Wenhao Zhan

WORKING

Mosaic AI, Databricks
Jan 2026-, Research Scientist
Reinforcement Learning and Large Language Model Post-training

GenAI, Meta
Jun 2025-Sep 2025, Research Intern
Reinforcement Learning for Tool-Integrated Reasoning Models

Ranking, Meta
May 2024-Oct 2024, Research Intern
Efficient Multi-Agent Offline Reinforcement Learning

EDUCATION

Princeton University
Aug 2021-Dec 2025
Ph.D. Student in Electrical and Computer Engineering
Advisors: Jason D. Lee, Yuxin Chen
- **Academic:** Overall GPA 4.00/4.00.

Princeton University
Aug 2021-May 2023
Master of Arts in Electrical and Computer Engineering
Advisors: Jason D. Lee, Yuxin Chen
- **Academic:** Overall GPA 4.00/4.00.

Tsinghua University
Bachelor of Electronic Engineering   *Aug 2017-Jul 2021*
- **Academic:** Major GPA 3.97/4.00, Overall GPA 3.92/4.00, rank 1/242.

RESEARCH

Reinforcement Learning, Large Language Models, Statistics, Optimization.

PUBLICATIONS

* = equal contributions, + = equal contributions and ordered randomly, # = equal contributions and ordered alphabetically

- K. Brantley, M. Chen#, Z. Gao#, J. D. Lee, W. Sun, **W. Zhan**#, X. Zhang, "Accelerating RL for LLM Reasoning with Optimal Advantage Regression ", 2025. The 39th Conference on Neural Information Processing Systems.

- **W. Zhan**, S. Fujimoto, Z. Zhu, J. D. Lee, D. R. Jiang, Y. Efroni, "Exploiting Structure in Offline Multi-Agent RL: The Benefits of Low Interaction Rank ", 2025. The 13th International Conference on Learning Representations.

Wenhao Zhan

- A. Huang, **W. Zhan**, T. Xie, J. D. Lee, W. Sun, A. Krishnamurthy, D. J. Foster, "Correcting the Mythos of KL-Regularization: Direct Alignment without Overoptimization via Chi-squared Preference Optimization", 2025. The 13th International Conference on Learning Representations, **Spotlight**.

- Z. Gao, **W. Zhan**, J. D. Chang, G. Swamy, K. Brantley, J. D. Lee, W. Sun "Regressing the Relative Future: Efficient Policy Optimization for Multi-turn RLHF", 2025. The 13th International Conference on Learning Representations.

- J. D. Chang*, **W. Zhan***, O. Oertell, K. Brantley, D. Misra, J. D. Lee, W. Sun, "Dataset Reset Policy Optimization for RLHF", 2024. Preprint.

- Z. Gao, J. D. Chang, **W. Zhan**, O. Oertell, G. Swamy, K. Brantley, T. Joachims, J. A. Bagnell, J. D. Lee, W. Sun, "REBEL: Reinforcement Learning via Regressing Relative Rewards", 2024. The 38th Conference on Neural Information Processing Systems.

- Z. Zhang, **W. Zhan**, Y. Chen, S. S. Du, J. D. Lee, "Optimal Multi-Distribution Learning", 2024. The 37th Annual Conference on Learning Theory.

- **W. Zhan**, M. Uehara, W. Sun, J. D. Lee, "Provable Reward-Agnostic Preference-Based Reinforcement Learning", 2024. The 12th International Conference on Learning Representations, **Spotlight**.

- **W. Zhan***, M. Uehara*, N. Kallus, J. D. Lee, W. Sun, "Provable Offline Preference-Based Reinforcement Learning", 2024. The 12th International Conference on Learning Representations, **Spotlight**.

- Y. Zhao[+], **W. Zhan**[+], X. Hu[+], H. Leung, F. Farnia, W. Sun, J. D. Lee, "Provably Efficient CVaR RL in Low-rank MDPs", 2024. The 12th International Conference on Learning Representations.

- G. Li*, **W. Zhan***, J. D. Lee, Y. Chi, Y. Chen, "Reward-agnostic Fine-tuning: Provable Statistical Benefits of Hybrid Reinforcement Learning", 2023. The 37th Conference on Neural Information Processing Systems.

- **W. Zhan***, S. Cen*, B. Huang, Y. Chen, J. D. Lee, Y. Chi, "Policy Mirror Descent for Regularized Reinforcement Learning: A Generalized Framework with Linear Convergence", 2023. SIAM Journal on Optimization.

- **W. Zhan**, M. Uehara, W. Sun, J. D. Lee, "PAC Reinforcement Learning for Predictive State Representations", 2023. The 11th International Conference on Learning Representations.

- **W. Zhan**, J. D. Lee, Z. Yang, "Decentralized Optimistic Hyperpolicy Mirror Descent: Provably No-Regret Learning in Markov Games", 2023. The 11th International Conference on Learning Representations.

- **W. Zhan**, B. Huang, A. Huang, N. Jiang, J. D. Lee, "Offline Reinforcement Learning with Realizability and Single-policy Concentrability", 2022. The 35th Annual Conference on Learning Theory.

- C. Z. Lee, L. P. Barnes, **W. Zhan**, A. Özgür, "Over-the-Air Statistical Estimation of Sparse Models", 2021. The 2021 IEEE Global Communications Conference.

- **W. Zhan**, H. Tang, J. Wang, "Delay Optimal Cross-Layer Scheduling Over Markov Channels with Power Constraint", 2020. The IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2020.

Wenhao Zhan

REVIEWING

- ICML 2022, 2023, 2024
- NeurIPS 2022, 2023, 2024
- ICLR 2024, 2025
- OPT 2022
- Operations Research (INFORMS)
- Mathematical Programming (Springer)
- Machine Learning (Springer)
- SIAM Journal on Mathematics of Data Science

TEACHING

- **Foundations of Reinforcement Learning**
  TA, Spring 2024, Princeton University
- **Special Topics in Information Sciences and Systems: Theory of Deep Weakly Supervised Learning**
  TA, Fall 2022, Princeton University

HONORS

- 2024 Award for Excellence by Princeton SEAS
- Honorable mention for the 2023 Jane Street Graduate Research Fellowship
- 2017-2020 Tsinghua Academic Excellence Award
- 2018-2020 Tsinghua Scientific Research Excellence Award
- 2018-2020 National Encouragement Scholarship

TALKS

- **Optimal Multi-Distribution Learning**
  Adaptive Learning in Complex Environments, TTIC Chicago Summer Workshop 2024