


XINYU LIAN

lian7@illinois.edu | xinyu86.github.io

RESEARCH INTEREST

I study machine learning and systems, focus on **efficient and elastic LLM (post)-training**, with a current emphasis on **RL training and efficient training kernels**. I am also a core contributor and maintainer to **DeepSpeed**  **Star 40k**. My work has been adopted by Microsoft, BigScience, Baidu, UC Berkeley, Oak Ridge and Argonne National lab for end-to-end training of real-world models up to 176B parameters.

EDUCATION

University of Illinois Urbana-Champaign

Ph.D. in Computer Science, GPA: 4.0/4.0

Aug. 2024 – Present

Advisor: Minjia Zhang

University of Illinois Urbana-Champaign

MS in Computer Science (Research-Oriented)

Aug. 2022 – May 2024

Advisors: Tianyin Xu and Darko Marinov

Zhejiang University

B.Eng. in Electronic and Computer Engineering

Sept. 2018 – June 2022

SELECTED PUBLICATIONS

- [1] **ASPLOS'26 SuperOffload: Unleashing the Power of Large-Scale LLM Training on Superchips**
Xinyu Lian, Masahiro Tanaka, Olatunji Ruwase, Minjia Zhang
31st ACM International Conference on Architectural Support for Programming Languages and Operating System
Featured by PyTorch Official Blog and invited for presentation at PyTorch Conference 2025
- [2] **ATC'25 Universal Checkpointing: Efficient and Flexible Checkpointing for Large Scale Distributed Training**
Xinyu Lian, Sam Ade Jacobs, Lev Kurilenko, Masahiro Tanaka, Stas Bekman, Olatunji Ruwase, Minjia Zhang
2025 USENIX Annual Technical Conference
Adopted by Microsoft for Phi-3.5-MoE (42B) Pre-Training, BigScience for BLOOM (176B) Pre-Training
Invited for presentation at PyTorch Day France 2025, FMS 2025, SDC 2025
- [3] **ICSE'25 Large Language Models as Configuration Validators**
Xinyu Lian, Yinfang Chen, Runxiang Cheng, Jie Huang, Parth Thakkar, Minjia Zhang, Tianyin Xu
47th International Conference on Software Engineering
- [4] **FSE'24 Ctest4J: A Practical Configuration Testing Framework for Java**
Shuai Wang, **Xinyu Lian**, Qingyu Li, Darko Marinov, Tianyin Xu
32nd ACM International Conference on the Foundations of Software Engineering, Demo Papers
- [5] **ICSE'23 Test Selection for Unified Regression Testing**
Shuai Wang, **Xinyu Lian**, Darko Marinov, Tianyin Xu
45th International Conference on Software Engineering

WORK EXPERIENCE

Research Intern at Microsoft

May 2025 – Aug. 2025

Topic: Quantization-Aware Training for OpenAI (style) models for efficient deployment on Azure

Increased the trainable model size by 80x and throughput by 10x; selected to present to entire organization and CVP.

AWARDS & HONORS

Amazon Fellowship (\$70K)	2025
USENIX Student Grant	2025
SIGSOFT CAPS Travel Grant	2025
Provincial Outstanding Graduate	2022
Academic Scholarship, Zhejiang University (¥10K)	2021

ACADEMIC SERVICES

Organizer:	Brett Daniel Software Engineering Seminar
Program Committee:	MSR'25
Artifact Evaluation Committee:	SOSP'24