JIUYAO LU

+1 667-910-3910 | jiuyaolu@wharton.upenn.edu | Google Scholar

EDUCATION

University of Pennsylvania, USA 2024 – present

Ph.D. program in Statistics and Data Science

Advisors (Department of Computer and Information Science): Aaron Roth, Michael Kearns

Amazon ASSET Fellow for Trustworthy AI

Johns Hopkins University, USA 2021 – 2024

Ph.D. program in Biostatistics

Sc.M. in Biostatistics (conferred)

Tsinghua University, China 2016 – 2021

B.S. in Pure and Applied Mathematics

SKILLS

Programming: Python (PyTorch, TensorFlow, NumPy, scikit-learn, pandas, SciPy), R, C++

Tools: LaTeX, Git, Linux, Bash, High-Performance Computing (HPC)

SELECTED RESEARCH EXPERIENCE

Neural Conditional Independence Testing via Generative Modeling

- **Discriminative independence test:** developed a framework for testing conditional independence by training discriminative neural network classifiers to distinguish between observed joint data and synthetic null samples.
- **Synthetic data generation:** leveraged neural density estimators (e.g., autoregressive models and normalizing flows) to approximate complex conditional distributions, enabling the generation of structure-preserving permutations that retain confounding relationships while breaking target dependencies.
- **Application to high-dimensional spatial data:** deployed the framework to identify latent structures in high-dimensional spatial data by conditioning on heterogeneous covariates, detecting non-linear spatial patterns missed by standard kernel-based statistical tests.

Computational Hardness and Overparameterization in Neural Networks

- **Proper vs. improper learning framework:** investigated the computational phase transition between bounded width networks and overparameterized networks, framing the width scaling as a modern analogue to the tractability gap between 3-term DNFs and 3-CNFs in classical PAC learning.
- **Hardness and convergence synthesis:** synthesized results linking the NP-hardness of training narrow networks to the global convergence of gradient descent for wide networks, characterizing how overparameterization allows optimization algorithms to circumvent worst-case hardness barriers.
- **Global convergence for general distributions:** overcame restrictive geometric data constraints in wide-network optimization by augmenting architectures with bias terms to enforce feature space separation, theoretically validating the improper learning story for general distributions.

Reliable Online Prediction for Multi-Agent Decision Making

- **One calibrated predictor, many decision makers:** leveraged algorithmic game theory and online learning to develop theory and algorithms for omniprediction, where a single forecaster's calibrated predictions enable many heterogeneous downstream decision makers to optimize their own objectives.
- **Polynomial sample complexity:** established the first omniprediction sample complexity bounds for Lipschitz losses that scale polynomially, rather than exponentially in the error parameter, without requiring convexity.

PAPERS

Machine Learning Theory (alphabetical author ordering)

- **Jiuyao Lu**, Aaron Roth, and Mirah Shi. Sample Efficient Omniprediction and Downstream Swap Regret for Non-Linear Losses. *Conference on Learning Theory (COLT)*, pp. 3829–3878. PMLR (2025).
- Natalie Collina, **Jiuyao Lu**, Georgy Noarov, and Aaron Roth. Optimal Lower Bounds for Online Multicalibration. *arXiv* preprint arXiv:2601.05245 (2026).
- Yahav Bechavod, **Jiuyao Lu**, and Aaron Roth. Dynamic Regret Bounds for Online Omniprediction with Long Term Constraints. *arXiv* preprint arXiv:2510.07266 (2025).
- Yahav Bechavod, **Jiuyao Lu**, and Aaron Roth. Online Omniprediction with Long-Term Constraints. *arXiv* preprint arXiv:2509.11357 (2025).

Machine Learning Applications

- Ruzhang Zhao, **Jiuyao Lu**, Weiqiang Zhou, Ni Zhao, and Hongkai Ji. A Systematic Evaluation of Highly Variable Gene Selection Methods for Single-Cell RNA-Sequencing. *Genome Biology*, accepted (2025).
- **Jiuyao Lu**, Glen A. Satten, Katie A. Meyer, Lenore J. Launer, Wodan Ling, and Ni Zhao. Identifying Unmeasured Heterogeneity in Microbiome Data via Quantile Thresholding (QuanT). *Microbiome*, accepted (2025).
- Wodan Ling, **Jiuyao Lu**, Ni Zhao, Anju Lulla, Anna M. Plantinga, Weijia Fu, et al. Batch Effects Removal for Microbiome Data via Conditional Quantile Regression. *Nature Communications* 13, no. 1: 5418 (2022).

Causal Inference and Experimental Design

- **Jiuyao Lu**, Tianruo Zhang, and Ke Zhu. Fast Rerandomization for Balancing Covariates in Randomized Experiments: A Metropolis–Hastings Framework. *Under review*.
- **Jiuyao Lu**, Daogao Liu, Zhanran Lin, and Xiaomeng Wang. Fast Rerandomization via the BRAIN. *arXiv* preprint arXiv:2312.17230 (2023).

HONORS AND AWARDS

American Statistical Association W. J. Youden Award in Interlaboratory Testing	2023
Ph.D. Examination Award (best performance on the JHU Biostatistics Ph.D. qualifying exam)	2022
National Scholarship (top 0.2%, highest student honor from Chinese government)	2020

PRESENTATIONS

NeurIPS 2025 Workshop on Constrained Optimization for Machine Learning (poster): Online Omniprediction with Long-Term Constraints.

ENAR 2022 (contributed paper): Benchmark Evaluation of Microbiome Differential Abundance Analysis.