JIAQI GU

Graduate Research Assistant ⋄ ECE Department ⋄ University of Texas at Austin
jqgu@utexas.edu ⋄ jeremiemelo.github.io

RESEARCH INTERESTS

Machine learning, optimization, algorithm and architecture design for next-generation AI, efficient neuromorphic computing with emerging technology, parallel computing/GPU acceleration for VLSI design automation

EDUCATION

University of Texas at Austin, TX, USA *Aug. 2018 – Present*
Ph.D. student, Department of Electrical and Computer Engineering
Advisor: David Z. Pan
Co-advisor: Ray T. Chen
(GPA 4.0/4.0)

Fudan University, Shanghai, China *Sep. 2014 – Jun. 2018*
B.E., Department of Microelectronic Science and Engineering (Eminent Engineer Program)
(GPA: 3.91/4.0)
(Rank top 2/71)

EXPERIENCE

University of Texas at Austin, TX, USA *Jan. 2019 – Present*
Graduate Research Assistant

- **Efficient Neural Arch. Design, NN Pruning**: Designed novel frequency-domain photonic neuromorphic computing architecture for area-efficient optical neural network; achieved 3-4× area reduction by using block-circulant matrices and structured pruning compared with previous ONN architectures
- **NN Quantization and Robustness**: Developed differentiable quantization-aware training scheme in the unitary manifold to enable robust optical neural networks with low-precision voltage controls; achieved better accuracy and robustness with limited control resolution and device-level variations
- **Photonics CNN and RNN Design**: Collaborated on designing high-throughput and low-power photonic CNN architectures; helped design and fabricate high-speed silicon-photonic RNN
- **NN On-chip Learning**: Developed customized CUDA extension for ONN simulation acceleration; proposed efficient on-chip learning algorithm for optical neural networks with stochastic zeroth-order optimization; achieved 3-4× higher learning efficiency, 10× better scalability, and better robustness than previous methods
- **NN Efficient Training Framework**: Collaborated on developing efficient training framework for reversible neural architectures via constrained optimization; our dynamic programming based scheduling achieves 5-20% speedup with comparable memory efficiency when training reversible NNs
- **Photonics Neural Chip Tapeout**: Worked on photonic neural chip tapeout for novel ONN architectures using Advanced Micro Foundry (AMF); collaborated on the full-stack schematic design, layout, validation, tape-out, and measurement of photonic neural chips using PyTorch, Lumerical toolkits, and Synopsys optodesigner
- **VLSI Detailed Placement Acceleration**: Collaborated on developing GPU-accelerated concurrent VLSI detailed placement with CUDA; implemented and optimized global swap and parallel auction algorithm for batched-based independent-set-matching; achieved >10× speedup than sequential implementations without quality degradation
- **VLSI Global Placement Acceleration**: Collaborated on high-performance VLSI analytical global placement acceleration with CUDA on GPUs; optimized wirelength and density computation operators with CUDA; developed parallel congestion map estimation for routability optimization; achieved 40× speedup in global placement
- **VLSI Global Placement Algorithm**: Developed multi-electrostatics-based robust VLSI placement framework DREAMPlace 3.0 with PyTorch/C++/CUDA; proposed multielectrostatic system for optimization under fence region constraints; developed divergence-aware optimizer for robust nonlinear global placement; achieved >13% HPWL improvement and >11% top5 overflow reduction compared with ISPD2015 contest winners
- **Efficient NN Learning and Optimization**: Proposed efficient ONN on-chip learning framework with two-level sparse optimizer and efficient power-aware optimization; achieved high convergence stability, ∼10× training efficiency improvement, and ∼10× power reduction than prior methods

University of Texas at Austin, TX, USA *Sep. 2018 – Jan. 2019*

Graduate Research Assistant

- **FPGA Emulation of RISC-V Core**: Projected RISC-V Rocket Core on Zynq FPGA with Chisel3 and achieved communication between them
- **Fault Injection**: Customized FIRRTL transformation and built infrastructure for fault injection and system state snapshot

Fudan University, Shanghai, China *Aug. 2017 – Jul. 2018*

Undergraduate Research Assistant

- **Medical Imaging Dataset**: Modified infant brain atlas and created complete tissue probability maps
- **MRI Reconstruction**: Developed two-stage reconstruction framework for infant thin-section MR image reconstruction by using GANs and CNN; research is developing brand new method to improve reconstruction performance by fusing multi-planar MR images, and improving PSNR, SSIM, and NMI by 26.2%, 93.4%, and 25.3% respectively compared to bicubic interpolation
- **Ultra-sonic Image Processing**: Collaborated on super-resolution reconstruction of ultra-sonic imaging using U-Net and GANs; improved the full width at half maximum (FWHM) of point targets by 3.23%

PUBLICATIONS

Journal Papers

[J8] Zhoufeng Ying, Chenghao Feng, Zheng Zhao, **Jiaqi Gu**, Richard Soref, David Z. Pan, and Ray T. Chen, "Sequential logic and pipelining in chip-based electronic-photonic digital computing," *IEEE Photonics Journal*, Oct. 2020.

[J7] **Jiaqi Gu**, Zheng Zhao, Chenghao Feng, Zhoufeng Ying, Mingjie Liu, Ray T. Chen, and David Z. Pan, "Towards Hardware-Efficient Optical Neural Networks: Beyond FFT Architecture via Joint Learnability," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, 2020.

[J6] Chenghao Feng, Zhoufeng Ying, Zheng Zhao, **Jiaqi Gu**, Ray T. Chen, and David Z. Pan, "Wavelength-division-multiplexing (WDM)-based integrated electronic–photonic switching network (EPSN) for high-speed data processing and transportation," *Nanophotonics*, Aug. 2020.

[J5] Yibo Lin, Zixuan Jiang, **Jiaqi Gu**, Wuxi Li, Shounak Dhar, Haoxing Ren, Brucek Khailany, and David Z. Pan, "DREAMPlace: Deep Learning Toolkit-Enabled GPU Acceleration for Modern VLSI Placement," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, Jun. 2020.

[J4] Zhoufeng Ying, Chenghao Feng, Zheng Zhao, Shounak Dhar, Hamed Dalir, **Jiaqi Gu**, Yue Cheng, Richard Soref, David Z. Pan, and Ray T. Chen, "Electronic-photonic Arithmetic Logic Unit for High-speed Computing," *Nature Communications*, Apr. 2020.

[J3] Yibo Lin, Wuxi Li, **Jiaqi Gu**, Mark Ren, Brucek Khailany, and David Z. Pan, "ABCDPlace: Accelerated Batch-based Concurrent Detailed Placement on Multi-threaded CPUs and GPUs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, Feb. 2020.

[J2] Ruoyao Wang, Zhenghan Fang, **Jiaqi Gu**, Yi Guo, Shicong Zhou, Yuanyuan Wang, Cai Chang, and Jinhua Yu, "High-resolution Image Reconstruction for Portable Ultrasound Imaging Devices," *EURASIP Journal on Advances in Signal Processing*, Dec. 2019.

[J1] **Jiaqi Gu**, Zeju Li, Yuanyuan Wang, Haowei Yang, Zhongwei Qiao, and Jinhua Yu, "Deep Generative Adversarial Networks for Thin-section Infant MR Image Reconstruction," *IEEE Access*, May 2019.

Conference Papers

[C18] **Jiaqi Gu**, Chenghao Feng, Zheng Zhao, Zhoufeng Ying, Ray T. Chen, and David Z. Pan, "Efficient On-Chip Learning for Optical Neural Networks Through Power-Aware Sparse Zeroth-Order Optimization," *Association for the Advancement of Artificial Intelligence (AAAI)*, Feb. 2021. (accepted)

[C17] Shubham Rai, Walter Lau Neto, Yukio Miyasaka, Xinpei Zhang, Mingfei Yu, Qingyang Yi, Masahiro Fujita, Guilherme B. Manske, Matheus F. Pontes, Leomar S. da Rosa Junior, Marilton S. de Aguiar, Paulo F. Butzen, Po-Chun Chien, Yu-Shan Huang, Hoa-Ren Wang, Jie-Hong R. Jiang, **Jiaqi Gu**, Zheng Zhao, Zixuan Jiang, David Z. Pan, *et al.*, "Logic Synthesis Meets Machine Learning: Trading Exactness for Generalization," *IEEE/ACM Proceedings Design, Automation and Test in Europe (DATE)*, Feb. 2021. (accepted)

[C16] **Jiaqi Gu**, Chenghao Feng, Zheng Zhao, Zhoufeng Ying, Mingjie Liu, Ray T. Chen, and David Z. Pan, "SqueezeLight: Towards Scalable Optical Neural Networks with Multi-Operand Ring Resonators," *IEEE/ACM Proceedings Design, Automation and Test in Europe (DATE)*, Feb. 2021. (accepted)

[C15] **Jiaqi Gu**, Zheng Zhao, Chenghao Feng, Zhoufeng Ying, Ray T. Chen, and David Z. Pan, "O2NN: Optical Neural Networks with Differential Detection-Enabled Optical Operands," *IEEE/ACM Proceedings Design, Automation and Test in Europe (DATE)*, Feb. 2021. (accepted)

[C14] Chenghao Feng, **Jiaqi Gu**, Zhoufeng Ying, Zheng Zhao, Ray T. Chen, and David Z. Pan, "Scalable fast-Fourier-transform-based (FFT-based) integrated optical neural network for compact and energy-efficient deep learning," *SPIE Photonics West*, Mar. 2021. (accepted)

[C13] Chenghao Feng, Zhoufeng Ying, Zheng Zhao, **Jiaqi Gu**, Ray T. Chen, and David Z. Pan, "Wavelength-division-multiplexing (WDM)-based integrated electronic–photonic switching network (EPSN) for high-speed data processing and transportation," *SPIE Photonics West*, Mar. 2021. (accepted)

[C12] **Jiaqi Gu**, Zixuan Jiang, and David Z. Pan, "DREAMPlace 3.0: Multi-Electrostatics Based Robustness VLSI Placement with Region Constraints," *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Nov. 2020.

[C11] Zixuan Jiang, Keren Zhu, Mingjie Liu, **Jiaqi Gu**, and David Z. Pan, "An Efficient Training Framework for Reversible Neural Architectures," *European Conference on Computer Vision (ECCV)*, Aug. 2020.

[C10] **Jiaqi Gu**, Zheng Zhao, Chenghao Feng, Wuxi Li, Ray T. Chen, and David Z. Pan, "FLOPS: Efficient On-Chip Learning for Optical Neural Networks Through Stochastic Zeroth-Order Optimization," *ACM/IEEE Design Automation Conference (DAC)*, Jul. 2020. (**Best Paper Candidate**)

[C9] Mario Miscuglio, Zibo Hu, Shurui Li, **Jiaqi Gu**, Aydin Babakhani, Puneet Gupta, Chee-Wei Wong, Hamed Dalir David Z. Pan Seth Bank, and Volker J. Sorger, "Million-channel parallelism Fourier-optic convolutional filter and neural network processor," *Conference on Lasers and Electro-Optics*, May 2020.

[C8] Chenghao Feng, Zhoufeng Ying, Zheng Zhao, **Jiaqi Gu**, Ray T. Chen, and David Z. Pan, "Integrated WDM-based Optical Comparator for High-speed Computing," *Conference on Lasers and Electro-Optics*, May 2020.

[C7] Chenghao Feng, Zheng Zhao, Zhoufeng Ying, **Jiaqi Gu**, David Z. Pan, and Ray T. Chen, "Compact design of On-chip Elman Optical Recurrent Neural Network," *Conference on Lasers and Electro-Optics*, May 2020.

[C6] **Jiaqi Gu**, Zheng Zhao, Chenghao Feng, Hanqing Zhu, Ray T. Chen, and David Z. Pan, "ROQ: A Noise-Aware Quantization Scheme Towards Robust Optical Neural Networks with Low-bit Controls," *IEEE/ACM Proceedings Design, Automation and Test in Europe (DATE)*, Mar. 2020.

[C5] Mingjie Liu, Keren Zhu, **Jiaqi Gu**, Linxiao Shen, Xiyuan Tang, Nan Sun, and David Z. Pan, "Towards Decrypting the Art of Analog Layout: Placement Quality Prediction via Transfer Learning," *IEEE/ACM Proceedings Design, Automation and Test in Europe (DATE)*, Mar. 2020.

[C4] Chenghao Feng, Zhoufeng Ying, Zheng Zhao, **Jiaqi Gu**, Ray T. Chen, and David Z. Pan, "Wavelength-division-multiplexing-based electronic-photonic network for high-speed computing," *SPIE, Smart Photonic and Optoelectronic Integrated Circuits XXII*, Feb. 2020.

[C3] **Jiaqi Gu**, Zheng Zhao, Chenghao Feng, Mingjie Liu, Ray T. Chen, and David Z. Pan, "Towards Area-Efficient Optical Neural Networks: an FFT-based architecture," *IEEE/ACM Asia and South Pacific Design Automation Conference (ASPDAC)*, Jan. 2020. (**Best Paper Award**)

[C2] Zheng Zhao, **Jiaqi Gu**, Zhoufeng Ying, Chenghao Feng, Ray T. Chen, and David Z. Pan, "Design Technology for Scalable and Robust Photonic Integrated Circuits," *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2019. (**Invited Paper**)

[C1] **Jiaqi Gu**, Ruoyao Wang, Jian Wang, Jinmei Lai, and Qinghua Duan, "Remote Embedded Simulation System for SW/HW Co-design Based On Dynamic Partial Reconfiguration," *International Conference on ASIC (ASICON)*, 2017.

RELATED COURSES

- EE382N: Computer Architecture — *Prof. Dam Sunwoo*
- EE382N: High-Speed Computer Arithmetic I — *Prof. Earl Swartzlander*
- EE382N: Computer Architecture: Parallelism/Locality — *Prof. Mattan Erez*
- CS395T: Parallel Algorithm Scientific Computing — *Prof. George Biros*
- CS394R: Reinforcement Learning: Theory and Practice — *Prof. Peter Stone and Prof. Scott Niekum*
- EE382M: VLSI I — *Prof. Jacob A. Abraham*
- EE382M: VLSI Physical Design Automation — *Prof. David Z. Pan*
- EE382V: Cross-layer ML Alg./HW Co-design — *Prof. Mattan Erez and Prof. Michael Orshansky*
- EE382M: VLSI CAD and Optimization — *Prof. David Z. Pan*
- EE381V: Combinatorial Optimization — *Prof. Constantine Caramanis*

SKILLS

Programming Languages
Python (PyTorch/TensorFlow), C/C++, CUDA, Matlab, Verilog

EDA Tools
Cadence Virtuoso, Synopsys Design Compiler, Hspice, Xilinx Vivado Design Suite, Synopsys Optodesigner

AWARDS AND HONORS

Best Poster Award at NSF Workshop on Machine Learning Hardware	NSF Workshop	2020
Gold Medal at ACM/SIGDA Student Research Competition	ACM/SIGDA	2020
7th Place at IWLS Contest on Machine Learning+Logic Synthesis	IWLS	2020
DAC Young Fellow	DAC	2020
Best Paper Finalist (1 out of 6)	DAC	2020
Best Paper Award	ASP-DAC	2020
4th Place, System Design Contest on Low Power Object Detection	DAC-SDC	2019
First Prize Scholarship	Fudan University	2017–2018
Top 5, HUAWEI & FUTURELAB AI Contest (CV Group)	Fudan University	2018
Top 11%, IEEEXtreme Global Programming Competition	IEEE	2017
2nd & 3rd Prize, National Mathematical Contest in Modeling	Fudan University	2016–2017