

# ANNUAL REPORT 2021



# CONTENTS

<b>1</b>	Annual Report 2021
<b>6</b>	Executive Summary
<b>8</b>	Organization
<b>10</b>	ZIB Structure
<b>12</b>	Quantum Computing: the Next Big Thing
<b>14</b>	The Transformation of HPC
<b>16</b>	Economic Situation in 2021
<b>23</b>	Spreading Dynamics
<b>32</b>	Inference and Simulation of Molecular and Cellular Processes
<b>42</b>	Multienergy Models for European Energy System Planning
<b>52</b>	Multiobjective Shortest Paths
<b>62</b>	Explainability and Interpretability in Deep Learning
<b>70</b>	Explosion in Space
<b>76</b>	Mastering Multidimensional Heterogeneity of Computers
<b>86</b>	References
<b>89</b>	Notes
<b>90</b>	Publications
<b>100</b>	Imprint





# EXECUTIVE SUMMARY

Following the onset of the COVID-10 pandemic in 2020, the year 2021 has again been an exceptional year in the history of ZIB. Over prolonged periods of time, the operation of ZIB was heavily burdened with the consequences of the crisis. However, ZIB's members and operational units adapted to the pandemic with a high level of motivation and commitment. It is due to the efforts of many that the transition to working from home and working from the office when required went very smoothly, and a significant number of excellent results were achieved also under these difficult circumstances.

The mission of the Zuse Institute Berlin (ZIB) is to advance Computational Thinking by leveraging data-driven Computational Science, application-oriented Mathematics, and High-Performance Computing. We develop computational methods for solving problems of relevance to society with partners from science, technology, and industry and provide high-end, state-of-the-art computing and data management infrastructures. 2021 was a very successful year for ZIB with many scientific highlights and very positive developments that make us confident that our institute has a bright future, and continues to be a place for excellent research and first-rate scientific services and infrastructure.

ZIB continues to be among the five cooperation partners of the Cluster of Excellence MATH+. In 2021, ZIB researchers were involved in almost 50% of the research projects in MATH+. The results obtained led to many publications in top-tier conferences and journals and achieved several prestigious awards. Simultaneous participation in four DFG Collaborative Research Centers, in two DFG Research Groups and in some highly visible BMBF funding formats, such as the Research Campus MODAL or flagship projects on COVID research, are proof of the high quality of research at ZIB. Furthermore, as one of the centers of the National High-Performance Computing (NHR) initiative and as a cornerstones of the mathematical network MaRDI within the National Research Data Infrastructure (NFDI) framework, ZIB is also at the forefront of research support services which facilitate science in Berlin and across Germany.

In the beginning of 2021, the National High-Performance Computing (NHR) alliance started with eight centers, including the NHR center at ZIB. In its first year, the alliance strengthened and interconnected the already established research and service activities at these eight centers, and established the NHR Association (NHR Verein e.V), a non-profit legal entity with the aim to coordinate between the centers. The board of the NHR Verein, presently chaired by ZIB's president Christof Schuette, will coordinate the activities of the NHR centers regarding research, infrastructure development, and advancement of the next generation of HPC researchers and experts. Within the NHR alliance, ZIB has the lead for life science applications, AI and machine learning (ML) methods, and heterogeneous architectures. Please find more information in the article "The Transformation of HPC".

In 2021, the Einstein Research Unit (ERU) "Perspectives of a quantum digital transformation: Near-term quantum computational devices and quantum processors" was created with funding through the Einstein Foundation and the Berlin University Alliance. This unit sets out to understand the precise computational power of near-term quantum devices and to devise building blocks for quantum processors. ZIB's Vice-President, Sebastian Pokutta, is among the founding PIs, and ZIB provides its key expertise in applied mathematics, optimization, machine learning, and quantum optics to the ERU. In the course of this development, ZIB has decided to bundle its own competencies in a new working

group that focuses on researching quantum computing algorithms as well as their verification and simulation on classical hardware, in particular also, in cooperation with the NHR center at ZIB, on supplementing quantum computing research with the necessary HPC resources. For more information see the article “Quantum Computing: the next big thing”.

This annual report gives a general overview of ZIB’s organization, research statistics, as well as key factors for its successful development. Most prominently, it provides insights into the broad variety of research activities and especially into some of its scientific highlights, which are presented in seven “Feature Articles”:

1. “**Inference and Simulation of Molecular and Cellular Processes**” illustrates that our research foci on computational drug design, multiscale modelling of cellular processes and learning cellular structures and processes from data have begun to merge into one larger research endeavor providing fascinating insights and completely new opportunities.
2. In “**Explainability & interpretability in deep learning**” insight is provided on explainable AI (XAI) aiming to make modern machine learning models more transparent, and how ZIB contributes by novel algorithms that tame the complexity of XAI by a Frank-Wolfe approach which, by design, constructs sparse iterates that accelerates many elementary operations and facilitates simply explanations due to the low-complexity structure of iterates.
3. “**Spreading Dynamics: Cultural Change, COVID-19, and Mobility**” reports on our research activities tailored towards understanding complex social processes with utmost relevance for society and in particular on how mathematical research opens up connections between seemingly completely different fields of application.
4. “**Multienergy Models for European Energy System Planning**” explains the role of mathematical Optimization in supporting the European Energy Transition with an integrated holistic energy system planning framework by developing an end-to-end planning tool to successfully increase the share of renewable

energy in the European energy system without compromising system reliability.

5. “**Multi-objective Shortest Paths**” reports on recent progress towards efficient solutions of multi-objective shortest path problems like the novel Multi-objective Dijkstra Algorithm that has been developed within the Research Campus MODAL at ZIB in cooperation with Lufthansa Systems in our joint effort to build a new generation of core-optimizers for flight planning.
6. “**Mastering Multidimensional Heterogeneity of Computers**” starts from the insight that heterogeneity has become the key element for building advanced and energy-efficient computing systems, and outlines how research at ZIB combines diversity in the processor landscape with technological innovations for byte-addressable storage as well as designing smarter networks.
7. Last but not least, “**Explosion in Space**” reports on the development of heuristic search methods as a powerful tool to tackle for search problems in galaxy-scale graphs and on how to use them to solve application-specific problems such as VLSI floor planning, multi agent planning, scheduling, and efficient code generation in compilers.



Prof. Dr. Christof Schütte  
President  
schuette@zib.de



Prof. Dr. Sebastian Pokutta  
Vice President  
pokutta@zib.de

# ORGANIZATION

## The Statutes

The new ZIB Statutes, adopted by the Board of Directors at its meeting on April 28, 2021, redefines ZIB's bodies and organizational structure based on the new ZIB law. Furthermore, it determines 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. Sabine Kunst**

President, Humboldt-Universität zu Berlin

**Prof. Dr. Christian Thomsen**

President, Technische Universität Berlin (Chairman)

**Prof. Dr. Günter Ziegler**

President, Freie Universität Berlin

**Christian Hingst**

Der Regierende Bürgermeister von Berlin, Senatskanzlei –  
Wissenschaft und Forschung

**Thomas Frederking**

Helmholtz-Zentrum Berlin für Materialien und Energie  
(HZB)

The Board of Directors met on April 28, 2021,  
and November 22, 2021.



# 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

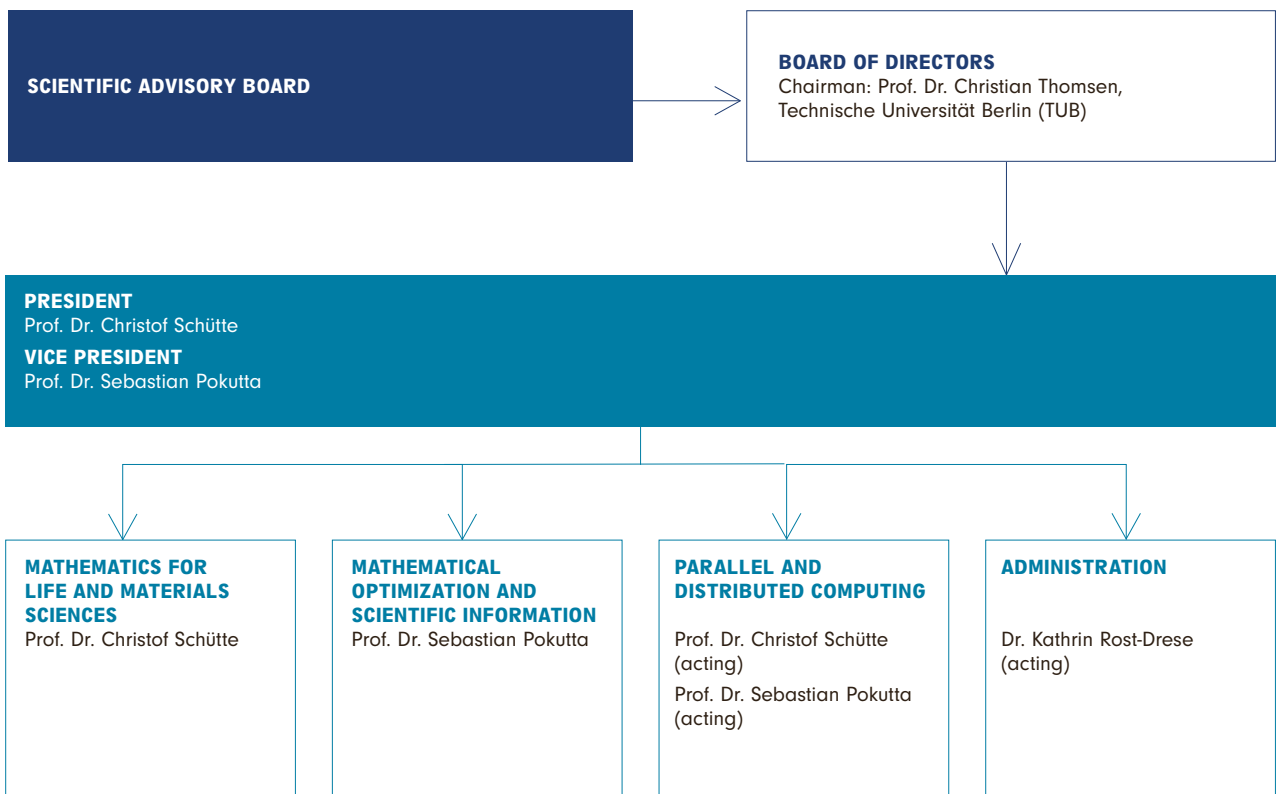
**Prof. Dr. Rolf Krause**  
Université della svizzera italiana, Lugano, Switzerland

**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 online on three days in June 2021.



# ZIB STRUCTURE

## Administrative Council

**President** (C. Schütte) | **Vice-President** (S. Pokutta) | **Managing Director** (K. Rost-Drese, act.)

## Research Departments

**Mathematics of Complex Systems**  
(C. Schütte)

**Mathematical Algorithmic**  
(S. Pokutta)

**Department  
"Modelling and  
Simulation of  
Complex Processes"**  
(M. Weiser)

**Department  
"Visual and  
Data-Centric  
Computing"**  
(T. Conrad)

**Department  
"AI in Society,  
Science, and  
Technology"**  
(S. Pokutta)

**Department  
"Applied  
Intelligence"**  
(T. Koch)

## Research Service Units

**IT and Data Services**  
(C. Schäuble)

**Digital Data and  
Information for  
Society, Science,  
and Culture**

**Administration** (K. Rost-Drese)

**Ombudsperson/Good Scientific Practice**  
(J. Sack)

**Staff Council Chairperson**  
(T. Hasler)

**Scientific Advisory Board**

**Intelligence**

**Parallel and Distributed Computing**

(C. Schütte, act. | S. Pokutta, act.)

**Algorithmic  
Methods"**

**Department  
"Network  
Optimization"**  
(R. Borndörfer)

**Department  
"Distributed  
Algorithms"**  
(F. Schintke)

**Department  
"Supercomputing"**  
(T. Steinke)

(B. Rusch, T. Koch)

**NHR Center**  
(T. Steinke, act.)

**Women's Representative**  
(M. Kostre)

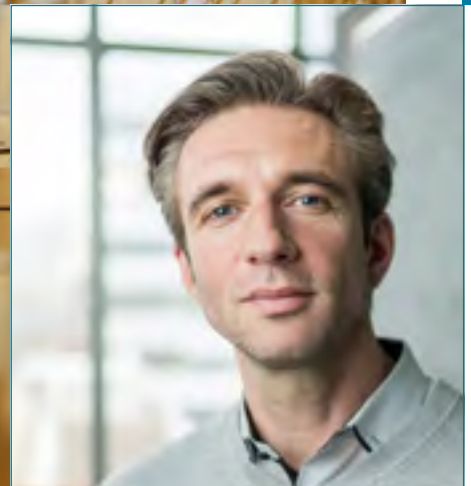
**Data Protection Officer**  
(O. Paetsch)

# QUANTUM COMPUTING: THE NEXT BIG THING

The recognition of the power of quantum physics dates back to fundamental works in physics in the 1930s when intellectual heavyweights such as Einstein, Bohr, Schrödinger, and Heisenberg opened up one Pandora's box after another. One such particular example is the Einstein-Podolsky-Rosen paradox (1935) and the later Bell inequalities (1964) where it was recognized that quantum systems can do things that go beyond what traditional physics and probability theory can explain. Initially, the EPR paradox led to significant confusion whether the observed effect was due to some form of incompleteness or whether it constituted a fundamentally new, "non-local" reality. The authenticity of these effects has since then been confirmed several times over in various experimental setups. Another example which has been also widely covered in popular media is Schrödinger's cat, a thought experiment that amplifies quantum effect to the macro scale of our world and provides an example what the quantum effects would look like in our macro world.

Harnessing these "strange effects" has since been a challenge, both methodologically "How do I design an algorithm that exploits quantum effects?" and practically "How do I build a quantum computer of sufficiently low error-rate and long durations of coherence?". Then there are very "mundane" yet no less formidable logistical and cryogenic challenges: Qubits typically operate at about 20 millikelvin which is approximately  $-273^{\circ}\text{C}$ , so "How do I keep this thing cold enough?". However, maybe similar to what has happened to machine learning about 10-15 years back, in very recent times some important advantages on all fronts were made suggesting that we might have everything that is needed to actually build the first (real) quantum computer and design algorithms for it. This has spurred an enormous research enterprise into quantum computing on all levels: hardware, software, and methodologies. Whether we will be successful in building such a quantum computer and operate it at scale both in size and numbers remains to be seen, but, and this is important to recognize, even if the answer would be no, the scientific advances that this endeavor has produced already up to today would have made this a very worthwhile effort. The next years are only going to further tilt the scale in favor.

With its mission being centered around the exploration of the interface between computing, mathematics, and computer science, the Zuse Institute Berlin does not only acknowledge the potential of quantum computing but understands quantum computing to be integrally tight to its research mission. As such, we have established a quantum computing research group, currently comprising six researchers, which focuses on researching quantum computing algorithms as well as their verification and simulation on classical hardware. These efforts are tightly integrated into our broader **Nationales Hochleistungsrechnen (NHR)** mission, where we operate a Tier-2 High-Performance Computing (HPC) system for the German research community and where, among other topics such as Artificial Intelligence and Simulation, we also focus on supplementing quantum computing research with the necessary HPC resources.



Prof. Dr. Sebastian Pokutta  
Phone +49 30 841 85-208  
pokutta@zib.de

# THE TRANSFORMATION

The German federal and state governments have agreed to jointly fund the National High-Performance Computing Alliance (NHR) with 600 million euros. NHR is a national network of tier-2 HPC centers with concerted efforts regarding the whole spectrum of HPC from infrastructure and software, to method development and applications. NHR focuses on excellent research, consulting, services, and training for the HPC community, understanding that scientific innovation with benefits to society often requires high-quality support and sufficient capacity rather than a race for the largest supercomputer.



# OF HPC

The NHR Alliance started in the beginning of 2021 with eight NHR centers, including the NHR center at ZIB. In its first year, the Alliance networked and strengthened the already established research and service activities at the eight centers, and established the NHR Association (NHR Verein e.V), a nonprofit legal entity for coordination between the centers. The NHR Verein, itself funded by the national NHR grant, will coordinate the procurement of HPC infrastructure, service, and consulting at the NHR centers, launch joint innovation projects, and consolidate the Alliance's support for young researchers into a graduate school. The board of the NHR Verein, presently chaired by ZIB's president Christof Schuette, will represent the interests of the Alliance to the funding agencies, and coordinates the activities of the NHR centers with the German tier-1 HPC centers, the Gauß-Allianz.

The NHR Alliance is broad enough to provide solutions for traditional HPC applications as well as for big data, AI, new computing paradigms, and new forms of demand from scientific communities with previously little HPC affinity. The leaps high-performance computing has made in computing power do not always correlate to improved insights, and the NHR Alliance believes that this simple fact calls for better support for and training of scientific users of high-performance infrastructure by HPC experts.

Within NHR, ZIB has the lead for life science applications, AI and machine learning (ML) methods, and heterogeneous

architectures. In regard to AI and ML, we are not only exploring new ways of infusing HPC with AI methods and practices to uncover new insights, but also to create better tools for nearly every stage of the modern HPC workflow to empower researchers with the ability to solve problems that were previously unsolvable. With the big-data boom, the immense amount of information represents tremendous opportunities for researchers, but it also provides a new set of challenges, as the boon of data requires new operations to maximize and create insight. Moreover, AI and big data demand new concepts for building advanced computing systems with a perfect balance of compute, memory access, and network to finally access data on persistent storage (see also the feature article "Mastering Multidimensional Heterogeneity of Computers" further below).

Of course, the NHR Alliance also wants to move into emergent fields that are not currently core to HPC. At ZIB, we focus on quantum computing and communication since we believe that this field has the potential to revolutionize HPC in the long term. Presently, we are mainly aiming at developing a workflow from the design of hybrid quantum-classical algorithms up to their simulation by means of quantum simulators on HPC hardware. We consider this specifically in the context of combinatorial optimization problems, because of the industrial relevance as well as first promising results from quantum computing for this problem class (see also the special topic "Quantum Computing").



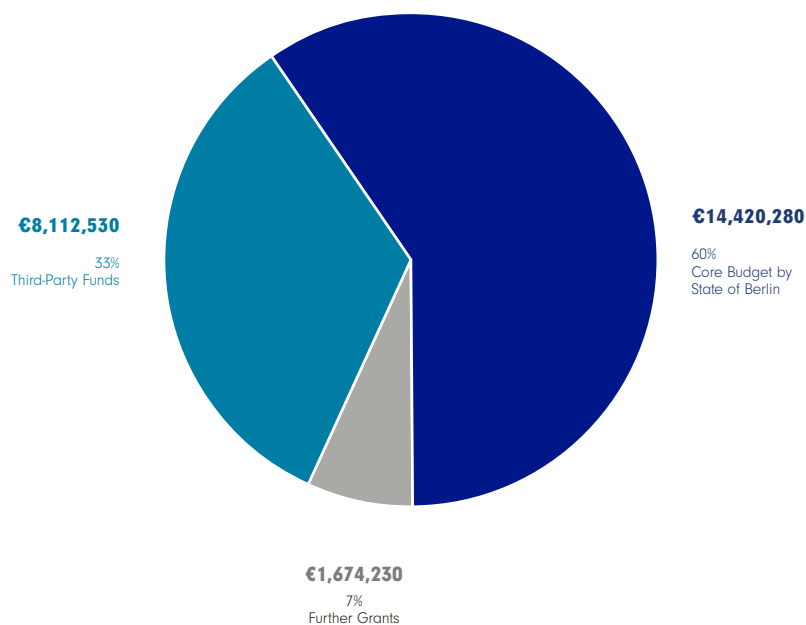
Prof. Dr. Christof Schütte  
Phone +49 30 841 85-101  
schuette@zib.de

# ECONOMIC SITUATION IN 2021

**In 2021, the total income of ZIB comprised 24.4 million euros.** The main part of this was made available by the Federal State of Berlin as the basic financial stock of ZIB (14.4 million euros) including investments as well as the budget of the NHR center at ZIB (4,7 million in 2021, made available by the Federal State of Berlin but co-funded by the federal government and the German Federal States participating in HLRN).

The second largest part of the budget resulted from third-party funds (8.1 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 research service budget of KOBV, summing up to almost 1.7 million euros in total.

## ZIB INCOME

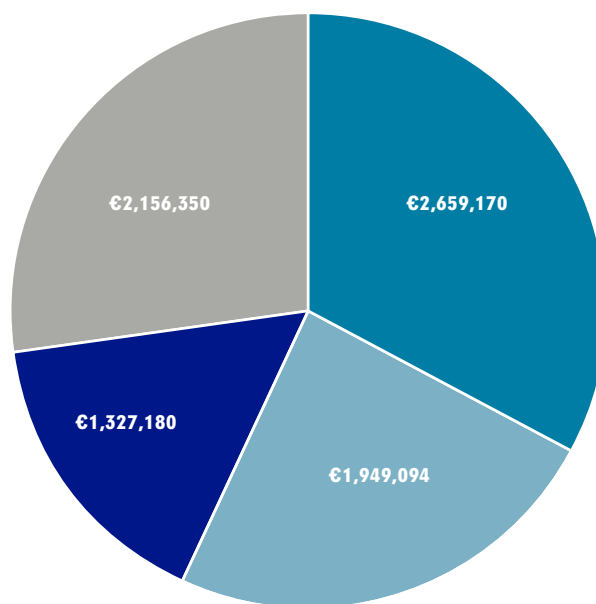




The Zuse Institute Berlin (ZIB) finances its scientific work via three main sources: the basic financial stock of the Federal State of Berlin and third-party funds from public sponsors and those of industrial cooperation contracts.

In 2021, ZIB raised third-party funding through a large number of projects. Project-related public third party funds increased from 5.2 million euros in 2020 to almost 6.8 million euros in 2021, industrial third-party projects declined from 1.8 million euros to 1.3 million euros. In total, 8.1 million euros in third-party funding marked an increase by more than 15% compared to 2020 despite the ongoing covid-19 crisis.

## ZIB THIRD-PARTY FUNDS BY SOURCE



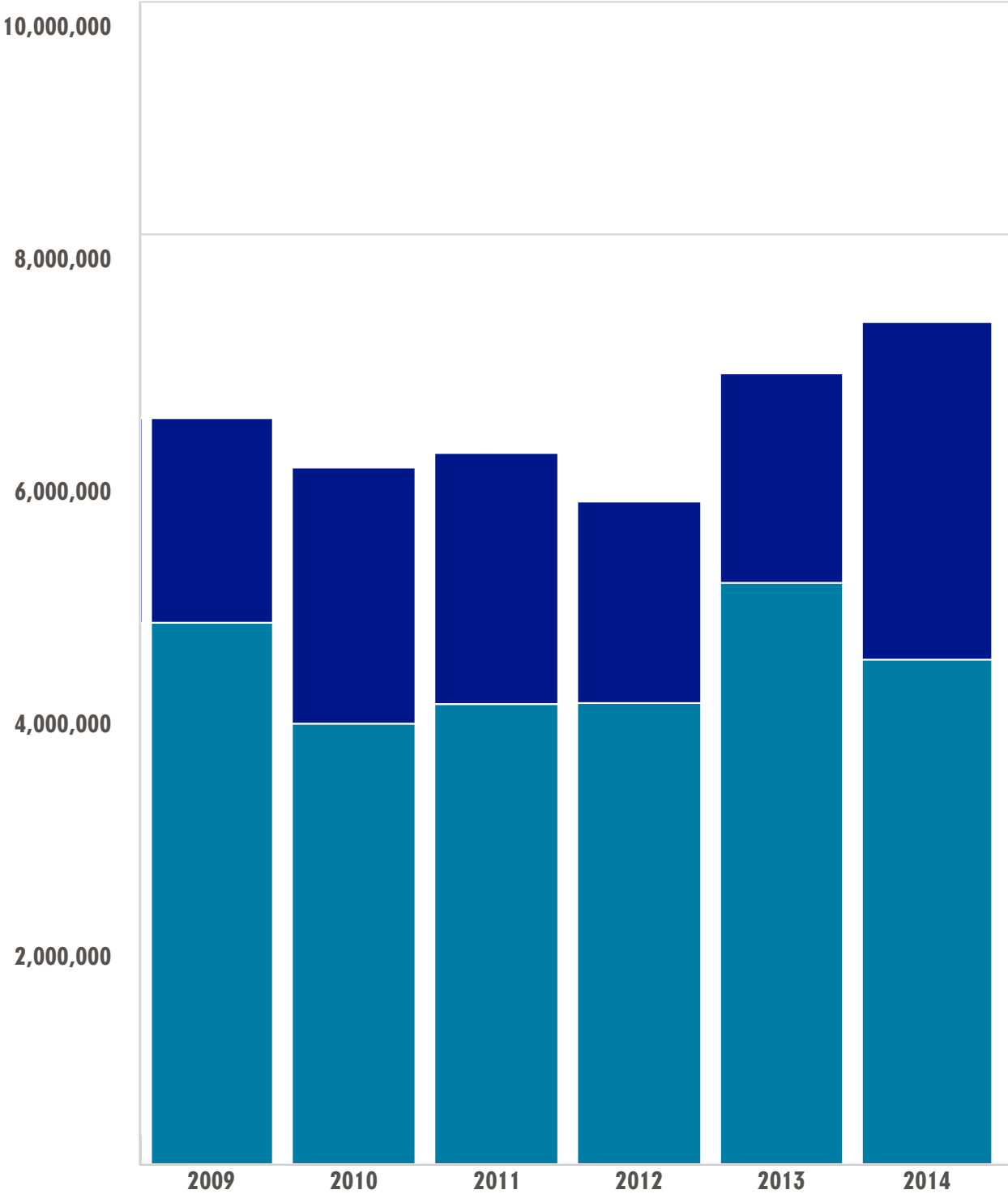
33% BMBF  
incl. FC Modal

16%  
Industry

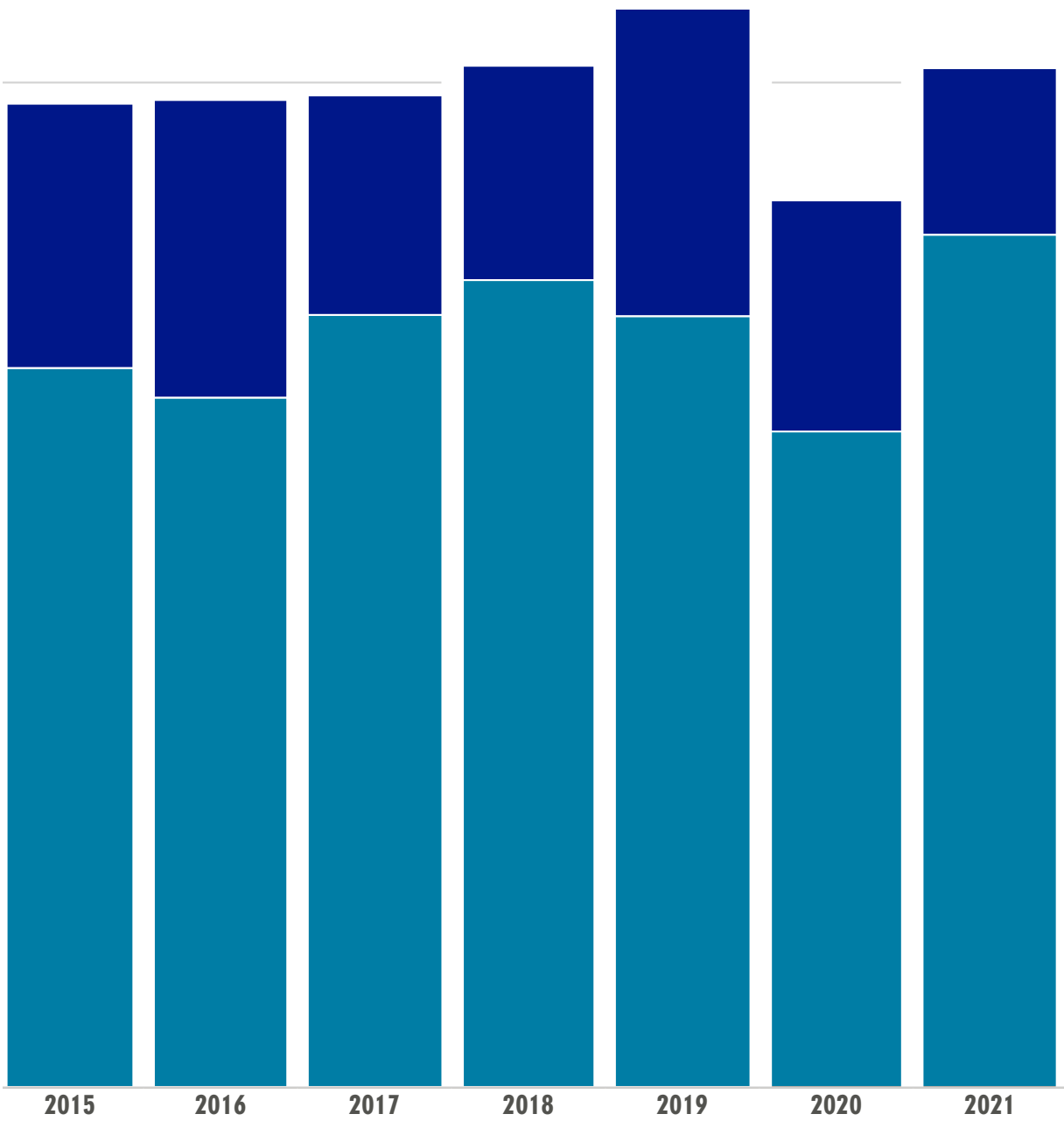
27% DFG  
incl. Math+

24%  
Other Public Funds

# ZIB THIRD-PARTY FUNDS IN EUROS



Economic Situation



■ Industry  
■ Public Funds

# Spin-Offs

## **Computing in Technology GmbH (CIT)**

1992 | [www.cit-wulkow.de](http://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](http://www.risk-consulting.de)

Database marketing for insurance companies

## **Intranetz GmbH**

1996 | [www.intranetz.de](http://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](http://www.visageimaging.com)

Advanced visualization solutions for diagnostic imaging

## **atesio GmbH**

2000 | [www.atesio.de](http://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](http://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](http://www.lenne3d.com)

3-D landscape visualization, software development, and services

## **JCMwave GmbH**

2005 | [www.jcmwave.com](http://www.jcmwave.com)

Simulation software for optical components

## **onScale solutions GmbH**

2006 | [www.onscale.de](http://www.onscale.de)

Software development, consulting, and services for parallel and distributed storage and computing systems

## **Laubwerk GmbH**

2009 | [www.laubwerk.com](http://www.laubwerk.com)

Construction of digital plant models

## **1000shapes GmbH**

2010 | [www.1000shapes.com](http://www.1000shapes.com)

Statistical shape analysis

## **Quobyte Inc.**

2013 | [www.quobyte.com](http://www.quobyte.com)

Quobyte develops carrier-grade storage software that runs on off-the-shelf hardware

## **Keylight GmbH**

2015 | [www.keylight.de](http://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

## **Exazyme**

2021 | [www.exazyme.com](http://www.exazyme.com)

Exazyme provides AI solutions for protein engineering

# Number of Employees

In 2021, 236 people were employed at ZIB; of these, 146 positions were financed by third-party funds. The number of employees grew in comparison to 2021, mainly because of progress in establishing the NHR center at ZIB and as a consequence of newly starting third-party funded projects.

1/1/2021			1/1/2022			
2	0	2*	2	0	2*	MANAGEMENT <small>*without temporary management</small>
28	86	114	31	89	120	SCIENTISTS
41	3	44	43	3	46	SERVICE PERSONNEL
14	1	15	14	2	16	KOBV HEADQUARTERS
0	44	44	0	52	52	STUDENTS
85	134	<b>219</b>	90	146	<b>236</b>	<b>Total</b>
Permanent	Temporary	<b>Total</b>	Permanent	Temporary	<b>Total</b>	





# SPREADING DYNAMICS

## Cultural Change, Covid-19 and a Mobility Transition

Understanding spreading processes and their dynamics is a topic of great interest in a variety of domains or applications, such as the spread of a virus in epidemics, innovation diffusion in cultural contexts, or information propagation in social networks. Increasing mobility and new communication technologies have transformed human interaction patterns from local to global scales. Because of this disrupting transformation, spreading processes are not a local phenomenon anymore, but are affecting people's lives worldwide. At ZIB, researchers are developing new mathematical tools and models for studying complex spreading mechanisms for real-world applications in the past, present, and future. Together with domain experts and research partners from various fields, new approaches and modeling techniques have been developed and applied successfully to multiple real-world systems, such as the cultural change of ancient societies, the spread of COVID-19, and mobility transitions of the future.



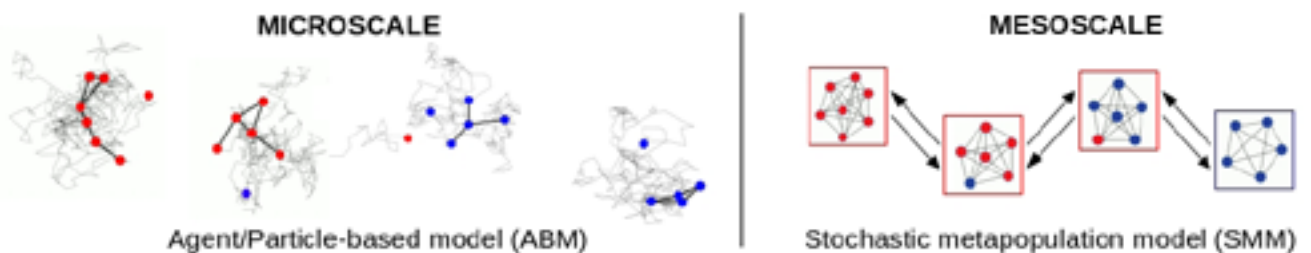
# Introduction

Can we predict and prevent epidemic outbreaks? Is it possible to control the spread of fake news? How can we accelerate the diffusion of knowledge? These are crucial questions related to spreading dynamics in real-world systems, such as the current COVID-19 pandemic. In order to gain more insight into fundamental principles underlying these dynamics, researchers often use mathematical models. Precise description of such models allows for their thorough mathematical analysis, and their abstract formulation makes them generalizable to a wide range of complex social systems.

How are mathematical models for spreading dynamics built? Basic elements of any spreading dynamics in a social system are individuals and their interactions. Every interaction creates a new opportunity for a change. During the history of humankind, types of interactions have changed significantly, from entirely physical interactions in ancient times to the possibility of contact between any two people in the world through the Internet nowadays. Recent technological advancement in obtaining large amounts of data characterizing social contagion, such as high-resolution population and mobility data, introduced new data-based descriptions of underlying complex systems.

Researchers at ZIB develop novel mathematical models for a variety of applications to better understand, predict, and control spreading dynamics on different spatiotemporal scales. The main focus is on macroscale, mesoscale, and microscale models, which will be introduced briefly below.

**Microscale:** The most detailed descriptions of real-world systems are achieved by models on a microscale. Through microscale descriptions of the spreading dynamics these models can capture global patterns on a large system scale. An example are agent-based models (ABMs) that consist of a set of rules describing the movement of agents and their interaction with each other. Interacting agents can alter their own or the other agents' properties, for example, by changing an opinion or transmitting an infection. Information about agents' mobility and status dynamics is typically deduced from available data sets and expert knowledge. Due to their versatility, ABMs are very attractive for modeling spatiotemporal population dynamics in various disciplines. However, ABMs are very costly to simulate and to recalibrate to new parameter sets, especially for large population numbers that appear in real-world systems.



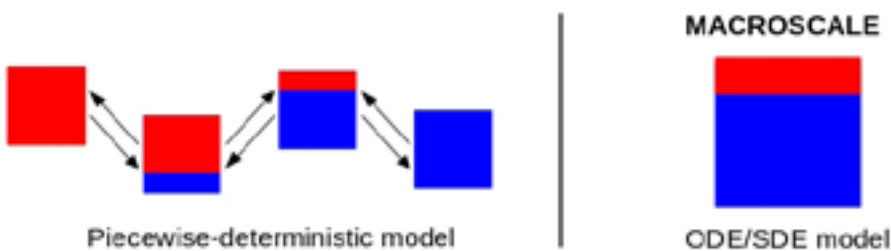
**Figure 1:** Hierarchy of modeling approaches. Microscale: dots refer to agents that have a status (blue or red), straight lines between dots indicate closeness of agents that enable interaction, and gray lines show movement trajectories of the agents. Mesoscale: boxes represent subpopulations, and connections between boxes denote possible transitions between subpopulations. On the mesoscale and macroscale, the population state is depicted by the proportion of blue and red.



**Mesoscale:** Models on a mesoscale are characterized by reduced granularity compared to the models on a microscale. This is achieved, for example, in stochastic metapopulation and piecewise-deterministic models, by dividing the spatial environment into areas, called subpopulations. Within each subpopulation, individuals of the same status are considered indistinguishable, and an agent can interact with all other agents of the same subpopulation anytime. On the contrary, interaction of agents from different subpopulations are considered rare in comparison. Despite the loss in spatial resolution compared to ABMs, mesoscale models offer a significant gain in computational efficiency and preserve important properties of the spreading dynamics. For this reason, they are often a model of choice for studying spreading processes in large populations.

**Macroscale:** Compared to models introduced above, models on a macroscale are the simplest representations of real systems, as they describe only the global change of the system over time. In large populations, the spreading process on a macroscale can be approximated by e.g. deterministic dynamics in the form of ordinary differential equations (ODEs). ODE models assume homogeneous mixing of the population, meaning that they do not consider any spatial component and thus do not involve mobility dynamics. This results in efficient simulations using ODE solvers, whose effort is independent of the population size. For this reason, ODE models of spreading dynamics are often used as surrogate representations of spatially resolved models.

More details about these modeling techniques will be described in the following text with respect to the individual applications in the past, present, and future social systems.



Dr. Nataša Djurdjevac Conrad  
Phone +49 30 841 85-439  
natasa.conrad @zib.de

# Modeling the Past: Cultural Change in Ancient Tunisia

Spreading dynamics in ancient times differs in many aspects from the spreading we observe in modern times, especially in the types of interactions, the mobility patterns, and spatiotemporal scales. Due to the lack of modern transportation and communication technology in ancient times, most human interactions were happening on a local spatial scale, for example, in the same region. Human migrations over large distances happened less often and were governed by strong influences, such as social conflicts, unsuitable environmental conditions, and climate change. Both interactions, within and between different geographical regions, led to the spreading of socioeconomic and technological influences that triggered cultural change [1]. However, modeling these dynamics is often hindered by the sparsity, incompleteness, and uncertainty of available data, due to a limited amount of archaeological traces. Obviously, these gaps and missing data are in most cases impossible to fill or obtain, as the past processes cannot be replicated or observed anymore.

Within the MATH+ project “Data-Driven Modeling of the Romanization Process of Northern Africa,” researchers at ZIB study cultural change in ancient Tunisia caused by the influence of the Roman Empire. Unlike infection dynamics, which usually assumes an interaction between two individuals, Romanization is a so-called complex meta process with different interacting layers, for example, on a social, political, and economic level. We study processes of change on these levels as important indicators of the Romanization process and use them to infer possible interaction and spreading dynamics [2]. The starting point of the Romanization process was in 146 BC, when the region, known nowadays as Tunisia, was annexed by Rome as the outcome of the Third Punic War. The African province expanded further in the following centuries, and we study its development until 350 AD.



Map 1



Map 2

Through a fruitful collaboration with the German Archaeological Institute, we were able to work with archaeological data on the Romanization in ancient Tunisia [2]. In this data set, the most important indicators for the cultural change are the settlement structures of different city types, reflecting the adaptation to the Roman administrative structure. To understand processes that triggered cultural change, the interaction dynamics governing the Romanization spreading needs to be uncovered. However, very little is known about the intercity and interregional interactions from this period and how these changed over time under Roman influence. Within this project, we approach the inference problem on the mesoscale, by considering interac-

tions between different regions in Tunisia [2]. Inferring the interaction patterns and their temporal evolution are hard problems that can have many different solutions, since the same Romanization spreading pattern could have resulted from diverse inter- and intraregional interactions. We reformulate the inference problem to a minimization problem of finding interactions producing the spreading patterns that are close to the ones in archaeological records [2].

Historical networks of influence resulting from this research help understand complex processes of cultural change in ancient times.



Map 3

**Figure 2:** Map 1: Settlement structure in 330 BC–30 BC; Map 2: Settlement structure in 30 BC–300 AD; Map 3: Settlement structure in 300 AD–640 AD; Maps 1–3 [13].

# Modeling the Present: Spreading Dynamics for COVID-19

Since the onset of the COVID-19 pandemic, many different types of models have been proposed to represent the spread of infection. A particular focus has been to study these models under a set of restrictions coming from implemented measures. Most of these models allow for the precise incorporation of real-life data, but face one major disadvantage: it is rather difficult to justify their use for predictions of scenarios with combinations of measures that have not been implemented in reality before. Agent-based models permit such accurate predictions of consequences of measures and thus facilitate modeling the spread of infection on the basis of the spatiotemporal agent interactions with realistic mobility and behavioral patterns. Researchers at ZIB together with partners from TU Berlin work on developing these kinds of ABMs to enable accurate assessment of entirely new (combinations of) measures. The mobility side of the model is based on real-world mobility data and contains complete daily activity chains for every agent [5,6]. The basic assumption is that an infection may occur when an infectious agent and a susceptible agent are in close proximity, for example, in the same room. In this case, the probability of an infection can be computed based on room properties, contact duration, and agent characteristics (see Fig. 3 and [5] for more details).



$$p(\text{infect}|\text{contact}) = 1 - \exp\left(-\Theta \sum_m sh_{m,t} \cdot ci_{nm,t} \cdot in_{n,t} \cdot \tau_{nm,t}\right).$$

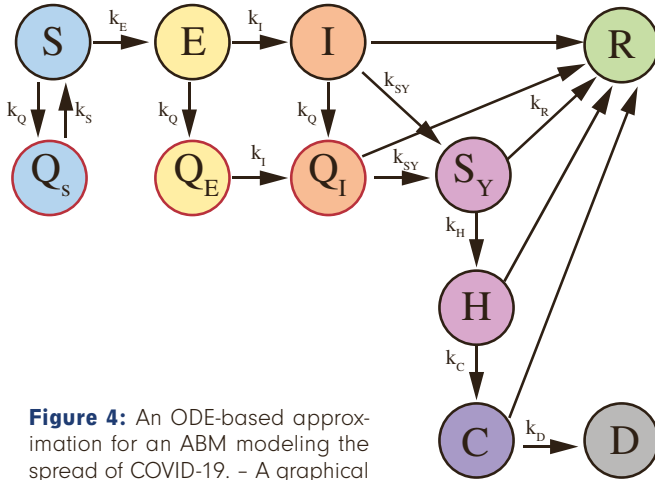
**Figure 3:** Infection model: **m** is the sum of all agents, **sh** the viral shedding of the infectious agent, **in** the intake,  **$\tau$**  is the duration of the contact, **ci** the contact intensity, and  **$\Theta$**  an overall scaling parameter. **sh** and **in** can be reduced by wearing a mask, the **ci** depends on the room size and air exchange, and  **$\tau$**  is based on the mobility model. See [5] for more details.



The main advantage of our approach is that different countermeasures and vaccination effects can be built in directly. Therefore, such ABMs are optimal candidates for micromodels of the spread of infection with full built-in, real-world scenarios. Within the project MODUS-COVID, this ABM has been implemented on ZIB's supercomputer for millions of agents. It has then been used for analyzing the infection dynamics in different regions of Germany (in particular Berlin and Cologne) and for making predictions of the effect of different combinations of countermeasures. Based on this, periodic reports to the German federal government were published (see <https://www.zib.de/projects/mobilitaetsmodelle-berlin> for the reports in German) that attracted significant public attention in national and international newspapers, radio and TV interviews (e.g. Tagesschau), blogs, and other websites.

While ABMs are a useful tool for modeling spatiotemporal population dynamics, their computational cost is very high since a lot of very complex simulations are required to reliably sample quantities of interest. In particular, large numbers of agents render the sampling infeasible. To overcome this challenge, we have developed so-called model reduction techniques, which lead to a significant gain in computational efficiency, while preserving important dynamic properties. For example, based on a precise mathematical description of spatiotemporal ABMs, we derived a reduced model in terms of stochastic partial differential equations (SPDEs) that describes the evolution of agent number densities for large (but finite) populations. Thus, the model complexity and simulation costs can be drastically reduced [7]. Other approaches such as stochastic metapopulation models [8] or networks of ODEs [8], [9] have also been derived successfully.

Once a reduced model has been found – for example, an ODE model (see Fig. 4) – optimal parameters can be determined by using simulation data from the original ABM, generated from high-performance computing approaches. That way, reduced models can be built that allow many different combinations of countermeasures to be mimicked and represent some of the building blocks within a target model. In [6], we have demonstrated that this strategy allows the use of machine-learning methods for finding the optimal parameter set for all feasible combinations of given countermeasures. As demonstrated in [6], this allows the multiobjective optimization problem of finding optimal combinations of countermeasures to be solved.



**Figure 4:** An ODE-based approximation for an ABM modeling the spread of COVID-19. – A graphical representation of the reduced ABM. Shown are the modeled compartments and their relations. The modeled population is divided into 11 compartments: susceptible  $S$ , quarantined  $Q_s$ , exposed  $E$ , noninfectious self-isolated  $Q_e$ , infectious  $I$ , infectious self-isolated  $Q_i$ , symptomatic  $S_y$ , hospitalized  $H$ , critical  $C$ , deceased  $D$ , and recovered  $R$ . Right: The differential equations describing the development of this system over time. For more details see [6].

$$\frac{dS}{dt} = -k_E \frac{S \cdot I}{N} + k_{S_y} Q_s - k_Q S$$

$$\frac{dE}{dt} = k_E \frac{S \cdot I}{N} - k_I E - k_Q E$$

$$\frac{dI}{dt} = k_I E - k_R I - k_{S_y} I - k_Q I$$

$$\frac{dS_y}{dt} = k_{S_y} I + k_{S_y} Q_i - k_R S_y - k_H S_y$$

$$\frac{dH}{dt} = k_H S_y - k_R H - k_C H$$

$$\frac{dC}{dt} = k_C H - k_R C - k_D C$$

$$\frac{dD}{dt} = k_D C$$

$$\frac{dR}{dt} = k_R I + k_R S_y + k_R H + k_R C + k_R Q_i$$

$$\frac{dQ_s}{dt} = -k_S Q_s + k_Q S$$

$$\frac{dQ_e}{dt} = -k_I Q_e + k_Q E$$

$$\frac{dQ_i}{dt} = k_I Q_e + k_Q I - k_{S_y} Q_i - k_R Q_i$$

# Modeling the Future: a Sustainable Mobility Transition

As for past and present social phenomena, spreading processes are an important element when analyzing potential future developments in social systems, such as whether or not sustainable behavior spreads through a population. Often, ABMs are used to model the underlying complex socio-ecological, -economic, and -technical systems. Beyond modeling alone, however, researching potential sustainability transitions requires dialogue processes with societal stakeholders, for example, from the policy and business worlds as well as citizens. Researchers at ZIB and FU Berlin, the Global Climate Forum, and Arizona State University are collaborating on the Decision Theater (DT), an instrument for combining models and discussion [10].

An example case is mobility in Germany: with transport emissions having remained at roughly the same level since 1990, DT participants discuss how to achieve the emission-reduction goals of the Federal Climate Change Act. They explore alternative policy options and expected future developments using the Mobility Transition Model (MoTMo); the modelers in turn benefit from the participants' practical knowledge of the real-world problem for further developing MoTMo.

ABMs are easy to explain to non-modelers and DT participants can identify with the modeled agents. In MoTMo, people choose a means of transport for fulfilling their mobility needs – as all of us do almost every day. Agents represent the German population, statistically matching distributions of, for example, age, income, and population density. The spread of opinion – for example, about the convenience of electric vehicles or car sharing – is part of the agents' mobility choice, as are collective decision-making

in households under common budget constraints, spatial developments like charging infrastructure rollout, and policy setting. Within MATH+, we develop a mathematical formulation of this large-scale computational ABM.

Even using the high-performance computing infrastructures at ZIB, MoTMo runs, based on millions of agents, take hours; however, precomputed results somewhat limit the interaction possibilities of DT participants with the model. As for COVID modeling, we thus need reduced models to forecast aggregate quantities. Here, these are, for example, the modal split, that is, how many people use which means of transport, or the resulting greenhouse gas emissions.

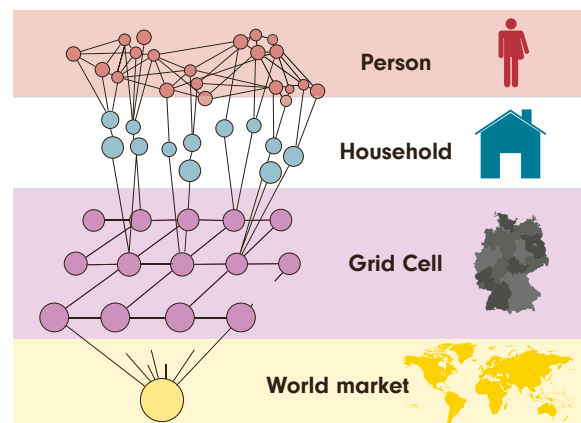
We can “automatically” obtain such reduced models using machine learning, based on several short simulations of the original system, for many different initial states. In this way, we get expected values of the dynamics – often not taking longer than a few complete simulations of the original model. For example, we compute a linear and finite-dimensional matrix  $L$  to approximate the infinite-dimensional infinitesimal generator of the Koopman operator using generator Extended Dynamic Mode Decomposition [11], [12]. We obtain a reduced dynamical model by minimizing

$$\|d\Psi_x - L^T \Psi_x\|_F$$

for the transformed data matrices  $d\Psi_x$  and  $\Psi_x$  for the data  $\mathbf{X}$ . If the initial points chosen cover the dynamics of the system well, the reduced model is a good approximation of the original system and can be used to forecast the behavior of the aggregate quantities in time.

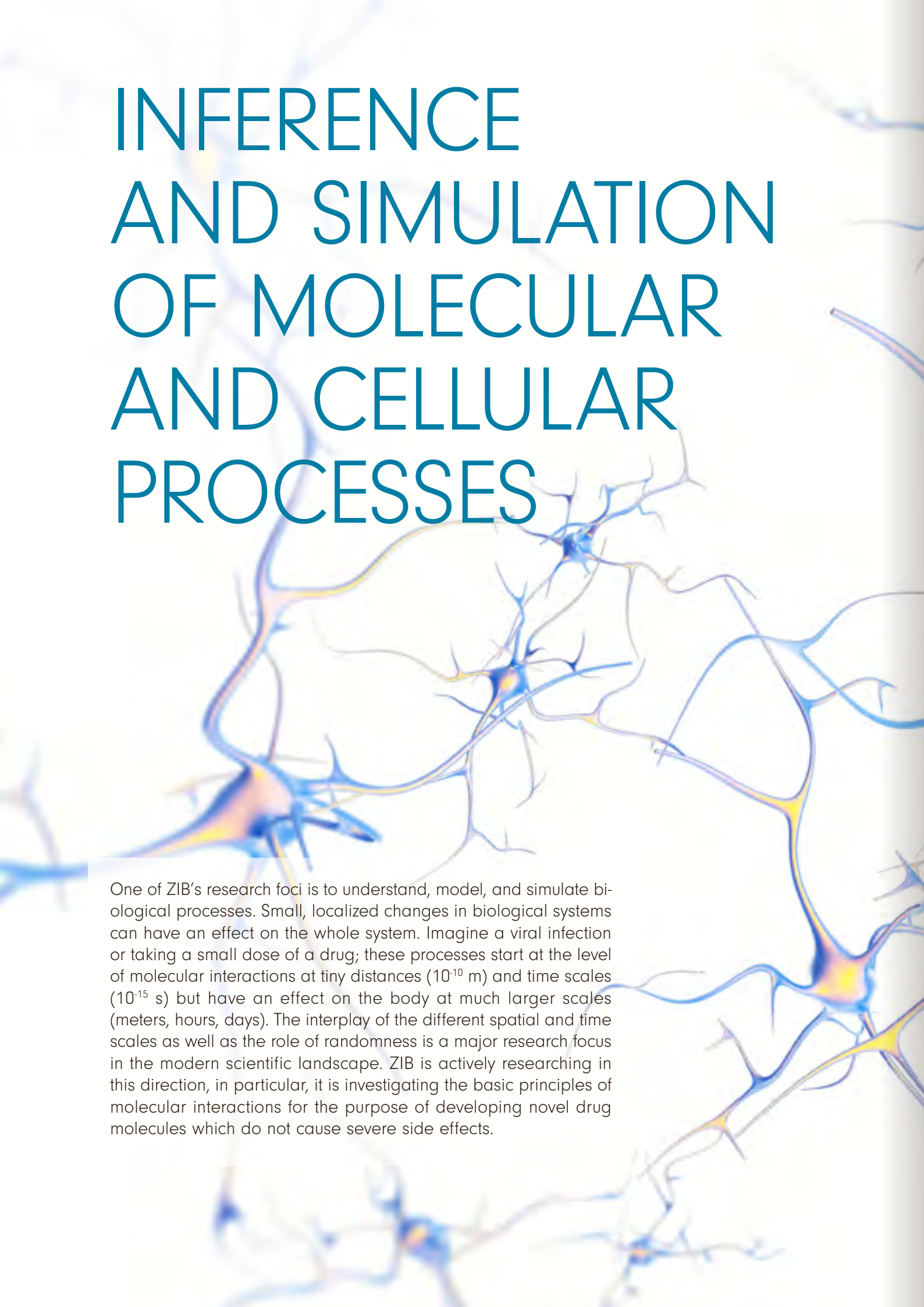


**Figure 5:** Participants composing scenarios during a Decision Theater workshop in the Studio Da Vinci at ZIB.



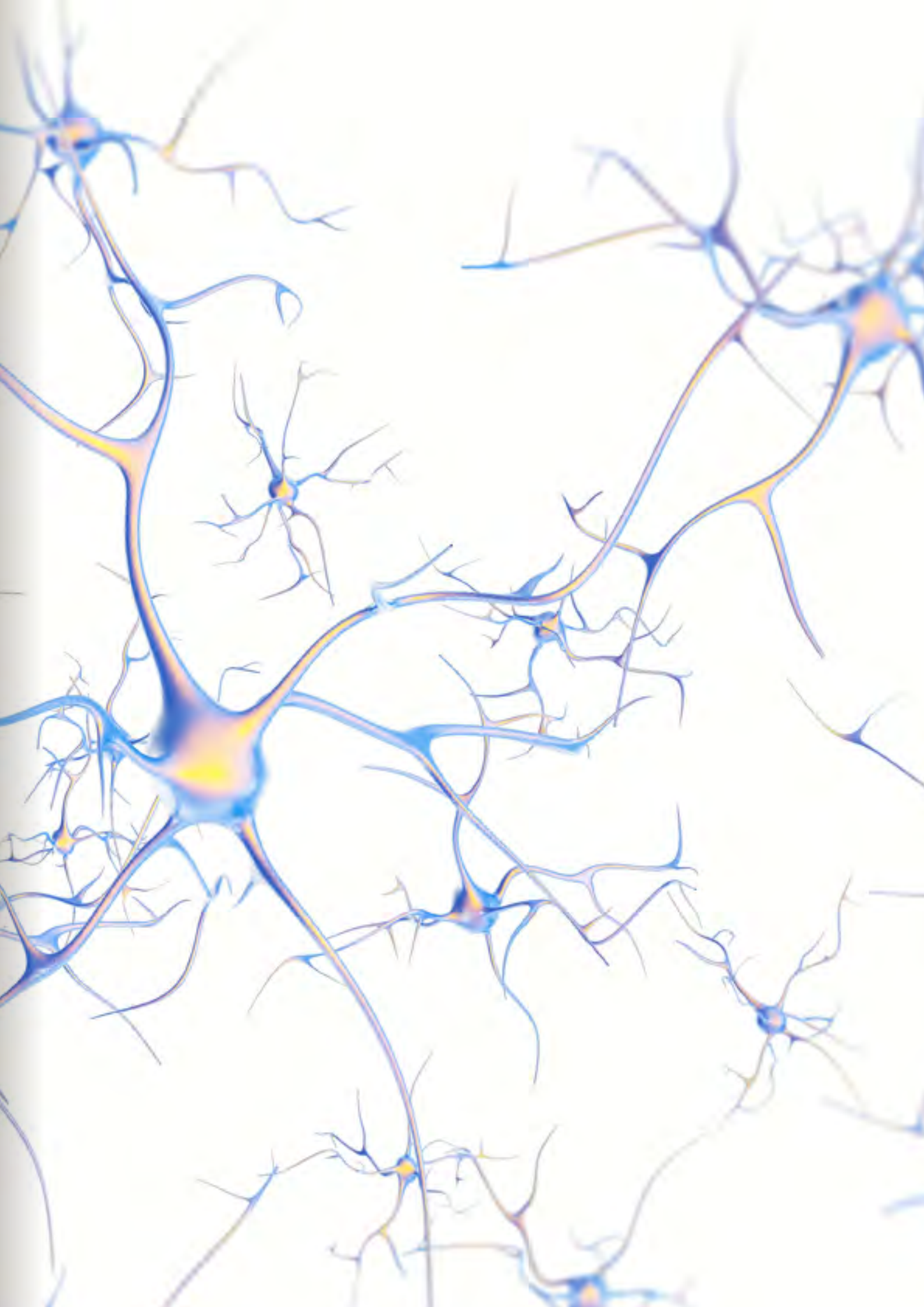
**Figure 6:** Structure of the Mobility Transition Model.

# INFERENCE AND SIMULATION OF MOLECULAR AND CELLULAR PROCESSES



One of ZIB's research foci is to understand, model, and simulate biological processes. Small, localized changes in biological systems can have an effect on the whole system. Imagine a viral infection or taking a small dose of a drug; these processes start at the level of molecular interactions at tiny distances ( $10^{-10}$  m) and time scales ( $10^{-15}$  s) but have an effect on the body at much larger scales (meters, hours, days). The interplay of the different spatial and time scales as well as the role of randomness is a major research focus in the modern scientific landscape. ZIB is actively researching in this direction, in particular, it is investigating the basic principles of molecular interactions for the purpose of developing novel drug molecules which do not cause severe side effects.





# Out of Equilibrium

Every biological process can be simplified by describing it as the net reaction of the interplay of **excitation processes** with **inhibitory processes**. The delay between excitation and inhibition dictates the equilibrium state the process will reach. The perturbation of such equilibria in our body causes diseases. For example, when inhibiting signals are lacking in the brain of individuals, they suffer from sporadic brain signaling and experience epileptic fits or seizures.

A disease perturbs the system and moves it into a new equilibrium. Perturbations need not be fatal, but rather occur over a long-time scale like in chronic disease and ageing. In such cases, by simply eliminating the disease, the system does not naturally need to return to the original equilibrium. For example, drugs can inhibit pain, but continued drug use has an effect on the patients' heart, the digestive system, and the liver, leading to an altered physical state that cannot simply be reversed after



the medication is discontinued. Hence, it is not sufficient to only stop the disease, but it is also necessary to drive the system back to its natural equilibrium. To achieve the latter, it is required to study the **multiple time scales** by which the disease evolves and its downstream **spatial effects**, and also to take the biological **stochastic effects** into account.

At ZIB, several research projects try to shed light on complex multiscale biological processes. This article reports on a group of projects focusing on a new model for neural transmission, which encapsulates the effect of a drug in-

tervention from signal transmission to pain inhibition. This model is the first to incorporate intracellular signaling, which is key for understanding the side effects of a drug. In the broader framework of this development, ZIB has aided its biologist partners toward better experimental design and data acquisition, and aims to enrich drug discovery in the future by also incorporating side effects into the drug-screening process through its mathematical models. In the longer term, ZIB aims to efficiently screen for new drugs without any major side effects and to cross-validate its mathematical findings in collaboration with their experimental partners in the wet lab.



PD Dr. Marcus Weber  
Phone +49 30 841 85-189  
weber@zib.de



Dr. Vikram Sunkara  
Phone +49 30 841 85-432  
sunkara@zib.de

# Building Knowledge from Data

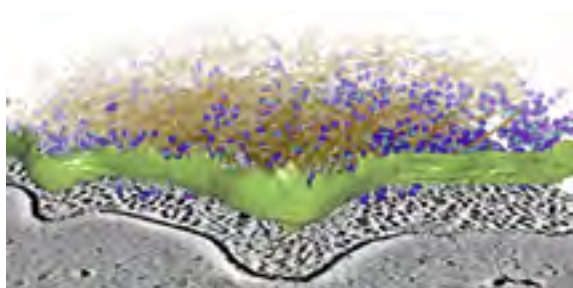
Curating and analyzing data is key to building a strong knowledge base from which realistic and meaningful mathematical models can be designed. ZIB is working with modern Machine Learning (ML) and Artificial Intelligence (AI) methods to help its wet lab biologist partners to enrich their experimental setup and analysis pipelines. For example, we use AI to segment molecules in a cell

from cryo-electron tomography to investigate their spatial distribution and interactions in the cell. Likewise, we use ML to look for small changes in the fluorophores that are attached to proteins to see their conformational changes. Such data sets, in turn, give ZIB an edge in accessing state-of-the-art data to feed into our mathematical models.

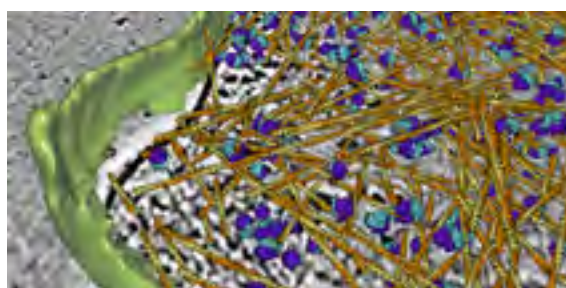
## Segmenting Nanometer-Scaled Structures in a Cell

Cells are the basic structural and functional units of all organisms. Each cell consists of a membrane enclosing the cytoplasm, which contains a large number of different molecular structures like proteins and nucleic acids. Depending on the type of organism, the cell type, and the particular function of a cell, the composition of molecular structures inside a cell can vary immensely. The understanding of cellular and molecular processes therefore, as a first step, demands the identification of all molecular structures as well as their interactions in living cells. Several imaging techniques can be applied to obtain this information. Among those techniques, electron tomography allows the acquisition of three-dimensional images with the highest resolution. However, these images are often noisy and suffer from artifacts. As a consequence, they are difficult to analyze.

Over the last decade, researchers at ZIB together with their collaborators from different institutes have developed methods to extract macromolecular structures, in particular, filamentous structures like microtubules [1] and actin filaments [2] (Figs. 1 and 2). Apart from the segmentation of such structures, the stitching of serial-section tomograms [3] has played an important role in allowing large subcellular structures like the spindle apparatus to be fully reconstructed [4]. However, the previously employed methods for the detection of macromolecular structures are restricted to rather large molecules (Figs. 1 and 2). In order to allow the identification of smaller structures, ZIB researchers have recently started to explore deep neural networks (DNNs) for this task with promising first results. The large amounts of training data required for DNNs represent a particular challenge for this task. Therefore, the use of synthetically created training data is currently being explored.



**Figure 1:** Crowded cell environment reconstructed from cryo-electron tomography. At the bottom, a cross section of the tomogram is shown. The green surface depicts part of the reconstructed cell membrane. The elongated structures are actin filaments and the violet-blue macromolecules are ribosomes. Data courtesy: Alexander Rigort (MPI). Image courtesy: Norbert Lindow (ZIB).



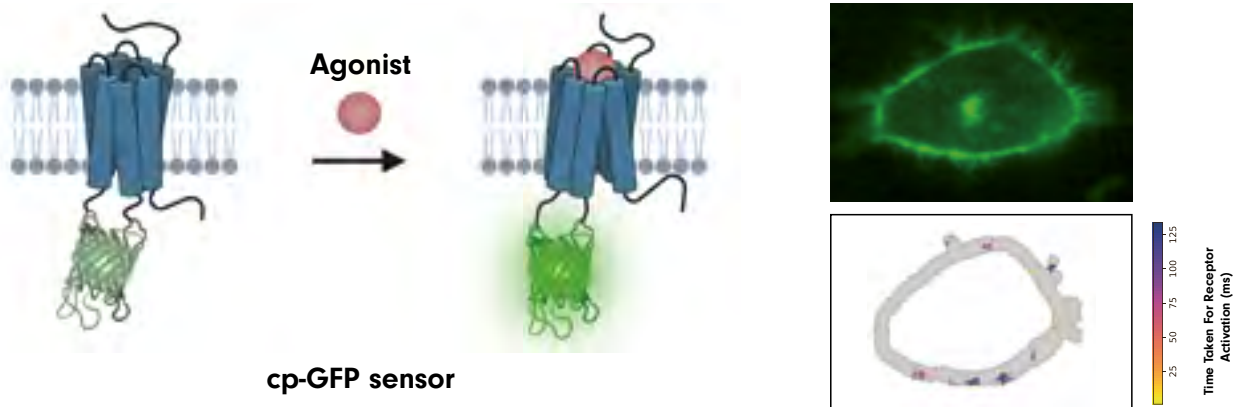
**Figure 2:** Different view and close-up of the crowded cell environment shown in Fig. 1. Data courtesy: Alexander Rigort (MPI). Image courtesy: Norbert Lindow (ZIB).

# Observing Real-Time Changes in Receptor Configurations

Cells are highly complex structures, in which many processes occur on a scale of milliseconds and even nanoseconds. Analyzing and examining them *in vivo* can be very difficult and intriguing. Difficulties mainly arise from the dynamic behavior of these processes and their sensitivity to factors like changes in temperature, drug induction, and more. However, one way of tackling these difficulties is to make use of single cell imaging, paired with fluorescent-based probes. This approach allows for the noninvasive *in vivo* observation of receptors in single cells at small time scales over time. Together with biologists from the Max Delbrück Center for Molecular Medicine (MDC), researchers at ZIB have developed a single-pixel level, fluorescent-based, sin-

gle cell imaging technique. Using this new technique, we have been able to uncover underlying heterogeneities, as well as coherency between small receptor instances and have discovered spatial clusters.

For example, we have been able to investigate the activation patterns of the  **$\mu$ -opioid receptor**, a drug target for pain inhibition in the brain; the **dopamine receptor**, a drug target for depression; and the **muscarinic acetylcholine receptor**, a drug target for heart disease. We were able to study the activation behavior of the individual receptors on a live cell (Fig. 3).



**Figure 3:** (Left) Sensor attached to the  $\mu$ -opioid receptor and the sensor being activated in the presence of an agonist. Generated using [BioRender.com](https://www.biorender.com) by Ali Isbilir (left) receptors emitting light after activation in a live cell (right). The spatial clusters found after spatial analysis.

## Mathematical Modeling

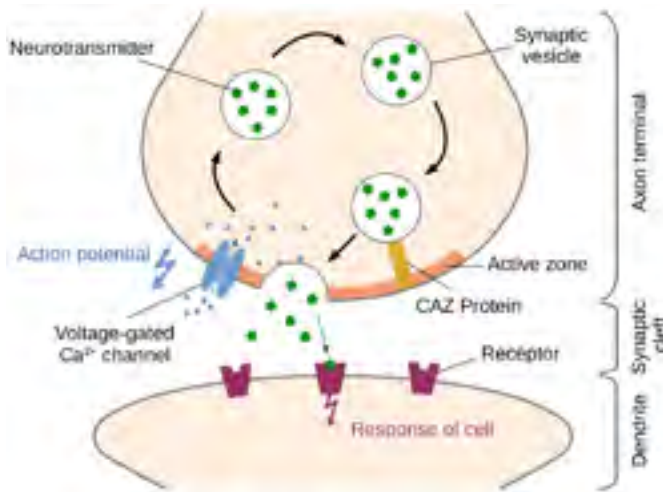
The aim of ZIB is to design new drugs with reduced side effects. To achieve this, ZIB is designing spatiotemporal mathematical models to describe the disease and its downstream effects. In the case of pain inhibition, ZIB has developed mathematical models which capture the pain signal in the presynaptic region and the drug activating the receptors in the postsynaptic region that receives the

pain signal, followed by the inhibition of calcium signaling through the activated receptor, which stops the pain signal from being passed on. This model spans from the nanoscale to the mesoscale, and all scales are important to understand the process, so that it is possible to design drugs that successfully interfere with the pain transmission while minimizing cascading side effects.

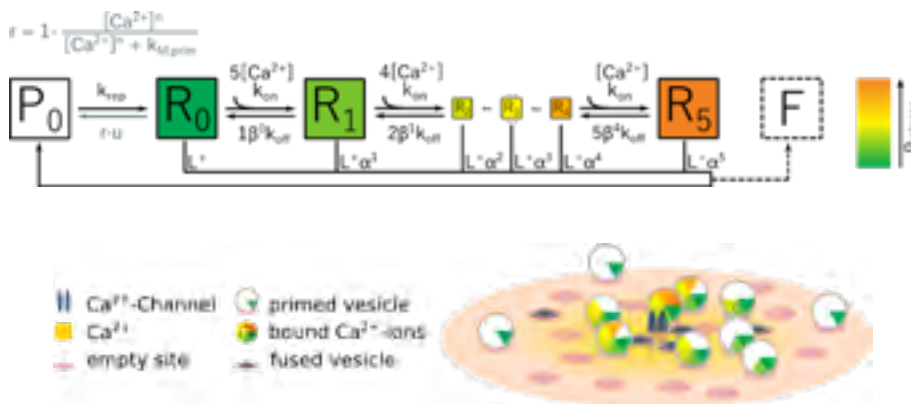
### NEURONAL SIGNAL TRANSMISSION MODEL – THE WAY PAIN WORKS

A central example for a biological process that contains several length and time scales and is of special significance for drug development is the process of neurotransmission. It describes the information transfer from one neuron to another, and its disorders are involved in many diseases. At chemical synapses, the arriving electric signal is translated into the release of messaging molecules (neurotransmitters) that diffuse across the synaptic cleft and set off an excitation in the postsynaptic neuron. The neurotransmitters are contained in vesicles located near the presynaptic membrane in a region called the active zone. An arriving action potential prompts the opening of voltage-gated calcium channels in the active zone, which causes an inflow of  $\text{Ca}^{2+}$ -ions into the presynaptic terminal. The ions dock to the primed vesicles, triggering their fusion with the presynaptic membrane, and thereby release the neurotransmitters into the synaptic cleft. See Figure 1a.

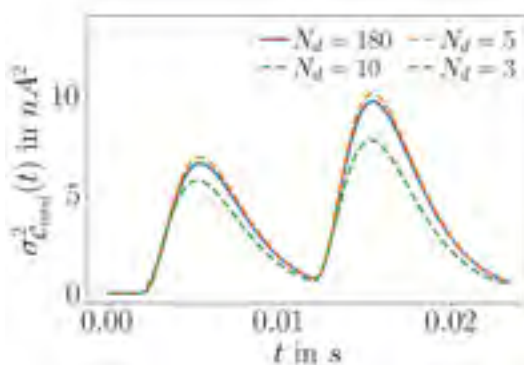
Vesicle release is a stochastic process in which the release probability – and thereby the postsynaptic current – is determined by the vesicle's distance to the calcium channels. An appropriate model of the process must be able to replicate both the mean and the variance of the postsynaptic currents observed in electrophysiological experiments. Previously, this required computationally expensive numerical simulations of stochastic processes. Researchers at ZIB have developed a method for direct computation of the exact first- and second-order moments for models that generate the postsynaptic current by convolving an impulse response function with the output signal of a linear reaction network [5]. This renders expensive stochastic simulations obsolete. Furthermore, a spatially extended vesicle fusion model has been developed, which allows analysis of the dependence of the dynamics on the spatial location. See Figure 1b/c.



**Figure 1a: Neurotransmission process at chemical synapses.** Upon arrival of an action potential, voltage-gated calcium channels open in the active zone. The inflowing  $\text{Ca}^{2+}$  ions bind to proteins on the surface of primed (docked and prepared) vesicles, causing them to release neurotransmitters into the synaptic cleft. After diffusing across the cleft, the molecules activate receptors in the postsynaptic membrane, triggering a new action potential.



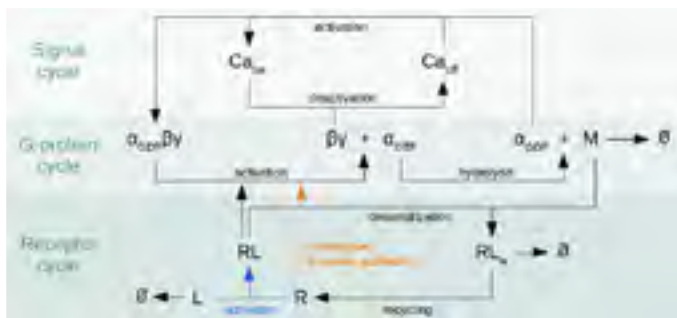
**Figure 1b: Model of synaptic vesicle fusion.** Each release site can either be empty (state  $P_0$ ), or there is a vesicle docked to it, which then switches itself between calcium-binding states:  $R_0, \dots, R_5$ . From each of these states, fusion can occur, leading back to an empty release site and an increase in fused vesicles  $F$ . A docked vesicle can detach again, thereby resetting the release site to the empty state (unpriming reaction, see gray arrow).



**Figure 1c: Variance of the current for different degrees of spatial coarseness.** Note how a coarseness of  $N_d = 10$  is sufficient to approximate a much finer spatial resolution of  $N_d = 180$ .

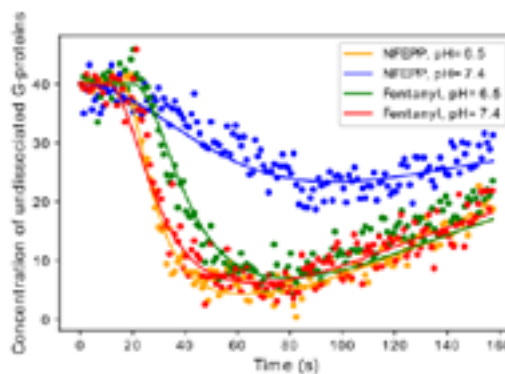
## POSTSYNAPTIC CALCIUM CHANNELS - A DRUG TARGET FOR PAIN RELIEF

Once neurotransmitters activate the receptors of the postsynaptic neuron, they evoke a multilevel signaling pathway, which have been modelled and analyzed at ZIB. It is composed of a biochemical reaction network which connects the receptor-ligand interactions to the G-protein cycle, and further to the signal cycle of calcium channel opening and closing, see Figure 2a. This model has been used for comparing the prototype opioid painkiller NFEPP, which had been designed at ZIB in previous years, to the conventional painkiller fentanyl. The studies showed a markedly diminished effect (i.e., less dissociated G-proteins and consequently a lower number of closed calcium channels) of NFEPP in non-injured environments (normal pH) compared to injured tissues (low pH), see Figure 2b. These insights confirm that NFEPP elicits less adverse side effects than conventional painkillers. Moreover, the studies indicated an enhanced constitutive G-protein activation but lower probability of ligand binding with increasing radical concentrations. The results suggest that, compared to radicals, low pH is a more important determinant of overall GPCR function in an inflamed environment. Such insights are far easier to investigate and explore through mathematical modelling than in the wet lab. Furthermore, ZIB aims to incorporate its mathematical modelling directly into the drug discovery pipeline, such that new drug screening can incorporate the downstream signaling insights into its search.



**Figure 2a:** Biochemical reaction network for the mu-opioid receptor signaling pathway, connecting the receptor cycle to the G-protein cycle and further to the signal cycle of membrane calcium channel modulation [6]. The focal point of the studies carried out at ZIB is the analysis of the impact that the rates for ligand-induced receptor activation (blue) and for constitutive G-protein activation (orange) have on the overall dynamics.

**Figure 2b:** Experimental in vitro data and optimally fitted ODE model. Dots represent the time course of ligand-induced G-protein subunit dissociation measured by FRET in HEK293 cells. FRET values were transformed into concentration of undissociated G-proteins by a scaling factor. Lines indicate the best-fit of the ODE model to the data



(using optimal parameters).

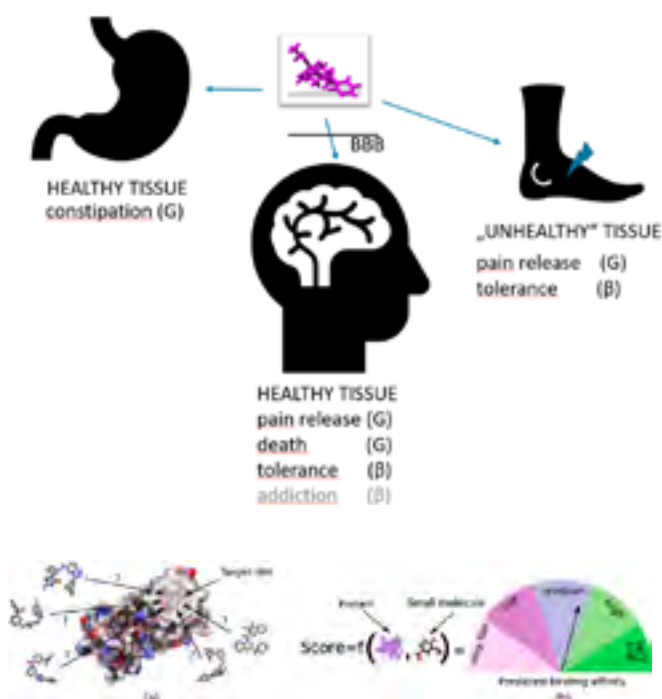


## DRUG DISCOVERY – STRIVING FOR MINIMAL SIDE EFFECTS

Throughout the last decade, ZIB has built up a deep understanding about cellular and subcellular structures, e.g., the distribution of different proteins on cellular surfaces. Even at the smallest length scales, at the level of molecular-protein interactions, we can derive structural information from experiments. On the basis of this knowledge, how proteins “look like”, new computational methods are developed which can figure out which kind of small drug-like molecules are able to interact with these proteins in a specific, desired way (*VirtualFlow*, [8]). The giant step forward during the last two years is given by the possibility to really find new drugs, which have the potential to reduce side effects, by applying *VirtualFlow*, (see Fig. 4).

*VirtualFlow* can search in a set of many million compounds for their interaction potential with given proteins. The interaction potential is mainly due to a high binding rate of the compound. *VirtualFlow* is one key to finding drugs. The other key is given by the idea of taking the chemical environment of the receptors into account, too [7]. In Figure 3 it is shown, what kind of drug design principles can be used in order to select drugs which allow for pain-relief without side-effects.

Simulating the molecular binding process for possible drug candidates includes many different time scales. The mathematical challenge when dealing with multiple timescales is to combine different mathematical models on different levels. The quivering of small molecules in the presence of proteins, their path into the protein’s active site in a stochastic back and forth motion, results in a quantity called the binding rate. This is how macro variables (binding rate) arise from the analysis of microsystems (molecular movement). This transition from microsystems to macro variables is from a mathematical point of view a projection, it has a projection error. This results in macrosystems having a kind of memory (called “rebinding effect”) and which are, thus, not Markov processes. As a consequence, using Markov models (e.g. stochastic differential equations) for modeling macrosystems without regarding the microsystem anymore, suffers from this projection error. Every projection to coarser scales amplifies this error. In the last year we developed a mathematical framework which accounts for a reduction of this error [9, 12]. We developed a projection method that takes into account information from different time scales [10] and a theoretical framework to analyse the effect of changing rates with regard to even coarser time and length scales [11].



**Figure 3:** A pain-relief drug (purple) can enter different parts of the body. There exist drugs which do not cross the blood brain barrier (BBB) and can not reach the brain. In different compartments of the body, the drug molecule generates different, listed effects. There exist drugs which only activate the G proteins and not the associated pathways. Furthermore, ZIB has designed NFEPF which only activates processes in inflamed (“unhealthy”) tissue. The “best” drug, however, has not been found yet: It would only activate G proteins and this only in inflamed tissue, because pain-relief is the wanted effect.

**Figure 4:** (a) With the aid of *VirtualFlow* it is possible to search for possible drug candidates in a huge space of chemical compounds. (b) The predicted binding affinity of drug molecules depends on the target protein (and its chemical environment) as well as on the selected small molecule (and its protonation state).

# MULTIENERGY EUROPEAN ENERGY



# MODELS FOR SYSTEM PLANNING

Decarbonization of electricity alone is not sufficient to meet CO<sub>2</sub> reduction targets! Sector coupling is essential for an effective future low-carbon energy system.





Figure 1. The plan4res Consortium

In the Paris Agreement<sup>1</sup> on Climate Change, the EU promised to reduce emissions and reach climate neutrality by 2050 while maintaining sustainability, security of supply, and competitiveness among the suppliers. The respective EU's 2030 and 2050 targets imply a steep increase in renewable energy sources (RES). This transition will induce higher volatility of the energy systems. To face this challenge, balancing between regions has to be strengthened, requiring an integrated pan-European system that promotes strong cooperation and alignment between the individual countries.

ZIB took part in the efforts of reaching the Paris Agreement targets with a consortium from France, Italy, the United Kingdom, Switzerland, and Germany (see Fig. 1) in the Horizon 2020 project plan4res<sup>2</sup>. The project's ambitious aim was to develop an end-to-end planning tool to successfully increase the share of renewable energy in the European Energy system without compromising system reliability.

1 <https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement>

2 Plan4res project web site: [www.plan4res.eu](http://www.plan4res.eu)



Prof. Dr. Thorsten Koch  
Phone +49 30 841 85-213  
[koch@zib.de](mailto:koch@zib.de)



Dr. Inci Yüksel-Ergün  
Phone +49 30 841 85-449  
[yueksel-erguen@zib.de](mailto:yueksel-erguen@zib.de)

# plan4res: Supporting the European Energy Transition with an Integrated Holistic Energy System Planning Framework

A massive transition of the pan-European energy system towards an increase in the share of RES will be required to meet EU's 2030 and 2050 targets. This transition will affect the energy production and transportation systems and all energy-consuming sectors. To achieve this transition involving the lowest cost, an integrated representation of the energy system is necessary to support the simulation of its expansion and operation. Such a representation involves diverse but highly interconnected models that work synergistically. It needs to retain modularity to tailor the tool for the needs of various stakeholders. Considering the increasing complexity of the energy models, developing and implementing this holistic modeling framework is technically challenging. The project plan4res addressed this challenge and aimed to fill the gaps between the available system analysis tools and the increasing com-

plexity of the future energy system planning and operational problems.

After more than three years of joint work of research and industrial partners, the plan4res consortium presents a modeling framework enabling the users to select from available models appropriate to their questions and access methods and algorithms to solve these models. In addition, the plan4res data library, data transformation methods, and formatting tools give a head start to the users generating their problem instances. Furthermore, a containerized IT platform facilitates energy system optimization and planning problems by providing an entire runtime environment. Finally, the system has demonstrated its validity in three case studies covering the diverse requirements of different stakeholders.

## Models for Holistic Energy System Planning

The *plan4res Modeling Framework* consists of interconnected layers of optimization models to enable integrated pan-European energy system modeling [1]. These layers impose various views on the energy transition problem to let users tailor the models to the requirements of different stakeholders. For example, one layer deals with investment decisions like grid expansion and new facility installation to support the optimal future energy mix. To account for the operational feasibility of these decisions, models in the next layer evaluate the results from the investment models in the operational planning context. Finally, the last layer consists of models for the electricity transmission grid and gas transport analysis with a high temporal and spatial resolution.

Researchers at ZIB implemented the *gas network optimization model (GNO)* [2]. It is a physical flow-based stationary model enabling us to represent the gas network by a bi-directional graph, and evaluate a *scenario*, the amount of gas flowing into/out of the gas network from entry/exit nodes, subject to the network. We model the gas flow ac-

ording to non-linear equations of the thermodynamic laws, with gas pressure decreasing as it flows through pipelines. Binary decision variables of the GNO notate states of the active components used to regulate the flow through the network. Thus, the GNO represents a mixed-integer non-linear program (MINLP). We validate a scenario with the GNO by investigating whether the inputs at the entries and exits of the network are technically feasible. However, extending the problem we have previously tackled for one of the most extensive German gas transport systems [2] to a pan-European scale was highly challenging. As seen in several similar activities, the main concern has been the scarcely available pan-European gas transport network data [3], [4]. Besides, the GNO had to be spatiotemporally aligned to the electricity transmission grid operation model (TGOM). We developed and implemented an analysis method consisting of preprocessing publicly available gas network data and capacitated network flow-based models for generating gas in-/out-flow scenarios for a restricted region in Europe to overcome these challenges.

# Open Data Enabling Integrated Energy System Planning

Reusability and public availability of the data in energy systems-related projects are essential to enable benchmark studies with the different models and tools developed for decision support. As participants of the plan4res, we acknowledged this requirement and took special care to document our data collection, transformation, and format-

ting [4]. We prepared a publically available data library of non-confidential data used by the case studies in a format ready to be used with the plan4res tool. Together with the library, our data-related documentation is published in the plan4res public data repository at Zenodo<sup>3</sup>.

## New Methods and Algorithms to Solve Huge Energy System Optimization Instances

The plan4res modeling framework involves a high level of detail regarding the high integration level of different technologies such as gas, electricity, transport, and heating. Besides, technology, market, and climate conditions impose a high degree of uncertainty. A holistic assessment of these models at the pan-European scale results in complex and large-sized problem instances. The state-of-the-art solution methods and algorithms were not capable of analyzing the energy systems with the aimed scale and detail level in plan4res. Hence, under the lead of ZIB, the consortium partners spent substantial efforts to enhance the state-of-the-art by:

- implementation of advanced decomposition algorithms to require less problem-specific tailoring and coding with the facilitation of SMS++ modeling capabilities;
- addressing non-differentiable optimization solvers by the NDOSolver/FiOracle project and integrating it with the SMS++ and releasing it as an open-source library during the project<sup>4</sup>;
- integration of the StOpt<sup>5</sup>, the stochastic optimization library for solving large-scale seasonal storage problems [5] interfaced in the SMS++ modeling library to facilitate its use for addressing uncertainty in the models;

- enhancing the SCIP solver as a readily available out-of-the-box tool for all LPs, and MILPs solved in the project focusing on the specific structure of energy system models to exploit time- and spatial-indexed decision variables more efficiently (released with the SCIP Optimization Suite v7.0 and v6.0);
- extending SCIP to exploit modern shared-memory architectures through parallelization and development of a parallel presolving library (The Parallel Presolving Library, PaPILO v1.0, was implemented and made available online with SCIP Optimization Suite v7.0 as a new addition<sup>6</sup>).

Finally, specific adaptations to fine-tune the methods and algorithms for the particular purpose of the plan4res models were demonstrated in the case studies. The selected solution methods were directly linked to implementation, e.g., via the plan4res IT platform.

<sup>3</sup> <https://zenodo.org/communities/plan4res>

<sup>4</sup> [https://gitlab.com/frangio68/ndosolver\\_fioracle\\_project](https://gitlab.com/frangio68/ndosolver_fioracle_project)

<sup>5</sup> <https://gitlab.com/stochastic-control/StOpt>

<sup>6</sup> <https://www.scipopt.org/>

# Plan4res Implementation and IT Platform

The plan4res IT platform<sup>7</sup> and tools for facilitating implementation are among the most important products of plan4res.

SMS++, a modeling suite to enable the solution of big-size complex energy problems based on decomposition methods, was publicly released for the first time in the project<sup>8</sup>. SMS++ allows users to develop a set of solution algorithms that can tackle very large-scale optimization problems under uncertainty. It was also integrated with SCIP during the project.

HPE developed the plan4res IT platform, a software container environment covering the entire lifecycle and all user contact points appearing during the project. Developers can use the platform to simultaneously build and deploy their energy modeling tools on different computational environments such as a laptop or a high-performance computing system. With its connection to the data store used by all project members, the container provides seamless access for different project members, even if they use different computational environments.

## MAKE IT WORK – CASE STUDIES

In three case studies, we have successfully demonstrated the plan4res tool's ability to model different viewpoints on the integrated energy system and solved the diverse resulting problem instances [6].

Case Study 1 modeled the transition of the multimodal pan-European energy system, including an integrated analysis of electricity and gas transport with perfect foresight, considering sector coupling of electricity, heat & cold, traffic, fuel/gas; and coupling to gas grids [7].

Case Study 2 focused on the strategic development of the pan-European transmission network without perfect foresight and considering long-term uncertainties in a transmission grid expansion model for the European Electricity System, through 2020 to 2060, with high shares of RES.

Case Study 3 assessed the feasibility of scenarios, cost of RES integration, and value of flexibilities for the European electricity system without perfect foresight [8]. In this case study, the value of different flexibility services was studied to analyze the amount of system cost reduction achieved by using the flexibilities.

<sup>7</sup> <https://gitlab.com/cerl/plan4res/p4r-env>

<sup>8</sup> <https://gitlab.com/smspp/smspp-project>



# Highlights from Case Study 1: An Integrated Analysis of European High-Pressure Gas Transport Network and Electricity Transmission Grid

Natural gas in the European Energy System: The gas network is a substantial enabler of the European energy system considering the amount of energy stored and transported in the gas network. Twice the amount of energy transported in the whole electricity grid in Germany<sup>9</sup> is transported in the gas grid. Its use is diverse – ranging from an energy source for heating and household usage over industrial use to electricity production.

EU's natural gas supplies are constituted mainly by imports from countries such as Russia and Norway, and LNG imports.

As the network is highly interconnected, the security of supply at one region in the European continent cannot be considered exempt from the gas flow in the rest of the network.

Natural gas as a flexibility option: With the increasing share of RES in electricity production, gas-powered plants act as a buffer when RES cannot supply enough energy because of insufficient wind or sunlight. A gas network also acts as a flexibility to the energy system during the energy transition by storing gas. For example, the gas network can allow transportation and storage of hydrogen or synthetic gases produced by power-to-gas (P2G) technology from surplus electricity. Gas can also be compressed by electric-driven gas network components and stored in the gas network [9].

With our partners Siemens and RWTH Aachen University, we at ZIB addressed the multimodal energy system at the pan-European level to analyze the optimal energy mix (see Fig.2). The first step was to define a transition pathway through 2020 to 2050, which considers costs, revenue streams, CO<sub>2</sub> targets and sector coupling, by projecting the energy mix, investment, and retirements along the pathway. The second step was a detailed bottom-up analysis of the year 2040 including electricity market analysis and TGOM. Finally, the resulting hourly gas supply/demand schedules were evaluated using the GNO and analysis methods at the pan-European scale.

As gas presents an important flexibility option of the energy system, whether it is feasible to transport the demanded gas within the gas transport system has an

ever-increasing value in the energy transition planning for mid-to-long-term decisions. We aimed to ask the question from the security of supply, existing capacities/infrastructures, and newly established services perspectives in Case Study 1.

The first challenge of the study was to align the GNO to the TGOM for interconnected operations. As observed from the substantially limited scales of the state-of-the-art academic studies [9], such an analysis is demanding even in the small-size test problems. Yet, in plan4res, we aimed to perform a pan-European scale analysis. We succeeded in the spatio-temporal integration of the two grid models at the data exchange level. This enabled us to provide feedback to upper stream models to improve the resulting gas schedules.

<sup>9</sup> <https://www.bmwi.de/Redaktion/EN/Artikel/Energy/gas-power-to-gas.html>

The scarce availability of the pan-European gas network data constituted the second significant challenge. Available data from public resources were insufficient to model the pan-European gas network as an input to the GNO [3]. Otherwise, we could have performed the gas network analysis at once by formatting the available data of the whole European gas network and running the GNO with it to evaluate the scenarios. Instead, using our gas network modeling know-how, we developed a novel methodology to extend it to a pan-European scale and gas network data preprocessing consistent with the real world. First, we as-

sembled a feasible data set from publicly available data sources for demonstration purposes consisting of:

- (i) a data set for pan-European gas supply and demand constructed from open historical flow data and forecasts [4], and
- (ii) an exemplary data set constructed for the German high-pressure gas transport network by improving another readily available data set and augmenting it with compressor and gas properties data [3].

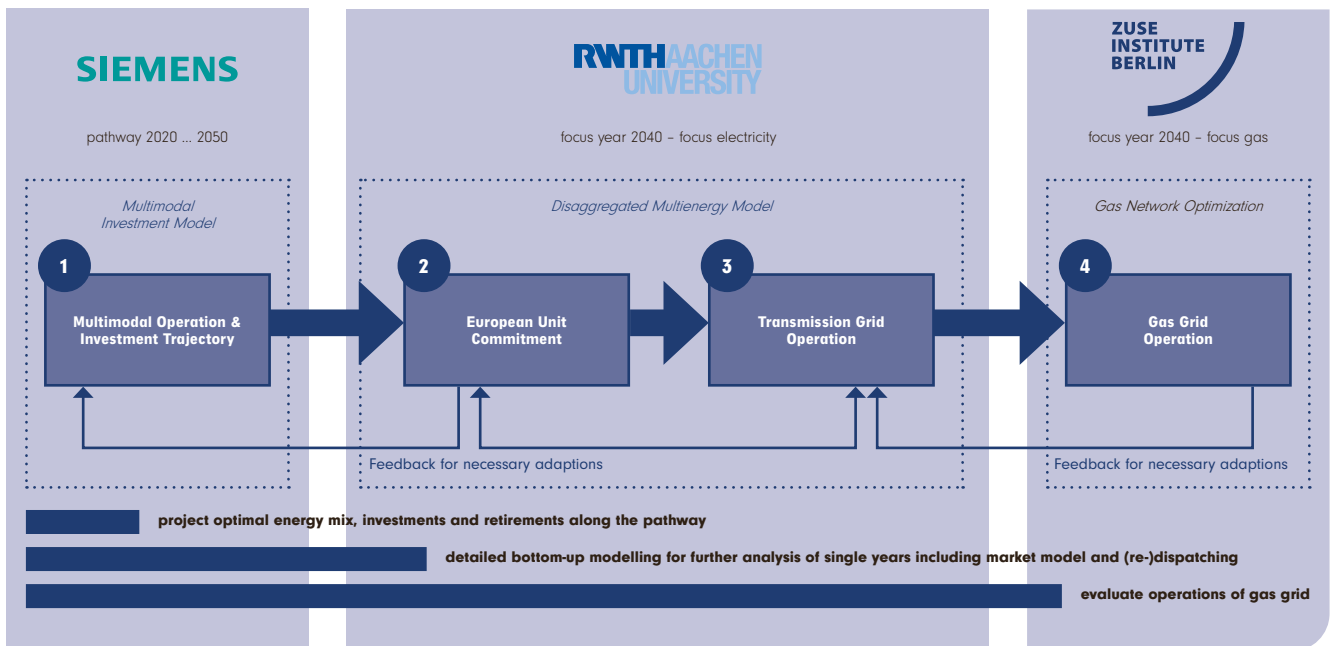


Figure 2. Case Study 1

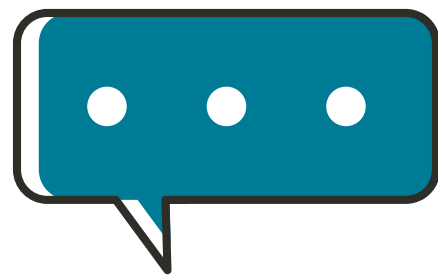
Next, we generated practically relevant gas in-/out-flow scenarios for Germany given the European high-level gas network topology, TGOM results, pan-European gas demand, and supply forecasts. We evaluated those scenarios by the GNO for the German gas network. Since Germany is in the center of Europe connected to several other countries, this is used as a feasibility analysis of the pan-European cross-border energy exchange. This method integrates electricity and gas transport considering pan-European gas import and forecasts and demand and generation (P2G) results from TGOM with high spatial resolution.

While coping with the well-known data (integration-)problems, the gas data preprocessing studies in plan4res resulted in new methods that we successfully implemented, such as

- (i) the sequential methodology that uses mathematical optimization model information to improve the network data and, hence, the analysis results, and
- (ii) a method to estimate compressor and driver data based on partially available public data and network topology using our gas network modeling know-how.

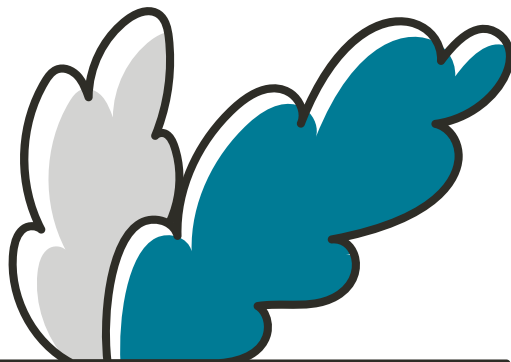
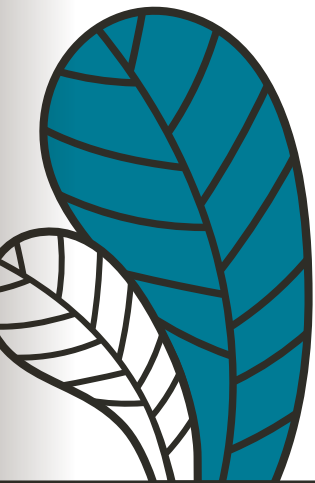
As the list of questions on how to plan the energy transition is not getting any shorter, the plan4res framework and tools will help answer some of them. Still, publically available data from transition system operators remains scarce and mostly limited to what transparency regulations require operators to share. A pan-European gas infrastructure data set, suitable to be used for physical-flow analyses, could provide the means to answer even more questions.

# MULTIOBJECTIVE SHORTEST PATHS



"It is difficult to trace back the history of the shortest path problem. One can imagine that even in very primitive (even animal) societies, finding short paths (for instance, to food) is essential. Compared with other combinatorial optimization problems, like shortest spanning tree, assignment and transportation, the mathematical research in the shortest path problem started relatively late."

[https://www.math.uni-bielefeld.de/documenta/vol-ismmp/32\\_schrijver-alexander-sp.pdf](https://www.math.uni-bielefeld.de/documenta/vol-ismmp/32_schrijver-alexander-sp.pdf)

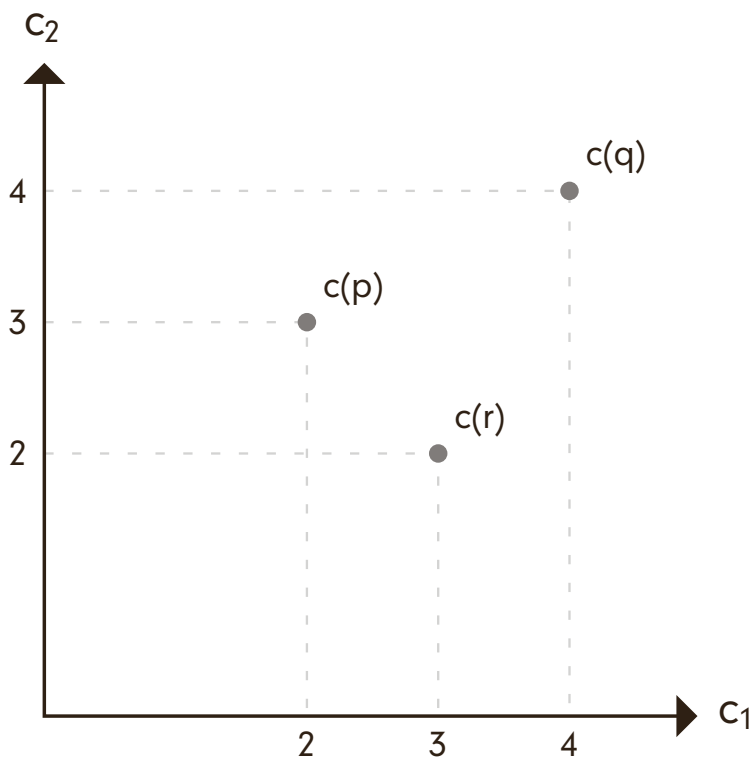


Even if mathematical research in the field started late, it has been a big success story in the last 60 years. The perceptibility of solution strategies for shortest path problems and their applicability in practice constitute a loop in which researchers keep adding features (resource constraints, time dependency, etc.) to the original shortest path problem while still controlling the theoretical problem complexity and computational running times. Each cycle in this loop brings shortest path problems closer to the real-world demands of industry or everyday users. One of these problem variants is the Multiobjective Shortest Path (MOSP) problem.

In the MOSP problem, we seek to find paths that are (Pareto) optimal with respect to multiple optimization criteria that are equally important and concurrent. If we consider two paths, that cost \$5 and \$6, respectively, it is clear that the first one is preferable if the minimization of costs is our only objective. However, if we consider not only the paths' price but also their duration as an optimization criterion,

things become more difficult. Assume that it takes 7 minutes to traverse the first path and only 4 minutes to traverse the second one. Then, the paths' cost vectors are (\$5, 7 minutes) and (\$6, 4 minutes) and it is not possible to decide which one is preferable since the first one is cheaper and the second one is faster. If these two paths were the only possible choice, they would both be considered to be optimal and thus the output of any MOSP algorithm. This output is often called a Pareto front of optimal or efficient paths.

For a company, a path's overall costs usually reflect the sum of multiple cost drivers defined arc-wise on the input search graph. If the different costs of each arc are summed up, the company ends up with a classical shortest path instance. Algorithms for the problem will output a single path with optimal overall costs and will thus deliver little flexibility in further planning stages. In contrast, refusing to sum up the arcs' costs leads to a MOSP instance. The solution set will contain an optimal single criteria path and further interesting solutions for the company. The Network



**Figure 1:** Paths  $p$ ,  $q$ , and  $r$  and their two-dimensional cost vectors. In multiobjective optimization,  $p$  and  $r$  are better than  $q$ . However, since both cost dimensions are equally relevant,  $p$  and  $r$  are both optimal paths.

Optimization department at ZIB has been successfully cooperating with Lufthansa Systems GmbH & Co. KG for many years, a market leader in flight-planning systems. Two of Lufthansa Systems' main cost drivers when planning aircraft trajectories are the flight duration and the fuel consumption. Often, an optimal solution for a single criterion scenario, using a weighted sum of time and fuel costs, is enough, but whenever infrastructures become overstrained and big delays become a threat, a Pareto front of optimal paths gives the possibility to choose a faster route even if more fuel is consumed. A solution to arrive "just in time" is probably in the Pareto front.

In general, multiobjective optimization solutions are mapped to multidimensional cost vectors that are then used to describe optimal sets of solutions. Consider two paths  $p$  and  $q$  connecting the same nodes in a graph.  $p$  is said to dominate  $q$  if  $p$ 's cost vector is at least as good as  $q$ 's cost vector in every dimension and there is at least one dimension in which  $p$ 's cost vector is better than  $q$ 's. Fig. 1

shows an example involving paths with two-dimensional cost vectors: the cost vector of  $p$  is better than  $q$ 's in every dimension. However, it is not possible to decide whether  $p$  dominates the path  $r$  or  $r$  dominates the path  $p$ . A path is said to be optimal or efficient if there is no other path that dominates it and connects the same end nodes.

An optimization problem is said to be intractable, if there is no algorithm that can solve all problem instances in polynomial time with respect to the size of the input. Even the easiest combinatorial optimization problems become intractable when their multiobjective siblings are considered. The reason is that even when two objectives are considered simultaneously, the number of optimal solutions can be exponential with respect to the input size. In the example in Fig. 1, it is easy to see that optimal paths can be added besides  $p$  and  $r$ , hence increasing the cardinality of a solution set. Clearly, if an algorithm is expected to output exponentially many solutions, it cannot terminate after a polynomial number of steps.



Pedro Maristany de las Casas  
Phone +49 30 841 85-232  
maristany@zib.de



Ricardo Euler  
Phone +49 30 841 85-247  
euler@zib.de



Prof. Dr. Ralf Borndörfer  
Phone +49 30 841 85-243  
borndoerfer@zib.de

We have discussed the practical advantages and the theoretical hurdles of the Multiobjective Shortest Path (MOSP) problem. Conventional wisdom in the operations-research community states that the former outweigh the latter. In other words: Even if the MOSP problem is hard in theory, it can be solved efficiently in practice. For this to hold and to scale properly as applications become bigger and more complex, the design of efficient MOSP algorithms is crucial.

This search for better algorithms can lead in two directions. First, the development of more efficient algorithms for general MOSP instances that can be used out of the box for a wide variety of problem variants. However, for many variants, especially those, that consider a large number of optimization criteria, the curse of dimensionality (more efficient solutions as the number of uncorrelated objectives increases) drives up computational complexity rapidly and specific problem structure must be exploited to achieve acceptable performance. This mandates the search for problem-specific algorithms.

## MULTIOBJECTIVE DIJKSTRA ALGORITHM

The input of a general MOSP instance consists of a directed graph with multidimensional costs on its arcs and a source node from which all optimal paths are meant to start. In this scenario, all nodes other than the source node are regarded as targets. The output of any MOSP algorithm is supposed to be a Pareto front of optimal paths from the source node to every target node. These sets are not unique because the graph possibly contains two paths with equivalent costs connecting the same nodes. For every optimal cost vector, the output is expected to contain

In the Network Optimization department at ZIB, constant progress is made in both directions. In the following, we showcase two examples of this: the novel Multiobjective Dijkstra Algorithm (MDA) for general MOSP problems [1,2,3] and a graph- and monoid-based approach for price-sensitive routing in public transportation [4]. The MDA has been developed in cooperation with Lufthansa Systems in our joint effort to build a new generation of their core-optimizer VOLAR for flight planning. This research is supported by the German Federal Ministry of Education and Research (BMBF) within the context of the MobilityLab of the Research Campus (MODAL), and also by the Berlin Mathematics Excellence Cluster Math+. Our second example, the price-sensitive routing algorithm was developed as a demonstrator in the collaborative research project Mobility Inside that was funded by the German Federal Ministry for Digital and Transport (BMVI).

one path with these costs. Until the output is generated, paths can be separated into three categories: unexplored, explored, and processed. Unexplored paths are not yet seen by the algorithm. Processed paths are paths for which the algorithm could already decide whether they are part of the output or not. Big algorithmic differences arise in the way explored paths are handled. These are the paths that the algorithm has already discovered without finally knowing if they can be discarded or not.



Classical (label-setting) algorithms for MOSP follow a very straightforward policy for the storage of explored paths: every discovered path is inserted into a single priority queue and remains therein until the algorithm is able to decide whether the path becomes part of the output or is finally discarded. We have already discussed that the output of a MOSP instance can contain exponentially many optimal paths and clearly the output is a subset of the paths that are explored throughout a MOSP algorithm. Thus, a priority queue containing all explored paths can become huge. Moreover, priority queues are sorted data structures. Keeping a container sorted when it is full consumes a lot of resources/time.

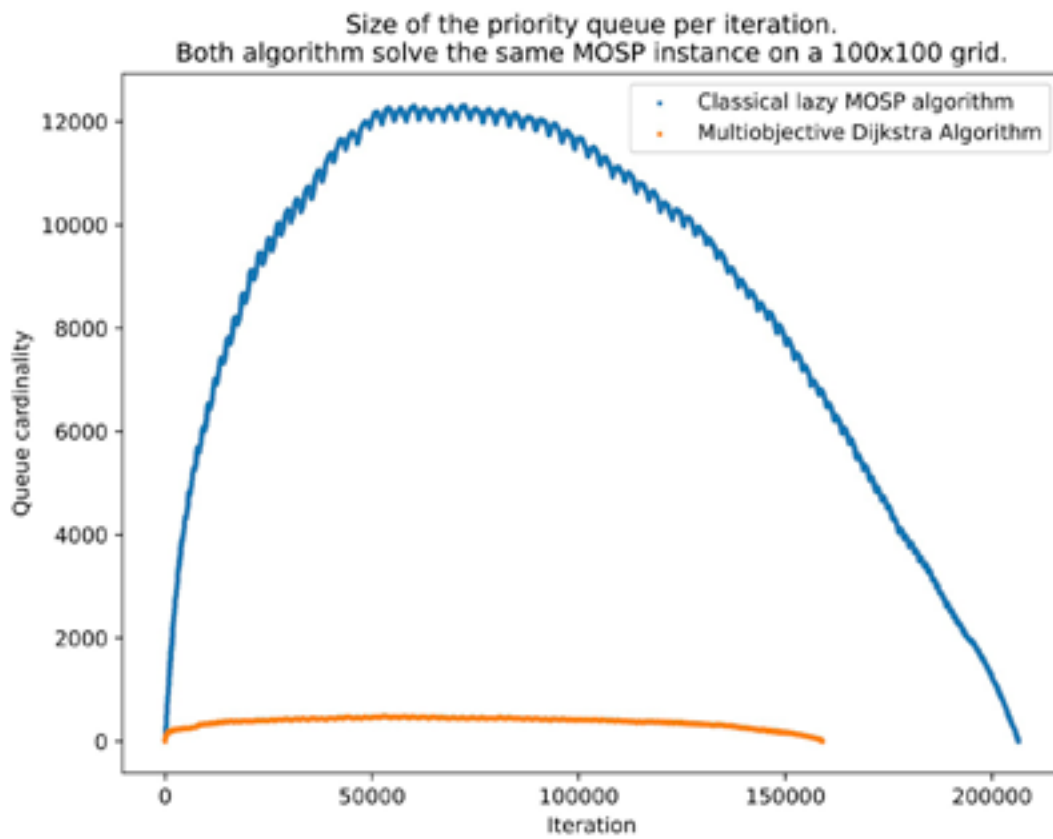
At ZIB we developed the Multiobjective Dijkstra Algorithm (MDA), a novel algorithm for MOSP problems that exploits structural properties to make the handling of explored paths more efficient in theory and practice [1,2,3]. Optimal MOSP solutions fulfill the so-called Bellman Equation. This property says that optimal solutions are built from optimal pieces. In our context, this means that an optimal path consists of optimal subpaths. Let  $v$  be a node and assume that the MDA just recognized an  $s$ - $v$ -path  $p$  to be optimal and part of the output. Then,  $p$  is extended along the outgoing arcs of  $v$  and explored paths from the source node  $s$  to the successor nodes of  $v$  are created. These paths'  $s$ - $v$  subpath – it's the path  $p$  – is optimal. Sometimes new explored paths obtained this way can directly be discarded. Classical MOSP algorithms also recognize this situation. The novelty of the MDA is how the newly explored paths are stored if a decision upon their relevance cannot be

made. Let  $w$  be a successor node of  $v$  and  $q$  the  $s$ - $w$  path via  $v$  that the MDA just explored. If the priority queue of the MDA does not contain any path ending at  $w$ ,  $q$  is inserted into it. However, if there is an  $s$ - $w$  path, say  $r$ , in the priority queue already, the most important invariant of the MDA is enforced: at any point in time and for every node  $w$  in the graph, there is at most one  $s$ - $w$  path in the priority queue. Which one to keep,  $q$  or  $r$ ? The answer depends on the lexicographic ordering of the paths' cost vectors: the priority queue must contain the path with the lexicographically smaller costs. If we assume that  $r$  is the path kept in the priority queue, we have to ask ourselves how the MDA ensures that  $q$  is reconsidered later. Recall that the  $s$ - $v$  subpath  $p$  of  $q$  is optimal and part of the output of the MDA. Thus, it has to be stored anyway for the duration of the algorithm. This means that the path  $q$  can be easily rebuilt by just repeating the extension of  $p$  along the arc  $(v, w)$ . At some point during the algorithm's execution, the path  $r$  or another  $s$ - $w$  path is extracted from the priority queue that is synonymous to label the path as processed in the MDA. Then, the queue is again allowed to contain a new  $s$ - $w$  path. This new path is built out of the processed paths ending at the predecessor nodes of  $w$ . The node  $v$  is among these nodes and the path  $p$  is a processed path ending at  $v$ . Thus, the path  $q$ , which is the extension of  $p$  along the arc connecting  $v$  and  $w$  is reconsidered and might be inserted into the priority queue.

Details apart, the key idea is that explored paths can be rebuilt as often as needed out of processed paths that have to be stored anyway since they are part of the final

output of the algorithm. This implies that throughout the algorithm, the size of the priority queue remains polynomially bounded by the number of nodes in the graph. Hence, the effort required to keep the queue sorted is smaller than in

classical algorithms for MOSP. Fig. 2 compares the size of the priority queues of the MDA and of a classical algorithm for MOSP while solving the same MOSP instance.



**Figure 2:** Size of the priority queue of the MDA and a classical algorithm for MOSP. Both algorithms solved the same MOSP instance.

We denote the number of nodes in the graph by  $n$ , the number of arcs by  $m$ , and the maximum cardinality of a Pareto front of optimal paths between the source node and a target node by  $k$ . The running time bound for the MDA is expressed in terms of the input size and the output size of any MOSP instance and improves the running time bounds derived so far for MOSP algorithms. In case only two objectives are considered simultaneously, the complexity drops to which is the complexity of Dijkstra's classical algorithm for shortest path problems times the maximum number of solutions to a single node. In other words: the intractability of MOSP problems forces any MOSP algorithm to possibly

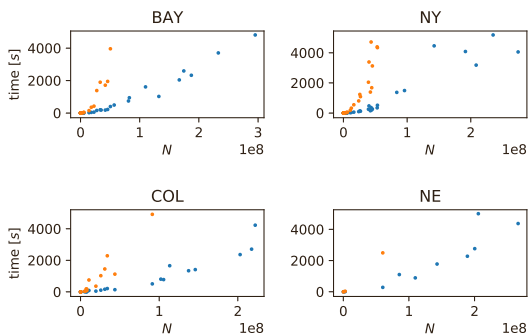
output exponentially many paths and in the biobjective case, the MDA does so by computing each solution as fast as Dijkstra's algorithm does. This is the motivation for the name Multiobjective Dijkstra Algorithm.

The results of experimental results obtained from the comparison of the MDA with the state-of-the-art algorithm for MOSP by Martins on different types of practically relevant graphs are summarized in Fig. 3, Fig. 4, and Fig. 5. We observe that the MDA outperforms Martins' algorithm in every scenario.

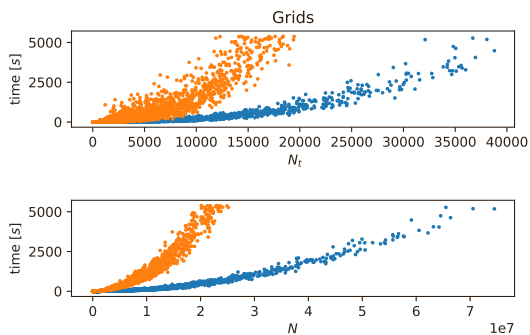
### PRICE-SENSITIVE ROUTING IN PUBLIC TRANSPORTATION

A traveler on public transportation usually aims to take the shortest route with the least number of transfers. However, additional criteria also play a role. For this problem, extremely fast specialized algorithms are available (such as RAPTOR). Arguably, the most important objective of a traveler next to travel time is the price of their journey and this objective is special. Namely, optimization criteria in MOSP can usually be represented as arc weights that are rational numbers of travel time functions that fulfill the FIFO (first-in first-out) property. Such arc weights then are simply summed up when calculating the cost of a route.

This is not possible, however, for public transit fares. The problem is that they are often calculated in an intricate fashion involving several parameters. Prices can depend on the set of visited fare zones, the traveled distance, and the number of transfers and stops. Some modes of transportation might incur surcharges that are only to be paid once even if they are used multiple times on a journey. It turns out that all these "surrogates" of factors can be handled in a mathematical structure called a positive, partially ordered monoid. Optimization over such a monoid can be done using a general MOSP algorithm. The actual prices can be calculated from the surrogates in a post-processing step under mild technical assumptions that are fulfilled in practical settings. The result contains a Pareto front of optimal routes but most likely also several superfluous routes.



**Figure 3:** MDA (blue) vs. Martins' (orange) algorithm on road networks with three-dimensional arc costs.



**Figure 4:** MDA (blue) vs. Martins' (orange) algorithm on grid graphs with three-dimensional arc costs.

However, there are some caveats: optimizing only for price can mean optimizing for three or more implicit criteria and many of the resulting routes will in the end not be optimal in price. Think of a fare system that bases prices on distances for short trips but for longer journeys on fare zones. Then, depending on the distance between origin and destination, a route of shortest distance is not necessarily a useful result. This outcome is caused by the fact that the dimensions of the monoid (i.e., the different optimization criteria) are interdependent. However, interdependences also have their pros – in fact, they can be exploited in the following way.

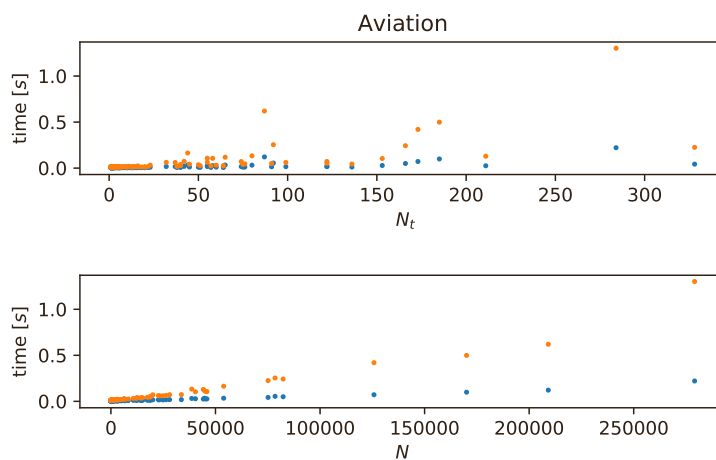
For every public transit journey, a ticket has to be bought. If we prolong a journey, this ticket might lose its validity and must be replaced by a more expensive ticket. We say the ticket transitions into another one. We model these transition relations in a ticket graph where nodes represent tickets and arcs represent ticket transitions. Each transition arc is annotated with a transition function that indicates in which state of the monoid the transition will be made. An example of such a ticket graph can be seen in Fig. 5.

We can identify every path in the routing graph with a path in the ticket graph giving its ticket and price. Path relations in the ticket path can then be used to determine dominance relations between tickets. For example, if we compare two paths with the same ticket without outgoing edges, we immediately know that it suffices to keep

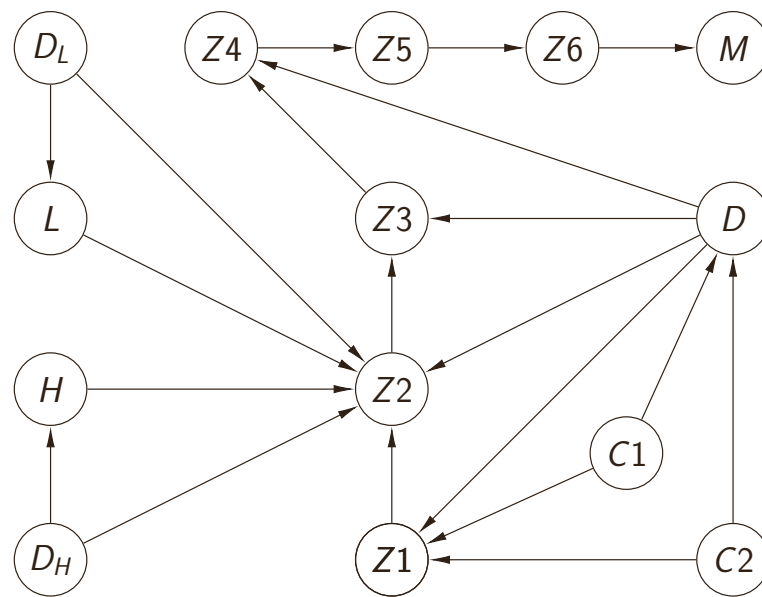
only one of them in the solution set. In general, the ticket graph can be used to infer domination relations between tickets that were previously incomparable. Additionally, price-based pruning schemes at the target node can be established [4].

Mitteldeutscher Verkehrsverbund (MDV), a public transport association in eastern Germany, services a network consisting of 5,576 routes and 4,371 stops. Ticket prices depend on a set of 56 fare zones, distance traveled in kilometers, the number of visited stops, and the number of transfers (see Fig. 5 for an illustration of the ticket graph). A naive MOSP approach optimizing only the number of transfers, the arrival time, and the number of visited fare zones took an average 4.6 seconds per query and found 47 routes out of which most are not optimal in price. The ticket graph approach, however, found the Pareto set in only 243 ms: a reduction of 95%.

Since travel time and price of a journey are usually correlated, our method can be combined with the bounded multiobjective RAPTOR algorithm recently introduced by Delling et al. (2019). Here, only routes whose travel time and transfers differ by at most a fixed distance from the Pareto front are taken into account. Even when using a large distance of 60 minutes and two transfers, computation times can be reduced to less than 30 ms while decreasing the size of the Pareto set from 2.62 to 1.83 on average.



**Figure 5:** MDA (blue) vs. Martins' (orange) algorithm on airway networks with two-dimensional costs.

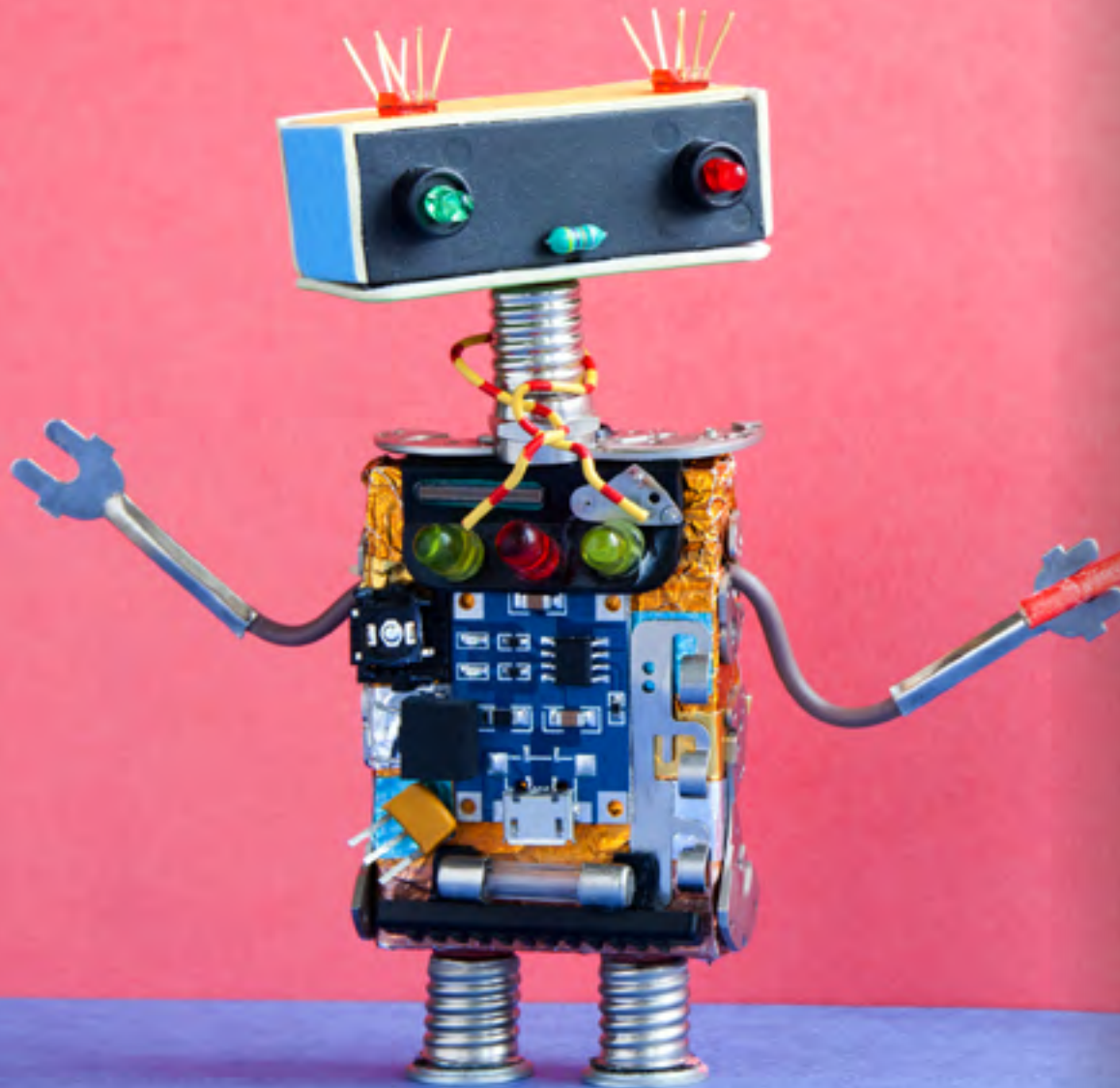


**Figure 6:** The ticket graph for the fare system of MDV (Saxony). There are seven tickets for zone-based fares, as well as fares for the cities of Halle (H) and Leipzig (L) and several smaller cities (C1, C2). Additionally, there are tickets with special discounts for short travel distances (D, DL, DK) and a maximal price (M). Arcs represent possible ticket transitions.

Multiobjective shortest path algorithms are currently becoming algorithmically feasible for a variety of important planning problems, in particular, but not limited to, traffic and transportation. The ability to choose from several Pareto-optimal solutions, balancing the trade-offs between competing objectives, provides managerial and operational degrees of freedom that translate directly into better decisions. Knowledge of the Pareto front yields insight and can be used to identify options for compromise as well as bottlenecks and opportunities for improvements. Of course,

there is a possibly very large number of criteria to consider. But, to close with a citation by Vilfredo Pareto, the inventor of multiobjective optimization, "For many events, roughly 80% of the effects come from 20% of the causes. [...] If you're Noah, and your ark is about to sink, look for the elephants first, because you can throw over a bunch of cats, dogs, squirrels, and everything else that is just a small animal and your ark will keep sinking. But if you can find one elephant to get overboard, you're in much better shape."

# EXPLAINABILITY AND INTERPRETABILITY IN DEEP LEARNING



The safe deployment of AI systems in high-stakes applications such as autonomous driving, medical imaging, or criminal justice requires that their decisions can be scrutinized by humans.

Traditional machine-learning models such as k-nearest neighbors, decision trees, or sparse linear models allow for a straightforward understanding of the reasoning process. However, state-of-the-art accuracy is generally achieved by training deep neural networks that are regarded as black-box models.

This has spawned the field of explainable AI (XAI), which aims to make modern machine-learning (ML) models more transparent.



A first step toward explaining the reasoning of a classifier is to decide which input variables were important for the decision and which were not. So-called salience attribution methods rate the input variables for their importance in an individual decision. A large number of heuristics have been proposed that visually display their attributions as a heat map, called a salience map. To measure the efficacy of these methods, there are two approaches:

1. Human-based evaluation: humans rate the saliency explanations directly, rating how useful they find the explanation.
2. Evaluation by proxy task: the explanation is used for a quantitative proxy task whose score is correlated with the quality of the salience map.

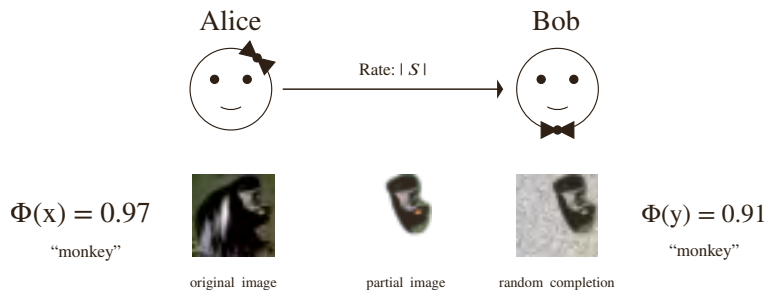


Stephan Wäldchen  
 Phone +49 30 84185-221  
 waeldchen@zib.de



Dr. Mathieu Besançon  
 Phone +49 30 841 85-221  
 besancon@zib.de





**Figure 1:** Probabilistic prime implicants as a communication problem. Alice and Bob share a classifier function  $\Phi$ . Alice aims to convince Bob that her image belongs to the "monkey" class while transmitting the fewest pixels as possible. Bob completes the image with noise to be able to use  $\Phi$ . There is a natural trade-off between the number of pixels (rate) and the difference in the classifier score (distortion).

## Rate-Distortion Framework

One of the most used proxy tasks is input obfuscation, described in [1]. The idea is that leaving out unimportant features from the input does not strongly influence the classifier decision while removing the important features does. Removing a feature is realized in this case by replacing it with a random variable from a baseline distribution. This can be seen as a rate-distortion problem, where the rate is the number of included features and the distortion is the expected deviation of the original classifier output. This is illustrated in Fig. 1 as a problem for two parties (Alice and Bob) that share the same classifier network. Alice tries to convince Bob of the true class of an image while transmitting as few pixels as possible. We can thus directly define the important input variables as a small subset  $S$  of the variables that ensures a distortion below some  $\epsilon > 0$ . There is a clear trade-off in that smaller  $\epsilon$  that will require a larger set of input variables.

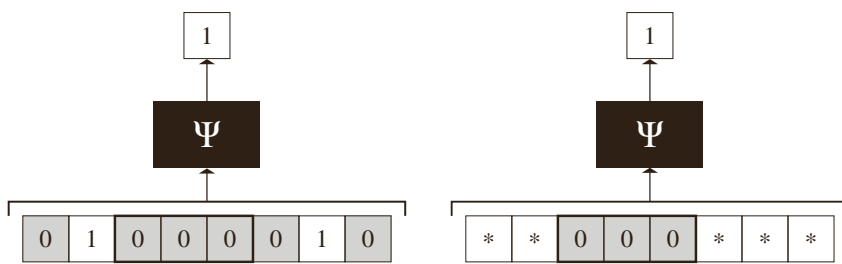
To practically solve this problem for highly nonlinear neural network classifiers, we rely on a convex relaxation of set membership. Instead of a binary decision to keep an input

fixed or replace it with noise, it is a convex combination of both.

To calculate the expectation of the classifier deviation, we make use of the Assumed Density Filtering (ADF) framework that propagates Gaussian distributions through the network layers, relying on a projection step after each case of nonlinearity. We have shown that there are no "cleverer" distributions that do not require such a projection step [2].

Combining the two ingredients of ADF and convex optimization of set membership leads us to our own saliency attribution method, called Rate-Distortion Explanation (RDE) [3]. We keep the continuous values from the optimization as the saliency map. Ordering them by size and thresholding them back to binary gives us a greedy strategy to select the most important features for any given rate. This raises the question how close our method comes to finding the optimal sets. However, due to the complexity of the task, it is not possible to show strong guarantees for our method or for any other computationally efficient method.

# Complexity of Explaining Neural Networks



**Figure 2:** Left: The Boolean function  $\Phi: \{0,1\}^8 \rightarrow \{0,1\}$  decides if a binary string contains a substring of three consecutive zeros. A relevant subset of input variables is highlighted. Right: The subset is relevant, since it is sufficient to predict the function output independent of all other input variables. In this example, the relevant set is not unique, as a second substring of three consecutive zeros exists.

This idea is strongly related to a concept in abductive logic, called prime implicants. We consider a  $d$ -ary Boolean function  $f$  and a given assignment to the  $d$  input variables  $x$ . An implicant is the smallest set of variables that, if assigned to their value in  $x$ , guarantees that  $f$  will output  $f(x)$  no matter what is assigned to the rest of the variables. A prime implicant is a subset-minimal implicant (i.e., a set that cannot be made smaller by omission while remaining an implicant (see Fig. 2. Applying this concept directly to classifiers in high-dimensional domains (images for example) is not practical, since any small part of an image can be manipulated to change the classifier decision. This can be achieved by so-called adversarial patches [T. Brown, D. Mané, A. Roy, M. Abadi, J. Gilmer, 2017]. Thus in [4],

we extended the concept to probabilistic prime implicants, where the implicant property only has to hold for a certain probability for all possible assignments.

It can be shown that solving the rate-distortion problem for a fixed distortion is equivalent to finding small probabilistic implicants, which we proved to be an NP<sup>PP</sup>-hard problem, a complexity class that lies even beyond the already challenging NP-class. Furthermore, finding implicants that only approximate the optimal set within any nontrivial approximation factor is an NP-hard problem. It is thus not possible to prove useful guarantees for computationally efficient methods. We will thus have to rely on numerical evaluation of the salience maps to demonstrate their merit.

# Sparser Relevance Maps and Orderings with Conditional Gradients

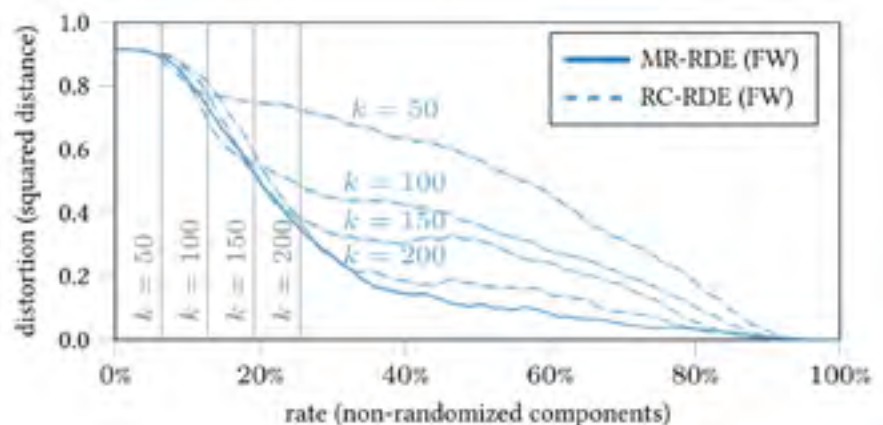
The initial formulation for RDE used an L1 penalty to bias the solution toward higher sparsity levels. Proximal methods, which are typical for such composite functions, tend to first produce dense iterates that are sparsified toward the end of the solution process. This may imply that the solution obtained after a finite number of iterations is not as sparse as expected, but also that each iteration is working on a dense iterate.

In [5], we replaced the penalty-based sparsity term with an explicit constraint on the L1 norm of the solution.

The obtained problem is then solved using a Frank-Wolfe approach, which, by design, constructs sparse iterates throughout the solution process. The sparse structure of the iterates accelerates many elementary operations performed at each iteration and saves on the memory footprint of the algorithm.

The values of the optimal relevance map are not as important per se in the interpretation as the ordering of the variables it induces. An alternative idea for RDE is therefore to directly compute optimal ordering of the variables. A given rate  $k$  would then induce a selection of the  $k$  first variables of the ordering. The set of permutation matrices defining an ordering is discrete but its convex hull, the Birkhoff polytope, can be used as a convex proxy with a well-defined linear minimization oracle calling the Hungarian algorithm, thus allowing us to leverage Frank-Wolfe to compute optimal ordering over the average of all rates.

Computational results show that the single-rate formulations perform well for their fixed rate but poorly when the rate is changed while the ordering-based formulation performs well across all rates.

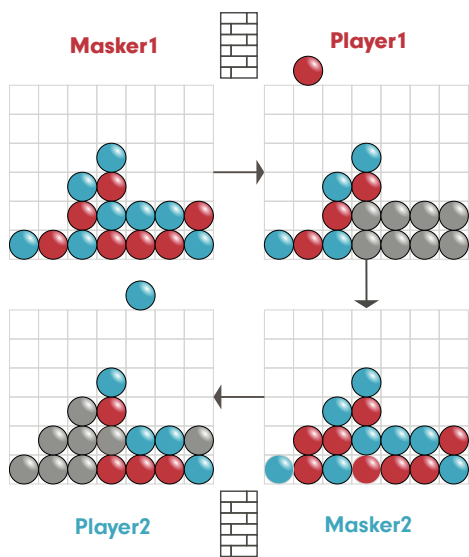


Compared to the penalty-based formulation, the constraint-based formulation using Frank-Wolfe performs better at lower rates and similarly at higher rates.

# Evaluating Relevance Maps

As mentioned, there are two approaches to evaluate XAI methods, one human-based and the other based on proxy tasks. While human-based evaluations directly measure what we want, namely that the interpretation is helpful to humans, they are difficult to scale and generalize from one setting to another. We thus aim to evaluate the different saliency methods with a fitting proxy task. Most of these tasks, like the input obfuscation described earlier, rely on evaluating the function off-manifold. The classifier has never been trained with parts of an image randomized, so inspecting the output of the function is not necessarily a good indicator of how the classifier reasons on-manifold (i.e., with complete pictures). To remedy this issue, we directly trained a neural-network-based agent on missing features.

We selected a reinforcement-learning problem of playing simple two-player games in [6]. We made use of the fact that neural networks emerged as the strongest models for reinforcement learning, for example, the first model to beat a human GO champion. Using games instead of image classification has the advantage that the input is low dimensional with weaker correlations between adjacent features. An image with random pixels hidden could likely be reconstructed using inpainting. We focused on the simple game of Connect Four illustrated Fig. 3.



The policy network of our agent maps a board state to a probability distribution over all possible moves, from which the next move is sampled. To allow the agent to always choose valid moves, we only hide the color features, not whether a field is occupied. The agent can directly be trained on missing information by randomly hiding the color features during training.

To compare different saliency methods, we devised the following setup: a masker and a player are paired against a second team of the same form. The masker, equipped with a saliency method, can present a limited amount of information to their player, who then selects the next move. We demonstrated that better information leads to the agents playing better. We then directly compare saliency methods by letting them play against each other. This setup has three advantages over image obfuscation:

1. It is canonically clear how missing features can be modeled.
2. There is no need to evaluate the network off-manifold.
3. We have a concrete task (winning the game) to use for comparison.

We show the win rate of commonly used saliency methods in a round-robin tournament in Fig. 4. The method based on RDE with a Frank-Wolfe optimizer is denoted as FW. As we can see, it is among the best performing methods. We thus demonstrated how concepts taken from abductive logic combined with convex optimization methods can lead to useful saliency methods that, while still being heuristics without guarantees, use heuristics commonly used in optimization that are generally well trusted.

**Figure 3:** Connect Four as a two-vs-two-player game. The masker chooses which color features to reveal to the player (at most half of the played pieces). The player then decides on their move based on this partial information.

Input	0.500	0.737	0.747	0.706	0.782	0.834	0.774	0.887	0.935
DeepShap	0.263	0.500	0.498	0.499	0.573	0.609	0.567	0.742	0.871
Guided Backprop	0.254	0.502	0.500	0.496	0.579	0.616	0.616	0.765	0.861
FW	0.294	0.501	0.503	0.500	0.587	0.604	0.586	0.742	0.848
Gradient	0.217	0.427	0.421	0.413	0.500	0.512	0.502	0.686	0.819
LRP- $\epsilon$	0.167	0.392	0.385	0.397	0.487	0.500	0.472	0.677	0.810
Deep Taylor	0.226	0.433	0.384	0.414	0.497	0.528	0.500	0.681	0.805
Smooth Grad	0.113	0.259	0.235	0.258	0.314	0.323	0.319	0.500	0.676
Random	0.065	0.130	0.140	0.152	0.181	0.190	0.195	0.324	0.500
	Input	DeepShap	Guided Backprop	FW	Gradient	LRP- $\epsilon$	Deep Taylor	Smooth Grad	Random

**Figure 4:** Results of the tournament between different saliency methods. The number represents the win rate of row vs. column over 1,000 games. Draws are being counted as a half win for each.



# EXPLOSION IN SPACE

Challenging the Intractable



Many problems in both computing and our daily life boil down to finding optimal solutions. This includes finding the quickest route from one place to another, dividing work between a number of people, or finding the best order in which to visit cities during a vacation trip, but also many more application-specific problems such as VLSI floor planning, multiagent planning, scheduling, and efficient code generation in compilers. Small instances of these problems are often easy to solve but for larger practical instances, the number of possible solutions grows exponentially, reaching quickly intractable dimensions that have become known as **explosion of the search space**. If a computer were to try and find the best solution by brute force (i.e., trying out every single one), it wouldn't finish for possibly thousands of years.

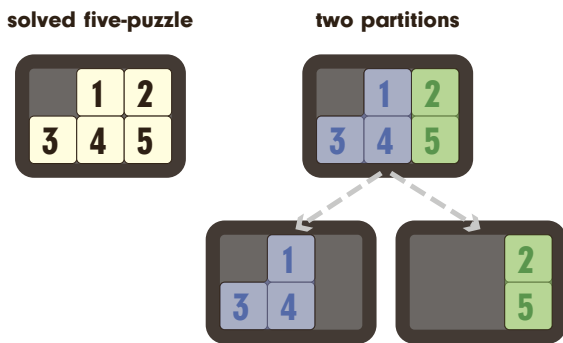
Can we still find an optimal solution within a few hours to a few days? By augmenting backtracking with problem-specific intelligence, we obtain **heuristic search** methods as a powerful tool to tackle these challenges. In our recent work, these methods have been refined through the development of better ways to make problem-specific knowledge guide the search to the best solution.

# Search in Galaxy-Scale Graphs

The heuristic search approach models search problems by finding the shortest path from an initial configuration (e.g., an empty schedule, an initial configuration of agents, etc.) to a target configuration (e.g., a completed schedule, a goal configuration of agents, etc.) by following a series of edges through a graph. Each traversal of an edge modifies some state of the problem and has a cost associated with how expensive it is to perform this modification. While the shortest path problem is usually considered fairly easy to solve, the state space graphs of real-world problems often have orders of magnitude with more vertices than there are atoms in the universe. Even an algorithm that is usually fast is ineffective with inputs of that size.

Heuristic search attacks the shortest path problem with a graph search algorithm augmented with a **heuristic function**, providing a lower bound on the cost needed to reach the goal from an arbitrary configuration. Being derived from extra knowledge about the problem space, the heuristic function allows us to quickly ignore bad edges and directs us toward the solution like a compass would direct you toward a destination. Due to this extra information, large swaths of the search space can immediately be disregarded and the set of vertices that actually need to be checked is reduced to a manageable amount.

## PATTERN DATABASES AND A MODEL PROBLEM



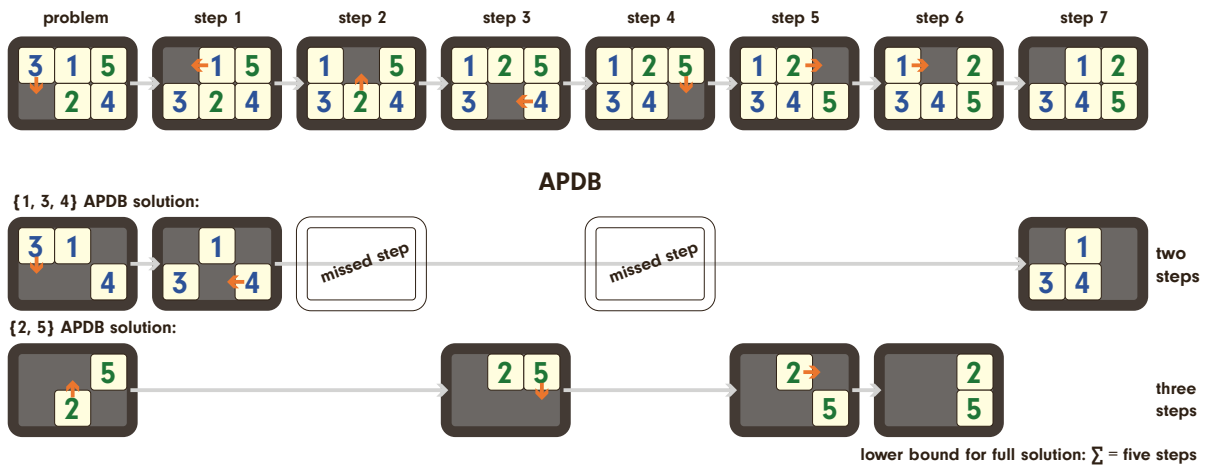
**Figure 1:** A five puzzle abstracted to the  $\{1, 3, 4\}$  and  $\{2, 5\}$  partitions. Additive pattern databases (APDBs) only consider the tiles in their partitions, treating the rest as free space.

As a model problem for this approach, our group and the community have long studied **sliding tile puzzles** such as the well-known 15 puzzle and its  $5 \times 5$  cousin, the 24 puzzle. The 24 puzzle is particularly useful in the study of heuristic search methods, as its problem instances are right at the edge of being intractable by current approaches. In addition, its regular structures enable convenient exploration of diverse strategies to improve heuristic search methods. The mapping of the 24 puzzle to a heuristic search problem is simple: each puzzle configuration is a vertex in the search space and each move is an edge of unity weight connecting the two configurations it transitions between.

The current state of the art for the construction of problem-specific heuristics is an approach called a **pattern database** or PDB. In this approach, the problem is abstracted into a simpler problem, usually by leaving out parts of the problem state. All instances of the abstracted problem are then solved by exhaustive search of the abstracted search space and a table of solutions is obtained. As the solutions of a properly abstracted problem cannot be longer than solutions to the full problem, this table serves as a lower bound and thus as an excellent heuristic for the problem we are trying to solve.

When demonstrating this idea on sliding-tile puzzles, such a simplification can be achieved by choosing a few tiles from the puzzle and discarding the others (see Fig. 1). A puzzle comprising only a few tiles has a lot less states than the full puzzle and its solutions can be tabulated. Since solving the full puzzle also involves moving each tile in the simplified puzzle to its goal position, a solution for the full puzzle cannot be shorter than a solution for the simplified puzzle, making it an admissible heuristic.





**Figure 2:** A puzzle instance is decomposed into the {1, 3, 4} and {2, 5} partitions. The pattern databases for each partition contain the shortest solutions for each configuration of tiles in their partitions. Added together, the lengths of these partial solutions predict the length of the full solution. The prediction may not be exact but is never longer than the correct length.

## Patterns and Partitions

If some pattern databases for the same problem then each consider disjoint sets of the problem state (e.g., in the case of the 24 puzzle, disjoint sets of tiles), the heuristics described by these PDBs are additive and can be summed up into an even more powerful heuristic. By finding such clever partitions of the problem state, the true potential of pattern database heuristics is found. For example, Fig. 2 shows how the solution to a five-puzzle instance is predicted through two pattern databases. The prediction is quite accurate but fails to account for two of the moves.

The main disadvantage of pattern databases is their memory consumption. Having previously studied the effectiveness of TB-sized pattern databases [4], our research has since found ways to reach similar or even better performance by combining several smaller pattern databases and a way to compress pattern databases by a factor of eight using differential encoding [3], paving the way to better resource utilization and the use of highly parallel but memory-starved computing models such as GPUs, caches for vector units, or limited high-bandwidth memory (HBM) of modern supercomputer architectures.

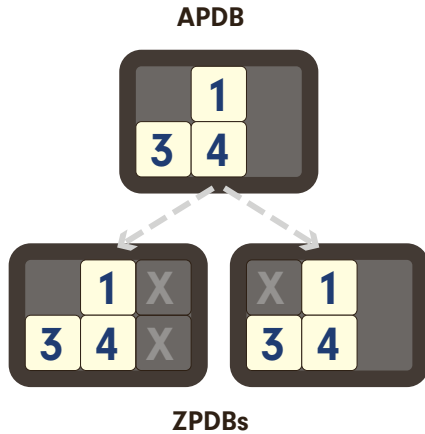


Dr. Florian Schintke  
Phone +49 30 841 85-306  
schintke@zib.de



Robert Clausecker  
Phone +49 30 841 85-263  
clausecker@zib.de

## IMPROVING PATTERN DATABASES



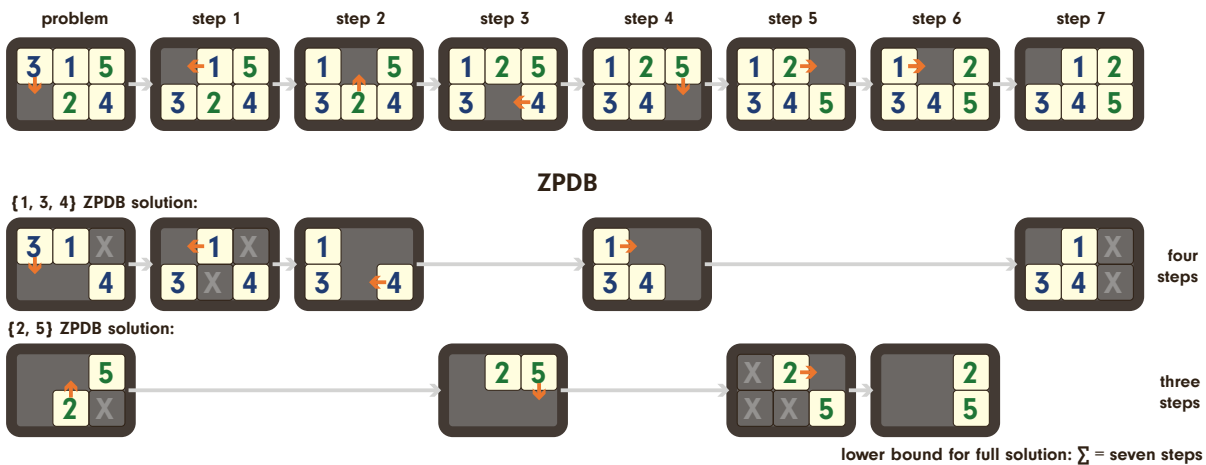
**Figure 3:** The ZPDB improves on the APDB by also tracking in which connected region of the tray the blank spot resides. In this case, it may either be to the left or to the right. Moving tiles into spots that are known not to contain the blank tile is forbidden.

To improve this idea of partition state, we have developed an extension to pattern databases called **zero-aware pattern databases (ZPDBs)** that improves the predictive power at a reasonable cost in size by taking into account not just the simplified problem state, but also classes of restrictions on the possible state transitions in the simplified

state coming from parts of the problem we have simplified away. Because these restrictions do not correspond to any particular bit of state but rather to the configuration as a whole, they can be taken into account in each pattern database of partitioning without affecting additivity.

Going back to the sliding-tile puzzles, we have implemented this idea by not only tracking the location of some tiles in a pattern database, but also which region of the tray the field not occupied by any tile in the full puzzle (blank) may reside (see Fig. 3). We can assume that all squares not occupied by tiles in the simplified puzzle, but not in the region we believe the blank to be in, are occupied by other tiles we did not track in this PDB. Hence, moves that would affect these regions cannot occur in the full puzzle and cannot be considered when solving the simplified puzzle. The result is a pattern database that occupies a bit more space due to the need to track a new bit of state, but with greatly enhanced predictive power. For example, the number of graph nodes expanded in the solution of a set of 24 puzzle instances was reduced by a remarkable 38% on average, just by tracking move restrictions due to the blank region.

Fig. 4 shows how this affects a concrete instance of the five puzzle. Compared to Fig. 2, the two pattern databases now predict the length of the solution correctly. The two moves previously not accounted for are visible to the {1, 3, 4} ZPDB in which the previous shorter solution is excluded due to the new move restrictions.



**Figure 4:** By tracking where the zero tile resides, the PDB's prediction becomes more accurate as it understands which moves that seem possible in the abstracted configuration are not actually possible once the other partitions' tiles are filled in.

## At the Frontier of the Unsolvable

Using pattern databases is like trading space for time. The larger the pattern database, the faster the search, which works well up to a certain level. But for a given database size, which pattern database has the most potential to gain speedups? In our recent publication [2], which won us the best-paper award at SoCS2021, we have proposed a novel metric to ascertain the effectiveness of heuristics for IDA\*, a popular heuristic search algorithm based on iterated depth-first search, and understood how lower bounds produced by the heuristic for different parts of the search space influence the speed of the search [1]. Based on this research, we can make informed decisions about what heuristics the memory budget is best spent on. In future research, we plan to use this knowledge to algorithmically tune the parameters of PDB heuristics for optimal performance at the same or lower memory consumption.

A major problem hampering the use of heuristic search in modern applications is the lack of parallelism in the standard search algorithms. In previous research, we have developed various approaches to parallelize IDA\* [5] and customize methods [6] based on the related A\* algorithm, a variant of breadth-first search. In future research, we would like to improve on these methods by reducing the communication and synchronization overhead and making the algorithms more scalable to modern highly parallel high-performance systems. This way, the envelope of what is tractable can be pushed further, enabling novel useful applications to be tackled.

Our work group looks back to a great portfolio of important improvements in the area of heuristic search and to a long tradition of international collaboration with key research groups in this area. To this end, ZIB has also provided a forum for the community to share their newest research in our 2018 “International Workshop on Pattern Databases and Large-Scale Search.”

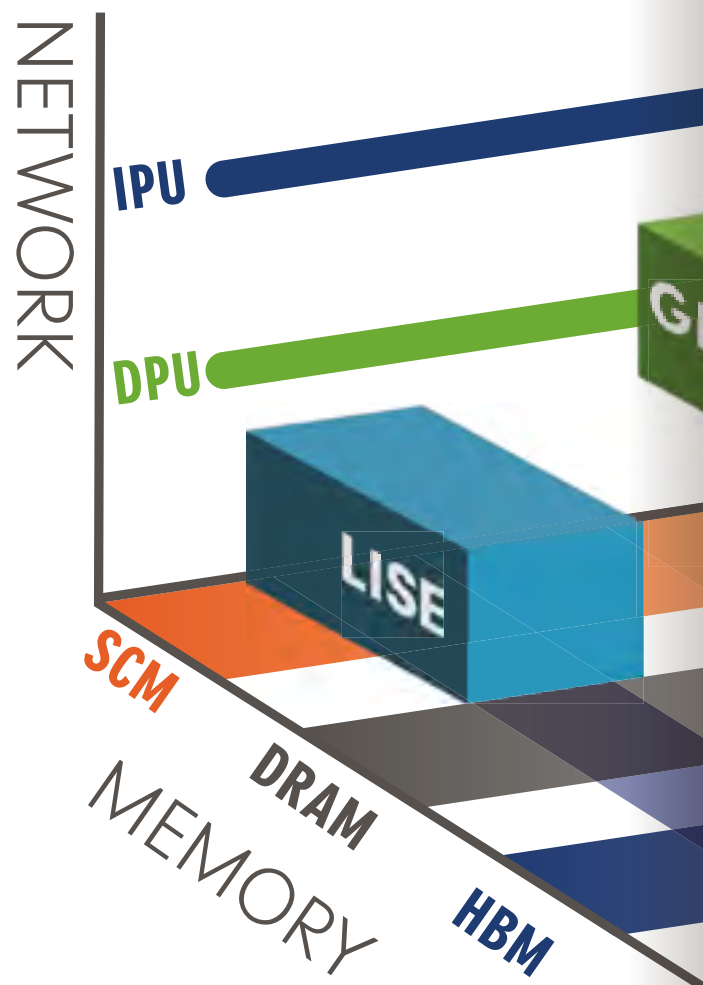
### FUTURE DIRECTIONS

The scalability and parallelization of heuristic search in large-scale systems remains challenging, but investigating new approaches for high-performance parallel heuristic search beyond the state of the art seems to offer quite some potential for further improvements.

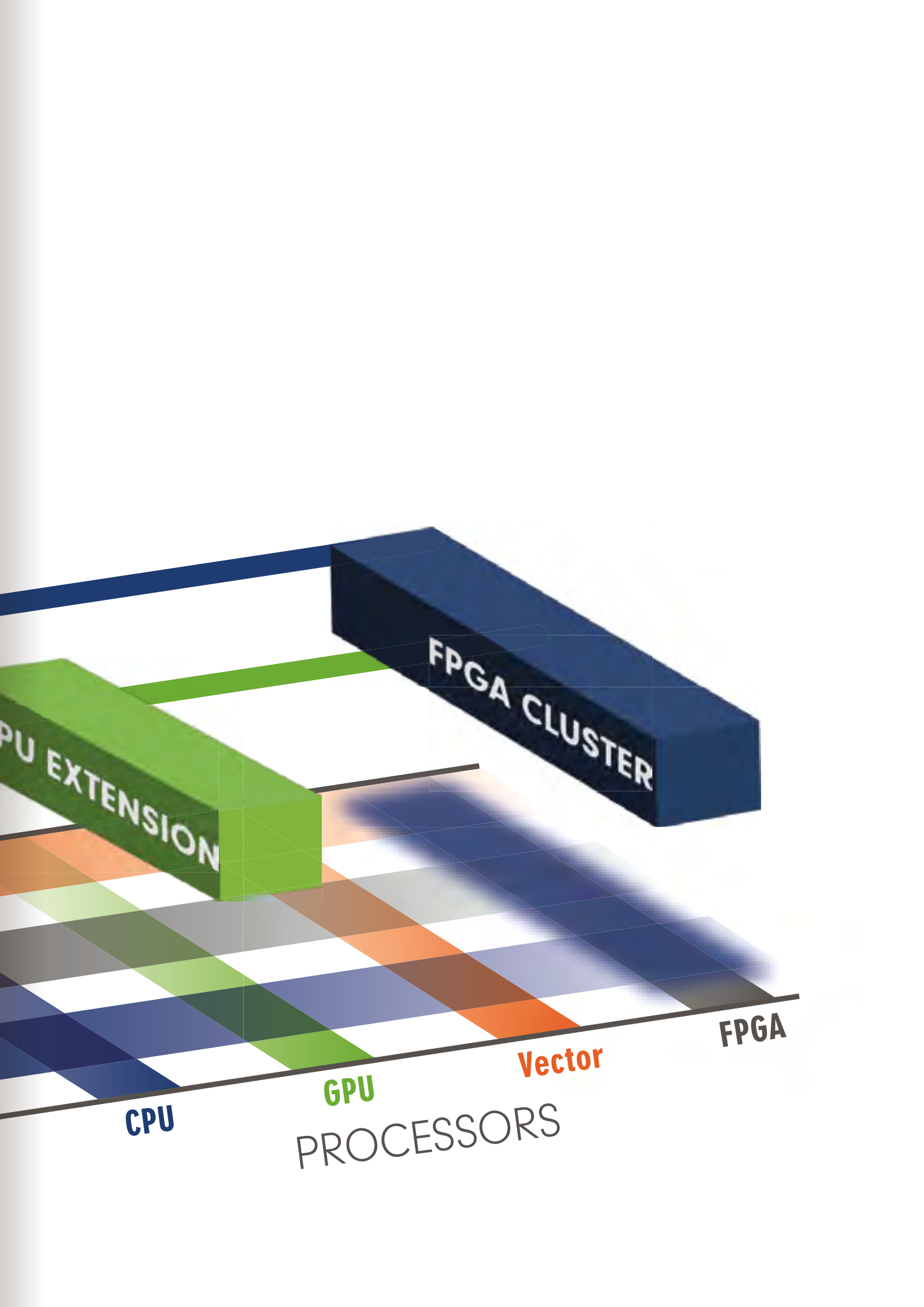
Another future field of research is the application of heuristic search techniques for program optimization and compilation. Using heuristic search methods, instruction selection and register allocation techniques could be improved with the eventual goal of being able to automatically find the optimal instruction sequence for a given piece of code.

# MASTERING MULTIDIMENSIONAL HETEROGENEITY OF COMPUTERS

Over the last years, heterogeneity has become the key element for building advanced and energy-efficient computing systems. Here, we see not only the diversity in the processor landscape but we also acknowledge technological innovations for byte-addressable storage as well as for designing smarter networks. These three areas with their technological evolutions are orthogonal as visualized in Fig. 1. For next-generation systems and for most of the theoretical combinations, real hardware is or will be available soon.



**Figure 1:** Illustration of the multidimensional space of heterogeneity in today's computer architectures (see text for explanation). Shown are (left) the current "LISE" HPC system of the NHR@ZIB, a CPU-only system comprising nodes with a traditional dynamic random-access memory (DRAM) configuration and some nodes with storage-class memory (SCM), (middle) a near-future GPU extension with high-bandwidth memory (HBM) on GPU package and potentially data-processing units (DPUs) in the network, and (right) an hypothetical FPGA cluster with FPGA boards comprising all three memory technologies and infrastructure-processing units (IPUs) as part of the smart network.



FPGA CLUSTER

CPU EXTENSION

CPU

GPU

Vector

FPGA

PROCESSORS

With these multiple options for processors, memory, and network, the aspect of a sufficient system balance becomes more challenging: (i) for the systems provider, if multiple user communities with a broad range of workloads need to be supported, and (ii) for code developers and users as problem sizes and simulation scenarios impact the overall performance depending on the chosen hardware configuration.

The balance of a HPC system – i.e., the balance of the performance of the individual architectural elements from compute over memory access and network (latency, bandwidth) to finally accessing data on persistent storage – was already one of the key design goals for supercomputers in the 1990s. At this time, system designers like Seymour Cray or Kenneth Batcher made their comments to emphasize the importance of this subject.

“A supercomputer is a device for turning compute-bound problems into I/O-bound problems.”

(Kenneth Edward Batcher)

“Anyone can build a fast CPU. The trick is to build a fast system.”

(Seymour Roger Cray)

One approach to estimate the balance of compute capabilities and access to memory is to run the famous STREAM [4] benchmark which already provides a defined metric of “Machine Balance” [<https://www.cs.virginia.edu/stream/>].

Here, we want to explain how this multidimensional heterogeneity can be mastered by users and code developers.



Dr. Steffen Christgau  
Phone +49 30 84185-214  
[christgau@zib.de](mailto:christgau@zib.de)



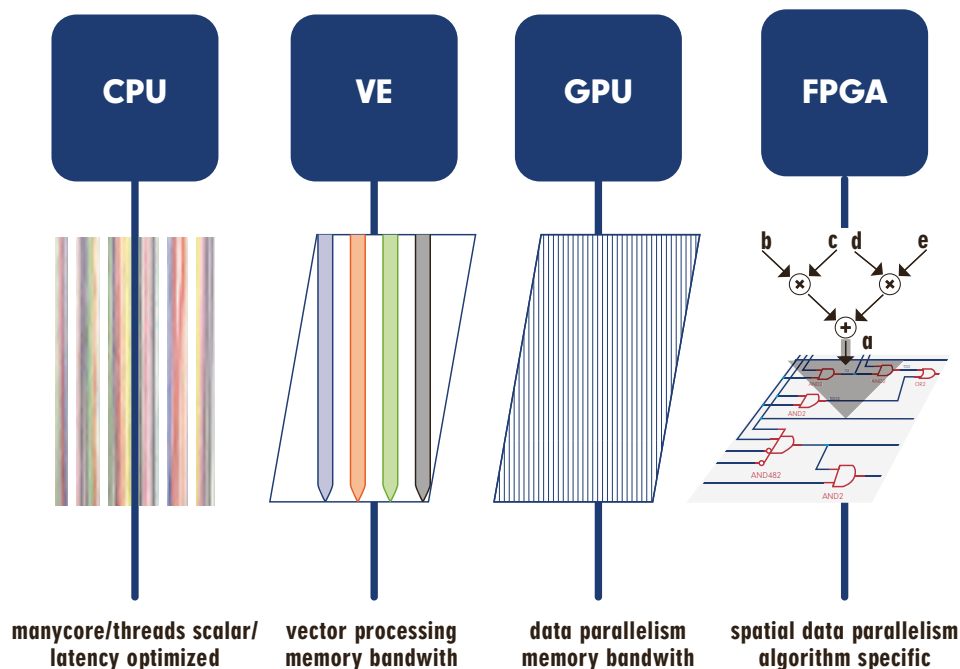
Dr. Thomas Steinke  
Phone +49 30 84185-114  
[steinke@zib.de](mailto:steinke@zib.de)

## Dimension 1: Multidevice/Multivendor Programming

Although in the last decade, general-purpose CPUs evolved to many-core processors with an increasing number of functional units, graphics-processing units (GPUs) made their way into mainstream computing architectures as compute building blocks to accelerate certain workloads. Another implementation of a highly parallel compute device architecture are vector processors (vector engines - VE) that became popular in the 1970s, dominated the HPC market till the 1990s, and are on the rise again today. If performance per watt is a key requirement, field-programmable gate arrays (FPGAs) are an alternative in power-restricted compute environments nowadays. Whereas GPUs, vector engines, and FPGAs can be efficiently used for different classes of workload, the need for highly energy-efficient solutions triggered the development of special-purpose processors, for example, particularly in the AI segment over the last few years.

The key motivation for this diversity in processor architecture is the increasing need for energy-efficient solutions (i.e., the highest performance per watt ratio). Strictly speaking, there was never a one-size-fits-all solution but one was more willing for compromises in the past. Today, the focus of providers and other stakeholders in the HPC segment is on energy efficiency as well as usability and security. Ideally, one would select the best fitting processor/accelerator and memory architecture for each workload. This idea led to the aforementioned diversity of mainstream processor architectures.

Given the new variety of computing hardware, the question arises how these devices are programmed. Related to that is the performance of the resulting code, its portability across architectures, and the time it takes to produce such a code, which is referred to as productivity.



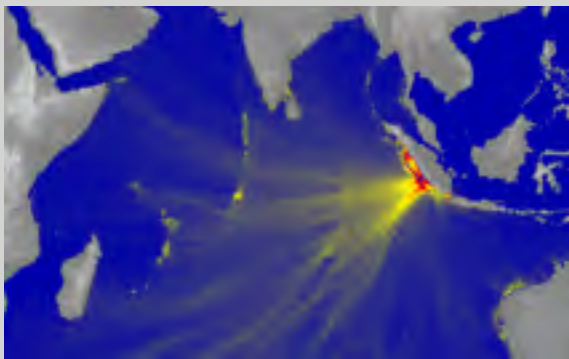
**Figure 2:** Schematic key characteristics of many-core processors (CPU), vector engines (VE), general-purpose graphics units (GPU), and field-programmable gate arrays (FPGA). The many CPU cores are latency optimized. Vector engines have multiple compute cores and are traditionally designed for exceptional memory bandwidth and balance. GPUs implement a very high count of smaller compute cores for parallel processing and usually a high memory bandwidth. In contrast to all these devices, FPGAs provide a fundamentally different approach, since spatially distributed elemental logic blocks can be algorithm-specific connected and thus hardware adapts to the algorithm.

From a pure language perspective, there is hardly any need to switch to new programming languages. Traditional procedural languages, such as C/C++ are often the foundation to program accelerators. For GPUs, vendors added syntax elements, such as implicit variables and keywords, to C/C++. These additions allow relatively low-level access to the GPU functionality and allow to specify kernels (i.e., per-item operations) which are later executed on the GPU. With the help of runtime functions, management tasks like memory operations and device selection are performed.

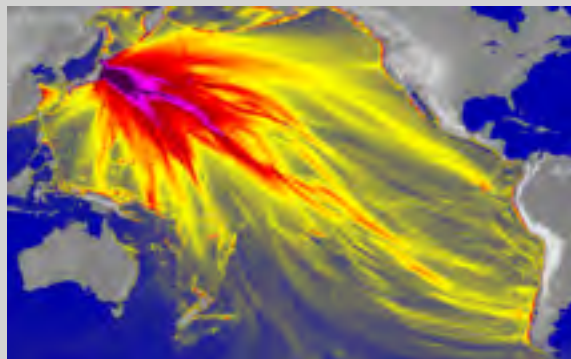
Historically, Nvidia's CUDA programming interface is an example of such a language extension to program their GPUs. Besides minor differences, the AMD counterpart Heterogeneous-Compute Interface for Portability (HIP) strongly resembles the extensions CUDA made to C/C++. This eases but does not fully provide code portability across devices from these and other vendors and their productivity can be rated as medium (see Table 1).

On the other side, vendor-neutral industry standards aim to achieve portability. The Open Compute Language (OpenCL) and SYCL are native libraries for C and C++, respectively. They allow cross-platform programming on a higher level of abstraction. Therefore, they provide more portability than, for example, CUDA. However, both approaches require support by runtime libraries and compilers. For SYCL, a large ecosystem of appropriate implementations has been established over time that currently supports the programming of GPUs, CPUs, and FPGAs from different vendors using the C++ language.

Nevertheless, SYCL is still quite new compared to CUDA and recently gained additional attention with the release of the oneAPI programming framework by Intel. While the code portability provided by SYCL is desirable, the efforts to port existing CUDA codes to another programming interface may hinder researchers from going in this direction. To lower the burden of this task, a compatibility tool was provided by Intel that assists developers in the migration from CUDA to Data-Parallel C++ (DPC++) which is Intel's SYCL implementation with some extensions.



ZIB collaborated with Intel on the Intel oneAPI framework. We examined the migration process and the portability of an existing code base across multiple hardware architectures and vendors using the tsunami simulation code *easyWave* as a demonstrator. Originally developed by the German GeoForschungs-Zentrum for CUDA and thus bound to GPUs from Nvidia, the stencil-based simulation was successfully ported to SYCL. Assisted by the oneAPI Compatibility Tool, a single code base was developed by the A4IA research group [3].



The new code is able to run on GPUs of different vendors including Intel and AMD, and Intel CPUs and FPGAs. The SYCL-based code can also run on Nvidia GPUs again, on which the performance loss is only 5% at most compared to the original CUDA version. Thus, *easyWave* demonstrates both code and performance portability on GPUs. The code even runs correctly on FPGAs although the employed algorithms and data structure are not well suited for such architectures. This shows that code portability does not always imply performance portability and code portability does not relieve programmers from the task of designing algorithms for particular hardware architectures.



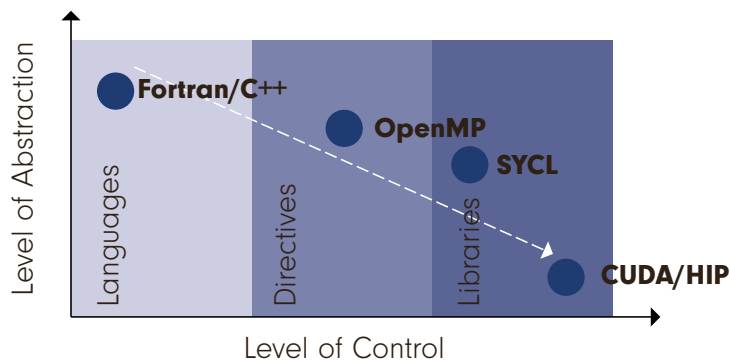
Although SYCL provides vendor-neutral code portability to different platforms, its usage is limited to C++. For code bases written in Fortran, this is hardly an option. For Fortran support, the directive-based Open Multi-Processing standard (OpenMP) allows to annotate code with compiler instructions. With such approaches, which can also be used in C/C++, code regions can be marked to be executed on accelerators and data movements to and from accelerators are marked in a similar way. This allows incremental migration of code without major rewrites and enables higher productivity. While the GPU support of OpenMP for GPUs matured over time and its functionality converged with the competing Open Accelerator (OpenACC) specification, the support for FPGAs was hardly investigated. Within the OKRA-HPC project, funded by the German Federal Ministry of Education and Research (BMBF), the working group Algorithms for Innovative Architectures (A4IA) showed that support for this class of devices is also possible with this directive-based standard.

The usage of directives simplifies the usage of accelerators, but still requires declaring data locations and which

pieces of a code are intended for accelerators. An even more elegant way to use devices is to let the compiler fully decide which code portions should be accelerated. Recent language changes in Fortran and C++, which allow expressing dependency-free loops, relieve programmers of the task of writing device-specific code.

With such a high level of abstraction from the targeted device, almost all control over the resulting code is handed over to the compiler. The performance may not be as optimal as code written with libraries closer to the hardware, like CUDA and SYCL.

This kind of approach is often helpful for getting codes to accelerators with only minimal efforts. However, if more control is needed, lower abstraction will generally be required. This opens an additional dimensionality in the first dimension of heterogeneity as illustrated in Fig. 3. Software developers must find an appropriate position to develop code for accelerating devices and choose a degree of abstraction, portability, and productivity (see Table 1).



**Figure 3:** Levels of abstraction and control for programming heterogeneous devices.

**Table 1:** Programming Languages and Standards for Heterogeneous Computing

Name	Portability	Productivity	Performance	Target Platforms
Nvidia CUDA	★	★★	★★★	Nvidia GPUs
AMD HIP	★★	★★	★★★	AMD GPUs
SYCL	★★★	★★	★★★	CPU, GPU, FPGA, and other compute devices
OpenMP	★★★	★★★	★★	CPU, GPU, FPGA, Vector Engines
OpenACC	★★	★★★	★	GPU
OpenCL	★★	★	★★★	CPU, GPU, FPGA
Standard Languages	★★★	★★★	★	CPU, GPU

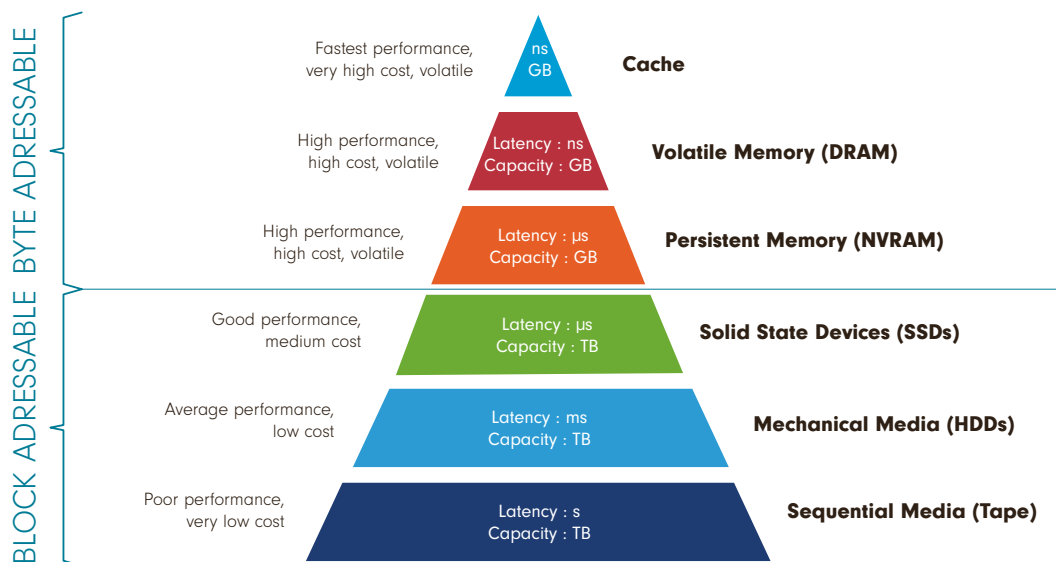
# Dimension 2: Choosing the Right Memory for Data

With the rise and evolution of the microprocessors, the gap between processor speed and memory performance began to increase. Innovations in on-chip packaging technology, the assembly of chiplets, and 3D stacking enabled the introduction of a new memory category close to the processors - High-Bandwidth Memory (HBM). It means that today's state of the art of GPUs, vector engines, and many-core CPUs provide another level in the memory hierarchy apart from the traditional dynamic random-access memory (DRAM).

Furthermore, with the Optane technology, a new class of Storage Class Memory (SCM) was introduced and com-

plements the DRAM memory modules on the main board. With its characteristics, the SCM represents an additional, new level in the hierarchy of byte-addressable on-node memory comprising Cache, HBM, DRAM, and SCM.

In mainstream programming environments, the memory hierarchy has almost ever been considered transparent. This is a valid model for systems with a single type of main memory. As the memory hierarchy changes (see Fig. 4) and new memory types are present in computer systems, the usage of memory in software needs to adapt.



**Figure 4:** Memory hierarchy with the new layer of persistent memory that is byte addressable.

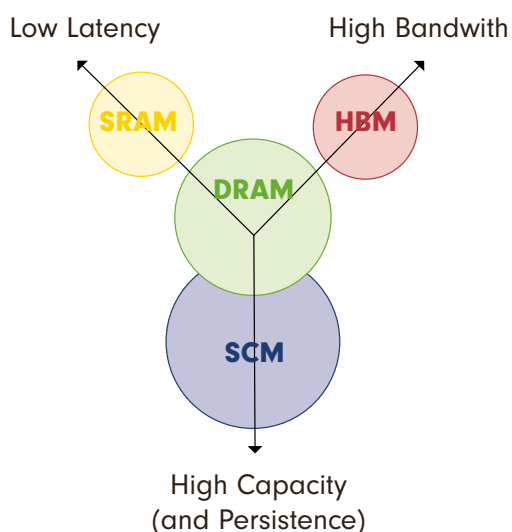
Besides the usually employed DRAM technology, which provides a good compromise between low latencies and high bandwidth, High Bandwidth Memory (HBM) provides, as its name suggests, high bandwidths but also undesirable high latencies. However, the usual memory capacity per core is quite low at just a few gigabytes. The same applies for DRAM. Both memory types are also volatile and data is lost when the containing system is powered off. Storage Class Memory (SCM), however, allows the encrypted data to be kept when turned off. It also provides multiple terabytes of capacity. The main drawback of SCM is, however, a reduced latency and bandwidth.

As shown in Fig. 5, this variety of properties leads to a three-dimensional decision space, where one can trade between capacity, latency, and bandwidth depending on the memory types installed in a system. This results in the initial challenge of choosing the appropriate memory type for application data. Since there is no single memory type that is well-suited to all data, application data needs to be partitioned between different memory types. At this point, performance-analysis tools provide assistance to code developers to check whether memory access patterns in application kernels would benefit from HBM or the usage

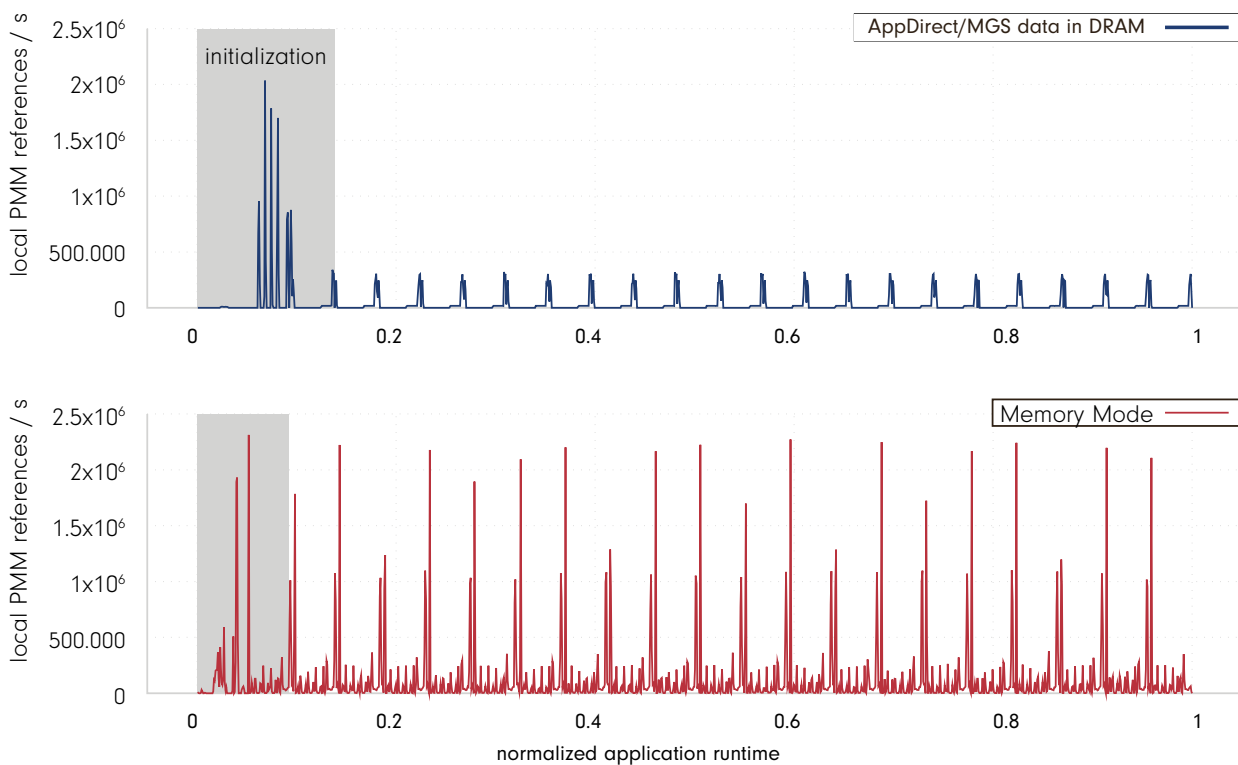
of SCM would not degrade performance too much. Based on such assisted observations, developers can make decisions between hot and cold data and which memory type is best suited for which data.

If the decision is made, data needs to be allocated in that particular memory. Even in low-level languages, additional efforts are required and libraries must be employed to control allocations in data. Libmemkind is an example for the C programming language. This library is also the foundation for additional allocators in C++. These languages share explicit memory management. Conversely, memory is usually allocated and freed automatically by the runtime in languages like Fortran. Thus, additional efforts are required to realize the intended partitioning.

Such efforts may include the addition of language keywords or modifiers that specify a preferred memory location for data. Alternatively, explicit function calls may be added to control the location of the memory allocation and strategy of the runtime library. The strategy may define which memory type is a hard requirement for certain allocations and which types might be tried second.



**Figure 5:** Three-dimensional decision space for choosing memory types.



**Figure 6:** By partitioning the application data across different memory technologies (here DRAM and SCM) according to their memory access pattern, a significant performance gain can be achieved. In this figure, memory access rates to the SCM (on the y-axis, PMM references per second) for two allocation modes are presented. (Bottom, in red) Transparent hardware-managed data partitioning and migration. (Top, in blue) Optimized partitioning by explicitly assigning the memory placement to SCM or DRAM.

The latter shows much less memory references to SCM and a performance gain of up to 8% is achieved (taken from [2]).

Such an approach was studied with the essential multigrid solver of the parallel large eddy simulation code PALM. The A4IA research group developed an allocator that was integrated into the employed Fortran runtime library to allocate data either in SCM or DRAM and is able to accept preferences (see Fig. 6). In this earlier study, we have shown that by partitioning the application data, a performance gain of up to 8% could be achieved compared to a transparent hardware solution that migrates data between memory types transparently [2].

To improve the usability of different memory types with higher productivity, the directive-based OpenMP standard defines so-called memory spaces. They allow memory characteristics intended for certain memory allocations to be specified. The decision as to which memory type should be employed is still left to the programmer. Within the NHR "PerfLab" project, ZIB together with other centers of the National High Performance Computing Alliance Erlangen, Aachen, Paderborn, and Göttingen are currently investigating how performance-analysis tools can help programmers to make a decision within that second dimension of heterogeneity.

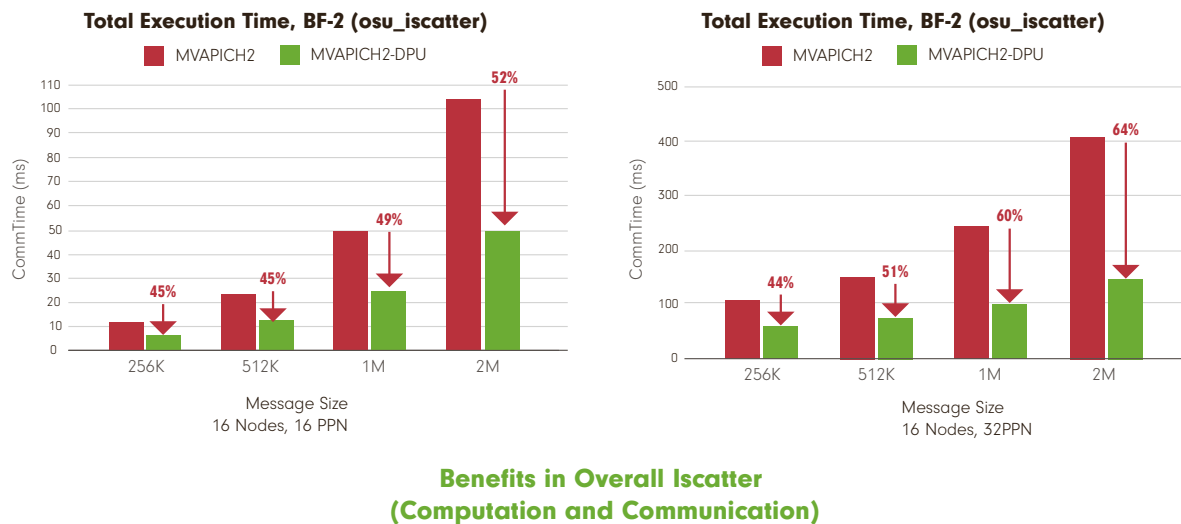
# Dimension 3: Smart Networks

Smart networks are a popular topic for designing the interplay of multiple active clients in an interconnected world. For data centers, innovative network configurations with active, partially programmable hardware devices operating at link speed can provide certain data (packet) processing capabilities. For high-performance computers, this means that, for example, MPI operations can be off-loaded from the compute units to network devices. Some of these active components are programmable by end users.

In multiuser environments like HPC resources with dynamically changing requirements regarding direct node-to-node communication, accessing data on shared file systems, object stores over fabric, and other services (data archives, remote data sources), the network is the heart of all services. Users and providers expect that it can deliver data according to the needs with suitable bandwidth and latency while fulfilling serious security requirements, easy maintainability, and flexible configurability. Thus, it is not surprising that attempts are increasingly being made to

integrate more intelligence into the network infrastructure with data-processing units (DPUs), infrastructure-processing units (IPUs), or others. The integration of security measures is an important application scenario for a provider. At the top, these smart network components comprise general-purpose CPU cores and FPGAs that are programmable at the provider.

A recent example demonstrates the possibility to migrate certain parts of collective communication operations onto a DPU (see Fig. 7 taken from [1]). The general concept is not new because in the past sophisticated HPC interconnects already implemented selected operations of common communication patterns into the fabric's networking chips (e.g. Cray Aries, HPE Slingshot, Fujitsu Tofu D). With the smart network devices, a vendor, provider, or end users may have the option to migrate user functionality of HPC and AI codes close to the data. But, it has to be broadly demonstrated which user scenarios benefit from these options.



**Figure 7:** Performance benefits with data-processing units (DPUs) for a collective non-blocking scatter operation of the message-passing interface (MPI). Shown are the total execution times for different message sizes using the standard MPI implementation (in red) and the version with DPU support (in green) on 16 nodes with 32 MPI processes each (lower is better). A significant performance improvement can be achieved for communication operations if they are handled by smart network devices (figure taken from [1]).

Increasing demands for AI training throughput and HPC simulations stimulate technology innovations and lead to increasing heterogeneity of major components on modern computer architectures – compute, storage, and network. We expect that this trend will continue and will thus impact the software ecosystems and programming models for future computer generations.

# REFERENCES

## Spreading Dynamics: Cultural Change, COVID-19, and a Mobility Transition

[1] N. Djurdjevac Conrad, L. Helfmann, J. Zonker, S. Winkelmann, and Ch. Schütte. *Human mobility and innovation spreading in ancient times: A stochastic agent-based simulation approach*. EPJ Data Science, 7(1):24, 2018.

[2] M. Kostre, V. Sunkara, and N. Djurdjevac Conrad. *Inference of historical influence networks*. International Conference on Complex Networks & Their Applications Proceedings, 110–113, 2021.

[3] <https://geoserver.dainst.org/maps/6002>

[4] N. Wulkow, R. Telgmann, K.-D. Hungenberg, C. Schütte, and M. Wulkow. *Deterministic and Stochastic Parameter Estimation for Polymer Reaction Kinetics I: Theory and Simple Examples*. Macromol. Theory Simul., 30: 2100017, 2021.

[5] S. A. Müller, M. Balmer, W. Charlton, R. Ewert, A. Neumann, C. Rakow, T. Schlenker, and K. Nagel. Predicting the effects of COVID-19 related interventions in urban settings by combining activity-based modelling, agent-based simulation, and mobile phone data. PLOS ONE, 16(10), 2021.

[6] H. Wulkow, T. Conrad, N. Djurdjevac Conrad, S. A. Müller, K. Nagel, and C. Schütte. Prediction of Covid-19 spreading and optimal coordination of counter-measures: From microscopic to macroscopic models to Pareto fronts. PLOS ONE, 16(4):1–29, 2021.

[7] L. Helfmann, N. Djurdjevac Conrad, A. Djurdjevac, S. Winkelmann, and C. Schütte. *From interacting agents to density-based modeling with stochastic PDEs*. Communications in Applied Mathematics and Computational Science, 16(1):1–32, 2021.

[8] S. Winkelmann, J. Zonker, C. Schütte, and N. Djurdjevac Conrad. Mathematical modeling of spatio-temporal population dynamics and application to epidemic spreading. Mathematical Biosciences, 336:108619, 2021.

[9] J.-H. Niemann, S. Winkelmann, S. Wolf, and C. Schütte. Agent-based modeling: Population limits and large timescales. Chaos: An Interdisciplinary Journal of Non-linear Science, 31(3):033140, 2021.

[10] Wolf, S., Fürst, S., Geiges, A., Laublichler, M., Mielke, J., Steudle, G., Winter, K., and Jaeger, C. *The Decision Theatre Triangle for societal challenges*. Global Climate Forum, 2021. [https://globalclimateforum.org/wp-content/uploads/2021/04/GCF-wp2\\_2021.pdf](https://globalclimateforum.org/wp-content/uploads/2021/04/GCF-wp2_2021.pdf).

[11] S. Klus, F. Nüske, S. Peitz, J.-H. Niemann, C. Clementi, and C. Schütte. *Data-driven approximation of the Koopman generator: Model reduction, system identification, and control*. Physica D: Nonlinear Phenomena, Volume 406:132416, 2020.

[12] J.-H. Niemann, S. Klus, and C. Schütte. *Data-driven model reduction of agent-based systems using the Koopman generator*. PLOS ONE, Vol. 16, 2021.

[13] Romanization data were made by F. Schweigart and include data provided by iDAI.world: iDAI.objects: <https://arachne.dainst.org>; iDAI.geoserver: <https://geoserver.dainst.org/maps/6002>; Roger Bagnall, et al. (eds.), Pleiades: A Gazetteer of Past Places, 2016, <http://pleiades.stoa.org/>; Trismegistos: [www.trismegistos.org](http://www.trismegistos.org); Digital Atlas of the Roman Empire (DARE): <https://dh.gu.se/dare/> (4.0 International license).

## Inference and Simulation of Molecular and Cellular Processes

[1] B. Weber, G. Greenan, S. Prohaska, D. Baum, H.-C. Hege, T. Müller-Reichert, A. Hyman, and J.-M. Verbovatz. *Automated Tracing of Microtubules in Electron Tomograms of Plastic Embedded Samples of Caenorhabditis Elegans Embryos*. Journal of Structural Biology, 178[2], pp.129–138, 2012. DOI: 10.1016/j.jsb.2011.12.004.

[2] A. Rigort, D. Günther, R. Hegerl, D. Baum, B. Weber, S. Prohaska, O. Medalia, W. Baumeister, and H.-C. Hege. *Automated segmentation of electron tomograms for a quantitative description of actin filament networks*. Journal of Structural Biology, 177[1], pp. 135–144, 2012. DOI: 10.1016/j.jsb.2011.08.012.

[3] N. Lindow, F. N. Brüning, V. J. Dercksen, G. Fabig, R. Kiewisz, S. Redemann, T. Müller-Reichert, S. Prohaska, and D. Baum. *Semi-automatic stitching of filamentous structures in image stacks from serial-section electron tomography*. Journal of Microscopy, 284[1], pp. 25–44, 2021. DOI: 10.1111/jmi.13039.

[4] S. Redemann, J. Baumgart, N. Lindow, M. Shelley, E. Nazockdast, A. Kratz, S. Prohaska, J. Brugués, S. Fürthauer, and T. Müller-Reichert. *C. Elegans Chromosomes Connect to Centrosomes by Anchoring into the Spindle Network*. Nature Communications, 8, 2017. DOI: t10.1038/ncomms15288.

[5] A. Ernst, C. Schütte, S. Sigrist, S. Winkelmann. *Variance of filtered signals: characterization for linear reaction networks and application to neurotransmission dynamics*. Mathematical biosciences, 343, 108760. 2022.

[6] S. Ray, A. Thies, V. Sunkara, H. Wulkow, M. Özgür Celik, F. Yergöz, C. Schütte, C. Stein, M. Weber, and S. Winkelmann. *Modelling altered signalling of G-protein coupled receptors in inflamed environment to advance drug design*. bioRxiv, 2021. DOI: 10.20.465073.

[7] S. Ray, V. Sunkara, C. 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*, 46(18):1443–1452, 2020.

[8] C. Gorgulla, S. Cinaroglu, P. D. Fischer, K. Fackeldey, G. Wagner, and H. Arthanari: *VirtualFlow Ants - Ultra-Large Virtual Screenings with Artificial Intelligence Driven Docking Algorithm Based on Ant Colony Optimization*. *Molecular Sciences, Special Issue Artificial Intelligence & Deep Learning Approaches for Structural Bioinformatics*, 22(11):5807, 2021.

[9] S. Roehl, M. Weber, and K. Fackeldey: *Computing the minimal rebinding effect for nonreversible processes*. *SIAM MMS*, 19(1):460–477, 2021.

[10] R. Sechi, A. Sikorski, and M. Weber: *Estimation of the Koopman generator by Newton's Extrapolation*. *SIAM MMS*, 19(2):758–774, 2021.

[11] A. Sikorski, M. Weber, and C. Schütte: *The Augmented Jump Chain - a sparse representation of time-dependent Markov jump processes*. *Advanced Theory and Simulations*, 4(4):2000274, 2021.

[12] R. Sechi, M. Weber, and K. Heyne: *MSM building and projection for the analysis of time-resolved spectra*. *PAMM*, 21(1):e202100102, 2021.

### **Multienergy Models for European Energy System Planning**

[1] D. Beulertz, S. Charousset, D. Most, S. Giannelos, and I. Yueksel-Erguen, *Development of a Modular Framework for Future Energy System Analysis*. UPEC 2019, 2019.

[2] T. Koch, B. Hiller, M. E. Pfetsch, and L. Schewe, Eds., *Evaluating Gas Network Capacities*. SIAM-MOS Series on Optimization, 2015.

[3] I. Yueksel-Erguen, J. Zittel, Y. Wang, F. Hennings, and T. Koch, *Lessons learned from gas network data preprocessing*. ZIB-Report 20-13, Zuse Institute Berlin, Berlin, 2020.

[4] M. Dell'Amico et al., *Documentation of the plan4res public dataset - deliverable D4.5 (Version 4.2)*. 2020.

[5] W. Ackooij van, H. Gevret, X. Warin, N. Langrené, A. Lelong, and J. Maheshwari, *plan4res Deliverable D5.2: New version of the StOpt library adapted to Plan4res framework*. 2019.

[6] D. Most et al., *A Novel Modular Optimization Framework for Modelling Investment and Operation of Energy Systems at European Level*. ZIB-Report 20-08, Zuse Institute Berlin, Berlin, 2020.

[7] D. Most, L. Wyrwoll, C. Schmitt, and I. Yueksel-Erguen, *plan4res D2.2 - Case Study 1 Report - Multimodal energy concept for achieving Europe's carbon reduction goals*. 2021.

[8] S. Charousset and D. Daniel, *plan4res D2.4: Assessing overall system costs due to massive RES integration*. 2021.

[9] R. Schwarz et al., *Network and Storage. Mathematical Optimization for Efficient and Robust Energy Networks*, N. S. Hadjidimitriou, A. Frangioni, T. Koch, and A. Lodi, Eds. AIRO Springer Series, 2021, pp. 89–105.

### **Multiobjective Shortest Paths**

[1] P. Maristany de las Casas, L. Kraus, R. Borndörfer, A. Sedeno-Noda, *Targeted Multiobjective Dijkstra Algorithm*. <http://arxiv.org/abs/2110.10978>.

[2] P. Maristany de las Casas, R. Borndörfer, L. Kraus, A. Sedeño-Noda, *An FPTAS for Dynamic Multiobjective Shortest Path Problems*. *Algorithms*, 14(2), pp. 1–22, 2021.

[3] P. Maristany de las Casas, A. Sedeno-Noda, R. Borndörfer, *An Improved Multiobjective Shortest Path Algorithm*. *Computers & Operations Research Volume 135*, November 2021, 105424.

[4] R. Euler, R. Borndörfer. *A Graph- and Monoid-based Framework for Price-Sensitive Routing in Local Public Transportation Networks*. 19th Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2019), pp. 12:1–12:15, vol. 75. OpenAccess Series in Informatics (OASICS), 2019.

### Explainability and Interpretability in Deep Learning

- [1] S. Wojciech, A. Binder, G. Montavon, S. Lapuschkin, and K.-R. Müller (2017). *Evaluating the visualization of what a deep neural network has learned*. IEEE Transactions on Neural Networks and Learning Systems 28(11):2660–2673.
- [2] J. Macdonald and S. Wäldchen (2022). *A Complete Characterisation of ReLU-Invariant Distributions*. Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS) 2022, Valencia [arxiv.org/abs/1905.11092](https://arxiv.org/abs/1905.11092).
- [3] J. Macdonald, S. Wäldchen, S. Hauch, and Gitta Kutyniok (2019). *A rate-distortion framework for explaining neural network decisions*. 37th International Conference on Machine Learning (ICML).
- [4] J. Macdonald, S. Wäldchen, S. Hauch, and Gitta Kutyniok (2020). *Explaining neural network decisions is hard*. [arXiv:2110.08105](https://arxiv.org/abs/2110.08105).
- [5] J. Macdonald, M. Besançon, and S. Pokutta (2021). *Interpretable Neural Networks with Frank-Wolfe: Sparse Relevance Maps and Relevance Orderings*. [arXiv:2202.11797](https://arxiv.org/abs/2202.11797).
- [6] S. Wäldchen, F. Huber, and S. Pokutta (2022). *Training Characteristic Functions with Reinforcement Learning: XAI-Methods Play Connect Four*.

### Explosion in Space

- [1] R. Clausecker. *The Quality of Heuristic Functions for IDA\**. ZIB-Report 20-17 and Master's Thesis HU Berlin.
- [2] R. Clausecker, Florian Schintke. *A Measure of Quality for IDA\* Heuristics*. Proceedings of the Fourteenth International Symposium on Combinatorial Search (SoCS 2021), pp. 55–63, 2021. Best paper award.
- [3] R. Clausecker, A. Reinefeld. *Zero-Aware Pattern Databases with 1-Bit Compression for Sliding Tile Puzzles*. Proceedings of the Twelfth International Symposium on Combinatorial Search (SoCS 2019), pp. 35–43, 2019.
- [4] R. Döbbelin, T. Schütt, A. Reinefeld. *Building Large Compressed PDBs for the Sliding Tile Puzzle*. Workshop on Computer Games, 2013 (preprint available as ZIB-Report 13–21).
- [5] A. Reinefeld, T. A. Marsland. *Enhanced Iterative-Deepening Search*. IEEE Transactions on Pattern Analysis and Machine Intelligence. vol 16, no 7, pp. 701–710. DOI 10.1109/34.297950.
- [6] T. Schütt, R. Döbbelin, A. Reinefeld. *Forward Perimeter Search with Controlled Use of Memory*. International Joint Conference on Artificial Intelligence, IJCAI-13, Beijing, 2013.
- Image source: NASA/JPL-Caltech/ESA/CXC/STScI <https://photojournal.jpl.nasa.gov/catalog/PIA12348>.

### Mastering Multidimensional Heterogeneity of Computers

- [1] D. K. Panda. *Designing High-Performance and Scalable Middleware for HPC, AI and Data Science Based on MPI: The MVAPICH2-Based Approach*. Keynote talk at ExaMPI '21, [https://mvapich.cse.ohio-state.edu/static/media/talks/slide/exampm\\_keynote\\_talk\\_dk.pdf](https://mvapich.cse.ohio-state.edu/static/media/talks/slide/exampm_keynote_talk_dk.pdf).
- [2] S. Christgau and T. Steinke. *Leveraging a Heterogeneous Memory System for a Legacy Fortran Code: The Interplay of Storage Class Memory, DRAM and OS*. 2020 IEEE/ACM Workshop on Memory Centric High Performance Computing (MCHPC), pp. 17–24, 2020.
- [3] S. Christgau and T. Steinke. *Porting a Legacy CUDA Stencil Code to oneAPI*. 2020 IEEE International Parallel and Distributed Processing Symposium Workshops, IPDPSW 2020, New Orleans, LA, USA, May 18–22, 2020, pp. 359–367, 2020.
- [4] J. McCalpin. *Memory Bandwidth and Machine Balance in High Performance Computers*. IEEE TCCA Newsletter, December 1995.



# NOTES

# PUBLICATIONS

## PEER-REVIEWED (174)

Doguscan Ahiboz, Elina Andresen, Phillip Manley, Ute Resch Genger, Christian Würth, Christiane Becker (2021). Enhanced Photon Upconversion Using Erbium-Doped Nanoparticles Interacting with Silicon Metasurfaces. In *IEEE 48th Photovoltaic Specialists Conference (PVSC)*, 2008–2010. <https://doi.org/10.1109/PVSC43889.2021.9518495>.

Doguscan Ahiboz, Elina Andresen, Phillip Manley, Ute Resch-Genger, Christian Würth, Christiane Becker (2021). Metasurface-Enhanced Photon Upconversion upon 1550 nm Excitation. *Adv. Opt. Mater.*, 9:2101285. <https://doi.org/10.1002/adom.202101285>.

Kevin-Martin Aigner, Jan-Patrick Clarner, Frauke Liers, Alexander Martin (2021). Robust Approximation of Chance Constrained DC Optimal Power Flow under Decision-Dependent Uncertainty. *European Journal of Operational Research* 301(1):318-333.

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 (2021). HighPerMeshes – A Domain-Specific Language for Numerical Algorithms on Unstructured Grids. *Euro-Par 2020: Parallel Processing Workshops.*, 185–196. [https://doi.org/10.1007/978-3-030-71593-9\\_15](https://doi.org/10.1007/978-3-030-71593-9_15).

Samer Alhaddad, Jens Förstner, Stefan Groth, Daniel

Grünewald, Yevgen Grynko, Frank Hannig, Tobias Kenter, F.J. Pfreundt, Christian Plessl, Merlind Schotte, Thomas Steinke, J. Teich, Martin Weiser, Florian Wende (2021). The HighPerMeshes Framework for Numerical Algorithms on Unstructured Grids. *Concurrency and Computation: Practice and Experience*. <https://doi.org/10.1002/cpe.6616>.

Ayman Alzaatreh, Mohammad Aljarrah, Ayanna Almagambetova, Nazgul Zakiyeva (2021). On the Regression Model for Generalized Normal Distributions. *Entropy*, 23(2):173. <https://doi.org/10.3390/e23020173>.

Felix Ambellan, Stefan Zachow, Christoph von Tycowicz (2021). Geodesic B-Score for Improved Assessment of Knee Osteoarthritis. In *Proc. Information Processing in Medical Imaging (IPMI)*, *Lecture Notes in Computer Science*, pages 177–188. [https://doi.org/10.1007/978-3-030-78191-0\\_14](https://doi.org/10.1007/978-3-030-78191-0_14).

Felix Ambellan, Stefan Zachow, Christoph von Tycowicz (2021). Rigid Motion Invariant Statistical Shape Modeling based on Discrete Fundamental Forms. *Medical Image Analysis*, 73. <https://doi.org/10.1016/j.media.2021.102178>.

Anna Andrlé, Philipp Hönicke, Grzegorz Gwalt, Philipp-Immanuel Schneider, Yves Kayser, Frank Siewert, Victor Soltwisch (2021). Shape- and Element-Sensitive Reconstruction of Periodic Nanostructures with Grazing Incidence X-ray Fluorescence Analysis and Machine Learning. *Nanoma-*

*terials*, 11:1647. <https://doi.org/10.3390/nano11071647>.

Anna Andrlé, Philipp Hönicke, Grzegorz Gwalt, Philipp-Immanuel Schneider, Yves Kayser, Frank Siewert, Victor Soltwisch (2021). GIXRF and Machine Learning as Metrological Tools for Shape and Element Sensitive Reconstructions of Periodic Nanostructures. *Proc. SPIE*, 11611:116110R. <https://doi.org/10.1117/12.2586082>.

Wolfgang Bauer, Marcus Weber, Eva Diehl-Wiesenecker, Noa Galtung, Monika Prpic, Rajan Somasundaram, Rudolf Tauber, Jochen Schwenk, Patrick Micke, Kai Kappert (2021). Plasma Proteome Fingerprints Reveal Distinctiveness and Clinical Outcome of SARS-CoV-2 Infection. *Viruses*, 13(12):2456. <https://doi.org/10.3390/v13122456>.

Daniel Baum, Felix Herter, John Møller Larsen, Achim Lichtenberger, Rubina Raja (2021). Revisiting the Jerash Silver Scroll: a New Visual Data Analysis Approach. *Digital Applications in Archaeology and Cultural Heritage*, 21:e00186. <https://doi.org/10.1016/j.daach.2021.e00186>.

Tobias Baumann, Matthias Noack, Thomas Steinke (2021). Performance Evaluation and Improvements of the PoCL Open-Source OpenCL Implementation on Intel CPUs. In *IWOCL'21: International Workshop on OpenCL*. <https://doi.org/10.1145/3456669.3456698>.

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>.

Mascha Berg, Julia Plöntzke, Heike Siebert, Susanna Röblitz (2021). Modelling Oscillatory Patterns in the Bovine Estrous Cycle with Boolean Delay Equations. *Bulletin of Mathematical Biology*, 83(12):1–25. <https://doi.org/10.1007/s11538-021-00942-z>.

Timo Berthold, Thorsten Koch, Yuji Shinano (2021). MILP. Try. Repeat. *Proceedings of the 13th International Conference on the Practice and Theory of Automated Timetabling – PATAT 2021*, 2.

Mathieu Besançon, Miguel F. Anjos, Luce Brotcorne (2021). Complexity of Near-Optimal Robust Versions of Multilevel Optimization Problems. *Optimization Letters*, 15(8):2597–2610. <https://doi.org/10.1007/s11590-021-01754-9>.

Fridtjof Betz, Felix Binkowski, Sven Burger (2021). RPEXpand: Software for Riesz Projection Expansion of Resonance Phenomena. *SoftwareX*, 15:100763. <https://doi.org/10.1016/j.softx.2021.100763>.

Felix Binkowski, Tong Wu, Philippe Lalanne, Sven Burger, Alexander O. Govorov (2021). Hot Electron Generation through Near-Field Excitation of Plasmonic Nanoresonators. *ACS Photonics*, 8:1243. <https://doi.org/10.1021/acsp Photonics.1c00231>.

Andreas Bittracher, Johann Moschner, Beate Kokschn, Roland Netz, Christof Schütte (2021). Exploring the Locking Stage of NFGAILS Amyloid Fibrillation via Transition Manifold Analysis. *The European Physical Journal B*, 94. <https://doi.org/10.1140/epjb/s10051-021-00200-0>.

Andreas Bittracher, Stefan Klus, Boumediene Hamzi, Christof Schütte (2021). Dimensionality Reduction of Complex Metastable Systems via Kernel Embeddings of Transition Manifolds. *Journal of Nonlinear Science*, 31:3. <https://doi.org/10.1007/s00332-020-09668-z>.

Andreas Bittracher, Christof Schütte (2021). A Probabilistic Algorithm for Aggregating Vastly Undersampled Large Markov Chains. *Physica D: Nonlinear Phenomena*, 416:132799. <https://doi.org/10.1016/j.physd.2020.132799>.

Heide Boeth, Robert Biesen, Jeremias Hollnagel, Sandra Herrmann, Rainald M. Ehrig, Lorenzo Pelli, William R. Taylor, Georg N. Duda, Frank Buttgerit (2021). Quantification of Morning Stiffness to Assess Disease Activity and Treatment Effects in Rheumatoid Arthritis. *Rheumatology*, 60(11):5282–5291. <https://doi.org/10.1093/rheumatology/keab323>.

Horst-Holger Boltz, Alexei Sirbu, Nina Stelzer, Martin J. Lohse, Christof Schütte, Paolo Annibale (2021). Quantitative Spectroscopy of Single Molecule Interaction Times. *Optic Letters*, 46(7):1538–1541. <https://doi.org/10.1364/OL.413030>.

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>.

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>.

Ralf Borndörfer, Katrin Casel, Davis Issac, Aikaterini Niklanovits, Stephan Schwartz, Ziena Zeif (2021). Connected k-Partition of k-Connected Graphs and c-Claw-Free Graphs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2021*, volume 207 of *LIPICs*, pages 27:1–27:14. <https://doi.org/10.4230/LIPICs.APPROX/RANDOM.2021.27>.

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 Schüllendorf, Peter Schütz, Tobias Therolf, Kerstin Waas, Steffen Weider (2021). Deutsche Bahn Schedules Train Rotations Using Hypergraph Optimization. *Infomers Journal on Applied Analytics*, 51(1):42–62. <https://doi.org/10.1287/inte.2020.1069>.

Ralf Borndörfer, Ricardo Euler, Marika Karbstein (2021). Ein Graphen-basiertes Modell zur Beschreibung von Preissystemen im öffentlichen Nahverkehr. *HEUREKA 21*, volume 002/127 of *FGSV Heureka*, pages 1–15.

Ralf Borndörfer, Jonas Denißen, Simon Heller, Torsten Klug, Michael Küpper, Niels Lindner, Markus Reuther, Thomas

Schlechte, Andreas Söhlke, William Steadman (2021). Microscopic Timetable Optimization for a Moving Block System. In *ICROMA RailBeijing 2021 conference proceedings*.

Alejandro Carderera, Sebastian Pokutta, Besançon Mathieu (2021). Simple Steps are All you Need: Frank-Wolfe and Generalized Self-Concordant Functions. In *Advances in Neural Information Processing Systems 34*:5390–5401, *NeurIPS 2021*.

Alejandro Carderera, Jelena Diakonikolas, Cheuk Yin Lin, Sebastian Pokutta (2021). Parameter-free Locally Accelerated Conditional Gradients. In *ICML 2021*: 1283–1293.

Pedro Maristany de las Casas, Ralf Borndörfer, Luitgard Kraus, Antonio Sedeño-Noda (2021). An FPTAS for Dynamic Multiobjective Shortest Path Problems. *Algorithms*, 14(2):1–22. <https://doi.org/10.3390/a14020043>.

Pedro Maristany de las Casas, Antonio Sedeno-Noda, Ralf Borndörfer (2021). An Improved Multiobjective Shortest Path Algorithm. *Computers & Operations Research*, 135. <https://doi.org/10.1016/j.cor.2021.105424>.

Fatemeh Chegini, Alena Kopanicakova, Rolf Krause, Martin Weiser (2021). Efficient Identification of Scars using Heterogeneous Model Hierarchies. *EP Europace*, 23:i113-i122. <https://doi.org/10.1093/europace/evaa402>.

Fatemeh Chegini, Alena Kopanicakova, Martin Weiser, Rolf Krause (2021). Quantitative Analysis of Nonlinear Multifidelity Optimization for Inverse

Electrophysiology. In *Domain Decomposition 26 Proceedings*.

Antonia Chmiela, Elias B. Khalil, Ambros Gleixner, Andrea Lodi, Sebastian Pokutta (2021). Learning to Schedule Heuristics in Branch and Bound. In *Advances in Neural Information Processing Systems 34*:24235–24246, *NeurIPS 2021*.

Antonia Chmiela, Gonzalo Muñoz, Felipe Serrano (2021). On the implementation and strengthening of intersection cuts for QCQPs. In *Integer Programming and Combinatorial Optimization: 22nd International Conference, IPCO 2021*, 22:134–147. [https://doi.org/10.1007/978-3-030-73879-2\\_10](https://doi.org/10.1007/978-3-030-73879-2_10).

Robert Clausecker, Florian Schintke (2021). A Measure of Quality for IDA\* Heuristics. In *Proceedings of the Fourteenth International Symposium on Combinatorial Search (SoCS 2021)*, 55–63.

Nada Cvetkovic, Tim Conrad, Han Cheng Lie (2021). A Convergent Discretisation Method for Transition Path Theory for Diffusion Processes. *Multi-scale Modeling & Simulation*, 19(1):242–266. <https://doi.org/10.1137/20M1329354>.

Chengxin Dai, Anja Füllgrabe, Julianus Pfeuffer, Elizaveta M. Solovyeva, Jingwen Deng, Pablo Moreno, Selvakumar Kamatchinathan, Deepti Jaiswal Kundu, Nancy George, Silvie Fexovy, Björn Grüning, Melanie Christine Föll, Johannes Griss, Marc Vaudel, Enrique Audain, Marie Locard-Paulet, Michael Turewicz, Martin Eisenacher, Julian Uszkoreit, Tim Van Den Bossche, Veit Schwämmle, Henry Webel, Stefan Schulze, David Bouyssy, Savita Jayaram, Vinay Kumar Duggineni,

- Patroklos Samaras, Mathias Wilhelm, Meena Choi, Mingxun Wang, Oliver Kohlbacher, Alvis Brazma, Irene Papatheodorou, Nuno Bandeira, Eric W. Deutsch, Juan Antonio Vizcaino, Mingze Bai, Timo Sachsenberg, Lev I. Levitsky, Yasset Perez-Riverol (2021). A Proteomics Sample Metadata Representation for Multiomics Integration and Big Data Analysis. *Nature Communications*, 12(5854). <https://doi.org/10.1038/s41467-021-26111-3>.
- Jo Devriendt, Ambros Gleixner, Jakob Nordström (2021). Learn to Relax: Integrating 0-1 Integer Linear Programming with Pseudo-Boolean Conflict-Driven Search. *Constraints*, 26:26–55. <https://doi.org/10.1007/s10601-020-09318-x>.
- Ana Djurdjevac (2021). Linear Parabolic Problems in Random Moving Domains. *SIAM/ASA Journal on Uncertainty Quantification*, 9(2):848–879. <https://doi.org/10.1137/19M1284889>.
- Tom Doerffel, Ariane Papke, Rupert Klein, Natalia Ernst, Piotr K. Smolarkiewicz (2021). Dynamics of Tilted Atmospheric Vortices under Asymmetric Diabatic Heating. *Theoretical and Computational Fluid Dynamics*, 35(6):831–873. <https://doi.org/10.1007/s00162-021-00591-x>.
- Joshua Driesen, Ziad El-Khatib, Niklas Wulkow, Mitchell Joblin, Iskriyana Vasileva, Andreas Glöcker, Valentin Kruspel, Catherine Vogel (2021). Data-Powered Positive Deviance during the SARS-CoV-2 Pandemic – An Ecological Pilot Study of German Districts. *International Journal of Environmental Research and Public Health*, 18(9765):1–29. <https://doi.org/10.3390/ijerph18189765>.
- Leon Eifler, Ambros Gleixner (2022). A Computational Status Update for Exact Rational Mixed Integer Programming. *Mathematical Programming*. <https://doi.org/10.1007/s10107-021-01749-5>.
- Leon Eifler, Ambros Gleixner, Jonad Pulaj (2021). A Safe Computational Framework for Integer Programming applied to Chvátal’s Conjecture. *ACM Transactions on Mathematical Software* 48(2):1–12. <https://doi.org/10.1145/3485630>.
- Leon Eifler, Ambros Gleixner (2021). A Computational Status Update for Exact Rational Mixed Integer Programming. *Integer Programming and Combinatorial Optimization: 22th International Conference, IPCO 2021*.
- Ariane Ernst, Christof Schütte, Stephan Sigrist, Stefanie Winkelmann (2022). Variance of Filtered Signals: Characterization for Linear Reaction Networks and Application to Neurotransmission Dynamics. *Mathematical Biosciences*, 343:108760. <https://doi.org/10.1016/j.mbs.2021.108760> (accepted for publication 2021-10-25)
- Laura Estacio, Moritz Ehlke, Alexander Tack, Eveling Castro-Gutierrez, Hans Lamecker, Rensso Mora, Stefan Zachow (2021). Unsupervised Detection of Disturbances in 2D Radiographs. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 367–370. <https://doi.org/10.1109/ISBI48211.2021.9434091>.
- Konstantin Fackeldey, Jonas Röhm, Amir Niknejad, Surahit Chewle, Marcus Weber (2021). Analyzing Raman Spectral Data without Separability Assumption. *Journal of Mathematical Chemistry*, 3(59):575–596. <https://doi.org/10.1007/s10910-020-01201-7>.
- Sophie Fischer, Rainald Ehrig, Stefan Schäfer, Enrico Tronci, Toni Mancini, Marcel Egli, Fabian Ille, Tillmann H. C. Krüger, Brigitte Leeners, Susanna Röblitz (2021). Mathematical Modelling and Simulation Provides Evidence for New Strategies of Ovarian Stimulation. *Frontiers in Endocrinology*, 12. <https://doi.org/10.3389/fendo.2021.613048>.
- Daniela Frömberg, Felix Höfling (2021). Generalized Master Equation for First-Passage Problems in Partitioned Spaces. *Journal of Physics A: Mathematical and Theoretical*, 54:215601. <https://doi.org/10.1088/1751-8121/abf2ec>.
- Gerwin Gamrath, Markus Reuther, Thomas Schlechte, Elmar Swarat (2021). An LP-based heuristic for Inspector Scheduling. In *Proceedings of the 13th International Conference on the Practice and Theory of Automated Timetabling – PATAT 2021: Volume I*, 1:77–86.
- Patrick Gelss, Stefan Klus, Ingmar Schuster, Christof Schütte (2021). Feature Space Approximation for Kernel-based Supervised Learning. *Knowledge-Based Systems*, 221. <https://doi.org/10.1016/j.knsys.2021.106935>.
- Masoud Gholami, Florian Schintke (2021). Combining XOR and Partner Checkpointing for Resilient Multilevel Checkpoint/Restart. In *2021 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 277–288. <https://doi.org/10.1109/IPDPS49936.2021.00036>.
- Abbas Gholami, Felix Höfling, Rupert Klein, Luigi Delle Site (2021). Thermodynamic Relations at the Coupling Boundary in Adaptive Resolution Simulations for Open Systems. *Advances Theory and Simulation*, 4:2000303. <https://doi.org/10.1002/adts.202000303>.
- 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 (2021). MIPLIB 2017: Data-Driven Compilation of the 6th Mixed-Integer Programming Library. *Mathematical Programming Computation*, 13(3):443–490. <https://doi.org/10.1007/s12532-020-00194-3>.
- Christoph Gorgulla, Süleyman Çınaroğlu, Patrick D. Fischer, Konstantin Fackeldey, Gerhard Wagner, Haribabu Arthanari (2021). VirtualFlow Ants—Ultra-Large Virtual Screenings with Artificial Intelligence Driven Docking Algorithm Based on Ant Colony Optimization. *Special Issue Artificial Intelligence & Deep Learning Approaches for Structural Bioinformatics*, 22(11):5807. <https://doi.org/10.3390/ijms22115807>.
- Christoph Gorgulla, Krishna M. Padmanabha Das, Kendra E Leigh, Marco Cesputigli, Patrick D. Fischer, Zi-Fu Wang, Guilhem Tesseyre, Shreya Pandita, Alex Shnapir, Anthony Calderaio, Colin Hutcheson, Minko Gechev, Alexander Rose, Noam Lewis, Erez Yaffe, Roni Luxenburg, Henry D. Herce, Vedat Durmaz, Thanos D. Halazonetis, Konstantin Fackeldey, Justin J. Patten, Alexander Chuprina, Igor Dziuba, Alla Plekhova, Yurii Moroz, Dmytro

- Radchenko, Olga Tarkhanova, Irina Yavnyuk, Christian C. Gruber, Ryan Yust, Dave Payne, Anders M. Näär, Mark N. Namchuk, Robert A. Davey, Gerhard Wagner, Jamie Kinney, Haribabu Arthanari (2021). A Multi-Pronged Approach Targeting SARS-CoV-2 Proteins Using Ultra-Large Virtual Screening. *iScience*, 24(2):102021. <https://doi.org/10.26434/chemrxiv.12682316>.
- Carl Martin Grewe, Tuo Liu, Christoph Kahl, Hildebrandt Andrea, Stefan Zachow (2021). Statistical Learning of Facial Expressions Improves Realism of Animated Avatar Faces. *Frontiers in Virtual Reality*, 2. <https://doi.org/10.3389/frvir.2021.619811>.
- Carl Martin Grewe, Tuo Liu, Andrea Hildebrandt, Stefan Zachow (2021). The Open Virtual Mirror Framework for Enfacement Illusions – Enhancing the Sense of Agency With Avatars That Imitate Facial Expressions. *Behavior Research Methods*.
- 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>.
- Martin Hanik, Mehmet Arif Demirtaş, Mohammed Amine Gharsallaoui, Islem Rekik (2021). Predicting cognitive scores with graph neural networks through sample selection learning. *Brain Imaging and Behavior*. <https://doi.org/10.1007/s11682-021-00585-7>.
- Luzie Helfmann, Jobst Heitzig, Péter Koltai, Jürgen Kurths, Christof Schütte (2021). Statistical Analysis of Tipping Pathways in Agent-based Models. *Eur. Phys. J. Spec. Top.*, 230:3249-3271. <https://doi.org/10.1140/epjs/s11734-021-00191-0>.
- 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 Mathematical Science*, 16(1):1–32. <https://doi.org/10.2140/camcos.2021.16.1>.
- Jessica Hembus, Felix Ambellan, Stefan Zachow, Rainer Bader (2021). Establishment of a Rolling-sliding Test Bench to Analyze Abrasive Wear Propagation of Different Bearing Materials for Knee Implants. *Applied Sciences*, 11(4). <https://doi.org/10.3390/app11041886>.
- Gregor Hendel, Daniel Anderson, Pierre Le Bodic, Marc Pfetsch (2021). Estimating the Size of Branch-And-Bound Trees. *INFORMS Journal on Computing*, 34(2):671–1304. <https://doi.org/10.1287/ijoc.2021.1103>.
- Felix Hennings, Lovis Anderson, Kai Hoppmann-Baum, Mark Turner, Thorsten Koch (2021). Controlling Transient Gas Flow in Real-world Pipeline Intersection Areas. *Optimization and Engineering*, 22:687–734. <https://doi.org/10.1007/s11081-020-09559-y>.
- Felix Hennings (2021). Large-scale Empirical Study on the Momentum Equation's Inertia term. *Journal of Natural Gas Science and Engineering*, 95:104153. <https://doi.org/10.1016/j.jngse.2021.104153>.
- Felix Herter, Hans-Christian Hege, Markus Hadwiger, Verena Lepper, Daniel Baum (2021). Thin-Volume Visualization on Curved Domains. *Computer Graphics Forum*, 40(3):147–157. <https://doi.org/10.1111/cgf.14296>.
- Benjamin Hiller, René Saitenmacher, Tom Walther (2021). Improved Models for Operation Modes of Complex Compressor Stations. *Mathematical Methods of Operations Research*, 94:171–195. <https://doi.org/10.1007/s00186-021-00745-x>.
- Kai Hoppmann-Baum, Felix Hennings, Ralf Lenz, Uwe Gotzes, Nina Heinecke, Klaus Spreckelsen, Thorsten Koch (2021). Optimal Operation of Transient Gas Transport Networks, 236–248. *Optimization and Engineering*, 22:735–781. <https://doi.org/10.1007/s11081-020-09584-x>.
- Kai Hoppmann-Baum (2021). On the Complexity of Computing Maximum and Minimum Min-Cost-Flows. *Networks*. <https://doi.org/10.1002/net.22060> (epub ahead of print 2021-05-25).
- Klaus-Dieter Hungenberg, Christian Schwede, Michael Wulkow, Niklas Wulkow (2021). Determination of Reactivity ratios for Acrylic Acid and its dimer from Classical Parameter Estimation and Bayesian Approach. *Canadian Journal of Chemical Engineering*, 1–10. <https://doi.org/10.1002/cjce.24330>.
- Sahar Irvani, Tim Conrad (2021). An Interpretable Deep Learning Approach for Biomarker Detection in LC-MS Proteomics Data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. <https://doi.org/10.1109/tcbb.2022.3141656> (epub ahead of print 2021-12-30).
- 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>.
- Klaus Jäger, Johannes Sutter, Martin Hamerschmidt, Philipp-Immanuel Schneider, Christiane Becker (2021). Prospects of Light Management in Perovskite/Silicon Tandem Solar Cells. *Nanophotonics*, 10:1991. <https://doi.org/10.1515/nanoph-2020-0674>.
- Ilja Klebanov, Björn Sprungk, T. J. Sullivan (2021). The Linear Conditional Expectation in Hilbert space. *Bernoulli*, 27(4):2299–2299. <https://doi.org/10.3150/20-BEJ1308>.
- Ilja Klebanov, Alexander Sikorski, Christof Schütte, Susanna Röblitz (2021). Objective Priors in the Empirical Bayes Framework. *Scandinavian Journal of Statistics*, 48(4):1212–1233. <https://doi.org/10.1111/sjos.12485>.
- Rupert Klein, Roya Ebrahimi Viand, Felix Höfling, Luigi Delle Site (2021). Nonequilibrium Induced by Reservoirs: Physico-mathematical Model and Numerical Tests. *Advances Theory and Simulation*, 4:2100071. <https://doi.org/10.1002/adts.202100071>.
- Margarita Kostre, Christof Schütte, Frank Noé, Mauricio del Razo Sarmina (2021). Coupling Particle-Based Reaction-Diffusion Simulations with Reservoirs Mediated by Reaction-Diffusion PDEs. *Multi-*

scale Modeling & Simulation, 19(4):1659–1683. <https://doi.org/10.1137/20M1352739>.

Margarita Kostre, Vikram Sunkara, Natasa Djurdjevic Conrad (2021). Inference of Historical Influence Networks. In *International Conference on Complex Networks & Their Applications Proceedings*, 110–113.

Rolf Krause, Martin Weiser (2021). Multilevel Augmented Lagrangian Solvers for Over-constrained Contact Formulations. In *ESAIM: ProcS*, 71:175–184. <https://doi.org/10.1051/proc/202171175>.

Ina Lantzsch, Che-Hang Yu, Yu-Zen Chen, Vitaly Zimyanin, Hossein Yazdkhasti, Norbert Lindow, Erik Szentgoeryi, Ariel M Pani, Steffen Prohaska, Martin Srayko, Sebastian Fürthauer, Stefanie Redemann (2021). Microtubule Reorganization During Female Meiosis in *C. Elegans*. *eLife*, 10:e58903. <https://doi.org/10.7554/eLife.58903>.

Thi Thai Le, Thorsten Koch (2021). Instability of a Tangential Discontinuity Surface in a Three-dimensional Compressible Medium. *The Physics of Fluids*, 33(1):016106. <https://doi.org/10.1063/5.0033753>.

Thi Thai Le (2021). Effect of Water Depth on Kelvin – Helmholtz instability in a Shallow-water Flow. *Journal of Mathematical Physics*, 62(10). <https://doi.org/10.1063/1.5145060>.

Thi Thai Le, Thorsten Koch (2021). Interface Stability of Compressible Fluids in Porous Media. *The Physics of Fluids*, 33(8):084102. <https://doi.org/10.1063/5.0059336>.

Ralf Lenz, Kai-Helge Becker (2022). Optimization of Capacities

Expansion in Potential-driven Networks including Multiple Looping – A comparison of Modelling Approaches. *OR Spectrum*, 44:179–224.

Ulf Leser, Marcus Hilbrich, Claudia Draxl, Peter Eisert, Lars Grunke, Patrick Hostert, Dagmar Kainmüller, Odej Kao, Birte Kehr, Timo Kehrer, Christoph Koch, Volker Markl, Henning Meyerhenke, Tilmann Rabl, Alexander Reinefeld, Knut Reinert, Kerstin Ritter, Björn Scheuermann, Florian Schintke, Nicole Schweikardt, Matthias Weidlich (2021). The Collaborative Research Center FONDA. *Datenbank-Spektrum*, 21(3):255–260. <https://doi.org/10.1007/s13222-021-00397-5>.

Jianning Li, Pedro Pimentel, Angelika Szengel, Moritz Ehlke, Hans Lamecker, Stefan Zachow, Laura Estacio, Christian Doenitz, Heiko Ramm, Haochen Shi, Xiaojun Chen, Franco Matzkin, Virginia Newcombe, Enzo Ferrante, Yuan Jin, David G. Ellis, Michele R. Aizenberg, Oldrich Kodym, Michal Spanel, Adam Herout, James G. Mainprize, Zachary Fishman, Michael R. Hardisty, Amirhossein Bayat, Suprosanna Shit, Bomin Wang, Zhi Liu, Matthias Eder, Antonio Pepe, Christina Gsaxner, Victor Alves, Ulrike Zefferer, Cord von Campe, Karin Pistracher, Ute Schäfer, Dieter Schmalstieg, Bjoern H. Menze, Ben Glocker, Jan Egger (2021). Autolplant 2020 – First MICCAI Challenge on Automatic Cranial Implant Design. *IEEE Transactions on Medical Imaging*, 40(9):2329–2342. <https://doi.org/10.1109/TMI.2021.3077047>.

YongTian Liang, Chengji Piao, Christine B. Beuschel, David Toppe, Laxmikanth Kollipara, Boris Bogdanow, Marta Ma-

glione, Janine Lützkendorf, Jason Chun Kit See, Sheng Huang, Tim Conrad, Ulrich Kintscher, Frank Madeo, Fan Liu, Albert Sickmann, Stephan J. Sigrist (2021). eIF5A Hypusination, Boosted by Dietary Spermidine, Protects from Premature Brain Aging and Mitochondrial Dysfunction. *Cell Reports*, 35(2):108941. <https://doi.org/10.1016/j.celrep.2021.108941>.

Han Cheng Lie, T. J. Sullivan, Aretha Teckentrup (2021). Error Bounds for some Approximate Posterior Measures in Bayesian Inference. *Numerical Mathematics and Advanced Applications ENUMATH 2019*, 275–283. [https://doi.org/10.1007/978-3-030-55874-1\\_26](https://doi.org/10.1007/978-3-030-55874-1_26).

Niels Lindner, Christian Liebchen, Berenike Masing (2021). Forward Cycle Bases and Periodic Timetabling. In *21st Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2021)*, volume 96 of *Open Access Series in Informatics (OASlcs)*, pages 2:1–2:14. <https://doi.org/10.4230/OASlcs.ATMOS.2021.2>.

Niels Lindner, Pedro Maristany de las Casas, Philine Schiewe (2021). Optimal Forks: Preprocessing Single-Source Shortest Path Instances with Interval Data. In *21st Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2021)*, volume 96 of *Open Access Series in Informatics (OASlcs)*, pages 7:1–7:15. <https://doi.org/10.4230/OASlcs.ATMOS.2021.7>.

Norbert Lindow, Florian N. Brüning, Vincent J. Dercksen, Gunar Fabig, Robert Kiewisz, Stefanie Redemann, Thomas Müller-Reichert, Steffen Pro-

haska, Daniel Baum (2021). Semi-automatic Stitching of Filamentous Structures in Image Stacks from Serial-section electron Tomography. *Journal of Microscopy*, 284(1):25–44. <https://doi.org/10.1111/jmi.13039>.

Yongtao Liu, Zhiguang Zhou, Fan Wang, Günter Kewes, Shihui Wen, Sven Burger, Majid Ebrahimi Wakiani, Peng Xi, Jiong Yang, Xusan Yang, Oliver Benson, Dayong Jin (2021). Axial Localization and Tracking of Self-interference Nanoparticles by Lateral Point Spread Functions. *Nat. Commun.*, 12:2019. <https://doi.org/10.1038/s41467-021-22283-0>.

Phillip Manley, Michele Segantini, Doguscan Ahiboz, Martin Hammerschmidt, Georgios Arnaoutakis, Rowan W. MacQueen, Sven Burger, Christiane Becker (2021). Double-layer Metasurface for Enhanced Photon Up-conversion. *APL Photon.*, 6:036103. <https://doi.org/10.1063/5.0040839>.

Kevin Martens, Felix Binkowski, Linh Nguyen, Li Hu, Alexander O. Govorov, Sven Burger, Tim Liedl (2021). Long- and Short-Ranged Chiral Interactions in DNA Assembled Plasmonic Chains. *Nat. Commun.*, 12:2025. <https://doi.org/10.1038/s41467-021-22289-8>.

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*, 33:611–637. <https://doi.org/10.1007/s00161-020-00932-x>.

Philippe Miron, Francisco Beron-Vera, Luzie Helfmann,

Péter Koltai (2021). Transition Paths of Marine Debris and the Stability of the Garbage Patches. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(3). <https://doi.org/10.1063/5.0030535>.

Philippe Moewis, René Kaiser, Adam Treczynski, Christoph von Tycowicz, Leonie Krahl, Ansgar Ilg, Johannes Holz, Georg N. Duda (2021). Patient Specific Resurfacing Implant Knee Surgery in Subjects with Early Osteoarthritis Results in Medial Pivot and Lateral Femoral Rollback During Flexion: A Retrospective Pilot Study. *Knee Surgery, Sports Traumatology, Arthroscopy*. <https://doi.org/10.1007/s00167-021-06749-8>.

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>.

Jan-Hendrik Niemann, Stefan Klus, Christof Schütte (2021). Data-driven Model Reduction of Agent-based Systems Using the Koopman Generator. *PLOS ONE*, 16(5). <https://doi.org/10.1371/journal.pone.0250970>.

Jan-Hendrik Niemann, Stefanie Winkelmann, Sarah Wolf, Christof Schütte (2021). Agent-based Modeling: Population Limits and Large Timescales. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(3). <https://doi.org/10.1063/5.0031373>.

Isabel Nitzke, Konstantin Fackeldey, Jadran Vrabec (2021). Long Range Corrections for Inhomogeneous Fluids Con-

taining a Droplet or a Bubble. *Molecular Simulation*, 1–14. <https://doi.org/10.1080/08927022.2021.1954639>.

Anton V. Pakhomov, Martin Hammerschmidt, Sven Burger, Thomas Pertsch, Frank Setzpfand (2021). Modeling of Surface-induced Second-harmonic Generation from Multilayer Structures by the Transfer Matrix Method. *Opt. Express*, 29:9098. <https://doi.org/10.1364/OE.417066>.

Yogyata Pathania, Dipanjan Chakraborty, Felix Höfling (2021). Continuous Demixing Transition of Binary Liquids: Finite-size Scaling from the Analysis of Sub-systems. *Advances Theory and Simulation*, 4:2000235. <https://doi.org/10.1002/adts.202000235>.

Jaap Pedersen, Birgit Schachler, Anya Heider, Guido Pleßmann (2021). Distribution System Planning with Battery Storage using Multiperiod Optimal Power Flow. In *14th International Renewable Energy Storage Conference 2020 (IRES 2020)*, volume 6 of *Atlantis Highlights in Engineering*, pages 41–48. <https://doi.org/10.2991/ahe.k.210202.007>.

Jaap Pedersen, Kai Hopmann-Baum, Janina Zittel, Thorsten Koch (2021). Blending Hydrogen into Natural Gas: An Assessment of the Capacity of the German Gas Grid. In *Operations Research Proceedings 2021*.

Felix Peppert, Max von Kleist, Christof Schütte, Vikram Sun-kara (2021). On the Sufficient Condition for Solving the Gap-Filling Problem Using Deep Convolutional Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*. <https://doi.org/10.1109/TNN-LS.2021.3072746>.

[doi.org/10.1109/TNN-LS.2021.3072746](https://doi.org/10.1109/TNN-LS.2021.3072746).

Dmitrii Perevoznik, Surajit Bose, Sven Burger, Ayhan Demircan, Uwe Morgner (2021). Writing Photonic Components in Polymers Using Femtosecond Pulses. In *Conference on Lasers and Electro-Optics (CLEO), OSA Technical Digest*, page ATH4P.4. [https://doi.org/10.1364/CLEO\\_AT.2021.ATH4P.4](https://doi.org/10.1364/CLEO_AT.2021.ATH4P.4).

Dmitrii Perevoznik, Surajit Bose, Sven Burger, Ayhan Demircan, Uwe Morgner (2021). Photonic Components in Polymers Made by Femtosecond Pulses. In *Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*, 1. <https://doi.org/10.1109/CLEO/Europe-EQEC52157.2021.9541906>.

Milena Petkovic, Ying Chen, Inken Gamrath, Uwe Gotzes, Natalia Selini Hadjidimitrou, Janina Zittel, Xiaofei Xu, Thorsten Koch (2022). A Hybrid Approach for High Precision Prediction of Gas Flows. *Energy Systems*, 13:383–408. <https://doi.org/10.1007/s12667-021-00466-4>.

Milena Petkovic, Thorsten Koch, Janina Zittel (2021). Deep Learning for Spatio-temporal Supply Anddemand forecasting in Natural Gas Transmission Networks. *Energy Science and Engineering*. <https://doi.org/10.1002/ese3.932> (epub ahead of print 2021-06-28).

Thomas Picht, Maxime Le Calve, Rosario Tomasello, Lucius Fekonja, Mohammad Fardin Gholami, Matthias Bruhn, Carola Zwick, Jürgen P. Rabe, Claudia Müller-Birn, Peter Vajkoczy, Igor M. Sauer, Stefan Zachow, John A. Nyakatura,

Patricia Ribault, Friedemann Pulvermüller (2021). A Note on Neurosurgical Resection and Why we Need to Rethink Cutting. *Neurosurgery*, 89(5):289–291. <https://doi.org/10.1093/neuros/nyab326>.

Matthias Plock, Sven Burger, Philipp-Immanuel Schneider (2021). Recent Advances in Bayesian Optimization with Applications to Parameter Reconstruction in Optical Nano-metrology. *Proc. SPIE*, 11783:117830J. <https://doi.org/10.1117/12.2592266>.

Felix Prause, Kai Hopmann-Baum, Boris Defourny, Thorsten Koch (2021). The Maximum Diversity Assortment Selection Problem. *Mathematical Methods of Operations Research*, 93:521–554. <https://doi.org/10.1007/s00186-021-00740-2>.

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>.

Alexia N. Raharinirina, Felix Peppert, Max von Kleist, Christof Schütte, Vikram Sun-kara (2021). Inferring Gene Regulatory Networks from Single-cell RNA-seq Temporal Snapshot Data Requires Higher-order Moments. *Patterns*, 2(9). <https://doi.org/10.1016/j.patter.2021.100332>.

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. *Sci-*

ence of the Total Environment, 751. <https://doi.org/10.1016/j.scitotenv.2020.141706>.

Mona Rams, Tim Conrad (2022). Dictionary Learning Allows Model-free Pseudotime Estimation of Transcriptomics Data. *BMC Genomics*, 23:56. 1. <https://doi.org/10.1186/s12864-021-08276-9>.

Mauricio J. del Razo, Manuel Dibak, Christof Schütte, Frank Noé (2021). Multiscale Molecular Kinetics by Coupling Markov State Models and Reaction-diffusion Dynamics. *The Journal of Chemical Physics*, 155(12). <https://doi.org/10.1063/5.0060314>.

Daniel Rehfeldt, Thorsten Koch (2021). Implications, Conflicts, and Reductions for Steiner Trees. *Mathematical Programming*. <https://doi.org/10.1007/s10107-021-01757-5> (epub ahead of print 2021-12-30).

Daniel Rehfeldt, Hannes Hobbie, David Schönheit, Thorsten Koch, Dominik Möst, Ambros Gleixner (2021). A Massively Parallel Interior-point Solver for LPs with Generalized Arrowhead Structure, and Applications to Energy System Models. *European Journal of Operational Research*, 296(1):60–71. <https://doi.org/10.1016/j.ejor.2021.06.063>.

Daniel Rehfeldt, Thorsten Koch (2022). On the Exact Solution of Prize-collecting Steiner Tree Problems. *INFORMS Journal on Computing*, 34(2):872–889. <https://doi.org/10.1287/ijoc.2021.1087>.

Daniel Rehfeldt, Thorsten Koch (2021). Implications, Conflicts, and Reductions for Steiner Trees. In *Integer Programming and Combinatorial Optimization: 22th International Con-*

*ference, IPCO 2021*, 473–487. [https://doi.org/10.1007/978-3-030-73879-2\\_33](https://doi.org/10.1007/978-3-030-73879-2_33).

Lucas Rickert, Timm Gao, Felix Urban, Jan Große, Nicole Srocka, Sven Rodt, Anna Musial, Kinga Zolnacz, Pawel Mergo, Kamil Dybka, Waclaw Urbanczyk, Grzegorz Sek, Sven Burger, Stephan Reitzenstein, Heindel Tobias (2021). A Quantum Key Distribution Testbed Using Plug&Play Telecom-Wavelength Single-Photons. In *Quantum Information and Measurement VI, OSA Technical Digest*, page M2C.1. <https://doi.org/10.1364/QIM.2021.M2C.1>.

Susanne Röhl, Marcus Weber, Konstantin Fackeldey (2021). Computing the Minimal Rebinding Effect for Non-reversible Processes. *Multiscale Modeling and Simulation*, 19(1):460–477. <https://doi.org/10.1137/20M1334966>.

Guillaume Sagnol, Daniel Schmidt genannt Waldschmidt (2021). Restricted Adaptivity in Stochastic Scheduling. *29th Annual European Symposium on Algorithms (ESA 2021)*, volume 204 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 79:1–79:14. <https://doi.org/10.4230/LIPIcs.ESA.2021.79>.

Guvenc Sahin, Amin Ahmadi Digehsara, Ralf Borndörfer (2021). Efficient Algorithms for the Multi-Period Line Planning Problem in Public Transportation. *Open Access Series in Informatics (OASICS), 21st Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2021)*, 96:17:1–17:6. <https://doi.org/10.4230/OASICS.ATMOS.2021.17>.

Manish Sahu, Anirban Mukhopadhyay, Stefan Zachow (2021). Simulation-to-Real Domain Adaptation with Teacher-student Learning for Endoscopic Instrument Segmentation. *International Journal of Computer Assisted Radiology and Surgery*, 16:849–859. <https://doi.org/10.1007/s11548-021-02383-4>.

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>.

Anton Schiela, Matthias Stöcklein, Martin Weiser (2021). A Primal Dual Projection Algorithm for Efficient Constraint Preconditioning. *SIAM Journal on Scientific Computing*.

Peter Schnauber, Johannes Schall, Samir Bounouar, Kartik Srinivasan, Marcelo Davanco, Jin-Dong Song, Sven Burger, Sven Rodt, Stephan Reitzenstein (2021). Deterministically Fabricated Quantum Dot – Waveguide Systems for On-Chip Quantum Optics. In *Proc. SPIE*, 11806:118060G. <https://doi.org/10.1117/12.2596042>

Jakob Schneck, Martin Weiser, Florian Wende (2021). Impact of Mixed Precision and Storage Layout on Additive Schwarz Smoothers. *Numerical Linear Algebra with Applications*, 28(4):e2366.

Philipp-Immanuel Schneider, Xavier Garcia-Santiago, Benjamin Wiegand, Oliver Anton, Markus Krutzik, Carsten Rockstuhl, Sven Burger (2021). Advances in Bayesian Optimization for Photonics and Quan-

tum Atom Optics Applications. In *OSA Advanced Photonics Congress, OSA Technical Digest*, page JTh1E.2. <https://doi.org/10.1364/IPRSN.2021.JTh1E.2>.

Florian Schäfer, T. J. Sullivan, Houman Owhadi (2021). Compression, Inversion, and Approximate PCA of Dense Kernel Matrices at Near-linear Computational Complexity. *Multiscale Modeling and Simulation*, 19(2):688–730. <https://doi.org/10.1137/19M129526X>.

Renata Sechi, Marcus Weber, Karsten Heyne (2021). MSM Building and Projection for the Analysis of Time-resolved Spectra. *Proceedings in Applied Mathematics and Mechanics*, e202100102. <https://doi.org/10.1002/pamm.202100102>.

Renata Sechi, Alexander Sikorski, Marcus Weber (2021). Estimation of the Koopman Generator by Newton’s Extrapolation. *Multiscale Modeling and Simulation*, 19(2):758–774. <https://doi.org/10.1137/20M1333006>.

Anjany Sekuboyina, Malek E. Hussein, Amirhossein Bayat, Maximilian Löffler, Hans Liebl, Hongwei Li, Giles Tetteh, Jan Kukačka, Christian Payer, Darko Štern, Martin Urschler, Mao-dong Chen, Dalong Cheng, Nikolas Lessmann, Yujin Hu, Tianfu Wang, Dong Yang, Daguang Xu, Felix Ambellan, Tamaz Amiranashvili, Moritz Ehlke, Hans Lamecker, Sebastian Lehnert, Marilia Lirio, Nicolás Pérez de Olague, Heiko Ramm, Manish Sahu, Alexander Tack, Stefan Zachow, Tao Jiang, Xinjun Ma, Christoph Angerman, Xin Wang, Kevin Brown, Alexandre Kirszenberg, Élodie Puybareau, Di Chen, Yiwei Bai, Brandon H. Rapazzo, Timyoas Yeah, Amber Zhang, Shangliang Xu, Feng Hou, Zhiqiang He, Chan Zeng,



Zheng Xiangshang, Xu Liming, Tucker J. Netherton, Raymond P. Mumme, Laurence E. Court, Zixun Huang, Chenhang He, Li-Wen Wang, Sai Ho Ling, Lê Duy Huynh, Nicolas Boutry, Roman Jakubicek, Jiri Chmelik, Supriti Mulay, Mohanasankar Sivaprakasam, Johannes C. Paetzold, Suprosanna Shit, Ivan Ezhov, Benedikt Wiestler, Ben Glocker, Alexander Valentinitich, Markus Rempfler, Björn H. Menze, Jan S. Kirschke (2021). VerSe: A Vertebrae Labelling and Segmentation Benchmark for Multi-detector CT Images. *Medical Image Analysis*, 73. <https://doi.org/10.1016/j.media.2021.102166>.

Upanshu Sharma, Wei Zhang (2021). Non-reversible Sampling Schemes on Submanifolds. *SIAM Journal on Numerical Analysis*, 59(6):2989–3031. <https://doi.org/10.1137/20M1378752>.

Alexander Sikorski, Marcus Weber, Christof Schütte (2021). The Augmented Jump Chain. *Advanced Theory and Simulations*, 4. <https://doi.org/10.1002/adts.202000274>.

Ivan Sipiran, Patrick Lazo, Cristian Lopez, Nihar Bagewadi, Benjamin Bustos, Hieu Dao, Shankar Gangisetty, Martin Hanik, Ngoc-Phuong Ho-Thi, Mike Holenderski, Dmitri Jarnikov, Arniel Labrada, Stefan Lengauer, Roxane Licandro, Dinh-Huan Nguyen, Thang-Long Nguyen-Ho, Luis A. Pérez Rey, Bang-Dang Pham, Minh-Khoi Pham, Reinhold Preiner, Tobias Schreck, Quoc-Huy Trinh, Loek Tonnaer, Christoph von Tycowicz, The-Anh Vu-Le (2021). SHREC 2021: Retrieval of Cultural Heritage Objects. *Computers and Graphics*, 100:1–20. <https://doi.org/10.1016/j.cag.2021.07.010>

Boro Sofranac, Ambros Gleixner, Sebastian Pokutta (2021). An Algorithm-Independent Measure of Progress for Linear Constraint Propagation. In *27th International Conference on Principles and Practice of Constraint Programming (CP 2021)*, volume 210 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 52:1–52:17. <https://doi.org/10.4230/LIPIcs.CP.2021.52> (epub ahead of print 2021-10-15).

Brian Stout, Remi Colom, Nicolas Bonod, Ross C McPhedran (2021). Spectral Expansions of Open and Dispersive Optical Systems: Gaussian Regularization and Convergence. *New J. Phys.*, 23:083004. <https://doi.org/10.1088/1367-2630/ac10a6>.

Arthur Straube, Stefanie Winkelmann, Christof Schütte, Felix Höfling (2021). Stochastic pH Oscillations in a Model of the Urea-urease Reaction Confined to Lipid Vesicles. *J. Phys. Chem. Lett.*, 12:9888–9893. <https://doi.org/10.1021/acs.jpcclett.1c03016>.

Vikram Sunkara, Gitta A. Heinz, Frederik F. Heinrich, Pawel Durek, Ali Mobasheri, Mir-Farzin Mashreghi, Annemarie Lang (2021). Combining Segmental Bulk- and Single-cell RNA-sequencing to Define the Chondrocyte Gene Expression Signature in the Murine Knee Joint. *Osteoarthritis and Cartilage*, 29(6):905–914. <https://doi.org/10.1016/j.joca.2021.03.007>.

Alexander Tack, Bernhard Preim, Stefan Zachow (2021). Fully automated Assessment of Knee Alignment from Full-Leg X-Rays employing a “YOLOv4 And Resnet Landmark regression Algorithm” (YARLA): Data from the Osteoarthritis Initiative. *Com-*

*puter Methods and Programs in Biomedicine*, 205(106080). <https://doi.org/10.1016/j.cmpb.2021.106080>.

Alexander Tack, Felix Ambellan, Stefan Zachow (2021). Towards Novel Osteoarthritis Biomarkers: Multi-criteria Evaluation of 46,996 Segmented Knee MRI Data from the Osteoarthritis Initiative. *PLOS One*, 16(10). <https://doi.org/10.1371/journal.pone.0258855>.

Alexander Tack, Alexey Shestakov, David Lüdke, Stefan Zachow (2021). A Deep Multi-task Learning Method for Detection of Meniscal Tears in MRI Data from the Osteoarthritis Initiative Database. *Frontiers in Bioengineering and Biotechnology, section Biomechanics*.

Nariaki Tateiwa, Yuji Shinano, Keiichiro Yamamura, Akihiro Yoshida, Shizuo Kaji, Masaya Yasuda, Katsuki Fujisawa (2021). CMAP-LAP: Configurable Massively Parallel Solver for Lattice Problems. In *EEE 28th International Conference on High Performance Computing, Data, and Analytics (HiPC)*, pages 42–52. <https://doi.org/10.1109/HiPC53243.2021.00018>.

Christine Tawfik, Bernard Gendron, Sabine Limbourg (2021). An Iterative Two-stage Heuristic Algorithm for a Bilevel Service Network Design and Pricing Model. *European Journal of Operational Research*, 300(2):512–526. <https://doi.org/10.1016/j.ejor.2021.07.052>.

Lasse Thostrup, Jan Skrzypczak, Matthias Jasny, Tobias Ziegler, Carsten Binnig (2021). DFI: The Data Flow Interface for High-Speed Networks. In *SIGMOD '21: International Conference on Management of*

*Data*, 1825–1837. <https://doi.org/10.1145/3448016.3452816>.

Pietro Tierno, Tom H. Johansen, Arthur Straube (2021). Thermally Active Nanoparticle Clusters Enslaved by Engineered Domain Wall Traps. *Nature Commun.*, 12:5813. <https://doi.org/10.1038/s41467-021-25931-7>.

Peter Tillmann, Klaus Jäger, Eugene Katz, Christiane Becker (2021). Relaxed Current-matching Constraints by Bifacial Operation and Luminescent Coupling in Perovskite/Silicon Tandem solar cells. In *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)*, 1086. <https://doi.org/10.1109/PVSC43889.2021.9518891>.

Peter Tillmann, Benedikt Bläsi, Sven Burger, Martin Hammerschmidt, Oliver Höhn, Christiane Becker, Klaus Jäger (2021). Optimizing Metal Grating Back Reflectors for III-V-on-silicon Multijunction Solar Cells. *Opt. Express*, 29:22517. <https://doi.org/10.1364/OE.426761>.

Cem Güney Torun, Philipp-Immanuel Schneider, Martin Hammerschmidt, Sven Burger, Joseph H.D. Munns, Tim Schröder (2021). Optimized Diamond Inverted Nanocones for Enhanced Color Center to Fiber Coupling. *Appl. Phys. Lett.*, 118:234002. <https://doi.org/10.1063/5.0050338>.

Husen M. Umer, Yafeng Zhu, Julianus Pfeuffer, Timo Sachsenberg, Janne Lehtiö, Rui Branca, Yasset Perez-Riverol (2021). Generation of ENSEMBL-based Proteogenomics Databases Boosts the Identification of Non-canonical Peptides. *Bioinformatics*, 1470–1472.

Sumit Kumar Vohra, Dimiter Prodanov (2021). The Active

Segmentation Platform for Microscopic Image Classification and Segmentation. *Brain Sciences*. 2021, 11(12):1645. <https://doi.org/10.3390/brainsci11121645>.

Junyang Wang, Jon Cockayne, Oksana Chkrebti, T. J. Sullivan, Chris Oates (2021). Bayesian Numerical Methods for Non-linear Partial Differential Equations. *Statistics and Computing*, 31:55. <https://doi.org/10.1007/s11222-021-10030-w>.

K. Weimann, Tim Conrad (2021). Transfer Learning for ECG Classification. *Scientific Reports*, 11:5251. <https://doi.org/10.1038/s41598-021-84374-8>.

Stefanie Winkelmann, Johannes Zonker, Christof Schütte, Natasa Djurdjevac Conrad (2021). Mathematical Modeling of Spatio-temporal Population Dynamics and Application to Epidemic Spreading. *Mathematical Biosciences*, 336:108619. <https://doi.org/10.1016/j.mbs.2021.108619>.

Jakob Witzig, Timo Berthold (2021). Conflict Analysis for

MINLP. *INFORMS Journal on Computing*, 33(2):421–435. <https://doi.org/10.1287/ijoc.2020.1050>.

Tong Wu, Denis Arrivault, Marc Duruflé, Alexandre Gras, Felix Binkowski, Sven Burger, Wei Yan, Philippe Lalanne (2021). Efficient Hybrid Method for the Modal Analysis of Optical Microcavities and Nanoresonators. *J. Opt. Soc. Am. A*, 38:1224. <https://doi.org/10.1364/JOSAA.428224>.

Hanna Wulkow, Tim Conrad, Natasa Djurdjevac Conrad, Sebastian A. Müller, Kai Nagel, Christof Schütte (2021). Prediction of Covid-19 Spreading and Optimal Coordination of Counter-measures: From Microscopic to Macroscopic Models to Pareto Fronts. *PLOS One*, 16(4). <https://doi.org/10.1371/journal.pone.0249676>.

Niklas Wulkow, Péter Koltai, Christof Schütte (2021). Memory-Based Reduced Modelling and Data-Based Estimation of Opinion Spreading. *Journal of Nonlinear Science*, 31:19. <https://doi.org/10.1007/s00332-020-09673-2>.

Niklas Wulkow, Regina Telgmann, Klaus-Dieter Hungenberg, Christof Schütte, Michael Wulkow (2021). Deterministic and Stochastic Parameter Estimation for Polymer Reaction Kinetics I: Theory and Simple Examples. *Macromolecular Theory and Simulations*, 30:2100017. <https://doi.org/10.1002/mats.202100017>.

Xiaofei Xu, Ying Chen, Ge Zhang, Thorsten Koch (2021). Modeling Functional Time Series and Mixed-type Predictors with Partially Functional Autoregressions. *Journal of Business and Economic Statistics*, 1–43. <https://doi.org/10.1080/0735015.2021.2011299>.

Ryohei Yokoyama, Hiroki Kamada, Yuji Shinano, Tetsuya Wakui (2021). A Hierarchical Optimization Approach to Robust Design of Energy Supply Systems Based on a Mixed-integer Linear Model. *Energy*, 229:120343. <https://doi.org/10.1016/j.energy.2021.120343>.

Ryohei Yokoyama, Kotaro Takeuchi, Yuji Shinano, Tetsuya Wakui (2021). Effect of Model Reduction by Time Aggrega-

tion in Multiobjective Optimal Design of Energy Supply Systems by a Hierarchical MILP Method. *Energy*, 228:120505. <https://doi.org/10.1016/j.energy.2021.120505> (epub ahead of print 2021-04-19).

Wei Zhang (2021). Some New Results on Relative Entropy Production, Time Reversal, and Optimal Control of Time-inhomogeneous Diffusion Processes. *Journal of Mathematical Physics*, 62(4):043302. <https://doi.org/10.1063/5.0038740>.

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>.

Boro Šofranac, Ambros Gleixner, Sebastian Pokutta (2022). Accelerating Domain Propagation: An Efficient GPU-parallel Algorithm over Sparse Matrices. *Parallel Computing*, 109:102874. <https://doi.org/10.1016/j.parco.2021.102874> (epub ahead of print 2021-11-26).

## BOOKS (1)

Mathias Lemke, Lewin Stein, Arne Hölder, Stefan Weinzierl (2021). *Adjungierten-basierte Synthese und Bestimmung optimaler Treiberfunktionen von Lautsprechern mit komplexer Richtcharakteristik*. Deutsche Gesellschaft für Akustik e.V.

## DISSERTATIONS (5)

Ralf Lenz (2021). *Optimization of Stationary Expansion Planning and Transient Network Control by Mixed-Integer Nonlinear Programming*. Technische Universität Berlin.

Marian Moldenhauer (2021). *Adaptive Algorithms in Optimization under PDE Constraints*. Freie Universität Berlin.

Anton Pakhomov (2021). *Efficient Modeling and Optimization of Surface Second-harmonic Generation from Nanophotonic Components*. Friedrich-Schiller-Universität Jena.

Daniel Rehfeldt (2021). *Faster Algorithms for Steiner Tree and Related Problems: From*

*Theory to Practice*. Technische Universität Berlin.

Xavier Garcia Santiago (2021). *Numerical Methods for Shape Optimization of Photonic Nanostructures*. Karlsruher Institut für Technologie.

# IMPRINT

## **Published by**

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

**Phone** +49 30 84185-0

**Fax** +49 30 84185-125

**Email** [info@zib.de](mailto:info@zib.de)

[www.zib.de](http://www.zib.de)

## **Layout**

Christiane Weimann Grafik + Produktion  
[www.schnittmarke.biz](http://www.schnittmarke.biz)

## **Printing**

Lochmann grafische Produktion GmbH

## **ISSN**

0934-5892

©2021 Konrad-Zuse-Zentrum für Informationstechnik Berlin







