

Zhen (Zoe) Qin

CONTACT INFORMATION

Address: Sammamish, WA (Seattle Area)
Personal Homepage: <https://zhenqinzoe.github.io/>

Email: zoezqin@gmail.com
Phone: +1 (626)-677-6551

EDUCATION

The Ohio State University

Ph.D. Candidate in Electrical and Computer Engineering

- Advisors: Prof. Jia (Kevin) Liu and Prof. Yingbin Liang

08/2023 – 05/2028 (Expected)

University of Southern California

M.A. in Applied Mathematics

01/2020 – 12/2021

Lanzhou University

B.S. in Mathematics and Applied Mathematics

09/2014 – 07/2018

RESEARCH INTERESTS

- **First- and Second-Order Stochastic Optimization:** Designing optimization algorithms to improve training efficiency and empirical performance, supported by rigorous theoretical analysis.
- **Distributed and Decentralized Learning:** Designing algorithms to tackle computational and communication efficiency challenges in distributed and decentralized learning.
- **Bilevel and Multi-Objective Optimization:** Designing efficient algorithms for bilevel learning and Pareto front exploration, with convergence and complexity guarantees.

SELECTED RESEARCH PROJECTS

Efficient Bilevel Optimization

2023 – Present

- Proposed **DUET**, a decentralized bilevel optimization algorithm that enables collaborative learning without requiring strongly convex lower-level objectives.
 - * **Key idea:** Using diminishing quadratic regularization to ensure unique regularized lower-level solutions, and gradient tracking to handle data heterogeneity.
 - * **Theoretical results:** Introduced a stationarity measure for the non-strongly-convex lower-level setting and established a finite-time convergence rate.
 - * **Empirical evaluation:** Validated DUET on decentralized meta-learning under both IID and non-IID data distributions, and on decentralized hyperparameter optimization.
- Proposed **BiMuon**, a Muon-style stochastic bilevel optimizer that incorporates the matrix geometry of hypergradients into upper-level updates.
 - * **Key idea:** Addressing hypergradient spectral imbalance induced by lower-level curvature through Muon-style momentum orthogonalization. Extending the method to multiple matrix blocks with learning rates balancing upper-level progress and lower-level sensitivity.
 - * **Theoretical results:** Established convergence rates for two BiMuon variants using exact orthogonalization and practical Newton–Schulz approximation, respectively, and derived a principled rule for blockwise learning rate allocation.
 - * **Empirical evaluation:** Validated matrix-parameter learning and blockwise learning rate allocation; practical Newton–Schulz updates achieved a $3.38\times$ speedup over exact orthogonalization across 60 updates.
- Proposed **DiSCO** for constrained stochastic decentralized bilevel optimization without lower-level strong convexity.
 - * **Key idea:** Combining proximal consensus updates with diminishing regularization and variance-reduced gradient tracking.
 - * **Theoretical results:** Established convergence rates and improved sample complexity over the deterministic counterpart.
 - * **Empirical evaluation:** Outperformed direct-projection and Prox-DSGD baselines on constrained decentralized bilevel learning tasks.
- **Related Publications:**
 - **Zhen Qin**, Zhuqing Liu, Songtao Lu, Yingbin Liang, Jia Liu. “DUET: Decentralized Bilevel Optimization without Lower-Level Strong Convexity.” *ICLR*, 2025.
 - **Zhen Qin**, Zhiyao Zhang, Yingbin Liang, Jia Liu. “BiMuon: A Geometry-Aware Stochastic Bilevel Optimization Approach.” *under review*.

- **Zhen Qin**, Zhuqing Liu, Yingbin Liang, Jia (Kevin) Liu. “Constrained Decentralized Bilevel Optimization over Networks: Beyond Lower-Level Strong Convexity.” *Manuscript*.

SELECTED
RESEARCH
PROJECTS
(CONTINUED)

Stochastic and Distributed Optimization

2022 – Present

- Proposed **DisMUSE**, a Muon-style distributed optimizer that combines variance reduction with spectral updates for efficient collaborative learning.
 - * **Key idea:** Addressing the noncommutativity of distributed aggregation and orthogonalization through variance reduction before spectral updates. Developing three variants: exact orthogonalization as the foundation, practical Newton–Schulz approximation to reduce computation, and low-rank momentum transmission to reduce communication.
 - * **Theoretical results:** Established convergence rates for the exact and practical Newton–Schulz variants under their respective assumptions, and quantified how accumulated low-rank compression errors affect convergence.
 - * **Empirical evaluation:** Achieved a $2.02\times$ training speedup with practical Newton–Schulz updates at comparable CIFAR-10 accuracy, and reduced communication payload by 40.5% through low-rank transmission with an accuracy trade-off.
- Investigated why dimension-free zeroth-order communication in federated learning does not directly extend to decentralized gossip.
 - * **Key idea:** Identifying historical information accumulated through gossip mixing that latest seed–scalar updates cannot capture without global model synchronization.
 - * **Theoretical results:** Proved latest neighbor seed–scalar exchange insufficient for exact gossip reconstruction and analyzed the computation cost of full-history reconstruction.
- Studied signSGD with practical random reshuffling and proposed **SignRVR**.
 - * **Key idea:** Using epoch-anchored variance reduction to control gradient sign errors.
 - * **Theoretical results:** Established convergence rates for SignRVR and conditions for SignRR.
 - * **Empirical evaluation:** Improved convergence over SignRR and signSGD across batch sizes, with test accuracy comparable to Adam.
- **Related Publications:**
 - **Zhen Qin**, Zhiyao Zhang, Yingbin Liang, Jia Liu. “DisMUSE: An Efficient Distributed Singular-Spectrum-Based Optimization Approach with Theoretical Convergence Guarantee.” *under review*.
 - **Zhen Qin**, Zhe Li, Bicheng Ying, Haibo Yang, Yingbin Liang, Jia Liu. “Why Dimension-Free Zeroth-Order Communication Does Not Directly Extend to Decentralized Gossip.” *Manuscript*.
 - **Zhen Qin**, Zhishuai Liu, Pan Xu. “Convergence of Sign-based Random Reshuffling Algorithms for Nonconvex Optimization.” *under review*.

OTHER
PUBLICATIONS

- Zhiyao Zhang, Myeung Suk Oh, **Zhen Qin**, Jiaxiang Li, Xin Zhang, Jia Liu. “A Tale of Two Problems: Multi-Task Bilevel Learning Meets Equality Constrained Multi-Objective Optimization.” *ICML*, 2026.

WORK
EXPERIENCES

Teaching Assistant, The Ohio State University

Fall 2026

ECE 8101: Nonconvex Optimization for Machine Learning (Prof. Jia Liu)

Graduate Research Associate, The Ohio State University

08/2023 – Present

Research under the supervision of Prof. Pan Xu, Duke University

01/2022 – 05/2023

PROGRAMMING
AND SKILLS

Machine Learning, Optimization Theory and Algorithms.
Python, LaTeX, PyTorch.

ACADEMIC
SERVICES

ICML Gold Reviewer, 2026

Conference reviewer: AISTATS 2024–2027; ICLR 2025, 2027; ICML 2026.

Journal reviewer: IEEE Trans. Cognitive Communications and Networking; IEEE/ACM Trans. Networking; IEEE Trans. Information Theory.