

QI SUN

ZJU100 Young Professor, Zhejiang University
qisunchn@zju.edu.cn, Hangzhou, Zhejiang, China

RESEARCH TOPICS

- Electronic Design Automation
- Computational Lithography
- Design Space Exploration
- Deep Learning Acceleration
- Large Language Model for DTCO

EXPERIENCE

Zhejiang University, Hangzhou, China ZJU100 Young Professor, College of Integrated Circuits, Zhejiang University	Aug. 2023 – Now
Cornell University, Ithaca, NY, USA Post-Doctoral Associate, School of ECE, Computer Systems Laboratory Supervisor: Prof. Zhiru Zhang	Sept. 2022 – Aug. 2023
SmartMore Technology Co., Ltd., Hong Kong Research Intern Heterogeneous Computing Group Topic: Deep Neural Network Hardware Acceleration	Mar. 2020 – Jun. 2022

EDUCATION

The Chinese University of Hong Kong, NT, Hong Kong Ph.D., Department of Computer Science & Engineering. Supervisor: Prof. Bei Yu	Aug. 2018 – Jul. 2022
Xidian University, Xi'an, P.R. China B.Eng., Software Engineering.	Aug. 2014 – Jul. 2018

PUBLICATIONS

Conference Papers

- [C35] Haomin Li, Fangxin Liu, Zongwu Wang, Dongxu Lyu, Shiyuan Huang, Ning Yang, **Qi Sun**, Zhuoran Song, Li Jiang, “TAIL: Exploiting Temporal Asynchronous Execution for Efficient Spiking Neural Networks with Inter-Layer Parallelism”, Design, Automation and Test in Europe Conference (**DATE**), Lyon, France, Mar. 31 - Apr. 02, 2025.
- [C34] Haonan Du, Chenyi Wen, Zhengrui Chen, Li Zhang, **Qi Sun**, Zheyu Yan, Cheng Zhuo, “Algorithm-Hardware Co-design of a Unified Accelerator for Non-linear Functions in Transformers”, Design, Automation and Test in Europe Conference (**DATE**), Lyon, France, Mar. 31 - Apr. 02, 2025.
- [C33] Bohao Li, Shizhang Wang, Tinghuan Chen, **Qi Sun**#, Cheng Zhuo, “Efficient Subgraph Matching Framework for Fast Subcircuit Identification”, 6th ACM/IEEE International Symposium on Machine Learning for CAD (**MLCAD**), Utah, Sept. 9-11, 2024.
- [C32] Xudong Lu, Yuqi Jiang, Haiwen Hong, **Qi Sun**#, Cheng Zhuo#, “DCAFuse: Dual-Branch Diffusion-CNN Complementary Feature Aggregation Network for Multi-Modality Image Fusion”, ACM International Conference on Multimedia (**MM**), Melbourne, Australia, Oct. 28-Nov. 01, 2024.
- [C31] Lvcheng Chen, Ying Wu, Chenyi Wen, Shizhang Wang, Li Zhang, Bei Yu, **Qi Sun**#, Cheng Zhuo#, “An Agile Framework for Efficient LLM Accelerator Development and Model Inference”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), New Jersey, Oct. 27–31, 2024. (**William J. McCalla Best Paper Award**)

- [C30] Qian Jin, Yuqi Jiang, Xudong Lu, Yumeng Liu, Yining Chen, Dawei Gao, **Qi Sun**#, Cheng Zhuo#, “SEM-CLIP: Precise Few-Shot Learning for Nanoscale Defect Detection in Scanning Electron Microscope Image”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), New Jersey, Oct. 27–31, 2024.
- [C29] Yuqi Jiang, Xudong Lu, Qian Jin, **Qi Sun**#, Hanming Wu, Cheng Zhuo#, “FabGPT: An Efficient Large Multimodal Model for Complex Wafer Defect Knowledge Queries”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), New Jersey, Oct. 27–31, 2024.
- [C28] Shangran Lin, Xinrui Zhu, Baohui Xie, Tinghuan Chen, Cheng Zhuo, **Qi Sun**, Bei Yu, “RISCsparse: Point Cloud Inference Engine on RISC-V Processor”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), New Jersey, Oct. 27–31, 2024.
- [C27] Yuanhang Gao, Donger Luo, Chen Bai, Bei Yu, Hao Geng, **Qi Sun**#, Cheng Zhuo, “Is Vanilla Bayesian Optimization Enough for High-Dimensional Architecture Design Optimization?”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), New Jersey, Oct. 27–31, 2024.
- [C26] Lancheng Zou, Wenqian Zhao, Shuo Yin, Chen Bai, **Qi Sun**#, Bei Yu#, “BiE: Bi-Exponent Block Floating-Point for Large Language Models Quantization”, International Conference on Machine Learning (**ICML**), Vienna, Jul. 21–27, 2024.
- [C25] Xianting Lu, Yunong Wang, Yu Fu, **Qi Sun**, Xuhua Ma, Xudong Zheng, Cheng Zhuo, “MISP: A Multimodal-based Intelligent Server Failure Prediction Model for Cloud Computing Systems”, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**ICDKDD**), Barcelona, Spain, Aug. 25–29, 2024.
- [C24] Yuqi Jiang, Qian Jin, Xudong Lu, **Qi Sun**#, Cheng Zhuo, “FabSage: A Large Multimodal Model for IC Defect Detection, Analysis, and Knowledge Querying”, 1st IEEE International Workshop on LLM-Aided Design (**LAD**), San Jose, CA, June 28–29, 2024.
- [C23] Donger Luo*, **Qi Sun***, Xinheng Li, Chen Bai, Bei Yu, Hao Geng, “Knowing The Spec to Explore The Design via Transformed Bayesian Optimization”, ACM/IEEE Design Automation Conference (**DAC**), San Francisco, Jun. 23–27, 2024.
- [C22] Yibo Qiao, Weiping Xie, Shunyu Lou, Qian Jin, Lichao Zeng, Yining Chen#, **Qi Sun**#, Cheng Zhuo#, “Minimizing Labeling, Maximizing Performance: A Novel Approach to Nanoscale Scanning Electron Microscope (SEM) Defect Segmentation”, ACM/IEEE Design Automation Conference (**DAC**), San Francisco, Jun. 23–27, 2024.
- [C21] Hongquan He, Guowen Kuang, **Qi Sun**, Hao Geng, “Point Cloud and Large Pre-trained Model Catch Mixed-type Wafer Defect Pattern Recognition”, IEEE/ACM Proceedings Design, Automation and Test in Europe (**DATE**), Valencia, Spain, Mar. 25–27, 2024.
- [C20] Donger Luo, **Qi Sun**, Qi Xu, Tinghuan Chen, Hao Geng, “Attention-Based EDA Tool Parameter Explorer: From Hybrid Parameters to Multi-QoR metrics”, IEEE/ACM Proceedings Design, Automation and Test in Europe (**DATE**), Valencia, Spain, Mar. 25–27, 2024.
- [C19] Chenyi Wen, Haonan Du, Zhengrui Chen, Li Zhang, **Qi Sun**, Cheng Zhuo, “PACE: A Piece-Wise Approximate and Configurable Floating-Point Divider for Energy-Efficient Computing”, IEEE/ACM Proceedings Design, Automation and Test in Europe (**DATE**), Valencia, Spain, Mar. 25–27, 2024.
- [C18] Erika S. Alcorta, Andreas Gerstlauer, Chenhui Deng, **Qi Sun**, Zhiru Zhang, Ceyu Xu, Lisa Wu Wills, Daniela Sánchez Lopera, Wolfgang Ecker, Siddharth Garg, and Jiang Hu, “Machine Learning for Embedded System Design”, International Conference on Hardware/Software Codesign and System Synthesis (**CODES+ISSS**), Sep. 2023. (Invited Paper)
- [C17] Yuxuan Zhao, **Qi Sun**#, Zhuolun He, Yang Bai, Bei Yu, “AutoGraph: Optimizing DNN Computation Graph for Parallel GPU Kernel Execution”, AAAI Conference on Artificial Intelligence (**AAAI**), Feb. 7–14, 2023.
- [C16] Xufeng Yao, Yang Bai, Xinyun Zhang, Yuechen Zhang, **Qi Sun**, Ran Chen, Ruiyu Li, Bei Yu, “PCL: Proxy-based Contrastive Learning for Domain Generalization”, IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), New Orleans, Jun. 19–24, 2022.
- [C15] **Qi Sun**, Xinyun Zhang, Hao Geng, Yuxuan Zhao, Yang Bai, Haisheng Zheng, Bei Yu, “GTuner: Tuning DNN Computations on GPU via Graph Attention Network”, ACM/IEEE Design Automation Conference

- (**DAC**), San Francisco, CA, Jul. 10–14, 2022.
- [C14] Xinyun Zhang, Binwu Zhu, Xufeng Yao, **Qi Sun**, Ruiyu Li, Bei Yu, “Context-based Contrastive Learning for Scene Text Recognition”, AAI Conference on Artificial Intelligence (**AAAI**), Feb. 22–Mar. 1, 2022.
 - [C13] Hao Geng, Tinghuan Chen, **Qi Sun**, Bei Yu, “Techniques for CAD Tool Parameter Auto-tuning in Physical Synthesis: A Survey”, IEEE/ACM Asian and South Pacific Design Automation Conference (**ASPDAC**), Jan. 17–20, 2022. (Invited Paper)
 - [C12] **Qi Sun**, Chen Bai, Tinghuan Chen, Hao Geng, Xinyun Zhang, Yang Bai, Bei Yu, “Fast and Efficient DNN Deployment via Deep Gaussian Transfer Learning”, International Conference on Computer Vision (**ICCV**), Oct. 11–17, 2021.
 - [C11] Wenqian Zhao, **Qi Sun**, Yang Bai, Haisheng Zheng, Wenbo Li, Bei Yu, Martin D.F. Wong, “A High-Performance Accelerator for Super-Resolution Processing on Embedded GPU”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), Nov. 1–4, 2021.
 - [C10] Chen Bai, **Qi Sun**, Jianwang Zhai, Yuzhe Ma, Bei Yu, Martin D.F. Wong, “BOOM-Explorer: RISC-V BOOM Microarchitecture Design Space Exploration Framework”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), Nov. 1–4, 2021. (**William J. McCalla Best Paper Award**)
 - [C9] Yang Bai, Xufeng Yao, **Qi Sun**, Bei Yu, “AutoGTCo: Graph and Tensor Co-Optimize for Image Recognition with Transformers on GPU”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), Nov. 1–4, 2021.
 - [C8] Tinghuan Chen, **Qi Sun**, Bei Yu, “Machine Learning in Nanometer AMS Design for Reliability”, IEEE International Conference on ASIC (**ASICON**), Kunming, Oct. 26–29, 2021. (Invited Paper)
 - [C7] **Qi Sun**, Tinghuan Chen, Siting Liu, Jin Miao, Jianli Chen, Hao Yu, Bei Yu, “Correlated Multi-objective Multi-fidelity Optimization for HLS Directives Design”, IEEE/ACM Proceedings Design, Automation and Test in Europe (**DATE**), Feb. 01–05, 2021. (**Best Paper Award Nomination**)
 - [C6] **Qi Sun**, Chen Bai, Hao Geng, Bei Yu, “Deep Neural Network Hardware Deployment Optimization via Advanced Active Learning”, IEEE/ACM Proceedings Design, Automation and Test in Europe (**DATE**), Feb. 01–05, 2021.
 - [C5] Siting Liu, **Qi Sun**, Peiyu Liao, Yibo Lin, Bei Yu, “Global Placement with Deep Learning-Enabled Explicit Routability Optimization”, IEEE/ACM Proceedings Design, Automation and Test in Europe (**DATE**), Feb. 01–05, 2021.
 - [C4] Tinghuan Chen, **Qi Sun**, Canhui Zhan, Changze Liu, Huatao Yu, Bei Yu, “Analog IC Aging-induced Degradation Estimation via Heterogeneous Graph Convolutional Networks”, IEEE/ACM Asian and South Pacific Design Automation Conference (**ASPDAC**), Jan. 18–21, 2021.
 - [C3] **Qi Sun**, Arjun Ashok Rao, Xufeng Yao, Bei Yu, Shiyan Hu, “Counteracting Adversarial Attacks in Autonomous Driving”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), Nov. 2–5, 2020. (Invited Paper)
 - [C2] **Qi Sun**, Tinghuan Chen, Jin Miao, Bei Yu, “Power-Driven DNN Dataflow Optimization on FPGA”, IEEE/ACM International Conference on Computer-Aided Design (**ICCAD**), Westminster, CO, Nov. 4–7, 2019. (Invited Paper)
 - [C1] Wei Li, Yuzhe Ma, **Qi Sun**, Yibo Lin, Iris Hui-Ru Jiang, Bei Yu, David Z. Pan, “OpenMPL: An Open Source Layout Decomposer”, IEEE International Conference on ASIC (**ASICON**), Chongqing, China, Oct. 29–Nov. 1, 2019. (Invited Paper)

Journal Papers

- [J11] Tinghuan Chen, Hao Geng, **Qi Sun**, Sanping Wan, Yongsheng Sun, Huatao Yu, Bei Yu, “Wages: The Worst Transistor Aging Analysis for Large-scale Analog Integrated Circuits via Domain Generalization”, accepted by ACM Transactions on Design Automation of Electronic Systems (**TODAES**).
- [J10] Chen Bai, **Qi Sun**, Jianwang Zhai, Yuzhe Ma, Bei Yu, Martin D.F. Wong, “BOOM-Explorer: RISC-V BOOM Microarchitecture Design Space Exploration”, ACM Transactions on Design Automation of Electronic Systems (**TODAES**), vol. 29, no. 01, pp. 1–23, 2024.

- [J9] Yang Bai, Xufeng Yao, **Qi Sun**, Wenqian Zhao, Shixin Chen, Zixiao Wang, Bei Yu, “GTCO: Graph and Tensor Co-Design for Transformer-based Image Recognition on Tensor Cores”, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**), vol. 43, no. 02, pp. 586–599, 2024.
- [J8] Wenqian Zhao, Yang Bai, **Qi Sun**, Wenbo Li, Haisheng Zheng, Nianjuan Jiang, Jiangbo Lu, Bei Yu, Martin D.F. Wong, “A High-Performance Accelerator for Super-Resolution Processing on Embedded GPU”, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**), vol. 42, no. 10, pp. 3210–3223, 2023.
- [J7] Qi Xu, Junpeng Wang, **Qi Sun**, Bo Yuan, Song Chen, Bei Yu, Yi Kang, Feng Wu, “Reliability-Driven Memristive Crossbar Design in Neuromorphic Computing Systems”, IEEE Transactions on Automation Science and Engineering (**TASE**), vol. 20, no. 01, pp. 74–87, 2023.
- [J6] **Qi Sun**, Tinghuan Chen, Siting Liu, Jianli Chen, Hao Yu, Bei Yu, “Correlated Multi-objective Multi-fidelity Optimization for HLS Directives Design”, ACM Transactions on Design Automation of Electronic Systems (**TODAES**), vol. 27, no. 4, 2022.
- [J5] **Qi Sun**, Xufeng Yao, Arjun Ashok Rao, Bei Yu, Shiyan Hu, “Counteracting Adversarial Attacks in Autonomous Driving”, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**), vol. 41, no. 12, pp. 5193–5206, 2022.
- [J4] Guojin Chen, Wanli Chen, **Qi Sun**, Yuzhe Ma, Haoyu Yang, Bei Yu, “DAMO: Deep Agile Mask Optimization for Full Chip Scale”, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**), vol. 41, no. 9, pp. 3118–3131, 2022.
- [J3] Tinghuan Chen, **Qi Sun**, Canhui Zhan, Changze Liu, Huatao Yu, Bei Yu, “Deep H-GCN: Fast Analog IC Aging-induced Degradation Estimation”, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**), vol. 41, no. 7, pp. 1990–2003, 2022.
- [J2] Tinghuan Chen, Bin Duan, **Qi Sun**, Meng Zhang, Guoqing Li, Hao Geng, Qianru Zhang, Bei Yu, “An Efficient Sharing Grouped Convolution via Bayesian Learning”, IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**), 2021.
- [J1] Wei Li, Yuzhe Ma, **Qi Sun**, Zhang Lu, Yibo Lin, Iris Hui-Ru Jiang, Bei Yu, David Z. Pan, “OpenMPL: An Open Source Layout Decomposer”, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**), vol. 40, no. 11, pp. 2331–2344, 2021.

SELECTED AWARDS AND HONORS

William J. McCalla Best Paper Award	ICCAD	2024
Qizhen Scholar Foundation	ZJU Education Foundation	2023
3rd Place of Student Research Competition	ICCAD	2021
William J. McCalla Best Paper Award	ICCAD	2021
Best Paper Award Nomination	DATE	2021
Full Postgraduate Studentship	CUHK	2018
Outstanding Graduate	Xidian University	2018
National Scholarship	Xidian University	2017
The First Class Scholarship	Xidian University	2014 - 2016

TEACHING ASSISTANT

Fall 2018	CSCI2720 Building Web Application
Spring 2019	CSCI1020 Hands-on Introduction to C++
Fall 2019	CSCI3230 Fundamentals of Artificial Intelligence
Fall 2020	CENG4480 Embedded System Development and Applications
Fall 2020	CMSC5743 Efficient Computing of Deep Neural Networks
Spring 2021	CENG5030 Energy Efficient (Deep Neural Network) Computing

SERVICE

Organization

Technical Program Committee

- ACM/IEEE Design Automation Conference (DAC), 2024, 2025
- International Symposium of EDA (ISED), 2025
- IEEE/ACM International Conference on Computer-Aided Design (ICCAD), 2024
- IEEE/ACM Asia and South Pacific Design Automation Conference (ASP-DAC), 2024
- AAAI Conference on Artificial Intelligence (AAAI), 2023
- IEEE International Symposium on VLSI (ISVLSI), 2023

Journal Review

- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems (TCAD)
- ACM Transaction on Design Automation of Electronic Systems (TODAES)
- IEEE Internet of Things Journal (IoT-J)

Conference Review

- ACM/IEEE Design Automation Conference (DAC)
- IEEE/ACM Asia and South Pacific Design Automation Conference (ASP-DAC)
- IEEE International Symposium on Circuits and Systems (ISCAS)
- IEEE International Conference on Computer Design (ICCD)
- ACM/IEEE Workshop on Machine Learning for CAD (MLCAD)
- IEEE International Conference on Embedded Software and Systems (ICCESS)
- IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC)
- ACM/IEEE International Workshop on System-Level Interconnect Pathfinding (SLIP)

TECHNICAL SKILLS

Languages	C/C++, Vivado HLS C++, Python, CUDA C/C++, \LaTeX
Operating Systems	Linux/UNIX