Lang Liu

Education

University of Washington (UW), *Seattle* Sep. 2017 – Dec. 2022 (expected)
Ph.D. in Statistics (Machine Learning and Big Data Track)
Advisors: Zaid Harchaoui & Soumik Pal

Tsinghua University, *Beijing* Sep. 2013 – Jun. 2017
B.S. in Mathematics and Applied Mathematics
Thesis: Bayesian Structure Learning for Stationary Time Series

Publications

Lang Liu, Mahdi Milani Fard, Sen Zhao. Distribution Embedding Networks for Generalization from a Diverse Set of Classification Tasks. *Transactions on Machine Learning Research*, 2022.

Lang Liu, Carlos Cinelli, Zaid Harchaoui. Orthogonal Statistical Learning with Self-Concordant Loss. *COLT*, 2022.

Lang Liu, Soumik Pal, Zaid Harchaoui. Entropy Regularized Optimal Transport Independence Criterion. *AISTATS*, 2022 (Oral, top 2.6% of the submissions).

Lang Liu, Krishna Pillutla, Sean Welleck, Sewoong Oh, Yejin Choi, Zaid Harchaoui. Divergence Frontiers for Generative Models: Sample Complexity, Quantization Effects, and Frontier Integrals. *NeurIPS*, 2021.

Lang Liu, Joseph Salmon, Zaid Harchaoui. Score-Based Change Detection for Gradient-Based Learning Machines. *ICASSP*, 2021.

Workshops

Lang Liu, Zaid Harchaoui. Likelihood Score under Generalized Self-Concordance. *NeurIPS Score-Based Methods Workshop*, 2022.

Zaid Harchaoui, **Lang Liu**, Soumik Pal. Discrete Schrödinger Bridges with Applications to Two-Sample Homogeneity Testing. *NeurIPS OTML Workshop*, 2021 (Best Paper Award).

Working papers

Lang Liu, Zaid Harchaoui. Confidence Sets under Generalized Self-Concordance. *Submitted*. Available at arXiv.

Krishna Pillutla*, **Lang Liu***, John Thickstun, Sean Welleck, Swabha Swayamdipta, Rowan Zellers, Sewoong Oh, Yejin Choi, Zaid Harchaoui (*Equal contribution). MAUVE Scores for Generative Models: Theory and Practice. *Submitted*. Available at arXiv.

Zaid Harchaoui, **Lang Liu**, Soumik Pal. Asymptotics of Discrete Schrödinger Bridges via Chaos Decomposition. *Major revision at Bernoulli*. Available at arXiv.

Jillian Fisher, **Lang Liu**, Krishna Pillutla, Yejin Choi, Zaid Harchaoui. Statistical and Computational Guarantees for Influence Diagnostics. *Submitted*. Available at arXiv.

Ronak Mehta, Vincent Roulet, Krishna Pillutla, **Lang Liu**, Zaid Harchaoui. Stochastic Optimization for Spectral Risk Measures. *Submitted*. Available at arXiv.

Software

Autodetect, autodiff-friendly change detection for monitoring machine learning models.

Research Experience

Graduate Research Assistant Jul. 2018 – Present
University of Washington, *Seattle*
Advisor: Zaid Harchaoui

Undergraduate Research Assistant Dec. 2015 – Jul. 2017
Tsinghua University, *Beijing*
Advisor: Xuegong Zhang

Research Intern Jul. 2016 – Sep. 2016
University of Washington, *Seattle*
Advisors: Emily Fox, Nicholas Foti

Professional Experience

Data Scientist Intern Jun. 2020 – Sep. 2020
Glassbox Machine Learning Team, *Google Research*
Hosts: Sen Zhao & Mahdi Milani Fard

- Proposed and implemented a novel meta-learning approach for applications where both the data distribution and the number of features could vary across tasks.
- Developed a novel methodology to massively simulate binary classification training tasks for the proposed approach.
- Demonstrated significant improvement compared to baselines in experiments.

Applied Scientist Intern Jun. 2019 – Sep. 2019
Music Machine Learning Team, *Amazon*
Manager & Mentor: Fabian Moerchen & Brandyn Kusenda

- Designed a deep track-query joint embedding model to search for relevant music entities given long tail queries.
- Collected, analyzed, and processed a dataset on music playbacks using PySpark, and used it to train the joint embedding model.
- Demonstrated significant improvements compared to existing approaches.

Honors and Awards

Z.W. Birnbaum Award, *Department of Statistics, University of Washington* 2022
Post-General Statistics Conference Travel Award, *University of Washington* 2022
Graduate Student Conference Presentation Award, *University of Washington* 2022
Best Paper Award, *NeurIPS OTML Workshop* 2021
Second Prize in the Mathematical Contest in Modeling, *CUMCM* 2016
Academic Excellence Award, *Department of Mathematics, Tsinghua University* 2015
Honorable Mention in the Mathematical Contest in Modeling, *COMAP* 2015
First Prize in the Math Olympiad, *Hunan Province, China* 2011 & 2012

Talks

Orthogonal Statistical Learning with Self-Concordant Loss

- *UW IFDS Seminar*, Oct. 2022.
- *IFDS Workshop on Distributional Robustness*, Aug. 2022.
- *COLT*, Jul. 2022.

Entropy Regularized Optimal Transport Independence Criterion

- *COMPSTAT*, Aug. 2022.
- *JSM*, Aug. 2022.
- *PIMS-IFDS-NSF Summer School on Optimal Transport*, Jun. 2022.
- *AISTATS*, Mar. 2022.
- *UW Kantorovich Retreat*, Mar. 2022.

Divergence Frontiers for Generative Models: Sample Complexity, Quantization Effects, and Frontier Integrals

- *SIAM MDS 2022*, Sep. 2022.
- *TRIPODS Meeting*, Sep, 2022.
- *UW IFDS Seminar*, Jan. 2022.
- *NeurIPS*, Dec. 2021.

- *Joint IFML/CCSI Symposium*, Nov. 2021.

Discrete Schrödinger Bridges with Applications to Two-Sample Homogeneity Testing.
- *NeurIPS OTML Workshop*, Dec. 2021.

Asymptotics of entropy-regularized optimal transport via chaos decomposition.
- *Joint Statistical Meeting*, Aug. 2021.
- *BIRS Workshop on Entropic Regularization of Optimal Transport and Applications*, Jun. 2021.
- *UW Probability Seminar*, Nov. 2020.
- *UW Machine Learning Retreat*, Nov. 2020.

Gradient-based monitoring of learning machines.
- *IEEE International Conference on Acoustics, Speech and Signal Processing*, Jun. 2021.
- *Symposium on Data Science and Statistics*, Jun. 2021.
- *IFDS Kickoff Meeting Poster Session*, Sep. 2020.
- *Google Statistics Journal Club*, Sep. 2020.
- *Google Research NYC and Athena Org Intern Talks*, Jul. 2020.
- *ICML Workshop on Challenges in Deploying and Monitoring Machine Learning Systems*, Jul. 2020.

Teaching

Teaching Assistant, *University of Washington*
- CSE 541: Interactive Learning 2022
- MATH 394: Probability I 2021
- STAT 516: Stochastic Modeling 2020
- STAT 538: Statistical Learning 2019 & 2020
- STAT 311: Elements of Statistical Methods 2017 & 2018

Guest lecture on statistical machine learning with random features, *STAT 538* 2022

Tutorial on optimal transport in computational neuroscience, *Neurohackademy* 2020

Tutor for mathematics, *Tsinghua University* 2015

Mentoring

Jillian Fisher (UW Ph.D. in Statistics, May 2022 – Dec. 2022)
Ronak Mehta (UW Ph.D. in Statistics, Sep. 2022 – Present)
Medha Agarwal (UW Ph.D. in Statistics, Oct. 2022 – Present)

Services

Co-founder of the Internship Preparation Program in Statistics at UW.
Reviewer for the Annals of Applied Probability.
Reviewer for the Journal of Machine Learning Research.
Reviewer for Statistics and Computing.
Reviewer for the Journal of Computational and Graphical Statistics.
Reviewer for the Journal of Optimization Theory and Applications.
Reviewer for the ICML 2021 & 2022, NeurIPS 2020 & 2021 & 2022, AISTATS 2022.

Reading Groups

Host Machine Learning and Mass Transportation working group at UW, 2021.

Skills

Python, PyTorch, R, C++, MATLAB