

# Quanquan Gu

---

CONTACT INFORMATION	Department of Computer Science University of California, Los Angeles EVI 382, 404 Westwood Plaza, Los Angeles, CA 90095	Phone: 01-217-898-9874 Email: qgu@cs.ucla.edu <a href="http://web.cs.ucla.edu/~qgu/">http://web.cs.ucla.edu/~qgu/</a>
RESEARCH INTERESTS	My research is in machine learning, with a focus on developing and analyzing nonconvex optimization algorithms for machine learning to understand large-scale, dynamic, complex, and multi-modal data, and build the foundations of deep learning and reinforcement learning. Recently, I use AI to empower scientific discovery in domains such as biology, medicine, chemistry, and public health.	
PROFESSIONAL EXPERIENCE	<b>Department of Computer Science, University of California, Los Angeles, CA, USA</b> 2022.7 - now <ul style="list-style-type: none"><li>• Associate Professor with Tenure</li></ul> <b>ByteDance Research, Los Angeles, CA, USA</b> 2023.7 - now <ul style="list-style-type: none"><li>• Research Scientist</li></ul> <b>Department of Computer Science, University of California, Los Angeles, CA, USA</b> 2018.7 - 2022.6 <ul style="list-style-type: none"><li>• Tenure-track Assistant Professor</li></ul> <b>Simons Institute for the Theory of Computing, Online</b> 2020.8 - 2020.12 <ul style="list-style-type: none"><li>• Long-term Participant in the Theory of Reinforcement Learning Program</li></ul> <b>Institute for Advanced Study, Princeton, NJ, USA</b> 2019.10 - 2019.11 <ul style="list-style-type: none"><li>• Short-term Visitor in the Special Year on Optimization, Statistics, and Theoretical Machine Learning</li></ul> <b>Simons Institute for the Theory of Computing, Berkeley, CA, USA</b> 2019.5 - 2019.8 <ul style="list-style-type: none"><li>• Research Fellow in the Foundations of Deep Learning Summer Program</li></ul> <b>Department of Computer Science, University of Virginia, VA, USA</b> 2016 - 2018.6 <ul style="list-style-type: none"><li>• Tenure-track Assistant Professor</li></ul> <b>Department of Systems and Information Engineering, University of Virginia, VA, USA</b> 2015 - 2017 <ul style="list-style-type: none"><li>• Tenure-track Assistant Professor</li></ul> <b>Department of Operations Research and Financial Engineering, Princeton University, NJ, USA</b> 2014 - 2015 <ul style="list-style-type: none"><li>• Postdoctoral Research Associate in Statistics</li></ul>	
EDUCATION	<b>University of Illinois at Urbana-Champaign, IL, USA</b> 2010 - 2014 <ul style="list-style-type: none"><li>• Ph.D. in Computer Science</li></ul> <b>Tsinghua University, Beijing, China</b> 2007 - 2010 <ul style="list-style-type: none"><li>• Master of Science in Control Science and Engineering</li></ul> <b>Tsinghua University, Beijing, China</b> 2003 - 2007 <ul style="list-style-type: none"><li>• Bachelor of Engineering in Automation</li></ul>	
HONORS AND AWARDS	<ul style="list-style-type: none"><li>• WSDM Test of Time Paper Award 2024</li><li>• Alfred P. Sloan Research Fellowship 2022</li><li>• JP Morgan Faculty Research Award 2022</li></ul>	

- AWS Machine Learning Research Award 2020
- IJCAI Early Career Talk 2020
- Simons Berkeley Research Fellowship 2019
- Salesforce Deep Learning Research Award 2018
- Adobe Data Science Research Award 2018
- NSF CAREER Award 2017
- UVa SEAS Research Innovation Award 2017
- Yahoo! Academic Career Enhancement Award 2015
- IBM PhD Fellowship 2013-2014
- UIUC Computer Science Department Fellowship 2010
- Best Master Thesis Award in Tsinghua University 2010

TEACHING  
EXPERIENCE

**Instructor, University of California, Los Angeles**

- CS 161: Fundamentals of AI (Winter 2020, Winter 2021, Winter 2022, Winter 2023)
- CS 260: Machine Learning Algorithms (Fall 2018, Spring 2020, Fall 2021, Fall 2022)
- CS 269: Foundations of Deep Learning (Winter 2019, Spring 2021, Spring 2022)
- CS 269: Optimization for Machine Learning (Spring 2024)

**Instructor, University of Virginia**

- CS 6316/SYS6016: Machine Learning (Spring 2018)
- CS 6501/SYS6003: Optimization for Machine Learning (Fall 2017)
- SYS 6003: Optimization Models and Methods (Fall 2015, 2016)
- SYS 4582/6016: Machine Learning (Spring 2017)
- SYS 3060: Stochastic Decision Models (Spring 2016)

**Teaching Assistant, Department of Computer Science, UIUC**

- CS 512: Data Mining: Principles and Algorithms (Spring 2013)
- CS 412: An Introduction to Data Warehousing and Data Mining (Fall 2012)

PUBLICATION

1. Zixiang Chen\*, Yihe Deng\*, Huizhuo Yuan\*, Kaixuan Ji and Quanquan Gu, Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models, in Proc. of the 41st International Conference on Machine Learning (**ICML**), Vienna, Austria, 2024.
2. Junkai Zhang\*, Weitong Zhang\*, Dongruo Zhou and Quanquan Gu, Uncertainty-Aware Reward-Free Exploration with General Function Approximation, in Proc. of the 41st International Conference on Machine Learning (**ICML**), Vienna, Austria, 2024.
3. Xuheng Li, Heyang Zhao and Quanquan Gu, Feel-Good Thompson Sampling for Contextual Dueling Bandits, in Proc. of the 41st International Conference on Machine Learning (**ICML**), Vienna, Austria, 2024.
4. Yue Wu, Tao Jin\*, Qiwei Di\*, Hao Lou, Farzad Farnoud and Quanquan Gu, Borda Regret Minimization for Generalized Linear Dueling Bandits, in Proc. of the 41st International Conference on Machine Learning (**ICML**), Vienna, Austria, 2024.
5. Chenlu Ye\*, Jiafan He\*, Quanquan Gu and Tong Zhang, Towards Robust Model-Based Reinforcement Learning Against Adversarial Corruption, in Proc. of the 41st International Conference on Machine Learning (**ICML**), Vienna, Austria, 2024.
6. Xinyou Wang\*, Zaixiang Zheng\*, Fei Ye, Dongyu Xue, Shujian Huang and Quanquan Gu, Diffusion Language Models Are Versatile Protein Learners, in Proc. of the 41st International Conference on Machine Learning (**ICML**), Vienna, Austria, 2024.
7. Yan Wang\*, Lihao Wang\*, Yuning Shen, Yiqun Wang, Huizhuo Yuan, Yue Wu and Quanquan Gu, Protein Conformation Generation via Force-Guided SE(3) Diffusion Models, in Proc. of the 41st International Conference on Machine Learning (**ICML**), Vienna, Austria, 2024.

8. Yue Huang, Lichao Sun, ... Quanquan Gu, ... and Yue Zhao, Position: TrustLLM: Trustworthiness in Large Language Models, in Proc. of the 41st International Conference on Machine Learning (**ICML**), Vienna, Austria, 2024.
9. Zichen Wang, Chuanhao Li, Chenyu Song, Lianghai Wang, **Quanquan Gu** and Huazheng Wang, Pure Exploration in Asynchronous Federated Bandits, in Proc. of the 40th International Conference on Uncertainty in Artificial Intelligence (**UAI**), Barcelona, Spain, 2024.
10. Yihe Deng\*, Ruochi Zhang\*, Pan Xu, Jian Ma and Quanquan Gu, Pre-trained Hypergraph Convolutional Neural Networks with Self-supervised Learning, Transaction on Machine Learning Research (**TMLR**), 2024.
11. Dongruo Zhou\*, Jinghui Chen\*, Yuan Cao\*, Ziyang Yang and Quanquan Gu, On the Convergence of Adaptive Gradient Methods for Nonconvex Optimization, Transaction on Machine Learning Research (**TMLR**), 2024.
12. Difan Zou\*, Jingfeng Wu\*, Vladimir Braverman, Quanquan Gu and Sham M. Kakade, Benign Overfitting of Constant-Stepsize SGD for Linear Regression, Journal of Machine Learning Research (**JMLR**), 2024.
13. Jingfeng Wu, Difan Zou, Zixiang Chen, Vladimir Braverman, Quanquan Gu and Peter L. Bartlett, How Many Pretraining Tasks Are Needed for In-Context Learning of Linear Regression?, in Proc. of the 11th International Conference on Learning Representations (**ICLR**), 2024.
14. Xiangxin Zhou, Xiwei Cheng, Yuwei Yang, Yu Bao, Liang Wang and Quanquan Gu, DecompOpt: Controllable and Decomposed Diffusion Models for Structure-based Molecular Optimization, in Proc. of the 11th International Conference on Learning Representations (**ICLR**), 2024.
15. Xuheng Li, Yihe Deng, Jingfeng Wu, Dongruo Zhou and Quanquan Gu, Risk Bounds of Accelerated SGD for Overparameterized Linear Regression, in Proc. of the 11th International Conference on Learning Representations (**ICLR**), 2024.
16. Kaixuan Ji\*, Qingyue Zhao\*, Jiafan He, Weitong Zhang and Quanquan Gu, Horizon-free Reinforcement Learning in Adversarial Linear Mixture MDPs, in Proc. of the 11th International Conference on Learning Representations (**ICLR**), 2024.
17. Zixiang Chen\*, Yihe Deng\*, Yuanzhi Li and Quanquan Gu, Understanding Transferable Representation Learning and Zero-shot Transfer in CLIP, in Proc. of the 11th International Conference on Learning Representations (**ICLR**), 2024.
18. Qiwei Di, Heyang Zhao, Jiafan He and Quanquan Gu, Pessimistic Nonlinear Least-Squares Value Iteration for Offline Reinforcement Learning, in Proc. of the 11th International Conference on Learning Representations (**ICLR**), 2024.
19. Qiwei Di, Tao Jin, Yue Wu, Heyang Zhao, Farzad Farnoud and Quanquan Gu, Variance-Aware Regret Bounds for Stochastic Contextual Dueling Bandits, in Proc. of the 11th International Conference on Learning Representations (**ICLR**), 2024.
20. Yiwen Kou\*, Zixiang Chen\* and **Quanquan Gu**, Implicit Bias of Gradient Descent for Two-layer ReLU and Leaky ReLU Networks on Nearly-orthogonal Data, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS**) 36, New Orleans, LA, USA, 2023.
21. Huizhuo Yuan\*, Chris Junchi Li\*, Gauthier Gidel, Michael I. Jordan, **Quanquan Gu** and Simon Shaolei Du, Optimal Extragradient-Based Algorithms for Stochastic Variational Inequalities with Separable Structure, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS**) 36, New Orleans, LA, USA, 2023.
22. Chenlu Ye, Rui Yang, Quanquan Gu and **Tong Zhang**, Corruption-Robust Offline Reinforcement Learning with General Function Approximation, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS**) 36, New Orleans, LA, USA, 2023.
23. Yihe Deng\*, Yu Yang\*, Baharan Mirzasoleiman and **Quanquan Gu**, Robust Learning with Progressive Data Expansion Against Spurious Correlation, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS**) 36, New Orleans,

- LA, USA, 2023.
24. Zixiang Chen\*, Junkai Zhang\*, Yiwen Kou, Xiangning Chen, Cho-Jui Hsieh and **Quanquan Gu**, Why Does Sharpness-Aware Minimization Generalize Better Than SGD? in Proc. of Advances in Neural Information Processing Systems (**NeurIPS**) 36, New Orleans, LA, USA, 2023.
  25. Heyang Zhao, Jiafan He, Dongruo Zhou, Tong Zhang and **Quanquan Gu**, Variance-Dependent Regret Bounds for Linear Bandits and Reinforcement Learning: Adaptivity and Computational Efficiency, in Proc. of the 36th Annual Conference on Learning Theory (**COLT**), Bangalore, India, 2023.
  26. Yuan Cao, Difan Zou, Yuanzhi Li and **Quanquan Gu**, The Implicit Bias of Batch Normalization in Linear Models and Two-layer Linear Convolutional Neural Networks, in Proc. of the 36th Annual Conference on Learning Theory (**COLT**), Bangalore, India, 2023.
  27. Lingxiao Wang, Bargav Jayaraman, David Evans and **Quanquan Gu**, Efficient Privacy-Preserving Stochastic Nonconvex Optimization, in Proc. of the 39th International Conference on Uncertainty in Artificial Intelligence (**UAI**), Pittsburgh, PA, USA, 2023.
  28. Yue Wu, Jiafan He and **Quanquan Gu**, Uniform-PAC Guarantees for Model-Based RL with Bounded Eluder Dimension, in Proc. of the 39th International Conference on Uncertainty in Artificial Intelligence (**UAI**), Pittsburgh, PA, USA, 2023.
  29. Weitong Zhang\*, Jiafan He\*, Dongruo Zhou, Amy Zhang and **Quanquan Gu**, Provably Efficient Representation Learning in Low-rank Markov Decision Processes: from online to offline RL, in Proc. of the 39th International Conference on Uncertainty in Artificial Intelligence (**UAI**), Pittsburgh, PA, USA, 2023.
  30. Jinghui Chen\*, Yuan Cao\* and **Quanquan Gu**, Benign Overfitting in Adversarially Robust Linear Classification, in Proc. of the 39th International Conference on Uncertainty in Artificial Intelligence (**UAI**), Pittsburgh, PA, USA, 2023.
  31. **Quanquan Gu**\*\* , Amin Karbasi\*\*, Khashayar Khosravi\*\*, Vahab Mirrokni\*\*, Dongruo Zhou\*\*, Batched Neural Bandits, ACM/IMS Journal of Data Science, 2023.
  32. Yiwen Kou\*, Zixiang Chen\*, Yuanzhou Chen and **Quanquan Gu**, Benign Overfitting for Two-layer ReLU Convolutional Neural Networks , in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  33. Jingfeng Wu\*, Difan Zou\*, Zixiang Chen\*, Vladimir Braverman, **Quanquan Gu** and Sham M. Kakade, Finite-Sample Analysis of Learning High-Dimensional Single ReLU Neuron, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  34. Chris Junchi Li\*, Huizhuo Yuan\*, Simon Du, Gauthier Gidel, **Quanquan Gu** and Michael I. Jordan, Cooperative Multi-Agent Reinforcement Learning: Asynchronous Communication and Linear Function Approximation, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  35. Yifei Min\*, Jiafan He\*, Tianhao Wang\* and **Quanquan Gu**, Cooperative Multi-Agent Reinforcement Learning: Asynchronous Communication and Linear Function Approximation, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  36. Junkai Zhang, Weitong Zhang and **Quanquan Gu**, Optimal Horizon-Free Reward-Free Exploration for Linear Mixture MDPs, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  37. Qiwei Di, Jiafan He, Dongruo Zhou and **Quanquan Gu**, On the Nearly Minimax Optimal Regret for Learning Linear Mixture Stochastic Shortest Path, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  38. Difan Zou, Yuan Cao, Yuanzhi Li and **Quanquan Gu**, The Benefits of Mixup for Feature Learning, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  39. Weitong Zhang, Jiafan He, Zhiyuan Fan and **Quanquan Gu**, On the Interplay

- Between Misspecification and Sub-optimality Gap in Linear Contextual Bandits, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
40. Heyang Zhao, Dongruo Zhou, Jiafan He and **Quanquan Gu**, Optimal Online Generalized Linear Regression with Stochastic Noise and Its Application to Heteroscedastic Bandits, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  41. Chenlu Ye, Wei Xiong, **Quanquan Gu** and Tong Zhang, Corruption-Robust Algorithms with Uncertainty Weighting for Nonlinear Contextual Bandits and Markov Decision Processes, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  42. Jiafan He, Heyang Zhao, Dongruo Zhou and **Quanquan Gu**, Nearly Minimax Optimal Reinforcement Learning for Linear Markov Decision Processes, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  43. Zaixiang Zheng\*, Yifan Deng\*, Dongyu Xue, Yi Zhou, Fei Ye and **Quanquan Gu**, Structure-informed Language Models Are Protein Designers, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  44. Jiaqi Guan\*, Xiangxin Zhou\*, Yuwei Yang, Yu Bao, Jian Peng, Jianzhu Ma, Qiang Liu, Liang Wang and **Quanquan Gu**, DecompDiff: Diffusion Models with Decomposed Priors for Structure-Based Drug Design, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  45. Yue Wu, Shuaicheng Zhang, Wenchao Yu, Yanchi Liu, **Quanquan Gu**, Dawei Zhou, Haifeng Chen and Wei Cheng, Personalized Federated Learning under Mixture of Distributions, in Proc. of the 40 th International Conference on Machine Learning (**ICML**), Hawaii, USA, 2023.
  46. Katriona Shea et al., Multiple Models for Outbreak Decision Support in the face of Uncertainty, in Proceedings of the National Academy of Sciences (**PNAS**), Volume 120, No. 18, 2023.
  47. Zixiang Chen\*, Chris Junchi Li\*, Huizhuo Yuan\*, **Quanquan Gu** and Michael I. Jordan, A General Framework for Sample-Efficient Function Approximation in Reinforcement Learning, in Proc. of the 10th International Conference on Learning Representations (**ICLR**), 2023.
  48. Difan Zou, Yuan Cao, Yuanzhi Li and **Quanquan Gu**, Understanding the Generalization of Adam in Learning Neural Networks with Proper Regularization, in Proc. of the 10th International Conference on Learning Representations (**ICLR**), 2023.
  49. Yiwen Kou, Zixiang Chen, Yuan Cao and **Quanquan Gu**, How Does Semi-supervised learning with Pseudo-labelers Work? A Case Study, in Proc. of the 10th International Conference on Learning Representations (**ICLR**), 2023.
  50. Xinzhe Zuo, Zixiang Chen, Huaxiu Yao, Yuan Cao and **Quanquan Gu**, Understanding Train-Validation Split in Meta-Learning with Neural Networks, in Proc. of the 10th International Conference on Learning Representations (**ICLR**), 2023.
  51. Yuan Cao\*, Zixiang Chen\*, Mikhail Belkin and **Quanquan Gu**, Benign Overfitting in Two-layer Convolutional Neural Networks, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'22**) 35, New Orleans, LA, USA, 2022.
  52. Dongruo Zhou and **Quanquan Gu**, Computationally Efficient Horizon-Free Reinforcement Learning for Linear Mixture MDPs, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'22**) 35, New Orleans, LA, USA, 2022.
  53. Difan Zou\*, Jingfeng Wu\*, Vladimir Braverman, **Quanquan Gu** and Sham M. Kakade, Risk Bounds of Multi-Pass SGD for Least Squares in the Interpolation Regime, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'22**) 35, New Orleans, LA, USA, 2022.
  54. Jingfeng Wu\*, Difan Zou\*, Vladimir Braverman, **Quanquan Gu** and Sham M. Kakade, The Power and Limitation of Pretraining-Finetuning for Linear Regression

- under Covariate Shift, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'22**) 35, New Orleans, LA, USA, 2022.
55. Jiafan He, Dongruo Zhou, Tong Zhang and **Quanquan Gu**, Nearly Optimal Algorithms for Linear Contextual Bandits with Adversarial Corruptions, in Proc. of Advances in Neural Information Processing Systems (NeurIPS'22) 35, New Orleans, LA, USA, 2022.
  56. Chris Junchi Li\*, Dongruo Zhou\*, **Quanquan Gu** and Michael I. Jordan, Learning Two-Player Mixture Markov Games: Kernel Function Approximation and Correlated Equilibrium, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'22**) 35, New Orleans, LA, USA, 2022.
  57. Jiafan He\*, Tianhao Wang\*, Yifei Min\*, **Quanquan Gu**, A Simple and Provably Efficient Algorithm for Asynchronous Federated Contextual Linear Bandits, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'22**) 35, New Orleans, LA, USA, 2022.
  58. Zixiang Chen, Yihe Deng, Yue Wu, **Quanquan Gu** and Yuanzhi Li, Towards Understanding the Mixture-of-Experts Layer in Deep Learning, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'22**) 35, New Orleans, LA, USA, 2022.
  59. Hao Lou, Tao Jin, Yue Wu, Pan Xu, **Quanquan Gu** and Farzad Farnoud, Active Ranking without Strong Stochastic Transitivity, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'22**) 35, New Orleans, LA, USA, 2022.
  60. Chonghua Liao\*, Jiafan He\* and **Quanquan Gu**, Locally Differentially Private Reinforcement Learning for Linear Mixture Markov Decision Processes, in Proc. of the 14th Asia Conference on Machine Learning (**ACML'22**), Hyderabad, India, 2022.
  61. Benjamin Hoar, Weitong Zhang, Shuangning Xu, Rana Deeba, Cyrille Costentin, **Quanquan Gu**, Chong Liu, Electrochemical Mechanistic Analysis from Cyclic Voltammograms Based on Deep Learning, ACS Measurement Science Au, 2022.
  62. Jingfeng Wu\*, Difan Zou\*, Vladimir Braverman, **Quanquan Gu** and Sham M. Kakade, Last Iterate Risk Bounds of SGD with Decaying Stepsize for Overparameterized Linear Regression, in Proc. of the 39th International Conference on Machine Learning (**ICML'22**), Baltimore, MD, USA, 2022.
  63. Yuanzhou Chen\*, Jiafan He\* and **Quanquan Gu**, On the Sample Complexity of Learning Infinite-horizon Discounted Linear Kernel MDPs, in Proc. of the 39th International Conference on Machine Learning (**ICML'22**), Baltimore, MD, USA, 2022.
  64. Yifei Min, Jiafan He, Tianhao Wang and **Quanquan Gu**, Learning Stochastic Shortest Path with Linear Function Approximation, in Proc. of the 39th International Conference on Machine Learning (**ICML'22**), Baltimore, MD, USA, 2022.
  65. Dongruo Zhou and **Quanquan Gu**, Dimension-free Complexity Bounds for High-order Nonconvex Finite-sum Optimization, in Proc. of the 39th International Conference on Machine Learning (**ICML'22**), Baltimore, MD, USA, 2022.
  66. Pan Xu, Zheng Wen, Handong Zhao and **Quanquan Gu**, Neural Contextual Bandits with Deep Representation and Shallow Exploration, in Proc. of the 10th International Conference on Learning Representations (**ICLR'22**), 2022.
  67. Yihan Wang, Zhouxing Shi, **Quanquan Gu** and Cho-Jui Hsieh, On the Convergence of Certified Robust Training with Interval Bound Propagation, in Proc. of the 10th International Conference on Learning Representations (**ICLR'22**), 2022.
  68. Yiling Jia, Weitong Zhang, Dongruo Zhou, **Quanquan Gu** and Hongning Wang, Learning Neural Contextual Bandits through Perturbed Rewards, in Proc. of the 10th International Conference on Learning Representations (**ICLR'22**), 2022.
  69. Yue Wu\*, Tao Jin\*, Hao Lou, Pan Xu, Farzad Farnoud and **Quanquan Gu**, Adaptive Sampling for Heterogeneous Rank Aggregation from Noisy Pairwise Comparisons, in Proc of the 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS'22**), 2022.

70. Yue Wu, Dongruo Zhou and **Quanquan Gu**, Nearly Minimax Optimal Regret for Learning Infinite-horizon Average-reward MDPs with Linear Function Approximation, in Proc of the 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS'22**), 2022.
71. Spencer Frei\*, Difan Zou\*, Zixiang Chen\* and **Quanquan Gu**, Self-training Converts Weak Learners to Strong Learners in Mixture Models, in Proc of the 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS'22**), 2022.
72. Jiafan He, Dongruo Zhou and **Quanquan Gu**, Near-optimal Policy Optimization Algorithms for Learning Adversarial Linear Mixture MDPs, in Proc of the 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS'22**), 2022.
73. Zixiang Chen\*, Dongruo Zhou\* and **Quanquan Gu**, Faster Perturbed Stochastic Gradient Methods for Finding Local Minima, in Proc. of the 33rd International Conference on Algorithmic Learning Theory (**ALT'22**), Paris, France, 2022.
74. Zixiang Chen, Dongruo Zhou and **Quanquan Gu**, Almost Optimal Algorithms for Two-player Zero-Sum Linear Mixture Markov Games, in Proc. of the 33rd International Conference on Algorithmic Learning Theory (**ALT'22**), Paris, France, 2022.
75. Jinghui Chen, Yu Cheng, Zhe Gan, **Quanquan Gu** and Jingjing Liu, Efficient Robust Training via Backward Smoothing, in Proc. of the 36th AAAI Conference on Artificial Intelligence (**AAAI'22**), Vancouver, BC, Canada, 2022.
76. Estee Y Cramer et al., Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the US, in Proceedings of the National Academy of Sciences (**PNAS**), Volume 119, No. 15, 2022.
77. Yuan Cao, **Quanquan Gu** and Mikhail Belkin, Risk Bounds for Over-parameterized Maximum Margin Classification on Sub-Gaussian Mixtures, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
78. Boxi Wu\*, Jinghui Chen\*, Deng Cai, Xiaofei He and **Quanquan Gu**, Do Wider Neural Networks Really Help Adversarial Robustness? in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
79. Spencer Frei and **Quanquan Gu**, Proxy Convexity: A Unified Framework for the Analysis of Neural Networks Trained by Gradient Descent, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
80. Difan Zou\*, Jingfeng Wu\*, Vladimir Braverman, **Quanquan Gu**, Dean P. Foster and Sham M. Kakade, The Benefits of Implicit Regularization from SGD in Least Squares Problems, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
81. Yinglun Zhu\*, Dongruo Zhou\*, Ruoxi Jiang\*, **Quanquan Gu**, Rebecca Willett and Robert Nowak, Pure Exploration in Kernel and Neural Bandits, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
82. Jiafan He, Dongruo Zhou and **Quanquan Gu**, Uniform-PAC Bounds for Reinforcement Learning with Linear Function Approximation, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
83. Jiafan He, Dongruo Zhou and **Quanquan Gu**, Minimax Optimal Reinforcement Learning for Discounted MDPs, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
84. Weitong Zhang, Dongruo Zhou, **Quanquan Gu**, Reward-Free Model-Based Reinforcement Learning with Linear Function Approximation, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
85. Tianhao Wang\*, Dongruo Zhou\* and **Quanquan Gu**, Provably Efficient Reinforcement Learning with Linear Function Approximation Under Adaptivity Constraints, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'21**) 34, 2021.
86. Yifei Min\*, Tianhao Wang\*, Dongruo Zhou and **Quanquan Gu**, Variance-Aware

- Off-Policy Evaluation with Linear Function Approximation, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS’21**) 34, 2021.
87. Luyao Yuan, Dongruo Zhou, Junhong Shen, Jingdong Gao, Jeffrey L Chen, **Quanquan Gu**, Ying Nian Wu and Song-Chun Zhu, Iterative Teacher-Aware Learning, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS’21**) 34, 2021.
  88. Hanxun Huang, Yisen Wang, Sarah Monazam Erfani, **Quanquan Gu**, James Bailey and Xingjun Ma, Exploring Architectural Ingredients of Adversarially Robust Deep Neural Networks, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS’21**) 34, 2021.
  89. Difan Zou\*, Jingfeng Wu\*, Vladimir Braverman, **Quanquan Gu** and Sham M. Kakade, Benign Overfitting of Constant-Stepsize SGD for Linear Regression, in Proc. of the 34th Annual Conference on Learning Theory (**COLT’21**), 2021.
  90. Dongruo Zhou, **Quanquan Gu** and Csaba Szepesvari, Nearly Minimax Optimal Reinforcement Learning for Linear Mixture MDPs, in Proc. of the 34th Annual Conference on Learning Theory (**COLT’21**), 2021.
  91. Tianyuan Jin, Pan Xu, Xiaokui Xiao and **Quanquan Gu**, Double Explore-then-Commit: Asymptotic Optimality and Beyond, in Proc. of the 34th Annual Conference on Learning Theory (**COLT’21**), 2021.
  92. Difan Zou, Pan Xu and **Quanquan Gu**, Faster Convergence of Stochastic Gradient Langevin Dynamics for Non-Log-Concave Sampling, in Proc. of the 37th International Conference on Uncertainty in Artificial Intelligence (**UAI’21**), 2021.
  93. Difan Zou and **Quanquan Gu**, On the Convergence of Hamiltonian Monte Carlo with Stochastic Gradients, in Proc. of the 38th International Conference on Machine Learning (**ICML’21**), 2021.
  94. Difan Zou\*, Spencer Frei\* and **Quanquan Gu**, Provable Robustness of Adversarial Training for Learning Halfspaces with Noise, in Proc. of the 38th International Conference on Machine Learning (**ICML’21**), 2021.
  95. Jiafan He, Dongruo Zhou and **Quanquan Gu**, Logarithmic Regret for Reinforcement Learning with Linear Function Approximation, in Proc. of the 38th International Conference on Machine Learning (**ICML’21**), 2021.
  96. Dongruo Zhou, Jiafan He and **Quanquan Gu**, Provably Efficient Reinforcement Learning for Discounted MDPs with Feature Mapping, in Proc. of the 38th International Conference on Machine Learning (**ICML’21**), 2021.
  97. Spencer Frei, Yuan Cao and **Quanquan Gu**, Agnostic Learning of Halfspaces with Gradient Descent via Soft Margins, in Proc. of the 38th International Conference on Machine Learning (**ICML’21**), 2021. (Long talk)
  98. Spencer Frei, Yuan Cao and **Quanquan Gu**, Provable Generalization of SGD-trained Neural Networks of Any Width in the Presence of Adversarial Label Noise, in Proc. of the 38th International Conference on Machine Learning (**ICML’21**), 2021.
  99. Tianyuan Jin, Pan Xu, Jieming Shi, Xiaokui Xiao and **Quanquan Gu**, MOTs: Minimax Optimal Thompson Sampling, in Proc. of the 38th International Conference on Machine Learning (**ICML’21**), 2021.
  100. Tianyuan Jin, Jing Tang, Pan Xu, Keke Huang, Xiaokui Xiao and **Quanquan Gu**, Almost Optimal Anytime Algorithm for Batched Multi-Armed Bandits, in Proc. of the 38th International Conference on Machine Learning (**ICML’21**), 2021.
  101. Johannes Bracher et al., Short-term Forecasting of COVID-19 in Germany and Poland during the Second Wave—A Preregistered Study, Nature Communications, 2021.
  102. Zixiang Chen\*, Yuan Cao\*, Difan Zou\* and **Quanquan Gu**, How Much Overparameterization Is Sufficient to Learn Deep ReLU Networks? in Proc. of the 9th International Conference on Learning Representations (**ICLR’21**), 2021.
  103. Weitong Zhang, Dongruo Zhou, Lihong Li and **Quanquan Gu**, Neural Thompson Sampling, in Proc. of the 9th International Conference on Learning Representations



- (**ICLR'21**), 2021.
104. Jingfeng Wu, Difan Zou, Vladimir Braverman and **Quanquan Gu**, Direction Matters: On the Implicit Regularization Effect of Stochastic Gradient Descent with Moderate Learning Rate, in Proc. of the 9th International Conference on Learning Representations (**ICLR'21**), 2021.
  105. Lingxiao Wang, Kevin Huang, Tengyu Ma, **Quanquan Gu** and Jing Huang, Variance-reduced First-order Meta-learning for Natural Language Processing Tasks, in Proc. of 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2021.
  106. Yuan Cao\*, Zhiying Fang\*, Yue Wu\*, Ding-Xuan Zhou and **Quanquan Gu**, Towards Understanding the Spectral Bias of Deep Learning, in Proc. of the 30th International Joint Conference on Artificial Intelligence (**IJCAI'21**), Montreal, Canada, 2021.
  107. Bargav Jayaraman, Lingxiao Wang, Katherine Knipmeyer, **Quanquan Gu** and David Evans, Revisiting Membership Inference Under Realistic Assumptions. 21st Privacy Enhancing Technologies Symposium (**PETS**), 2021.
  108. Bao Wang\*, Difan Zou\*, **Quanquan Gu**, Stanley Osher, Laplacian Smoothing Stochastic Gradient Markov Chain Monte Carlo, SIAM Journal on Scientific Computing, 2020.
  109. Yue Wu, Weitong Zhang, Pan Xu and **Quanquan Gu**, A Finite Time Analysis of Two Time-Scale Actor Critic Methods, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'20**) 33, 2020.
  110. Spencer Frei, Yuan Cao and **Quanquan Gu**, Agnostic Learning of a Single Neuron with Gradient Descent, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'20**) 33, 2020.
  111. Zixiang Chen, Yuan Cao, **Quanquan Gu** and Tong Zhang, A Generalized Neural Tangent Kernel Analysis for Two-layer Neural Networks, in Proc. of Advances in Neural Information Processing Systems (**NeurIPS'20**) 33, 2020.
  112. Dongruo Zhou, Lihong Li and **Quanquan Gu**, Neural Contextual Bandits with UCB-Based Exploration in Proc. of the 37th International Conference on Machine Learning (**ICML'20**), 2020.
  113. Pan Xu and **Quanquan Gu**, A Finite-Time Analysis of Q-Learning with Neural Network Function Approximation in Proc. of the 37th International Conference on Machine Learning (**ICML'20**), 2020.
  114. Yonatan Dukler, **Quanquan Gu** and Guido Montufar, Optimization Theory for ReLU Neural Networks Trained with Normalization Layers in Proc. of the 37th International Conference on Machine Learning (**ICML'20**), 2020.
  115. Fabrice Harel-Canada, Lingxiao Wang, Muhammad Ali Gulzar, **Quanquan Gu** and Miryung Kim, Is Neuron Coverage a Meaningful Measure for Testing Deep Neural Networks? in Proc of ACM SIGSOFT International Symposium on the Foundations of Software Engineering (**ESEC/FSE'20**), Sacramento, California, USA, 2020.
  116. Jinghui Chen and **Quanquan Gu**, RayS: A Ray Searching Method for Hard-label Adversarial Attack in Proc of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD'20**), San Diego, CA, USA, 2020.
  117. Dongruo Zhou, Pan Xu and **Quanquan Gu**, Stochastic Nested Variance Reduction for Nonconvex Optimization, Journal of Machine Learning Research (**JMLR**), 2020.
  118. Jinghui Chen, Dongruo Zhou, Yiqi Tang, Ziyang Yang, Yuan Cao and **Quanquan Gu**, Closing the Generalization Gap of Adaptive Gradient Methods in Training Deep Neural Networks in Proc. of the 29th International Joint Conference on Artificial Intelligence (**IJCAI'20**), Yokohama, Japan , 2020.
  119. Bao Wang, **Quanquan Gu**, March Boedihardjo, Lingxiao Wang, Farzin Barekat and Stanley J. Osher, DP-LSSGD: A Stochastic Optimization Method to Lift the Utility in Privacy-Preserving ERM In Proc of the Mathematical and Scientific Machine Learning Conference (**MSML'20**), Princeton, New Jersey, USA, 2020.

120. Xiao Zhang\*, Jinghui Chen\*, **Quanquan Gu** and David Evans, Understanding the Intrinsic Robustness of Image Distributions using Conditional Generative Models, In Proc of the 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS'20**), Palermo, Sicily, Italy, 2020.
121. Dongruo Zhou, Yuan Cao and **Quanquan Gu**, Accelerated Factored Gradient Descent for Low-Rank Matrix Factorization, In Proc of the 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS'20**), Palermo, Sicily, Italy, 2020.
122. Dongruo Zhou and **Quanquan Gu**, Stochastic Recursive Variance-Reduced Cubic Regularization Methods, In Proc of the 23rd International Conference on Artificial Intelligence and Statistics (**AISTATS'20**), Palermo, Sicily, Italy, 2020.
123. Difan Zou, Philip M. Long and **Quanquan Gu**, On the Global Convergence of Training Deep Linear ResNets, in Proc. of the 8th International Conference on Learning Representations (**ICLR'20**), Addis Ababa, Ethiopia, 2020.
124. Pan Xu, Felicia Gao and **Quanquan Gu**, Sample Efficient Policy Gradient Methods with Recursive Variance Reduction, in Proc. of the 8th International Conference on Learning Representations (**ICLR'20**), Addis Ababa, Ethiopia, 2020.
125. Lingxiao Wang, Jing Huang, Kevin Huang, Ziniu Hu, Guangtao Wang and **Quanquan Gu**, Improving Neural Language Generation with Spectrum Control, in Proc. of the 8th International Conference on Learning Representations (**ICLR'20**), Addis Ababa, Ethiopia, 2020.
126. Yisen Wang\*, Difan Zou\*, Jinfeng Yi, James Bailey, Xingjun Ma and **Quanquan Gu**, Improving Adversarial Robustness Requires Revisiting Misclassified Examples, in Proc. of the 8th International Conference on Learning Representations (**ICLR'20**), Addis Ababa, Ethiopia, 2020.
127. Jinghui Chen, Dongruo Zhou, Jinfeng Yi and **Quanquan Gu**, A Frank-Wolfe Framework for Efficient and Effective Adversarial Attacks, in Proc. of the 34th AAAI Conference on Artificial Intelligence (**AAAI'20**), New York, New York, USA, 2020.
128. Yuan Cao and **Quanquan Gu**, Generalization Error Bounds of Gradient Descent for Learning Over-parameterized Deep ReLU Networks, in Proc. of the 34th AAAI Conference on Artificial Intelligence (**AAAI'20**), New York, New York, USA, 2020.
129. Lingxiao Wang and **Quanquan Gu**, A Knowledge Transfer Framework for Differentially Private Sparse Learning, in Proc. of the 34th AAAI Conference on Artificial Intelligence (**AAAI'20**), New York, New York, USA, 2020.
130. Tao Jin\*, Pan Xu\*, **Quanquan Gu** and Farzad Farnoud, Rank Aggregation via Heterogeneous Thurstone Preference Models, in Proc. of the 34th AAAI Conference on Artificial Intelligence (**AAAI'20**), New York, New York, USA, 2020. Oral presentation
131. Difan Zou, Pan Xu, **Quanquan Gu**, Stochastic Gradient Hamiltonian Monte Carlo Methods with Recursive Variance Reduction, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'19**) 32, Vancouver, Canada, 2019.
132. Difan Zou, **Quanquan Gu**, An Improved Analysis of Training Over-parameterized Deep Neural Networks, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'19**) 32, Vancouver, Canada, 2019.
133. Difan Zou\*, Ziniu Hu\*, Yewen Wang, Song Jiang, Yizhou Sun, **Quanquan Gu**, Layer-Dependent Importance Sampling for Training Deep and Large Graph Convolutional Networks, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'19**) 32, Vancouver, Canada, 2019.
134. Yuan Cao, **Quanquan Gu**, Generalization Bounds of Stochastic Gradient Descent for Wide and Deep Neural Networks, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'19**) 32, Vancouver, Canada, 2019. (Spotlight Presentation)
135. Yuan Cao, **Quanquan Gu**, Tight Sample Complexity of Learning One-hidden-layer Convolutional Neural Networks, In Proc. of Advances in Neural Information

- Processing Systems (**NeurIPS'19**) 32, Vancouver, Canada, 2019.
136. Spencer Frei, Yuan Cao, **Quanquan Gu**, Algorithm-Dependent Generalization Bounds for Overparameterized Deep Residual Networks, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'19**) 32, Vancouver, Canada, 2019.
  137. Difan Zou\*, Yuan Cao\*, Dongruo Zhou and **Quanquan Gu**, Gradient Descent Optimizes Over-parameterized Deep ReLU Networks, Machine Learning Journal (**MLJ**), 2019.
  138. Dongruo Zhou, Pan Xu and **Quanquan Gu**, Stochastic Variance-Reduced Cubic Regularized Newton Methods, Journal of Machine Learning Research (**JMLR**), 2019.
  139. Pan Xu, Felicia Gao and **Quanquan Gu**, An Improved Convergence Analysis of Stochastic Variance-Reduced Policy Gradient, in Proc. of the 35th International Conference on Uncertainty in Artificial Intelligence (**UAI'19**), Tel Aviv, Israel, 2019.
  140. Lingxiao Wang and **Quanquan Gu**, Differentially Private Iterative Gradient Hard Thresholding for Sparse Learning, in Proc. of the 28th International Joint Conference on Artificial Intelligence (**IJCAI'19**), Macao, China , 2019.
  141. Dongruo Zhou, **Quanquan Gu**, Lower Bounds for Smooth Nonconvex Finite-Sum Optimization, in Proc. of the 36th International Conference on Machine Learning (**ICML'19**), Long Beach, CA, USA, 2019.
  142. Yisen Wang, Xingjun Ma, James Bailey, Jinfeng Yi, Bowen Zhou, **Quanquan Gu**, On the Convergence and Robustness of Adversarial Training, in Proc. of the 36th International Conference on Machine Learning (**ICML'19**), Long Beach, CA, USA, 2019.
  143. Xiao Zhang\*, Yaodong Yu\*, Lingxiao Wang\* and **Quanquan Gu**, Learning One-hidden-layer ReLU Networks via Gradient Descent, In Proc. of the 22nd International Conference on Artificial Intelligence and Statistics (**AISTATS'19**), Naha, Okinawa, Japan, 2019.
  144. Difan Zou, Pan Xu and **Quanquan Gu**, Sampling from Non-Log-Concave Distributions via Variance-Reduced Gradient Langevin Dynamics, In Proc. of the 22nd International Conference on Artificial Intelligence and Statistics (**AISTATS'19**), Naha, Okinawa, Japan, 2019.
  145. Dongruo Zhou, Pan Xu and **Quanquan Gu**, Stochastic Nested Variance Reduction for Nonconvex Optimization, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'18**) 31, Montreal, Canada, 2018. Spotlight
  146. Pan Xu\*, Jinghui Chen\*, Difan Zou and **Quanquan Gu**, Global Convergence of Langevin Dynamics Based Algorithms for Nonconvex Optimization, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'18**) 31, Montreal, Canada, 2018. Spotlight
  147. Yaodong Yu\*, Pan Xu\* and **Quanquan Gu**, Third-order Smoothness Helps: Even Faster Stochastic Optimization Algorithms for Finding Local Minima, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'18**) 31, Montreal, Canada, 2018.
  148. Bargav Jayaraman, Lingxiao Wang, David Evans and **Quanquan Gu**, Distributed Learning without Distress: Privacy-Preserving Empirical Risk Minimization, In Proc. of Advances in Neural Information Processing Systems (**NeurIPS'18**) 31, Montreal, Canada, 2018.
  149. Yang Wang, **Quanquan Gu** and Donald Brown, Differentially Private Hypothesis Transfer Learning, In Proc. of 28th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML/PKDD'18**), Dublin, Ireland, 2018.
  150. Wenjun Jiang, Qi Li, Lu Su, Chenglin Miao, **Quanquan Gu**, Wenyao Xu: Towards Personalized Learning in Mobile Sensing Systems. In Proc. of 38th IEEE International Conference on Distributed Computing Systems (**ICDCS'18**), Vienna, Austria, pp.321-333, 2018.
  151. Yang Yang, **Quanquan Gu**, Takayo Sasaki, Rachel O'neill, David Gilbert and

- Jian Ma, Continuous-trait probabilistic model for comparing multi-species functional genomic data, *Cell Systems*, in press, 2018.
152. Difan Zou\*, Pan Xu\* and **Quanquan Gu**, Subsampled Stochastic Variance-Reduced Gradient Langevin Dynamics, in Proc. of the 34th International Conference on Uncertainty in Artificial Intelligence (**UAI'18**), Monterey, California, 2018.
  153. Xiao Zhang\*, Simon S. Du\* and **Quanquan Gu**, Fast and Sample Efficient Inductive Matrix Completion via Multi-Phase Procrustes Flow, in Proc. of the 35th International Conference on Machine Learning (**ICML'18**), Stockholm, Sweden, 2018.
  154. Xiao Zhang\*, Lingxiao Wang\*, Yaodong Yu and **Quanquan Gu**, A Primal-Dual Analysis of Global Optimality in Nonconvex Low-Rank Matrix Recovery, in Proc. of the 35th International Conference on Machine Learning (**ICML'18**), Stockholm, Sweden, 2018.
  155. Difan Zou\*, Pan Xu\* and **Quanquan Gu**, Stochastic Variance-Reduced Hamilton Monte Carlo Methods, in Proc. of the 35th International Conference on Machine Learning (**ICML'18**), Stockholm, Sweden, 2018.
  156. Jinghui Chen, Pan Xu, Lingxiao Wang, Jian Ma and **Quanquan Gu**, Covariate Adjusted Precision Matrix Estimation via Nonconvex Optimization, in Proc. of the 35th International Conference on Machine Learning (**ICML'18**), Stockholm, Sweden, 2018.
  157. Pan Xu\* and Tianhao Wang\* and **Quanquan Gu**, Continuous and Discrete-Time Accelerated Stochastic Mirror Descent for Strongly Convex Functions, in Proc. of the 35th International Conference on Machine Learning (**ICML'18**), Stockholm, Sweden, 2018.
  158. Dongruo Zhou, Pan Xu and **Quanquan Gu**, Stochastic Variance-Reduced Cubic Regularized Newton Method, in Proc. of the 35th International Conference on Machine Learning (**ICML'18**), Stockholm, Sweden, 2018.
  159. Pan Xu\* and Tianhao Wang\* and **Quanquan Gu**, Accelerated Stochastic Mirror Descent: From Continuous-time Dynamics to Discrete-time Algorithms, in Proc of the 21st International Conference on Artificial Intelligence and Statistics (**AISTATS'18**), Playa Blanca, Lanzarote, Canary Islands, 2018.
  160. Xiao Zhang\* and Lingxiao Wang\* and **Quanquan Gu**, A Unified Framework for Nonconvex Low-Rank plus Sparse Matrix Recovery, in Proc of the 21st International Conference on Artificial Intelligence and Statistics (**AISTATS'18**), Playa Blanca, Lanzarote, Canary Islands, 2018.
  161. Yang Yang, **Quanquan Gu**, Takayo Sasaki, Rachel O'neill, David Gilbert and Jian Ma, Continuous-trait Probabilistic Model for Comparing Multi-species Functional Genomic Data, in Proc. of the 22nd Annual International Conference on Research in Computational Molecular Biology ( **RECOMB'18**), 2018.
  162. Pan Xu and Jian Ma and **Quanquan Gu**, Speeding Up Latent Variable Gaussian Graphical Model Estimation via Nonconvex Optimization, In Proc. of Advances in Neural Information Processing Systems (**NIPS'17**) 30, Long Beach, CA, USA, 2017.
  163. Jinghui Chen and **Quanquan Gu**, Fast Newton Hard Thresholding Pursuit for Sparsity Constrained Nonconvex Optimization, in Proc of the 23rd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD'17**), Halifax, Nova Scotia, Canada, 2017.
  164. Aditya Chaudhry, Pan Xu and **Quanquan Gu**, Uncertainty Assessment and False Discovery Rate Control in High-Dimensional Granger Causal Inference, in Proc. of the 34th International Conference on Machine Learning (**ICML'17**), Sydney, Australia, 2017.
  165. Rongda Zhu, Lingxiao Wang, Chengxiang Zhai, **Quanquan Gu**, Variance-Reduced Stochastic Gradient High-Dimensional Expectation Maximization Algorithm, in Proc. of the 34th International Conference on Machine Learning (**ICML'17**), Sydney, Australia, 2017.

166. Lingxiao Wang\*, Xiao Zhang\*, **Quanquan Gu**, A Unified Variance-Reduction Based Framework for Nonconvex Low-Rank Matrix Recovery, in Proc. of the 34th International Conference on Machine Learning (**ICML'17**), Sydney, Australia, 2017. (\*: equal contribution)
167. Lingxiao Wang, **Quanquan Gu**, Robust Gaussian Graphical Model Estimation with Arbitrary Corruption, in Proc. of the 34th International Conference on Machine Learning (**ICML'17**), Sydney, Australia, 2017.
168. Lu Tian and **Quanquan Gu**, Communication-efficient Distributed Sparse Linear Discriminant Analysis, in Proc of the 20th International Conference on Artificial Intelligence and Statistics (**AISTATS'17**), Fort Lauderdale, Florida, USA, 2017.
169. Lingxiao Wang\* and Xiao Zhang\* and **Quanquan Gu**, A Unified Computational and Statistical Framework for Nonconvex Low-Rank Matrix Estimation, in Proc of the 20th International Conference on Artificial Intelligence and Statistics (**AISTATS'17**), Fort Lauderdale, Florida, USA, 2017. (\*: equal contribution)
170. Pan Xu and Tingting Zhang and **Quanquan Gu**, Efficient Algorithm for Sparse Tensor-variate Gaussian Graphical Models via Gradient Descent, in Proc of the 20th International Conference on Artificial Intelligence and Statistics (**AISTATS'17**), Fort Lauderdale, Florida, USA, 2017.
171. Dezhi Hong and **Quanquan Gu** and Kamin Whitehouse, High-dimensional Time Series Clustering via Cross-Predictability, in Proc of the 20th International Conference on Artificial Intelligence and Statistics (**AISTATS'17**), Fort Lauderdale, Florida, USA, pp.642-651, 2017.
172. Pan Xu and **Quanquan Gu**, Semiparametric Differential Graph Models, In Proc. of Advances in Neural Information Processing Systems (**NIPS'16**) 29, Barcelona, Spain, pp.1064-1072, 2016.
173. Lu Tian\*, Bargav Jayaraman\*, **Quanquan Gu**, David Evans, Aggregating Private Sparse Learning Models Using Multi-Party Computation, NIPS Workshop on Private Multi-Party Machine Learning, Barcelona, Spain, 2016. (\*: equal contribution)
174. Dechao Tian, **Quanquan Gu**, and Jian Ma. Identifying gene regulatory network rewiring using latent differential graphical models. Nucleic Acids Research, 2016.
175. Jinghui Chen and **Quanquan Gu**, Accelerated Stochastic Block Coordinate Gradient Descent for Sparsity Constrained Nonconvex Optimization, in Proc of the 32th International Conference on Uncertainty in Artificial Intelligence (**UAI'16**), New York/New Jersey, USA, 2016.
176. Lu Tian, Pan Xu and **Quanquan Gu**, Forward Backward Greedy Algorithms for Multi-Task Learning with Faster Rates, in Proc of the 32th International Conference on Uncertainty in Artificial Intelligence (**UAI'16**), New York/New Jersey, USA, 2016.
177. Huan Gui, Jiawei Han and **Quanquan Gu**, Towards Faster Rates and Oracle Property for Low-Rank Matrix Estimation, in Proc. of the 33th International Conference on Machine Learning (**ICML'16**), New York, USA, pp.2300-2309, 2016.
178. Zhaoran Wang, **Quanquan Gu**, and Han Liu, Statistical Limits of Convex Relaxations, in Proc. of the 33th International Conference on Machine Learning (**ICML'16**), New York, USA, pp.1368-1377, 2016.
179. Qingyun Wu, Huazheng Wang, **Quanquan Gu** and Hongning Wang. Contextual Bandits in A Collaborative Environment. The 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR'16**), Pisa, Tuscany, Italy, pp.529-538, 2016.
180. **Quanquan Gu** and Zhaoran Wang and Han Liu, Low-Rank and Sparse Structure Pursuit via Alternating Minimization, in Proc of the 19th International Conference on Artificial Intelligence and Statistics (**AISTATS'16**), Cadiz, Spain, pp.600-609, 2016.
181. Lingxiao Wang, Xiang Ren and **Quanquan Gu**, Precision Matrix Estimation in High Dimensional Gaussian Graphical Models with Faster Rates, in Proc of the 19th International Conference on Artificial Intelligence and Statistics (**AISTATS'16**),

- Cadiz, Spain, pp.177-185, 2016.
182. Renkun Ni and **Quanquan Gu**, Optimal Statistical and Computational Rates for One Bit Matrix Completion, in Proc of the 19th International Conference on Artificial Intelligence and Statistics (**AISTATS'16**), Cadiz, Spain, pp.426-434, 2016.
  183. Zhaoran Wang and **Quanquan Gu** and Han Liu, Sharp Computational-Statistical Phase Transitions via Oracle Computational Model, arXiv:1512.08861, 2015.
  184. Haiping Wang, **Quanquan Gu**, Jia Wei, Zhiwei Cao and Qi Liu, Mining drug–disease relationships as a complement to medical genetics-based drug repositioning: Where a recommendation system meets genome-wide association studies, *Clinical Pharmacology & Therapeutics*, pp.451-454, 2015.
  185. Zhaoran Wang, **Quanquan Gu**, Yang Ning, and Han Liu, High Dimensional Expectation-Maximization Algorithm: Statistical Optimization and Asymptotic Normality, In Proc. of Advances in Neural Information Processing Systems (**NIPS'15**) 28, Montreal, Quebec, Canada, pp. 2521-2529, 2015.
  186. Chang Wan, Ben Kao, Xiao Yu, **Quanquan Gu**, David W.L Cheung, Jiawei Han, Classification with Active Learning and Meta-paths in Heterogeneous Information Networks, In Proc of the 24th ACM Conference on Information and Knowledge Management (**CIKM'15**), Melbourne, VIC, Australia, pp.443-452, 2015.
  187. Shi Zhi and Jiawei Han and **Quanquan Gu**, Robust Classification of Information Networks by Consistent Graph Learning, In Proc. of 25th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML/PKDD'15**), Porto, Portugal, pp.752-767, 2015.
  188. Rongda Zhu and **Quanquan Gu**, Towards a Lower Sample Complexity for Robust One-bit Compressed Sensing, In Proc. of the 32nd International Conference on Machine Learning (**ICML'15**), Lille, France, pp. 739-747, 2015.
  189. Lu An Tang, Xiao Yu, **Quanquan Gu**, Jiawei Han, Guofei Jiang, Alice Leung and Thomas F. La Porta: A Framework of Mining Trajectories from Untrustworthy Data in Cyber-Physical System. *ACM Transactions on Knowledge Discovery from Data* 9(3): 16:1-16:35 (2015)
  190. **Quanquan Gu**, Yuan Cao, Yang Ning and Han Liu, Local and Global Inference for High Dimensional Gaussian Copula Graphical Models, arXiv:1502.02347., 2015.
  191. Jialu Liu, Chi Wang, Jing Gao, **Quanquan Gu**, Charu Aggarwal, Lance Kaplan and Jiawei Han, GIN: A Clustering Model for Capturing Dual Heterogeneity in Networked Data, in Proc. of SIAM Int. Conf. on Data Mining (**SDM'15**), Vancouver, Canada, pp.388-396, 2015.
  192. **Quanquan Gu**, Zhaoran Wang and Han Liu, Sparse PCA with Oracle Property, in Proc. of Advances in Neural Information Processing Systems (**NIPS'14**) 27, Montreal, Quebec, Canada, pp.1529-1537, 2014.
  193. **Quanquan Gu\***, Huan Gui\* and Jiawei Han, Robust Tensor Decomposition with Gross Corruption, in Proc. of Advances in Neural Information Processing Systems (**NIPS'14**) 27, Montreal, Quebec, Canada, pp.1422-1430, 2014. (\*: equal contribution)
  194. **Quanquan Gu**, Tong Zhang and Jiawei Han, Batch-Mode Active Learning via Error Bound Minimization, in Proc. of the 30th Conference on Uncertainty in Artificial Intelligence (**UAI'14**), Quebec City, Quebec, Canada, pp.300-309, 2014.
  195. Yiyi Liu, **Quanquan Gu**, Jack P Hou, Jiawei Han and Jian Ma, A network-assisted co-clustering algorithm to discover cancer subtypes based on gene expression, *BMC Bioinformatics*, 2014.
  196. Xiang Ren, Jialu Liu, Xiao Yu, Urvashi Khandelwal, **Quanquan Gu**, Lidan Wang, Jiawei Han: ClusCite: effective citation recommendation by information network-based clustering. in Proc of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD'14**), New York City, USA, pp. 821-830, 2014.
  197. Xiao Yu, Xiang Ren, Yizhou Sun, **Quanquan Gu**, Bradley Sturt, Urvashi Khan-

- delwal, Brandon Norick and Jiawei Han: Personalized entity recommendation: a heterogeneous information network approach. in Proc. of the 7th ACM International Conference on Web Search and Data Mining (**WSDM'14**), pp.283-292, 2014.
198. **Quanquan Gu**, Charu Aggarwal, Jialu Liu and Jiawei Han, Selective Sampling on Graphs for Classification, in Proc of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD'13**), Chicago, USA, pp.131-139, 2013.
199. Lu-An Tang, Xiao Yu, **Quanquan Gu**, Jiawei Han, Alice Leung and Thomas F. La Porta, Mining lines in the sand: on trajectory discovery from untrustworthy data in cyber-physical system, in Proc of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD'13**), Chicago, USA, pp.410-418, 2013.
200. **Quanquan Gu**, Charu Aggarwal and Jiawei Han, Unsupervised Link Selection in Networks, in Proc of the 16th International Conference on Artificial Intelligence and Statistics (**AISTATS'13**), Scottsdale, AZ, 2013
201. **Quanquan Gu** and Jiawei Han, Clustered Support Vector Machine, in Proc of the 16th International Conference on Artificial Intelligence and Statistics (**AISTATS'13**), Scottsdale, AZ, USA, 2013.
202. Peng Wei, **Quanquan Gu** and Dengfeng Sun, Wireless sensor network data collection by connected cooperative UAVs, in Proc of the 2013 American Control Conference (**ACC'13**), Washington, DC, USA, pp.5911-5916, 2013.
203. Xiao Yu, Xiang Ren, Yizhou Sun, Bradley Sturt, Urvashi Khandelwal, **Quanquan Gu**, Brandon Norick and Jiawei Han, Recommendation in heterogeneous information networks with implicit user feedback. in Proc of the 7th ACM Conference on Recommender Systems (**RecSys'13**), Hong Kong, China, pp.347-350, 2013.
204. Lu-An Tang, Xiao Yu, Sangkyum Kim, **Quanquan Gu**, Jiawei Han, Alice Leung and Thomas La Porta, Trustworthiness Analysis of Sensor Data in Cyber-Physical Systems, Special Issue on Data Warehousing and Knowledge Discovery from Sensors and Streams, Journal of Computer and System Sciences, pp.383-401, 2013.
205. **Quanquan Gu** and Jiawei Han, Towards Active Learning on Graphs: An Error Bound Minimization Approach, in Proc of the 12th IEEE International Conference on Data Mining (**ICDM'12**), Brussels, Belgium, pp.882-887, 2012.
206. **Quanquan Gu**, Tong Zhang, Chris Ding and Jiawei Han, Selective Labeling via Error Bound Minimization, in Proc of Advances in Neural Information Processing Systems (**NIPS'12**) 25, Lake Tahoe, Nevada, United States, pp.332-340, 2012.
207. Xiao Yu, **Quanquan Gu**, Mianwei Zhou and Jiawei Han, Citation Prediction in Heterogeneous Bibliographic Networks, in Proc of the 12th SIAM International Conference on Data Mining (**SDM'12**), Anaheim, CA, USA, pp.1119-1130, 2012.
208. Lu-An Tang, **Quanquan Gu**, Xiao Yu, Jiawei Han, Thomas La Porta, Alice Leung, Tarek F. Abdelzaher and Lance M. Kaplan, IntruMine: Mining Intruders in Untrustworthy Data of Cyber-physical Systems, in Proc of the 12th SIAM International Conference on Data Mining (**SDM'12**), Anaheim, CA, USA, pp.600-611, 2012.
209. Zhijun Yin, Liangliang Cao, **Quanquan Gu**, and Jiawei Han, A Probabilistic Model of Community-Based Latent Topic Analysis, ACM Transactions on Intelligent Systems and Technology (**TIST**), 2012.
210. **Quanquan Gu**, Marina Danilevsky, Zhenhui Li and Jiawei Han, Locality Preserving Feature Learning, in Proc of the 15th International Conference on Artificial Intelligence and Statistics (**AISTATS'12**), La Palma, Canary Islands, pp.477-485, 2012.
211. **Quanquan Gu** and Jiawei Han, Towards Feature Selection in Network, in Proc of the 20th ACM Conference on Information and Knowledge Management (**CIKM'11**), Glasgow, Scotland, UK, pp.1175-1184, 2011.
212. **Quanquan Gu**, Zhenhui Li and Jiawei Han, Correlated Multi-label Feature Selection, in Proc of the 20th ACM Conference on Information and Knowledge Management

- (**CIKM'11**), Glasgow, Scotland, UK, pp.1087-1096, 2011.
213. **Quanquan Gu**, Zhenhui Li and Jiawei Han, Linear Discriminant Dimensionality Reduction, in Proc of the 21st European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML/PKDD'11**), Athens, Greece, pp.549-564, 2011.
  214. **Quanquan Gu**, Zhenhui Li and Jiawei Han, Learning a Kernel for Multi-Task Clustering, in Proc of the 25th AAAI Conference on Artificial Intelligence (**AAAI'11**), San Francisco, California, USA, 2011.
  215. **Quanquan Gu**, Zhenhui Li and Jiawei Han, Generalized Fisher Score for Feature Selection, in Proc of the 27th Conference on Uncertainty in Artificial Intelligence (**UAI'11**), Barcelona, Spain, 2011.
  216. **Quanquan Gu**, Zhenhui Li and Jiawei Han, Joint Feature Selection and Subspace Learning, in Proc of the 22nd International Joint Conference on Artificial Intelligence (**IJCAI'11**), Barcelona, Spain, pp.1294-1299, 2011.
  217. **Quanquan Gu**, Chris Ding and Jiawei Han, On Trivial Solution and Scale Transfer Problems in Graph Regularized NMF, in Proc of the 22nd International Joint Conference on Artificial Intelligence (**IJCAI'11**), Barcelona, Spain, pp.1288-1293, 2011.
  218. Han Hu, **Quanquan Gu** and Jie Zhou, HTF: a novel feature for general crack detection. in Proc. of the 17th IEEE International Conference on Image Processing (**ICIP'10**), Hong Kong, China, pp.1633-1636, 2010.
  219. **Quanquan Gu**, Jie Zhou and Chris Ding, Collaborative Filtering: Weighted Nonnegative Matrix Factorization Incorporating User and Item Graphs, in Proc of the 10th SIAM International Conference on Data Mining (**SDM'10**), Columbus, Ohio, pp.199-210, 2010.
  220. **Quanquan Gu** and Jie Zhou, Co-clustering on Manifolds, in Proc of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD'09**), Paris, France, pp.359-368, 2009.
  221. **Quanquan Gu** and Jie Zhou, Learning the Shared Subspace for Multi-Task Clustering and Transductive Transfer Classification, in Proc of the 9th IEEE International Conference on Data Mining (**ICDM'09**), Miami, Florida, USA, pp.159-168, 2009.
  222. **Quanquan Gu** and Jie Zhou, Local Relevance Weighted Maximum Margin Criterion for Text Classification, in Proc of the 9th SIAM International Conference on Data Mining (**SDM'09**), John Ascuaga's Nugget, Sparks, Nevada, USA, pp.1135-1146, 2009.
  223. **Quanquan Gu** and Jie Zhou, Subspace Maximum Margin Clustering, in Proc of the 18th ACM Conference on Information and Knowledge Management (**CIKM'09**), Hong Kong, China, pp.1337-1346, 2009.
  224. **Quanquan Gu** and Jie Zhou, Local Learning Regularized Nonnegative Matrix Factorization, in Proc of the 21st International Joint Conference on Artificial Intelligence (**IJCAI'09**), Pasadena, California, USA, pp.1046-1051, 2009.
  225. **Quanquan Gu** and Jie Zhou, Transductive Classification via Dual Regularization, in Proc of the 19th European Conference on Machine Learning (**ECML/PKDD'09**), Bled, Slovenia, pp.439-454, 2009.
  226. **Quanquan Gu** and Jie Zhou, Multiple Kernel Maximum Margin Criterion, in Proc. of the 16th IEEE International Conference on Image Processing (**ICIP'09**), Cairo, Egypt, pp.2049-2052, 2009.
  227. **Quanquan Gu** and Jie Zhou, Two Dimensional Nonnegative Matrix Factorization, in Proc. of the 16th IEEE International Conference on Image Processing (**ICIP'09**), Cairo, Egypt, pp.2069-2072, 2009.
  228. **Quanquan Gu** and Jie Zhou, Neighborhood Preserving Nonnegative Matrix Factorization, in Proc. of the 20th British Machine Vision Conference (**BMVC'09**), London, 2009.
  229. Han Hu, **Quanquan Gu**, Lei Deng and Jie Zhou, Multiframe Motion Segmentation



- via Penalized MAP Estimation and Linear Programming, in Proc. of the 20th British Machine Vision Conference (**BMVC'09**), London, 2009.
230. **Quanquan Gu** and Jie Zhou, Regular Simplex Criterion: A Novel Feature Extraction Criterion, in Proc. of the 34th IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP'09**), Taipei, Taiwan, pp.1581-1584, 2009.
231. **Quanquan Gu** and Jie Zhou, Two Dimensional Maximum Margin Criterion, in Proc. of the 34th IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP'09**), Taipei, Taiwan, pp.1621-1624, 2009.
232. **Quanquan Gu** and Jie Zhou, A Similarity Measure under Log-Euclidean Metric for Stereo Matching, in Proc. of the 19th IEEE International Conference on Pattern Recognition (**ICPR'08**), Tampa, Florida, USA, pp.1-4, 2008.
233. **Quanquan Gu** and Jie Zhou, Belief propagation on Riemannian manifold for stereo matching, in Proc. of the 15th IEEE International Conference on Image Processing (**ICIP'08**), San Diego, California, USA, pp.1788-1791, 2008.
234. **Quanquan Gu** and Jie Zhou, A novel similarity measure under Riemannian metric for stereo matching, in Proc. of the 33rd IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP'08**), Las Vegas, Nevada, USA, pp.1073-1076, 2008.
235. Lin Zhu, Jie Zhou, Jingyan Song, Zhenlei Yan and **Quanquan Gu**. A practical algorithm for learning scene information from monocular video. Optics Express, Vol. 16(3), pp. 1448-1459, 2008.

#### INVITED TALKS

1. "Self-Play Preference Optimization for Language Model Alignment", 2nd RL Theory Workshop, June, 2024.
2. "Self-Play Preference Optimization for Language Model Alignment", Workshop on Applied Algorithms for Machine Learning, June, 2024.
3. "Self-Play Preference Optimization for Language Model Alignment", MBZUAI Workshop on Machine Learning for Large Models, June, 2024.
4. "Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models", UCI Workshop on AI and Cloud Computing, May, 2024.
5. "Guided Discrete Diffusion for Electronic Health Record Generation", Generative AI for Healthcare Workshop at UCLA, April, 2024.
6. "Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models", Princeton Language and Intelligence Seminar, March, 2024.
7. "Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models", Annual Conference on Information Sciences and Systems, March, 2024.
8. "Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models", Machine Learning Seminar at the University of Minnesota, March, 2024.
9. "Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models", Information Theory and Applications Workshop, Feb, 2024.
10. "Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models", UBC Centre for Artificial Intelligence Decision-making and Action Seminar, Feb, 2024.
11. "Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models", AAAI 2024 Workshop on AI in Finance for Social Impact, Feb, 2024.
12. "Self-Play Fine-Tuning Converts Weak Language Models to Strong Language Models", Statistics Empowering Data Science Conference at USC, January, 2024.
13. "Why Does Sharpness-Aware Minimization Generalize Better Than SGD?", Google Efficient Machine Learning Workshop, Nov, 2023.
14. "Why Does Sharpness-Aware Minimization Generalize Better Than SGD?", UCSD CSE AI Seminar, Oct, 2023.
15. "Nearly Minimax Optimal Reinforcement Learning for Linear Markov Decision Processes", IEM Distinguished Lecturers Seminar Series at EPFL, Oct, 2023.

16. “Variance-Dependent Regret Bounds for Linear Bandits and Reinforcement Learning: Adaptivity and Computational Efficiency”, 1st RL Theory Workshop, May, 2023.
17. “Towards Understanding the Mixture-of-Experts Layer in Deep Learning”, Future of Large-Scale Machine Learning Workshop at Rice University, April, 2023.
18. “Diffusion Models with Decomposed Priors for Structure-Based Drug Design”, UCLA Level Set Seminar, April, 2023.
19. “Benign Overfitting in Two-layer ReLU Networks”, SILO Seminar at the UW-Madison, March, 2023.
20. “Benign Overfitting in Two-layer Convolutional Neural Networks”, SJTU AI+Math Colloquia, Nov, 2022.
21. “Towards Understanding the Mixture-of-Experts Layer in Deep Learning”, Shanghai AI Lab Seminar, Oct, 2022.
22. “Benign Overfitting in Machine Learning”, Deep Learning Autumn School, Oct, 2022.
23. “Benign Overfitting of Constant-Stepsize SGD for Linear Regression”, SIAM Conference on Mathematics of Data Science, Sep, 2022.
24. “Towards Understanding the Mixture-of-Experts Layer in Deep Learning”, ByteDance AI Lab Seminar, July, 2022.
25. “Nearly Minimax Optimal Reinforcement Learning for Linear Mixture Markov Decision Processes”, From Statistics to Artificial Intelligence Workshop, July, 2022.
26. “Benign Overfitting in Two-layer Convolutional Neural Networks”, ASA Statistical Learning and Data Science Webinar, May, 2022.
27. “Benign Overfitting in Two-layer Convolutional Neural Networks”, Microsoft Research ML Foundations Seminar, May, 2022.
28. “Benign Overfitting in Two-layer Convolutional Neural Networks”, Information Theory and Applications Workshop, May, 2022. (Plenary Talk)
29. “Stochastic Gradient Descent: Benign Overfitting and Implicit Regularization”, Berkeley BLISS Seminar, May, 2022
30. “Epidemic Model Guided Machine Learning for COVID-19 Forecasts”, UCLA ACM AI Seminar, Feb, 2022.
31. “Benign Overfitting of Constant-Stepsize SGD for Linear Regression”, Chinese operations research society online seminar series, Dec, 2021.
32. “Stochastic Gradient Descent: Benign Overfitting and Implicit Regularization”, CMU AI Seminar, Nov, 2021.
33. “Epidemic Model Guided Machine Learning for COVID-19 Forecasts”, ODSC WEST, Nov, 2021
34. “Faster Perturbed Stochastic Gradient Methods for Finding Local Minima”, INFORMS, Oct, 2021.
35. “Benign Overfitting of Constant-Stepsize SGD for Linear Regression”, Google Learning Theory Workshop, Oct, 2021.
36. “Understanding, Improving and Evaluating Adversarial Robustness in Deep Learning”, UCLA CS1 Seminar, Oct, 2021.
37. “On the Convergence of Monte Carlo Methods with Stochastic Gradients”, Simons Institute Workshop on Sampling Algorithms and Geometries on Probability Distributions, Oct, 2021
38. “Stochastic Variance-Reduced High-order Optimization for Nonconvex Optimization”, ICML 2021 Workshop on Beyond first order methods in machine learning systems, July, 2021
39. “Epidemic Model Guided Machine Learning for COVID-19 Forecasts”, ICLR 2021 Workshop on Machine Learning for Preventing and Combating Pandemics, May, 2021
40. “Benign Overfitting: From Minimum-norm Interpolator to Stochastic Gradient Descent”, BAAI Conference, May, 2021.
41. “On the Implicit Bias of Stochastic Gradient Descent with Moderate Learning Rate”, UCSD Halcioğlu Data Science Institute Seminar, April, 2021

42. “Stochastic Variance-Reduced Cubic Regularized Newton Methods for Nonconvex Optimization”, SIAM CSE Mini-symposium on beyond first-order algorithms in modern machine learning systems, March, 2021
43. “Understanding Overparameterized Deep Neural Networks: From Optimization To Generalization”, IJCAI Early Career Talk, January 2021
44. “Epidemic Model Guided Machine Learning for COVID-19 Forecas”, Center for Interdisciplinary Scientific Computation Seminar at Illinois Tech, November, 2020
45. “Learning Wide Neural Networks: Polylogarithmic Over-parameterization and A Mean Field Perspective”, Northwestern University IDEAL Theory of Deep Learning Seminar, October, 2020
46. “Epidemic Model Guided Machine Learning for COVID-19 Forecast”, UCSB Second Annual Responsible Machine Learning Summit, October, 2020
47. “Understanding, Improving and Evaluating Adversarial Robustness in Deep Learning”, Johns Hopkins University Machine Learning Seminar, September, 2020
48. “A Generalized Neural Tangent Kernel Analysis for Two-layer Neural Networks”, Simons Institute Deep Learning Reunion Workshop, August, 2020
49. “Understanding, Improving and Evaluating Adversarial Robustness in Deep Learning”, KDD 2020 Workshop on Adversarial Learning Methods for Machine Learning and Data Mining, August, 2020
50. “Epidemic Model Guided Machine Learning for COVID-19 Forecast”, 2020 KDD Workshop on Applied Data Science for Healthcare, August, 2020
51. “Epidemic Model Guided Machine Learning for COVID-19 Forecast”, D. E. Shaw Technical Talk Forum, June, 2020
52. “Epidemic Model Guided Machine Learning for COVID-19 Forecast”, Institute for Digital Research and Education (IDRE), June, 2020
53. “Epidemic Model Guided Machine Learning for COVID-19 Forecast”, UCLA Computer Science Department Seminar, May, 2020
54. “Epidemic Model Guided Machine Learning for COVID-19 Forecast”, AI for COVID-19 in LA Symposium, May, 2020
55. “Learning Over-parameterized Neural Networks: From Neural Tangent Kernel to Mean-field Analysis”, IPAM Workshop on PDE and Inverse Problem Methods in Machine Learning, April, 2020
56. “Learning Over-parameterized Neural Networks: From Neural Tangent Kernel to Mean-field Analysis”, UCSD AI Seminar, February, 2020
57. “On the Optimization and Generalization of Neural Networks: A Mean-Field Perspective”, Information Theory and Applications Workshop, February, 2020
58. “Towards Understanding Overparameterized Deep Neural Networks: From Optimization To Generalization”, TTIC Workshop on “Recent Trends in Clustering and Classification”, September 2019.
59. “Two facets of stochastic optimization: continuous-time dynamics and discrete-time algorithms”, Workshop on “Interplay between Control, Optimization, and Machine Learnin” at American Control Conference, July 2019.
60. “Towards Understanding Overparameterized Deep Neural Networks: From Optimization To Generalization”, Machine Learning Theory Workshop at Peking University, June 2019.
61. “Towards Understanding Overparameterized Deep Neural Networks: From Optimization To Generalization”, Statistics Department Colloquium, University of California, Los Angeles, April 2019.
62. “New Variance Reduction Algorithms for Nonconvex Finite-Sum Optimization”, Machine Learning Seminar, University of Southern California, Nov 2018.
63. “Closing the Generalization Gap of Adaptive Gradient Methods in Training Deep Neural Networks: Algorithms and Theory”, AI Seminar, USC Information Science Institute, Nov 2018.
64. “The Power and Promise of Nonconvex Optimization for Machine Learning”, School

- of Electrical and Computer Engineering, Cornell University, March 2018.
65. “The Power and Promise of Nonconvex Optimization for Machine Learning”, School of Computational Science and Engineering, Georgia Institute of Technology, March 2018.
  66. “Two Facets of Stochastic Optimization: Continuous-time Dynamics and Discrete-time Algorithms”, Machine Learning Department, Carnegie Mellon University, Sep 2017.
  67. “Blessing of Nonconvexity: Nonconvex Statistical Learning Methods”, University of Virginia, Quantitative Psychology Lecture Series, Feb 2017.
  68. “Blessing of Nonconvexity: Nonconvex Statistical Learning Methods”, Virginia Tech (NVC), CS Seminar, Oct 2016.
  69. “Blessing of Nonconvexity: Nonconvex Statistical Learning Methods”, University of Virginia, SIE Colloquium, Sep 2016.
  70. “Blessing of Nonconvexity: Nonconvex Statistical Learning Methods”, University of Virginia, CS Seminar, Sep 2016.
  71. “Distributed Inference for High Dimensional Semi-parametric Elliptical Graphical Models”, ENAR Spring Meeting, Austin, Texas, March 2016.
  72. “On Longitudinal Gaussian Graphical Models: Estimation and Asymptotic Inference”, ENAR Spring Meeting, Austin, Texas, March 2016
  73. “Asymptotic Inference for High Dimensional Gaussian Copula Graphical Models”, University of Virginia, Statistics Colloquium, Feb 2016.
  74. “Local and Global Inference for High-Dimensional Gaussian Copula Graphical Models”, Joint Statistical Meetings, Seattle, Aug 2015.
  75. “Big Network Analytics: Online and Active Learning Approaches” Michigan State University, CS Colloquium, March 2014.
  76. “Big Network Analytics: Online and Active Learning Approaches”, University of Oregon, CIS Colloquium, March 2014.
  77. “Big Network Analytics: Online and Active Learning Approaches”, University of Utah, CS Colloquium, March 2014.
  78. “Big Network Analytics: Online and Active Learning Approaches”, University of Virginia, SIE Colloquium, Feb 2014.
  79. “Big Network Analytics: Online and Active Learning Approaches”, University of Illinois at Urbana-Champaign, AIIS Seminar, Feb 2014.
  80. “Big Network Analytics: Online and Active Learning Approaches”, University of Illinois at Urbana-Champaign, DAIS Seminar, Feb 2014.
  81. “Selective Labeling via Error Bound Minimization”, University of Illinois at Urbana-Champaign, DAIS Seminar, Oct 2012.

## RESEARCH GRANTS

- PI: National Science Foundation (NSF) (Award #2312094) “Collaborative Research: CPS: Medium: Provably Safe and Robust Multi-Agent Reinforcement Learning with Applications in Urban Air Mobility”, Total award amount: \$399,999.00, Duration: 6/01/23 - 05/31/26, My share: 100%.
- PI: National Science Foundation (NSF) (Award #2323113) “Collaborative Research: Towards the Foundation of Approximate Sampling-Based Exploration in Sequential Decision Making”, Total award amount: \$300,000.00, Duration: 10/01/23 - 09/30/26, My share: 100%.
- Co-PI: National Science Foundation (NSF) (Award #2247426) “Automated Electrochemical Research based on Deep Learning”, Total award amount: \$600,000.00, Duration: 9/01/21 - 08/31/23, My share: 50%.
- PI: National Science Foundation (NSF) (Award #2403400) “Collaborative Research: III: Medium: Neural Networks Meet Bandits: Theory, Algorithms and Applications”, Total award amount: \$600,000.00, Duration: 10/01/24 - 09/30/28, My share: 50%.
- Co-PI: National Science Foundation (NSF) (Award #2140762) “EAGER: ADAPT:

- AI-based Categorization to Decipher Reaction Mechanisms from Cyclic Voltammetry”, Total award amount: \$299,999.00, Duration: 9/01/21 - 08/31/23, My share: 50%.
- Single PI: National Science Foundation (NSF) (Award #2008981) “III: Small: Towards the Foundations of Training Deep Neural Networks: New Theory and Algorithms”, Total award amount: \$500,000.00, Duration: 10/01/20 - 09/30/23.
  - Lead PI: National Science Foundation (NSF) (Award #1911168) “CIF: Small: Collaborative Research: Rank Aggregation with Heterogeneous Information Sources: Efficient Algorithms and Fundamental Limits” with UVa PI Farzad Farnoud, Total award amount: \$500,000.00, Duration: 10/01/2019-9/30/2022, My share: 50%.
  - Lead PI: National Science Foundation (NSF) (Award #1741342/#1855099) “BIG-DATA: F: Collaborative Research: Taming Big Networks via Embedding” with Co-PI Jiawei Han, Total award amount: \$900,000.00, Duration: 01/01/2018-12/31/2021, My share: 55%.
  - PI: National Science Foundation (NSF) (Award #1717950) “SaTC: CORE: Small: Multi-Party High-dimensional Machine Learning with Privacy” with Co-PI David Evans, Total award amount: \$498,624.00, Duration: 08/01/2017-07/31/2020, My share: 50%. (Switched to Co-PI after I moved to UCLA in 2018.)
  - Lead PI: National Science Foundation (NSF) (Award #1717206/#1903202) “III: Small: Collaborative Research: High-Dimensional Machine Learning Methods for Personalized Cancer Genomics” with Co-PI Jian Ma, Total award amount: \$500,000, Duration: 08/01/2017-07/31/2020, My share: 60%.
  - Single PI: National Science Foundation (NSF) (Award #1652539/#1906169) “CA-REER: Scaling Up Knowledge Discovery in High-Dimensional Data via Nonconvex Statistical Optimization”, Total award amount: \$515,835.00, Duration: 2017/8/1-2022/7/31.
  - PI: National Science Foundation (NSF) (Award #1618948/#1904183) “III: Small: Collaborative Learning with Incomplete and Noisy Knowledge” with Co-PI Hongning Wang, Total award amount: \$500,000, Duration: 2016/8/1-2019/7/31, My share: 50%.
  - PI: UVa SEAS Cyber-security Initiative Award “Multi-party Machine Learning with Privacy” with David Evans (Co-PI). Total award amount: \$75,000 (2016-2017). My share: 50%.
  - Co-PI: UVa SEAS Research Innovation Award “Machine Learning in Adversarial Contexts” with David Evans (PI), Homa Alemzadeh, Mohammad Mahmoody, Yanjun Qi. (2017-2018)
  - Co-PI: UVa SEAS Research Innovation Award “Robust Data Fusion in Dynamic Environments” with Farzad Hassanzadeh (PI). (2017-2018)
  - Co-PI: UVa Brain Institute Transformative, Collaborative Neuroscience Pilot Grant “SysNimDB: a public resource for characterizing neuroimmunological cell types and disease states in heterogeneous transcriptomic datasets” with Chris Overall (PI), Yanjun Qi and Abigail Flower. Total award amount: \$75,000 (2017-2018).

#### PROFESSIONAL SERVICES

- Editorial Board Member
  - Artificial Intelligence Journal
  - Information Processing and Management
  - Pattern Recognition and Artificial Intelligence
- Associate Editor/Section Editor/Action Editor
  - Journal of Machine Learning Research, 2024-present
  - Artificial Intelligence Journal, 2023-present
  - Journal of Artificial Intelligence Research, 2022-present
  - PLOS ONE
- Journal Reviewer for
  - IEEE Transaction on Pattern Analysis and Machine Intelligence

- IEEE Transaction on Knowledge and Data Engineering
- IEEE Transactions on Neural Networks and Learning Systems
- IEEE Transaction on Systems, Man and Cybernetics - Part B
- IEEE Transaction on Cybernetics
- ACM Transaction on Knowledge Discovery from Data
- Journal of Machine Learning Research
- Machine Learning
- Statistica Sinica
- Data Mining and Knowledge Discovery
- Information Processing and Management
- Neurocomputing
- Computer Vision and Image Understanding
- Artificial Intelligence
- Knowledge and Information Systems
- PLOS ONE
- Journal of Mathematical Imaging and Vision
- Multimedia Systems
- Computational Intelligence
- PC Chair/PC Vice-Chair for
  - IEEE International Conference on Big Data (ICBD) 2019
- Senior Area Chair
  - International Conference on Machine Learning (ICML) 2024
  - Annual Conference on Neural Information Processing Systems (NeurIPS) 2023, 2024
  - International Conference on Artificial Intelligence and Statistics (AISTATS), 2024
  - AAAI Conference on Artificial Intelligence (AAAI) 2023, 2024
- Area Chair
  - International Conference on Machine Learning (ICML) 2019, 2020, 2021, 2022, 2023
  - Annual Conference on Neural Information Processing Systems (NeurIPS) 2019, 2021, 2022
  - Annual Conference on Learning Theory (COLT) 2023, 2024
  - AAAI Conference on Artificial Intelligence (AAAI) 2020, 2021, 2022
  - International Conference on Artificial Intelligence and Statistics (AISTATS) 2020, 2021, 2022, 2023
  - International Conference on Learning Representations (ICLR) 2020, 2021, 2022, 2023, 2024
  - International Conference on Uncertainty in Artificial Intelligence (UAI) 2022, 2023
  - International Conference on Data Mining (ICDM) 2021, 2024
  - Conference on Language Modeling (COLM) 2024
- Senior PC Member for
  - International Joint Conferences on Artificial Intelligence (IJCAI) 2019, 2020, 2021, 2024
  - Asian Conference on Machine Learning (ACML) 2019
- PC Member for
  - International Conference on Machine Learning (ICML) 2015, 2016, 2017, 2018
  - International Conference on Artificial Intelligence and Statistics (AISTATS) 2015, 2016, 2017, 2018, 2019, 2020
  - International Conference on Uncertainty in Artificial Intelligence (UAI) 2016, 2017, 2018, 2019
  - ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) 2015, 2016, 2017, 2018, 2019
  - World Wide Web (WWW) Conference 2017, 2018

- ACM International Conference on Web Search and Data Mining (WSDM) 2023
- Very Large Data Bases (VLDB) Conferences 2017, 2018, 2019
- SIAM Conference on Data Mining (SDM) 2017, 2018
- International Joint Conferences on Artificial Intelligence (IJCAI) 2013, 2015, 2016, 2017
- AAAI Conference on Artificial Intelligence (AAAI) 2018, 2019
- International Conference on Learning Representations (ICLR) 2018, 2019
- Asian Conference on Machine Learning (ACML) 2012
- Natural Language Processing and Chinese Computing 2015
- Reviewer for
  - Annual Conference on Neural Information Processing Systems (NIPS) 2014, 2015, 2016, 2017, 2018
  - Conference on Learning Theory (COLT) 2018, 2019
  - Conference on Algorithmic Learning Theory 2024
- Section Chair for
  - International Conference on Machine Learning (ICML) 2019, 2021
  - Conference on Learning Theory (COLT) 2021
  - ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) 2016
  - ENAR Spring Meeting 2016
- Grant Proposal Review:
  - NSF III, Panelist, 2017
  - NSF RI, Panelist, 2017
  - NSF III, Panelist, 2018
  - NSF III, Panelist, 2019
  - NSF CIF, SBIR, Panelist, 2020
  - NSF SBIR, Panelist, 2021

## REFERENCES

Available upon request.