HANQING ZHU

Graduate Research Assistant - University of Texas at Austin
hqzhu@utexas.edu ⋄ (512)200-6791 ⋄ https://zhuhanqing.github.io/

RESEARCH INTERESTS

Efficient and robust AI computing system with emerging technology (photonics), hardware-efficient machine learning, and machine learning for VLSI design automation.

EDUCATION

The University of Texas at Austin (UT-Austin), TX, USA *Aug. 2020 – - Present*
Ph.D. student, Department of Electrical and Computer Engineering
Advisor: David Z. Pan
Co-advisor: Ray T. Chen
(GPA: 3.95/4.00)

Shanghai Jiao Tong University (SJTU), Shanghai, China *Sept. 2016 - Jun. 2020*
B.E., Dept. of Microelectronics Science and Engineering
(GPA: 3.81/4.00)
(Rank: 2^{nd}/57)

INTERNSHIP

Google LLC., CA, USA *Jul. 2022 – Present*
Student Researcher
Google Brain Team, advised by Dr. Wenjie (Joe) Jiang
Topic: Reinforcement Learning for Macro Placement

HONORS AND AWARDS

Winner at Robert S. Hilbert Memorial Optical Design Competition	Synopsys	2022
DAC Young Fellow	DAC	2021
Shanghai Outstanding Graduate	Shanghai City	2020
Department Excellent Undergraduate Thesis	Shanghai Jiao Tong University	2020
Hongyi Scholarship	Shanghai Jiao Tong University	2019
Outstanding Undergraduate Scholarship	Shanghai Jiao Tong University	2019
Samsung Scholarship	Shanghai Jiao Tong University	2018
Zhiyuan College Honors Scholarship	Shanghai Jiao Tong University	2018
1st Prize, National Mathematical Contest in Modeling	Shanghai Division	2018
Academic Excellence Scholarship	Shanghai Jiao Tong University	2017-2019

RESEARCH EXPERIENCE

The University of Texas at Austin, TX, USA *Aug. 2020 – Present*
Graduate Research Assistant, advised by Prof. David Z. Pan and Prof Ray T. Chen

- **ML for EDA**
  - **Reinforcement learning for Macro placement**: Intern project.
  - **DREAMPlace quality improvement**: Use reinforcement learning to tune the parameters of DREAMPlace to boost its performance.
- **AI Computing System with Emerging technology** [C1, C4, C5, C6, C9, C10, J1, J2, J3, P1]
  - **Efficient Analog Activation Unit for NN Accelerators** [C9, C10]: Devised an energy-efficient analog CAM-based activation unit for emerging NN accelerators; proposed a differentiable search framework to enable on-chip activation workloads distribution on a mixed activation system; demonstrated over 60% activation and A/D energy saving with comparable accuracy.
  - **Aging-resilient Photonic In-memory Neurocomputing** [J2, C5]: Proposed an aging-aware optimization framework for photonic in-memory computing; achieved over 40× dynamic energy cost and 20× programming operations reduction; significantly enhanced executing lifetime of photonic accelerators.

  - **AI-assisted Automatic Photonic Tensor Core Design** [C6]: Enabled the usage of AI to automatically search Photonic tensor core (PTC) circuit topology in a differentiable way; achieved 2×-30× higher footprint compactness with competitive matrix representability on searched photonic designs.
  - **NN On-chip/On-device Learning** [C4]: Proposed an efficient first-order based on-chip learning protocol for optical neural networks; devised a subspace learning procedure with multi-level sparsity to enable *in-situ* gradient evaluation and low computation cost; achieved 3-order-of-magnitude higher scalability and over 30× better efficiency than previous optical on-chip training tools.
  - **Photonics Neural Chip Tape-out**: [P1] Worked on photonic neural chip tape-out 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.
- **Hardware-efficient Machine Learning** [C1, C3, C5, J2]
  - **Memory-efficient NN Design** [C3]: Designed memory-efficient multi-level low-rank weight generation methodology with mixed-precision quantization; saved 10×-20× on-chip memory cost for emerging NN accelerators;.
  - **NN Quantization and Robustness** [C1, C5, J2]: Developed differentiable quantization-aware training scheme with uniform/logarithmic quantization operators considering limited control resolution for different types of optical neural networks; also enabled noise-injection-based variation-aware training to improve robustness with limited control resolution and device-level variations.

PROFESSIONAL SERVICE

Reviewer

- Photonic Network Communications
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- IEEE/ACM International Conference on Computer-Aided Design (ICCAD)
- IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)

SKILLS

Programming Languages
Python (PyTorch), C++, CUDA, Verilog

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

PUBLICATIONS

Journal Papers

[J3] Jiaqi Gu, Chenghao Feng, **Hanqing Zhu**, Zheng Zhao, Zhoufeng Ying, Mingjie Liu, Ray T. Chen and David Z. Pan, "SqueezeLight: A Multi-Operand Ring-Based Optical Neural Network with Cross-Layer Scalability," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, Jul., 2022.

[J2] **Hanqing Zhu**, Jiaqi Gu, Chenghao Feng, Mingjie Liu, Zixuan Jiang, Ray T. Chen, and David Z. Pan, "ELight: Towards Efficient and Aging-Resilient Photonic In-Memory Neurocomputing," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, Jun., 2022.

[J1] Jiaqi Gu, Chenghao Feng, **Hanqing Zhu**, Ray T. Chen and David Z. Pan, "Light in AI: Toward Efficient Neurocomputing with Optical Neural Networks - A Tutorial," in *IEEE Transactions on Circuits and Systems–II: Express Briefs (TCAS-II)*, Apr., 2022.

Conference Papers

[C11] Jiaqi Gu, Zhengqi Gao, Chenghao Feng, **Hanqing Zhu**, Ray Chen, Duane S Boning, and David Z. Pan, " NeurOLight: A Physics-Agnostic Neural Operator Enabling Parametric Photonic Device Simulation," in *Conference on Neural Information Processing Systems (NeurIPS)*, Nov 26 - Dec 4, 2022. (Accepted)

[C10] Harrison Jin, **Hanqing Zhu**, Keren Zhu, Thomas Leonard, Mahshid Alamdar, David Z. Pan, and Jean Anne C. Incorvia, "Design of Domain Wall-Magnetic Tunnel Junction Analog Content Addressable Memory using Current and Projected Prototype Data," in *Annual Conference on Magnetism and Magnetic Materials (MMM)*, Minneapolis, MN, October 31 - November 4, 2022. (Accepted)

[C9] **Hanqing Zhu**, Keren Zhu, Jiaqi Gu, Harrison Jin, Ray Chen, Jean Anne Incorvia and David Z. Pan, "Fuse and Mix: MACAM-Enabled Analog Activation for Energy-Efficient Neural Acceleration" in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Oct., 2022

[C8] Chenghao Feng, Jiaqi Gu, **Hanqing Zhu**, Zhoufeng Ying, Zheng Zhao, David Z. Pan, and Ray T. Chen, "Optoelectronically Interconnected Hardware-Efficient Deep Learning using Silicon Photonic Chips," in *Smart Photonic and Optoelectronic Integrated Circuits (SPIE)*, Mar., 2022

[C7] Chenghao Feng, Jiaqi Gu, **Hanqing Zhu**, David Z. Pan, and Ray T. Chen, "Design and Experimental Demonstration of A Hardware-Efficient Integrated Optical Neural Network," in *Smart Photonic and Optoelectronic Integrated Circuits (SPIE)*, Mar., 2022

[C6] Jiaqi Gu, **Hanqing Zhu**, Chenghao Feng, Zixuan Jiang, Mingjie Liu, Shuhan Zhang, Ray T. Chen, and David Z. Pan, "ADEPT: Automatic Differentiable DEsign of Photonic Tensor Cores," in *ACM/IEEE Design Automation Conference (DAC)*, Jul., 2022

[C5] **Hanqing Zhu**, Jiaqi Gu, Chenghao Feng, Mingjie Liu, Zixuan Jiang, Ray T. Chen, and David Z. Pan, "ELight: Enabling Efficient Photonic In-Memory Neurocomputing with Life Enhancement," in *IEEE/ACM Asia and South Pacific Design Automation Conference (ASP-DAC)*, Jan., 2022.

[C4] Jiaqi Gu, **Hanqing Zhu**, Chenghao Feng, Zixuan Jiang, Ray T. Chen, and David Z. Pan, "L2ight: Enabling On-Chip Learning for Optical Neural Networks via Efficient in-situ Subspace Optimization," in *Conference on Neural Information Processing Systems (NeurIPS)*, Dec., 2021.

[C3] Jiaqi Gu, **Hanqing Zhu**, Chenghao Feng, Mingjie Liu, Zixuan Jiang, Ray T. Chen, and David Z. Pan, "Towards Memory-Efficient Neural Networks via Multi-Level in situ Generation," in *International Conference on Computer Vision (ICCV)*, Oct., 2021.

[C2] Chenghao Feng, Jiaqi Gu, **Hanqing Zhu**, David Z. Pan, and Ray T. Chen, "Experimental Demonstration of a WDM-based Integrated Optical Decoder for Compact Optical Computing," in *Conference on Lasers and Electro-Optics*, May, 2021.

[C1] 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," in *IEEE Design, Automation & Test in Europe Conference & Exhibition (DATE)*, Mar., 2020.

Preprint Papers

[P1] Chenghao Feng*, Jiaqi Gu*, **Hanqing Zhu**, Zhoufeng Ying, Zheng Zhao, David Z. Pan, and Ray T. Chen, "Silicon photonic subspace neural chip for hardware-efficient deep learning," in *arXiv preprint 2111.06705*, 2021.