

Sijia Liu

Assistant Professor

Department of Computer Science & Engineering
Michigan State University, East Lansing, MI

Affiliated Professor

MIT-IBM Watson AI Lab, IBM Research

Email: liusiji5@msu.edu

Tel: (315)-744-6778 (Mobile)

[\[Personal Website\]](#)

[\[OPTML Lab Website\]](#)

[\[Google Scholar\]](#)

PRIMARY RESEARCH AREAS

Trustworthy ML: Adversarial ML, model explanation, fairness, security & privacy

Scalable ML: Zeroth-order optimization, deep model compression, distributed ML, automated ML

EDUCATION

Ph.D., Electrical and Computer Engineering, Syracuse University Mar. 2016

All University Doctoral Prize; Advisors: [Pramod Varshney](#) and [Makan Fardad](#)

M. A. Sc., Electrical Engineering, Xi'an Jiaotong University May 2011

B.S., Electrical Engineering, Xi'an Jiaotong University May 2008

PROFESSIONAL EXPERIENCE

Assistant Professor, CSE, Michigan State University Jan. 2021 – present

Affiliated Professor, MIT-IBM Watson AI Lab, IBM Research Oct. 2021 – present

Research Staff Member, MIT-IBM Watson AI Lab, IBM Research Jan. 2018 – Dec. 2020

Postdoc Research Fellow, University of Michigan, Ann Arbor July 2016 – Dec. 2017

Supervisors: [Alfred Hero](#) (EECS) and [Indika Rajapakse](#) (Computational Medicine & Bioinformatics)

HONORS AND RECOGNITION

Withrow Rising Scholar Award, Michigan State University, 2025

— *This prestigious award annually recognizes junior faculty for excellence in instruction, scholarship, and distinguished service*

IBM PhD Fellowship Award, PhD Advisor of the Award Recipient [Yihua Zhang](#), 2025

NAIRR Pilot Award (Artificial Intelligence and Intelligent Systems), 2024

— *For the project titled “Enhancing Large Language Model Unlearning Across the Lifecycle”*

Amazon Research Award (AI for Information Security), 2024

— *For the project titled “Fostering Trustworthy Generative AI: The Role of Machine Unlearning”*

National Science Foundation (NSF) CAREER Award, 2024

— *For the project titled “Zeroth-Order Machine Learning: Foundations and Emerging AI Applications”*

AAAI’23 New Faculty Highlights on “General and Scalable Optimization for Robust AI”, 2023

Best Paper Runner-Up Award at 38th Conference on Uncertainty in Artificial Intelligence (UAI), 2022

— *For the paper titled “Distributed Adversarial Training to Robustify Deep Neural Networks at Scale”*

IBM Pat Goldberg Best Paper Award Finalist, 2020

— *For the AAAI’20 paper titled “An ADMM Based Framework for AutoML Pipeline Configuration”, the key enabling technique in the IBM Watson Studio Automated ML System*

Three IBM Outstanding Research Accomplishments, 2019

— *Trustworthy AI; Toward Automating the AI Lifecycle with AutoAI; Deep Learning on Graphs*

Best Student Paper Award at 42nd ICASSP, 2017

— For the paper titled “Ultra-fast Robust Compressive Sensing Based on Memristor Crossbars”

Best Student Paper Award Finalist at Asilomar Conference on Signals, Systems, and Computers, 2013

— For the paper titled “Adaptive Non-myopic Quantizer Design for Target Tracking in Wireless Sensor Networks”

First Class Award in National Mathematics Olympiad, 2004

SELECTED PUBLICATIONS

Full publication list can be found at [Google Scholar](#) (11,149 citations as of April 13, 2025). **CSRanking score**: 80

* denotes equal contribution; † denotes student authors under my supervision.

Five Representative Publications in *Trustworthy ML*:

- T5. **S. Liu**, Y. Yao*, J. Jia^{†,*}, S. Casper, N. Baracaldo, P. Hase, Y. Yao[†], C. Y. Liu, X. Xu, H. Li, K. R. Varshney, M. Bansal, S. Koyejo, Y. Liu, “Rethinking Machine Unlearning for Large Language Models.” *Nature Machine Intelligence*, 2025, pp.181–194
- T4. Y. Zhang^{†,*}, J. Jia^{†,*}, X. Chen, A. Chen[†], Y. Zhang[†], J. Liu[†], K. Ding, **S. Liu**, “To Generate or Not? Safety-Driven Unlearned Diffusion Models Are Still Easy To Generate Unsafe Images ... For Now.” *European Conference on Computer Vision (ECCV)*, 2024, pp.385–403
- T3. C. Fan^{†,*}, J. Liu^{†,*}, Y. Zhang[†], E. Wong, D. Wei, **S. Liu**, “SalUn: Empowering Machine Unlearning via Gradient-based Weight Saliency in Both Image Classification and Generation.” *International Conference on Learning Representations (ICLR)*, 2024 (**Spotlight**)
- T2. J. Jia^{†,*}, J. Liu^{†,*}, P. Ram, Y. Yao[†], G. Liu, Y. Liu, P. Sharma, **S. Liu**, “Model Sparsity Can Simplify Machine Unlearning.” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023, pp.51584–51605 (**Spotlight**)
- T1. Y. Zhang^{†,*}, G. Zhang^{†,*}, P. Khanduri, M. Hong, S. Chang, **S. Liu**, “Revisiting and advancing fast adversarial training through the lens of bi-level optimization.” *International Conference on Machine Learning (ICML)*, 2022, pp.26693–26712

Five Representative Publications in *Scalable ML*:

- S5. A. Chen^{†,*}, Y. Zhang^{†,*}, J. Jia[†], J. Diffenderfer, J. Liu[†], K. Parasyris, Y. Zhang[†], Z. Zhang, B. Kailkhura, **S. Liu**, “DeepZero: Scaling up Zeroth-Order Optimization for Deep Model Training.” *International Conference on Learning Representations (ICLR)*, 2024
- S4. Y. Zhang^{*,†}, Y. Yao^{*,†}, P. Ram, P. Zhao, T. Chen, M. Hong, Y. Wang, **S. Liu**, Advancing Model Pruning via Bi-level Optimization, *Advances in Neural Information Processing Systems (NeurIPS)*, 2022, pp.18309–18326
- S3. G. Zhang^{†,*}, S. Lu*, Y. Zhang[†], X. Chen, P.-Y. Chen, Q. Fan, L. Martie, L. Horesh, M. Hong, **S. Liu**, “Distributed Adversarial Training to Robustify Deep Neural Networks at Scale.” *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2022, pp.2353–2363 (**the Best Paper Runner-Up Award**)
- S2. **S. Liu***, S. Lu*, X. Chen*, Y. Feng, K. Xu, A. Al-Dujaili, M. Hong, U.-M. O’Reilly, “Min-Max Optimization without Gradients: Convergence and Applications to Adversarial ML.” *International Conference on Machine Learning (ICML)*, 2020, pp.6282–6293
- S1. **S. Liu**, P.-Y. Chen, B. Kailkhura, G. Zhang, A. O. Hero, P. K. Varshney, “A Primer on Zeroth-Order Optimization in Signal Processing and Machine Learning.” *IEEE Signal Processing Magazine*, 2020, pp.43–54

SELECTED TALKS/PRESENTATIONS

- T1. “Machine Unlearning in Computer Vision: Foundations and Applications.” *CVPR’24 Tutorial*, 06/2024
- T2. “Zeroth-Order Machine Learning: Fundamental Principles and Emerging Applications in Foundation Models.” *AAAI’24 Tutorial*, 02/2024
- T3. “Reverse Engineering of Deceptions: Foundations and Applications”, *CVPR’23 Tutorial*, 06/2023
- T4. “Bi-level Optimization in Machine Learning: Foundations and Applications.” *AAAI’23 Tutorial*, 02/2023
- T5. “Foundational Robustness of Foundation Models”, *NeurIPS’22 Tutorial*, 12/2022