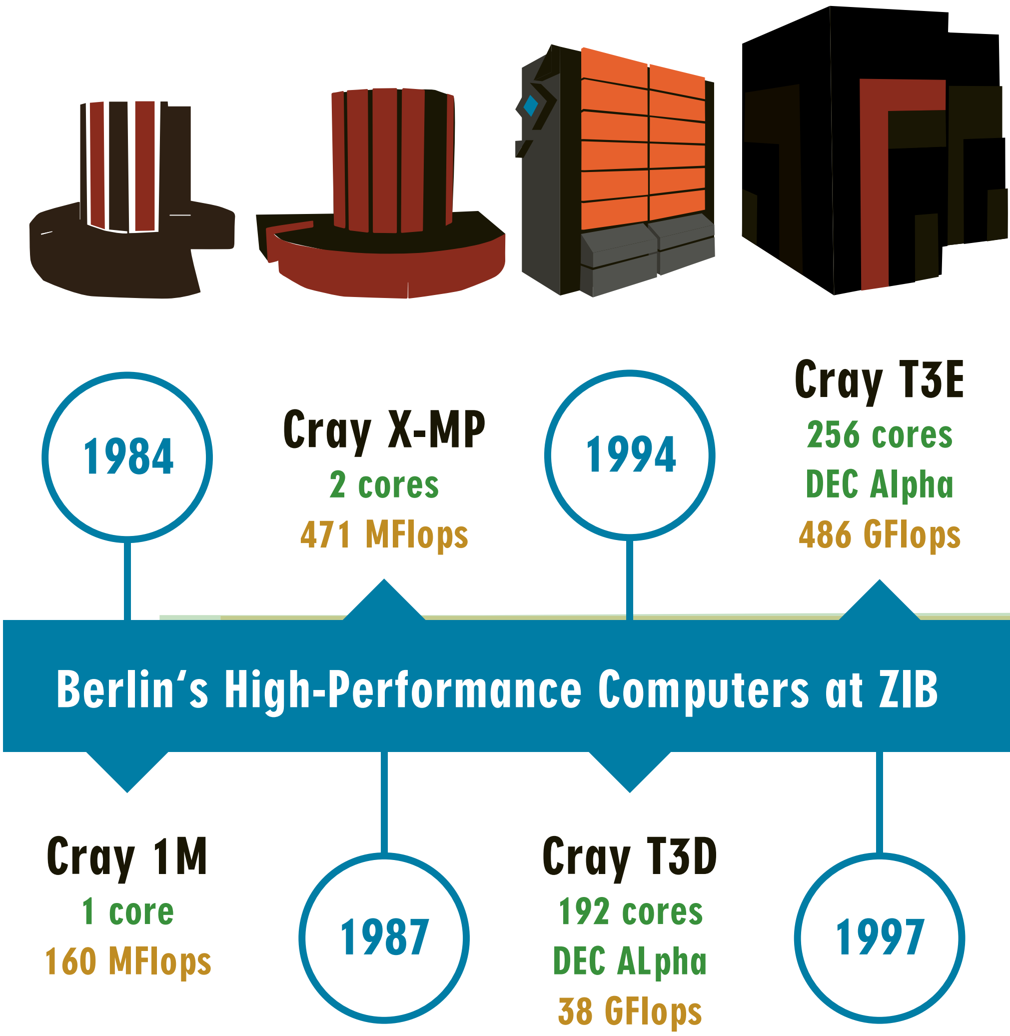


Figure 14: Numerical experiments confirm the order of the theoretical error bound.

NHR@ZIB: TAKING THE NEXT

The NHR Center at ZIB provides high-performance compute and storage resources to support the German science community.

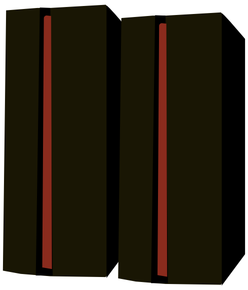
Thomas Steinke | steinke@zib.de | +49 30 841 85-144



The German federal and state governments decided to jointly fund the National High Performance Computing Alliance (NHR). This will create a nationally coordinated network of high-performance computing centers. Eight centers, including ZIB, were selected on the basis of a competitive and science-driven process involving a review by the German Research Foundation (DFG) and an evaluation by an independent strategy committee appointed by Germany's Joint Science Conference. For the next ten years, funding of more than 70 million euros was awarded for the NHR Center at ZIB.

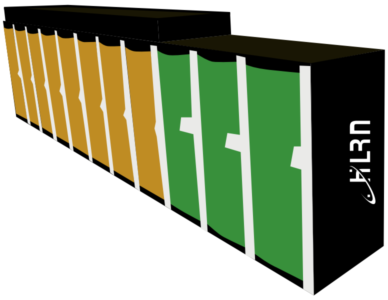
STEP IN HPC

HLRN Supercomputers at ZIB



2002

IBM p690
384 cores
IBM Power4
2.5 TFlops



2008

SGI ICE, XE
10,240 cores
Intel Harpertown,
Nehalem
160 TFlops

2013

Cray XC40/30
44,928 cores
Intel Ivy Bridge,
Haswell
1.3 PFlops

2019

Atos/Intel
121,920 cores
Intel Cascade
Lake AP
8.2 PFlops

HPC Services at ZIB – a Long Journey

Scientific computing, nowadays complemented by artificial-intelligence applications, is the undisputed third pillar of science today. This is the result of evolution over the last decades in many areas of science, technology, manufacturing, and, last but not least, the enthusiasm of young scientists and engineers in recognizing the potential of performing experiments "in silico." In the beginning, high-performance computer or supercomputer systems were special designs and used by a few groups of scientists who needed to understand the hardware and software technology behind the scene. Operation of supercomputers was nontrivial from the beginning, and the need to provide special support to the domain scientists soon became clear to ensure that these powerful but expensive resources were used efficiently.

Consequently, in 1984, the Zuse Institute Berlin was founded as an institute for scientific computing as well as a provider of high-performance computing (HPC) resources. From the very beginning, an important goal was to establish first-class scientific support provided by HPC consultants to promote HPC in the science communities in Berlin; later, this was expanded to northern Germany, and now nationwide in the context of NHR. So, since the first days, an essential aspect has been the seamless integration of scientific and technical expertise with the excellent administrative support of users across different science domains. Physics, chemistry, and engineering evolved as traditional HPC science domains in the 80s and therefore the early team of HPC consultants at ZIB included senior scientists from these areas. Nowadays, other science disciplines like life science and data science have joined.

Sharing different supercomputer models and best-practice solutions among different academic HPC sites is another promising way to cope with the increasing demand for HPC services. As early as with its newly installed CRAY 1M in 1984, ZIB together with the universities in Kiel and Hannover founded the North-German Vector computer Network (NVV)[1] to establish one of the first HPC alliances across German states. The NVV not only shared the hardware resources, at that time dominated by vector-computer architectures with new installations rotating every two years, but also a distributed network of HPC consultants was implemented together with cross-site user management of the joint-user base.

1 The NVV (German: Norddeutscher Vektorrechnerverbund) was based on a state treaty between the federal states of Berlin, Schleswig-Holstein, and Lower Saxony.

In 2001, the next major step was taken: The North German Supercomputing Alliance[2] (HLRN) was founded. With the experiences of the NVV having been well accepted by scientists and uniquely visible by German federal and state bodies, the seven federal states Berlin, Brandenburg (since 2012), Bremen, Hamburg, Mecklenburg-Western Pomerania, Lower Saxony, and Schleswig-Holstein now became partners to foster HPC for researchers in northern Germany. The HLRN alliance jointly operated a distributed supercomputer system at the two HLRN sites: one HPC system at ZIB in Berlin and its sister system in Lower Saxony at the Leibniz University Hannover (until 2018) and the Georg-August-Universität Göttingen (from 2018), respectively. With the joint forces of seven states, not only very competitive HPC system resources could be provided to the scientific-research community in Germany, but the users of the HLRN could also resort to a transregional and interdisciplinary competence network consisting of application experts. This network involved all boards and institutions in the HLRN.

So far, four generations of HPC systems have been operated, with differences in their computer architecture including processor models and high-speed network types (see box). The "Lise" system is the current flagship system at ZIB with a peak performance of about 8,000,000,000,000,000 (Peta) floating-point operations per second (PFLOPS). Such a class of HPC system provides many levels of concurrency of processing data to achieve this tremendous performance. And clearly, it is not easy to exploit this potential as it requires a sufficient understanding of the system architecture, the software tools, and application packages. Again, this is part of the daily scientific services of the HPC team at ZIB to the users.

2 Norddeutscher Verbund für Hoch- und Höchstleistungsrechnen – HLRN.

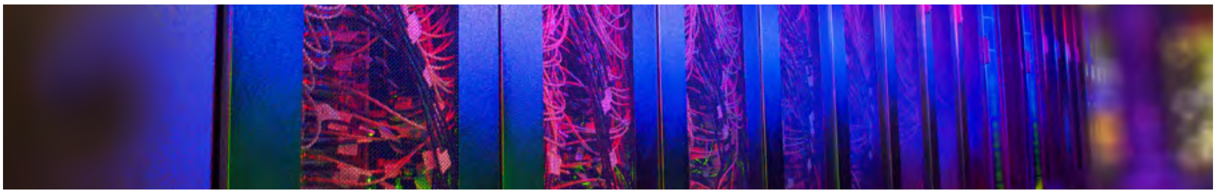
The Next Step: The NHR Center at ZIB

Over the last half decade, it has become evident that nationwide further steps have to be taken to foster and extend high-performance computing infrastructure and competence at a broader level in the German science community. Not only the increasing demands for computational resources has become increasingly challenging, but also the skill set now required for using these resources efficiently has required more and more attention. For example, as laptops and smartphones become parallel computers at the lower end, and large, highly parallel, and very powerful computer systems at the top develop toward exascale computing, developing high-quality scientific software for these machines becomes increasingly more complex, as a deep technical understanding is a must. The gap between the performance capabilities of the hardware and the performance achieved with real-world codes has to be minimized. Simultaneously, the gap between the required skills for using Tier-3 (universities, research groups) and Tier-1 (pre-/exascale) systems has to be closed, too – a call for better coordinated actions for tier-2 HPC sites is needed. Furthermore, the operational costs of a Tier-2 HPC system over its five- to six-year life span become as expensive as the amount invested.

In order to meet the increasing importance of and demand for high-performance computing, in November 2018, the German Joint Science Conference[3] (GWK) agreed on the joint funding of a coordinated National High-Performance Computing Alliance (NHR) by the federal and state governments. Importantly, in addition to the provision of computing capacities, one focus of the NHR network, according to the agreement, was the strengthening of methodological competence through coordinated education and training of users and, in particular, young scientists.

In January 2020, the call for proposals to become a NHR Centers was published. In an open, competitive, science-driven process involving a review by the German Research Foundation (DFG) and an evaluation by an independent strategy committee appointed by the GWK, the new Tier-2 HPC centers were selected. In November 2020, the GWK announced eight HPC centers as new NHR member sites with the ZIB as one of them. The NHR Center at ZIB has been in operation since January 2021.

3 Gemeinsame Wissenschaftskonferenz – GWK.



The Next Years as NHR Center

The HPC world is facing a few challenges in the coming years. The NHR network offers the potential for solving some of them by combining the diverse expertise and working jointly on some of the demanding challenges. At the technological and methodological level, one important challenge is how future Tier-2-level HPC system architectures will look like. The answer to this is mostly driven by economical demands to operate HPC facilities. Heterogeneity is one promising key word that is part of the solution. Diversity in the worlds of processors and accelerators (general-purpose graphic processing units, vector engines, field-programmable gate arrays, special chips for artificial intelligence operations, quantum computers) and memory technologies (e.g., high-bandwidth memory, traditional DRAM, storage-class memory, memory over fabric) offer greater space for designing innovative system architectures. But there is no one-size-fits-all architecture providing maximum performance while staying in a given power envelope. NHR offers the potential for a coordinated and balanced specialization in terms of future system architectures and the way the science domains, with their software codes, can be supported on the NHR systems.

The NHR Center at ZIB will take the lead in selected application domains (e.g., life sciences with a focus on model-driven simulations, advanced integration of machine learning and simulation, or chemistry/material science) together with other NHR centers. As one of the first joint-activity steps of the NHR centers, ZIB, the Friedrich-Alexander University Erlangen-Nuremberg (FAU), and the Paderborn University (UPB) are establishing the NHR Competence Center for Atomistic Simulations to support the broad scope of atom-based simulation techniques in the areas of physics, chemistry, and engineering. As a horizontal method, machine learning/artificial-intelligence workflows will be increasingly integrated into the model-driven simulation schemes. For the next-generation HPC architecture, one can try to maximize the sustained application performance for important software packages of the science domains highlighted above. Such a heterogeneous system architecture may significantly comprise GPUs that are complemented by FPGAs or available AI processors. So, as the next challenge, the question arises as to what extent existing application software can take advantage of such a heterogeneous system architecture, and which program codes that are important in the science communities need to be migrated and optimized for these next-generation systems. Again, the NHR Alliance provides opportunities for collaborative work on optimizing such codes here, for example, for accelerators. With FAU, UPB, and RWTH Aachen, ZIB is initiating an open NHR project called “Performance Lab” in which these problems can be tackled.

Funding for the NHR Alliance was secured until 2030. For the next ten years, we are looking forward to welcoming the NHR user community at ZIB.

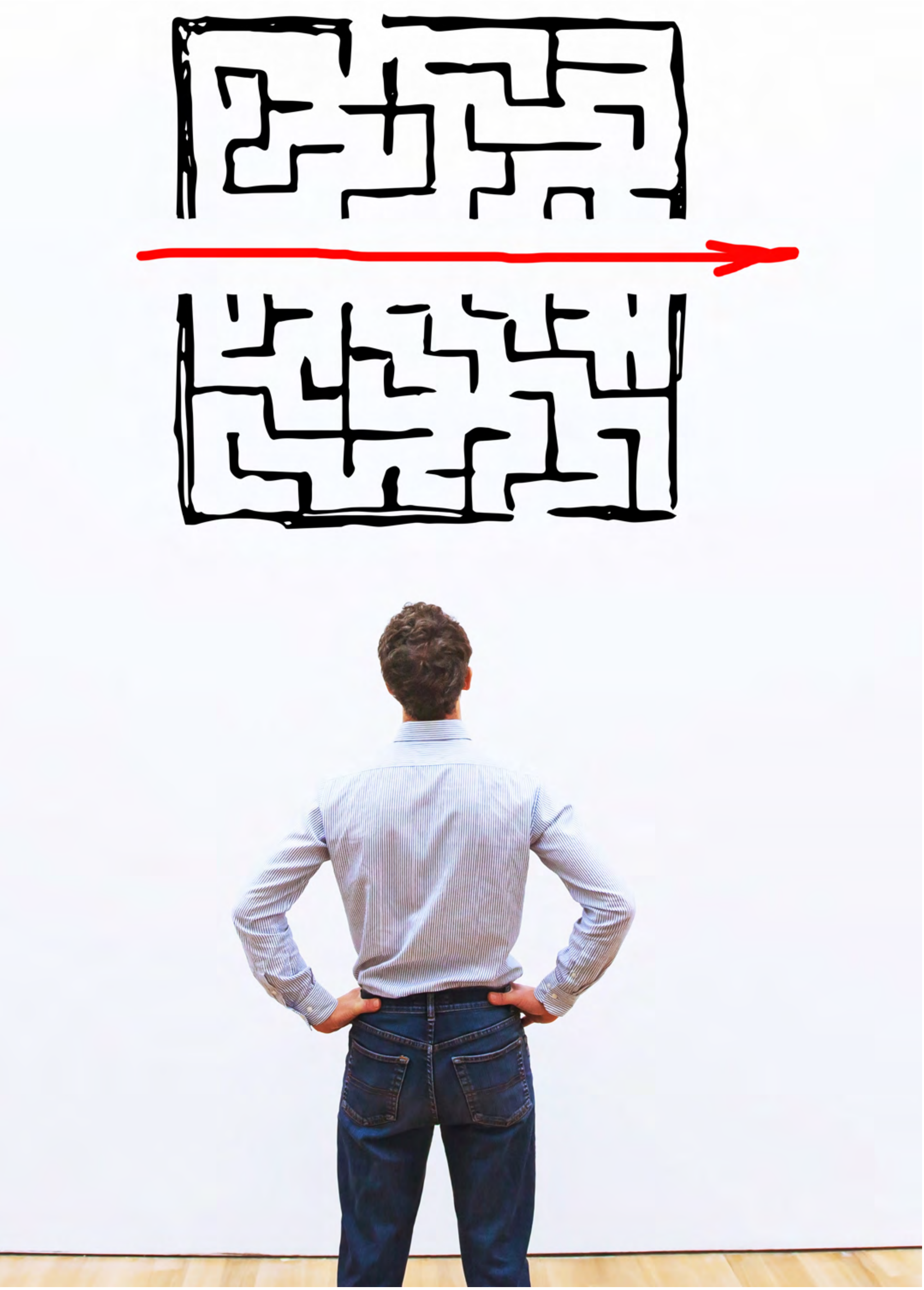
SIMPLE EXPLANATIONS, SPARSITY, AND CONDITIONAL GRADIENTS

Christoph Spiegel | spiegel@zib.de | +49 30 841 85 - 436

Motivation

Almost all questions can have a complex answer or a simple answer. This rings true in particular in science, and we typically strive to find the most parsimonious explanation, i.e., the simplest explanation that is consistent with an observed phenomenon:

> "EVERYTHING SHOULD BE MADE AS SIMPLE AS POSSIBLE, BUT NO SIMPLER."
>
> Albert Einstein.



A simple explanation is more stable, reacts better to noise, and does not induce "structure" or "phenomena" beyond what is in the data. Moreover, also from a Bayesian statistics perspective, the most parsimonious model is (almost) always the "best" answer. Now that we are entering an age where we use more and more machine learning (ML) systems to gain insights into the world, this imperative is as important as ever: nobody can interpret the more than 1 billion parameters of modern deep-learning systems. While we do often need a large number of parameters to describe a phenomenon, we would still like to adhere to the principle of the simplest sufficient model in order to explain what is going on.

At the core of ML training is often the solution to an optimization problem and the parsimoniousness from above translates into the "sparsity" of the solution. For example, the famous LASSO regression problem aims for regression, where the number of explaining factors in the final solution is kept small. More generally, one might strive for solutions to these ML optimization problems that are *made up of few (potentially complex) base objects; these base objects are often called atoms. First*-order methods are essential to many modern optimization problems, from smooth and convex regression problems in lower-dimensional spaces to highly overparameterized and very non-convex deep neural networks. For many applications, the parameter space is often constrained, either due to natural restrictions or because a particular regularizing effect is desired. Frank-Wolfe algorithms (Frank and Wolfe, 1956), one of the simplest and earliest known classes of iterative algorithms and also often referred to as conditional gradient algorithms (Levitin and Polyak, 1966), have been the center of enormous renewed interest over the last several years (see, e.g., Jaggi, 2013) and have increasingly established themselves as the method of choice for such constrained optimization problems when we seek sparsity of the final solution. Not only can they handle constraints very efficiently as they maintain feasibility of the iterates without requiring projections, they also naturally induce sparsity of the final solution as it is built up as a convex combination of a few atoms. Additionally, these methods are computationally very efficient, so that rather large learning problems can be solved efficiently and in particular the computational cost is often much cheaper than that of corresponding projection-based methods (Combettes and Pokutta, 2021).

In the following, we present a few examples of our work in the area of conditional gradients with a particular focus on sparsity.

Conditional Gradient Algorithms for Learning Dynamical Systems

As mentioned earlier, the modern age of machine learning and big data that we live in has enabled the rise of deep-learning models, often possessing billions of parameters, which can be trained to classify or predict with high accuracy when given enough input data. However, these models are oblivious to how the underlying data with which they are trained is generated, and often fail to predict or classify out-of-sample data, coming from a regime that has not been encountered in the training process.

This approach stands in contrast to how many of the phenomena in physics were described or explained in previous centuries, that is, through the use of differential equations. These equations have helped build our understanding in fields as wide-ranging as quantum mechanics, fluid mechanics, or electromagnetism, and have enabled a deep understanding of how the data coming from a physical system has been generated. Given the current availability of data, there has been a recent trend to find these differential equations through sparse regression methodologies. Such methods attempt to find the best sparse fit to observational data in a large library of potential nonlinear models and often perform extremely well with noise-free data but fail to produce accurate or sparse solutions when noise is present. Please see the feature article "Understanding and Modeling Complex Biological Systems" for more details on such methods for learning dynamical laws from data.

Conditional gradient methods are an exceptionally attractive family of algorithms for these types of applications because of their sparsity-inducing properties. They produce more accurate and sparser solutions than other competing algorithms in the presence of noise, allowing us to find governing differential equations with less training data, for example. This means that the learned governing equations often contain fewer spurious terms and noise is less likely to be explained by phantom terms. Moreover, due to the fact that these families of methods are easily able to deal with linear constraints (and more generally convex constraints), we can add constraints to our learning problem to impose certain symmetries on our learned model, or to guarantee certain conservation properties are satisfied, thereby producing models that are consistent with the underlying physical phenomena (Carderera et al. 2021). We call this conditional gradient-based recovery method CINDy (short for conditional gradient-based identification of nonlinear dynamics), in homage to SINDy (Brunton et al., 2016). On numerical examples with noisy data, CINDy outperforms several alternative approaches in terms of (1) prediction accuracy, (2) sparsity, and (3) replication of dynamics trajectories. One reason is that CINDy blends orthogonal matching-pursuit-like steps gradually into basis-pursuit-like steps by picking up atoms sequentially, which are only accepted into the set of active atoms if the improvement of training accuracy significantly outweighs the loss in sparsity. This leads to extremely sparse iterates, often close to the information-theoretic recovery limit.

The Kuramoto system, consisting of a set of weakly coupled oscillators with slightly different natural frequencies, is a model for studying synchronization phenomena. The natural dictionary contains trigonometric functions and is overcomplete. This results in other approaches selecting too many atoms and producing dense solutions (Figure 1). The extra atoms do not only hide the actual structure of the Kuramoto model, but also lead to inferior model predictions and evolution of trajectories (Figures 1 and 2).

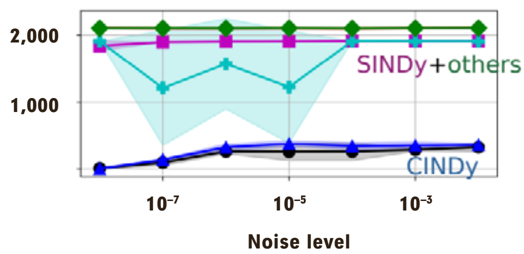


Figure 1: Number of wrongly selected atoms for the Kuramoto example over the noise level in the trajectory data.

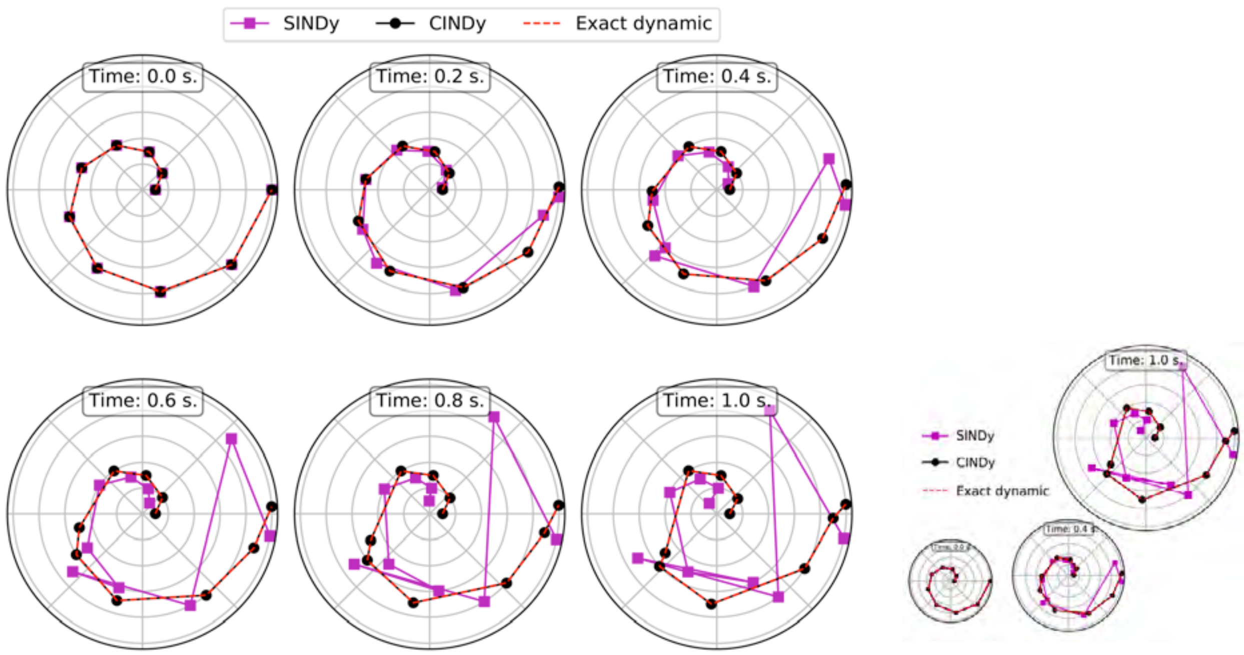


Figure 2: Trajectory snapshots of dynamics learned by the SINDy and CINDy algorithms with data from the Kuramoto model under the presence of noise. The ordinary differential equation learned by the CINDy algorithm produces a trajectory that well matches that of the exact dynamic.

Conditional Gradients for Everyone

As the aforementioned conditional gradient algorithms are an essential part of our own work but are also heavily used within the broader research community, we are currently developing a new high-performance software library implementing major Frank-Wolfe variants in Julia. The algorithms can be used from within other common programming languages (such as C, C++, Python) and frameworks (such as MathOptInterface and JuMP in Julia) through respective wrappers.

We had several specific design aspects in mind, when designing the toolbox. In particular, we tried to push the envelope on the usual "speed vs. robustness vs. scale" trade-off triangle:

1. **It should be fast. That is why we opted for Julia in first place as** the language and the overall design is geared toward speed with specialized data structures.

2. **It should be robust. When solving optimization problems, numerical instabilities can be a huge problem when high-accuracy solutions are sought after, or the data is ill-conditioned. We cater to this by supporting various floating-point modes out of the box allowing for very-high-precision computing** (e.g., using BigFloats).

3. **It should work at scale. In order to work at** scale for very large-scale models with several billion variables, memory management becomes very important. At that scale, allocations of memory cost real time and you are close to the physical memory limit of modern machines, so that unnecessary allocations can easily lead to out-of-memory blowups. We paid particular attention to this by supporting a memory-emphasis mode in all algorithms that significantly reduces the memory consumption. We also implemented a variety of constant-storage sparse representations for common feasible regions (independent of the problem dimension).

4. **It should be flexible. We support almost arbitrary number types and element types for the iterates out** of the box. Notable examples are rational arithmetic, so that we can, for example, solve the approximate Carathéodory problem to actual optimality.

5. **It should be easy to use. The best code is worth nothing if it cannot be used. As such**, we provide ample interfaces and hide all complexities within the framework, so that the user only needs to supply a linear optimization oracle for the feasible region as well as a first-order oracle for computing gradients. Gradient computations can also be transparently performed using GPUs or other hardware accelerators.

To put things into numbers, the following table provides statistics for various components of the basic Frank-Wolfe algorithms, here solving a LASSO problem over 1e9 (i.e., one billion variables).

Size of single vector (Float64): 7629.39453125 MB

		Time			Allocations		
Tot / % measured:		278s / 31.4%			969GiB / 30.8%		
Section	ncalls	time	%tot	avg	alloc	%tot	avg
update (blas)	10	36.1s	41.3%	3.61s	149GiB	50.0%	14.9GiB
lmo	10	18.4s	21.1%	1.84s	0.00B	0.00%	0.00B
grad	10	12.8s	14.6%	1.28s	74.5GiB	25.0%	7.45GiB
f	10	12.7s	14.5%	1.27s	74.5GiB	25.0%	7.45GiB
update (memory)	10	5.00s	5.72%	500ms	0.00B	0.00%	0.00B
dual gap	10	2.40s	2.75%	240ms	0.00B	0.00%	0.00B

Each gradient here is about 7.6 GB of memory and a single gradient computation takes about 1.28 seconds. We can also see that while an update of the iterate in optimized linear algebra requires about 15.2 GB of memory, in memory emphasis mode, it can be computed without allocating memory, so that it ends up being faster than the optimized linear algebra operations at that scale due to the large cost of allocating memory.

Vanilla Frank-Wolfe Algorithm.

EMPHASIS: Memory STEPSIZE: Agnostic EPSILON: 1.0e-7 MAXITERATION: 1,000 TYPE: Float64
MOMENTUM: Nothing GRADIENTTYPE: Nothing
WARNING: In memory emphasis mode iterates are written back into x0!

Type	Iteration	Primal	Dual	Dual Gap	Time	It/sec
I	0	1.000000e+00	-1.000000e+00	2.000000e+00	8.456415e+00	0.000000e+00
FW	100	1.326732e-02	-1.326733e-02	2.653465e-02	4.628479e+02	2.160537e-01
FW	200	6.650080e-03	-6.650086e-03	1.330017e-02	9.171626e+02	2.180638e-01
FW	300	4.437059e-03	-4.437064e-03	8.874123e-03	1.371521e+03	2.187352e-01
FW	400	3.329174e-03	-3.329180e-03	6.658354e-03	1.825769e+03	2.190858e-01
FW	500	2.664003e-03	-2.664008e-03	5.328011e-03	2.280048e+03	2.192936e-01
FW	600	2.220371e-03	-2.220376e-03	4.440747e-03	2.734262e+03	2.194376e-01
FW	700	1.903401e-03	-1.903406e-03	3.806807e-03	3.188507e+03	2.195385e-01
FW	800	1.665624e-03	-1.665629e-03	3.331253e-03	3.642893e+03	2.196057e-01
FW	900	1.480657e-03	-1.480662e-03	2.961319e-03	4.097155e+03	2.196646e-01
FW	1,000	1.332665e-03	-1.332670e-03	2.665335e-03	4.551414e+03	2.197119e-01
Last	1,000	1.331334e-03	-1.331339e-03	2.662673e-03	4.556530e+03	2.196847e-01

4557.523048 seconds (6.90 M allocations: 112.096 GiB, 0.01% gc time)

Here, these 1,000 iterations were performed in about 4,557 seconds (i.e., about 4.55 seconds per iteration requiring about 112.096 GB of memory allocation, which works out to just 112 MB per iteration), which is considerably less than the size of a single vector (7.629 GB) and we maintain two of these: the gradient and the iterate. The reason is that we aggressively reuse memory, so that the only real memory allocations that are required are those for computing information for reporting. It should also be noted that conditional gradients offer dimension-independent convergence guarantees, with consistent convergence speed irrespective of the problem size.

Conditional Gradients in Deep Learning

Another natural application of conditional gradient methods is in the context of deep learning, where the sparsity-inducing properties of these methods – together with an appropriately chosen feasible region – can significantly affect how information is represented within the neural network and how decisions are made by the network. For example, in Figure 3, we can see that depending on the method and feasible region used for training, the network relies more on positive evidence (green pixels) or negative evidence (red pixels). Here, evidence relates whether the presence of a pixel (green) is evidence *for* a given class (here the digits 0 or 5) or whether the presence of a pixel (red) is evidence *against* a given class. As such, conditional gradient algorithms can significantly change the statistical properties of the trained neural network altering, for example, the false-positive vs. false-negative trade-off.

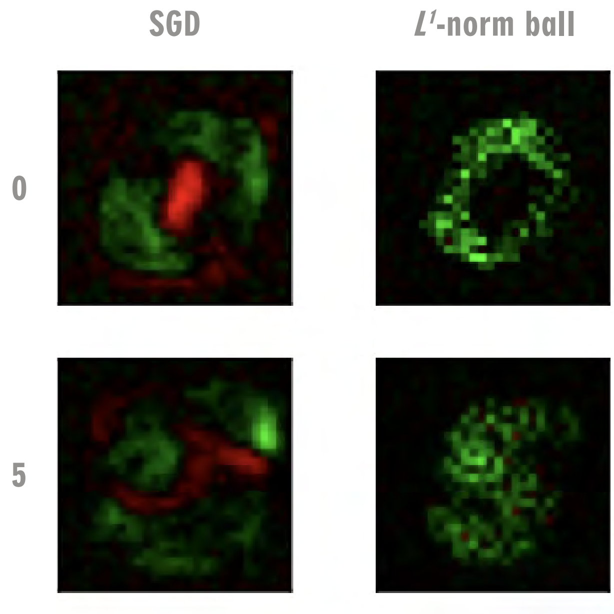
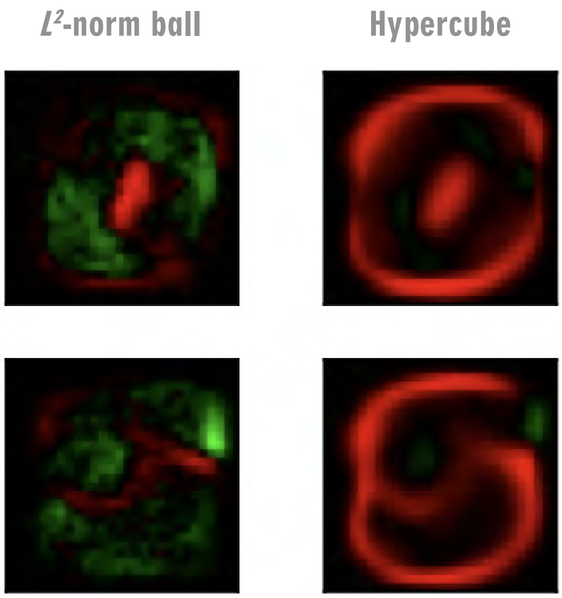


Figure 3: Visualization of the weights in a fully connected no-hidden-layer classifier trained on the MNIST dataset corresponding to the digits 0 and 5. Red corresponds to negative and green to positive weights.



In order to be able to try our methods on state-of-the-art neural-network training problems, we have likewise implemented several Frank-Wolfe methods in the two most common frameworks used for the optimization of neural networks, TensorFlow and PyTorch. In our experiments, we have demonstrated that stochastic Frank-Wolfe methods can achieve state-of-the-art test accuracy results on several well-studied benchmark datasets, namely CIFAR-10, CIFAR-100 and ImageNet. We furthermore show that the chosen feasible region significantly affects the encoding of information into the networks both through a simple visualization and by studying the number of active weights of networks trained on MNIST with various types of constraints, as already discussed above (see also Figures 1 and 2).

RESEARCH CAMPUS MODAL: SUCCESS STORIES

Ralf Borndörfer | borndoerfer@zib.de | +49 30 841 85-243

Funding for Research Campus MODAL Renewed

After successful completion of its first funding phase from 2014 to 2020, the Research Campus MODAL was evaluated by an expert review panel and the grand jury of the German Federal Ministry of Research and Education (BMBF). MODAL's proposal for a second phase was granted (2020–2025). The grant includes 10 million euros funding by the BMBF and a considerably larger amount by the 35 participating private companies. MODAL hosts more than 60 researchers from academia and industry under one roof. This article reports on the plans for the second funding period, presents success stories from the first funding period, and highlights research and practice training events of MODAL.


TEXT

EnergyLab

Smart control of energy transport networks

MobilityLab

Optical planning for traffic networks

MedLab

Smart solutions for digital precision medicine

MODAL is a public-private partnership project of the research partners Zuse Institute Berlin (ZIB) and Freie Universität Berlin with more than 30 industrial partners, funded by the German Federal Ministry of Education and Research within the "Research Campus – Public-Private Partnership for Innovation" funding initiative. The research campus is hosted by ZIB, and conducts research and development of digital systems for the optimization of data-driven processes in the fields of energy, health, mobility, and communication.

The developed data-driven methods make rapid cross-technology development possible and advance innovation processes by digital decision-support systems. Present projects aim at the solution of previously unavailable planning and control problems for complex supply and mobility networks or for digital health care.

In the first funding phase, the MODAL research campus selected mobility, sustainable energy supply, secure communication, and personalized medicine as the main topics of its research program. During the first funding period, MODAL produced a series of success stories in transferring basic research in Applied Mathematics into highly relevant technological innovations with productive utilization in industry and a proven value of several million euros per year. Two of these success stories advanced to the 2020 finals of two of the most renowned awards competitions in research transfer. We will report on these success stories in detail below.

The different application areas considered in MODAL face similar challenges requiring mathematical methodological competence with a focus on Modeling, Simulation and Optimization (MSO), Artificial Intelligence (AI), and High-Performance Computing (HPC). In MODAL, these methodological core competencies are complemented by application expertise related to specific data-driven economic processes in the areas of energy, mobility, health, and nanotechnology. The resulting competence matrix for the second funding phase is illustrated in Figure 1 and constitutes both the strategic research program and the unique selling point of the Research Campus. MODAL is strong in communicating this unique mix of research excellence and proven application competence in industry to the next generation of researchers. One of MODAL's main activities in this field, the school on "Combinatorial Optimization at Work" (CO@Work), was organized for the sixth time in 2020, and attracted more than 1,000 participants . We report on this event in detail below.

NanoLab

Smart design of resonant nanostructures

HPCLab

Solutions for high-performance optimization

SynLab

High-performance optimization software

Figure 1: In its second funding phase, the Research Campus MODAL comprises six labs. The EnergyLab is based on the former GasLab, the MobilityLab extends the former RailLab, while MedLab and SynLab just shifted focus. With HPCLab and NanoLab, new aspects were introduced.

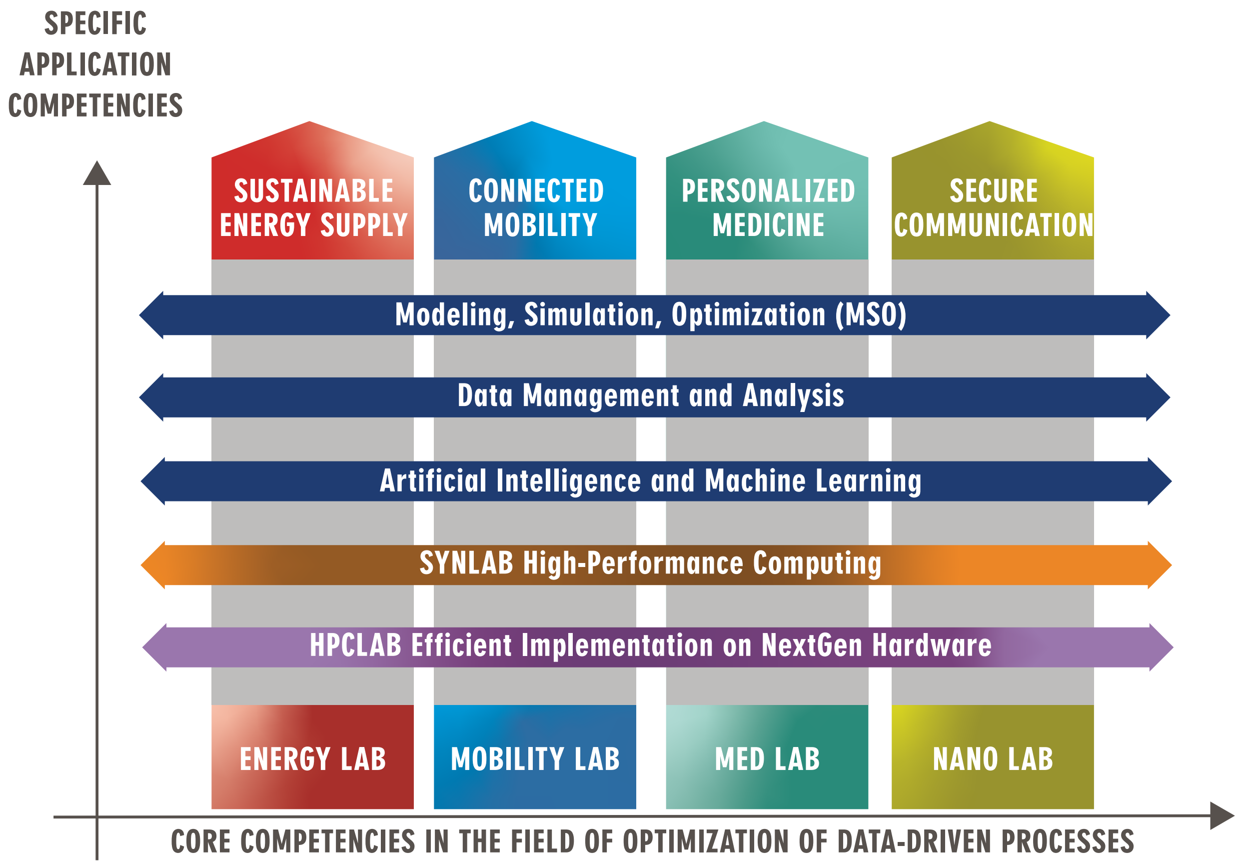


Figure 2: Mobility, sustainable energy supply, secure communication, and personalized medicine are the main topics of MODAL's research and research transfer program. Despite this broad range of topics, MODAL is characterized by strict coherence in terms of content, which results from the mathematical methodological competence with a focus on Modeling, Simulation, Optimization (MSO), and AI. These core competencies (shown horizontally here) are complemented by (vertical) application competencies related to specific data-driven economic processes. The resulting competence matrix illustrates both the strategic research program and the unique selling proposition of the research campus.

MobilityLab Enters Finals of Franz Edelman Competition 2020 – Mobility

A MODAL team with members of Deutsche Bahn (DB), ZIB, and LBW Optimization advanced to the final of the renowned Franz Edelman Award 2020 for Achievement in Operations Research and the Management Sciences with FEO, a decision-support system to schedule the train rotations of DB. Based on completely new mathematical methods of hypergraph optimization and efficient algorithms based on it, both developed within MODAL, FEO has greatly increased DB's planning agility and productivity, and produced significant direct savings of roughly 75 million euros per year.

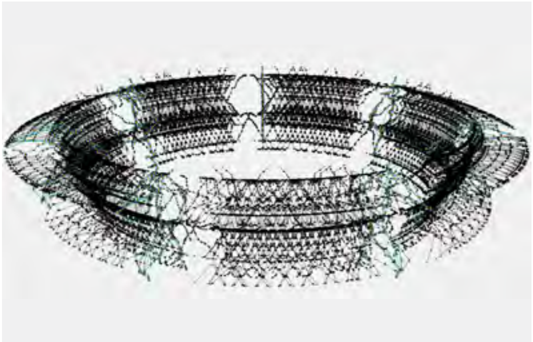
The Franz Edelman Award for Achievement in Operations Research and the Management Sciences is the "Nobel Prize of Operations Research." Presented by the Institute for Operations Research and the Management Sciences (INFORMS), it is organized as an international competition of successful OR applications. A winner is chosen among six finalist teams, who present their contributions in an illustrious gala at the INFORMS Conference on Applied Analytics.

The combined impact of all Edelman finalist contributions since 1974 is impressive, amounting to a total of $302 billion. The contributions cover a range of (up to today) 143 different application areas. In this way, the records of the Edelman Award serve not only as an objective indicator of a hard, monetary "Value of Operations Research," but also as a showcase for the practical viability of groundbreaking new mathematical ideas.

The 2020 Edelman gala was originally scheduled to take place in Denver. Due to the coronavirus situation, however, it was rescheduled and held, for the first time, as a virtual online event on September 29. Five teams participated in the final: Carnival & plc, the world market leader in cruises, with the YODA (Yield Optimization and Demand Analytics) revenue management system; IBMServices (the Global Technology Services Unit within IBM) identified devices at risk from outages before they occur; Intel improved their supply chain; Walmart developed an innovative markdown approach for perishable goods; and Deutsche Bahn scheduled its 26,000 trips per day using the FEO (Fleet Employment Optimization decision-support system.

FEO is designed as a modular software system centered around adaptable optimization kernels, focusing on using the best possible OR methods. FEO is designed to deal with complex and large-scale scenarios. DB is the largest European cargo and regional passenger train operator, and the second largest in long-distance passenger rail transport, carrying 2 billion passengers per year. DB Cargo owns over 2,500 locomotives (shunters included), distributed over 17 European countries, of which around 800 locomotives serve the core network in central Europe. DB Regio owns over 4,000 railcars and over 850 locomotives (shunters included). DB Fernverkehr owns over 270 intercity express (ICE) high-speed railcars and over 240 locomotives (shunters included). At a price of up to 33 million euros per piece (ICE 3 type 407/Velaro), the efficient use of this rolling stock is critical.

Figure 3: A team comprising Deutsche Bahn, LBW Optimization, and ZIB advanced to the final of the 2020 Franz Edelman Competition.



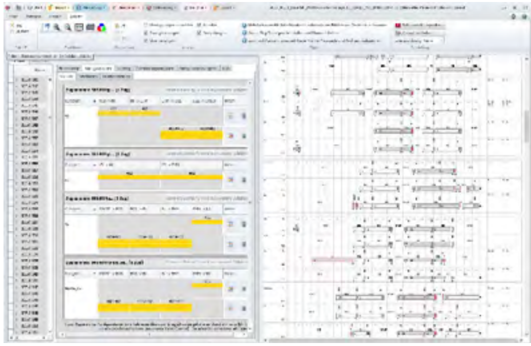
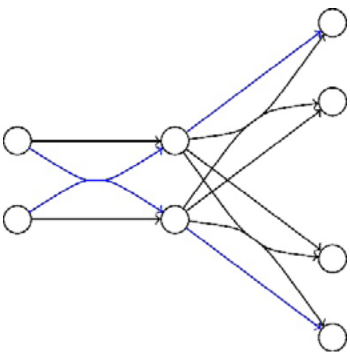


Figure 4: Timetabled ICE 1 trips in a cyclic standard week (left). FEO GUI screenshot (right).

A key problem is train rotation planning (i.e., the construction of the rolling-stock rotations in order to implement a given timetable [BK+18]). In contrast to public transit or air traffic [BGJ10], where individual vehicles can be scheduled using network-flow approaches, several units of rolling stock are combined to form trains. The types (e.g., country-specific power or safety systems), positions (e.g., restaurant in the middle), and orientations (e.g., tick = first class in front, tock = second class in front) of the units in a train follow elaborate rules and are not easily changed. This means that vehicle rotations and train compositions have to be planned simultaneously in order to arrive at an implementable solution. The same holds for maintenance intervals and parking assignments. The main objectives are cost minimization (number of vehicles and deadhead mileage) and regularity (i.e., similarity of operation over several days of the planning horizon). Typical planning scenarios include the cyclic "standard week" in long-term planning (cf. Figure 4 left) and fully dated planning over four to six weeks.

Traditionally, train rotations are planned in an incremental fashion that was straightforward to implement. Today, as a result of the deregulation of the European railway sector, train services change much more dynamically due to competition, and resource allocation is mainly driven by cost minimization. This results in highly challenging scheduling problems in large networks, with complex technical constraints, over multiple time scales, integrating several stages of operations. These could not be tackled using established methods. In fact, the inability to solve the key problem of train rotation scheduling was a major obstacle that prevented OR methods for the longest time from blazing a trail of success in the railway industry comparable to the one in the airline industry.

The FEO optimization core is based on completely new mathematical methods of algorithmic hypergraph theory developed within MODAL at ZIB [BR+11, BR+12]. This approach ensures the integrated treatment of the main operational requirements of train-rotation creation, train composition, maintenance inter-



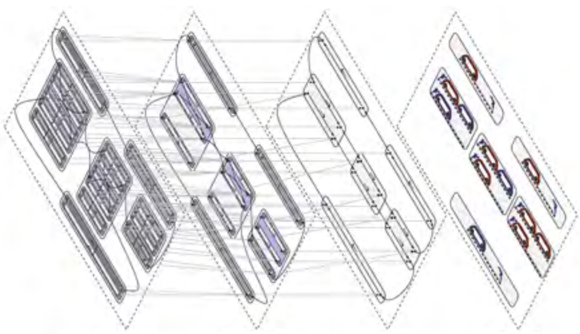


Figure 5: Idea behind the hyperflow train-rotation model (left). The three layers of the coarse-to-fine method model compositions, configurations, and vehicle flows (right).

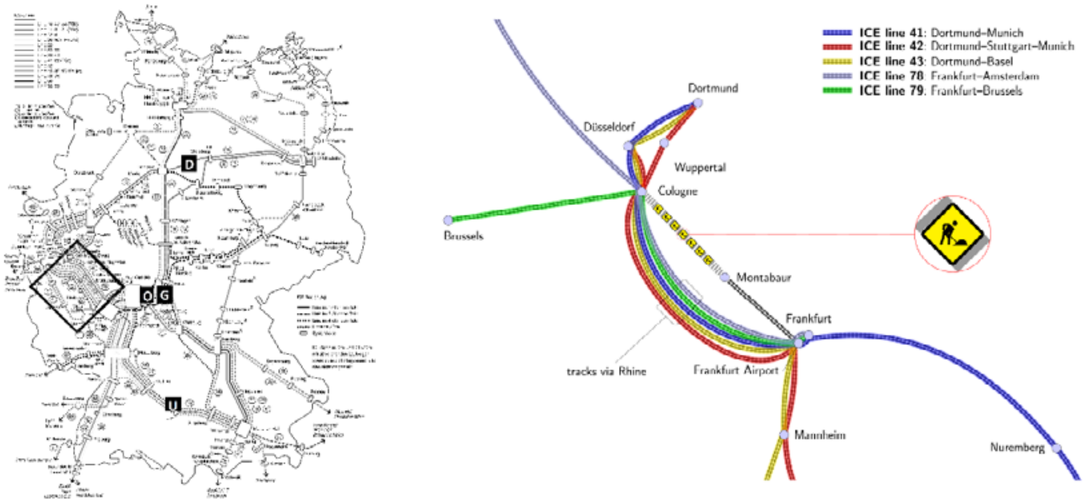


Figure 6: The ICE rotations through the 2015 Cologne-Rhine-Main construction site were managed by FEO (square left and right). The impact of FEO was analyzed for four dedicated construction sites marked D, D, O, U (left).

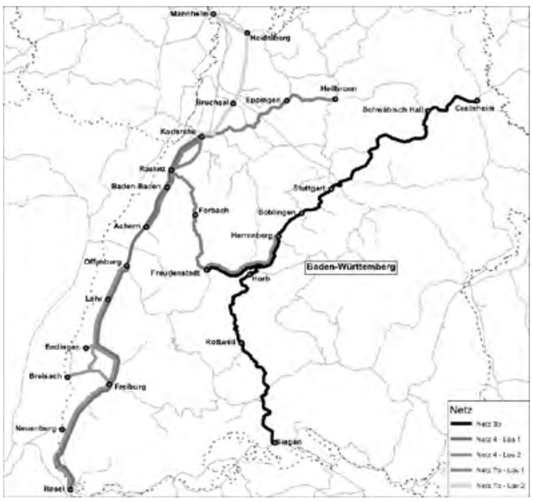
vals, parking, and regularity, all within one generic hypergraph model (cf. Figure 5 left). Exploiting the special "graph-based" structure is the key to the development of a novel coarse-to-fine (C2F) algorithm for the solution of large-scale instances with up to one-hundred-million hyperarcs (= variables) [BRS14] (cf. Figure 5 right). C2F works on coarse layers where possible and resorts to finer layers when needed to compute a provably optimal linear programming solution and in turn an integer solution of provable quality [BR+15]. The hypergraph train-rotation model also gave rise to deep theoretical developments in combinatorial optimization, including a generalized Hall theorem for normal hypergraphs [BB18], a tight-cut-decomposition algorithm for matching covered uniformizable hypergraphs [B19], and a purely combinatorial hypergraph network simplex algorithm [BG+17, BG+17, B18]. Since 2014, the transfer research within MODAL's RailLab has been focused on developing efficient algorithms based on hypergraph theory; these algorithms are now in productive use in the core of FEO.

DB Fernverkehr first used FEO for supporting decisions about the fleet and train sizes of the new ICE4 fleet. Since the rollout of FEO, its main use has been in the management of construction sites, currently 850 per day. The impact of FEO was analyzed in four dedicated cases from the fourth quarter of 2019 (see Figure 6), where service cancellations of approximately 360,600 total train kilometers were prevented. Extrapolating this to an annual value, FEO allows DB Fernverkehr to drive roughly one million additional train kilometers with the existing fleet. This equals the annual mileage of two ICE double-traction train sets and would generate additional fare revenue of approximately 24 million euros per year or equate to the transport of 784,000 additional travelers. Each train journey saves around 44 kilograms of CO_2 in comparison to the amount that would have been expended had each passenger driven an automobile. Thus, FEO saves approximately 34,000 tons of CO_2 emissions per year.

DB Regio uses FEO in managing public tenders for regional rail services all over Germany. In these tenders, the train-operating company that requires the lowest subsidy wins the contract for a period of up to 15 years. Three cases were documented in which FEO was instrumental in winning tenders in the state of Baden-Württemberg (see Figure 7), which together cover mileage of roughly 11 million train kilometers per year and an annual revenue of 23 million euros. In all cases, pivotal optimization opportunities were identified by human ingenuity, and their precise impact analyzed using FEO: in one case, the repositioning of a vehicle could be squeezed into a reserve time, and in another, using vehicles with superior acceleration capabilities gave room to implement an operational concept involving heavy coupling and uncoupling, while exactly the opposite idea turned out to be cheaper in the third case. Assessing these concepts would have been impossible without FEO.

DB Cargo uses FEO for strategic and operational locomotive scheduling in their highly demand-driven and volatile business. A particular problem is cross-border operations, which constitute more than 60% of DB Cargo's transports. These require either special locomotives for every country or expensive so-called interoperable locomotives. The Alps, in particular, can be crossed using the Gotthard or the Lötschberg/Simplon passes in Switzerland ("over the Alps"), or the Lötschberg/Simplon and the new Gotthard Basis tunnels ("under the Alps"). The feasible train compositions depend on the inclination of the track, the weight and the length of the train, and the delivery deadline. In a series of optimization projects, it became clear that an integrated operation of cross-border transports, including a timetable improvement, was superior to operating a diverse locomotive fleet. The new concept leads to an increase in efficiency of 3% to 5%, which amounts to savings of 27 million euros per year.

The Edelman Competition is named after Franz Edelman, who fled from Germany to England in the late 1930s as a teenager, on to Canada, and finally to the US. He worked for 30 years at the Radio Corporation of America, where he set up one of the earliest industrial OR/MS groups. "However, the most powerful incentive will continue to be effective support of the firm's managerial and professional activities," said Edelman [E82]. The FEO system achieves exactly that for the vehicle-rotation operations at DB.



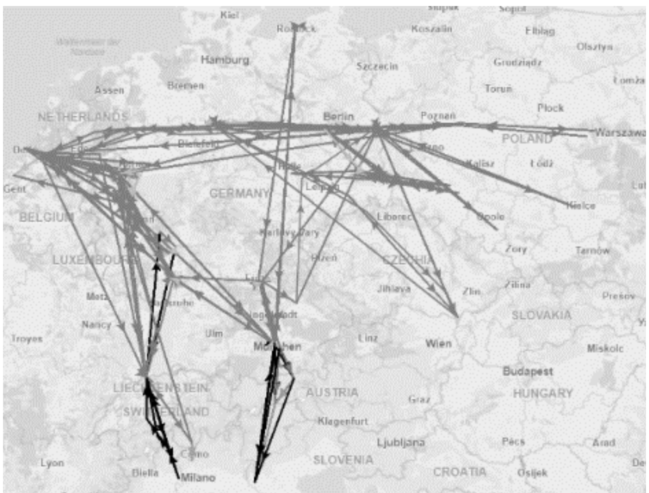


Figure 7: DB Regio used FEO for successful tendering in Baden-Württemberg (left) and DB Cargo for cross-border freight-traffic operations (right).

EnergyLab Enters Finals of the INFORMS 2020 Innovative Applications in Analytics Award

The MODAL EnergyLab team from ZIB, TU, LBW Optimization, and Open Grid Europe made it to the finals of the INFORMS 2020 Innovative Applications in Analytics award. A natural-gas-dispatching decision-support system, developed by the team, employs a novel combination of analytics to compute smart and forward-looking recommendations for safe and efficient control of one of the largest German natural gas networks. This system is the starting point to an optimized control of the network, which will allow the integration of new technologies, especially those related to hydrogen.

The "Innovative Applications in Analytics" award, presented by the Analytics Society of INFORMS, recognizes the creative combination of descriptive, predictive, and prescriptive analytics in innovative applications to provide insights and/or business value. In 2020, the jury selected six finalists out of 33 contributions from across the globe. The MODAL EnergyLab team with researchers from ZIB and TU, LBW Optimization, and Open Grid Europe is proud to be one of them.

The team presented its work on optimal control of the German gas network with its manifold connections to the European network. Open Grid Europe (OGE)

Figure 8: The MODAL team presents its contribution in a video presentation and a video conference with the jury.

is a transmission-system operator responsible for the delivery of more than 25% of the German primary energy consumption and transporting about the same amount to Western and Southern Europe. To achieve this, they operate a natural-gas network of comparable length to the German Autobahn. The 24/7 dispatching center operates 92 compressor units, almost 300 regulators, and more than 3,000 valves to control the network to meet all demands. This is a network that has been built in the course of the last 100 years; the oldest still-active part was built in 1916. There is no computer-readable description and it is not stuffed full of sensors. This is solid engineering made of thousands of tons of steel. Thus, the network is operated mainly based on measured data, which is only available in the "rearview mirror" and the dispatcher's expert knowledge. To improve this situation and to anticipate and prevent critical situations in the network, the MODAL team has developed a smart, forward-looking, analytics-based decision-support system. This was one of the main research projects of MODAL's GasLab, the "ancestor" of the EnergyLab in MODAL's first funding period. It is the main success story of the GasLab that five years of research have led to a software suite that is in productive use at OGE for controlling the gas network.

For this to work, it was necessary to utilize three types of analytics:

- **Descriptive:** modeling and simulating the gas flow in the network.
- **Predictive:** predicting future gas supply and demand at the entries and exits of the network.
- **Prescriptive:** recommending network-control measures to ensure safe and efficient operation of the network.

Every 15 minutes, the optimization core computes recommendations based on the current state of the network, its past, and its technical capabilities. A single station in the network can have more than 1,250 discrete operational modes. Each mode may include target values for continuous quantities, such as the target pressure of a regulator. First, an hourly forecast for the more than 1,000 entry and exit points of the network for the next 24 hours is computed, employing a mixture of machine learning and optimi-

The Combinatorics of Gernsheim

30,000,000,000,000,000 mathematically possible combinations of valve and compressor states.

200,000 feasible operation modes identified based on practitioners knowledge.

1,285 relevant operation modes extracted using analytical evaluation of historical data.

Figure 9: Largest individual compressor station at Gernsheim.

zation on a preselected set of features most suitable for each entry or exit. Based on these forecasts, the recommended operational measures are computed. However, an exact transient model, including all discrete and continuous variables for the full network, is computationally intractable. Therefore, we employ a two-phase approach, decomposing the different aspects of the problem by exploiting the gas network's topology: a coarse model computes amounts and directions of flow, computing when to transport how much gas on which path through the network. Then, detailed models for the individual compressor stations are solved in parallel. The most complex one is shown in Figure 9.

The solutions of the detailed models validate and compliment the amounts and directions calculated by the coarse model and compute precise action recommendations. The objective is to fulfill demands while minimizing costly mode changes considering many additional constraints needed to ensure the practical feasibility of the recommended action. Finally, the results are summarized and displayed to the dispatchers on dedicated iPads, providing them with a set of directions.

As the MODAL GasLab has progressed to the EnergyLab, integrating new technologies into this system has started. On a larger scale, companies have started to produce hydrogen from excess power. This hydrogen can be injected into the natural-gas infrastructure. Since the maximum concentration of hydrogen anywhere in the network is limited, detailed tracking of the gas composition becomes necessary, making the network's operation even more difficult. Also, pure hydrogen transport networks are needed to fulfill, for example, future industrial and mobility demands. Due to the lower energy density of hydrogen, such a network holds only about a quarter of the line-pack. Thus, velocities need to increase to fulfill the same demand, and the control becomes much more dynamic. Employees of the gas dispatchers cannot scale such increasing complexity and scope at will. The MODAL EnergyLab navigation system will provide solutions for these changing demands of energy dispatching.

CO@WORK Online 2020

CO@Work is one of the most prominent training and outreach activities of MODAL. In 2020, this time in the form of a massive online open course, it attracted 1,154 participants from 68 countries with a rich program including 50 lectures by 38 lecturers as well as daily hands-on exercise sessions via Jupyter Notebooks running in parallel in 1,200 instances in a Kubernetes cloud hosted at ZIB.

The CO@WORK summer school addresses master's students (in their final year), PhD students, postdocs, and everyone else interested in combinatorial optimization and mathematical programming in industrial applications. We originally planned to hold it as a physical event with approximately one-hundred participants in Berlin like several times before. However, motivated by the COVID-19 crisis, the whole 2020 summer school was transferred to a MOOC (massive online open course).

CO@WORK 2020 had 1,154 participants from 68 countries in 17 different time zones (see Figures 10 and 11). About 54% of participants were PhD students, 22% undergrad students, 15% faculty, and 9% from industry.

To make CO@Work accessible worldwide, we made it free of charge and undertook significant program changes. While for on-site workshops, one can have eight straight hours of lecturing (interleaved with breaks), this is not expedient when lecturers and students are spread worldwide. The team at MODAL, particularly the lead organizers Dr. Timo

Figure 10: Distribution of CO@Work participants around the globe.

Figure 11: Snapshot of 275 out of more than 1,000 participants.

Berthold (FICO) and Professor Thorsten Koch (ZIB), came up with some unique solutions to transfer the mix of theoretical lectures, panel discussions, and hands-on programming experience into an online format.

The 50 lectures by 38 lecturers were all prerecorded and released as daily content. This included 29 classes from the MODAL partner institutions Siemens, SAP, Gurobi, FICO, GAMS, FU Berlin, and ZIB. This daily video feed enabled the students to watch videos when it suited them and at their own pace. Twice a day, with an eleven-hour time shift, we had a 30-minute Q&A session on the lectures of the previous day and a 90-minute online exercise session.

We uploaded all lectures to a YouTube channel, where they are still available for interested viewers (https://www.youtube.com/channel/UCphLz_BXrOAInHozAlTsigA/playlists). Additionally, the talks were transferred to bilibili.com, a Chinese video-streaming site, as YouTube is blocked in China. The most popular presentations received more than 2,000 views, some of them still getting more than 100 views per week as of March 2021. To reduce barriers, we produced proper subtitles for all 50 lecture videos; around two-thirds of the students reported that they made use of this feature. During the preparation of the course, we made extensive experiments with automatic extraction and translation of subtitles, as well as with using synthesized speech generated from subtitles. Some courses were then used as a test bed to get feedback on the respective user experience.

Q&A and the exercise sessions took place as Zoom webinars. The live programming exercises were accompanied by four virtual breakout rooms, where tutors helped students interactively when they got stuck. We wanted to guarantee that all students would have ready access to the required software and an identical setup to not lose any time and motivation through tedious technical support. Therefore, we implemented all exercises via Jupyter Notebooks, which ran inside a Docker image. The IT team at ZIB implemented a system to run 1,200 instances of this Docker image in parallel in a Kubernetes cloud hosted at ZIB. Each student received their own personal, identical instance to work in.

The summer school was very well received, and we all learned a lot about online teaching. It is quite possible that next time, CO@Work will take place as a hybrid event to combine the accessibility feats of an online format with the networking possibilities of a physical event.

REFERENCES

Research to Fight COVID-19

[1] Sebastian A Müller, Michael Balmer, Billy Charlton, Ricardo Ewert, Andreas Neumann, Christian Rakow, Tilmann Schlenther, and Kai Nagel. (2020) *Using mobile Phone Data for Epidemiological Simulations of Lockdowns: Government Interventions, Behavioral Changes, and Resulting Changes of Reinfections.* medRxiv, doi:10.1101/2020.07.22.20160093

[2] Hanna Wulkow, Tim Conrad, Natasa Djurdjevac Conrad, Sebastian A. Mueller, Kai Nagel, and Christof Schuette (2020) *Prediction of COVID-19 Spreading and Optimal Coordination of Counter-Measures: from Microscopic to Macroscopic Models to Pareto Fronts.* medRxev, doi: 10.1101/2020.12.01.20241885

Understanding and Modeling Complex Biological Systems

[Weimann2021] Weimann, K. and Conrad, T. O. F. (2021) Transfer Learning for ECG Classification. Scientific Reports. ISSN 2045-2322 (In Press)

[Melnyk2020] Melnyk, Kateryna and Klus, S. and Montavon, Grègoire and Conrad, T. O. F. (2020) Graph Kernel Koopman Embedding for Human Microbiome Analysis. Applied Network Science, 5 (96)

[Rams2020] Rams, Mona and Conrad, T. O. F. (2020) Dictionary Learning for Transcriptomics Data Reveals Type-Specific Gene Modules in a Multi-Class Setting. Information Technology

[Iravani2019] Iravani, Sahar and Conrad, T. O. F. (2019) Deep Learning for Proteomics Data for Feature Selection and Classification. Lecture Notes in Computer Science, 11713 .

[Zhang2019] Zhang, W. and Klus, S. and Conrad, T. O. F. and Schütte, Ch. (2019) Learning Chemical Reaction Networks from Trajectory Data. *SIAM Journal on Applied Dynamical Systems*, 18(4). ISSN 1536-0040

[Donati2020] L. Donati, M. Weber, B. G. Keller: Markov Models from the Square Root Approximation of the Fokker-Planck Equation: Calculating the Grid-Dependent Flux. *Journal of Physics: Condensed Matter,* https://doi.org/10.1088/1361-648X/abd5f7, 2020

[Ernst2020] N. Ernst, K. Fackeldey, A. Volkamer, O. Opitz, M. Weber: Computation of Temperature-Dependent Dissociation Rates of Metastable Protein-Ligand Complexes. *Molecular Simulation,* 45(11):904-911, 2019.

[Rabben2020] R. J. Rabben, S. Ray, M. Weber: ISOKANN: Invariant Subspaces of Koopman Operators Learned by a Neural Network. *The J. of Chem. Phys.,* 153(11):114109, 2020.

[Ray2020] S. Ray, V. Sunkara, Ch. Schütte, and M. Weber: How to Calculate pH-Dependent Binding Rates for Receptor-Ligand Systems Based on Thermodynamic Simulations with Different Binding Motifs. *Molecular Simulation,* Vol. 46, Issue 18, 2020

[Bittracher2020a] Andreas Bittracher and Christof Schütte, "A Weak Characterization of Slow Variables in Stochastic Dynamical Systems," *Advances in Dynamics, Optimization and Computation*, ed. Oliver Junge et al., Studies in Systems, Decision and Control (Springer International Publishing), 132–50, 2020, https://doi.org/10.1007/978-3-030-51264-4_6.

[Bittracher2020b] Andreas Bittracher, Stefan Klus, Boumediene Hamzi, Péter Kolta and Christof Schütte, "Dimensionality Reduction of Complex Metastable Systems via Kernel Embeddings of Transition Manifolds," *Journal of Nonlinear Science,* 31, no. 1, 2020, https://doi.org/10.1007/s00332-020-09668-z.

[Bittracher2021] Andreas Bittracher and Christof Schütte, "A Probabilistic Algorithm for Aggregating Vastly Undersampled Large Markov Chains," *Physica D: Nonlinear Phenomen*a 416 (February 1, 2021): 132799, https://doi.org/10.1016/j.physd.2020.132799.

[Brunton2016a] S.L. Brunton, J.L. Proctor, and J.N. Kutz: Discovering Governing Equations from Data by Sparse Identification of Nonlinear Dynamical Systems. Proceedings of the national academy of sciences, 113(15):3932–3937, 2016.

[Brunton2016b] S. L. Brunton, B. W. Brunton, J. L. Proctor, and J. Kutz: Koopman Invariant Subspaces and Finite Linear Representations of Nonlinear Dynamical Systems for Control. PLoS ONE, 11(2):e0150171, 2016.

[Klus2018] St. Klus, F. Nüske, P. Koltai, H. Wu, I. Kevrekidis, Ch. Schütte, and F. Noé: Data-Driven Model Reduction and Transfer Operator Approximation. Nonlinear Science 28(3), pp.985–1010, 2018

[Gelß2019] P. Gelß, St. Klus, J. Eisert, Ch. Schütte: Multidimensional Approximation of Nonlinear Dynamical Systems. J. Comput. Nonlinear Dynam. 14(6), 061006, 2019

[Wulkow2020] N. Wulkow, P. Koltai, Ch. Schütte: Memory-Based Reduced Modeling and Data-Based Estimation of Opinion Spreading. Nonlinear Science 31(19), 2020

[Klus2020] St. Klus, F. Nüske, S. Peitz, J.-H. Niemann, C. Clementi, Ch. Schütte: Data-Driven Approximation of the Koopman Generator: Model Reduction, System Identification, and Control. Physica D: Nonlinear Phenomena, Volume 406:132416, 2020

[Carderera2021] A. Carderera, S. Pokutta, C. Schütte, M. Weiser: CINDy: Conditional Gradient-Based Identification of Non-linear Dynamics – Noise-Robust Recovery. arXiv:2101.02630, 2021.

Wind under Your Wings – Discrete-Continuous Free Flight Planning

Oxley, D., & Jain, C. (2015). Chapter 1.4 Global Air Passenger Market: Riding Out Periods of Turbulence, Travel and Tourism Competitiveness Report 2015. International Air Transport Association. World Economic Forum.

C.A. Wells, P.D. Williams, N.K. Nichols, D. Kalise, I. Poll. Reducing transatlanticflight emissions by fuel-optimised routing. Env. Res. Letters 16:025002, 2021.

M. Blanco, R. Borndörfer, N.-D. Hoang, A. Kaier, A. Schienle, T. Schlechte, and S. Schlobach. Solving Time Dependent Shortest Path Problems on Airway Networks Using Super-Optimal Wind. 16th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS 2016).

R. Borndörfer, F. Danecker, and M. Weiser. A Discrete-Continuous Algorithm for Free Flight Planning. Algorithms 14(1):4, 2021.

R. Borndörfer, F. Danecker, and M. Weiser. A Priori Error Bounds for Graph-Based Trajectory Optimization in Free Flight Planning. In preparation 2021.

P. Deuflhard, M. Weiser. Numerische Mathematik 3: Adaptive Lösung partieller Differentialgleichungen. de Gruyter 2020.

C.A. Wells, P.D. Williams, N.K. Nichols, D. Kalise, and I. Poll. Reducing Transatlantic Flight Emissions by Fuel-Optimised Routing. Env. Res. Letters 16:025002, 2021.

IATA, https://www.iata.org/en/iata-repository/publications/economic-reports/global-air-passenger-markets-riding-out-periods-of-turbulence/

Simple Explanations, Sparsity, and Conditional Gradients

(Braun et al., 2017) G. Braun, S. Pokutta, and D. Zink. Lazifying Conditional Gradient Algorithms. *Proceedings of the 34th International Conference on Machine Learning,* pp. 566–575, 2017.

(Braun et al., 2019) G. Braun, S. Pokutta, D. Tu, and S. Wright. Blended Conditional Gradients: the Unconditioning of Conditional Gradients. *Proceedings of the 36th International Conference on Machine Learning,* 2019.

(Brunton et al., 2016) Steven. L. Brunton, Joshua L. Proctor, and J. Nathan Kutz (2016). Discovering Governing Equations from Data by Sparse Identification of Nonlinear Dynamical Systems. *Proceedings of the national academy of sciences,* 113(15), pp. 3932–3937, 2016.

(Carderera et al. 2021) Alejandro Carderera, Sebastian Pokutta, Christof Schütte, and Martin Weiser (2021). CINDy: Conditional Gradient-Based Identification of Non-linear Dynamics – Noise-Robust Recovery. Preprint.

(Combettes et al., 2020) Cyrille W. Combettes, Christoph Spiegel, and Sebastian Pokutta. Projection-Free Adaptive Gradients for Large-Scale Optimization. Preprint, 2020.

(Combettes and Pokutta, 2021) Cyrille W. Combettes and Sebastian Pokutta. Complexity of Linear Minimization and Projection on Some Sets. Preprint, 2021.

(Defazio and Bottou, 2019) Aaron Defazio and Leon Bottou. On the Ineffectiveness of Variance Reduced Optimization for Deep Learning. *Advances in Neural Information Processing Systems 32*, pp. 1755–1765, 2019.

(Duchi et al., 2011) John C. Duchi, Elad Hazan, and Yoram Singer. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research*, 12(61), pp. 2121–2159, 2011.

(Frank and Wolfe, 1956) Marguerite Frank and Philip Wolfe. An Algorithm for Quadratic Programming. *Naval Research Logistics Quarterly*, 3(1-2), pp. 95–110, 1956.

(Grant et al., 2006) Michael Grant, Stephen Boyd, and Yinyu Ye. Disciplined Convex Programming, Chapter in Global Optimization: From Theory to Implementation, L. Liberti and N. Maculan, eds., in the book series Nonconvex Optimization and Applications, Springer, pp. 155–210, 2006.

(Hazan and Luo, 2016) Elad Hazan and Haipeng Luo. Variance-Reduced and Projection-Free Stochastic Optimization. *Proceedings of the 33rd International Conference on Machine Learning*, pp. 1263–1271, 2016.

(Holloway, 1974) C. A. Holloway. An Extension of the Frank and Wolfe Method of Feasible Directions. *Mathematical Programming*, 6, pp. 14–27, Dec. 1974. doi:10.1007/BF01580219.

(see, e.g., Jaggi, 2013) Martin Jaggi. Revisiting Frank-Wolfe: Projection-Free Sparse Convex Optimization. *International Conference on Machine Learning*, pp. 427-435, 2013 PMLR.

(Lan and Zhou, 2016) Guanghui Lan and Yi Zhou. Conditional Gradient Sliding for Convex Optimization. *SIAM Journal on Optimization*, 26(2), pp.1379–1409, 2016.

(Levitin and Polyak, 1966) Evgeny S. Levitin and Boris T. Polyak. Constrained Minimization Methods. *USSR Computational mathematics and mathematical physics* 6.5, pp. 1–50, 1966.

(McMahan and Streeter, 2010) H. Brendan McMahan and Matthew Streeter. Adaptive Bound Optimization for Online Convex Optimization. *Proceedings of the 23rd Annual Conference on Learning Theory*, 2010.

(Négiar et al., 2020) Geoffrey Négiar, Gideon Dresdner, Alicia Tsai, Laurent El Ghaoui, Francesco Locatello, Robert M. Freund, Fabian Pedregosa. Stochastic Frank-Wolfe for Constrained Finite-Sum Minimization. *Proceedings of the 37th International Conference on Machine Learning,* pp. 7253–7262, 2020.

(Xie et al., 2020) Jiahao Xie, Zebang Shen, Chao Zhang, Boyu Wang, and Hui Qian. Efficient Projection-Free Online Methods with Stochastic Recursive Gradient. *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, pp. 6446–6453, 2020.

(Yurtsever et al, 2019) Alp Yurtsever, Suvrit Sra, and Volkan Cevher. Conditional Gradient Methods via Stochastic Path-Integrated Differential Estimator. *Proceedings of the 36th International Conference on Machine Learning,* pp. 7282–7291, 2019.

Research Campus MODAL: Success Stories

[B18] Isabel Beckenbach. *A Hypergraph Network Simplex Algorithm*. Operations Research Proceedings 2017, P. 309–316, 2018.

[B19] Isabel Beckenbach. *Matchings and Flows in Hypergraph*s. Dissertation, Freie Universität Berlin, 2019.

[BB18] Isabel Beckenbach, Ralf Borndörfer. *Hall's and Kőnig's Theorem in Graphs and Hypergraphs*. Discrete Mathematics 341(10), P. 2753–2761, 2018.

[BGJ10] Ralf Borndörfer, Martin Grötschel, Ulrich Jäger. *Planning Problems in Public Transit*. Production Factor Mathematics, acatech und Springer, P. 95–122, 2010.

[BG+17] Ralf Borndörfer, Boris Grimm, Markus Reuther, Thomas Schlechte. *Template-Based Re-optimization of Rolling Stock Rotations*. Public Transport 9(1-2), P. 365–383, 2017

[BK+18] Ralf Borndörfer, Torsten Klug, Leonardo Lamorgese, Carlo Mannino, Markus Reuther, Thomas Schlechte (Eds.). *Handbook of Optimization in the Railway Industry*. Springer International Publishing, 2018.

[BM+14] Ralf Borndörfer, Julika Mehrgardt, Markus Reuther, Thomas Schlechte, Kerstin Waas. *Re-Optimization of Rolling Stock Rotations*. Operations Research Proceedings 2013, P. 49–55, 2014.

[BRS14] Ralf Borndörfer, Markus Reuther, Thomas Schlechte. *A Coarse-to-Fine Approach to the Railway Rolling Stock Rotation Problem*. 14th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems, OpenAccess Series in Informatics (OASIcs), (42) P. 79–91, 2014.

[BR+11] Ralf Borndörfer, Markus Reuther, Thomas Schlechte, Steffen Weider. *A Hypergraph Model for Railway Vehicle Rotation Planning*. 11th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems, OpenAccess Series in Informatics (OASIcs) (20), P. 146–155, 2011.

[BR+12] Ralf Borndörfer, Markus Reuther, Thomas Schlechte, Steffen Weider. *Vehicle Rotation Planning for Intercity Railways*. Proceedings of Conference on Advanced Systems for Public Transport 2012 (CASPT12), 2012.

[BR+15] Ralf Borndörfer, Markus Reuther, Thomas Schlechte, Kerstin Waas, Steffen Weider. *Integrated Optimization of Rolling Stock Rotations for Intercity Railways*. Transportation Science, 50(3), P. 863–877, 2015.

[E82] Franz Edelman. *Managers, Computer Systems, and Productivity.* Interfaces 12(5), P. 35–46, 1982.

PUBLICATIONS

PEER-REVIEWED (156)

*Samer Alhaddad, Jens Förstner, Stefan Groth, Daniel Grünewald, Yevgen Grynko, Frank Hannig, Tobias Kenter, Franz-Josef Pfreundt, Christian Plessl, Merlind Schotte, Thomas Steinke, Jürgen Teich, Martin Weiser, Florian Wende (2020). HighPerMeshes – A Domain-Specific Language for Numerical Algorithms on Unstructured Grids. *Euro-Par 2020*. (Joint publication: Modeling and Simulation of Complex Processes, Distributed Algorithms and Supercomputing) (accepted for publication 2020-07-22)

Sónia Alves, Rainald Ehrig, Peter C. Raffalt, Alwina Bender, Georg N. Duda, Alison N. Agres (2020). Quantifying Asymmetry in Gait: A Weighted Universal Symmetry Index to Evaluate 3D Ground Reaction Forces. *Frontiers in Bioengineering and Biotechnology*. (accepted for publication 2020-09-18)

N. Anari, N. Haghtalab, S. Naor, Sebastian Pokutta, M. Singh, A. Torrico (2020). Structured Robust Submodular Maximization: Offline and Online Algorithms. *INFORMS Journal on Computing*.

*H. Anzt, F. Bach, S. Druskat, F. Löffler, A. Loewe, B. Y. Renard, G. Seemann, A. Struck, E. Achhammer, F. Appell, M. Bader, L. Brusch, C. Busse, G. Chourdakis, P. W. Dabrowski, P. Ebert, B. Flemisch, S. Friedl, B. Fritzsch, M. D. Funk, V. Gast, F. Goth, J.-N. Grad, Sibylle Hermann, F. Hohmann, S. Janosch, D. Kutra, J. Linxweiler, T. Muth, Wolfgang Peters-Kottig, F. Rack, F. H. C. Raters, S. Rave, G. Reina, M. Reißig, T. Ropinski, J. Schaarschmidt, H. Seibold, J. P. Thiele, B. Uekermann, S. Unger, R. Weeber (2020). An environment for sustainable research software in Germany and beyond: current state, open challenges, and call for action. *F1000Research*. https://doi.org/10.12688/f1000research.23224.1 (Joint publication: AI in Society, Science, and Technology, Digital Data and Information for Society, Science, and Culture)

Bahareh Banyassady, Luis Barba, Wolfgang Mulzer (2020). Time-Space Trade-Offs for Computing Euclidean Minimum Spanning Trees. *Journal of Computational Geometry (JoCG)*, 11(1):525-547.

Christiane Becker, Klaus Jäger, Sven Burger (2021). Nanophotonics for Solar Energy. *Sol. Energy Mater. Sol. Cells*, 221:110916. https://doi.org/10.1016/j.solmat.2020.110916 (epub ahead of print 2020-12-15)

Fabian Becker, Natasa Djurdjevac Conrad, Raphael A. Eser, Luzie Helfmann, Brigitta Schütt, Christof Schütte, Johannes Zonker (2020). The Furnace and the Goat—A spatio-temporal model of the fuelwood requirement for iron metallurgy on Elba Island, 4th century BCE to 2nd century CE. *PLOS ONE*, 15:1-37. https://doi.org/10.1371/journal.pone.0241133

Felix Binkowski, Fridtjof Betz, Rémi Colom, Martin Hammerschmidt, Lin Zschiedrich, Sven Burger (2020). Quasinormal mode expansion of optical far-field quantities. *Phys. Rev. B*, 102:035432. https://doi.org/10.1103/PhysRevB.102.035432

Felix Binkowski, Fridtjof Betz, Rémi Colom, Martin Hammerschmidt, Lin Zschiedrich, Sven Burger (2020). Modal expansion of optical far-field quantities using quasinormal modes. In *EPJ Web Conf.*, 238:05007. https://doi.org/10.1051/epjconf/202023805007

Felix Binkowski, Lin Zschiedrich, Sven Burger (2020). A Riesz-projection-based method for nonlinear eigenvalue problems. *J. Comput. Phys.*, 419:109678. https://doi.org/10.1016/j.jcp.2020.109678

Andreas Bittracher, Christof Schütte (2020). A weak characterization of slow variables in stochastic dynamical systems. In Oliver Junge, O. Schütze, Gary Froyland, S. Ober-Blobaum, E. Padberg-Gehle, editors, *Advances in Dynamics, Optimization and Computation. Series: Studies in Systems, Decision and Control.* 304:132-150.

Horst-Holger Boltz, Jorge Kurchan, Andrea J. Liu (2021). Fluctuation distributions of energy minima in complex landscapes. *Physical Review Research*, 3(1):013061. https://doi.org/10.1103/PhysRevResearch.3.013061 (accepted for publication 2020-12-18)

Luc Bonnet, Jean-Luc Akian, Éric Savin, T. J. Sullivan (2020). Adaptive reconstruction of imperfectly-observed monotone functions, with applications to uncertainty quantification. *Algorithms*, 13(8):196. https://doi.org/10.3390/a13080196

*Ralf Borndörfer, Fabian Danecker, Martin Weiser (2021). A Discrete-Continuous Algorithm for Free Flight Planning. *Algorithms*, 14(1):4. https://doi.org/10.3390/a14010004 (Joint publication: Modeling and Simulation of Complex Processes, Network Optimization) (epub ahead of print 2020-12-25)

Ralf Borndörfer, Heide Hoppmann, Marika Karbstein, Niels Lindner (2020). Separation of cycle inequalities in periodic timetabling. *Discrete Optimization*, 100552. https://doi.org/10.1016/j.disopt.2019.100552

Ralf Borndörfer, Niels Lindner, Sarah Roth (2020). A Concurrent Approach to the Periodic Event Scheduling Problem. *Journal of Rail Transport Planning & Management*, 100175. https://doi.org/10.1016/j.jrtpm.2019.100175

Ralf Borndörfer, Ricardo Euler, Marika Karbstein (2020). Ein Graphen-basiertes Modell zur Beschreibung von Preissystemen im öffentlichen Nahverkehr. In Volume 002/127 of *FGSV Heureka*, pages 1-15.

Ralf Borndörfer, Thomas Eßer, Patrick Frankenberger, Andreas Huck, Christoph Jobmann, Boris Krostitz, Karsten Kuchenbecker, Kai Moorhagen, Philipp Nagl, Michael Peterson, Markus Reuther, Thilo Schang, Michael Schoch, Hanno Schülldorf, Peter Schütz, Tobias Therolf, Kerstin Waas, Steffen Weider (2020). Deutsche Bahn Schedules Train Rotations Using Hypergraph Optimization. *Informs Journal on Applied Analytics*. (accepted for publication 2020-11-05)

Jan Brüning, Thomas Hildebrandt, Werner Heppt, Nora Schmidt, Hans Lamecker, Angelika Szengel, Natalja Amiridze, Heiko Ramm, Matthias Bindernagel, Stefan Zachow, Leonid Goubergrits (2020). Characterization of the Airflow within an Average Geometry of the Healthy Human Nasal Cavity. *Scientific Reports*, 3755(10). https://doi.org/10.1038/s41598-020-60755-3

Jens Buchmann, Bernhard Kaplan, Samuel Powell, Steffen Prohaska, Jan Laufer (2020). Quantitative PA tomography of high resolution 3-D images: experimental validation in tissue phantoms. *Photoacoustics*, 17:100157. https://doi.org/10.1016/j.pacs.2019.100157

Júlia Chaumel, Merlind Schotte, Joseph J. Bizzarro, Paul Zaslansky, Peter Fratzl, Daniel Baum, Mason N. Dean (2020). Co-aligned chondrocytes: Zonal morphological variation and structured arrangement of cell lacunae in tessellated cartilage. *Bone*, 134:115264. https://doi.org/10.1016/j.bone.2020.115264 (epub ahead of print 2020-02-11)

Fatemeh Chegini, Alena Kopanicakova, Rolf Krause, Martin Weiser (2020). Efficient Identification of Scars using Heterogeneous Model Hierarchies. *EP Europace*. https://doi.org/10.1093/europace/euaa402 (accepted for publication 2020-12-09)

Ying Chen, Nazgul Zakiyeva, Bangzhu Zhu, Thorsten Koch (2020). Modeling and Forecasting the Dynamics of the Natural Gas Transmission Network in Germany with the Demand and Supply Balance Constraint. *Applied Energy*. https://doi.org/10.1016/j.apenergy.2020.115597

Ying Chen, Xiuqin Xu, Thorsten Koch (2020). Day-ahead high-resolution forecasting of natural gas demand and supply in Germany with a hybrid model. *Applied Energy*, 262(114486). https://doi.org/https://doi.org/10.1016/j.apenergy.2019.114486

Surahit Chewle, Franziska Emmerling, Marcus Weber (2020). Effect of choice of solvent on crystallization pathway of Paracetamol: An experimental and theoretical case study. *Crystals*, 10(12):1107. https://doi.org/10.3390/cryst10121107

Steffen Christgau, Bettina Schnor (2020). Comparing MPI Passive Target Synchronization on a Non-Cache-Coherent Shared-Memory Processor. *Mitteilungen – Gesellschaft für Informatik e. V., Parallel-Algorithmen und Rechnerstrukturen, ISSN 0177-0454, 28. PARS-Workshop*, 121-132.

Steffen Christgau, Thomas Steinke (2020). Porting a Legacy CUDA Stencil Code to oneAPI. In *2020 IEEE International Parallel and Distributed Processing Symposium Workshops, IPDPSW 2020, New Orleans, LA, USA, May 18-22, 2020*, 359-367. https://doi.org/10.1109/IPDPSW50202.2020.00070

Steffen Christgau, Thomas Steinke (2020). Leveraging a Heterogenous Memory System for a Legacy Fortran Code: The Interplay of Storage Class Memory, DRAM and OS. In *2020 IEEE/ACM Workshop on Memory Centric High Performance Computing (MCHPC)*, 17-24. https://doi.org/10.1109/MCHPC51950.2020.00008

Rémi Colom, Felix Binkowski, Fridtjof Betz, Martin Hammerschmidt, Lin Zschiedrich, Sven Burger (2020). Quasi-normal mode expansion as a tool for the design of nanophotonic devices. In *EPJ Web Conf.*, 238:05008. https://doi.org/10.1051/epjconf/202023805008

CyrilleW Combettes, Sebastian Pokutta (2020). Boosting Frank-Wolfe by Chasing Gradients. In *Proceedings of ICML*.

Zsolt Csizmadia, Timo Berthold (2020). The confined primal integral: a measure to benchmark heuristic MINLP solvers against global MINLP solvers. *Mathematical Programming*. https://doi.org/10.1007/s10107-020-01547-5

Wolfgang Dalitz, Wolfram Sperber, Hagen Chrapary (2020). swMATH: A Publication-based Approach to Mathematical Software. *SIAM Newsletter*, Volume 53 (Number 06 | July/August 2020). https://doi.org/10.12752/8009 (accepted for publication 2020-05-31)

Jo Devriendt, Ambros Gleixner, Jakob Nordström (2020). Learn to Relax: Integrating 0-1 Integer Linear Programming with Pseudo-Boolean Conflict-Driven Search. In *Integration of AI and OR Techniques in Constraint Programming. CPAIOR 2020*, Volume 12296 of *LNCS*, pages xxiv-xxvi.

Jelena Diakonikolas, Alejandro Carderera, Sebastian Pokutta (2020). Locally Accelerated Conditional Gradients. In *Proceedings of AISTATS*.

Yannic Ege, Christian Foth, Daniel Baum, Christian S. Wirkner, Stefan Richter (2020). Making spherical-harmonics-based Geometric Morphometrics (SPHARM) approachable for 3D images containing large cavity openings using Ambient Occlusion – a study using hermit crab claw shape variability. *Zoomorphology*, 139:421-432. https://doi.org/10.1007/s00435-020-00488-z

Gunar Fabig, Robert Kiewisz, Norbert Lindow, James A. Powers, Vanessa Cota, Luis J. Quintanilla, Jan Brugués, Steffen Prohaska, Diana S. Chu, Thomas Müller-Reichert (2020). Sperm-specific meiotic chromosome segregation in C. elegans. *eLife*, 9:e50988. https://doi.org/10.7554/eLife.50988

Konstantin Fackeldey, Jonas Röhm, Amir Niknejad, Surahit Chewle, Marcus Weber (2021). Analyzing Raman Spectral Data without Separabiliy Assumption. *Journal of Mathematical Chemistry*, 3(59):575-596. https://doi.org/10.1007/s10910-020-01201-7 (accepted for publication 2020-11-25)

Nando Farchmin, Martin Hammerschmidt, Philipp-Immanuel Schneider, Matthias Wurm, Bernd Bodermann, Markus Bär, Sebastian Heidenreich (2020). Efficient Bayesian inversion for shape reconstruction of lithography masks. *J. Micro/Nanolith. MEMS MOEMS*, 19:024001. https://doi.org/10.1117/1.JMM.19.2.024001

Joshua Feis, Dominik Beutel, Julian Köpfler, Xavier Garcia Santiago, Carsten Rockstuhl, Martin Wegener, Ivan Fernandez-Corbaton (2020). Helicity-Preserving Optical Cavity Modes for Enhanced Sensing of Chiral Molecules. *Phys. Rev. Lett.*, 124:033201. https://doi.org/10.1103/PhysRevLett.124.033201

Bernhard Fröhler, Tim Elberfeld, Torsten Möller, Hans-Christian Hege, Jan De Beenhouwer, Jan Sijbers, Johann Kastner, Christoph Heinzl (2020). Analysis and comparison of algorithms for the tomographic reconstruction of curved fibres. *Nondestructive Testing and Evaluation*, 35(3):328-341. https://doi.org/10.1080/10589759.2020.1774583

Gerald Gamrath, Timo Berthold, Domenico Salvagnin (2020). An exploratory computational analysis of dual degeneracy in mixed-integer programming. *EURO Journal on Computational Optimization*, 241-246. https://doi.org/10.1007/s13675-020-00130-z

Patrick Gemander, Wei-Kun Chen, Dieter Weninger, Leona Gottwald, Ambros Gleixner (2020). Two-row and two-column mixed-integer presolve using hashing-based pairing methods. *EURO Journal on Computational Optimization*, 8(3-4):205-240. https://doi.org/10.1007/s13675-020-00129-6

Ambros Gleixner, Stephen Maher, Benjamin Müller, João Pedro Pedroso (2020). Price-and-verify: a new algorithm for recursive circle packing using Dantzig–Wolfe decomposition. *Annals of Operations Research*, 284(2):527-555. https://doi.org/10.1007/s10479-018-3115-5

Ambros Gleixner, Nils-Christian Kempke, Thorsten Koch, Daniel Rehfeldt, Svenja Uslu (2020). First Experiments with Structure-Aware Presolving for a Parallel Interior-Point Method. In *Operations Research Proceedings 2019*, 105-111. https://doi.org/10.1007/978-3-030-48439-2_13

Ambros Gleixner, Daniel Steffy (2020). Linear Programming using Limited-Precision Oracles. *Mathematical Programming*, 183(1-2):525-554. https://doi.org/10.1007/s10107-019-01444-6

Ambros Gleixner, Gregor Hendel, Gerald Gamrath, Tobias Achterberg, Michael Bastubbe, Timo Berthold, Philipp M. Christophel, Kati Jarck, Thorsten Koch, Jeff Linderoth, Marco Lübbecke, Hans Mittelmann, Derya Ozyurt, Ted Ralphs, Domenico Salvagnin, Yuji Shinano (2020). MIPLIB 2017: Data-Driven Compilation of the 6th Mixed-Integer Programming Library. *Mathematical Programming Computation*. https://doi.org/10.1007/s12532-020-00194-3 (accepted for publication 2020-09-10)

Christoph Gorgulla, Konstantin Fackeldey, Gerhard Wagner, Haribabu Arthanari (2020). Accounting of Receptor Flexibility in Ultra-Large Virtual Screens with VirtualFlow Using a Grey Wolf Optimization Method. *Supercomputing Frontiers and Innovations*, 7(3):4-12. https://doi.org/10.14529/jsfi200301

Christoph Gorgulla, Andras Boeszoermnyi, Zi-Fu Wang, Patrick D. Fischer, Paul Coote, Krishna M. Padmanabha Das, Yehor S. Malets, Dmytro S. Radchenko, Yurii Moroz, David A. Scott, Konstantin Fackeldey, Moritz Hoffmann, Iryna Iavniuk, Gerhard Wagner, Haribabu Arthanari (2020). An open-source drug discovery platform enables ultra-large virtual screens. *Nature*, 580:663-668. https://doi.org/https://doi.org/10.1038/s41586-020-2117-z

Uwe Gotzes, Kai Hoppmann-Baum (2020). Bounding the final rank during a round robin tournament with integer programming. *Operational Research*, 1866-1505. https://doi.org/https://doi.org/10.1007/s12351-020-00546-w

Andreas Griewank, Tom Streubel, Caren Tischendorf (2020). On the abs-polynomial expansion of piecewise smooth functions. *Optimization Methods and Software*. https://doi.org/10.1080/10556788.2020.1817448 (epub ahead of print 2020-09-01)

Philipp Gutsche, Xavier Garcia Santiago, Philipp-Immanuel Schneider, Kevin McPeak, Manuel Nieto-Vesperinas, Sven Burger (2020). Role of Geometric Shape in Chiral Optics. *Symmetry*, 12:158. https://doi.org/10.3390/sym12010158

Sebastian Götschel, Anton Schiela, Martin Weiser (2021). Kaskade 7 – a Flexible Finite Element Toolbox. *Computers and Mathematics with Applications*, 81:444-458. https://doi.org/10.1016/j.camwa.2020.02.011 (epub ahead of print 2020-03-07)

Martin Hammerschmidt, Sandra Döpking, Sven Burger, Sebastian Matera (2020). Field Heterogeneities and Their Impact on Photocatalysis: Combining Optical and Kinetic Monte Carlo Simulations on the Nanoscale. *J. Phys. Chem. C*, 124:3177. https://doi.org/10.1021/acs.jpcc.9b11469

Martin Hanik, Hans-Christian Hege, Christoph von Tycowicz (2020). Bi-invariant Two-Sample Tests in Lie Groups for Shape Analysis. In *Shape in Medical Imaging*, 44-54. https://doi.org/10.1007/978-3-030-61056-2_4

Martin Hanik, Hans-Christian Hege, Anja Hennemuth, Christoph von Tycowicz (2020). Nonlinear Regression on Manifolds for Shape Analysis using Intrinsic Bézier Splines. In *Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 617-626. https://doi.org/10.1007/978-3-030-59719-1_60

Luzie Helfmann, Natasa Djurdjevac Conrad, Ana Djurdjevac, Stefanie Winkelmann, Christof Schütte (2021). From interacting agents to density-based modeling with stochastic PDEs. *Communications in Applied Mathematics and Computational Science*, 16(1):1-32. https://doi.org/10.2140/camcos.2021.16.1 (accepted for publication 2020-11-01)

Luzie Helfmann, Enric Ribera Borrell, Christof Schütte, Peter Koltai (2020). Extending Transition Path Theory: Periodically Driven and Finite-Time Dynamics. *Journal of Nonlinear Science*, 30:3321-3366. https://doi.org/https://doi.org/10.1007/s00332-020-09652-7

Felix Hennings, Lovis Anderson, Kai Hoppmann-Baum, Mark Turner, Thorsten Koch (2020). Controlling transient gas flow in real-world pipeline intersection areas. *Optimization and Engineering*. https://doi.org/https://doi.org/10.1007/s11081-020-09559-y (epub ahead of print 2020-10-03)

Margaux Hoang, Philippe Garnier, Jeremie Lasue, Henri Rème, Maria Teresa Capria, Kathrin Altwegg, Matthias Läuter, Tobias Kramer, Martin Rubin (2020). Investigating the Rosetta/RTOF observations of comet 67P/Churyumov-Gerasimenko using a comet nucleus model: Influence of dust mantle and trapped CO. *Astronomy & Astrophysics*, 638:A106. https://doi.org/10.1051/0004-6361/201936655

Kai Hoppmann-Baum, Gioni Mexi, Oleg Burdakov, Carl Johan Casselgren, Thorsten Koch (2020). Minimum Cycle Partition with Length Requirements. In Emmanuel Hebrard, Nysret Musliu, editors, *Integration of Constraint Programming, Artificial Intelligence, and Operations Research*, Volume 12296 of *Lecture Notes in Computer Science*, pages 273-282, Springer International Publishing. https://doi.org/10.1007/978-3-030-58942-4_18

Felix Höfling, Siegfried Dietrich (2020). Finite-size corrections for the static structure factor of a liquid slab with open boundaries. *The Journal of Chemical Physics*, 153:054119. https://doi.org/10.1063/5.0017923

Carmen Juds, Johannes Schmidt, Michael Weller, Thorid Lange, Tim Conrad, Hans Boerner (2020). Combining Phage Display and Next-generation Sequencing for Materials Sciences: A Case Study on Probing Polypropylene Surfaces. *Journal of the American Chemical Society*. https://doi.org/https://pubs.acs.org/doi/10.1021/jacs.0c03482

Klaus Jäger, Peter Tillmann, Eugene A. Katz, Christiane Becker (2021). Perovskite/silicon tandem solar cells: Effect of luminescent coupling and bifaciality. *Sol. RRL*, 5:2000628. https://doi.org/10.1002/solr.202000628 (epub ahead of print 2020-12-12)

Klaus Jäger, Peter Tillmann, Christiane Becker (2020). Detailed illumination model for bifacial solar cells. *Opt. Express*, 28:4751. https://doi.org/10.1364/OE.383570

Klaus Jäger, Peter Tillmann, Eugene A. Katz, Christiane Becker (2020). Simulating bifacial perovskite/silicon tandem solar cells in large PV fields. In *OSA Advanced Photonics Congress, OSA Technical Digest*, page PvTh1G.3. https://doi.org/10.1364/PVLED.2020.PvTh1G.3

Markus Kantner, Theresa Höhne, Thomas Koprucki, Sven Burger, Hans-Jürgen Wünsche, Frank Schmidt, Alexander Mielke, Uwe Bandelow (2020). Multi-dimensional modeling and simulation of semiconductor nanophotonic devices. In Michael Kneissl, Andreas Knorr, Stephan Reitzenstein, Axel Hoffmann, editors, *Semiconductor Nanophotonics*, Volume 194 of *Springer Series in Solid-State Sciences*, pages 241-283, Springer. https://doi.org/10.1007/978-3-030-35656-9_7

Hans Kersting, T. J. Sullivan, Philipp Hennig (2020). Convergence rates of Gaussian ODE filters. *Statistics and Computing*, 30:1791-1816. https://doi.org/10.1007/s11222-020-09972-4

Larousse Khosravi Khorashad, Lucas V. Besteiro, Miguel A. Correa-Duarte, Sven Burger, Zhiming M. Wang, Alexander O. Govorov (2020). Hot electrons generated in chiral plasmonic nanocrystals as a mechanism for surface photochemistry and chiral growth. *J. Am. Chem. Soc.*, 142:4193. https://doi.org/10.1021/jacs.9b11124

Ilja Klebanov, Ingmar Schuster, T. J. Sullivan (2020). A rigorous theory of conditional mean embeddings. *SIAM Journal on Mathematics of Data Science*, 2(3):583-606. https://doi.org/10.1137/19M1305069

Ilja Klebanov, Alexander Sikorski, Christof Schütte, Susanna Röblitz (2020). Objective priors in the empirical Bayes framework. *Scandinavian Journal of Statistics*. https://doi.org/10.1111/sjos.12485

Stefan Klus, Feliks Nüske, Sebastian Peitz, Jan-Hendrik Niemann, Cecilia Clementi, Christof Schütte (2020). Data-driven approximation of the Koopman generator: Model reduction, system identification, and control. *Physica D*, 406. https://doi.org/10.1016/j.physd.2020.132416 (epub ahead of print 2020-03-06)

Thorsten Koch, Ying Chen, Kian Guan Lim, Xiaofei Xu, Nazgul Zakiyeva (2020). A review study of functional autoregressive models with application to energy forecasting. *WIREs Computational Statistics*. https://doi.org/10.1002/wics.1525

Tobias Kramer, Mirta Rodríguez (2020). Effect of disorder and polarization sequences on two-dimensional spectra of light harvesting complexes. *Photosynthesis Research*, 144:147-154. https://doi.org/10.1007/s11120-019-00699-6

Tobias Kramer (2020). Transient capture of electrons in magnetic fields, or: comets in the restricted three-body problem. In *Journal of Physics: Conference Series*, 1612:012019. https://doi.org/10.1088/1742-6596/1612/1/012019

Daniel J. Laydon, Vikram Sunkara, Lies Boelen, Charles R. M. Bangham, Becca Asquith (2020). The relative contributions of infectious and mitotic spread to HTLV-1 persistence. *PLOS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1007470

Mathias Lemke, Lewin Stein (2020). Adjoint-Based Identification of Sound Sources for Sound Reinforcement and Source Localization. In *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, Volume 145 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, pages 263-278, Springer, Cham. https://doi.org/https://doi.org/10.1007/978-3-030-52429-6_17

Han Cheng Lie, T. J. Sullivan, Aretha Teckentrup (2020). Error bounds for some approximate posterior measures in Bayesian inference. Numerical Mathematics and Advanced Applications ENUMATH 2019. https://arxiv.org/abs/1911.05669 (accepted for publication 2020-06-15)

Niels Lindner, Christian Liebchen (2020). Determining all integer vertices of the PESP polytope by flipping arcs. In Dennis Huisman, Christos D. Zaroliagis, editors, *20th Symposium on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS 2020)*, Volume 85 of *OpenAccess Series in Informatics (OASIcs)*, pages 5:1-5:18. https://doi.org/10.4230/OASIcs.ATMOS.2020.5

Niels Lindner (2020). Hypersurfaces with defect. *Journal of Algebra*, 555:1-35. https://doi.org/10.1016/j.jalgebra.2020.02.022

Matthias Läuter, Tobias Kramer, Martin Rubin, Kathrin Altwegg (2020). The gas production of 14 species from comet 67P/Churyumov-Gerasimenko based on DFMS/COPS data from 2014-2016. *Monthly Notices of the Royal Astronomical Society*, 498(3):3995-4004. https://doi.org/10.1093/mnras/staa2643

Fabian Löbel, Niels Lindner, Ralf Borndörfer (2020). The Restricted Modulo Network Simplex Method for Integrated Periodic Timetabling and Passenger Routing. In Janis S. Neufeld, Udo Buscher, Rainer Lasch, Dominik Möst, Jörn Schönberger, editors, *Operations Research Proceedings 2019*, 757-763. https://doi.org/https://doi.org/10.1007/978-3-030-48439-2_92

Heinz-Eberhard Mahnke, Tobias Arlt, Daniel Baum, Hans-Christian Hege, Felix Herter, Norbert Lindow, Ingo Manke, Tzulia Siopi, Eve Menei, Marc Etienne, Verena Lepper (2020). Virtual unfolding of folded papyri. *Journal of Cultural Heritage*, 41:264-269. https://doi.org/10.1016/j.culher.2019.07.007

Phillip Manley, Sebastian Walde, Sylvia Hagedorn, Martin Hammerschmidt, Sven Burger, Christiane Becker (2020). Nanopatterned Sapphire Substrates in Deep-UV LEDs: Is there an Optical Benefit?. *Opt. Express*, 28:3619. https://doi.org/10.1364/OE.379438

Kateryna Melnyk, Grègoire Montavon, Stefan Klus, Tim Conrad (2020). Graph Kernel Koopman Embedding for Human Microbiome Analysis. *Applied Network Science*, 5(96). https://doi.org/10.1007/s41109-020-00339-2

Alexander Mielke, Alberto Montefusco, Mark A. Peletier (2021). Exploring families of energy-dissipation landscapes via tilting: three types of EDP convergence. *Continuum Mechanics and Thermodynamics*. https://doi.org/10.1007/s00161-020-00932-x (accepted for publication 2020-10-01)

Gunther Mohr, Simon J. Altenburg, Alexander Ulbricht, Philipp Heinrich, Daniel Baum, Christiane Maierhofer, Kai Hilgenberg (2020). In-situ defect detection in laser powder bed fusion by using thermography and optical tomography – comparison to computed tomography. *Metals*, 10(1):103. https://doi.org/10.3390/met10010103

Mattes Mollenhauer, Ingmar Schuster, Stefan Klus, Christof Schütte (2020). Singular Value Decomposition of Operators on Reproducing Kernel Hilbert Spaces. In Oliver Junge, O. Schütze, Gary Froyland, S. Ober-Blobaum, K. Padberg-Gehle, editors, *Advances om Dynamics, Optimization and Computation. Series: Studies in Systems, Decision and Control.* 304:109-131.

Alberto Montefusco, Mark A. Peletier, Hans Christian Öttinger (2021). A Framework of Nonequilibrium Statistical Mechanics. II. Coarse-Graining. *Journal of Non-Equilibrium Thermodynamics*, 46(1):15-33. https://doi.org/10.1515/jnet-2020-0069 (accepted for publication 2020-09-14)

Hassan Mortagy, Swati Gupta, Sebastian Pokutta (2020). Walking in the Shadow: A New Perspective on Descent Directions for Constrained Minimization. In *Proceedings of NeurIPS*.

Paweł Mrowiński, Peter Schnauber, Arsenty Kaganskiy, Johannes Schall, Sven Burger, Sven Rodt, Stephan Reitzenstein (2020). Directional single-photon emission from deterministic quantum dot waveguide structures. *Phys. Status Solidi RRL*, 14:2000115. https://doi.org/10.1002/pssr.202000115

Anna Musiał, Kinga Żołnacz, Nicole Srocka, Oleh Kravets, Jan Große, Jacek Olszewski, Krzysztof Poturaj, Grzegorz Wojcik, Paweł Mergo, Kamil Dybka, Mariusz Dyrkacz, Michal Dlubek, Kristian Lauritsen, Andreas Bülter, Philipp-Immanuel Schneider, Lin Zschiedrich, Sven Burger, Sven Rodt, Wacław Urbanczyk, Grzegorz Sek, Stephan Reitzenstein (2020). Front Cover: Plug&Play Fiber-Coupled 73 kHz Single-Photon Source Operating in the Telecom O-Band (Adv. Quantum Technol. 6/2020). *Adv. Quantum Technol.*, 3:2070061. https://doi.org/10.1002/qute.202070061

Anna Musiał, Kinga Żołnacz, Nicole Srocka, Oleh Kravets, Jan Große, Jacek Olszewski, Krzysztof Poturaj, Grzegorz Wojcik, Paweł Mergo, Dybka Kamil, Mariusz Dyrkacz, Michal Dlubek, Kristian Lauritsen, Andreas Bülter, Philipp-Immanuel Schneider, Lin Zschiedrich, Sven Burger, Sven Rodt, Wacław Urbanczyk, Grzegorz Sek, Stephan Reitzenstein (2020). Plug&play fibre-coupled 73 kHz single-photon source operating in the telecom O-band. *Adv. Quantum Technol.*, 3:2000018. https://doi.org/10.1002/qute.202000018

Jan Möller, Ali Isbilir, Titiwat Sungkaworn, Brenda Osberg, Christos Karathanasis, Vikram Sunkara, Eugene O Grushevsky, Andreas Bock, Paolo Annibale, Mike Heilemann, Christof Schütte, Martin J. Lohse (2020). Single molecule mu-opioid receptor membrane-dynamics reveal agonist-specific dimer formation with super-resolved precision. *Nature Chemical Biology*, 16:946-954. https://doi.org/10.1038/s41589-020-0566-1

Benjamin Müller, Felipe Serrano, Ambros Gleixner (2020). Using two-dimensional Projections for Stronger Separation and Propagation of Bilinear Terms. *SIAM Journal on Optimization*, 30(2):1339-1365. https://doi.org/https://doi.org/10.1137/19M1249825

Benjamin Müller, Gonzalo Muñoz, Maxime Gasse, Ambros Gleixner, Andrea Lodi, Felipe Serrano (2020). On Generalized Surrogate Duality in Mixed-Integer Nonlinear Programming. In *Integer Programming and Combinatorial Optimization: 21th International Conference, IPCO 2020*, 322-337. https://doi.org/10.1007/978-3-030-45771-6_25

Esfandiar Nava-Yazdani, Hans-Christian Hege, T. J. Sullivan, Christoph von Tycowicz (2020). Geodesic Analysis in Kendall's Shape Space with Epidemiological Applications. *Journal of Mathematical Imaging and Vision*, 62(4):549-559. https://doi.org/10.1007/s10851-020-00945-w

Chris Oates, Jon Cockayne, Dennis Prangle, T. J. Sullivan, Mark Girolami (2020). Optimality criteria for probabilistic numerical methods. In F. J. Hickernell, P. Kritzer, editors, *Multivariate Algorithms and Information-Based Complexity*, 27:65-88. https://doi.org/10.1515/9783110635461-005

Mohamed Omari, Alexander Lange, Julia Plöntzke, Susanna Röblitz (2020). Model-based exploration of the impact of glucose metabolism on the estrous cycle dynamics in dairy cows. *Biology Direct*, 15. https://doi.org/10.1186/s13062-019-0256-7

Anton Pakhomov, Franz Löchner, Lin Zschiedrich, Sina Saravi, Martin Hammerschmidt, Sven Burger, Thomas Pertsch, Frank Setzpfand (2020). Far-field polarization signatures of surface optical nonlinearity in noncentrosymmetric semiconductors. *Sci. Rep.*, 10:10545. https://doi.org/10.1038/s41598-020-67186-0

Marc Pfetsch, Sebastian Pokutta (2020). IPBoost – Non-Convex Boosting via Integer Programming. In *Proceedings of ICML*.

Mika Pflüger, R. Joseph Kline, Analia Fernández Herrero, Martin Hammerschmidt, Victor Soltwisch, Michael Krumrey (2020). Extracting dimensional parameters of gratings produced with self-aligned multiple patterning using grazing-incidence small-angle x-ray scattering. *J. Micro Nanolithogr. MEMS MOEMS*, 19:014001. https://doi.org/10.1117/1.JMM.19.1.014001

Pedro Pimentel, Angelika Szengel, Moritz Ehlke, Hans Lamecker, Stefan Zachow, Laura Estacio, Christian Doenitz, Heiko Ramm (2020). Automated Virtual Reconstruction of Large Skull Defects using Statistical Shape Models and Generative Adversarial Networks. Jianning Li, Jan Egger, editors, *Towards the Automatization of Cranial Implant Design in Cranioplasty*, 12439. (accepted for publication 2020-09-21)

Sebastian Pokutta, M. Singh, A. Torrico (2020). On the Unreasonable Effectiveness of the Greedy Algorithm: Greedy Adapts to Sharpness. In *Proceedings of ICML*.

Sebastian Pokutta (2020). Restarting Algorithms: Sometimes there is Free Lunch. In *Proceedings of CPAIOR*.

Luc Pronzato, Guillaume Sagnol (2021). Removing inessential points in c- and A-optimal design. *Journal of Statistical Planning and Inference*, 213:233-252. https://doi.org/10.1016/j.jspi.2020.11.011 (epub ahead of print 2020-12-08)

Robert Julian Rabben, Sourav Ray, Marcus Weber (2020). ISOKANN: Invariant subspaces of Koopman operators learned by a neural network. *The Journal of Chemical Physics*, 153(11):114109. https://doi.org/10.1063/5.0015132

Elham Ramin, Ksenia Bestuzheva, Carina Gargalo, Danial Ramin, Carina Schneider, Pedram Ramin, Xavier Flores-Alsina, Maj M. Andersen, Krist V. Gernaey (2021). Incremental design of water symbiosis networks with prior knowledge: The case of an industrial park in Kenya. *Science of the Total Environment*, 751. https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.141706 (epub ahead of print 2020-08-26)

Mona Rams, Tim Conrad (2020). Dictionary Learning for transcriptomics data reveals type-specific gene modules in a multi-class setting. *it – Information Technology*, 62(3-4). https://doi.org/https://doi.org/10.1515/itit-2019-0048

Sourav Ray, Vikram Sunkara, Christof Schütte, Marcus Weber (2020). How to calculate pH-dependent binding rates for receptor-ligand systems based on thermodynamic simulations with different binding motifs. *Molecular Simulation*, 46(18):1443-1452. https://doi.org/10.1080/08927022.2020.1839660

Sven Rodt, Philipp-Immanuel Schneider, Lin Zschiedrich, Tobias Heindel, Samir Bounouar, Markus Kantner, Thomas Koprucki, Uwe Bandelow, Sven Burger, Stephan Reitzenstein (2020). Deterministic Quantum Devices for Optical Quantum Communication. In Michael Kneissl, Andreas Knorr, Stephan Reitzenstein, Axel Hoffmann, editors, *Semiconductor Nanophotonics*, Volume 194 of *Springer Series in Solid-State Sciences*, pages 285-359, Springer. https://doi.org/10.1007/978-3-030-35656-9_8

Susanne Röhl, Marcus Weber, Konstantin Fackeldey (2020). Computing the minimal rebinding effect for non-reversible processes. *Multiscale Modeling and Simulation*. (accepted for publication 2020-11-30)

Ansgar Rössig, Milena Petkovic (2020). Advances in Verification of ReLU Neural Networks. *Journal of Global Optimization*. https://doi.org/10.1007/s10898-020-00949-1

Manish Sahu, Ronja Strömsdörfer, Anirban Mukhopadhyay, Stefan Zachow (2020). Endo-Sim2Real: Consistency learning-based domain adaptation for instrument segmentation. In *Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI), Part III*, Volume 12263 of *Lecture Notes in Computer Science*. https://doi.org/https://doi.org/10.1007/978-3-030-59716-0_75

Manish Sahu, Angelika Szengel, Anirban Mukhopadhyay, Stefan Zachow (2020). Surgical phase recognition by learning phase transitions. *Current Directions in Biomedical Engineering (CDBME)*, 6(1). https://doi.org/https://doi.org/10.1515/cdbme-2020-0037

Farouk Salem, Florian Schintke, Thorsten Schütt, Alexander Reinefeld (2020). Scheduling data streams for low latency and high throughput on a Cray XC40 using Libfabric. *Concurrency and Computation Practice and Experience*, 32(20):1-14. https://doi.org/10.1002/cpe.5563

Xavier Garcia Santiago, Sven Burger, Carsten Rockstuhl, Philipp-Immanuel Schneider (2021). Bayesian optimization with improved scalability and derivative information for efficient design of nanophotonic structures. *J. Light. Technol.*, 39:167. https://doi.org/10.1109/JLT.2020.3023450 (epub ahead of print 2020-09-11)

Gerhard Scholtz, David Knötel, Daniel Baum (2020). D'Arcy W. Thompson's Cartesian transformations: a critical evaluation. *Zoomorphology*, 139:293-308. https://doi.org/10.1007/s00435-020-00494-1

Merlind Schotte, Júlia Chaumel, Mason N. Dean, Daniel Baum (2020). Image analysis pipeline for segmentation of a biological porosity network, the lacuno-canalicular system in stingray tesserae. *MethodsX*, 7:100905. https://doi.org/10.1016/j.mex.2020.100905 (epub ahead of print 2020-05-01)

Philip Scott, Xavier Garcia Santiago, Dominik Beutel, Carsten Rockstuhl, Martin Wegener, Ivan Fernandez-Corbaton (2020). On enhanced sensing of chiral molecules in optical cavities. *Appl. Phys. Rev.*, 7:041413. https://doi.org/10.1063/5.0025006

Felipe Serrano, Robert Schwarz, Ambros Gleixner (2020). On the relation between the extended supporting hyperplane algorithm and Kelley's cutting plane algorithm. *Journal of Global Optimization*, 78:161-179. https://doi.org/10.1007/s10898-020-00906-y

Felipe Serrano, Gonzalo Muñoz (2020). Maximal Quadratic-Free Sets. In *Integer Programming and Combinatorial Optimization: 21th International Conference, IPCO 2020*, 307-321. https://doi.org/10.1007/978-3-030-45771-6_24

Liping Shi, Andrey B. Evlyukhin, Carsten Reinhardt, Ihar Babushkin, Vladimir A. Zenin, Sven Burger, Radu Malureanu, Boris N. Chichkov, Uwe Morgner, Milutin Kovacev (2020). Progressive Self-Boosting Anapole-Enhanced Deep-Ultraviolet Third Harmonic During Few-Cycle Laser Radiation. *ACS Photonics*, 7:1655. https://doi.org/10.1021/acsphotonics.0c00753

Yuji Shinano, N. Tateiwa, S. Nakamura, A. Yoshida, M. Yasuda, S. Kaji, K. Fujisawa (2020). Massive Parallelization for Finding Shortest Lattice Vectors Based on Ubiquity Generator Framework. In *2020 SC20: International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, 834-848. https://doi.org/10.1109/SC41405.2020.00064

Thomas Siefke, Carol Hurtado, Johannes Dickmann, Walter Dickmann, Tim Käseberg, Jan Meyer, Sven Burger, Uwe Zeitner, Bernd Bodermann, Stefanie Kroker (2020). Quasi-bound states in the continuum for deep subwavelength structural information retrieval for DUV nano-optical polarizers. *Opt. Express*, 28:23122. https://doi.org/10.1364/OE.396044

Jan Skrzypczak, Florian Schintke, Thorsten Schütt (2020). RMWPaxos: Fault-Tolerant In-Place Consensus Sequences. *IEEE Transactions on Parallel and Distributed Systems*, 31(10):2392-2405. https://doi.org/10.1109/TPDS.2020.2981891

Jan Skrzypczak, Florian Schintke (2020). Towards Log-Less, Fine-Granular State Machine Replication. *Datenbank Spektrum*, 20(3):231-241. https://doi.org/10.1007/s13222-020-00358-4

Ana Prates Soares, Daniel Baum, Bernhard Hesse, Andreas Kupsch, Bernd Müller, Paul Zaslansky (2020). Scattering and phase-contrast X-ray methods reveal damage to glass fibers in endodontic posts following dental bur trimming. *Dental Materials*. https://doi.org/10.1016/j.dental.2020.10.018 (epub ahead of print 2020-12-11)

Boro Sofranac, Ambros Gleixner, Sebastian Pokutta (2020). Accelerating Domain Propagation: an Efficient GPU-Parallel Algorithm over Sparse Matrices. In *Proceedings of IA^3 at SC20*.

Lewin Stein, Florian Straube, Stefan Weinzierl, Mathias Lemke (2020). Directional sound source modeling using the adjoint Euler equations in a finite-difference time-domain approach. *Acoustical Society of America*, 148(5):3075-3085. https://doi.org/https://doi.org/10.1121/10.0002425

*Ralph L. Stoop, Arthur Straube, Tom H. Johansen, Pietro Tierno (2020). Collective directional locking of colloidal monolayers on a periodic substrate. *Phys. Rev. Lett.*, 124:058002. https://doi.org/10.1103/PhysRevLett.124.058002 (Joint publication: Modeling and Simulation of Complex Processes, Distributed Algorithms and Supercomputing)

*Arthur Straube, Bartosz G. Kowalik, Roland R. Netz, Felix Höfling (2020). Rapid onset of molecular friction in liquids bridging between the atomistic and hydrodynamic pictures. *Commun. Phys.*, 3:126. https://doi.org/10.1038/s42005-020-0389-0 (Joint publication: Modeling and Simulation of Complex Processes, Distributed Algorithms and Supercomputing)

Tom Streubel, Caren Tischendorf, Andreas Griewank (2020). Piecewise Polynomial Taylor Expansions – The Generalization of Faà di Bruno's Formula. *Modeling, Simulation and Optimization of Complex Processes HPSC 2018*, 63-82. https://doi.org/10.1007/978-3-030-55240-4_3

Yahui Sun, Daniel Rehfeldt, Marcus Brazil, Doreen Thomas, Saman Halgamuge (2020). A Physarum-Inspired Algorithm for Minimum-Cost Relay Node Placement in Wireless Sensor Networks. *IEEE/ACM Transactions on Networking*. https://doi.org/10.1109/TNET.2020.2971770

Alexander Tesch (2020). A Polyhedral Study of Event-Based Models for the Resource-Constrained Project Scheduling Problem. *Journal of Scheduling*.

Peter Tillmann, Klaus Jäger, Christiane Becker (2020). Minimising levelised cost of electricity of bifacial solar panel arrays using Bayesian optimisation. *Sustain. Energy Fuels*, 4:254. https://doi.org/10.1039/C9SE00750D

Philipp Tockhorn, Johannes Sutter, Rémi Colom, Lukas Kegelmann, Amran Al-Ashouri, Marcel Roß, Klaus Jäger, Thomas Unold, Sven Burger, Steve Albrecht, Christiane Becker (2020). Improved Quantum Efficiency by Advanced Light Management in Nanotextured Solution-Processed Perovskite Solar Cells. *ACS Photonics*, 7:2589. https://doi.org/10.1021/acsphotonics.0c00935

Christoph von Tycowicz (2020). Towards Shape-based Knee Osteoarthritis Classification using Graph Convolutional Networks. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI 2020)*. https://doi.org/10.1109/ISBI45749.2020.9098687

Roya Ebrahimi Viand, Felix Höfling, Rupert Klein, Luigi Delle Site (2020). Theory and simulation of open systems out of equilibrium. *The Journal of Chemical Physics*, 153:101102. https://doi.org/10.1063/5.0014065

José Villatoro, Martin Zühlke, Daniel Riebe, Toralf Beitz, Marcus Weber, Hans-Gerd Löhmannsröben (2020). Sub-ambient pressure IR-MALDI ion mobility spectrometer for the determination of low and high field mobilities. *Analytical and Bioanalytical Chemistry*, 412:5247-5260. https://doi.org/10.1007/s00216-020-02735-0

Marcus Weber, Weitere Autoren (2020). DIN SPEC 2343: Übertragung von sprachbasierten Daten zwischen Künstlichen Intelligenzen – Festlegung von Parametern und Formaten. Christine Reichardt, editor, *Beuth Verlag*.

Marie-Christin Weber, Lisa Fischer, Alexandra Damerau, Igor Ponomarev, Moritz Pfeiffenberger, Timo Gaber, Sebastian Götschel, Jens Lang, Susanna Röblitz, Frank Buttgereit, Rainald Ehrig, Annemarie Lang (2020). Macroscale mesenchymal condensation to study cytokine-driven cellular and matrix-related changes during cartilage degradation. *Biofabrication*, 12(4). https://doi.org/10.1088/1758-5090/aba08f

Selina Weiss, Carl Martin Grewe, Sally Olderbak, Benjamin Goecke, Laura Kaltwasser, Andrea Hildebrandt (2020). Symmetric or not? A holistic approach to the measurement of fluctuating asymmetry from facial photographs. *Personality and Individual Differences*, 166:1-12. https://doi.org/https://doi.org/10.1016/j.paid.2020.110137

Daniel Werdehausen, Sven Burger, Isabelle Staude, Thomas Pertsch, Manuel Decker (2020). General design formalism for highly efficient flat optics for broadband applications. *Opt. Express*, 28:6452. https://doi.org/10.1364/OE.386573

Daniel Werdehausen, Xavier Garcia Santiago, Sven Burger, Isabelle Staude, Thomas Pertsch, Carsten Rockstuhl, Manuel Decker (2020). Modeling Optical Materials at the Single Scatterer Level: The Transition from Homogeneous to Heterogeneous Materials. *Adv. Theory Simul.*, 3:2000192. https://doi.org/10.1002/adts.202000192

Daniel Werdehausen, Sven Burger, Isabelle Staude, Thomas Pertsch, Manuel Decker (2020). Flat optics in high numerical aperture broadband imaging systems. *J. Opt.*, 22:065607. https://doi.org/10.1088/2040-8986/ab8ea2

Maraike Willsch, Frank Friedrich, Daniel Baum, Ivo Jurisch, Michael Ohl (2020). A comparative description of the mesosomal musculature in Sphecidae and Ampulicidae (Hymenoptera, Apoidea) using 3D techniques. *Deutsche Entomologische Zeitschrift*, 67(1):51-67. https://doi.org/10.3897/dez.67.49493

Jakob Witzig, Timo Berthold (2020). Conflict-Free Learning for Mixed Integer Programming. In *Integration of AI and OR Techniques in Constraint Programming. CPAIOR 2020*, *LNCS*, pages 521-530. https://doi.org/10.1007/978-3-030-58942-4_34

Jakob Witzig, Timo Berthold (2020). Conflict Analysis for MINLP. *INFORMS Journal on Computing*.

Jakob Witzig, Ambros Gleixner (2020). Conflict-Driven Heuristics for Mixed Integer Programming. *INFORMS Journal on Computing*. https://doi.org/10.1287/ijoc.2020.0973 (epub ahead of print 2020-10-08)

Niklas Wulkow, Péter Koltai, Christof Schütte (2020). Memory-based reduced modeling and data-based estimation of opinion spreading. *Journal of Nonlinear Science*. (accepted for publication 2020-12-12)

*Hanna Wulkow, Tim Conrad, Natasa Djurdjevac Conrad, Sebastian A. Müller, Kai Nagel, Christof Schütte (2020). Prediction of COVID-19 spreading and optimal coordination of counter-measures: From microscopic to macroscopic models to Pareto fronts. *PLOS One*. https://doi.org/https://doi.org/10.1101/2020.12.01.20241885 (Joint publication: Modeling and Simulation of Complex Processes, Visual and Data-centric Computing)

Christian Würth, Phillip Manley, Robert Voigt, Doguscan Ahiboz, Christiane Becker, Ute Resch-Genger (2020). Metasurface enhanced sensitized photon upconversion: towards highly efficient low power upconversion applications and nano-scale E-field sensors. *Nano Lett.*, 20:6682. https://doi.org/10.1021/acs.nanolett.0c02548

Nazgul Zakiyeva, X. Xu (2020). Nonlinear network autoregressive model with application to natural gas network forecasting. *Mathematics Japonica*. (accepted for publication 2020-06-29)

Ralf F. Ziesche, Tobias Arlt, Donal P. Finegan, Thomas M.M. Heenan, Alessandro Tengattini, Daniel Baum, Nikolay Kardjilov, Henning Markötter, Ingo Manke, Winfried Kockelmann, Dan J.L. Brett, Paul R. Shearing (2020). 4D imaging of lithium-batteries using correlative neutron and X-ray tomography with a virtual unrolling technique. *Nature Communications*, 11:777. https://doi.org/10.1038/s41467-019-13943-3

Hans Christian Öttinger, Alberto Montefusco, Mark A. Peletier (2021). A Framework of Nonequilibrium Statistical Mechanics. I. Role and Types of Fluctuations. *Journal of Non-Equilibrium Thermodynamics*, 46(1):1-13. https://doi.org/10.1515/jnet-2020-0068 (accepted for publication 2020-08-13)

BOOKS (4)

Peter Deuflhard, Martin Weiser (2020). Numerische Mathematik 3. Adaptive Lösung partieller Differentialgleichungen. de Gruyter.

Natalia Selini Hadjidimitriou, Antonio Frangioni, Andrea Lodi, Thorsten Koch (2020). Mathematical Optimization for Efficient and Robust Energy Networks. Springer.

Bernhard Reuter (2020). Generalisierte Markov-Modellierung. Springer Spektrum, Wiesbaden.

Stefanie Winkelmann, Christof Schütte (2020). Stochastic Dynamics in Computational Biology. Springer International Publishing.

DISSERTATIONS (5)

Moritz Ehlke (2020). 3D Reconstruction of Anatomical Structures from 2D X-ray Images. Technische Universität Berlin.

Neveen Ali Salem Eshtewy (2020). Mathematical Modeling of Metabolic-Genetic Networks. Freie Universität Berlin.

Nico Kruber (2020). Approximate Distributed Set Reconciliation with Defined Accuracy. Humboldt-Universität zu Berlin.

Xavier Garcia Santiago (2020). Numerical methods for shape optimization of photonic nanostructures. Karlsruher Institut für Technologie.

Narendra Lagumaddepalli Venkatareddy (2020). Revealing secrets of mussel-glue mimetic peptides – From advanced NMR to computational process modeling. Humboldt-Universität zu Berlin.

IMPRINT

Imprint

Published by
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustrasse 7 | 14195 Berlin | Germany

Phone +49-30-841 85-0
Fax +49-30-841 85-125
E-mail info@zib.de
www.zib.de

Layout
Christiane Weimann Grafik + Produktion
www.schnittmarke.biz

Printing
Lochmann grafische Produktion GmbH

ISSN
0934-5892