HARSH SHARMA

harshari.github.io • harsh.sharma@wsu.edu

SUMMARY

Doctoral candidate in ML-Computing systems for Energy Efficient Design. The focus of my research is AI-driven design and optimization of *chiplet system*. Specific topics include:

- Enabling server-scale systems with Dataflow-aware PIM architectures (First work employing Space-filling Curves for interconnect design) **[Best Paper Award]**
- Hardware/Software co-design and energy efficient training/inferencing large AI models.
- Experience with run-time thermal management and limiting the impact of voltage drooping using machine learning.
- Yield and traceability of chiplet systems. Published work on high-performance interconnection design even in the presence of defective chiplets.

EDUCATION

Ph.D. in Computer Engineering **Expected Fall 2025**

Advisor: Partha Pande; Jana Doppa

Washington State University *Pullman, Washington*

Coursework: • Advanced Computer Architecture • Machine Learning • HW/SW Codesign • Neural Network Design & Analysis • SoC Design and Test • VLSI Systems Design

Bachelor of Engineering, Electronics **2017–2021**

NSIT, Delhi University *New Delhi, India*

Department ranker (Top 5%)

AWARDS AND HONORS

- **Outstanding RA Award**, Voiland College of Engineering and Architecture, WSU, 2024.
- **NSF Travel Grant**, 2023, 2024.
- **Best Graduate Researcher Award**, EECS, 2024.
- Harvard Scholar at the HPAIR Conference, Kazakhstan. Technology Track
- ACM SIGDA Richard Newton Young PhD Fellowship, 2022.

SELECTED PUBLICATIONS

1. **H. Sharma**, A. Kalyanraman, U. Ogras, P. Pande. A Dataflow-Aware Network-on-Interposer for CNN Inferencing in the Presence of Defective Chiplets. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD-IC)*, 2024.**[Best Paper Award]; Top 1%**
2. **H. Sharma**, G. Narang, J. Doppa, U. Ogras, P. Pande. Dataflow-aware Interconnect design for DNN accelerators. *Design Automation and Test in Europe Conference (DATE)*, Spain, 2024.
3. **H. Sharma**, L. Pfromm, R. Topaloglu, J. Doppa, U. Ogras, A. Kalyanraman, P. Pande. Florets for Chiplets: Data Flow-aware High-Performance and Energy-efficient Network-on-Interposer for CNN Inference Tasks. *ACM Transactions on Embedded Computing Systems (TECS)*, 2023. **[Best Paper Award]; Top 1% in ACM**
4. **H. Sharma**, S. Mandal, J. Doppa, U. Ogras, P. Pande. Achieving Datacenter-scale Performance through Chiplet-based Manycore Architectures. *Design Automation and Test in Europe Conference (DATE)*, Belgium, 2023.
5. **H. Sharma**, S. Mandal, J. Doppa, U. Ogras, P. Pande. SWAP: A Server-Scale Communication-Aware Chiplet-Based Manycore PIM Accelerator. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD-IC)*, 2022. **[Best Paper Award]; Top 1% in IEEE**

INDUSTRIAL EXPERIENCE

Co-Op Ph.D. Intern **August 2024–Present**

AMD Advanced Packaging Research Group *Austin, Texas*

- Black-Box Design Space Exploration for next generation of MI products. (NDA).
- Interface for enabling end-to-end Traceability using graph structures.
- Key skills: Neural Networks, Machine Learning, GNN, Database Design, Snowflake, Statistics, Efficient Design-Space-Exploration, Optimization.
- Rating: Exceeded expectations.

SELECTED PROFESSIONAL AND OUTREACH ACTIVITIES

Invited Talks

- *Towards Large-Scale composable chiplet system design for Datacenters* at AMD, Austin, Nov. 24.
- *2.5D system design employing defective chiplets* at UNC, Raleigh, Oct. 2024.
- *Heterogenous architecture-package co-optimizations for Datacenters* at Microsoft, Bellevue, Sept. 24.
- *Accelerating the Future of Electronics and Beyond*, IISc, Bangalore, Jan 2024. [‡]
- *Florets for Chiplets to Accelerate DNNs* at ESWEEK, Germany, Oct. 2023.
- *Server-scale Communication-Aware Chiplet Systems* at Boston University, Apr 2023. [§]

JC Bose Science Dialogues at India Science Festival, IISER, Pune, Jan 2024

- Engaged in invitation-only roundtable for open discussion among science and technology leaders on the call on India's academic institutions to evolve.
- **Outcome**: Contributed to a three-page document outlining key insights and strategies, submitted to the Prime Minister's Office and the Office of the Principal Scientific Advisor, influencing national science and technology policies and actions.

Journal and Conference Reviewer

- Design Automation Conference (DAC), DATE, AAAI, ESWEEK, ICCAD
- TCAD-IC, TODAES, TECS, TVLSI, Design and Test

SKILLS

- **Programming Languages.** Python, C/C++, MATLAB
- **Tools/Packages.** BookSim2, COMPASS, MEMSYS, Gem5, HeteroGarnet, Ansys, State-Space modelling, ASTRA-SIM, RAMULATOR, Hotspot 6.0, NeuroSim, PyTorch, Python data science tools, ARM Microcontrollers, Xilinx FPGA, Deep Learning, Javascript, FrontEnd Design, Raspberry Pi, Embedded System Design, Design-space exploration, Topology construction, GNNs.

[‡]Link `https://www.csa.iisc.ac.in/~skmandal/speakers.html`

[§]Based on `https://medium.com/@harshari/accelerating-the-future-of-electronics-e23cc42d9d39`