

# Ziniu Li

✉ ziniuli@link.cuhk.edu.cn  
🌐 <http://www.liziniu.org>  
🏠 <https://scholar.google.com/citations?user=8oUnKQQAAAAJ&hl=en>  
📄 <https://www.zhihu.com/people/mo-fei-10-41>  
✂ <https://x.com/ZiniuLi>

## Research

My research focuses on the algorithm design and theoretical analysis of machine learning models, particularly in reinforcement learning. Currently, I am primarily working on large language models to pursue the development of an agent capable of intelligent multi-task performance. My previous research outcomes are highlighted below:

- In the field of large language models, my work spans several key areas: data selection (NeurIPS 2023, Spotlight), diversity-preserving supervised fine-tuning (ICLR 2025), computationally efficient RLHF (ICML 2024), and hallucination mitigation (ICLR2 2025).
- In the field of imitation learning and reinforcement learning, I am interested in the theory of sample complexity (NeurIPS 2020, TPAMI 2021, UAI 2023, Oral), as well as applications in robotics (ICLR 2024 Blog) and signal processing (TSP 2024).
- I also work on optimization-centric topics with other researchers, including understanding Adam in training Transformers (NeurIPS 2024), memory-efficient optimizers (ICLR 2025), zero-order optimization (IJCAI 2020), and prompt-tuning (EMNLP 2024).

## Education

08/2020 - Present    📖 **Ph.D.** at The Chinese University of Hong Kong, Shenzhen  
Advisor: Zhi-Quan (Tom) Luo

08/2015 - 06/2020    📖 **B.E.** at Xi'an Jiaotong University  
Advisor: Zhiyuan Liu.

## Employment History

10/2021 - 08/2022    📖 **Algorithm Engineer (Intern)** at Tencent, Shenzhen  
Advisor: Peilin Zhao

06/2020 - 09/2020    📖 **Algorithm Engineer (Intern)** at Cardinal Operations, Shanghai

07/2019 - 06/2020    📖 **Research Assistant** at Nanjing University, Nanjing  
Advisor: Yang Yu

12/2018 - 02/2019    📖 **Algorithm Engineer (Intern)** at Tianrang Technology Inc., Hangzhou  
Advisor: Zhenhui (Jessie) Li

## Selected Publications

\* indicates equal contribution.

- 1 **Ziniu Li**, Congliang Chen, Tian Xu, Zeyu Qin, Jiancong Xiao, Ruoyu Sun, Zhi-Quan Luo. *Entropic Distribution Matching in Supervised Fine-tuning of LLMs: Less Overfitting and Better Diversity*. In Proceedings of the 13th International Conference on Learning Representations (ICLR), 2025.
  - This paper is selected as an **oral** presentation (acceptance rate < 5%) at NeurIPS FITML Workshop, 2024.

- 2 **Ziniu Li**, Tian Xu, Yushun Zhang, Zhihang Lin, Yang Yu, Ruoyu Sun, Zhi-Quan Luo. *ReMax: A Simple, Effective, and Efficient Reinforcement Learning Method for Aligning Large Language Models*. In Proceedings of the 41st Conference on International Conference on Machine Learning (ICML), 2024.
- 3 **Ziniu Li\***, Tian Xu\*, Yang Yu. *When is RL better than DPO in RLHF? A Representation and Optimization Perspective*. In Proceedings of the 12th International Conference on Learning Representations (ICLR) (Tiny Paper Track), 2024.
  - This paper is selected as an **oral** presentation at ICLR, 2024.
- 4 **Ziniu Li\***, Tian Xu\*, Zeyu Qin, Yang Yu, Zhi-Quan Luo. *Imitation Learning from Imperfection: Theoretical Justifications and Algorithms*. ConfereIn Advances in Neural Information Processing System 37 (NeurIPS), 2023.
  - This paper is selected as an **spotlight** presentation (acceptance rate < 5%) at NeurIPS, 2023.
- 5 Tian Xu\*, **Ziniu Li\***, Yang Yu, Zhi-Quan Luo. *Provably Efficient Adversarial Imitation Learning with Unknown Transitions*. In Proceedings of the 39th Conference on Uncertainty in Artificial Intelligence (UAI), 2023.
  - This paper is selected as an **oral** presentation (acceptance rate < 3%) at UAI, 2023.
- 6 **Ziniu Li**, Yingru Li, Yushun Zhang, Tong Zhang, Zhi-Quan Luo. *HyperDQN: A Randomized Exploration Method for Deep Reinforcement Learning*. In Proceedings of the 10th International Conference on Learning Representations (ICLR), 2022.
  - This paper is selected as an **oral** presentation in Workshop on Ecological Theory of Reinforcement Learning at NeurIPS, 2021.
- 7 Tian Xu, **Ziniu Li**, Yang Yu. *Error Bounds of Imitating Policies and Environments for Reinforcement Learning*. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2021.

## Full Publications

### Journal Articles

- [Fan+24] 1 Youlin Fan, Bo Jiu, Wenqiang Pu, **Ziniu Li**, Kang Li, and Hongwei Liu. “Sensing Jamming Strategy From Limited Observations: An Imitation Learning Perspective”. In: *IEEE Transactions on Signal Processing* (2024).
- [XLY21] 2 Tian Xu, **Ziniu Li**, and Yang Yu. “Error bounds of imitating policies and environments for reinforcement learning”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.10 (2021), pp. 6968–6980.

### Conference Proceedings

- [Li+25] 1 **Ziniu Li**, Congliang Chen, Tian Xu, Zeyu Qin, Jiancong Xiao, Ruoyu Sun, and Zhi-Quan Luo. “Entropic distribution matching in supervised fine-tuning of LLMs: Less overfitting and better diversity”. In: *International Conference on Learning Representations*. 2025.
- [Yan+25] 2 Tianyun Yang, **Ziniu Li**, Juan Cao, and Chang Xu. “Mitigating Hallucination in Large Vision-Language Models via Modular Attribution and Intervention”. In: *International Conference on Learning Representations*. 2025.
- [Zha+25] 3 Yushun Zhang, Congliang Chen, **Ziniu Li**, Tian Ding, Chenwei Wu, et al. “Adam-mini: Use fewer learning rates to gain more”. In: *International Conference on Learning Representations*. 2025.
- [LXY24] 4 **Ziniu Li**, Tian Xu, and Yang Yu. “When is RL better than DPO in RLHF? A Representation and Optimization Perspective”. In: *The Second Tiny Papers Track at ICLR 2024*. 2024.

- [Zha+24a] **5** Heshen Zhan, Congliang Chen, Tian Ding, **Ziniu Li**, and Ruoyu Sun. “Unlocking Black-Box Prompt Tuning Efficiency via Zeroth-Order Optimization”. In: *Findings of the Association for Computational Linguistics: EMNLP 2024*. 2024, pp. 14825–14838.
- [Zha+24b] **6** Yushun Zhang, Congliang Chen, Tian Ding, **Ziniu Li**, Ruoyu Sun, and Zhi-Quan Luo. “Why transformers need adam: A hessian perspective”. In: *Advances in Neural Information Processing Systems* 37. 2024.
- [Li+23b] **7** **Ziniu Li**, Tian Xu, Yushun Zhang, Zhihang Lin, Yang Yu, Ruoyu Sun, and Zhi-Quan Luo. “Remax: A simple, effective, and efficient reinforcement learning method for aligning large language models”. In: *Forty-first International Conference on Machine Learning*. 2023.
- [Xu+23] **8** Tian Xu, **Ziniu Li**, Yang Yu, and Zhi-Quan Luo. “Provably Efficient Adversarial Imitation Learning with Unknown Transitions”. In: *Proceedings of the 39th Conference on Uncertainty in Artificial Intelligence*. 2023, pp. 2367–2378.
- [Li+22] **9** **Ziniu Li**, Tian Xu, Yang Yu, and Zhi-Quan Luo. “Rethinking valuedice-does it really improve performance”. In: *ICLR Blog Track*. 2022.
- [Li+21] **10** **Ziniu Li**, Yingru Li, Yushun Zhang, Tong Zhang, and Zhi-Quan Luo. “HyperDQN: A randomized exploration method for deep reinforcement learning”. In: *International Conference on Learning Representations*. 2021.
- [LC20] **11** **Ziniu Li** and Xiong-Hui Chen. “Efficient Exploration by Novelty-Pursuit”. In: *Proceedings of the 2nd International Conference on Distributed Artificial Intelligence*. 2020, pp. 85–102.
- [LLQ20] **12** Fei-Yu Liu, **Zi-Niu Li**, and Chao Qian. “Self-Guided Evolution Strategies with Historical Estimated Gradients.” In: *International Joint Conferences on Artificial Intelligence*. 2020, pp. 1474–1480.
- [XLY20] **13** Tian Xu, **Ziniu Li**, and Yang Yu. “Error Bounds of Imitating Policies and Environments”. In: *Advances in Neural Information Processing Systems* 33. 2020, pp. 15737–15749.

## arXiv Preprints

- [Tan+25] **1** Zhengyang Tang, **Ziniu Li**, Zhenyang Xiao, Tian Ding, Ruoyu Sun, et al. *Enabling Scalable Oversight via Self-Evolving Critic*. In: *arXiv preprint arXiv:2501.05727* (2025).
- [Jia+24] **2** Chengxing Jia, Pengyuan Wang, **Ziniu Li**, Yi-Chen Li, Zhilong Zhang, Nan Tang, and Yang Yu. *BWArea Model: Learning World Model, Inverse Dynamics, and Policy for Controllable Language Generation*. In: *arXiv preprint arXiv:2405.17039* (2024).
- [Xia+24] **3** Jiancong Xiao, **Ziniu Li**, Xingyu Xie, Emily Getzen, Cong Fang, Qi Long, and Weijie J. Su. *On the Algorithmic Bias of Aligning Large Language Models with RLHF: Preference Collapse and Matching Regularization*. In: *arXiv preprint arXiv:2405.16455* (2024).
- [Li+23a] **4** **Ziniu Li**, Ke Xu, Liu Liu, Lanqing Li, Deheng Ye, and Peilin Zhao. *Deploying Offline Reinforcement Learning with Human Feedback*. In: *arXiv preprint arXiv:2303.07046* (2023).
- [LXY22] **5** **Ziniu Li**, Tian Xu, and Yang Yu. *A Note on Target Q-learning For Solving Finite MDPs with A Generative Oracle*. In: *arXiv preprint arXiv:2203.11489* (2022).
- [Xu+22] **6** Tian Xu, **Ziniu Li**, Yang Yu, and Zhi-Quan Luo. *Understanding Adversarial Imitation Learning in Small Sample Regime: A Stage-coupled Analysis*. In: *arXiv preprint arXiv:2208.01899* (2022).

## Service

- Organizers: Reinforcement Learning Weekly Seminar, Deep Learning Seminar at CUHKSZ.

- T-PAMI, NeurIPS'2024, ICML'2024, ICLR'2024, NeurIPS'2023, ICML'2023, NeurIPS'2022 (Top Reviewer), ICML'2022 (Outstanding Reviewer), ICLR'2022 (Highlighted Reviewer), DAI'2020.
- Technical Program Committee Members: RL4RealLife@NeurIPS2022, DAI'2022.

## Awards

---

- |      |   |   |
|------|---|---|
| 2021 | 📖 | Best paper presentation award at the Doctoral and Postdoctoral Forum of Shenzhen Research Institute of Big Data |
| 2017 | 📖 | Honor student of Xi'an Jiaotong University  |
|      | 📖 | Second prize of National Electrical Math Modeling Competition   |
| 2016 | 📖 | Second prize of Chinese Mathematics Competition, Shaanxi Province   |