

# Yixiao HUANG

yixiaoh@berkeley.edu | homepage | +1 7348828969 | Last updated on May 2, 2026

---

## EDUCATION

### University of California, Berkeley

**Ph.D** in Computer Science    Advisor: Prof. Somayeh Sojoudi    GPA: 3.92 / 4.0    Aug, 2024 - Jun 2029

**Core courses taken:**    Convex Optimization, Theoretical Statistics, High-dimensional Statistics, Random Processes, Stochastic Systems.

### City University of Hong Kong

**B.E.** in Electrical Engineering    CGPA: 3.98 / 4.3 (**rank 1/91**)    Aug, 2019 – Jun, 2023

## SELECTED WORKS

7. Ziheng Cheng\*, [Yixiao Huang\\*](#), Hanlin Zhu, Haoran Geng, Somayeh Sojoudi, Jitendra Malik, Pieter Abbeel, and Xin Guo. **“Multi-Objective Learning for Diffusion Models: A Statistical Theory under Semi-Supervised Learning”**. The 43rd International Conference on Machine Learning (ICML), 2026.

6. Xutao Ma\*, [Yixiao Huang\\*](#), Hanlin Zhu, and Somayeh Sojoudi. **“Breaking the reversal curse in autoregressive language models via identity bridge”**. The 43rd International Conference on Machine Learning (ICML), 2026 (**Spotlight**).[pdf]

5. Samuel Pfrommer, [Yixiao Huang](#), Somayeh Sojoudi, **“Reinforcement Learning for Flow-Matching Policies”**. Preprint, 2025. [pdf]

4. [Yixiao Huang\\*](#), Hanlin Zhu\*, Tianyu Guo\*, Jiantao Jiao, Somayeh Sojoudi, Michael I. Jordan, Stuart Russell, Song Mei, **“Generalization or Hallucination? Understanding Out-of-Context Reasoning in Transformers”**. The Thirty-ninth Annual Conference on Neural Information Processing Systems (NeurIPS), 2025. [pdf]

3. Ziheng Cheng\*, [Yixiao Huang\\*](#), Hui Xu, Somayeh Sojoudi, Xuandong Zhao, Dawn Song, and Song Mei, **“OVERT: A Benchmark for Over-Refusal Evaluation on Text-to-Image Model”**. The Thirty-ninth Annual Conference on Neural Information Processing Systems Datasets and Benchmarks Track (NeurIPS D&B), 2025. [pdf]

2. Muhammed Emrullah Ildiz, [Yixiao Huang](#), Yingcong Li, Ankit Singh Rawat, and Samet Oymak. **“From Self-Attention to Markov models: Unveiling the Dynamics of Generative Transformers.”**. The 41st International Conference on Machine Learning (ICML), 2024. [pdf]

1. Yingcong Li\*, [Yixiao Huang\\*](#), Muhammed Emrullah Ildiz, Ankit Singh Rawat, Samet Oymak, **“Mechanics of Next Token Prediction with Self-Attention”**. The 27th International Conference on Artificial Intelligence and Statistics (AISTATS), 2024. (\* equal contribution) [pdf]

## HONORS AND AWARDS

### Optiver PhD Quant Focus Program

*Received return offer for Quantitative Research Intern*    Apr, 2026

**Tinker Research Grant (\$5,000 compute credits)**    Dec, 2025

**Anthropic Fellows Program for AI safety research - Finalist**    Oct, 2025

**2025 NeurIPS Scholar Award**    Oct, 2025

**2024 ICML Travel Award**    Jun, 2024

## SKILLS AND CAPABILITIES

**Programming Language:** Python (Advanced), C++ (Advanced), PyTorch (Advanced), Jax (Proficient), SQL (Proficient).

**Language:** Mandarin (Native), English (Proficient, TOEFL: 105 (S:24), GRE: 326), and Cantonese (Proficient).