Zaiyan Xu

Contact Information

Graduate Research Assistant
Department of Electrical and Computer Engineering
Texas A&M University

Email: zxu43@tamu.edu
Webpage: https://www.zaiyanxu.com
GitHub: https://github.com/zaiyan-x

Research Interests

Reinforcement Learning, Distributionally Robust Optimization, Reinforcement Learning from Human Feedback, Large Language Model (LLM) Alignment

Education

Texas A&M University, College Station, TX — Aug. 2020 - Present
Ph.D. in Computer Engineering — Anticipated graduation 05/26
Advisor: Dr. Dileep Kalathil — GPA: 3.86

University of Illinois at Urbana-Champaign, IL — Aug. 2015 - Jul. 2020
B.S. in Statistics & Computer Science and Actuarial Science — GPA: 3.92
Cum Laude, Highest Distinction in CS and Statistics, High Distinction in Actuarial Science

Honors and Achievements

- NeurIPS 2023 Top Reviewer
- Dept. of Electrical and Computer Engineering Graduate Merit Fellowship, TAMU, 2020
- Willis Towers Watson Actuarial Science Scholarship, Dept. of Mathematics, UIUC, 2018

Work Experience

Amazon AGI, New York, NY — Aug. 2025 - Present
Applied Scientist Intern (Host: Dr. Cole Hawkins)
Improving the training stability and scalability of reasoning models.

Mitsubishi Electric Research Laboratories, Cambridge, MA — May. 2023 - Aug 2023
Research Intern (Host: Dr. Mouhacine Benosman)
I worked on developing a distributionally robust reinforcement learning algorithm that also satisfies the the safety constraint. I proposed a Lyaponuv function-based method which characterizes cost functions which satisfy both safety and distributional robustness constraints.

National Center for Supercomputing Application, Champaign, IL — Jun. 2019 - May 2020
Undergraduate Researcher (NCSA SPIN Program)
Worked on speech recognition and auto-captioning with a focus on engineering lectures. Developed several wrappers for CMU Sphinx engine and streamlined model training process by automating audio slicing, caption partitioning.

Publications

Distributionally Robust Large Language Model Finetuning

1. Zaiyan Xu, Sushil Vemuri, Kishan Panaganti, Dileep Kalathil, Rahul Jain, Deepak Ramachandran. "Robust LLM alignment via distributionally robust direct preference optimization", *accepted to NeurIPS 2025*, arXiv:2502.01930, 2025. [ArxiV link].

Communication-efficient Federated Reinforcement Learning

2. Min Cheng, Ruida Zhou, Zaiyan Xu, Chao Tian, P. R. Kumar. "Communication-efficient Federated Natural Policy Gradient for Reinforcement Learning", *under review*.

Sample-efficient Robust Reinforcement Learning

3. Kishan Panaganti, Zaiyan Xu, Dileep Kalathil, Mohammad Ghavamzadeh. "Bridging Distributionally Robust Learning and Offline RL: An Approach to Mitigate Distribution Shift and Partial Data Coverage", in *7th Annual Learning for Dynamics & Control Conference (L4DC)*, 2025. [Publication link].

4. Zaiyan Xu*, Kishan Panaganti*, Dileep Kalathil. "Improved Sample Complexity Bounds For Distributionally Robust Reinforcement Learning", in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2023. [Publication link].

Robust Reinforcement Learning with Neural Network Function Approximation

5. Kishan Panaganti, Zaiyan Xu, Dileep Kalathil, Mohammad Ghavamzadeh. "Robust Reinforcement Learning Using Offline Data", in *Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS)*, 2022. [Publication link].

Sample-efficient Distributionally Robust Imitation Learning

6. Kishan Panaganti*,Zaiyan Xu*, Dileep Kalathil. "Distributionally Robust Behavioral Cloning for Robust Imitation Learning", in *the 62nd IEEE Conference on Decision and Control (CDC)*, 2023. [Publication link].

Reinforcement Learning for Hardward Security

7. Chen Chen, Zaiyan Xu, Mohamadreza Rostami, David Liu, Dileep Kalathil, Ahmad-Reza Sadeghi, Jeyavijayan Rajendran. "ReFuzz: Reusing Tests for Processor Fuzzing with Contextual Bandits", *submitted to NDSS '26*.

(* denotes equal contribution)

Skills

Languages and Platforms: Python, C, C++, R, SQL, PyTorch, Ray, Huggingface TRL, `OpenRLHF`, `verl`, Gymnasium, MuJoCo, CVXPY

Cloud Services: Amazon EC2, Google Computer Enginer (GCE)

Summary of Selected Reserach

Distributionally Robust Large Language Model Finetuning 2024-2025

I contributed to the development of two novel distributionally robust direct preference optimization algorithms, Wasserstein DPO (WDPO) and Kullback-Leibler DPO (KLDPO), to address preference distribution shift in LLM alignment. These methods leverage principled minimax approaches with scalable gradient descent-style learning, ensuring robust performance under shifting user preferences across diverse regions, demographics, and cultural trends. We provided formal sample complexity guarantees and demonstrated substantial alignment improvements in empirical experiments on LLaMA-3.2-1B and LLaMA-3.1-8B models.

Sample-efficient Robust Reinforcement Learning 2023

I contributed to the development of Robust Phased Value Learning (RPVL), a distributionally robust RL algorithm designed for tabular episodic learning and capable of handling mismatch between training and testing environments. Our method achieves an $\tilde{O}(|S||A|H^5)$ sample complexity—improving upon existing bounds by a factor of $|S|$—and supports multiple divergence-based uncertainty sets (including total variation, chi-square, KL, and Wasserstein).

Robust Reinforcement Learning with Neural Network Function Approximation 2022

I implemented a distributionally robust reinforcement learning algorithm to systematically handle uncertainties due to distributional shifts in the MDP transition model between training and testing environments. Designed a neural network architecture using PyTorch to explicitly encode the dual variables arising from the distributionally robust optimization formulation. This architecture was trained simultaneously with the robust Q-function, eliminating the need for external optimization solvers. Empirically validated the approach on MuJoCo benchmarks, demonstrating strong robustness and superior performance of the proposed Robust Fitted Q-Iteration (RFQI) algorithm.

Professional Services

Conference reviewer: AAAI (2025), ICLR (2024, 2025), NeurIPS (**2023 Top Reviewer**, 2024), ICML (2023, 2024, 2025), AISTATS (2023, 2024), American Control Conference (2023), IEEE Conference on Decision and Control (2023, 2024, 2025), L4DC (2023, 2024, 2025)

References

Dr. Dileep Kalathil
Dept. of Electrical and Computer Engineering
Texas A&M University, College Station, TX
Email: dileep.kalathil@tamu.edu