

Yi Zhou

Position Assistant Professor
Affiliation Department of Electrical and Computer Engineering
University of Utah, Salt Lake City, UT, USA
Phone +1 (315) 751 6542
Email yi.zhou@utah.edu
Google Scholar <https://scholar.google.com/citations?user=4fK8bYIAAAAJ&hl=en>
Homepage <https://sites.google.com/site/yizhouhomepage/home>

Education

Aug. 2013 - Ph.D. in Electrical and Computer Engineering
Aug. 2018 The Ohio State University, Columbus, OH, USA
Thesis: Nonconvex Optimization in Machine Learning: Convergence, Landscape, and Generalization
Advisor: Prof. Yingbin Liang

Sep. 2009 - B.S. in Optoelectronics
May. 2013 Beijing Institute of Technology, Beijing, China

Research Interests

Machine learning: Reinforcement learning, statistical & deep learning, multi-agent intelligence system
Optimization: Nonconvex optimization, stochastic optimization, distributed optimization
High dimensional statistics: Statistical signal processing, detection and estimation

Academic Appointments

Aug. 2019 - **Assistant Professor, University of Utah**
present Department of ECE

Sep. 2018 - **Postdoctoral Research Associate, Duke University**
Jul. 2019 Advisor: Vahid Tarokh

Aug. 2013 - **Research Assistant, The Ohio State University**
Dec. 2017 Advisor: Yingbin Liang

Dec. 2014 - **Visiting Scholar, Carnegie Mellon University**
Sep. 2016 Advisor: Eric P. Xing

Jan. 2011 - **Research Intern, IBM, Beijing, China**
Sep. 2012 Advisors: Guangyu Sun, Kun Wang, Yu Zhang

Student Supervision

Cheng Chen Ph.D. student, fall 2019 - present
Ziyi Chen Ph.D. student, fall 2019 - present
Shaocong Ma Ph.D. student, fall 2019 - present
Chedi Morchdi Ph.D. student, Fall 2022 - present
Yile Li Ph.D. student, Fall 2022 - present

Wallace Wang Undergrad student, summer 2021
Aaron Kramer Undergrad student, Fall 2021
Aaron Pettit Master student, Fall 2021

Research Grants

- 2021 **NSF grant:** CCF-2106216, small
Award amount: \$ 411,248
Project title: Self-Adaptive Optimization Algorithms with Fast Convergence
via Geometry-Adapted Hyper-Parameter Scheduling
Project period: 07/01/2021 – 06/30/2024
PI Role: Leading PI (single)
- 2021 **NSF grant:** DMS-2134223, medium
Award amount: \$ 850,000
Project title: SCALE MoDL: Advancing Theoretical Minimax Deep Learning:
Optimization, Resilience, and Interpretability
Project period: 09/01/2021 – 08/31/2024
PI Role: Leading PI

Publications

(“*” represents equal contribution, “[Blue](#)” denotes student (co)authors)

Submitted Conference Papers

- [C1] [S. Ma](#), [Z. Chen](#), K. Ji, **Y. Zhou**, Y. Liang. “Data Sampling Affects the Complexity of Online SGD over Dependent Data”, submitted to *Uncertainty in Artificial Intelligence (UAI)*, 2022.
- [C2] [Z. Chen](#), **Y. Zhou**. “A Fast and Convergent Proximal Algorithm for Regularized Nonconvex and Non-smooth Bi-level Optimization.” submitted to *International Conference on Machine Learning (ICML)*, 2022.
- [C3] [Z. Chen](#), Q. Li, **Y. Zhou**. “Finding Local Minimax Points via (Stochastic) Cubic-Regularized GDA: Global Convergence and Complexity.” submitted to *International Conference on Machine Learning (ICML)*, 2022.
- [C4] [Z. Chen](#), S. Zou, R. Chen, **Y. Zhou**. “Sample and Communication-Efficient Decentralized Actor-Critic Algorithms with Finite-Time Analysis.” submitted to *International Conference on Machine Learning (ICML)*, 2022.

Published Conference Papers

- [C1] [Z. Chen](#), [S. Ma](#), **Y. Zhou**. “Sample Efficient Stochastic Policy Extragradient Algorithm for Zero-Sum Markov Game”, Proc. *International Conference on Learning Representations (ICLR)*, 2022.
- [C2] J. Cho, M. Liu, **Y. Zhou**, R. Chen. “Communication-Free Two-Stage Multi-AgentDDPG under Partial States and Observations”, Proc. *Asilomar*, 2021.
- [C3] Y. Wang, S. Zou, **Y. Zhou**. “Non-Asymptotic Analysis for Two Time-scale TDC with General Smooth Function Approximation”, Proc. *Neural Information Processing Systems (NeurIPS)*, 2021.
- [C4] [C. Chen](#), B. Kailkhura, R. Goldhahn, **Y. Zhou**. “Certifiably-Robust Federated Adversarial Learning via Randomized Smoothing”, Proc. *IEEE MASS*, 2021.
- [C5] [Z. Chen](#), **Y. Zhou**, T. Xu, Y. Liang. “Proximal Gradient Descent-Ascent: Variable Convergence under KL Geometry”, Proc. *International Conference on Learning Representations (ICLR)*, 2021.
- [C6] [S. Ma](#), [Z. Chen](#), **Y. Zhou**, S. Zou. “Greedy-GQ with Variance Reduction: Finite-time Analysis and Improved Complexity”, Proc. *International Conference on Learning Representations (ICLR)*, 2021.

- [C7] [C. Chen](#), [Z. Chen](#), [Y. Zhou](#), B. Kailkhura. “FedCluster: Boosting the Convergence of Federated Learning via Cluster-Cycling”, *IEEE Bigdata Conference* 2020.
- [C8] [S. Ma](#), [Y. Zhou](#), S. Zou. “Variance-Reduced Off-Policy TDC Learning: Non-Asymptotic Convergence Analysis”, *Neural Information Processing Systems (NeurIPS)*, 2020.
- [C9] B. Kailkhura, J. Thiagarajan, Q. Li, J. Zhang, [Y. Zhou](#), P. Bremer. “A Statistical Mechanics Framework for Task-Agnostic Sample Design in Machine Learning”, *Neural Information Processing Systems (NeurIPS)*, 2020.
- [C10] [C. Chen](#), [Y. Zhou](#). “Neural Network Training Techniques Regularize Optimization Trajectory: An Empirical Study”, *IEEE Bigdata Conference* 2020.
- [C11] [Y. Zhou](#), K. Ji, W. Zhe, Y. Liang, V. Tarokh. “Proximal Gradient Algorithm with Momentum and Flexible Parameter Restart for Nonconvex Optimization”, Proc. *International Joint Conference on Artificial Intelligence (IJCAI)* 2020.
- [C12] K. Ji, Z. Wang, [Y. Zhou](#), Y. Liang. “History-Gradient Aided Batch Size Adaptation for Variance Reduced Algorithms”, *International Conference on Machine Learning (ICML)* 2020.
- [C13] [S. Ma](#), [Y. Zhou](#). “Understanding the Impact of Model Incoherence on Convergence of Incremental SGD with Random Reshuffle”. *International Conference on Machine Learning (ICML)* 2020.
- [C14] C. P. Le, [Y. Zhou](#), J. Ding, V. Tarokh. “SupervisedD Encoding for Discrete Representation Learning.” *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2020.
- [C15] C. Cannella, J. Ding, M. Soltani, [Y. Zhou](#), V. Tarokh. “Perception-Distortion Trade-off with Restricted Boltzmann Machines,” *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2020.
- [C16] T. Xu, Z. Wang, [Y. Zhou](#), Y. Liang. “Reanalysis of Variance Reduced Temporal Difference Learning,” *Proc. International Conference on Learning Representations (ICLR)*, 2020.
- [C17] Y. Feng, [Y. Zhou](#), V. Tarokh. “Recurrent Neural Network-Assisted Adaptive Sampling for Approximate Computing,” *IEEE Bigdata conference* 2019.
- [C18] [Y. Zhou](#), Y. Feng, V. Tarokh, V. Gintautas, J. McClelland, D. Garagic, “Multi-level Mean-shift Clustering for Single-channel Radio Frequency Signal Separation,” *Machine Learning for Signal Processing (MLSP)*, 2019.
- [C19] J. Regatti, G. Tendolkar, [Y. Zhou](#), A. Gupta, Y. Liang. “Distributed SGD Generalizes Well Under Asynchrony,” *Annual Allerton Conference*, 2019.
- [C20] K. Ji, Z. Wang, [Y. Zhou](#), Y. Liang, V. Tarokh. “Improved Zeroth-Order Variance Reduced Algorithms and Analysis for Nonconvex Optimization,” in *Proc. International Conference on Machine Learning (ICML)*, 2019.
- [C21] Z. Wang, K. Ji, [Y. Zhou](#), Y. Liang, V. Tarokh. “SpiderBoost: A Class of Faster Variance-reduced Algorithms for Nonconvex Optimization,” *Neural Information Processing Systems (NeurIPS)*, 2019.
- [C22] [Y. Zhou](#), W. Zhe and Y. Liang. “Convergence of Cubic Regularization for Nonconvex Optimization under KL Property”. *Neural Information Processing Systems (NIPS)*, 2018 (Spotlight paper).
- [C23] [Y. Zhou](#), J. Yang, H. Zhang, Y. Liang. “SGD Converges to Global Minimum in Deep Learning via Star-convex Path,” *International Conference on Learning Representations (ICLR)*, 2019.
- [C24] W. Dai, [Y. Zhou](#), N. Dong, H. Zhang, E. P Xing. “Toward Understanding the Impact of Staleness in Distributed Machine Learning”, *International Conference on Learning Representations (ICLR)*, 2019.
- [C25] Z. Wang, [Y. Zhou](#), Y. Liang, G. Lan. “Cubic Regularization with Momentum for Nonconvex Optimization,” *Uncertainty in AI (UAI)*, 2019.
- [C26] Z. Wang*, [Y. Zhou*](#), Y. Liang. “Sample Complexity of Stochastic Variance-Reduced Cubic Regularization for Nonconvex Optimization,” *Artificial Intelligence and Statistics (AISTATS)*, 2019.

- [C27] **Y. Zhou**, Y. Liang. “Critical Points of Neural Networks: Analytical Forms and Landscape Properties,” *Proc. International Conference on Learning Representations (ICLR)*, Vancouver, Canada, Apr. 2018.
- [C28] **Y. Zhou**, Y. Liang. “Characterization of Gradient Dominance and Regularity Conditions for Neural Networks,” in *Proc. Optimization Workshop in Proc. Advances in Neural Information Processing Systems (NIPS)*, California, USA, Dec. 2017.
- [C29] P. Xie, Y. Deng, **Y. Zhou**, A. Kumar, Y. Yu, J. Zou, E. P. Xing. “Learning Latent Space Models with Angular Constraints,” in *Proc. International Conference on Machine Learning (ICML)*, Sydney, Australia, Aug. 2017.
- [C30] Q. Li*, **Y. Zhou***, Y. Liang. “Convergence Analysis of Proximal Gradient with Momentum for Nonconvex Optimization,” in *Proc. International Conference on Machine Learning (ICML)*, Sydney, Australia, Aug. 2017.
- [C31] P. Xie, J. Kim, **Y. Zhou**, Q. Ho, A. Kumar, Y. Yu, E. Xing. “Lighter-Communication Distributed Machine Learning via Sufficient Factor Broadcasting,” in *Proc. Conference on Uncertainty in Artificial Intelligence (UAI)*, New York, USA, Jun. 2016.
- [C32] **Y. Zhou**, Y. Yu, W. Dai, Y. Liang, E. Xing. “On Convergence of Model Parallel Proximal Gradient Algorithm for Stale Synchronous Parallel System,” in *Proc. International Conference on Artificial Intelligence and Statistics (AISTATS)*, Cadiz, Spain, May. 2016.
- [C33] H. Zhang*, **Y. Zhou***, Y. Liang. “Analysis of Robust PCA via Local Incoherence,” in *Proc. Advances in Neural Information Processing Systems (NIPS)*, Montreal, Canada, Dec. 2015.
- [C34] **Y. Zhou**, Y. Liang. “Demixing Sparse Signals via Convex Optimization,” in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, USA, Mar. 2017.
- [C35] **Y. Zhou**, H. Zhang, Y. Liang. “Accelerated Gradient Descent for Non-convex Phase Retrieval,” in *Proc. 55th Allerton Conference on Communication, Control, and Computing*, Monticello, USA, Oct. 2016.
- [C36] **Y. Zhou**, H. Zhang, Y. Liang. “On Compressive Orthonormal Sensing,” in *Proc. 55th Allerton Conference on Communication, Control, and Computing*, Monticello, USA, Oct. 2016.
- [C37] **Y. Zhou**, C. Zhang, G. Sun, K. Wang, Y. Zhang. “Asymmetric-access Aware Optimization for STT-RAM Caches with Process Variations,” in *Proc. Great Lakes Symposium on VLSI (GLSVLSI)*, Paris, France, May. 2013.

Submitted Journal Papers

- [J1] [Z. Chen](#), **Y. Zhou**. “Momentum with Variance Reduction for Nonconvex Composition Optimization.” Submitted to TNNLS.
- [J2] Y. Zhang, **Y. Zhou**, K. Ji, M. Zavlanos. “Boosting One-Point Derivative-Free Online Optimization via Residual Feedback.” Submitted to TAC.

Published Journal Papers

- [J1] Y. Zhang, **Y. Zhou**, K. Ji, M. Zavlanos. “A New One-Point Residual-Feedback Oracle For Black-Box Learning and Control”, *Automatica*, 2021.
- [J2] Q. Li, B. Kailkhura, R. Anirudh, **Y. Zhou**, Y. Liang, P. Varshney. “MR-GAN: Manifold Regularized Generative Adversarial Networks for Scientific Data”, *SIAM Journal on Mathematics of Data Science*, 2021.
- [J3] **Y. Zhou**, H. Zhang, Y. Liang. “Understanding generalization error of SGD in nonconvex optimization”, *Machine Learning Journal*, 2021.
- [J4] K. Ji*, **Y. Zhou***, Y. Liang. “Understanding Estimation and Generalization Error of Generative Adversarial Networks”, *IEEE Transaction on Information Theory*, vol. 67, no. 5, pp. 3114-3129, 2021.
- [J5] T. Xu, **Y. Zhou**, K. Ji, Y. Liang. “When Will Gradient Methods Converge to Max-margin Classifier under ReLU Models?”, in *Stats*, Stat. 2021;10:e354. <https://doi.org/10.1002/sta4.354>, 2021.

- [J6] Z. Wang, **Y. Zhou**, Y. Liang, G. Lan. “A Note on Inexact Condition for Cubic Regularized Newton’s Method”, in *Operations Research Letters*, vol 47, no. 2, 2019.
- [J7] **Y. Zhou**, Y. Yu, W. Dai, Y. Liang, E. Xing. “Distributed Proximal Gradient Algorithm for Partially Asynchronous Computer Clusters,” *Journal of Machine Learning Research*, vol 19, no. 1, 2018.
- [J8] 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. 141, 2017.
- [J9] **Y. Zhou**, L. Shen, Y. Liang. “A Simple Convergence Analysis of Bregman Proximal Gradient Algorithm,” in *Computational Optimization and Applications*, vol 73, no. 3, 2017.

Presentations

- “Optimization Meets Reinforcement Learning”, tutorial talk at *IEEE BigData Conference*, 2021.
- “Recent Advances in Reinforcement Learning Theory”, tutorial talk at *IEEE International Symposium on Information Theory (ISIT)*, 2021.
- “Proximal Gradient Descent-Ascent: Variable Convergence under KL Geometry”, presented at *International Conference on Learning Representations (ICLR)*, Apr 27th, 2021, virtual conference.
- “Greedy-GQ with Variance Reduction: Finite-time Analysis and Improved Complexity”, presented at *International Conference on Learning Representations (ICLR)*, Apr 26th, 2021, virtual conference.
- “Proximal Gradient Algorithm with Momentum and Flexible Parameter Restart for Nonconvex Optimization,” oral presentation, International Joint Conference on Artificial Intelligence, 2021.
- “FedCluster: Boosting the Convergence of Federated Learning via Cluster-Cycling”, presented at *IEEE Conference on BigData*, Dec 11th, 2020, virtual conference.
- “Neural Network Training Techniques Regularize Optimization Trajectory: An Empirical Study”, presented at *IEEE Conference on BigData*, Dec 12th, 2020, virtual conference.
- “Variance-Reduced Off-Policy TDC Learning: Non-Asymptotic Convergence Analysis”, presented at *Neural Information Processing Systems (NeurIPS)*, Dec 8th, 2020, virtual conference.
- “A Statistical Mechanics Framework for Task-Agnostic Sample Design in Machine Learning”, presented at *Neural Information Processing Systems (NeurIPS)*, Dec 8th, 2020, virtual conference.
- “Momentum with Variance Reduction for Nonconvex Compositional Optimization with Near-optimal Sample Complexity,” presented at *ITA Workshop*, February 4th, 2020, San Diego, USA.
- “Proximal Gradient Algorithm with Momentum and Flexible Parameter Restart for Nonconvex Optimization”, presented at *International Joint Conference on Artificial Intelligence (IJCAI)*, Jan 10th, 2020, virtual conference.
- “History-Gradient Aided Batch Size Adaptation for Variance Reduced Algorithms”, presented at *International Conference on Machine Learning (ICML)*, July 10th, 2020, virtual conference.
- “Understanding the Impact of Model Incoherence on Convergence of Incremental SGD with Random Reshuffle”, presented at *International Conference on Machine Learning (ICML)*, July 11th, 2020, virtual conference.
- “Reanalysis of Variance Reduced Temporal Difference Learning”, presented at *International Conference on Learning Representations (ICLR)*, Apr 28th, 2020, virtual conference.
- “Multi-level Mean-shift Clustering for Single-channel Radio Frequency Signal Separation”, presented at *Machine Learning for Signal Processing (MLSP)*, October 15th, 2019, Pittsburgh, USA.
- “SGD Converges to Global Minimum in Deep Learning via Star-convex Path”, presented at *International Conference on Learning Representations (ICLR)*, May 7th, 2019, New Orleans, USA.

- “Toward Understanding the Impact of Staleness in Distributed Machine Learning”, presented at *International Conference on Learning Representations (ICLR)*, May 7th, 2019, New Orleans, USA.
- “Convergence of Cubic Regularization for Nonconvex Optimization under KL Property”, presented at *Neural Information Processing Systems (NIPS)*, Dec 10th, 2018, Montreal, Canada.
- “Critical Points of Neural Networks: Analytical Forms and Landscape Properties,” presented at *International Conference on Learning Representations (ICLR)*, 2018, Vancouver, Canada.
- “Demixing Sparse Signals via Convex Optimization”, presented at *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar 8th, 2017, New Orleans, USA.
- “On Convergence of Model Parallel Proximal Gradient Algorithm for Stale Synchronous Parallel System”, presented at *International Conference on Artificial Intelligence and Statistics (AISTATS)*, May 9th, 2016, Cadiz, Spain.
- “Analysis of Robust PCA via Local Incoherence”, presented at *Advances in Neural Information Processing Systems (NIPS)*, Dec 9th, 2015, Montreal, Canada.

Honors and Awards

2018 Spotlight paper award, NeurIPS

Teaching Experiences

Spring 2022 *Fundamentals of Signals and Systems*
 Fall 2021 *Optimization Algorithms for Machine Learning and Signal Processing*
 Spring 2021 *Fundamentals of Signals and Systems*
 Fall 2020 *Random Process*
 Spring 2020 *Fundamentals of Signals and Systems*
 Fall 2016 Guest lecturer for course *Optimization for Machine Learning*
 Topic: Accelerated gradient descent methods
 Fall 2014 Guest lecturer for course *Advanced Topics in Learning*
 Topic: Theoretical guarantees for matrix completion
 Fall 2013 Guest lecturer for course *Statistical Learning Theory*

Internal Service

Undergrad 2020-2022 ECE undergrad committee member
 committee
 Faculty candidate 2020 ECE faculty search committee member.
 search committee

External Service

Workshop Chair IEEE Bigdata 2021 workshop on scalable reinforcement learning with big data: theory and applications

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 Transaction on Information Theory
 IEEE Transaction on Signal Processing

Automatica
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.

Program ICML 2018 workshop on theoretical foundations and applications of deep generative models
Committee NeurIPS 2019 workshop on Learning with Rich Experience: Integration of Learning Paradigms.
IEEE Bigdata 2020 workshop on Systems for High Performance ML and Numerical Computing.
ICML 2021 Workshop on Federated Learning for User Privacy and Data Confidentiality

Reference List

Dr. Vahid Tarokh: Professor
Department of Electrical and Computer Engineering
Duke University
Phone: (919) 660-7594
Email: vahid.tarokh@duke.edu

Dr. Yingbin Liang: Professor
Department of Electrical and Computer Engineering
The Ohio State University
Phone: (609) 658-1330
Email: liang.889@osu.edu

Dr. Guanghui Lan: Professor
Industrial & Systems Engineering
Georgia Institute of Technology
Phone: (404) 385-5402
Email: george.lan@isye.gatech.edu