

# YIFAN HAO

✉ [yifanh12@illinois.edu](mailto:yifanh12@illinois.edu) ✉ [Homepage](#)  
🏛 [University of Illinois Urbana-Champaign](#)

## RESEARCH INTERESTS

My research interests lie at the **intersection of machine learning and statistics**. I am particularly interested in **developing efficient algorithms** for modern machine learning systems (such as large language models) and in exploring the **theoretical foundations** that explain their effectiveness.

## EDUCATION

<b>University of Illinois Urbana-Champaign</b> Ph.D student, Computer Science Advisor: Prof. <a href="#">Tong Zhang</a>	2024.9 – present
<b>The Hong Kong University of Science and Technology</b> MPhil, Mathematics Advisor: Prof. <a href="#">Tong Zhang</a>	2022.9 – 2024.8
<b>University of Science and Technology of China</b> Bachelor of Science, School of the Gifted Young Ranking: 5/91 in Statistics	2018.9 – 2022.6

## RESEARCH EXPERIENCE

<b>Ph.D. Student, University of Illinois Urbana-Champaign</b> Advisor: Prof. Tong Zhang	2024.9 – present
<ul style="list-style-type: none"><li>Proposed the <b>Distributional Input Projection Network (DIPNeT)</b>, which demonstrates strong performance across diverse scenarios, including reasoning tasks and adversarial robustness.</li><li>Developed <b>GVM-RAFT</b>, a prompt-specific dynamic sample allocation strategy designed to minimize stochastic gradient variance under computational budget constraints, achieving a 2-4× speedup and significant accuracy improvements over vanilla RAFT.</li><li>Identified and analyzed the <b>feature decoupling phenomenon</b> in Transformers, and established theoretical results on their expressiveness for modeling hidden Markov processes.</li><li>Provided a theoretical explanation for the benefits of <b>ensemble methods</b> in supervised fine-tuning, attributing their effectiveness to reduced overadaptation.</li></ul>	
<b>Master, The Hong Kong University of Science and Technology</b> Advisor: Prof. Tong Zhang	2022.9 – 2024.8
<ul style="list-style-type: none"><li>Demonstrated the <b>harmful effects of benign overfitting</b> on adversarial robustness in overparameterized models, and characterized the trade-off between standard and adversarial risks.</li></ul>	

- In the out-of-distribution (OOD) regime, provided theoretical insights into how **spurious feature diversification and overparameterization** improve OOD generalization.

## PUBLICATIONS AND PREPRINTS

---

### Publications

- **Optimizing Chain-of-Thought Reasoners via Gradient Variance Minimization in Rejection Sampling and RL.** [[NeurIPS 2025](#)]  
Jiarui Yao\*, [Yifan Hao](#)\*, Hanning Zhang, Hanze Dong, Wei Xiong, Nan Jiang, Tong Zhang
- **Understanding Overadaptation in Supervised Fine-Tuning: The Role of Ensemble Methods.** [[ICML 2025](#)]  
[Yifan Hao](#)\*, Xingyuan Pan\*, Hanning Zhang\*, Chenlu Ye, Rui Pan, Tong Zhang
- **Spurious feature diversification improves out-of-distribution generalization.** [[ICLR 2024](#)]  
Yong Lin\*, Lu Tan\*, [Yifan Hao](#)\*, Honam Wong, Hanze Dong, Weizhong Zhang, Yujia Yang and Tong Zhang
- **Monte Carlo Sampling without Isoperimetry: A Reverse Diffusion Approach.** [[ICLR 2024](#)]  
Xunpeng Huang\*, Hanze Dong\*, [Yifan Hao](#), Yi-An Ma and Tong Zhang.
- **SFQRA: Scaled Factor-augmented Quantile Regression with Aggregation in Conditional Mean Forecasting.** [[Journal of Multivariate Analysis](#)]  
Lei Shu, [Yifan Hao](#), Yu Chen and Qing Yang

### Pre-prints

- **Towards Better Generalization via Distributional Input Projection Network .** [[Arxiv](#)]  
[Yifan Hao](#)\*, Yanxin Lu\*, Hanning Zhang, Xinwei Shen, Tong Zhang
- **Transformers as Multi-task Learners: Decoupling Features in Hidden Markov Models.** [[Arxiv](#)]  
[Yifan Hao](#)\*, Chenlu Ye\*, Chi Han, Tong Zhang
- **The Surprising Harmfulness of Benign Overfitting for Adversarial Robustness.** [[Arxiv](#)]  
[Yifan Hao](#) and Tong Zhang
- **On the Benefits of Over-parameterization for Out-of-Distribution Generalization.** [[Arxiv](#)]  
[Yifan Hao](#)\*, Yong Lin\*, Difan Zou and Tong Zhang
- **Efficient Model Editing with Task Vector Bases: A Theoretical Framework and Scalable Approach.** [[Arxiv](#)]  
Siqi Zeng, Yifei He, Weiqiu You, [Yifan Hao](#), Yao-Hung Hubert Tsai, Makoto Yamada, Han Zhao

where \* means alpha-order or equal contribution.

## SELECTED HONORS AND AWARDS

---

Hong Kong PhD Fellowship ( approx. 90,000 USD over two years)	2022
Outstanding Graduation of Anhui Province	2022
Outstanding Graduation of University of Science and Technology of China	2022
China National Scholarship (top 1.8% by China' s Ministry of Education)	2021
Gold Prize for Outstanding Student Scholarship	2020, 2018

## TEACHING ASSISTANT EXPERIENCE

---

MATH 2011 - Introduction to Multivariable Calculus	Fall 2023
MATH 2421 - Probability	Spring 2023, Spring 2024
Mathematical Analysis B3	Fall 2021
Complex Analysis	Spring 2021

## PROFESSIONAL ACTIVITIES

---

Conference Reviewer: NeurIPS, ICLR, ICML, AISTATS, AAAI  
Journal Reviewer: SIAM Journal on Mathematics of Data Science, JMLR

## TECHNICAL SKILLS

---

**Programming:** Python, Pytorch, MATLAB, R, C/C++