HUI YUAN

huiyuan@princeton.edu | Google Scholar

EDUCATION

Princeton University 2020 - 2025 (Expected)
Ph.D. in Electrical and Computer Engineering
- Adviser: Prof. Mengdi Wang
- GPA to date: 3.96/4.00

University of Science and Technology of China 2016 - 2020
B.S. in Statistics

AWARDS AND HONORS

- **Anthony Ephremides Fellowship in Electrical Engineering** 2021
  Princeton University
- **Best Undergraduate Thesis in Statistics** 2020
  University of Science and Technology of China

RESEARCH INTERESTS

General Interest Foundations of machine learning methods (recent focus: generative models) and their applications to real-world decision making.

Overarching Goal
- Provide theoretical understating for machine learning methods
- Design better algorithms in a principle and math-backed way

PREPRINTS*

[1] **A Common Pitfall of Margin-based Language Model Alignment: Gradient Entanglement**
H Yuan*, Y Zeng*, Y Wu*, H Wang, M Wang, L Liu*. (* *leading contributors*)
arXiv: 2410.13828

[2] **Training Guided Diffusion Solver for Combinatorial Optimization with Imperfect Data**
In submission to ICLR 2025.

PUBLICATIONS

Generative Models

[1] **Gradient Guidance for Diffusion Models: An Optimization Perspective**
Y Guo*, **H Yuan***, Y Yang, M Chen, M Wang. (* *equal contribution*)
Conference on Neural Information Processing Systems (NeurIPS) 2024.

[2] **MaxMin-RLHF: Towards Equitable Alignment of Large Language Models with Diverse Human Preferences**
S Chakraborty, J Qiu, **H Yuan**, A Koppel, D Manocha, F Huang, A Bedi, M Wang.

* Titles for anonymous submissions are altered to reflect the topics only.

The Forty-First International Conference on Machine Learning (ICML) 2024.

[3] **Reward-Directed Conditional Diffusion: Provable Distribution Estimation and Reward Improvement**
H Yuan, K Huang, C Ni, M Chen, M Wang.
Conference on Neural Information Processing Systems (NeurIPS) 2023.

Decision Making

[1] **Conversational Dueling Bandits in Generalized Linear Models**
S Yang, **H Yuan**, X Zhang, M Wang, H Zhang, H Wang.
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2024.

[2] **Adversarial Attacks to Online Learning to Rank**
Z Wang, R Balasubramanian, **H Yuan**, C Song, M Wang, H Wang.
arXiv:2305.19218

[3] **Tree Search-Based Evolutionary Bandits for Protein Sequence Optimization**
J Qiu*, **H Yuan***, J Zhang*, W Chen, H Wang, M Wang *(* equal contribution).*
AAAI Conference on Artificial Intelligence (AAAI) 2024.

[4] **Unified Off-Policy Learning to Rank: A Reinforcement Learning Perspective**
Z Zhang, Y Su, **H Yuan**, Y Wu, R Balasubramanian, Q Wu, H Wang, M Wang.
Conference on Neural Information Processing Systems (NeurIPS) 2023.

[5] **Bandit Theory and Thompson Sampling-Guided Directed Evolution for Sequence Optimization**
H Yuan, H Wang, C Ni, X Zhang, L Cong, C Szepesvári, M Wang.
Conference on Neural Information Processing Systems (NeurIPS) 2022.

Statistical Machine Learning

[1] **Learning Entangled Single-Sample Gaussians in the Subset-of-Signals Model**
Y Liang, **H Yuan.**
Annual Conference on Learning Theory (COLT) 2020.

[2] **Learning Entangled Single-Sample Distributions via Iterative Trimming**
H Yuan, Y Liang.
International Conference on Artificial Intelligence and Statistics (AISTATS) 2020.

[3] **Uniform Joint Screening for Ultra-High Dimensional Graphical Models**
Z Zheng, H Shi, Y Li, **H Yuan.**
Journal of Multivariate Analysis 179 (2020): 104645.

TEACHING AND MENTORING

- **Teaching Assistant |** ECE 535 Machine Learning and Pattern Recognition, Fall 2022, Princeton
- **Research Mentor for Junior Graduate Students** | Summer 2023, Summer 2022, Princeton

PROFESSIONAL SERVICES

- **Reviewer |** NeurIPS 2023, ICML 2023, ICLR 2023, Operation Research
- **Conference Session Organizer |** CISS 2022
- **Committee Member |** SWE (Society of Women Engineers) @ Princeton