Saminul Haque
saminulh@stanford.edu
(650) 665-4376
Website: https://web.stanford.edu/~saminulh
Google Scholar: https://tinyurl.com/samhscholar

EDUCATION

Stanford University, Stanford, California
Ph.D. in Computer Science, advised by Prof. John Duchi
September 2021 – Present

University of Toronto, Toronto, Ontario
H.B.Sc. in Computer Science Specialist with focus in Artificial Intelligence
Major in Mathematics, Minor in Statistics
September 2016 – May 2020 GPA: 4.0 (96.5% average)

PUBLICATIONS

A fast and slightly robust covariance estimator
- John Duchi, Saminul Haque, Rohith Kuditipudi. URL: https://arxiv.org/abs/2502.20708

Near Exact Privacy Amplification for Matrix Mechanisms
- Christopher A. Choquette-Choo, Arun Ganesh, Saminul Haque, Thomas Steinke, Abhradeep Thakurta. In International Conference on Learning Representations 2025.
  URL: https://arxiv.org/abs/2410.06266

An information-theoretic lower bound in time-uniform estimation
- John Duchi, Saminul Haque. In Conference on Learning Theory 2024.
  URL: https://arxiv.org/abs/2402.08794

Universally Instance-Optimal Mechanisms for Private Statistical Estimation
- Hilal Asi, John C. Duchi, Saminul Haque, Zewei Li, Feng Ruan. In Conference on Learning Theory 2024.
  URL: https://proceedings.mlr.press/v247/asi24a.html

A Fast Algorithm for Adaptive Private Mean Estimation
- John Duchi, Saminul Haque, Rohith Kuditipudi. In Conference on Learning Theory 2023. **Won the Mark Fulk best student paper award.**
  URL: https://arxiv.org/abs/2301.07078.pdf

Undersampling is a Minimax Optimal Robustness Intervention in Nonparametric Classification.
- Niladri S. Chatterji*, Saminul Haque*, Tatsunori Hashimoto. In Transactions on Machine Learning Research, June 2023.
  URL: https://arxiv.org/abs/2205.13094.pdf

Is Importance Weighting Incompatible with Interpolating Classifiers?
- Ke Alexander Wang*, Niladri S. Chatterji*, Saminul Haque, Tatsunori Hashimoto. In International Conference on Learning Representations, 2022.
  URL: https://arxiv.org/abs/2112.12986.pdf

Preventing Gradient Attenuation in Lipschitz Constrained Convolutional Networks.
- Qiyang Li*, Saminul Haque*, Cem Anil, James R Lucas, Roger B Grosse, and Jörn-Henrik Jacobsen. In Advances in Neural Information Processing Systems, 2019. URL: https://arxiv.org/abs/1911.00937.pdf

A comparison of two classifications of solvable Lie algebras.

- Nicolas Bryenton, Cameron Davies, Andrew Douglas, Saminul Haque et al. In Journal of Mathematical Physics 59, no. 12 (2018): 121701.
  URL: `https://aip.scitation.org/doi/abs/10.1063/1.5050789`

EXPERIENCE

Student Researcher **June 2024 – Sept. 2024**
Google

- **Privacy amplification for Matrix Mechanisms:** In this work we tackled the problem analyzing the privacy amplification of various sub-sampling schemes applied to matrix mechanisms for the purposes of differentially-private optimization. While exact theoretical guarantees on the amplification of arbitrary sub-sampling schemes are often hard to compute, we are able to obtain nearly exact amplification by an MCMC sampling-based approach. Moreover, we introduce the balls-in-bins sub-sampling scheme which is simultaneously practically usable (and amenable to our amplification analysis method), and gets comparable amplification utility to Poisson sub-sampling.

Ph.D. Student advised by Prof. John Duchi **Sept. 2021 – Present**
Stanford University

- **Differentially-private and robust statistics:** In this line of work, we investigate covariance-adaptive mean estimation in the differentially-private and robust settings. In the private setting, we developed an algorithm for the problem of differentially private mean estimation that is adaptive to the population covariance. Previous work either had suboptimal sample-complexity or were intractable. Our algorithm has near-optimal sample-complexity and runs in near-linear time. In the robust setting, we focus instead on covariance estimation in relative-operator norm (which can then be used for mean estimation). We develop an algorithm that filters based on inner products and is able to learn the covariance in the low contamination setting. While polynomial-time SoS algorithms already exist in this setting, our algorithm is much faster, with runtime on slightly slower than matrix multiplication.
- **Instance-optimal private statistical estimation:** In this work, we investigated the problem of private statistical estimators, with sample complexity adapting to the difficulty of the problem. Adapting Birgé's T-estimators to the private setting, we develop an estimator that is instance optimal in the low-dimensional setting.
- **Lower bounds in time-uniform estimation:** In this work, we investigated the cost of time-uniform estimation. Using a novel information-theoretic technique, we recover the typical lower bound derived from the law-of-iterated-logarithm (LIL) without the assumptions required for LIL to hold.
- **Importance weighting and undersampling:** In this line of work, we investigated the problem of handling distribution shift in machine learning settings. In our first paper, we demonstrate that the popular "importance weighting" technique is ineffectual for interpolating models, rendering it useless as a means of correcting for distribution shift. In our second work, we demonstrate that the naïve undersampling algorithm is minimax-optimal; thereby necessitating the need for additional problem structure to possibly better handle distribution shift.

In Progress Projects:

- **DP benchmarking platform:** With John Duchi and Gavin Brown, I am developing a testing platform to enable the system testing and benchmarking of differentially private algorithms for practical applications.

- **Robust covariance estimation:** With John Duchi and Rohith Kuditipudi, we are extending our work on robust covariance estimation to larger contamination regimes.
- **Estimation under missingness:** With John Duchi, Rohith Kuditipudi, Ankit Pensia, and Kabir Verchand, we are investigating efficient algorithms for estimation under missingness assumptions beyond the typical missing completely at random assumption.
- **Privacy of HMC:** With John Duchi and Abhradeep Thakurta, we are developing tools to track privacy loss throughout Hamiltonian Monte Carlo, so that HMC algorithms can be used privately without requiring convergence.

Research Assistant under Prof. Roger Grosse **Oct. 2020 – Jun. 2021**
University of Toronto

- **Self-Tuning Networks (STNs):** I analyzed convergence properties of the Δ-STN method and found asymptotic optimality with exact updates under reasonable conditions. I then devised modifications to the method that retains asymptotic optimality and has finite asymptotic variance under stochastic updates.

Research Intern **Jun. 2020 – Aug. 2020**
Jane Street Capital

- Developed various detectors for anomalies in market data.

Research Student under Prof. Roger Grosse **Feb. 2019 – Jan. 2020**
University of Toronto

- **Bayesian Neural Networks Group:** Worked in a group investigating the problem of cost-sensitive Bayesian optimization. The project focused on benchmarking various criteria and methods for this task.

Data Science Intern **May – Sept. 2018**
Department of National Defence

- Used a variety of regression and classification techniques to predict the delays in maintenance projects for the Royal Canadian Navy.

AWARDS

- Governor General's Silver Model (May 2020)
- The Chancellor Northrop Frye Gold Medal (June 2020)
- Roberts Scholarship (Sept. 2016 – Sept. 2019)