

George Bisbas

Zurich (Switzerland) | [mail me here](#) | [Google Scholar](#) | [LinkedIn](#) | (GR) +30 6979 41 27 58
(CH) +41 765 498 054 | <https://georgebisbas.github.io/> | github.com/georgebisbas

Professional Summary

Senior researcher and developer with deep expertise (10 years) in compiler design, for high-performance computing (HPC). Full stack compiler contributions from high-level DSL design and efficient compilation, all the way down to automated HPC-ready code generation for scientific applications. Proven track record in scalable code generation, performance engineering, and open-source leadership. Experienced in cross-disciplinary collaboration and mentoring. Seeking impactful roles in HPC software engineering or compiler development.

Technical Skills

Programming: Python, C, C++, MPI, OpenMP, CUDA, SYCL, OpenACC, Matlab (Expert); Julia, Java, Cilk, AVR/MIPS Assembly, Cerebras CSL (Proficient)

Compiler/IR: MLIR, LLVM (familiar), Devito DSL, PyPTO(Huawei)

HPC/DevOps/CI/CD: GitHub Actions, Travis, Jenkins, Slurm, PBS, Docker, Azure, all kind of HPC envs

HPC Tooling: Intel VTune, Advisor, Trace Analyzer & Collector, NVIDIA Nsight (Compute, Systems), Likwid, writing documentation and tutorials

Application-Oriented Skills

- **LLM Training and inference compiler infrastructure:** Compiler infrastructure that enable distributed training and inference (Data, Tensor, Expert, Sequence, Pipeline Parallelism) of large-scale LLM models on multi-chip clusters. Abstractions for distributed computing, compiler IRs and lowerings, collective communication primitives (AR, AG, RS, etc), DSL API design, distributed IR types and compiler passes, orchestration code generation, runtime integration, performance benchmarking.
- **Seismic Imaging and Wave Propagation:** High-fidelity seismic waveform simulation, inversion workflows, isotropic acoustic, isotropic elastic, anisotropic acoustic (TTI), visco-acoustic, visco-elastic, distributed-memory code generation and temporal cache optimization. Large-scale scientific and industrial seismic and subsurface exploration projects by accelerating finite-difference solvers within the Devito framework.
- **Computational Science and Engineering Workflows:** Developed scalable, production-quality, scalable software stacks (Devito, xDSL) and reproducible pipelines for PDE-based simulations, supporting climate, fluid dynamics, and multiphysics modeling. Automated complex parallel and distributed computation workflows with robust CI/CD, thorough testing, documentation, active contribution, code review, and tutorial materials, facilitating collaborative research and industrial adoption.
- **3D Point Cloud Processing:** Density-based clustering, feature-preserving data reduction. Portable, cost-effective solutions for applications in computer graphics, digital reconstruction, and smart manufacturing, addressing challenges in accuracy, speed, and geometric fidelity.

Education

PhD, High Performance, Embedded and Distributed Systems, Imperial College London 2023

- Thesis [\[Online\]](#): *Automated cache optimizations of stencil computations for partial differential equations*
- Supervisor: Prof. P. H.J. Kelly (with F. Luporini, G. J. Gorman)

MSc, Intelligent Systems/Methods of Computational Intelligence and Applications, Aristotle University of Thessaloniki, Greece 2019

- Grade: 9.59/10 (Distinction)
- Thesis [\[Online \(GR\)\]](#): *On developing and accelerating point cloud simplification methods*
- Supervisor: Prof. N. P. Pitsianis

Diploma of Engineering, Electrical and Computer Engineering, Aristotle 2017

University of Thessaloniki, Greece

- Grade: 8.75/10 (Distinction)
- Thesis [\[Online \(GR\)\]](#): Forecast demand using Extended Discrete Fourier Transform
- Supervisor: Prof. N. P. Pitsianis

Professional Experience

Senior Research Engineer, Huawei Research Center, Zurich, Switzerland

Nov 2025 – Now

- **Compiler & runtime development for distributed training and inference on Ascend NPUs.** Contributing to Huawei's compiler ecosystem (PyPTO/simpler/pto-isa), building the multi-chip distributed communication stack that enables large-scale models (e.g., DeepSeek, MoE architectures) to train and serve across clusters of Ascend NPUs 910B/910C.
- Designed and implemented distributed abstractions, compiler IRs and lowerings, for collective communication primitives (AllReduce, Broadcast, AllGather, ReduceScatter, etc) in PyPTO'S DSL, runtime and ISA. Contributed in compiler development across the full stack: DSL API design, distributed IR types and compiler passes, orchestration code generation, runtime integration and performance benchmarking.
- Repositories: [PyPTO \(compiler\)](#) · [simpler \(runtime\)](#) · [pto-isa \(virtual ISA\)](#) · [PTOAS \(assembler/optimizer\)](#)

Airman (Informatics), Hellenic Air Force, Larissa

May 2025 – Nov 2025

- Six-month mandatory military service as required by the Greek government.

Post-Doctoral Research Associate, Imperial College London, UK

Jun 2023 – Mar 2025

Post-Doctoral Research Assistant, Imperial College London, UK

May 2022 – Jun 2023

PhD student, Imperial College London, UK

Oct 2018 – Apr 2022

With a strong focus on automation, performance engineering, and open-source collaboration, my work consistently delivers scalable solutions for large-scale scientific computing. Across multiple projects, I have advanced the state of the art in compiler infrastructure and code generation for PDE-based applications, developed tools and methodologies that enable domain-specific optimizations and efficient distributed-memory parallelism, and collaborated across teams to deliver solutions now adopted by the HPC community. All work and research are open-source and available, following extensive software testing, code verification, CI/CD, documentation, tutorials, PR code review, and HPC reproducibility. Selected highlights of my work:

- **Temporal cache Blocking for Sparse –off the FD grid –Operators:** Developed and implemented a novel methodology for temporal blocking optimizations in stencil computations with sparse, off-the-grid operators (such as source injections and receiver interpolations) within the Devito DSL and compiler framework. Automated the transformation and precomputation of sparse operator effects, aligning them with the computational grid to unlock significant cache-locality and performance improvements (up to 1.6x speedup) for large-scale wave-propagation kernels used in scientific and seismic imaging applications.
- **Automated Distributed-Memory Parallelism and MPI Code Generation:** Contributed to and integrated automated distributed-memory parallelism and MPI code generation techniques within the Devito DSL and compiler framework, enabling researchers to efficiently scale finite-difference solvers for large-scale scientific simulations. Leveraged advanced compiler infrastructure to abstract away low-level parallel programming details, drastically reducing execution time and developer effort, and delivering highly competitive scaling on modern HPC clusters (CPUs/GPUs) for real-world PDE-based applications.
- **Unified Compilation Stack and MLIR-Based Pipelines:** Led the development of a unified compilation stack (XDSL/Devito Project) enabling distributed-memory parallelism for stencil DSLs, integrating advanced compiler IRs and scalable code generation techniques for high-performance scientific computing. Developed and extended MLIR-based compiler pipelines and custom dialects to enable domain-specific optimizations and scalable distributed-memory parallelism for stencil DSLs, leveraging MLIR's modular infrastructure for advanced IR transformations and extensible compilation workflows. Finally, participated (mostly supervising) an end-to-end framework to bridge the gap between high-level DSLs and the unique architectural constraints of wafer-scale Cerebras CSL.
- Repositories: [Devito](#) · [xDSL Devito](#) · [xdsl](#)

Research Assistant, Aristotle University of Thessaloniki

Nov 2017 – Oct 2018

- Working for the [DigiPro project](#), developed advanced algorithms using Mean Shift clustering to efficiently

simplify 3D point clouds, improving accuracy and speed for large-scale datasets. Implemented solutions in MATLAB and C/Cilk, creating a cost-efficient, portable method for photo-realistic 3D digitization of rigid objects. Addressed challenges in preserving geometric features, balancing data reduction with model fidelity, with applications in computer graphics, 3D modeling, and digital reconstruction.

Teaching Experience

Graduate Teaching Assistant , Imperial College London	Nov 2018 – Oct 2022
Coursework preparation, lecture support, marking for:	
• (ACSE-6) Parallel Programming using the Message Passing Interface (MPI)	Jan 2021 - Mar 2021
• (COMP60001) Advanced Computer Architecture	Nov 2018 - Jan 2021
• Second Year Laboratory program (C++ Picture Processing/Pintos)	Nov 2018 - Jun 2019
• (COMP50006) Compilers	Jan 2021 - Jan 2022
• (COMP60017) Performance Engineering	Jan 2021 - Mar 2021
• (COMP40006) Reasoning about Programs	Jan 2021 - Mar 2021
DiRAC Training Course Mentor , University College London, DIRAC HPC Facility, Department of Physics & Astronomy DiRAC Training Course Mentor	Nov 2023
• Foundation HPC-Skills course: FM01 Bash Shell: Using the Command Line, FM02 Version Control with Git, FM03 Principles of Software Engineering, FM04 Testing, Documenting, and Reviewing Code, FM05 Principles of Code Scaling	

Publications

Conference Publications

- Nicolai Stawinoga*, David Katz*, Anton Lydike, Justs Zarins, Nick Brown, **George Bisbas**, Tobias Grosser (2026). An MLIR Lowering Pipeline for Stencils at Wafer-Scale. In Proceedings of the 31st ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2 (ASPLOS '26), Vol. 2. Association for Computing Machinery, New York, NY, USA, Article 1, 16 pages.
[\[ACM\]](#) [\[Available on Arxiv\]](#)
*equal contribution
- **George Bisbas**, Rhodri Nelson, Mathias Louboutin, Fabio Luporini, Paul H.J. Kelly, Gerard Gorman (2024). Automated MPI-X code generation for scalable finite-difference solvers. In 39th IEEE International Parallel & Distributed Processing Symposium (IPDPS), Milano, Italy, 2025, pp. 689-701.
[\[IEEE\]](#) [\[Available on Arxiv\]](#) [\[Presentation slides\]](#) [\[Poster\]](#)
*Finalist for IPDPS'25 Open Source Contribution Award
- **George Bisbas***, Anton Lydike*, Emilien Bauer*, Nick Brown*, Mathieu Fehr, Lawrence Mitchell, Gabriel Rodriguez-Canal, Maurice Jameson, Paul H.J. Kelly, Michel Steuwer, Tobias Grosser (2024). A shared compilation stack for distributed-memory parallelism in stencil DSLs. In Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 3 (ASPLOS '24), Vol. 3. Association for Computing Machinery, New York, NY, USA, 38–56.
[\[ACM\]](#) [\[Available on Arxiv\]](#) [\[Presentation Slides\]](#) [\[Poster\]](#)
*equal contribution
- **George Bisbas**, Fabio Luporini, Mathias Louboutin, Rhodri Nelson, George Bisbas, Gerard Gorman, Paul H.J. Kelly. 2020. Temporal blocking of finite-difference stencil operators with sparse "off-the-grid" sources. (2020). In 35th IEEE International Parallel & Distributed Processing Symposium (IPDPS), Portland, OR, USA, 2021, pp. 497-506.
[\[IEEE\]](#) [\[Available on Arxiv\]](#) [\[Presentation Slides\]](#)

Workshop Publications

- Joao Speglich, Navjot Kukreja, **George Bisbas**, Atila Saraiva, Jan Hückelheim, Fabio Luporini, John Washbourne (2024). Optimizing wavefield storage with high-speed media. In ESSA'24, IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)
[\[Paper available online\]](#)

In preparation/Submitted/Pre-prints

- Mathias Louboutin, Fabio Luporini, Philipp Witte, Rhodri Nelson, **George Bisbas**, Jan Thorbecke, Felix J. Herrmann, and Gerard Gorman. 2020. Scaling through abstractions – high-performance vectorial wave simulations for seismic inversion with Devito. (2020).
[\[Available on Arxiv\]](#)
- Rhodri Nelson, Fabio Luporini, Mathias Louboutin, **George Bisbas**, Gerard Gorman (2020). TheMatrix: An automated cross-platform benchmarking suite. Submitted to The Journal of Open Source Software
[\[Available on Github\]](#)

Talks and Presentations (Selected)

- **G. Bisbas**, A. Lydike, E. Bauer, N. Brown, M. Fehr, P. H.J. Kelly, T. Grosser *A shared compilation stack for distributed-memory parallelism in stencil DSLs.* (2024). Presented at HiPEAC 2025, Barcelona, Spain
[\[Slides\]](#)
- **G. Bisbas**, A. Lydike, E. Bauer, N. Brown, M. Fehr, P. H.J. Kelly, T. Grosser *A shared compilation stack for HPC stencil DSLs*, Presented at PASC24, Minisymposium: Motif-Based Automated Performance Engineering for HPC.
[\[Slides\]](#)
- **G. Bisbas**, F. Luporini, M. Louboutin, R. Nelson, G. Gorman, P.H.J. Kelly *Automated Temporal Blocking in the Devito Compiler*, Presented at Stencil Computation for Scientific Applications MiniSymposium, SIAM CSE 2023 conference.
[\[Slides\]](#)
- **G. Bisbas**, F. Luporini, M. Louboutin, R. Nelson, G. Gorman, P.H.J. Kelly *Temporal blocking for wave propagation with sparse off-the-grid sources* Presented at Rice Oil and Gas HPC (OGHPC 2021) conference.
[\[Available: Youtube\]](#)
- **G. Bisbas**, F. Luporini, M. Louboutin, R. Nelson, G. Gorman, P.H.J. Kelly *Temporal blocking of finite-difference stencil operators with sparse “off-the-grid” sources* Presented at 21st Workshop on Compilers for Parallel Computing (CPC21, Porto) conference.
[\[Slides\]](#)
- **G. Bisbas**, F. Luporini, M. Louboutin, R. Nelson, G. Gorman, P.H.J. Kelly *Temporal blocking of finite-difference stencil operators with sparse non-grid-aligned sources and receivers in Devito*, Presented at Domain-Specific Languages in High-Performance Computing 2020.
[\[Available: Youtube\]](#)

Poster Presentations (Selected)

- **G. Bisbas**, R. Nelson, M. Louboutin, P. H.J. Kelly, F. Luporini, G. Gorman *Automated MPI-X code generation for scalable finite-difference solvers* Poster presented at Rice Energy HPC 2024.
[\[Poster\]](#)
- **G. Bisbas**, F. Luporini, M. Louboutin, G. Gorman, P.H.J. Kelly *Accelerating real-world stencil computations using temporal blocking: handling sparse sources and receivers* Poster presented at the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC 2019).
[\[Poster\]](#)

Selected Contributions to open-source Projects

- My merged PRs in <https://github.com/hw-native-sys/pypto>
- My merged PRs in <https://github.com/hw-native-sys/simpler>
- My merged PRs in <https://github.com/devitocodes/devito>
- Lead Maintainer/Developer in <https://github.com/xdslproject/devito>
- My merged PRs in <https://github.com/xdslproject/xdsl>

Selected Personal Projects

GreekProphet (Owner/Developer/Maintainer), greekprophet.gr

2026 – Now

- Developed a **probabilistic forecasting platform** for the Greek Super League, combining statistical modelling and simulation to predict match outcomes, team strength, and season trajectories.
- Designed and calibrated a **Monte Carlo-based forecasting engine** with rating systems and score models, emphasizing uncertainty quantification, scenario analysis, and automated forecast generation from real match data.
- Engineered the full analytics stack in Python, optimizing simulation-heavy workloads for high-throughput execution while also building the surrounding data pipelines, backend services, deployment, and live user-facing product.

Peer Reviewing

Conferences

- ACM Transactions on Parallel Computing (2025)
- Supercomputing Conference (SC), Reproducibility Committee Member (2024, 2025)
- Euro-Par 2025, Program Committee Member
- International Conference on Parallel Processing (ICPP), Poster Committee Member (2024); Program Committee (2021)
- JuliaCon Conference Proceedings, Reviewer [\[Reviews\]](#)
- International Workshop on Polyhedral Compilation Techniques (IMPACT), Reviewer (2024)
- PPOPP 2020, Artifact Evaluation Committee Member
- JupyterCon 2020, Proposal Community Reviewer

Journals

- The Journal of Supercomputing, Springer
- Future Generation Computer Systems, Elsevier

Student Supervision

MEng individual project (Imperial college)

- **N. Duer**, *Temporal Tiling for Distributed Parallel Solution of Partial Differential Equations*, Department of Computing, Imperial College London, 2023, in collaboration with P. H.J. Kelly. [\[Thesis Online\]](#)
Distinguished undergraduate project award for 2022-2023 cohort.

Honours and Awards

- Finalist, IPDPS'25 Open Source Contribution Award
- Invited to ACM Student Research Competition and awarded the SRC Travel Award (500\$)
- PhD student position, fully funded by a joint HiPEDS/DoC scholarship

- Scholarship for achieving the best grade in the 1st semester of MSc in Advanced Computer and Communication Systems (650 Euros)
- Eurobank Ergasias monetary prize for excellence in Panhellenic exams (1000 Euros)

Certifications on Online Courses (selected)

- Workshop/Fundamentals of Deep Learning, earned on September 26, 2024 [\[Show credential\]](#)
- Workshop/Fundamentals of Accelerated Computing with CUDA C/C++, earned on June 26, 2024 [\[Show credential\]](#)

Memberships

- SIAM Early Career Membership, #020900949
- IEEE Student Member, #93014477 Greece Section
- ACM Professional Member, #1563304
- Technical Chamber of Greece, Member

References

[Paul HJ Kelly](#), Faculty of Engineering, Department of Computing, Imperial College London
p.kelly@imperial.ac.uk , +44 (0)20 7594 8332

[Gerard J. Gorman](#), Faculty of Engineering, Department of Earth Science & Engineering, Imperial College London
g.gorman@imperial.ac.uk, +44 (0)20 7594 9985

[Fabio Luporini](#), CTO at Devito Codes Ltd.
fabio@devitocodes.com

[Nikos P. Pitsianis](#), Electrical and Computer Engineering, Aristotle University of Thessaloniki
Nikos.P.Pitsianis@Duke.edu, +30 (2310) 994369