

# ARYAN MOKHTARI

The University of Texas at Austin  
2501 Speedway, EER 6.826  
Austin, TX 78712

Phone: (512)-232-8112  
Email: [mokhtari@austin.utexas.edu](mailto:mokhtari@austin.utexas.edu)  
[Homepage](#) and [Google Scholar Profile](#)

---

## RESEARCH INTERESTS

**Foundations of AI and ML:** Machine Unlearning, Efficient Fine-Tuning, Distillation, Representation Learning, Multi-task Learning, In-Context Learning,

**Optimization Theory and Algorithms:** Convex and Nonconvex Optimization, Minimax Optimization, Bilevel Optimization, Parameter-free and Adaptive Optimization

## APPOINTMENTS

**Department of Electrical and Computer Engineering, UT Austin** Austin, TX  
Assistant Professor August 2019 - Present  
Jack Kilby/Texas Instruments Endowed Faculty Fellow in Computer Engineering

**ML Theory Group, Google Research** Austin, TX  
Visiting Faculty Researcher January 2025 - Present

**Laboratory for Information and Decision Systems (LIDS), MIT** Cambridge, MA  
Postdoctoral Associate January 2018 - July 2019  
Hosts: Prof. Asu Ozdaglar and Prof. Ali Jadbabaie

**Simons Institute for the Theory of Computing, UC Berkeley** Berkeley, CA  
Research Fellow August 2017 - December 2017  
Program: "Bridging Continuous and Discrete Optimization"

## EDUCATION

**University of Pennsylvania** Philadelphia, PA  
Ph.D. in Electrical & Systems Engineering August 2017  
Advisor: Prof. Alejandro Ribeiro

**University of Pennsylvania (The Wharton School)** Philadelphia, PA  
A.M. in Statistics August 2017

**University of Pennsylvania** Philadelphia, PA  
M.Sc. in Electrical Engineering May 2014

**Sharif University of Technology** Tehran, Iran  
B.Sc. in Electrical Engineering June 2011

## HONORS and AWARDS

- Google Research Scholar Award 2024
- NSF CAREER Award 2024
- IEEE Senior Member 2024
- UT Austin ECE Junior Faculty Excellence in Teaching Award 2023
- Member of \$20M NSF AI Institute of Foundations of Machine Learning (IFML) 2022-present
- Member of \$20M NSF AI Inst. for Future Edge Networks & Distributed Intelligence (AI-EDGE) 2021-present
- Texas Instruments/Kilby Fellow 2021-present
- ARO Early Career Program Award (Previously known as ARO YIP Award) 2021

- Joseph and Rosaline Wolf Best Doctoral Dissertation Award (Awarded by the ESE Department of the University of Pennsylvania) 2018
- Research Fellowship from the Simons Institute at UC Berkeley 2017

## PUBLICATIONS

## Book

1. G. Braun, A. Carderera, C. W. Combettes, H. Hassani, A. Karbasi, A. Mokhtari, S. Pokutta, “[Conditional Gradient Methods](#),” 2023.

## Conference Papers

1. R. Jiang, D. Maladkar, A. Mokhtari. “Convergence Analysis of Adaptive Gradient Methods under Refined Smoothness and Noise Assumptions,” *Conference on Learning Theory (COLT)*, 2025. [[pdf](#)]
2. Q. Fruytier, A. Mokhtari, S. Sanghavi. “Learning Mixtures of Experts with EM,” *Int. Conference on Machine Learning (ICML)*, 2025. [[pdf](#)]
3.  $(\alpha, \beta)$  R. Jiang, A. Mokhtari, F. Patitucci. “Improved Complexity for Smooth Nonconvex Optimization: A Two-Level Online Learning Approach with Quasi-Newton Methods,” *Symposium on Theory of Computing (STOC)*, 2025. [[pdf](#)]
4. B. Li, L. Zhang, A. Mokhtari, N. He. “On the Crucial Role of Initialization for Matrix Factorization,” *International Conference on Learning Representations (ICLR)*, 2025. [[pdf](#)]
5. L. Collins\*, A. Parulekar\*, A. Mokhtari, S. Sanghavi, S. Shakkottai. “In-Context Learning with Transformers: Softmax Attention Adapts to Function Lipschitzness,” *Neural Information Processing Systems (NeurIPS)*, 2024. [[pdf](#)] ([Spotlight](#))
6. Q. Jin, R. Jiang, A. Mokhtari. “Non-asymptotic Global Convergence Analysis of BFGS with the Armijo-Wolfe Line Search,” *Neural Information Processing Systems (NeurIPS)*, 2024. [[pdf](#)] ([Spotlight](#))
7. R. Jiang, A. Kavis, Q. Jin, S. Sanghavi, A. Mokhtari. “Adaptive and Optimal Second-order Optimistic Methods for Minimax Optimization,” *Neural Information Processing Systems (NeurIPS)*, 2024. [[pdf](#)]
8. J. Cao, R. Jiang, E. Yazdandoost Hamedani, A. Mokhtari. “An Accelerated Gradient Method for Simple Bilevel Optimization with Convex Lower-level Problem,” *Neural Information Processing Systems (NeurIPS)*, 2024. [[pdf](#)]
9. R. Jiang, M. Derezhinski, A. Mokhtari. “Stochastic Newton Proximal Extragradient Method,” *Neural Information Processing Systems (NeurIPS)*, 2024. [[pdf](#)]
10. L. Collins, H. Hassani, M. Soltanolkotabi, A. Mokhtari, S. Shakkottai. “Provable Multi-Task Representation Learning by Two-Layer ReLU Neural Networks,” *Int. Conference on Machine Learning (ICML)*, 2024. ([Oral](#)) [[pdf](#)]
11. R. Jiang, P. Raman, S. Sabach, A. Mokhtari, M. Hong, and V. Cevher, “Krylov Cubic Regularized Newton: A Subspace Second-Order Method with Dimension-Free Convergence Rate,” *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2024. [[pdf](#)]
12. R. Jiang, A. Mokhtari. “Accelerated Quasi-Newton Proximal Extragradient: Faster Rate for Smooth Convex Optimization.” *Neural Information Processing Systems (NeurIPS)*, 2023. ([Spotlight](#)) [[pdf](#)]
13. N. Rajaraman, Devvrit, A. Mokhtari, K. Ramchandran. “Greedy Pruning with Group Lasso Provably Generalizes for Matrix Sensing,” *Neural Information Processing Systems (NeurIPS)*, 2023. [[pdf](#)]
14. J. Cao, R. Jiang, N. Abolfazli, E.Y. Hamedani, A. Mokhtari. “Projection-Free Methods for Stochastic Simple Bilevel Optimization with Convex Lower-level Problem.” *Neural Information Processing Systems (NeurIPS)*, 2023. [[pdf](#)]
15. R. Jiang, Q. Jin, A. Mokhtari. “Online Learning Guided Curvature Approximation: A Quasi-Newton Method with Global Non-Asymptotic Superlinear Convergence.” *Conference on Learning Theory (COLT)*, 2023. [[pdf](#)]

16. A. Parulekar, L. Collins, K. Shanmugam, A. Mokhtari, S. Shakkottai. "InfoNCE Loss Provably Learns Cluster-Preserving Representations." *Conference on Learning Theory (COLT)*, 2023. [\[pdf\]](#)
17. R. Jiang, N. Abolfazli, A. Mokhtari, E. Yazdandoost Hamedani. "A Conditional Gradient-based Method for Simple Bilevel Optimization with Convex Lower-level Problem." *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2023. [\[pdf\]](#)
18. P. Hedge, G. de Veciana, A. Mokhtari. "Network Adaptive Federated Learning: Congestion and Lossy Compression." *IEEE International Conference on Computer Communications (INFOCOM)*, 2023. [\[pdf\]](#)
19. L. Collins, H. Hassani, A. Mokhtari, S. Shakkottai. "FedAvg with Fine Tuning: Local Updates Lead to Representation Learning." *Neural Information Processing Systems (NeurIPS)*, 2022. [\[pdf\]](#)
20. L. Collins, A. Mokhtari, S. Oh, S. Shakkottai "MAML and ANIL Provably Learn Representations." *International Conference on Machine Learning (ICML)*, 2022. [\[pdf\]](#)
21. Q. Jin, A. Koppel, K. Rajawat, A. Mokhtari. "Sharpened Quasi-Newton Methods: Faster Superlinear Rate and Larger Local Convergence Neighborhood." *Int. Conference on Machine Learning (ICML)*, 2022. [\[pdf\]](#)
22. M. Faw\*, I. Tziotis\*, C. Caramanis, A. Mokhtari, S. Shakkottai, R. Ward. "The Power of Adaptivity in SGD: Self-Tuning Step Sizes with Unbounded Gradients and Affine Variance." *Conference on Learning Theory (COLT)*, 2022. [\[pdf\]](#)
23. A. Adibi, A. Mokhtari, and H. Hassani. "Minimax Optimization: The Case of Convex-Submodular," *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2022. (Oral presentation: Top 2.6% of the submitted papers) [\[pdf\]](#)
24. M. Ye, R. Jiang, H. Wang, D. Choudhary, X. Du, B. Bhushanam, A. Mokhtari, A. Kejariwal, and Q. Liu. "Future Gradient Descent for Adapting the Temporal Shifting Data Distribution in Online Recommendation System," *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2022.
25. L. Collins, A. Mokhtari, S. Shakkottai, "How Does the Task Landscape Affect MAML Performance?," *Conference on Lifelong Learning Agents*, 2022. [\[pdf\]](#)
26. Q. Jin and A. Mokhtari. "Exploiting Local Convergence of Quasi-Newton Methods Globally: Adaptive Sample Size Approach," *Neural Information Processing Systems (NeurIPS)*, 2021. [\[pdf\]](#)
27. A. Fallah, A. Mokhtari, and A. Ozdaglar. "Generalization of Model-Agnostic Meta-Learning Algorithms: Recurring and Unseen Tasks," *Neural Information Processing Systems (NeurIPS)*, 2021. [\[pdf\]](#)
28. A. Fallah, K. Georgiev, A. Mokhtari, and A. Ozdaglar. "Provably Convergent Policy Gradient Methods for Model-Agnostic Meta-Reinforcement Learning," *Neural Information Processing Systems (NeurIPS)*, 2021. [\[pdf\]](#)
29. L. Collins, H. Hassani, A. Mokhtari, S. Shakkottai, "Exploiting Shared Representations for Personalized Federated Learning," *International Conference on Machine Learning (ICML)*, 2021. [\[pdf\]](#)
30. F. Haddadpour, M. M. Kamani, A. Mokhtari, and M. Mahdavi. "Federated Learning with Compression: Unified Analysis and Sharp Guarantees," *Int. Conf. on Artificial Intelligence and Statistics (AISTATS)*, 2021. [\[pdf\]](#)
31. L. Collins, A. Mokhtari, and S. Shakkottai. "Task-Robust Model-Agnostic Meta-Learning," *Neural Information Processing Systems (NeurIPS)*, 2020. [\[pdf\]](#)
32. I. Tziotis, C. Caramanis, and A. Mokhtari. "Second Order Optimality in Decentralized Non-Convex Optimization via Perturbed Gradient Tracking," *Neural Information Processing Systems (NeurIPS)*, 2020. [\[pdf\]](#)
33. A. Fallah, A. Mokhtari, and A. Ozdaglar. "Personalized Federated Learning with Theoretical Guarantees: A Model-Agnostic Meta-Learning Approach," *Neural Information Processing Systems (NeurIPS)*, 2020. [\[pdf\]](#)
34. A. Adibi, A. Mokhtari, and H. Hassani. "Submodular Meta-Learning," *Neural Information Processing Systems (NeurIPS)*, 2020. [\[pdf\]](#)
35. H. Taheri, A. Mokhtari, H. Hassani, R. Pedarsani. "Quantized Decentralized Stochastic Learning over Directed Graphs," *International Conference on Machine Learning (ICML)*, 2020. [\[pdf\]](#)
36. M. Zhang, Z. Shen, A. Mokhtari, H. Hassani, and A. Karbasi. "One Sample Stochastic Frank-Wolfe," *Int.*

- Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020. [\[pdf\]](#)
37. S. Soori, K. Mischenko, A. Mokhtari, M. Dehnavi, and M. Gurbuzbalaban. "DAve-QN: A Distributed Averaged Quasi-Newton Method with Local Superlinear Convergence Rate," *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020. [\[pdf\]](#)
  38. M. Zhang, L. Chen, A. Mokhtari, H. Hassani, and A. Karbasi. "Quantized Frank-Wolfe: Faster Optimization, Lower Communication, and Projection Free," *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020. [\[pdf\]](#)
  39. A. Mokhtari, A. Ozdaglar, and S. Pattathil. "A Unified Analysis of Extra-gradient and Optimistic Gradient Methods for Saddle Point Problems: Proximal Point Approach," *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020. [\[pdf\]](#)
  40. A. Fallah, A. Mokhtari, and A. Ozdaglar. "On the Convergence Theory of Gradient-Based Model-Agnostic Meta-Learning Algorithms," *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020. [\[pdf\]](#)
  41. A. Reisizadeh, A. Mokhtari, H. Hassani, A. Jadbabaie, and R. Pedarsani, "FedPAQ: A Communication-Efficient Federated Learning Method with Periodic Averaging and Quantization," *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020. [\[pdf\]](#)
  42. M. Jahani, X. He, C. Ma, A. Mokhtari, D. Mudigere, A. Ribeiro, M. Takac. "Efficient Distributed Hessian Free Algorithm for Large-scale Empirical Risk Minimization via Accumulating Sample Strategy," *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020. [\[pdf\]](#)
  43. H. Hassani, A. Karbasi, A. Mokhtari, Z. Shen. "Stochastic Continuous Greedy ++: When Upper and Lower Bounds Match," *Neural Information Processing Systems (NeurIPS)*, 2019. [\[pdf\]](#)
  44. A. Reisizadeh, H. Taheri, A. Mokhtari, H. Hassani, and R. Pedarsani, "Robust and Communication-Efficient Collaborative Learning," *Neural Information Processing Systems (NeurIPS)*, 2019. [\[pdf\]](#)
  45. A. Mokhtari, A. Ozdaglar, and A. Jadbabaie, "Efficient Nonconvex Empirical Risk Minimization via Adaptive Sample Size Methods," *Int. Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019. [\[pdf\]](#)
  46. J. Zhang, C. Uribe, A. Mokhtari, and A. Jadbabaie, "Achieving Acceleration in Distributed Optimization via Direct Discretization of the Heavy-Ball ODE," *American Control Conference (ACC)*, 2019. [\[pdf\]](#)
  47. A. Mokhtari, A. Ozdaglar, and A. Jadbabaie, "Escaping Saddle Points in Constrained Optimization," *Neural Information Processing Systems (NeurIPS)*, pp. 3533-3643, 2018. (Spotlight: Top 4% of the submitted papers) [\[pdf\]](#) [\[Supplementary Material\]](#)
  48. J. Zhang, A. Mokhtari, S. Sra, and A. Jadbabaie, "Direct Runge-Kutta Discretization Achieves Acceleration," *Neural Information Processing Systems (NeurIPS)*, pp. 3901-3910, 2018. (Spotlight: Top 4% of the submitted papers) [\[pdf\]](#) [\[Supplementary Material\]](#)
  49. A. Reisizadeh, A. Mokhtari, H. Hassani, and R. Pedarsani, "Quantized Decentralized Consensus Optimization," *IEEE 57th Conference on Decision and Control (CDC)*, 2018. [\[pdf\]](#)
  50. S. Paternain, A. Mokhtari, and A. Ribeiro, "A Newton Method for Faster Navigation in Cluttered Environments," *IEEE 57th Conference on Decision and Control (CDC)*, 2018. [\[pdf\]](#)
  51. A. Mokhtari, H. Hassani, and A. Karbasi, "Decentralized Submodular Maximization: Bridging Discrete and Continuous Settings", *International Conference on Machine Learning (ICML)*, PMLR 80:3613-3622, 2018. (Long talk) [\[pdf\]](#) [\[Supplementary Material\]](#)
  52. Z. Shen, A. Mokhtari, H. Qian, P. Zhao, and T. Zhou, "Towards More Efficient Stochastic Decentralized Learning: Faster Convergence and Sparse Communication", *International Conference on Machine Learning (ICML)*, PMLR 80:4631-4640, 2018. [\[pdf\]](#) [\[Supplementary Material\]](#)
  53. A. Mokhtari, H. Hassani, and A. Karbasi, "Conditional Gradient Method for Stochastic Submodular Maximization: Closing the Gap", *International Conference on Artificial Intelligence and Statistics (AISTATS)*, PMLR 84:1886-1895, 2018. [\[pdf\]](#) [\[Supplementary Material\]](#)
  54. M. Eisen, A. Mokhtari, and A. Ribeiro, "Large Scale Empirical Risk Minimization via Truncated Adaptive Newton Method", *International Conference on Artificial Intelligence and Statistics (AISTATS)*, PMLR 84:1447-

- 1455, 2018. [[pdf](#)] [[Supplementary Material](#)]
55. A. Koppel, A. Mokhtari, and A. Ribeiro, "Parallel Stochastic Successive Convex Approximation Method for Large-Scale Dictionary Learning," *Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, Calgary, Alberta, Canada, 2018. [[pdf](#)]
  56. A. Mokhtari and A. Ribeiro, "First-Order Adaptive Sample Size Methods to Reduce Complexity of Empirical Risk Minimization", *Neural Information Processing Systems (NeurIPS) 2017*, pp. 2057-2065, Long Beach, CA, December 4-9, 2017. [[pdf](#)] [[Supplementary Material](#)]
  57. M. Eisen, A. Mokhtari, and A. Ribeiro, "A Primal-Dual Quasi-Newton Method for Consensus Optimization", in *51th Asilomar Conference on Signals, Systems and Computers*, 2017, pp. 298-302. [[pdf](#)]
  58. A. Mokhtari, M. Eisen, and A. Ribeiro, "An Incremental Quasi-Newton Method with a Local Superlinear Convergence Rate," *Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, New Orleans, LA, 2017, pp. 4039-4043. [[pdf](#)]
  59. A. Mokhtari, M. Gürbüzbalaban, and A. Ribeiro, "A Double Incremental Aggregated Gradient Method with Linear Convergence Rate for Large-Scale Optimization," *Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, New Orleans, LA, 2017, pp. 4696-4700. [[pdf](#)]
  60. A. Mokhtari, A. Koppel, G. Scutari, and A. Ribeiro, "Large-Scale NonConvex Stochastic Optimization by Doubly Stochastic Successive Convex Approximation," *Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, New Orleans, LA, 2017, pp. 4701-4705. [[pdf](#)]
  61. A. Mokhtari, and A. Ingber, "A Diagonal-Augmented Quasi-Newton Method with Application to Factorization Machines," *Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, New Orleans, LA, 2017, pp. 2671-2675. [[pdf](#)]
  62. A. Mokhtari, H. Daneshmand, A. Lucchi, T. Hofmann, and A. Ribeiro, "Adaptive Newton Method for Empirical Risk Minimization to Statistical Accuracy", *Neural Information Processing Systems (NeurIPS) 2016*, pp. 4062-4070, Barcelona, Spain, Dec. 5-10, 2016. [[pdf](#)] [[Supplementary Material](#)]
  63. T. Chen, A. Mokhtari, X. Wang, A. Ribeiro, and G. B. Giannakis, "A Data-driven Approach to Stochastic Network Optimization", *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Washington DC, DC, USA, 2016, pp. 510-514. [[pdf](#)]
  64. H. Zhang, W. Shi, A. Mokhtari, A. Ribeiro, and Q. Ling, "Decentralized Constrained Consensus Optimization with Primal-Dual Splitting Projection", *2016 IEEE Global Conf. on Signal and Inform. Process. (GlobalSIP)*, Washington DC, DC, USA, 2016, pp. 565-569. [[pdf](#)]
  65. M. Eisen, A. Mokhtari, and A. Ribeiro, "An Asynchronous Quasi-Newton Method for Consensus Optimization", *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Washington DC, DC, USA, 2016, pp. 570-574. [[pdf](#)]
  66. A. Mokhtari, W. Shi, and Qing Ling, "ESOM: Exact Second-Order Method for Consensus Optimization," *50th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, 2016, pp. 783-787. [[pdf](#)]
  67. A. Koppel, A. Mokhtari, and A. Ribeiro, "Doubly Stochastic Algorithms for Large-Scale Optimization," *50th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, 2016, pp. 1705-1709. [[pdf](#)]
  68. A. Mokhtari, S. Shahrampour, A. Jadbabaie, and A. Ribeiro, "Online Optimization in Dynamic Environments: Improved Regret Rates for Strongly Convex Problems", *IEEE 55th Conf. on Decision and Control (CDC)*, pp. 7195-7201, Las Vegas, NV, 2016. [[pdf](#)]
  69. A. Mokhtari, W. Shi, Q. Ling, and A. Ribeiro, "A Decentralized Second-Order Method for Dynamic Optimization", *IEEE 55th Conf. on Decision and Control (CDC)*, pp. 6036-6043, Las Vegas, NV, 2016. [[pdf](#)]
  70. M. Eisen, A. Mokhtari, and A. Ribeiro, "A Decentralized Quasi-Newton Method for Dual Formulations of Consensus Optimization," *IEEE 55th Conf. on Decision and Control (CDC)*, pp. 1951-1958, Las Vegas, NV, 2016. [[pdf](#)]
  71. A. Simonetto, A. Koppel, A. Mokhtari, G. Leus, and A. Ribeiro, "A Quasi-Newton Prediction-Correction Method for Decentralized Dynamic Convex Optimization," *European Control Conference (ECC)*, pp. 1934-1939, Aalborg, Denmark, 2016. [[pdf](#)]



72. A. Mokhtari, A. Koppel, and A. Ribeiro, "Doubly Random Parallel Stochastic Methods for Large Scale Learning," *American Control Conference (ACC)*, pp. 4847-4852, 2016. [\[pdf\]](#)
73. A. Simonetto, A. Mokhtari, A. Koppel, G. Leus, and A. Ribeiro, "A Decentralized Prediction-Correction Method for Networked Time-Varying Convex Optimization," *IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 509-512, 2015. [\[pdf\]](#)
74. A. Mokhtari, W. Shi, Q. Ling, and A. Ribeiro, "Decentralized Quadratically Approximated Alternating Direction Method of Multipliers", *IEEE Global Conf. on Signal and Inform. Process.*, pp. 795-799, Orlando, FL, Dec. 14-16, 2015. [\[pdf\]](#)
75. A. Koppel, A. Simonetto, A. Mokhtari, G. Leus, and A. Ribeiro, "Target Tracking with Dynamic Convex Optimization," *IEEE Global Conf. on Signal and Inform. Process.*, pp. 1210-1214, Orlando, FL, Dec. 14-16, 2015. [\[pdf\]](#)
76. A. Mokhtari and A. Ribeiro, "Decentralized Double Stochastic Averaging Gradient," *Asilomar Conference on signals, systems, and computers*, pp. 406-410, 2015. [\[pdf\]](#)
77. A. Simonetto, A. Koppel, A. Mokhtari, G. Leus, and A. Ribeiro, "Prediction-Correction Methods for Time-Varying Convex Optimization," *Asilomar Conference on signals, systems, and computers*, pp. 666-670, Pacific Grove, CA, Nov. 8-11, 2015. [\[pdf\]](#)
78. A. Mokhtari, Q. Ling, and A. Ribeiro, "An Approximate Newton Method for Distributed Optimization," *Proc. Int. Conf. Acoustics Speech Signal Process.*, pp. 2959-2963, 2015. [\[pdf\]](#)
79. A. Mokhtari, Q. Ling, and A. Ribeiro, "Network Newton," *Asilomar Conf. on signals, systems, and computers*, pp. 1621-1625, Pacific Grove, CA, Nov. 2-5, 2014. [\[pdf\]](#)
80. A. Mokhtari and A. Ribeiro, "A Quasi-Newton Method for Large Scale Support Vector Machines," *Int. Conf. Acoustics Speech Signal Process.*, pp. 8302-8306, 2014. [\[pdf\]](#)
81. A. Mokhtari and A. Ribeiro, "Regularized Stochastic BFGS algorithm," *IEEE Global Conf. on Signal and Inform. Process.*, pp. 1109-1112, Austin, TX, Dec, 2013. [\[pdf\]](#)
82. A. Mokhtari and A. Ribeiro, "A Dual Stochastic DFP Algorithm for Optimal Resource Allocation in Wireless Systems," *IEEE Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, pp. 21-25, Darmstadt, Germany, June 16-19, 2013. [\[pdf\]](#)

## Journal Papers

1. R. Jiang, A. Mokhtari. "Generalized Optimistic Methods for Convex-Concave Saddle Point Problems." *SIAM Journal on Optimization*, 2025. [\[pdf\]](#)
2. Q. Jin, T. Ren, N. Ho, A. Mokhtari, "Statistical and Computational Complexities of BFGS Quasi-Newton Method for Generalized Linear Models," *Transactions on Machine Learning Research (TMLR)*, 2024. [\[pdf\]](#)
3. I. Tziotis, Z. Shen, R. Pedarsani, H. Hassani, A. Mokhtari. "Straggler-Resilient Personalized Federated Learning," *Transactions on Machine Learning Research (TMLR)*, 2023. [\[pdf\]](#)
4. M. Fereydounian, A. Mokhtari, R. Pedarsani, H. Hassani. "Provably Private Distributed Averaging Consensus: An Information-Theoretic Approach," *IEEE Transactions on Information Theory*, 2023. [\[pdf\]](#)
5. Q. Jin, A. Mokhtari. "Non-asymptotic Superlinear Convergence of Standard Quasi-Newton Methods," *Mathematical Programming (MAPR)*, 2022. [\[pdf\]](#)
6. A. Reisizadeh, I. Tziotis, H. Hassani, A. Mokhtari, R. Pedarsani, "Straggler-Resilient Federated Learning: Leveraging the Interplay Between Statistical Accuracy and System Heterogeneity," *IEEE Journal on Selected Areas in Information Theory (JSAT)*, 2022. [\[pdf\]](#)
7. A. Mokhtari and A. Ribeiro. "Stochastic Quasi-Newton Methods," *Proceedings of the IEEE*, vol. 108, no. 11, pp. 1906-1922, 2020. [\[pdf\]](#) [survey paper]
8. A. Mokhtari, A. Ozdaglar, and S. Pattathil. "Convergence Rate of  $O(1/k)$  for Optimistic Gradient and Extra-gradient Methods in Smooth Convex-Concave Saddle Point Problems," *SIAM Journal on Optimization*, vol. 30, no. 4, pp. 3230-3251, 2020. [\[pdf\]](#)

9. H. Hassani, A. Karbasi, A. Mokhtari, and Z. Shen. "Stochastic Conditional Gradient++: (Non-)Convex Minimization and Continuous Submodular Maximization," *SIAM Journal on Optimization*, vol. 30, no. 4, pp. 3315-3344, 2020. [\[pdf\]](#)
10. A. Mokhtari, H. Hassani, and A. Karbasi, "Stochastic Conditional Gradient Methods: From Convex Minimization to Submodular Maximization," *Journal of Machine Learning Research (JMLR)*, vol. 21, no. 105, pp. 1-49, 2020. [\[pdf\]](#)
11. A. Mokhtari, A. Koppel, M. Takac, and A. Ribeiro, "A Class of Parallel Doubly Stochastic Algorithms for Large-Scale Learning," *Journal of Machine Learning Research (JMLR)*, vol. 21, no.120, pp.1-51, 2020. [\[pdf\]](#)
12. S. Paternain, A. Mokhtari, and A. Ribeiro, "A Newton-based Method for Nonconvex Optimization with Fast Evasion of Saddle Points," *SIAM Journal on Optimization*, vol. 29, no. 1, pp. 343-368, 2019. [\[pdf\]](#)
13. A. Reisizadeh, A. Mokhtari, H. Hassani, and R. Pedarsani, "An Exact Quantized Decentralized Gradient Descent Algorithm," *IEEE Transactions on Signal Processing*, vol. 67, no. 19, pp. 4934-4947, 2019. [\[pdf\]](#)
14. M. Eisen, A. Mokhtari, and A. Ribeiro. "A Primal-Dual Quasi-Newton Method for Exact Consensus Optimization," *IEEE Transactions on Signal Processing*, vol. 67, no. 23, pp. 5983-5997., 2019. [\[pdf\]](#)
15. A. Mokhtari, M. Eisen, and A. Ribeiro, "IQN: An Incremental Quasi-Newton Method with Local Superlinear Convergence Rate," *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 1670-1698, 2018. [\[pdf\]](#)
16. A. Mokhtari, M. Gürbüzbalaban, and A. Ribeiro, "Surpassing Gradient Descent Provably: A Cyclic Incremental Method with Linear Convergence Rate," *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 1420-1447, 2018. [\[pdf\]](#)
17. A. Simonetto, A. Koppel, A. Mokhtari, G. Leus, and A. Ribeiro, "Decentralized Prediction-Correction Methods for Networked Time-Varying Convex Optimization," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5724-5738, Nov. 2017. [\[pdf\]](#)
18. T. Chen, A. Mokhtari, X. Wang, A. Ribeiro, and G. B. Giannakis, "Stochastic Averaging for Constrained Optimization with Application to Online Resource Allocation," *IEEE Transactions on Signal Processing*, vol. 65, no. 12, pp. 3078-3098, June 15, 15 2017. [\[pdf\]](#)
19. M. Eisen, A. Mokhtari, and A. Ribeiro, "Decentralized Quasi-Newton Methods," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2613-2628, May15, 15 2017. [\[pdf\]](#)
20. A. Mokhtari, Q. Ling, and A. Ribeiro, "Network Newton Distributed Optimization Methods," *IEEE Transactions on Signal Processing*, vol. 65, no. 1, pp. 146-161, Jan.1, 1 2017. [\[pdf\]](#)
21. A. Mokhtari and A. Ribeiro, "DSA: Decentralized Double Stochastic Averaging Gradient Algorithm," *Journal of Machine Learning Research*, 17(61):1-35, 2016. [\[pdf\]](#)
22. A. Mokhtari, W. Shi, Q. Ling, and A. Ribeiro, "A Decentralized Second-Order Method with Exact Linear Convergence Rate for Consensus Optimization," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 507-522, Dec. 2016. [\[pdf\]](#)
23. A. Mokhtari, W. Shi, Q. Ling, and A. Ribeiro, "DQM: Decentralized Quadratically Approximated Alternating Direction Method of Multipliers," *IEEE Transactions on Signal Processing*, vol. 64, no. 19, pp. 5158-5173, Oct. 1, 2016. [\[pdf\]](#)
24. A. Simonetto, A. Mokhtari, A. Koppel, G. Leus, and A. Ribeiro, "A Class of Prediction-Correction Methods for Time-Varying Convex Optimization," in *IEEE Transactions on Signal Processing*, vol. 64, no. 17, pp. 4576-4591, Sept.1, 1 2016. [\[pdf\]](#)
25. A. Mokhtari and A. Ribeiro, "Global Convergence of Online Limited Memory BFGS," *Journal of Machine Learning Research*, vol. 16, pp. 3151-3181, 2015. [\[pdf\]](#)
26. A. Mokhtari and A. Ribeiro, "RES: Regularized Stochastic BFGS Algorithm," *IEEE Trans. Signal Process.*, vol. 62, no. 23, pp. 6089 - 6104, December 2014. [\[pdf\]](#)

## Thesis

1. A. Mokhtari, "Efficient Methods for Large-Scale Empirical Risk Minimization," Ph.D. Dissertation, University

of Pennsylvania, 2017. (**Joseph and Rosaline Wolf Best Doctoral Dissertation Award**). [\[pdf\]](#)

## Preprints

1. J. Block, S. Srinivasan, L. Collins, A. Mokhtari, S. Shakkottai, "Meta-Learning Adaptable Foundation Models," under review. [\[pdf\]](#)
2. R. Jiang, A. Mokhtari, "Online Learning Guided Quasi-Newton Methods with Global Non-Asymptotic Convergence," *Mathematical Programming*, under review. [\[pdf\]](#)
3. Q. Jin, R. Jiang, A. Mokhtari, "Non-asymptotic Global Convergence Rates of BFGS with Exact Line Search," *Mathematical Programming*, under review. [\[pdf\]](#)
4. R. Jiang, A. Mokhtari, "Generalized Optimistic Methods for Convex-Concave Saddle Point Problems," *SIAM Journal on Optimization*, under review. [\[pdf\]](#)

## MAJOR INVITED TALKS

- "Accelerating Nonconvex Optimization via Online Learning"  
**Keynote talk at Horizons for Optimization in AI Workshop, The Institute for Learning-enabled Optimization at Scale (TILOS)**, San Diego, April 2025.
- "Improved Complexity for Smooth Nonconvex Optimization"  
**Learning Theory Seminar, Google Research**, remote, January 2025.
- "Online Learning Guided Quasi-Newton Methods: Improved Global Non-asymptotic Guarantees"  
**Optimization for Machine Learning Workshop, NeurIPS 2024**, December 2024  
**UPenn Optimization Seminar, Wharton, University of Pennsylvania**, November 2024  
**Machine Learning Seminar, The University of Washington, Seattle**, October 2024  
**Algorithms and Randomness Center (ARC) Seminar Series, Georgia Tech**, September 2024  
**Foundations of Data Science (FDS) Seminar Series, ETH Zurich**, May 2024  
**ECE Seminar Series, EPFL**, May 2024.  
**Learning and Optimization in Luminy (LOL 24) Conference**, Luminy, Marseille, France, June 2024  
**Learning Theory Seminar Series, Google Research**, NY, April 2024.  
**Operations Research Center (ORC) Seminar Series, MIT**, March 2024.  
**Systems, Information, Learning, Optimization (SILO) Seminar Series, UW-Madison**, March 2024.  
**IFML Workshop on Generative AI, UT Austin**, December 2023.  
**ORIE Department Colloquium, Cornell University**, October 2023.
- "In-Context Learning with Transformers: Softmax Attention Adapts to Function Lipschitzness"  
**ITA Workshop**, February 2024.
- "Representation Learning via MAML and its applications in Networking"  
**NSF AI-EDGE Institute Annual Meeting at Northeastern University**, July 2023.
- "A Conditional Gradient Method for Simple Bilevel Optimization with Convex Lower-level Problem"  
**SIAM Conference on Optimization**, June 2023.
- "FedAvg with Fine Tuning: Local Updates Lead to Representation Learning"  
**2023 ITA Workshop**, February 2023.  
**2022 Google Workshop on Federated Learning and Analytics, Google**, November 2022.  
**NSF AI-EDGE Institute Annual Meeting at The Ohio State University**, September 2022.
- "The Power of Adaptivity in Representation Learning: From Meta-Learning to Federated Learning"  
**IFML Seminar Series, UT Austin**, November 2022.  
**TILOS and OPTML++ Seminar Series, MIT**, October 2022.  
**Machine Learning Seminar Series, College of Science and Engineering, U. of Minnesota**, October 2022.
- "Representation Learning with Model-Agnostic Meta-Learning (MAML)"  
**Data Science Institute for Applied Math and CS, Texas A&M University**, October 2022.  
**NSF AI-EDGE Institute Annual Meeting at The Ohio State University**, September 2022.  
**Information Theory and Applications (ITA) Workshop**, May 2022.  
**6G@UT Kick-off Event**, May 2022.
- "Optimistic High-order Methods for Saddle Point Problems"  
**INFORMS Annual Meeting**, October 2021.  
**SIAM Conference on Optimization (OP21)**, July 2021.



- “Towards Personalized Federated Learning via Representation Learning and Meta-Learning”  
**Amazon Research**, July 2021  
**NSF-TRIPODS Workshop on Communication Efficient Distributed Optimization**, April 2021.
- “Exploiting Fast Local Convergence of Second-Order Methods Globally: Adaptive Sample Size Methods”  
**SIAM Conference on Computational Science and Engineering (CSE21)**, March 2021.
- “Gradient-Based Model-Agnostic Meta-Learning Algorithms”  
**Informa Annual Meeting**, October 2020.
- “Communication-Efficient Federated Learning with Periodic Averaging and Quantization”  
**Information Theory and Applications (ITA) Workshop**, February 2020.
- “Understanding the Role of Optimism in Minimax Optimization”  
**Bridging Game Theory and Deep Learning Workshop at NeurIPS**, December 2019.
- “Large-scale Optimization for Machine Learning and Data Science”  
**University of Illinois at Urbana-Champaign**, Computer Science Dept., March 2019.  
**Johns Hopkins University**, Mathematical Institute for Data Science (MINDS), March 2019.  
**Georgia Tech**, Schools of Electrical and Computer Eng. & Industrial and Systems Eng., March 2019.  
**University of Washington**, Industrial and Systems Eng. Dept., Feb 2019.  
**Rutgers Business School**, Management Science & Information Systems Dept., Feb 2019.  
**University of Texas at Austin**, Electrical & Computer Engineering Dept., Feb 2019.  
**Rensselaer Polytechnic Inst. (RPI)**, Electrical, Computer, and Systems Engineering Dept., Jan 2019.  
**Purdue University**, Schools of Electrical and Computer Eng. & Industrial Eng., Jan 2019.

#### RESEARCH GRANTS

- Google Research Scholar Award
  - Title: “Representation Learning and Fine-tuning of Large Models in Federated Environments”
  - Role: PI
  - Duration: 4/1/2024-no end date
  - Total funding: \$60,000.00
- NSF: Division of Computing and Communication Foundations (CCF)
  - Title: “CAREER: Structured Minimax Optimization: Theory, Algorithms, and Applications in Robust Learning”
  - Role: PI
  - Duration: 08/01/2024-07/30/2029
  - Total funding: \$659,959.00
- Qualcomm
  - Title: “Representation Learning and Meta-Learning for Multi-Task Learning”
  - Role: Co-PI
  - Duration: 02/01/2022-02/01/2024
  - Total funding: \$100,000 out of \$200,000
- NSF – National Artificial Intelligence (AI) Research Institute
  - Title: “AI Institute: Symbiotic Foundations for AI and Network Research”
  - Role: Co-PI
  - Duration: 10/01/2021-09/30/2026
  - Total funding: \$600,000 out of \$1,800,000 [Institute budget:  $\approx$  \$20,000,000]
- NSF, Division of Electrical, Communication and Cyber Systems (ECCS)
  - Title: “Collaborative Research: Computationally Efficient Algorithms for Large-scale Bilevel Optimization Problems”
  - Role: PI
  - Duration: 09/15/2021-08/31/2024

- Total funding: \$224,375 out of \$448,750
- ARO Early Career Program (ECP) Award
  - Title: “Accelerating Adaptivity under Limited Data and Computation: A Meta-Learning Approach”
  - Role: sole-PI
  - Duration: 05/17/2021-11/16/2025
  - Total funding: \$360,000
- Machine Learning Laboratory at UT Austin
  - Title: “Learning to Learn Fast: Theory and Algorithm”
  - Role: sole-PI
  - Duration: 06/01/2021-05/30/2022
  - Total funding: \$50,000
- NSF, Division of Computing and Communication Foundations (CCF)
  - Title: “CIF Small: Computationally Efficient Second-Order Optimization Algorithms for Large-Scale Learning”
  - Role: sole-PI
  - Duration: 07/01/2020-06/30/2025
  - Total funding: \$500,000

#### CURRENT POSTDOCS

- Ali Kavis (Nov. 2023 – present) [co-advised with Sujay Sanghavi]

#### CURRENT PH.D. STUDENTS

- Ruichen Jiang (Sep. 2020 – present)
- Jincheng Cao (Sep. 2022 – present)
- Jacob Block (Sep. 2023 – present) [co-advised with Sanjay Shakkottai]
- Sundararajan Srinivasan (Sep. 2023 – present) [co-advised with Sanjay Shakkottai]
- Quentin Fruytier (Sep. 2023 – present) [co-advised with Sujay Sanghavi]
- Francisco Patitucci Perez (Sep. 2023 – present)

#### PAST GRADUATE STUDENTS

- Qiujiang Jin (Jan. 2020 – May 2024)
  - Degree: Ph.D.
  - Next position: Goldman Sachs
- Devyani Maladkar (Jan. 2023 – May 2024)
  - Degree: M.Sc.
  - Next position: Goldman Sachs
- Liam Collins (Sep. 2019 – Aug. 2024) [co-advised with Sanjay Shakkottai]
  - Degree: Ph.D.
  - Next position: Snap Research
- Isidoros Tziotis (Jan. 2020 – Aug. 2024)
  - Degree: Ph.D.
  - Next position: Meta

## PROFESSIONAL SERVICE

- Senior Area Chair for:
  - Neural Information Processing Systems Conference (NeurIPS 2024, 2025)
  - International Conference on Machine Learning (ICML 2025)
- Area Chair for:
  - Neural Information Processing Systems Conference (NeurIPS 2021–2023)
  - International Conference on Machine Learning (ICML 2021–2024)
  - International Conference on Artificial Intelligence and Statistics (AISTATS 2021, 2024)
  - Conference on Learning Theory (COLT 2023–2025)
- Journal refereeing:
  - SIAM Journal on Optimization (SIOPT)
  - Mathematical Programming Journal
  - Journal of Machine Learning Research (JMLR)
  - Journal of Optimization Theory and Applications
  - IEEE Transactions on Information Theory
  - IEEE Transactions on Signal Processing (TSP)
  - IEEE Transactions on Automatic Control (TAC)
  - IEEE Transactions on Control of Network Systems (TCNS)
  - IEEE Transactions on Signal and Information Processing over Networks (TSIPN)
  - Journal of Selected Topics in Signal Processing (JSTSP)
  - IEEE Transactions on Network Science and Engineering (TNSE)
  - IEEE Signal Processing Letters (SPL)
- Conference refereeing:
  - Annual Symposium on Foundations of Computer Science (FOCS 2022, 2025)
  - Conference on Learning Theory (COLT 2020, 2022, 2023, 2024, 2025)
  - Neural Information Processing Systems Conference (NeurIPS 2018–2020)
  - International Conference on Machine Learning (ICML 2019, 2020)
  - International Conference on Artificial Intelligence and Statistics (AISTATS 2019, 2020, 2024)
  - International Conference on Learning Representations (ICLR 2021, 2022, 2025)
  - Innovations in Theoretical Computer Science Conference (ITCS 2024)
  - IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP 2021–2024)
  - IEEE Int. Symposium on Information Theory (ISIT 2020, 2021, 2022)