Gerasimos Gerogiannis

gg24@illinois.edu
Office 4238, Thomas M. Siebel Center
201 N. Goodwin Avenue, Urbana, IL, 61801, USA

Research Interests

- Computer architecture
- High-performance computing
- Accelerator-centric architectures for machine learning, graph analytics and scientific workloads at scale
- Modern accelerator aspects such as as flexibility, scalability, performance modeling, and decision-making

Education

University of Illinois at Urbana-Champaign (UIUC) Urbana, IL, USA
PhD - Electrical and Computer Engineering; expected graduation date 06/2026 *Aug. 2021 - present*

- Advisor: Prof. Josep Torrellas
- Current GPA: 4.0/4.0

University of Patras Patras, Greece
Diploma in Electrical and Computer Engineering (B.Eng. + M.Eng.) *Sep. 2016 - June 2021*

- GPA: 9.83/10.00 (first of my class)
- Thesis: *'Reinforcement Learning for Task Offloading in Next Generation Networks: Algorithms and Hardware Acceleration'*
  Thesis Supervisor: Prof. Alexios Birbas

Work Experience

Graduate Intern January 2024 - December 2025 (ongoing)
Intel Corporation *(remote) Champaign, IL, USA*

- CPU architecture redesign for emerging Machine Learning applications such as Large Language Models.
- Research conducted in internship led to one publication in MICRO'25, one paper under review, and four filled US patents.

Graduate Intern May 2023 - August 2023
Intel Corporation *(remote) Champaign, IL, USA*

- Heterogeneous accelerator architectures to accelerate sparse matrix - dense matrix multiplication.
- Research conducted in internship led to publication in HPCA'24.

Graduate Intern May 2022 - August 2022
Intel Corporation *(remote) Champaign, IL, USA*

- Synergies between hardware and software to maximize performance on the Intel Programmable Integrated Unified Memory Architecture (PIUMA).
- Research conducted in internship led to publication in ISPASS'23.

Research Assistant August 2021 - present
i-acoma group, University of Illinois at Urbana-Champaign *Champaign, IL, USA*

Awards and Honors

MICRO Conference Travel Grant IEEE/ACM, USA, 2025

Mavis Future Faculty Fellowship UIUC, The Grainger College of Engineering, USA, 2025

- Awarded to selected senior PhD students interested in pursuing engineering faculty careers.

Wen-mei W. Hwu Award UIUC, ECE Department, USA, 2025

- For demonstrating research expertise in the area of Computer Engineering.

Rambus Computer Engineering Fellowship UIUC, ECE Department, USA, 2024

- For demonstrating outstanding research performance in the area of Computer Engineering.

IEEE Micro Top Picks from Computer Architecture Conferences Top Pick IEEE, USA, 2024

- Awarded to the 12 most significant papers in computer architecture published in the previous year based on novelty and potential for long-term impact. Top Pick awarded for my paper: 'Micro-Armed Bandit: Lightweight & Reusable Reinforcement Learning for Microarchitecture Decision-Making'.

IEEE Micro Top Picks from Computer Architecture Conferences Honorable Mention IEEE, USA, 2024

- Awarded to the 12 most significant papers in computer architecture published in the previous year based on novelty and potential for long-term impact. Honorable Mention awarded for my paper: 'SPADE: A Flexible and Scalable Accelerator for SpMM and SDDMM'.

ASPLOS Conference Travel Grant ACM, USA, 2024

HPCA Conference Travel Grant IEEE, USA, 2024

MICRO Conference Travel Grant IEEE/ACM, USA, 2023

ISCA Conference Travel Grant IEEE/ACM, USA, 2023

Student Excellence Award University of Patras, Greece, 2021

- For the highest GPA among all 2020-2021 graduates of the Electrical and Computer Engineering Department of the University of Patras.

Excellence Award State Scholarships Foundation (IKY), Greece, 2021

- For the highest GPA among all 2020-2021 graduates of the Electrical and Computer Engineering Department of the University of Patras.

Skouras Foundation Award Skouras Foundation, Greece, 2020

- For the highest GPA among all undergraduate students in the Electrical and Computer Engineering Department of the University of Patras.

Freshman Excellence Award University of Patras, Greece, 2016

- For the highest grade in the National University Entrance Exams (Panhellenic Exams) among the freshmen students in the Electrical and Computer Engineering Department of the University of Patras.

Distinctions in national high-school student competitions 2013-2016

- Replacement member of the student team that represented Greece in the 47th International Physics Olympiad (Switzerland-Liechtenstein, 2016).
- 7th place in phase A and phase B of the National Student Physics Competition organized by the University of Athens (2016).
- 10th place in the National Student Physics Competition organized by the University of Athens in cooperation with Greek Physicists Union (2015).
- 12th place in the National Student Chemistry Competition organized by the Greek Chemists Union (2015).
- Thales Prize in the National Student Mathematics Competition organized by the Greek Mathematical Society (2014-2015).
- Thales Prize in the National Student Mathematics Competition organized by the Greek Mathematical Society (2013-2014).
- Member of the team that finished 9th in the Greek Qualifiers for the European Union Science Olympiad (EUSO) (2015).

Conference Publications

[C12] *'GRANII: Selection and Ordering of Primitives in GRAph Neural Networks using Input Inspection'*
Damitha Lenadora, Vimarsh Sathia, **Gerasimos Gerogiannis**, Serif Yesil, Josep Torrellas, and Charith Mendis
To appear in Proceedings of the International Symposium on Code Generation and Optimization (CGO), Feburary 2026

[C11] *'NetSparse: In-Network Acceleration of Distributed Sparse Kernels'*
Gerasimos Gerogiannis, Dimitrios Merkouriadis, Charles Block, Annus Zulfiqar, Filippos Tofalos, Muhammad Shahbaz, and Josep Torrellas
To appear in Proceedings of the International Symposium on Microarchitecture (MICRO), October 2025

[C10] *'DECA: A Near-Core LLM Decompression Accelerator Grounded on a 3D Roofline Model'*
Gerasimos Gerogiannis, Stijn Eyerman, Evangelos Georganas, Wim Heirman, and Josep Torrellas
To appear in Proceedings of the International Symposium on Microarchitecture (MICRO), October 2025

[**C9**] *'Micro-MAMA: Multi-Agent Reinforcement Learning for Multicore Prefetching'*
Charles Block, **Gerasimos Gerogiannis** and Josep Torrellas
To appear in Proceedings of the International Symposium on Microarchitecture (MICRO), October 2025

[**C8**] *'COGNATE: Learning-Based Acceleration of Sparse Tensor Programs on Emerging Hardware'*
Chamika Sudusinghe, **Gerasimos Gerogiannis**, Damitha Lenadora, Charles Block, Josep Torrellas, and Charith Mendis
In Proceedings of the International Conference on Machine Learning (ICML), July 2025

[**C7**] *'MeshSlice: Efficient 2D Tensor Parallelism for Distributed DNN Training'*
Hyoungwook Nam, **Gerasimos Gerogiannis**, and Josep Torrellas,
In Proceedings of the International Symposium on Computer Architecture (ISCA), June 2025

[**C6**] *'Distributed-Memory Parallel Algorithms for Sparse Matrix and Sparse Tall-and-Skinny Matrix Multiplication'*
Isuru Ranawaka, Md Taufique Hussain, Charles Block, **Gerasimos Gerogiannis**, Josep Torrellas, and Ariful Azad
In Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC), November 2024

[**C5**] *'Two-Face: Combining Collective and One-Sided Communication for Efficient Distributed SpMM'*
Charles Block*, **Gerasimos Gerogiannis***, Charith Mendis, Ariful Azad, and Josep Torrellas
*Co-first authors, order is alphabetical
In Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), April 2024

[**C4**] *'HotTiles: Accelerating SpMM with Heterogeneous Accelerator Architectures'*
Gerasimos Gerogiannis, Sriram Aananthakrishnan, Josep Torrellas, and Ibrahim Hur
In Proceedings of the International Symposium on High Performance Computer Architecture (HPCA), March 2024

[**C3**] *'Micro-Armed Bandit: Lightweight & Reusable Reinforcement Learning for Microarchitecture Decision-Making'*
Gerasimos Gerogiannis and Josep Torrellas
In Proceedings of the International Symposium on Microarchitecture (MICRO), October 2023
Selected as a Top Pick in IEEE Micro Top Picks from Computer Architecture Conferences

[**C2**] *'SPADE: A Flexible and Scalable Accelerator for SpMM and SDDMM'*
Gerasimos Gerogiannis, Serif Yesil, Damitha Lenadora, Dingyuan Cao, Charith Mendis, and Josep Torrellas
In Proceedings of the International Symposium on Computer Architecture (ISCA), June 2023
Selected as a Honorable Mention in IEEE Micro Top Picks from Computer Architecture Conferences

[**C1**] *'Characterizing the Scalability of Graph Convolutional Networks on Intel® PIUMA'*
Matthew Adiletta, Jesmin Jahan Tithi, Emmanouil-Ioannis Farsarakis, **Gerasimos Gerogiannis**, Robert Adolf, Robert Benke, Sidharth Kashyap, Samuel Hsia, Kartik Lakhotia, Fabrizio Petrini, Gu-Yeon Wei, and David Brooks
In Proceedings of the International Symposium on Performance Analysis of Systems and Software (ISPASS), April 2023

Journal Publications

[**J2**] *'Practical Online Reinforcement Learning for Microprocessors with Micro-Armed Bandit'*
Gerasimos Gerogiannis and Josep Torrellas
IEEE Micro Magazine, Top Picks in Computer Architecture Special Issue, July-August 2024

[**J1**] *'Deep Reinforcement Learning Acceleration for Real-Time Edge Computing Mixed Integer Programming Problems'*
Gerasimos Gerogiannis, Michael Birbas, Aimilios Leftheriotis, Eleftherios Mylonas, Nikolaos Tzanis, and Alexios Birbas
IEEE Access, January 2022

Workshop Publications (with proceedings)

[**W1**] *'Automated Data Selection for Efficient Cost Model Training to Optimize Sparse Matrix Kernels on Emerging Hardware Accelerators'*
Chamika Sudusinghe, **Gerasimos Gerogiannis**, Damitha Lenadora, Charles Block, Josep Torrellas, and Charith Mendis
In Proceedings of the Exploration in AI Today Workshop at ICML 2025

US Patents

[**P4**] *'Unified Transfer Engine for Near-Core Compute Accelerators'*
Gerasimos Gerogiannis, Stijn Eyerman, and Wim Heirman

[**P3**] *'Speculative Invocation of Accelerators in Out-of-Order Core Pipelines'*
Gerasimos Gerogiannis, Stijn Eyerman, and Wim Heirman

[**P2**] *'Efficient On-chip Memory Bandwidth Configurations and Distribution for Artificial Intelligence (AI) Workloads'*
Wim Heirman, Stijn Eyerman, and **Gerasimos Gerogiannis**

[**P1**] *'Methods and Apparatus for a Machine Learning Model Decompression Accelerator'*
Gerasimos Gerogiannis, Stijn Eyerman, Wim Heirman, and Evangelos Georganas
US Patent App. 18/927,638 November 2024

Posters

[**O3**] *'HotTiles: Accelerating SpMM with Heterogeneous Accelerator Architectures'*
International Symposium on High Performance Computer Architecture (HPCA) Poster Session, March 2024

[**O2**] *'SPADE: A Flexible and Scalable Accelerator for SpMM and SDDMM'*
- ACE Center for Evolvable Computing Spring Meeting, June 2023
- ACE Center for Evolvable Computing Annual Meeting, October 2023

[**O1**] *'Micro-Armed Bandit: Lightweight & Reusable Reinforcement Learning for Microarchitecture Decision-Making'*
International Symposium on Microarchitecture (MICRO) Poster Session, October 2023

Conference Talks

[**N7**] *'NetSparse: In-Network Acceleration of Distributed Sparse Kernels'*
To be given at the International Symposium on Microarchitecture (MICRO), Session 5C: Sparsity - 2, October 21 2025

[**N6**] *'DECA: A Near-Core LLM Decompression Accelerator Grounded on a 3D Roofline Model'*
To be given at the International Symposium on Microarchitecture (MICRO), Session 2A: Systems for AI (LLMs) - 2, October 20 2025

[**N5**] *'Sparsity at Scale: Towards Efficient Distributed Sparse Accelerators'*
SRC TECHCON 2024,
Session 24: Digital Processing: Artificial Intelligence, September 10 2024

[**N4**] *'Two-Face: Combining Collective and One-Sided Communication for Efficient Distributed SpMM'*
International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), Session 1B: Optimizing ML Communication, April 29 2024

[**N3**] *'HotTiles: Accelerating SpMM with Heterogeneous Accelerator Architectures'*
International Symposium on High Performance Computer Architecture (HPCA), Session 10A: Accelerator, March 6 2024

[**N2**] *'Micro-Armed Bandit: Lightweight & Reusable Reinforcement Learning for Microarchitecture Decision-Making'*
International Symposium on Microarchitecture (MICRO), Session 5B: Prefetching, October 31 2023

[**N1**] *'SPADE: A Flexible and Scalable Accelerator for SpMM and SDDMM'*
International Symposium on Computer Architecture (ISCA), Session 1B: CPU Microarchitecture, June 19 2023

Other Talks

[**T5**] *'Micro-Armed Bandit: Lightweight & Reusable Reinforcement Learning for Microarchitecture Decision-Making'*
Intel uArch Monthly Meeting, July 16 2024

[**T4**] *'Micro-Armed Bandit: Lightweight & Reusable Reinforcement Learning for Microarchitecture Decision-Making'*
Intel Archfest, May 17 2024

[**T3**] *'Micro-Armed Bandit: Lightweight & Reusable Reinforcement Learning for Microarchitecture Decision-Making'*
ACE Center for Evolvable Computing Theme Meeting, December 1 2023

[**T2**] *'Domain-Specific Hardware and Software for Mixed Sparse-Dense Algebra at Scale'*
ACE Center for Evolvable Computing Theme Meeting, August 4 2023

[**T1**] *'Accelerators for Irregular Applications'*
IBM-Illinois Discovery Accelerator Institute Annual Meeting, March 1 2023

Selected Projects

miniSPADE 2022-2023

- Designed the microarchitecture and taped-out a simplified ASIC prototype of the accelerator described in our ISCA'23 paper using TSMC 65nm technology.

Teaching Experience

Private Tutor 2016 – 2020
Patras, Greece

Tutor 2020 – 2021
Prosimo Education Center *Patras, Greece*

- Gave lectures and tutored undergraduate students in preparation for their exams in a variety of electrical and computer engineering subjects.

Grant Writing Experience

NSF 2023 – 2028
PPoSS: LARGE: General-Purpose Scalable Technologies for Fundamental Graph Problems *Amount: $5.000.000*

- Lead PI: Josep Torrellas, number of PIs: 10
- Assisted in the preparation and writing of the NSF Grant proposal, focusing on hardware support for scalable graph algorithms and on the hardware-software interaction.

Professional Memberships

Student Member May 2023 – present
ACM

Reviewer Service

IEEE Transactions on Computers 2025
Invited Reviewer

Technical Skills

Programming Languages: C/C++, Python, Matlab, Shell/Bash scripting, Perl, VHDL, Verilog
Parallel Programming: OpenMP, MPI, Posix Threads
Frameworks/Libraries: Simulink, Vivado, Vivado HLS, Cadence EDA tools, TensorFlow, PyTorch, Deep Graph Library
Microarchitectural Simulators and Tools: SST, Sniper, gem5, ChampSim, Accel-Sim, DRAMSim, CACTI, McPAT
Development Tools: git, CMake

Hobbies

- FIDE rated chess player - member of the team that won the bronze medal in the National Student Chess Competition (2013).
- Guitar and Greek traditional instruments player.

References Available Upon Request

Josep Torrellas, Professor, University of Illinois at Urbana-Champaign, USA

Charith Mendis, Assistant Professor, University of Illinois at Urbana-Champaign, USA

Ibrahim Hur, Senior Principal Architecture Lead, Microsoft, USA

Stijn Eyerman, Research Scientist, Intel, Belgium

Alexios Birbas, Professor, University of Patras, Greece

Stavros Koubias, Professor Emeritus and Former Rector, University of Patras, Greece