AKHILESH V. BALASINGAM

avb03@stanford.edu

EDUCATION

I am a master's student at Stanford University studying Computer Science. I earned my BS in Electrical Engineering from Stanford University, where I graduated with departmental honors. My research interests are in computer architecture, compilers, and the design of hardware-software systems for high-performance computing and machine learning.

Stanford University, *M.S., Computer Science (Focus: Artificial Intelligence)* *Apr 2025 – Present*
Stanford University, *B.S. with Honors, Electrical Engineering (Focus: Hardware-Software Systems)* *Sep 2021 - Jun 2025*

WORK & RESEARCH EXPERIENCE

Mapping Recurrences to Systolic Arrays *Mar 2024 – Present*
Mentored by Prof. Fredrik Kjolstad, Stanford University

- Developing a systematic approach to map recurrence equations onto systolic array architectures for parallel computing. Compiler to lower high-level descriptions of recurrences to custom systolic chiplets, that can be integrated into larger HW designs.

SWE Intern at Apple, CoreOS, Vector and Numerics Team *June 2024 – Sept 2024, June 2025 – Sept 2025*
Machine Learning Subgroup

- I implemented a fast and energy-efficient ML model, consisting of fully-connected and convolutional layers, by taking advantage of new HW features that supported a compressed datatype. Showed close-to-theoretical performance improvement (1.8x) and memory improvement (2x) over previous implementation.
- Built a Python library to quantize PyTorch models to custom microscaling low-precision representations. Extended compiler framework to support quantization for internal models.

Evaluation & Optimization of Sparse Tensor Algebra Kernels on CGRAs *May 2023 – Mar 2024*
Mentored by Prof. Priyanka Raina & Prof. Fredrik Kjolstad, Stanford University

- Developed tiling pipeline for set of tensor algebra expressions. Investigated tiling optimizations to improve performance relative to CPU baselines using state-of-the-art sparse computation libraries. Built infrastructure to compare performance of sparse and dense acceleration flows for matrices of varying sparsity.

Coupled Bubbles Simulation and Sound Synthesis *Mar 2023 – Mar 2024*
Mentored by Prof. Doug James, Stanford University

- Implemented matrix factorization and update algorithms in CUDA to reduce sound synthesis compute time for large systems of coupled bubbles (e.g. oceans). Achieved up to 717x speed-up over single-threaded CPU implementation. Implemented bubble tracking algorithms for constraint bubble-based sound synthesis.

Mapping End-to-End Convolutional Neural Networks (CNNs) on CGRAs *Jan 2023 – May 2023*
Mentored by Prof. Priyanka Raina, Stanford University

- Implemented algorithms and schedules for non-linear layers of CNNs in Halide. Optimized schedules to maximize processing element (PE) utilization and parallelism.

Impact of Multi-Level RRAM Cell (MLC) Errors on DNN Accuracy *Mar 2022 – Nov 2023*
Mentored by Prof. Priyanka Raina, Stanford University

- Used Pytorch to train quantization-aware convolutional neural networks. Created framework to inject MLC errors and measure inference accuracy degradation. Investigated different encoding schemes.

Multipole NEM-Relays for Programmable Routing *May 2021 – Dec 2021*
Mentored by Prof. Priyanka Raina, Stanford University

- Built simulation pipeline in Perl and HSPICE to model NEM relay one-hot multiplexers, which, in this design, are integrated into the switchboxes of CGRAs to control the signal routing.

Selected Coursework (Completed, Current, & Planned)

Programming Abstractions, Computer Systems, Operating Systems, Probability and Statistics, Multivariable Calculus, Linear Algebra, Optics, Ordinary Differential Equations, Partial Differential Equations, Signal Processing, Engineering Design Optimization, Circuits, Neural Models for 3D Geometry, Digital System Design, Machine Learning, Graph Neural Networks, Computer Vision, Intro to Linear Dynamical Systems, Computational Imaging, Graphics: Animation & Simulation, Applied Quantum Mechanics, Robot Autonomy and Perception, Convex Optimization, Digital Signal Processing, Digital Systems Architecture, Compilers, Applied Quantum Mechanics.

Presentations & Publications

[1] "Systolic-Sim: A Framework for the Rapid Prototyping of Custom Systolic Architectures for Recurrence-Based Computation," Electrical Engineering Honors Thesis, Stanford University. **Akhilesh Balasingam**.

[2] "Onyx: A 12nm 756 GOPS/W Coarse-Grained Reconfigurable Array for Accelerating Dense and Sparse Applications," IEEE VLSI 2024. Kalhan Koul, Maxwell Strange, Jackson Melchert, Alex Carsello, Yuchen Mei, Olivia Hsu, Taeyoung Kong, Po-Han Chen, Huifeng Ke, Keyi Zhang, Qiaoyi Liu, Gedeon Nyengele, **Akhilesh Balasingam**, Jayashree Adivarahan, Ritvik Sharma, Zhouhua Xie, Christopher Torng, Joel Emer, Fredrik Kjolstad, Mark Horowitz, Priyanka Raina.

[3] "Onyx: A Programmable Accelerator for Sparse Tensor Algebra" IEEE Hot Chips Symposium. Kalhan Koul, Maxwell Strange, Jackson Melchert, Alex Carsello, Yuchen Mei, Olivia Hsu, Taeyoung Kong, Po-Han Chen, Huifeng Ke, Keyi Zhang, Qiaoyi Liu, Gedeon Nyengele, **Akhilesh Balasingam**, Jayashree Adivarahan, Ritvik Sharma, Zhouhua Xie, Christopher Torng, Joel Emer, Fredrik Kjolstad, Mark Horowitz, Priyanka Raina.

[4] A. Levy, M. Oduoza, **A. Balasingam**, R. T. Howe, P. Raina, "3-D coarse-grained reconfigurable array using multi-pole NEM relays for programmable routing, *Integration*, 2022, ISSN 0167-9260.

[5] A. Levy, M. Oduoza, **A. Balasingam**, R. T. Howe and P. Raina, "Efficient Routing in Coarse-Grained Reconfigurable Arrays Using Multi-Pole NEM Relays," *2022 27th Asia and South Pacific Design Automation Conference (ASP-DAC), 2022*, pp. 472-478

[6] **Balasingam, A.,** Levy, A., Li, H., & Raina, P. (2020). Monte Carlo simulation of a three-terminal RRAM with applications to neuromorphic computing. *2020 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*.

Awards & Honors

- **NSF Graduate Research Fellowship (GRFP).** Awarded scholarship for research potential. *2025*
- **Member, Tau Beta Pi, Stanford University Chapter** *2023*
  Inducted into engineering honors society for the top 1/8 of engineering juniors at Stanford.
- **Finalist, Regeneron Science Talent Search (STS)** *2021*
  Named 1 of 40 finalists in the nation's oldest and most prestigious science competition.
- **National AP Scholar & National Merit Finalist** *2021*
  Perfect Score of 5 on all 12 AP exams taken.
- **Valedictorian, Archbishop Mitty High School, Class of 2021** *2021*

Activities

President, Stanford Institute of Electrical and Electronics Engineers (IEEE) Chapter *2023 – present*
Curriculum Developer, Sci-MI, EE Division *2024 – present*
Course Instructor, Stanford Splash! *2022 - present*
Taught computer vision course to high-school students from under-resourced communities in the Bay Area.
Core Member, Stanford Association of Computing Machinery (ACM) Machine Learning Lab *2022 – present*

Skills

C/C++, Matlab, Python, Java, Pytorch, Tensorflow, CUDA, OpenCV, OpenGL, HLS, Javascript, HTML, Rust, ARM Assembly, HSPICE, Vivado Design Suite, Perl, Unity, HoudiniFX, Robotics Operating System (ROS), Breadboarding.