Richard Schulze

Einsteinstraße 62
48149 Münster, Germany
☎ *+49 (0) 251 83-32744*
✉ *r.schulze@uni-muenster.de*

University Education

since 2018 **Ph.D. studies**, *University of Münster*, Münster, Germany.
Supervisor: Prof. Sergei Gorlatch
My PhD research focuses on compiler technologies (code generation and optimization) with a special interest in implementing portable high-performance code for deep learning applications (e.g., linear algebra routines and stencil computations). My overall research goal is to provide *performance*, *portability*, and *productivity* for these computations, making their use in real-world applications feasible. To achieve my goal, I am one of the designers of a holistic code *generation*, *optimization*, and *execution* approach, consisting of three major sub projects:

1. *Multi-Dimensional Homomorphisms (MDH)* – a novel algebraic formalism toward expressing and formally reasoning about data-parallel computations; in particular, this project includes the formal design and specification of a Domain-Specific Language (DSL) for expressing MDH functions, as well as the design and implementation of a compiler for this DSL – the compiler enables automatically generating code for MDHs (e.g., in CUDA, OpenMP, or OpenCL) that can be automatically optimized (auto-tuned) for state-of-the-art GPUs, CPUs, etc.
2. *Auto-Tuning Framework (ATF)* – a general-purpose auto-tuning approach that automatically optimizes parallel programs, based on numerical search techniques and optimized processes of generating, storing, and exploring the optimization spaces of modern parallel programs
3. *Host Code Abstraction (HCA)* – a high-level programming abstraction that simplifies implementing and optimizing so-called host code which is required in modern programming approaches (e.g., CUDA and OpenCL) to execute parallel code on the devices of distributed, heterogeneous systems.

2015 – 2018 **Master of Science in computer science**, *University of Münster*, Münster, Germany, *Final grade in computer science: excellent with distinction (100%)*.
Thesis title: Design and Implementation of a Performance-Portable BLAS Library Using the md_hom Pattern.
Grade for thesis: excellent
Awarded with the German IHK award (Chamber of Commerce and Industry) for its contribution in the area of deep learning

Technical Skills

I am the lead developer of the MDH+ATF+HCA approach. As such, I translate complex mathematical concepts into executable program code (mainly in C++ and Python) and am responsible for all technical decisions regarding the implementation of our projects. Moreover, I coordinated several master and bachelor theses that worked with and contributed code to our projects.

Awards & Achievements

2023 *Performance Bonus* from University of Münster for *Extraordinary Achievements in Research and Teaching (2500 EUR)*

2021 *Best Research Poster Finalist* at SC (*International Conference for HPC, Networking, Storage, and Analysis*) for our work titled: *Code Generation & Optimization for Deep-Learning Computations on GPUs via Multi-Dimensional Homomorphisms*

2019 *Best Poster Award 2019* at PUMPS+AI (Programming and Tuning Massively Parallel Systems + Artificial Intelligence) for our poster: *"Performance, Portability, and Productivity for Data-Parallel Applications on Multi- and Many-Core Architectures"*

2018 *IHK Price 2018* awarded by the German Chamber of Commerce and Industry for my master thesis titled: *"Design and Implementation of a Performance-Portable BLAS Library Based on Multi-Dimensional Homomorphisms"*

Fundings & Grants

2022-2025 **DFG Project Funding (**606.271 **EUR)**, *Performance, Portability, and Productivity for Deep-Learning Computations on Multi- and Many-Core Architectures (PPP-DL)*.
This project aims at achieving *Performance*, *Portability*, and *Productivity* (PPP) for Deep-Learning (DL) computations on multi- and many-core architectures (GPUs, CPUs, etc) via our approaches of Multi-Dimensional Homomorphisms (MDH) and the Auto-Tuning Framework (ATF).

Research Visits

2022 **University of Copenhagen (1 week)**, Copenhagen, Denmark, participants: Mary Hall (University of Utah), Cosmin Oancea (University of Copenhagen – Host), Ari Rasch (University of Münster), Richard Schulze (University of Münster), Denys Shabalin (Google Zurich).
This meeting was focused on discussing and designing programming abstractions for expressing low-level code optimizations.

Invited Talks

2022 AccML - *4th Workshop on Accelerated Machine Learning* - Talk title: *Code Generation and Optimization for Deep-Learning Computations on GPUs via Multi-Dimensional Homomorphisms*

Publications

2023 [1] A. Rasch, R. Schulze, D. Shabalin, A. Elster, S. Gorlatch, and M. Hall. "(De/Re)-Compositions Expressed Systematically via MDH-Based Schedules". In: *ACM SIGPLAN International Conference on Compiler Construction (CC)* (2023).

2022 [2] A. Rasch, R. Schulze, and S. Gorlatch. "Expressing Hierarchical Code Optimizations via MDH-Based Schedules". In: *Workshop on Hierarchical Parallelism for Exascale Computing (HiPar)@SC'22* (2022), (WIP paper).

2021 [3] R. Schulze, A. Rasch, and S. Gorlatch. "Code Generation & Optimization for Deep-Learning Computations on GPUs via Multi-Dimensional Homomorphisms". In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC)* (2021), (short paper).

[4] A. Rasch, R. Schulze, M. Steuwer, and S. Gorlatch. "Efficient Auto-Tuning of Parallel Programs With Interdependent Tuning Parameters via Auto-Tuning Framework ATF". In: *ACM Transactions on Architecture and Code Optimization (TACO)* (2021), (original work).

2020 [5] A. Rasch, R. Schulze, and S. Gorlatch. "md_poly: A Performance-Portable Polyhedral Compiler Based on Multi-Dimensional Homomorphisms". In: *10th International Workshop on Polyhedral Compilation Techniques (IMPACT)* (2020), (WIP paper).

2019 [6] A. Rasch, R. Schulze, and S. Gorlatch. "Generating Portable High-Performance Code via Multi-Dimensional Homomorphisms". In: *The 28th International Conference on Parallel Architectures and Compilation Techniques (PACT)* (2019).

[7] A. Rasch, J.Bigge, M. Wrodarczyk, R. Schulze, and S. Gorlatch. "dOCAL: high-level distributed programming with OpenCL and CUDA". In: *The Journal of Supercomputing (JOS)* (2019).

[8] A. Rasch, R. Schulze, and S. Gorlatch. "Developing High-Performance, Portable OpenCL Code via Multi-Dimensional Homomorphisms". In: *7th International Workshop on OpenCL (IWOCL)* (2019), (extended abstract).

2018 [9] A. Rasch, R. Schulze, M. Gorus, J. Hiller, S. Bartholomäus, and S. Gorlatch. "High-Performance Probabilistic Record Linkage via Multi-Dimensional Homomorphisms". In: *The 34th ACM/SIGAPP Symposium On Applied Computing (SAC)* (2018).

[10] A. Rasch, M. Wrodarczyk, R. Schulze, and S. Gorlatch. "OCAL: An Abstraction for Host-Code Programming with OpenCL and CUDA". In: *The 24th IEEE International Conference on Parallel and Distributed Systems (ICPADS)* (2018).

[11] A. Rasch, R. Schulze, and S. Gorlatch. "Portable Parallel Performance via Multi-Dimensional Homomorphisms". In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC)* (2018), (short paper).

Attended Academic Events

I presented our research (in form of talks and/or posters) at different conferences and events:

2023 PLDI conference (at FCRC) - *ACM SIGPLAN Conference on Programming Language Design and Implementation*, Orlando FL, USA

CC conference - *ACM SIGPLAN International Conference on Compiler Construction*, Montreal, Canada

C4ML workshop - *Compilers for Machine Learning*, Montreal, Canada

2022 HiPEAC conference - *European Forum for Experts in Computer Architecture, Programming Models, Compilers and Operating Systems for Embedded and General-Purpose Systems*, Budapest, Hungary

Lorentz Center workshop - *Generic Auto-Tuning Technologies for GPU Applications*, Leiden, Netherlands

PRACE Course - *Modern C++ Software Design (Advanced)*, Stuttgart, Germany

2021 SC conference - *International Conference for High Performance Computing, Networking, Storage, and Analysis*, St. Louis MO, USA (remote participation)

HiPEAC conference - *European Forum for Experts in Computer Architecture, Programming Models, Compilers and Operating Systems for Embedded and General-Purpose Systems*, Budapest, Hungary (shifted to online event)

PRACE Course - *Modern C++ Software Design (Intermediate)*, online

2020 ACACES summer school (organized by HiPEAC) - *Sixteenth International Summer School on Advanced Computer Architecture and Compilation for High-Performance and Embedded Systems*, Fiuggi, Italy (shifted to online event)

GTC conference - *NVIDIA GPU Technology Conference*, San Jose CA, USA (shifted to online event)

CGO conference - *International Symposium on Code Generation and Optimization*, San Diego CA, USA

C4ML workshop - *Compilers for Machine Learning*, San Diego CA, USA

IMPACT workshop - *International Workshop on Polyhedral Compilation Techniques*, Bologna, Italy

HiPEAC conference - *European Forum for Experts in Computer Architecture, Programming Models, Compilers and Operating Systems for Embedded and General-Purpose Systems*, Bologna, Italy

2019 Google Compiler and Programming Language Summit - München, Germany

LCPC workshop - *Workshop on Languages and Compilers for Parallel Computing*, Atlanta, USA

PRACE course - *Deep Learning and GPU programming using OpenACC*, Stuttgart, Germany

2018 Google Compiler and Programming Language Summit - München, Germany

ACACES summer school (organized by HiPEAC) - *Fourteenth International Summer School on Advanced Computer Architecture and Compilation for High-Performance and Embedded Systems*, Fiuggi, Italy

Euro-Par conference - *International European Conference on Parallel and Distributed Computing*, Turin, Italy

Memberships

ACM Student Member (no.: 8879847)

Supervised Master and Undergraduate Students

Master Arne Wilp: *Accelerating Neural Networks using the MDH Approach*

Lukas Rosendahl: *Evaluating the MDH Approach using Benchmark Suites Parboil and Rodinia*

Sebastian Kock: *Evaluating the MDH Approach for Multi-Device Systems*

Timo Hoth: *Generating High-Performance Code for FFTs via Multi-Dimensional Homomorphisms*

Martin Wrodarczyk: *ECC Classification via Support Vector Machines and Multilayer Perceptron*

Bachelor Dominique Bönninghof: *Design and Implementation of a Directive-Based Code Generation Approach for Multi-Dimensional Homomorphisms*

Gabriel Borrelli: *Visualizing Multi-Layered, Multi-Dimensional Parallel Computations for Modern Processors Based on the MDH Approach*

Luis Wetzel: *Evaluating Multi-Dimensional Homomorphisms via Ensemble Classifier Chains*

Lars Hunloh: *Evaluating the MDH Approach using CUTLASS and PPCG*

Waldemar Gorus: *pyATF: Auto-Tuning Interdependent Tuning Parameters in Python*

Karl Heimes: *Design and Implementation of a Visualization Tool for MDH computations*

Luke Thienemann: *Visualizing Multi-Layered, Multi-Dimensional Parallel Computations via a Multi-Transparent Cube-Based Approach*

Moritz Tätweiler: *Evaluating the MDH Approach Based on Frameworks Kokkos, Raja, Occa, and SYCL*

Julian Bigge: *Extending the OCAL Library for Clusters*

Teaching

Summer 2021	Supervisor of a student project: *Code Generation and Optimization for Deep-Learning Computations on Modern Processors*
Summer 2020	Supervisor of a student project: *Code Generation and Optimization for Deep-Learning Computations on Modern Processors*
Winter 2019	Teaching assistant for the course: *Operating systems*
Summer 2019	Teaching assistant for the course: *Multi-core and GPU: Parallel Programming*
Winter 2018	Supervisor of a student project: *Implementation of Multi-Dimensional Homomorphisms in Low-Level Programming Models*
Summer 2018	Supervisor of a student project: *Automatic Program Optimization via Auto-Tuning and Machine Learning*
Winter 2016	Student assistant for the course: *Operating systems*