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

Education

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

08/2015 – 06/2019 — **B.E.** at Xi'an Jiaotong University
Major: Electrical Engineering
Advisor: Zhiyuan Liu

Employment History

05/2025 – Present — **Algorithm Engineer (Intern)** at Bytedance Seed (Horizon), Beijing
Advisor: Ge Zhang
Outcome: contributed to Seed-OSS-36B model; co-authored 2 research papers on online reinforcement learning; developing large-scale, high-quality chain-of-thought reasoning datasets (ongoing)

10/2021 – 08/2022 — **Algorithm Engineer (Intern)** at Tencent AI-Lab, Shenzhen
Advisor: Peilin Zhao
Outcome: secured 2 patents for offline reinforcement learning algorithms; first-authored research paper on offline-to-online RL adaption

06/2020 – 09/2020 — **Algorithm Engineer (Intern)** at Cardinal Operations, Shanghai
Outcome: contributed to dynamic pricing optimization project

07/2019 – 06/2020 — **Research Assistant** at Nanjing University, Nanjing
Advisor: Yang Yu
Outcome: co-authored theory paper on imitation learning; first-authored research paper on exploration strategies in reinforcement learning

12/2018 – 02/2019 — **Algorithm Engineer (Intern)** at Tianrang Technology Inc., Hangzhou
Advisor: Zhenhui (Jessie) Li
Outcome: contributed to reinforcement learning applications in traffic control systems

Research

My research focuses on the **algorithm design** and **theoretical foundation** of **machine learning** models, particularly in (large-scale) **reinforcement learning**. Currently, I am primarily working on **large language models** to pursue the development of an agent capable of universal intelligence.

My passion is to unlock the full potential of reinforcement learning through the integration of training stability, system efficiency, and data design, thereby expanding the frontiers of its capabilities.

I have published 10+ papers in top-tier machine learning and artificial intelligence conferences, including **ICML**, **NeurIPS**, and **ICLR**, as well as prestigious journals such as **IEEE TPAMI** (the most influential AI journal), **IEEE TSP** (a leading signal processing journal), and **JASA** (a premier statistics journal).

My work has received several honors, including **Best Paper Runner-Up** (NeurIPS 2024 FITML Workshop), **Oral Presentations** (ICLR 2024 Tiny Paper Track, UAI 2023, NeurIPS 2021 EcoRL Workshop), and **Spotlight Presentation** (NeurIPS 2023). Below, I highlight key contributions from my previous research:

- 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, NeurIPS 2024 FITML Workshop Best Paper Runner-up), generalization of RLHF (ICLR2024 Tiny Paper Oral), foundation of RL in LLM (ICML 2024), scalable oversight (COLM 2025), latent action model (ICML 2025), 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), efficient exploration (ICLR 2022, NeurIPS 2021 EcoRL Workshop Oral, DAI 2020), 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 Finding).

Selected Publications

* indicates equal contribution.

1. **Ziniu Li**, Congliang Chen, Tianyun Yang, Tian Ding, Ruoyu Sun, Ge Zhang, Wenhao Huang, Zhi-Quan Luo. *Knapsack RL: Unlocking Exploration of LLMs via Optimizing Budget Allocation*. arXiv preprint: 2509.25849, 2025.

2. Chengpeng Li*, Zhengyang Tang*, **Ziniu Li***, Mingfeng Xue, Keqin Bao, Tian Ding, Ruoyu Sun, Benyou Wang, Xiang Wang, Junyang Lin, Dayiheng Liu. *CoRT: Code-integrated Reasoning within Thinking.* Advances in Neural Information Processing Systems (NeurIPS), 2025.

3. Zhengyang Tang*, **Ziniu Li***, Zhenyang Xiao*, Tian Ding, Ruoyu Sun, Benyou Wang, Dayiheng Liu, Fei Huang, Tianyu Liu, Bowen Yu, Junyang Lin. *Self-Evolving Critique Abilities in Large Language Models.* In Proceedings of the 2nd Conference on Language Modeling (COLM), 2025.

4. **Ziniu Li**, Congliang Chen, Tian Xu, Zeyu Qin, Jiancong Xiao, Zhi-Quan Luo, Ruoyu Sun. *Preserving Diversity in Supervised Fine-tuning of Large Language Models.* In Proceedings of the 13th International Conference on Learning Representations (ICLR), 2025.
    - This paper is selected as an **best paper runner-up** at NeurIPS FITML Workshop, 2024.

5. **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.
    - This paper builds the foundation of REINFORCE-based algorithms for LLMs. Subsequent methods, such as GRPO, fall under our framework.

6. **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.

7. **Ziniu Li***, Tian Xu*, Zeyu Qin, Yang Yu, Zhi-Quan Luo. *Imitation Learning from Imperfection: Theoretical Justifications and Algorithms.* Advances in on Neural Information Processing System 37 (NeurIPS), 2023.
    - This paper is selected as an **spotlight** presentation (acceptance rate $< 5\%$) at NeurIPS, 2023.

8. 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.

9 **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.

Full Publications

Journal Articles

[Xia+25] 1 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: *Journal of the American Statistical Association* (2025).

[Fan+24] 2 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] 3 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

[Jia+25] 1 Chengxing Jia, **Ziniu Li**, Pengyuan Wang, Yi-Chen Li, Zhenyu Hou, Yuxiao Dong, and Yang Yu. "Controlling Large Language Model with Latent Actions". In: *International Conference on Machine Learning*. 2025.

[Li+25a] 2 Chengpeng Li, Zhengyang Tang, **Ziniu Li**, Mingfeng Xue, Keqin Bao, et al. "CoRT: Code-integrated Reasoning within Thinking". In: *Advances in Neural Information Processing Systems*. 2025.

[Li+25c] 3 **Ziniu Li**, Congliang Chen, Tian Xu, Zeyu Qin, Jiancong Xiao, Zhi-Quan Luo, and Ruoyu Sun. "Preserving Diversity in Supervised Fine-tuning of Large Langguage Models". In: *International Conference on Learning Representations*. 2025.

[Tan+25b] 4 Zhengyang Tang, **Ziniu Li**, Zhenyang Xiao, Tian Ding, Ruoyu Sun, et al. "Self-Evolving Critique Abilities in Large Language Models". In: *Conference on Language Modeling*. 2025.

[Yan+25b] 5 Tianyun Yang, **Ziniu Li**, Juan Cao, and Chang Xu. "Understanding and Mitigating Hallucination in Large Vision-Language Models via Modular Attribution and Intervention". In: *International Conference on Learning Representations*. 2025.

[Zha+25a] 6 Xueyao Zhang, Yuancheng Wang, Chaoren Wang, **Ziniu Li**, Zhuo Chen, and Zhizheng Wu. "Advancing Zero-shot Text-to-Speech Intelligibility across Diverse Domains via Preference Alignment". In: *The 63rd Annual Meeting of the Association for Computational Linguistics*. 2025.

[Zha+25b] 7 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] 8 **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 of International Conference on Learning Representations*. 2024.

[Yan+24] 9 Tianyun Yang, **Ziniu Li**, Juan Cao, and Chang Xu. "Pruning for Robust Concept Erasing in Diffusion Models". In: *Neurips Safe Generative AI Workshop 2024*. 2024.

[Zha+24a] 10 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] 11 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] 12 **Ziniu Li**, Tian Xu, Zeyu Qin, Yang Yu, and Zhi-Quan Luo. "Imitation learning from imperfection: Theoretical justifications and algorithms". In: *Advances in Neural Information Processing Systems 36*. 2023.

[Li+23c] 13 **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: *International Conference on Machine Learning*. 2023.

[Xu+23] 14 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] 15 **Ziniu Li**, Tian Xu, Yang Yu, and Zhi-Quan Luo. "Rethinking valuedice-does it really improve performance". In: *ICLR Blog Track*. 2022.

[Li+21] 16 **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] 17 **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] 18 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] 19 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

[Che+25] 1 Peter Chen, Xiaopeng Li, **Ziniu Li**, Xi Chen, and Tianyi Lin. *Spectral Policy Optimization: Coloring your Incorrect Reasoning in GRPO*. In: *arXiv preprint arXiv:2505.11595* (2025).

[Li+25b] 2 Yizhi Li, Qingshui Gu, Zhoufutu Wen, **Ziniu Li**, Tianshun Xing, et al. *TreePO: Bridging the Gap of Policy Optimization and Efficacy and Inference Efficiency with Heuristic Tree-based Modeling*. In: *arXiv preprint arXiv:2508.17445* (2025).

[Li+25d] 3 **Ziniu Li**, Congliang Chen, Tianyun Yang, Tian Ding, Ruoyu Sun, Ge Zhang, Wenhao Huang, and Zhi-Quan Luo. *Knapsack RL: Unlocking Exploration of LLMs via Optimizing Budget Allocation*. In: *arXiv preprint arXiv:2509.25849* (2025).

[Tan+25a] 4 Zhengyang Tang, **Ziniu Li**, Zhenyang Xiao, Tian Ding, Ruoyu Sun, et al. *RealCritic: Towards Effectiveness-Driven Evaluation of Language Model Critiques*. In: *arXiv preprint arXiv:2501.14492* (2025).

[Wan+25] 5 Ren-Jian Wang, Ke Xue, Zeyu Qin, **Ziniu Li**, Sheng Tang, Hao-Tian Li, Shengcai Liu, and Chao Qian. *Quality-Diversity Red-Teaming: Automated Generation of High-Quality and Diverse Attackers for Large Language Models*. In: *arXiv preprint arXiv:2506.07121* (2025).

[Yan+25a] 6 Tianyun Yang, Yunwen Li, **Ziniu Li**, Zhihang Lin, Ruoyu Sun, and Tian Ding. *Bridging Formal Language with Chain-of-Thought Reasoning to Geometry Problem Solving*. In: *arXiv preprint arXiv:2508.09099* (2025).

[Li+23a] 7 **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] 8 **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] 9 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.
- ICML'2025 (Outstanding Reviewer), 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

2024 Best paper runner-up award at NeurIPS 2024 Workshop on Fine-Tuning in Modern Machine Learning: Principles and Scalability

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