ZIHAN ZHANG

zihanz46@uw.edu

RESEARCH INTERESTS

Reinforcement learning theory, bandit learning theory, online PAC learning, multi-distribution learning.

WORKING EXPERIENCE

University of Washington, WA, USA Sept 2024 - Now
Postdoc in Paul G. Allen School of Computer Science & Engineering
Mentors: Prof. Yuxin Chen, Prof. Simon S. Du, and Prof. Jason D. Lee

Princeton University, NJ, USA Apr 2023 - Aug 2024
Postdoc in Department of ECE
Mentors: Prof. Yuxin Chen, Prof. Simon S. Du, and Prof. Jason D. Lee

EDUCATION

Tsinghua University, Beijing, China Aug 2017 - Oct 2022
PhD of Engineering in Control Science and Engineering
Advisor: Prof. Xiangyang Ji

Tsinghua University, Beijing, China Aug 2013 - Jul 2017
Bachelor of Engineering in Automation

SELECTED PUBLICATIONS

Settling the Sample Complexity of Online Reinforcement Learning (*COLT 2024*)
Zihan Zhang, Yuxin Chen, Jason D. Lee, Simon S. Du
We present a new model-basedalgorithm for learning in the fundamental paradigm of reinforcement learning: the tabular Markov Decision Process (MDP) setting. Our algorithm achieves a regret bound of $\min\{\sqrt{SAH^3K}, HK\}$ *where S is the number of states, A is the number of actions, H is the number of horizons and K is the sample of samples. Notably, this is the first-ever result that matches the minimax lower bound for the entire range of sample size $K \geq 1$, settling a 30-year-old open problem.*

Almost Optimal Model-Free Reinforcement Learning via Reference-Advantage Decomposition (*NeurIPS 2020*)
Zihan Zhang, Yuan Zhou, Xiangyang Ji
We propose a novel model-free algorithm with regret bound on the order $\sqrt{SAH^3K}$ for large enough K. The first such bound that matches the information-theoretic lower bound in the large K regime. The reference-advantage technique developed in this paper has been widely used in the literature.

Optimal Multi-Distribution Learning (*COLT 2024*)
Zihan Zhang, Wenhao Zhan, Yuxin Chen, Simon S. Du, Jason D. Lee
We propose a novel algorithm that achieves a sample complexity on the order of $\frac{d+k}{\epsilon^2}$ for the multi-distribution learning problem where d is the VC-dimension, k is the number of distributions, and ϵ is the target accuracy, This bound matches the best-known lower bound, and settles an open problem in COLT 2023.

FULL PUBLICATION LIST

Achieving Tractable Minimax Optimal Regret in Average Reward MDPs
Victor Boone, **Zihan Zhang**.
NeurIPS 2024, 38th Conference on Neural Information Processing Systems

Settling the sample complexity of online reinforcement learning
Zihan Zhang, Yuxin Chen, Jason D. Lee, Simon S. Du.
COLT 2024, 37th Annual Conference on Learning Theory

Optimal Multi-Distribution Learning
Zihan Zhang, Wenhao Zhan, Yuxin Chen, Simon S. Du, Jason D. Lee.
COLT 2024, 37th Annual Conference on Learning Theory

Horizon-Free Regret for Linear Markov Decision Processes
Zihan Zhang, Jason D. Lee, Yuxin Chen, Simon S. Du
ICLR 2024, 12th International Conference on Learning Representations

Sharper Model-free Reinforcement Learning for Average-reward Markov Decision Processes
Zihan Zhang, Qiaomin Xie
COLT 2023, 36th Annual Conference on Learning Theory

Sharp variance-dependent bounds in reinforcement learning: Best of both worlds in stochastic and deterministic environments
Runlong Zhou, **Zihan Zhang**, Simon S. Du.
ICML 2023, Proceedings of the 40th International Conference on Machine Learning

Near-Optimal Regret Bounds for Multi-batch Reinforcement Learning
Zihan Zhang, Yuhang Jiang, Yuan Zhou, Xiangyang Ji.
NeurIPS 2022, 36th Conference on Neural Information Processing Systems

Horizon-Free Reinforcement Learning in Polynomial Time: the Power of Stationary Policies
Zihan Zhang, Xiangyang Ji, Simon S. Du.
COLT 2022, 35th Annual Conference on Learning Theory

Improved Variance-Aware Confidence Sets for Linear Bandits and Linear Mixture MDP
Zihan Zhang*, Jiaqi Yang*, Xiangyang Ji, Simon S. Du.
NeurIPS 2021, 35th Conference on Neural Information Processing Systems (: equal contribution)*

Is Reinforcement Learning More Difficult Than Bandits? A Near-optimal Algorithm Escaping the Curse of Horizon
Zihan Zhang, Xiangyang Ji, Simon S. Du.
COLT 2021, 34th Annual Conference on Learning Theory

Model-Free Reinforement Learning: from Pseudo-Regret to Sample Complexity
Zihan Zhang, Yuan Zhou, Xiangyang Ji.
ICML 2021, Proceedings of the 38th International Conference on Machine Learning

Near Optimal Reward-Free Reinforcement Learning
Zihan Zhang, Simon S. Du, Xiangyang Ji.
ICML 2021, Proceedings of the 38th International Conference on Machine Learning (long talk)

Almost Optimal Model-Free Reinforcement Learning via Reference-Advantage Decomposition
Zihan Zhang , Yuan Zhou, Xiangyang Ji.
NeurIPS 2020, 34th Conference on Neural Information Processing Systems

Regret Minimization for Reinforcement Learning by Evaluating the Optimal Bias Function
Zihan Zhang, Xiangyang Ji.
NeurIPS 2019, 33rd Conference on Neural Information Processing Systems

TEACHING EXPERIENCE

Teaching Assistant: Foundations of Machine Learning (80250943), 2018 Fall and 2019 Spring, Tsinghua University (Instructor: Prof. Xiangyang Ji).

REVIEW SERVICES

I have served as a reviewer for top conferences in the field, including NeurIPS, ICML, ICLR, COLT, AISTATS and ALT.