

Xinwen(Ellen) Zhang

Ph.D. Student in Temple University

✉: ellenz@temple.edu

🖥: [Personal Site](#)

🔗: [LinkedIn](#)

EDUCATION

Temple University

Ph.D. in Computer and Information Science

Advisor: Dr. Hongchang Gao

Research Interest: Design large-scale optimization algorithms, including stochastic optimization beyond minimization, distributed optimization, and efficient optimizers for LLMs.

Philadelphia, USA

Aug.2022 - May.2027

Tongji University

B.S. in Mathematics

Shanghai, China

Sep.2017 - Jun.2021

TECHNICAL SKILLS

- **Programming Language:** Python, C++, Matlab, SAS, SQL, LaTeX.
- **Tools and Frameworks:** PyTorch, TensorFlow, JAX, Scikit-learn, MPI, PyTorch DDP, CUDA, Docker, Git, Weights & Biases, Matplotlib, Seaborn.
- **Sound Background in Computer Science:** Advanced understanding in Machine Learning, Optimization Theory, Neural Computation, Design & Analysis of Algorithm, Artificial Intelligence, Data Mining, Graph, Large Language Model, etc. Experienced in large-scale model implementation, GPU-based distributed training, and experimental reproducibility on multi-GPU clusters and Linux environments.
- **Well-established Knowledge in Mathematics:** Proficiency in the Abstract Algebra, Real Analysis, Complex Analysis, Differential Equations, Number Theory, Topology, Differential Geometry, with expertise in constructing and understanding mathematical proofs and theorems. In-depth knowledge of optimization theory, probability, and stochastic processes, with research experience in convergence analysis for stochastic and distributed algorithms.
- **Applied Modeling and Data Analysis:** Experience in designing and evaluating large-scale learning systems for deep AUC maximization, multi-instance learning, hyperparameter optimization, and model pruning tasks. Applied proposed algorithms to image, text, and medical datasets, involving models such as CNNs, RNNs, Transformers, and LLMs, to address challenges including data imbalance, heterogeneity, and heavy-tailed noise. Skilled in experiment tracking, visualization, and reproducibility analysis using structured pipelines and scientific visualization tools.

RESEARCH EXPERIENCE

Ph.D. Research: Large-Scale Optimization

Aug 2022 - Present

Advisor: Prof. HongChang Gao (School of Computer and Information Science, Temple University)

- **Convergence Analysis for Stochastic Optimization Beyond Minimization:** (*Paper published; Tutorial*)
 - Investigating the convergence behavior of stochastic optimization algorithms beyond minimization, including minimax, compositional, bilevel, and multi-level problems.
 - Studying convergence acceleration techniques, such as moving-average momentum, variance reduction, and smoothing, and exploring the connection between loss landscape geometry and generalization to better understand optimization stability.
 - Applying the developed theoretical frameworks to diverse machine learning tasks, including deep AUC maximization, multi-instance learning, hyperparameter optimization, and model pruning tasks.
- **Large-Scale Distributed Optimization:** (*Paper published; Tutorial*) Co-Advisor: Prof. Richard Souvenir
 - Designing efficient algorithms for large-scale distributed models, such as federated and decentralized learning, with an emphasis on handling decentralized topologies and heterogeneous data distributions cross clients.
 - Proposed provably convergent methods with improve computation and communication efficiency, while maintaining strong robustness, demonstrating consistent performance across both federated and decentralized settings.
- **Efficient Optimizers for Large Language Models:** (*Preprint*)
 - Studying the design and theoretical foundations of modern optimizers for large language models. Explored a wide range of emerging algorithms, including Adam-based, sign-based, matrix-based, parameter-free methods, and variance-reduced methods, aiming to improve convergence efficiency and training stability for large-scale neural networks.
 - Particularly focus on Muon, investigating its convergence stability and robustness under heavy-tailed noise, as well as its behavior in distributed and large-batch training environments.

RESEARCH PUBLICATIONS

[NeurIPS 2025] Xinwen Zhang, Hongchang Gao. "On the Convergence of Stochastic Smoothed Multi-Level Compositional Gradient Descent Ascent." In *Proceedings of the 39th Conference on Neural Information Processing Systems, 2025*.

[ICDM 2025] Xinwen Zhang, Hongchang Gao. "Sharpness-Aware Optimization Through Variance Suppression on Deep AUC Maximization." In *2025 IEEE International Conference on Data Mining, 2025*.

[IWQoS 2025] Jiyao Liu, **Xinwen Zhang**, Xinliang Wei, Xuanzhang Liu, Yuzhou Chen, Hongchang Gao, Yu Wang. “Joint Swapping and Purification with Failures for Entanglement Distribution in Quantum Networks.” *In 2025 IEEE/ACM 33rd International Symposium on Quality of Service*, 2025.

[ICML 2024] **Xinwen Zhang**, Ali Payani, Myungjin Lee, Richard Souvenir, Hongchang Gao. “A Federated Stochastic Multi-Level Compositional Minimax Algorithm for Deep AUC Maximization.” *In Proceedings of the 41th International Conference on Machine Learning*, 2024.

[NeurIPS 2023] **Xinwen Zhang**, Yihan Zhang, Tianbao Yang, Richard Souvenir, Hongchang Gao. “Federated Compositional Deep AUC Maximization.” *In Proceedings of the 37th Conference on Neural Information Processing Systems*, 2023.

Preprints

[arXiv 2025] **Xinwen Zhang**, Yihan Zhang, Hongchang Gao. “Nonconvex Decentralized Stochastic Bilevel Optimization under Heavy-Tailed Noises.” *arXiv preprint arXiv:2509.15543*.

[arXiv 2025] **Xinwen Zhang**, Hongchang Gao. “On Provable Benefits of Muon in Federated Learning.” *arXiv preprint arXiv:2510.03866*.

TUTORIALS & WORKSHOPS

[ICDM’25 Tutorial] Hongchang Gao, **Xinwen Zhang**. “Federated Stochastic Compositional and Bilevel Optimization.” *In 2025 IEEE International Conference on Data Mining*, 2025.

[IJCAI’25 Tutorial] Hongchang Gao, **Xinwen Zhang**. “Federated Stochastic Compositional and Bilevel Optimization.” *In Proceedings of the 34th International Joint Conference on Artificial Intelligence*, 2025.

[WiML’23 Workshop] **Xinwen Zhang**, Richard Souvenir, Hongchang Gao. “A Fast Federated Stochastic Compositional Minimax Optimization Algorithm with Variance Reduction.” *18th Women in Machine Learning Workshop*, 2023.

[KDD’23 Tutorial] Hongchang Gao, **Xinwen Zhang**. “Distributed Optimization for Big Data Analytics: Beyond Minimization.” *In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2023.

WORK EXPERIENCE

Teaching Assistant in Temple University

Philadelphia, Aug 2022 - Dec 2024

- CIS3715 Principles of Data Science (Spring 2023, Spring 2024)
- CIS1051 Introduction to Problem Solving and Programming in Python (Fall 2022, Spring 2023, Fall 2023, Spring 2024, Fall 2024)

Data Analysis Intern in Zhongyuan Consumer Finance

Shanghai, Mar 2021 - Jun 2021

- Developed Python and SQL pipelines for automated credit monitoring and risk assessment, and built user profiling and targeting models from multi-source behavioral and media data to optimize marketing strategy.
- Enhanced the flexibility and automation of the credit evaluation process by introducing hierarchical detection based on advertising schedules and temporal patterns.
- Conducted multivariate statistical analysis on user engagement data from a festival campaign, demonstrating that cash incentive activities significantly improved customer retention.

SERVICES

• Conference Reviewer:

- 2026: ICLR, AISTATS.
- 2025: AAAI, ICML, NeurIPS, KDD, ICLR, AISTATS.
- 2024: AAAI, ICML, NeurIPS.

- **Journal Reviewer:** Journal of Combinatorial Optimization.

AWARDS

- KDD 2023 Student Travel Award, NeurIPS 2023 Scholar Award, WiML 2023 Travel Funding, ICML 2024 Travel Award.
- Second Prize, Shanghai Mathematical Model Competition.
- Academic Scholarships, Three Consecutive Years in Tongji University.