

Joshua Park

B.A. in Computer Science and Mathematics, M.A. in Statistics

jhpark@college.harvard.edu

Updated June 4, 2025

TECHNICAL INTERESTS

HW-SW Co-Design, Machine Learning Model Compression, Computer Architecture, GPU Computing

EDUCATION

Master of Arts (A.M.), Statistics

Harvard University - Cambridge, Massachusetts

GPA: 3.96/4.0

Aug. 2023 - May 2025

Bachelor of Arts (A.B.), Computer Science and Mathematics

Summa Cum Laude, Phi Beta Kappa

Harvard College - Cambridge, Massachusetts

Advisor: [Prof. Gu-Yeon Wei](#) and [David Brooks](#)

GPA: 4.0/4.0

Aug. 2021 - May 2025

AWARDS & HONORS

Sophia Freund Prize, Harvard College

May 2025

Awarded annually to the students in the senior class of Harvard College who are graduating summa cum laude with the highest grade point average.

U.S. Presidential Scholar, U.S. Department of Education

May 2021

Each year, up to 161 students are named as Presidential Scholars, one of the nation's highest honors for high school students.

Ky and Yu-Fen Fan Scholarship, American Mathematical Society

Aug. 2020

Presented by the American Mathematical Society to a mathematically talented student to attend Program in Mathematics for Young Scientists (PROMYS).

TECHNICAL SKILLS

Programming Languages/Frameworks: C/C++, Python, CUDA

Hardware Description Languages: Verilog

DL Frameworks/Libraries: PyTorch, cuDNN

WORK EXPERIENCE

Deep Learning Software Engineering Intern

Summer 2024, 2025

NVIDIA Corporation - Santa Clara, CA

- Optimized **cuDNN** kernel selection to improve performance of machine learning workloads in PyTorch, TensorFlow, and JAX.
- Enhanced cuDNN forward compatibility capabilities for future GPU architectures by extending the support surface.
- Designed and implemented a compile-time reduction feature to accelerate execution of shape-similar operations.

Research Assistant

Oct. 2021 - May 2025

Architecture, Circuits and Compilers Laboratory, Harvard University - Cambridge, MA

- Contributed to [GoldenEye](#), a testing framework for evaluating neural network resiliency under custom number systems.
- Wrote and optimized C driver code for ATROPOS, a BERT inference accelerator, with roughly 64× speedup compared to CPU.
- Theoretically proved optimal sparsity-quantization ordering; proposed and validated Quantization-Aware Sparsification (QAS) for improved accuracy in quantized models.

	Quantitative Trader Intern	Winter 2024
	Five Rings Capital - New York, NY	
	<ul style="list-style-type: none"> • Participated in mock trading to obtain familiarity with how the market operates at the level of individual orders. • Competed in automated trading strategy and design competition. • Received classroom-style instruction on a wide range of financial concepts. 	
TEACHING EXPERIENCE	CS 124: Data Structures and Algorithms	Spring 2025
	Teaching Fellow, Instructors: Madhu Sudan and Sitan Chen	
	CS 181: Machine Learning	Spring 2024, 2025
	Teaching Fellow, Instructors: Finale Doshi-Velez and David Alvarez-Melis	
	Statistics 110: Probability	Fall 2024
	Teaching Fellow, Instructor: Joe Blitzstein	
	CS 61: Systems Programming and Machine Organization	Fall 2023
	Teaching Fellow, Instructor: James Mickens	
	CS 51: Abstraction and Design in Computation	Spring 2023
	Teaching Fellow, Instructor: Stuart Shieber	
	CS 50: Introduction to Computer Science	Fall 2022
	Teaching Fellow, Instructor: David Malan	
	Art of Problem Solving	May - Oct. 2022
	Grader and Teaching Assistant	
PUBLICATIONS	<p>Joseph Zuckerman, Martin Cochet, Maico Cassel dos Santos, Erik Jens Loscalzo, Karthik Swaminathan, Tianyu Jia, Davide Giri, Thierry Tambe, Jeff Jun Zhang, Alper Buyuktosunoglu, Kuan-Lin Chiu, Giuseppe Di Guglielmo, Paolo Mantovani, Luca Piccolboni, Gabriele Tombesi, David Trilla, John-David Wellman, En-Yu Yang, Aporva Amarnath, Ying Jing, Bakshree Mishra, Joshua Park, Vignesh Suresh, Samira Zaliasl, Michael Lekas, Stephen Cahill, Hesam Sadeghi, Joseph Meyer, Noah Sturcken, Sarita Adve, David Brooks, Gu-Yeon Wei, Kenneth L. Shepard, Pradip Bose, and Luca P. Carloni. 2025. EPOCHS-1: A 12nm Highly Heterogeneous Open-Source SoC with Distributed Coin Based Power Management and Integrated Hybrid Voltage Regulation. <i>IEEE Journal of Solid-State Circuits (JSSC)</i>.</p> <p>Maico Cassel dos Santos, Tianyu Jia, Joseph Zuckerman, Martin Cochet, Davide Giri, Erik Jens Loscalzo, Karthik Swaminathan, Thierry Tambe, Jeff Jun Zhang, Alper Buyuktosunoglu, Kuan-Lin Chiu, Giuseppe Di Guglielmo, Paolo Mantovani, Luca Piccolboni, Gabriele Tombesi, David Trilla, John-David Wellman, En-Yu Yang, Aporva Amarnath, Ying Jing, Bakshree Mishra, Joshua Park, Vignesh Suresh, Sarita Adve, Pradip Bose, David Brooks, Luca P. Carloni, Kenneth L. Shepard, Gu-Yeon Wei. 2024. A 12nm Linux-SMP-Capable RISC-V SoC with 14 Accelerator Types, Distributed Hardware Power Management and Flexible NoC-based Data Orchestration. <i>IEEE International Solid-State Circuits Conference (ISSCC)</i>, Vol. 67, pp. 262-264.</p>	
TALKS	Joshua Park. 2025. Quantization, Sparsity, and Reliability. <i>Mohamed bin Zayed University of Artificial Intelligence</i> .	
ACTIVITIES	Student and Researcher	Summer 2019, 2020
	Program in Mathematics for Young Scientists	