

QIUSHI LIN

Email | *Scholar*

EDUCATION

Tsinghua University

Bachelor in Electronic Engineering (EE)

GPA: 3.9/4.0

Sep. 2022—Present

PUBLICATIONS

1. **TD-Orch: Scalable Load-Balancing for Distributed Systems with Applications to Graph Processing** (*Paper*)
Qiushi Lin*, Yiwei Zhao*, Hongbo Kang, Guy E. Blelloch, Laxman Dhulipala, Charles McGuffey and Phillip B. Gibbons.
Under Review at OSDI'26.
2. **Breaking the Memory Wall: A Survey of DRAM-based Processing-In-Memory Architectures and Systems** (*Paper*)
Qiushi Lin*, Zhenhua Zhu+, Tongxin Xie, Yiwei Zhao, Guy E. Blelloch, Phillip B. Gibbons, Guohao Dai, Mingyu Gao, Yuan Xie and Yu Wang.
Under Review at CSUR.
3. **How Do Errors Impact NN Accuracy on Non-Ideal Analog PIM? Fast Evaluation via an Error-Injected Robustness Metric** (*Paper* | *Slides*)
Lidong Guo, Zhenhua Zhu+, **Qiushi Lin**, Yuan Xie, Huazhong Yang, Wangyang Fu+, and Yu Wang+
ICCAD'25.
4. **HPIM-NoC: A Priori-Knowledge-Based Optimization Framework for Heterogeneous PIM-Based NoC** (*Paper* | *Code*)
Shuai Yuan, Angxin Cai, **Qiushi Lin**, Guoxing Wang, Yu Wang, Zhenhua Zhu, Yanan Sun.
DAC'25.
5. **Deep Neural Network Inference Partitioning in Embedded Analog-Digital Hybrid Systems** (*Paper*)
Fabian Kreß, Julian Hoefer, **Qiushi Lin**, Patrick Schmidt, Zhenhua Zhu, Yu Zhu, Tanja Harbaum, Yu Wang, Jurgen Becker.
ISQED'25.

SELECTED RESEARCH EXPERIENCE

Task-Data Orchestration in Distributed Memory Systems (Under Review at OSDI'26)

Nov. 2024—Oct. 2025

Carnegie Mellon University | Advisor: Prof. Phillip B. Gibbons, Prof. Guy E. Blelloch

- Designed a skew-aware load-balancing orchestration algorithm to optimize task-data allocation in distributed memory systems and built a graph processing engine on top of the framework.
- Formulated the theoretical foundations of the scheduling algorithm, and architected the system core with ~10,000 lines of C++/MPI code, conducting extensive testing and optimization in distributed environments.
- Co-first-authored a paper under submission, where our framework TD-Orch achieves up to $2.7\times$ speedup over distributed scheduling baselines and the derived system TDO-GP achieves $4.1\times$ average speedup over prior state-of-the-art distributed graph processing frameworks.

Scheduling Algorithm and Microarchitecture Design for Processing Near Memory (Under Review at CSUR; Preparing for SOSP'26)

Apr. 2025—Present

Carnegie Mellon University | Advisor: Prof. Phillip B. Gibbons, Prof. Guy E. Blelloch and Prof. Yu Wang

- Conducted a comprehensive survey on DRAM-based processing-in-memory (PIM), covering hardware architectures, system-level optimizations, programming models, and security challenges, and summarized future directions for heterogeneous computing. Under review at CSUR.
- Co-designed a QoS- and tail-latency-aware scheduling framework for PIM, addressing skewed access patterns, hot data regions, and multi-tenant workloads. Built a microservice-inspired execution model that jointly optimizes data placement, parallelism, and memory-access affinity across PIM modules.
- Developed a general-purpose PIM simulation framework that models a wide range of DRAM-PIM architectures and integrates host-CPU timing, enabling unified full-system evaluation of kernel latency, data movement, and QoS-aware scheduling policies across CPU and PIM cores.

NIPA: A Non-slicing Accuracy Evaluation Framework for Analog PIM (ICCAD'25)

Sep. 2024—Mar. 2025

Tsinghua University | Advisor: Prof. Yu Wang

- Contributed to the theoretical framework for unifying diverse hardware errors (device variation, quantization, ADC) into a single weight-level error model, enabling joint error analysis and bypassing traditional bit-and-crossbar slicing simulations.
- Developed the Non-Ideal PIM Accuracy (NIPA) evaluation model and its simulation framework, which leverages the unified error metric to rapidly assess relative accuracy and achieved a high correlation (up to 0.91) with ground-truth results.
- Implemented the complete non-slicing absolute accuracy evaluation method, which injects the unified error directly into model weights, achieving up to a 105.8 \times speedup over conventional PIM simulators with an average evaluation error as low as 0.29%.

Simulation and Optimization of Digital–Analog Heterogeneous PIM-NoC (DAC'25)

May 2024—Nov. 2024

Tsinghua University | Advisor: Prof. Yu Wang

- Constructed a heterogeneous multi-core PIM-NoC simulation framework with flexible tile parameterization and configurable load ordering, providing an experimental platform for heterogeneous PIM system optimization.
- Authored the majority of the codebase, including over 4,000 lines of Python to streamline simulation workflows and data analysis.
- Designed an efficient search-based optimization algorithm on top of the simulator, achieving 2–3 \times faster design-space exploration and up to 37.4% FoM improvement over homogeneous PIM-NoC baselines.

Network Partitioning for Heterogeneous Analog PIM–Digital Accelerators (ISQED'25)

Jan. 2024—May 2024

Tsinghua University | Advisor: Prof. Yu Wang

- Designed an automated exploration architecture for embedded analog in-memory computing-digital accelerator hybrid chiplet systems, optimized deployment across multiple accelerators and breakpoints, and performed neural network searches based on defined optimization objectives to develop near-optimal partitioning strategies.
- Developed a simulation model to validate different partitioning strategies, analyzed experimental data, and optimized computational load distribution to enhance system performance.
- Implemented multiple optimization techniques for rapid performance evaluation and refined code execution for efficiency, enabling up to 52% latency reduction with less than 1% accuracy loss in hybrid analog-digital PIM accelerator systems

TALKS

MIT Parallel Reading Group (Oct. 2025)

Tsinghua NICSEFC (Sep. 2025)

DAC Presentation (Jun. 2025)

AWARDS & RECOGNITION

Tsinghua University Comprehensive Scholarship (Top 8%, 2023)
Tsinghua University Comprehensive Scholarship (Top 5%, 2024)
Tsinghua University Comprehensive Scholarship (Top 2%, 2025)
1st Place, 2024 EE Dept Innovation Project Evaluation Annual Conference, Tsinghua

PROFESSIONAL SKILLS

Programming: C/C++, Python, Verilog, Java, Shell, LaTeX

Systems & Simulation: CUDA, MPI, OpenMP, Parlaylib, Gem5, MNSIM, BookSim, SPICE

Tools & Infrastructure: Docker, Git, GDB, perf, CMake

ML & Analysis: PyTorch, TensorFlow, QAT, MATLAB

Languages: English (TOEFL 107), Chinese (Native)