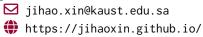
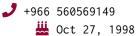
Jihao Xin









Introduction

I'm a PhD student in computer science, working with Prof. Marco Canini at SANDS Lab. My research focuses on accelerating large scale AI through algorithm-system co-designed approaches. I'm recently working on quantization, GPU communication, and efficient LLM inference.

Education

January 2023 - Present

Ph.D. student in Computer Science, KAUST, Saudi Arabia.

August 2021 – December 2022

Master of Science in Computer Science, KAUST, Saudi Arabia.

Dean's List Award

GPA = 3.88/4.0

October 2020 - October 2021

Master of Science in Applied Computational Science and Engineering, Imperial College London, UK.

Distinction Degree

September 2016 – June 2020

Bachelor of Engineering in Computer Science and Technology, Shandong University, China.

GPA = 89.34/100

Employment & Internship

September 2023 – March 2024

Research Intern, Microsoft Research Asia, Beijing, China. Supervised by Dr. Changho Hwang Efficient GPU communication for distributed AI tasks.

May 2022 – August 2022

Research Intern, ISTAustria, Klosterneuburg, Austria.
Supervised by Prof. Dan Alistarh
TopK compression with efficient CUDA-based threshold selection.

August 2019 – November 2019

Engineering Intern, Tencent Cloud, Shenzhen, China.
Tencent international website backend development by GoLang.

June 2019 – August 2019

Research Intern, The University of Western Australia, Remote. YOLO-based splash detection of Western Australia's coastal highway by Python.

Research Publications

- J. Weng, B. Han, D. Gao, R. Gao, W. Zhang, A. Zhong, C. Xu, **J. Xin**, Y. Luo, L. W. Wills, and M. Canini, "Assassyn: A unified abstraction for architectural simulation and implementation," in *Proceedings of the 52nd Annual International Symposium on Computer Architecture*, ser. ISCA '25, Association for Computing Machinery, 2025, pp. 1464–1479, ISBN: 9798400712616. ODI: 10.1145/3695053.3731004.
- J. Xin, M. Canini, P. Richtárik, and S. Horváth, "Global-qsgd: Allreduce-compatible quantization for distributed learning with theoretical guarantees," in *Proceedings of the 5th Workshop on Machine Learning and Systems*, ser. EuroMLSys '25, Rotterdam, Netherlands: Association for Computing Machinery, 2025, ISBN: 979-8-4007-1538-9/2025/03. ODI: 10.1145/3721146.3721932.

- J. Xin, S. Bae, K. Park, M. Canini, and C. Hwang, "Immediate communication for distributed ai tasks," in *The 2nd Workshop on Hot Topics in System Infrastructure*, 2024. OURL: https://hotinfra24.github.io/papers/hotinfra24-final2.pdf.
- M. Mao, X. Feng, **J. Xin**, and T. W. S. Chow, "A convolutional neural network-based maximum power point voltage forecasting method for pavement pv array," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–9, 2023. ODI: 10.1109/TIM.2022.3227552.
- J. Xin, I. Ilin, S. Zhang, M. Canini, and P. Richtárik, "Kimad: Adaptive gradient compression with bandwidth awareness," in *Proceedings of the 4th International Workshop on Distributed Machine Learning*, ser. DistributedML '23, Paris, France: Association for Computing Machinery, 2023, pp. 35–48, ISBN: 9798400704475. ODI: 10.1145/3630048.3630184.
- N. Xue, D. Guo, J. Zhang, **J. Xin**, Z. Li, and X. Huang, "Openfunction for software defined iot," in 2021 International Symposium on Networks, Computers and Communications (ISNCC), 2021, pp. 1–8. DOI: 10.1109/ISNCC52172.2021.9615751.

Skills

Programming C++/CUDA, Python/PyTorch, GoLang

Systems Linux, GPU Programming, MPI, NCCL

Languages English (Professional), Chinese (Native)

Miscellaneous Experience

Teaching Asistant

2025 Spring KAUST CS398 Graduate Seminar

2022 Fall, 2024 Fall KAUST CS240: Computing Systems and Concurrency

Saudi Ministry of Interior Training Program CS229: Machine

Learning

2022 Summer, 2023 Summer UESTC&KAUST Summer Camp

Seasonal Programs

2019 Summer Research Program, The University of Western Australia.

2019 Winter Robotics & AI Winter School, Imperial College London Hamlyn Center.

Leadership

2024-2025 KAUST Graduate Student Council Representative (12/1683).