Zhihan Xiong

✉ zhihanx@cs.washington.edu
https://homes.cs.washington.edu/~zhihanx/

Research Interests

Reinforcement Learning, Bandit Problems, Large Language Models

Education

University of Washington, Seattle, WA — *Sep 2020–Aug 2025*
Ph.D. in Computer Science & Engineering, advisor: **Maryam Fazel** — **GPA**: 4.0/4.0

Stanford University, Stanford, CA — *Sep 2018–Jun 2020*
M.S. in Statistics — **GPA**: 4.0/4.0

University of Illinois at Urbana-Champaign, Champaign, IL — *Aug 2014–May 2018*
B.S. in Mathematics (**Summa Cum Laude and with Highest Distinction**) — **GPA**: 4.0/4.0
B.S. in Engineering Physics (**with Highest Honors**)
Minors in Computer Science and Statistics

Work/Research Experience

Meta — **Seattle, WA**
Visiting Researcher, FAIR Labs — *Oct 2022–Sep 2024*
supervised by Dr. **Lin Xiao**

- Proposed a primal-dual-based policy optimization framework that not only enjoys linear convergence in continuous state-action space, but also incorporates popular practical methods as special cases.

Bytedance — **Seattle, WA**
Research Scientist Intern, Applied Machine Learning (AML) Group — *Jun 2021–Sep 2021*
supervised by Dr. **Yingxiang Yang** and Dr. **Chong Wang**

- Developed Fourier learning method that utilizes periodic pattern in online streaming data distribution for large-scale recommendation systems.
- Published "*Fourier Learning with Cyclical Data*" at ICML 2022.

Zillow — **Seattle, WA**
Applied Scientist Intern, Personalization Team — *Jun 2019–Sep 2019*
supervised by Dr. **Luca Cazzanti**

- Participated in designing contextual bandit tree model for in-app personalized recommendation.
- Developed a computationally efficient method for off-policy evaluation.

Publications/Preprints (Google Scholar)

(* indicates equal contributions)

Hybrid Preference Optimization for Alignment: Faster Convergence Rates by Combining Offline Preferences with Online Exploration
- Avinandan Bose, **Zhihan Xiong**, Aadirupa Saha, Simon S. Du, Maryam Fazel
- *In submission.* [arXiv]

Language Model Preference Evaluation with Multiple Weak Evaluators
- Zhengyu Hu, Jieyu Zhang, **Zhihan Xiong**, Alexander Ratner, Hui Xiong, Ranjay Krishna
- *In submission.* [arXiv]

Dual Approximation Policy Optimization
- **Zhihan Xiong**, Maryam Fazel, Lin Xiao
- *Workshop on Aligning Reinforcement Learning Experimentalists and Theorists, ICML 2024.* [arXiv]
- *In submission to JMLR.*

LoRe: Personalizing LLMs via Low-Rank Reward Modeling
- Avinandan Bose, **Zhihan Xiong**, Yuejie Chi, Simon S. Du, Lin Xiao, Maryam Fazel
- *Conference on Language Modeling (COLM), 2025.* [arXiv] [Tweet]

A/B Testing and Best-arm Identification for Linear Bandits with Robustness to Non-stationarity
- **Zhihan Xiong***, Romain Camilleri*, Maryam Fazel, Lalit Jain, Kevin Jamieson
- *International Conference on Artificial Intelligence and Statistics (AISTATS), 2024.* [paper]
- *Conference on Digital Experimentation @ MIT (CODE@MIT), 2023*

A Black-box Approach for Non-stationary Multi-agent Reinforcement Learning
- Haozhe Jiang, Qiwen Cui, **Zhihan Xiong**, Maryam Fazel, Simon S. Du
- *International Conference on Learning Representations (ICLR), 2024.* [arXiv]

Offline Congestion Games: How Feedback Type Affects Data Coverage Requirement
- Haozhe Jiang*, Qiwen Cui*, **Zhihan Xiong**, Maryam Fazel, Simon S. Du
- *International Conference on Learning Representations (ICLR), 2023.* [paper]

Learning in Congestion Games with Bandit Feedback
- Qiwen Cui*, **Zhihan Xiong***, Maryam Fazel, Simon S. Du
- *Advances in Neural Information Processing Systems (NeurIPS), 2022.* [paper]

Near-Optimal Randomized Exploration for Tabular Markov Decision Processes
- **Zhihan Xiong***, Ruoqi Shen*, Qiwen, Cui*, Maryam Fazel, Simon S. Du
- *Advances in Neural Information Processing Systems (NeurIPS), 2022.* [paper]

Fourier Learning with Cyclical Data
- Yingxiang Yang*, **Zhihan Xiong***, Tianyi Liu*, Taiqing Wang, Chong Wang
- *International Conference on Machine Learning (ICML), 2022.* [paper]

Selective Sampling for Online Best-arm Identification

- Romain Camilleri*, **Zhihan Xiong***, Maryam Fazel, Lalit Jain, Kevin Jamieson
- *Advances in Neural Information Processing Systems* (*NeurIPS*), *2021.* [paper]

Parameterized Indexed Value Function for Efficient Exploration in Reinforcement Learning
- Tian Tan*, **Zhihan Xiong***, Vikranth R. Dwaracherla
- *Association for the Advancement of Artificial Intelligence* (*AAAI, Oral*), *2020.* [arXiv]

Teaching Experience

Graduate Teaching Assistant — **University of Washington, WA**
CSE 541: Interactive Learning — *Spring 2025*
EE/CSE 578: Convex Optimization — *Winter 2025*

Graduate Course Assistant — **Stanford University, CA**
CS 229: Machine Learning — *Spring 2020*
CS 234: Reinforcement Learning — *Winter 2020*
CS 229: Machine Learning — *Autumn 2019*

Awards & Honors

Meta AI Mentorship Program (2-year Funding) — *2022-2024*
NeurIPS 2022 Top Reviewers — *Autumn 2022*
ICML 2022 Travel Award — *Summer 2022*
IFDS Research Assistantship — *Winter 2021, Summer 2025*
Yee Seung Ng Award — *Spring 2017*

Professional Services

- Reviewer of ICML (2021, 2022, 2023, 2024), NeurIPS (2021, 2022, 2023) and ICLR (2022, 2023, 2024).
- UW CSE PhD Application Reviewer of 2021, 2022, 2023.

Skills

Programming: Python (NumPy and PyTorch), LaTeX, C++
Languages: Chinese (native), English (professional proficiency)