Mufan Li

Email: mufan.li@uwaterloo.ca Website: mufan-li.github.io

ACADEMIC POSITIONS

Assistant Professor, University of Waterloo — 2025–Present
Department of Statistics and Actuarial Science

Postdoctoral Research Associate, Princeton University — 2023–2025
Department of CSML and ORFE
Supervised by Boris Hanin

DEGREES

Ph.D. Statistics, University of Toronto — 2017–2023
Thesis: *Analysis of Learning Algorithms via Diffusion Limits*
Supervised by Daniel M. Roy and Murat A. Erdogdu

M.Sc. Statistics, University of Toronto — 2015–2016

B.A.Sc. Engineering Science, University of Toronto — 2010–2015

PUBLISHED ARTICLES

See also my Google Scholar or Semantic Scholar pages.

1. Nolan Dey, Bin Claire Zhang, Lorenzo Noci, **M. Li**, Blake Bordelon, Shane Bergsma, Cengiz Pehlevan, Boris Hanin, and Joel Hestness, *Don't be lazy: CompleteP enables compute-efficient deep transformers*. NeurIPS (2025). arXiv:2505.01618.
2. Sinho Chewi, Murat A. Erdogdu, **M. Li**, Ruoqi Shen, and Matthew Zhang, *Analysis of Langevin Monte Carlo from Poincaré to Log-Sobolev*. Foundations of Computational Mathematics (2024). COLT 2022 Extended Abstract. arXiv:2112.12662.
3. Yunbum Kook, Matthew S. Zhang, Sinho Chewi, Murat A. Erdogdu, and **M. Li**, *Sampling from the Mean-Field Stationary Distribution*. COLT 2024. arXiv:2402.07355.
4. **M. Li** and Mihai Nica, *Differential Equation Scaling Limits of Shaped and Unshaped Neural Networks*. TMLR 2024. arXiv:2310.12079.
5. Blake Bordelon, Lorenzo Noci, **M. Li**, Boris Hanin, and Cengiz Pehlevan, *Depthwise Hyperparameter Transfer in Residual Networks: Dynamics and Scaling Limit*. ICLR 2024. M3L Workshop Oral Presentation. arXiv:2309.16620.
6. Lorenzo Noci*, Chuning Li*, **M. Li***, Bobby He, Thomas Hofmann, Chris Maddison, and Daniel M. Roy, *The Shaped Transformer: Attention Models in the Infinite Depth-and-Width Limit*. NeurIPS 2023. arXiv:2306.17759.
7. Matthew Zhang, Sinho Chewi, **M. Li**, Krishnakumar Balasubramanian, and Murat A. Erdogdu, *Improved Discretization Analysis for Underdamped Langevin Monte Carlo*. COLT 2023. arXiv:2302.08049.
8. **M. Li** and Murat A. Erdogdu, *Riemannian Langevin Algorithm for Solving Semidefinite Programs*. Bernoulli (2023). arXiv:2010.11176.
9. **M. Li**, Mihai Nica, and Daniel M. Roy, *The Neural Covariance SDE: Shaped Infinite Depth-and-Width Networks at Initialization*. NeurIPS 2022 (Selected for Oral, Nominated for Outstanding Paper Award). arXiv:2206.02768.
10. Raphaël Berthier and **M. Li**, *Acceleration of Gossip Algorithms through the Euler–Poisson–Darboux Equation*. IMA Journal of Applied Mathematics (2022). arXiv:2202.10742.

*Equal Contribution.

11. **M. Li**, Mihai Nica, and Daniel M. Roy, *The Future is Log-Gaussian: ResNets and Their Infinite-Depth-and-Width Limit at Initialization.* NeurIPS 2021. arXiv:2106.04013.

PREPRINTS

1. Yihe Dong, Lorenzo Noci, Mikhail Khodak, and **M. Li**, *Attention Retrieves, MLP Memorizes: Disentangling Trainable Components in the Transformer.* Preprint (2025). arXiv:2506.01115.
2. **M. Li**, and Maxime Gazeau, *Higher Order Generalization Error for First Order Discretization of Langevin Diffusion.* Preprint (2021). arXiv:2102.06229.

AWARDS

Princeton DataX Postdoctoral Fellowship	2024–2025
NSERC Postdoctoral Fellowship (Declined)	2024
Doctoral Award, University of Toronto	2023
Ontario Graduate Scholarship	2019–2023
Student Research Presentation Award, Stat. Soc. of Canada	2021
MITACS Accelerate Fellowship, with Borealis AI	2018–2019
Undergraduate Summer Research Fellowship, University of Toronto	2012

RECENT INVITED TALKS

Gatsby Seminar, University College London *The Proportional Scaling Limit of Neural Networks*	October 2025
Stochastic Processes and Applications (SPA) Conference *The Proportional Scaling Limit of Neural Networks*	July 2025
CRM Workshop: Random Matrices and High-Dim. Learning Dynamics *The Proportional Scaling Limit of Neural Networks*	June 2025
Deep Learning Theory Seminar, University of Tokyo *The Proportional Scaling Limit of Neural Networks*	May 2025
Math ML Seminar, MPI MiS and UCLA *The Proportional Scaling Limit of Neural Networks*	May 2025
Probability Seminar, University of Washington *The Proportional Scaling Limit of Neural Networks*	February 2025
INFORMS Annual Meeting *The Proportional Scaling Limit of Neural Networks*	October 2024
STATQAM Seminar, UQAM *The Proportional Scaling Limit of Neural Networks*	September 2024
Cerebras Systems *Infinite-Depth Neural Networks as Depthwise Stochastic Processes*	June 2024
Transformers Seminar, Flatiron Institute *Neural Covariance SDE and the Shaped Transformer*	April 2024
Alg-ML Seminar, Princeton University *Neural Covariance SDE and Its Limiting Spectrum*	April 2024
One World Mathematics of ML Seminar (Video) *Infinite-Depth Neural Networks as Depthwise Stochastic Processes*	April 2024
LCDS Seminar, Brown University *Geometric Dyson Brownian Motion and the Free Log-Normal for Minor of Products of Random Matrices*	November 2023

EDITORIAL SERVICE

International Conference on Learning Representations (ICLR), Area Chair 2024-2025

	Algorithmic Learning Theory Conference (ALT), Area Chair	2024-2025

WORK EXPERIENCE

Research Intern, Borealis AI	Aug 2018–Apr 2019
Investment Analyst, Ontario Teachers' Pension Plan	Jul 2016–Jul 2017
Electronic Trading Intern, RBC Capital Markets	May 2013–Aug 2014

PEER REVIEW

Annals of Applied Probability (AoAP).
Foundations of Computational Mathematics (FoCM)
Journal of Machine Learning Research (JMLR)
Transactions on Machine Learning Research (TMLR) Expert Reviewer
SIAM Journal on Mathematics of Data Science (SIMODS)
Journal of Computational and Graphical Statistics (JCGS)
Neural Information Processing Systems (NeurIPS)
International Conference on Learning Representations (ICLR)
International Conference on Machine Learning (ICML)