Yi Zhou, Ph.D.

✉ yi.zhou@utah.edu   ✆ +1(315)-751-6542

G https://scholar.google.com/citations?user=4fK8bYIAAAAJ&hl=en

⌂ https://sites.google.com/site/yizhouhomepage/home

Academic Appointments

2019 – present   **Assistant Professor, University of Utah**
Department of Electrical and Computer Engineering

2018 – 2019   **Postdoctoral Research Associate, Duke University**
Department of Electrical and Computer Engineering
Advisor: Prof. Vahid Tarokh

Education

2013 – 2018   **Ph.D. Electrical & Computer Engineering, The Ohio State University**
Thesis title: *Nonconvex Optimization in Machine Learning: Convergence, Landscape, and Generalization.*
Advisor: Prof. Yingbin Liang

2009 – 2013   **B.S. Electrical Engineering, Beijing Institute of Technology**

Research Interest

Machine Learning   Reinforcement learning, deep learning, AI for science and engineering

Optimization   Nonconvex optimization, stochastic optimization, distributed optimization

Signals & Statistics   Statistical signal processing, robustness and uncertainty

Research Grants

2022   **NSF CAREER-2237830 ($ 549,712)**
PI Role: Leading PI (single)
Project title: Reinforcement Learning-Based Control of Heterogeneous Multi-Agent Systems in Structured Environments: Algorithms and Complexity
Project period: 07/01/2023 – 06/30/2028

DOE-10068105 ($ 40,000)
PI Role: Leading PI (single)
Project title: Black-box PDE Solvers Meet Deep Learning: End-to-End Mesh Optimization for Efficient Fluid Flow Prediction
Project period: 01/01/2023 – 09/30/2023

2021   **NSF CCF-2106216 ($ 411,248)**
PI Role: Leading PI (single)
Project title: Self-Adaptive Optimization Algorithms with Fast Convergence via Geometry Adapted Hyper-Parameter Scheduling
Project period: 07/01/2021 – 06/30/2024

NSF DMS-2134223 ($ 850,000)
PI Role: Leading PI
Project title: Advancing Theoretical Minimax Deep Learning: Optimization, Resilience, and Interpretability
Project period: 09/01/2021 – 08/31/2024

Research Publications

Journal Articles (∗ student author)

1   C. Chen*, J. Zhou, J. Ding, and **Y. Zhou**, "Assisted learning for organizations with limited imbalanced data," *Transactions on Machine Learning Research (TMLR)*, 2023. 🔗 URL: https://openreview.net/forum?id=SEDWlhcFWA.

2   Z. Chen*, Z. Hu, Q. Li, Z. Wang, and **Y. Zhou**, "A cubic regularization approach for finding local minimax points in nonconvex minimax optimization," *Transactions on Machine Learning Research (TMLR)*, 2023. 🔗 URL: https://openreview.net/forum?id=jVMMdg31De.

3   Z. Chen* and **Y. Zhou**, "An accelerated proximal algorithm for regularized nonconvex and nonsmooth bi-level optimization," *Machine Learning*, 2023. 🔗 URL: https://link.springer.com/article/10.1007/s10994-023-06329-6.

4   Y. Zhang, **Y. Zhou**, K. Ji, and M. Zavlanos., "Boosting one-point derivative-free online optimization via residual feedback," *To appear in IEEE Transactions on Automatic Control (TAC)*, 2023. 🔗 URL: https://arxiv.org/abs/2010.07378.

5   Z. Chen*, **Y. Zhou**, and R. Chen, "Multi-agent off-policy tdc with near-optimal sample and communication complexities," *Transactions on Machine Learning Research (TMLR)*, 2022. 🔗 URL: https://openreview.net/forum?id=tnPjQpYk7D&referrer=%5BTMLR%5D(%2Fgroup%3Fid%3DTMLR).

6   Y. Zhang, **Y. Zhou**, K. Ji, and M. M. Zavlanos, "A new one-point residual-feedback oracle for black-box learning and control," *Automatica*, vol. 136, no. C, 2022, ISSN: 0005-1098. 🔗 URL: https://doi.org/10.1016/j.automatica.2021.110006.

7   **Y. Zhou**, Y. Liang, and H. Zhang, "Understanding generalization error of sgd in nonconvex optimization," *Machine Learning*, vol. 111, no. 1, pp. 345–375, 2022. 🔗 URL: https://doi.org/10.1007/s10994-021-06056-w.

8   K. Ji, **Y. Zhou**, and Y. Liang, "Understanding estimation and generalization error of generative adversarial networks," *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 3114–3129, 2021. 🔗 URL: https://ieeexplore.ieee.org/document/9330788.

9   Q. Li, **Y. Zhou**, R. Anirudh, *et al.*, "Mr-gan: Manifold regularized generative adversarial networks for scientific data," *SIAM Journal on Mathematics of Data Science*, vol. 3, no. 4, pp. 1197–1222, 2021. 🔗 URL: https://doi.org/10.1137/20M1344299.

**10**   T. Xu, **Y. Zhou**, K. Ji, and Y. Liang, "When will gradient methods converge to max-margin classifier under relu models?" *Stat*, vol. 10, no. 1, 2021. 🔗 URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/sta4.354.

11   Z. Wang, **Y. Zhou**, Y. Liang, and G. Lan, "A note on inexact gradient and hessian conditions for cubic regularized newton's method," *Operations Research Letters*, vol. 47, no. 2, pp. 146–149, 2019, ISSN: 0167-6377. 🔗 URL: https://www.sciencedirect.com/science/article/pii/S016763771830470X.

12   **Y. Zhou**, Y. Liang, and L. Shen, "A simple convergence analysis of bregman proximal gradient algorithm," *Computational Optimization and Applications*, vol. 73, no. 3, pp. 903–912, 2019. 🔗 URL: https://doi.org/10.1007/s10589-019-00092-y.

13   **Y. Zhou**, Y. Liang, Y. Yu, W. Dai, and E. P. Xing, "Distributed proximal gradient algorithm for partially asynchronous computer clusters," *Journal of Machine Learning Research*, vol. 19, no. 19, pp. 1–32, 2018. 🔗 URL: http://jmlr.org/papers/v19/17-444.html.

14   H. Zhang, **Y. Zhou**, Y. Liang, and Y. Chi, "A nonconvex approach for phase retrieval: Reshaped wirtinger flow and incremental algorithms," *Journal of Machine Learning Research*, vol. 18, no. 1, pp. 5164–5198, 2017. 🔗 URL: https://jmlr.org/papers/v18/16-572.html.

Conference Proceedings

1 C. Chen*, J. Zhang, J. Ding, and **Y. Zhou**, "Assisted unsupervised domain adaptation," in *IEEE International Symposium on Information Theory (ISIT)*, 2023, pp. 2482–2487. 🔗 URL: https://ieeexplore.ieee.org/document/10206737.

2 Z. Chen*, **Y. Zhou**, Y. Liang, and Z. Lu, "Generalized-smooth nonconvex optimization is as efficient as smooth nonconvex optimization," in *International Conference on Machine Learning (ICML)*, 2023. 🔗 URL: https://proceedings.mlr.press/v202/chen23ar.

3 J. Cho, M. Liu, Yi Zhou, and R.-R. Chen, "Multi-agent recurrent deterministic policy gradient with inter-agent communication (mardpg-iac)," in *Asilomar Conference on Signals, Systems, and Computers*, 2023.

4 Z. Guan, **Y. Zhou**, and Y. Liang, "Online nonconvex optimization with limited instantaneous oracle feedback," in *Conference on Learning Theory (COLT)*, vol. 195, 2023, pp. 3328–3355. 🔗 URL: https://proceedings.mlr.press/v195/guan23a.html.

5 Z. Li, Q. Li, **Y. Zhou**, W. Zhong, G. Zhang, and C. Wu, "Edge-cloud collaborative learning with federated and centralized features," in *International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2023, pp. 1949–1953. 🔗 URL: https://doi.org/10.1145/3539618.3591976.

6 C. Morchdi*, **Y. Zhou**, J. Ding, and B. Wang, "Exploring gradient oscillation in deep neural network training," in *Allerton Conference on Communication, Control, and Computing*, 2023.

7 Y. Zhou, **Y. Zhou**, J. Ding, and B. Wang, "Visualizing and analyzing the topology of neuron activations in deep adversarial training," in *ICML Workshop on Topology, Algebra, and Geometry in Machine Learning*, 2023. 🔗 URL: https://openreview.net/forum?id=Q692Q3dPMe.

8 Z. Chen*, S. Ma*, and **Y. Zhou**, "Accelerated proximal alternating gradient-descent-ascent for nonconvex minimax machine learning," in *IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 672–677. 🔗 URL: https://ieeexplore.ieee.org/document/9834691.

9 Z. Chen*, S. Ma*, and **Y. Zhou**, "Finding correlated equilibrium of constrained markov game: A primal-dual approach," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. 🔗 URL: https://openreview.net/forum?id=2-CflpDkezH.

10 Z. Chen*, S. Ma*, and **Y. Zhou**, "Sample efficient stochastic policy extragradient algorithm for zero-sum markov game," in *International Conference on Learning Representations (ICLR)*, 2022. 🔗 URL: https://openreview.net/forum?id=IvepFxYRDG.

11 Z. Chen*, **Y. Zhou**, R.-R. Chen, and S. Zou, "Sample and communication-efficient decentralized actor-critic algorithms with finite-time analysis," in *International Conference on Machine Learning (ICML)*, vol. 162, 2022, pp. 3794–3834. 🔗 URL: https://proceedings.mlr.press/v162/chen22ah.html.

12 S. Ma*, Z. Chen*, **Y. Zhou**, K. Ji, and Y. Liang, "Data sampling affects the complexity of online sgd over dependent data," in *Conference on Uncertainty in Artificial Intelligence (UAI)*, vol. 180, 2022, pp. 1296–1305. 🔗 URL: https://proceedings.mlr.press/v180/ma22a.html.

13 Y. Wang, Y. Wang, **Y. Zhou**, A. Velasquez, and S. Zou, "Data-driven robust multi-agent reinforcement learning," in *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2022, pp. 1–6. 🔗 URL: https://ieeexplore.ieee.org/document/9943500.

14 C. Chen*, B. Kailkhura, R. Goldhahn, and **Y. Zhou**, "Certifiably-robust federated adversarial learning via randomized smoothing," in *IEEE International Conference on Mobile Ad Hoc and Smart Systems (MASS)*, 2021, pp. 173–179. 🔗 DOI: 10.1109/MASS52906.2021.00032.

15 Z. Chen*, **Y. Zhou**, T. Xu, and Y. Liang, "Proximal gradient descent-ascent: Variable convergence under kŁ geometry," in *International Conference on Learning Representations (ICLR)*, 2021. 🔗 URL: https://openreview.net/forum?id=LVotkZmYyDi.

[16] J. Cho, M. Liu, **Y. Zhou**, and R.-R. Chen, "Communication-free two-stage multi-agent ddpg under partial states and observations," in *Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 459–463. 🔗 URL: https://ieeexplore.ieee.org/document/9723197.

[17] S. Ma*, Z. Chen*, **Y. Zhou**, and S. Zou, "Greedy-gq with variance reduction: Finite-time analysis and improved complexity," in *International Conference on Learning Representations*, 2021. 🔗 URL: https://openreview.net/forum?id=6t_dLShIUyZ.

[18] Y. Wang, S. Zou, and **Y. Zhou**, "Non-asymptotic analysis for two time-scale TDC with general smooth function approximation," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2021. 🔗 URL: https://openreview.net/forum?id=SBNs7EULzqq.

[19] C. Cannella, J. Ding, M. Soltani, **Y. Zhou**, and V. Tarokh, "Perception-distortion trade-off with restricted boltzmann machines," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 4022–4026. 🔗 DOI: 10.1109/ICASSP40776.2020.9052991.

[20] C. Chen*, Z. Chen*, **Y. Zhou**, and B. Kailkhura, "Fedcluster: Boosting the convergence of federated learning via cluster-cycling," in *IEEE International Conference on Big Data (Big Data)*, 2020, pp. 5017–5026. 🔗 DOI: 10.1109/BigData50022.2020.9377960.

[21] C. Chen*, J. Yang, and **Y. Zhou**, "Neural network training techniques regularize optimization trajectory: An empirical study," in *IEEE International Conference on Big Data (Big Data)*, 2020, pp. 141–146. 🔗 DOI: 10.1109/BigData50022.2020.9378359.

[22] K. Ji, Z. Wang, B. Weng, **Y. Zhou**, W. Zhang, and Y. Liang, "History-gradient aided batch size adaptation for variance reduced algorithms," in *International Conference on Machine Learning*, vol. 119, 2020, pp. 4762–4772. 🔗 URL: https://proceedings.mlr.press/v119/ji20a.html.

[23] B. Kailkhura, J. Thiagarajan, Q. Li, J. Zhang, **Y. Zhou**, and T. Bremer, "A statistical mechanics framework for task-agnostic sample design in machine learning," in *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 11 925–11 935. 🔗 URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/8a7129b8f3edd95b7d969dfc2c8e9d9d-Paper.pdf.

[24] C. P. Le, **Y. Zhou**, J. Ding, and V. Tarokh, "Supervised encoding for discrete representation learning," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 3447–3451. 🔗 DOI: 10.1109/ICASSP40776.2020.9054118.

[25] S. Ma* and **Y. Zhou**, "Understanding the impact of model incoherence on convergence of incremental SGD with random reshuffle," in *International Conference on Machine Learning*, vol. 119, 2020, pp. 6565–6574. 🔗 URL: https://proceedings.mlr.press/v119/ma20e.html.

[26] S. Ma*, **Y. Zhou**, and S. Zou, "Variance-reduced off-policy tdc learning: Non-asymptotic convergence analysis," in *Neural Information Processing Systems*, 2020. 🔗 URL: https://dl.acm.org/doi/abs/10.5555/3495724.3496964.

[27] T. Xu, Z. Wang, **Y. Zhou**, and Y. Liang, "Reanalysis of variance reduced temporal difference learning," in *International Conference on Learning Representations (ICLR)*, 2020. 🔗 URL: https://openreview.net/forum?id=S1ly10EKDS.

[28] **Y. Zhou**, Z. Wang, K. Ji, Y. Liang, and V. Tarokh, "Proximal gradient algorithm with momentum and flexible parameter restart for nonconvex optimization," in *International Joint Conference on Artificial Intelligence (IJCAI)*, Jul. 2020, pp. 1445–1451. 🔗 URL: https://doi.org/10.24963/ijcai.2020/201.

[29] W. Dai, **Y. Zhou**, N. Dong, H. Zhang, and E. Xing, "Toward understanding the impact of staleness in distributed machine learning," in *International Conference on Learning Representations*, 2019. 🔗 URL: https://openreview.net/forum?id=BylQV305YQ.

30 Y. Feng, **Y. Zhou**, and V. Tarokh, "Recurrent neural network-assisted adaptive sampling for approximate computing," in *IEEE International Conference on Big Data (Big Data)*, 2019, pp. 2240–2246. 🔗 DOI: 10.1109/BigData47090.2019.9006504.

31 K. Ji, Z. Wang, **Y. Zhou**, and Y. Liang, "Improved zeroth-order variance reduced algorithms and analysis for nonconvex optimization," in *International Conference on Machine Learning*, vol. 97, 2019, pp. 3100–3109. 🔗 URL: https://proceedings.mlr.press/v97/ji19a.html.

32 J. Regatti, G. Tendolkar, **Y. Zhou**, A. Gupta, and Y. Liang, "Distributed sgd generalizes well under asynchrony," in *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2019, pp. 863–870. 🔗 DOI: 10.1109/ALLERTON.2019.8919791.

33 Z. Wang, K. Ji, **Y. Zhou**, Y. Liang, and V. Tarokh, "Spiderboost and momentum: Faster variance reduction algorithms," in *Advances in Neural Information Processing Systems*, vol. 32, 2019. 🔗 URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/512c5cad6c37edb98ae91c8a76c3a291-Paper.pdf.

34 Z. Wang, **Y. Zhou**, Y. Liang, and G. Lan, "Cubic regularization with momentum for nonconvex optimization," in *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2019. 🔗 URL: http://auai.org/uai2019/proceedings/papers/118.pdf.

35 Z. Wang, **Y. Zhou**, Y. Liang, and G. Lan, "Stochastic variance-reduced cubic regularization for nonconvex optimization," in *International Conference on Artificial Intelligence and Statistics*, vol. 89, 2019, pp. 2731–2740. 🔗 URL: https://proceedings.mlr.press/v89/wang19d.html.

36 **Y. Zhou**, Y. Feng, V. Tarokh, V. Gintautas, J. McClelland, and D. Garagic, "Multi-level mean-shift clustering for single-channel radio frequency signal separation," in *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2019, pp. 1–6. 🔗 DOI: 10.1109/MLSP.2019.8918879.

37 **Y. Zhou**, J. Yang, H. Zhang, Y. Liang, and V. Tarokh, "SGD converges to global minimum in deep learning via star-convex path," in *International Conference on Learning Representations*, 2019. 🔗 URL: https://openreview.net/forum?id=BylIciRcYQ.

38 **Y. Zhou** and Y. Liang, "Critical points of linear neural networks: Analytical forms and landscape properties," in *International Conference on Learning Representations*, 2018. 🔗 URL: https://openreview.net/forum?id=SysEexbRb.

39 **Y. Zhou**, Z. Wang, and Y. Liang, "Convergence of cubic regularization for nonconvex optimization under kl property," in *Advances in Neural Information Processing Systems*, vol. 31, 2018. 🔗 URL: https://proceedings.neurips.cc/paper_files/paper/2018/file/b4568df26077653eeadf29596708c94b-Paper.pdf.

40 Q. Li, **Y. Zhou**, Y. Liang, and P. K. Varshney, "Convergence analysis of proximal gradient with momentum for nonconvex optimization," in *International Conference on Machine Learning*, vol. 70, 2017, pp. 2111–2119. 🔗 URL: https://proceedings.mlr.press/v70/li17g.html.

41 P. Xie, Y. Deng, **Y. Zhou**, *et al.*, "Learning latent space models with angular constraints," in *International Conference on Machine Learning*, vol. 70, 2017, pp. 3799–3810. 🔗 URL: https://proceedings.mlr.press/v70/xie17a.html.

42 **Y. Zhou** and Y. Liang, "Characterization of gradient dominance and regularity conditions for neural networks," in *Optimization Workshop in Advances in Neural Information Processing Systems*, 2017. 🔗 URL: http://opt-ml.org/oldopt/papers/OPT2017_paper_22.pdf.

43 **Y. Zhou** and Y. Liang, "Demixing sparse signals via convex optimization," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 4202–4206. 🔗 DOI: 10.1109/ICASSP.2017.7952948.

44 P. Xie, J. K. Kim, **Y. Zhou**, *et al.*, "Lighter-communication distributed machine learning via sufficient factor broadcasting," in *Conference on Uncertainty in Artificial Intelligence*, 2016, pp. 795–804.

45 **Y. Zhou**, Y. Yu, W. Dai, Y. Liang, and E. Xing, "On convergence of model parallel proximal gradient algorithm for stale synchronous parallel system," in *International Conference on Artificial Intelligence and Statistics*, vol. 51, 2016, pp. 713–722. URL: https://proceedings.mlr.press/v51/zhou16.html.

46 **Y. Zhou**, H. Zhang, and Y. Liang, "Geometrical properties and accelerated gradient solvers of non-convex phase retrieval," in *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2016, pp. 331–335. DOI: 10.1109/ALLERTON.2016.7852249.

47 **Y. Zhou**, H. Zhang, and Y. Liang, "On compressive orthonormal sensing," in *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2016, pp. 299–305. DOI: 10.1109/ALLERTON.2016.7852244.

48 H. Zhang, **Y. Zhou**, and Y. Liang, "Analysis of robust pca via local incoherence," in *Advances in Neural Information Processing Systems*, vol. 28, 2015. URL: https://proceedings.neurips.cc/paper_files/paper/2015/file/43baa6762fa81bb43b39c62553b2970d-Paper.pdf.

49 **Y. Zhou**, C. Zhang, G. Sun, K. Wang, and Y. Zhang, "Asymmetric-access aware optimization for stt-ram caches with process variations," in *International Conference on Great Lakes Symposium on VLSI*, 2013, pp. 143–148. DOI: 10.1145/2483028.2483079.

Honors and Awards

2023	U.S. National Science Foundation (NSF) Career Award
2021	Top instructor in the College of Engineering, University of Utah
2018	Spotlight paper award, Neural Information Processing Systems (NeurIPS)

Teaching Experience

Fall 2020-2023	Random Process
Spring 2020-2023	Fundamentals of Signals and Systems
Fall 2021, 2023	Optimization Algorithms for Machine Learning and Signal Processing
Fall 2016	Optimization for Machine Learning (guest lecturer)
Fall 2014	Advanced Topics in Learning (guest lecturer)
Fall 2013	Statistical Learning Theory (guest lecturer)

Student Supervision

Current students	Cheng Chen, Ph.D. student, 2019 - present
	Shaocong Ma, Ph.D. student, 2019 - present
	Chedi Morchdi, Ph.D. student, 2022 - present
	Yufeng Yang, Ph.D. student, 2023 - present
Former students	Ziyi Chen, Ph.D., 2019 - 2023
	Thesis title: Convergence Analysis of Minimax Optimization and Multi-agent Reinforcement Learning
	Yile Li, M.Sc, 2022 - 2023

Internal Service

Dept. of ECE	2020-2023, undergraduate committee member

Internal Service (continued)

- 2020, ECE faculty search committee member

External Service

Workshop Chair
- 2021, workshop on Scalable Reinforcement Learning with Big Data: Theory and Applications at IEEE Bigdata Conference

Conference Area Chair
- 2023, Artificial Intelligence and Statistics (AISTATS)

Program Committee
- 2021, workshop on Federated Learning for User Privacy and Data Confidentiality at International Conference on Machine Learning (ICML)
- 2020, workshop on Systems for High Performance ML and Numerical Computing at IEEE Bigdata Conference
- 2019, workshop on Learning with Rich Experience: Integration of Learning Paradigms at Advances in Neural Information Processing Systems (NeurIPS)
- 2018, workshop on Theoretical Foundations and Applications of Deep Generative Models at International Conference on Machine Learning (ICML)

Reviewer
- Neural Information Processing Systems (NeurIPS)
- International Conference on Machine Learning (ICML)
- International Conference on Learning Representations (ICLR)
- AAAI Conference on Artificial Intelligence (AAAI)
- IEEE International Symposium on Information Theory (ISIT)
- International Joint Conference on Artificial Intelligence (IJCAI)
- IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
- International Conference on Algorithmic Learning Theory (ALT)
- IEEE Transaction on Information Theory
- IEEE Transaction on Signal Processing
- Automatica
- Mathematics of Operations Research
- Proceedings of IEEE
- IEEE Internet of Things
- Computational Optimization and Applications
- IEEE Transactions on Neural Networks and Learning Systems
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- IEEE Transactions on Control of Network Systems