Licong Lin

✉ liconglin@berkeley.edu | 🏠 licong-lin.github.io | 437 Evans Hall, UC Berkeley, CA, 94720

Education

University of California at Berkeley   *Berkeley, CA*
PhD student in Statistics   *Aug 2021 - Current*
- Advised by Song Mei and Peter Bartlett.

Peking University   *Beijing, China*
B.S. in Statistics   *Sep 2017 - July 2021*
- GPA: **3.82**/4 (rank: **2**/45); Honors graduate; Advised by Cheng Zhang.

Research Interests

- Theory of AI and deep learning, AI safety, high-dimensional statistics, statistical inference.

- My work focuses on **understanding the statistical foundations of AI** and on **developing theory-inspired algorithms for AI**. I use and extend tools from statistical learning theory, high-dimensional statistics, and optimization to study, e.g., foundations of multimodal generative AI, scaling laws, Transformers, in-context learning, and to build mathematically motivated algorithms for AI safety.

Work Experience

Amazon Web Services   *Santa Clara, CA*
applied scientist intern, MLSys; manager: Jiaji Huang.   *May 2024 - August 2024*

Google Research   *New York, NY*
student researcher, BigML; manager: Aditya Menon.   *Jan 2025 - May 2025*

The Voleon Group   *Berkeley, CA*
quantitative research intern.   *June 2025 - August 2025*

Teaching Experience

UC Berkeley   *Berkeley, CA*
Graduate Student Instructor (TA)   *Jan 2022 - Dec 2023*
- STAT 153: introduction to time series (Instructor: Ruoqi Yu, Spring 2022).
- STAT 210B: theoretical statistics II (Instructor: Song Mei, Spring 2023).
- STAT 135: concepts of statistics (Instructor: Adam Lucas, Fall 2023).
- Received Outstanding Graduate Student Instructor Award in 2023.

Services

- **Journal Reviewer**: Annals of Statistics (AoS); Journal of the American Statistical Association (JASA); SIAM Journal on Mathematics of Data Science (SIMODS); Journal of Machine Learning Research (JMLR); Transactions on Machine Learning Research (TMLR); IEEE Information Theory Workshop (ITW).

- **Conference Reviewer**: International Conference on Machine Learning (ICML); Conference on Neural Information Processing Systems (NeurIPS); Conference on Learning Theory (COLT); International Conference on Learning Representations (ICLR); International Conference on Artificial Intelligence and Statistics (AISTATS).

- **Mentorship**: Student mentor at the Statistics Undergraduate Mentorship (SUM) program and Statistics Graduate Student Association (SGSA) mentorship program, UC Berkeley, 2022–2025.

Talks

08/2025	Joint Statistical Meeting (JSM)	*Nashville, TN*
10/2024	INFORMS Annual Meeting	*Seattle, WA*
08/2024	Joint Statistical Meeting (JSM)	*Portland, OR*
04/2024	UIUC Machine Learning Reading Group (online)	*Urbana, IL*

Honors & Awards

2023	**Outstanding Graduate Student Instructor**, UC Berkeley	*Berkeley, CA*
2021	**Huaixin Bachelor**, Peking University	*China*
2021	**Honor graduate of Applied Mathematics and Statistics Program**, Peking University	*China*
2018-20	**Academic Excellence Award**, Peking University	*China*
2019,20	**Peking University Scholarship**, Peking University	*China*
2020	**Gold Medal in Probability & Statistics**, *S.-T. Yau College Student Mathematics Contest, placed 1st nationally*	*China*
2018	**1st Prize**, Beijing College Student Mathematics Competition	*China*
2016	**2nd Prize**, China National Mathematical Olympiad	*China*

Preprints (∗ for a co-first author or alphabetical order)

A Statistical Theory of Contrastive Learning via Approximate Sufficient Statistics
Licong Lin, Song Mei
arXiv preprint arXiv:2503.17538 (2025). 2025

Improved Scaling Laws in Linear Regression via Data Reuse
Licong Lin, Jingfeng Wu, Peter L Bartlett
arXiv preprint arXiv:2506.08415 (2025). 2025

A Statistical Theory of Contrastive Pre-training and Multimodal Generative AI
Kazusato Oko*, **Licong Lin***, Yuhang Cai*, Song Mei*
arXiv preprint arXiv:2501.04641 (2025). 2025

Minimax Optimal Convergence of Gradient Descent in Logistic Regression via Large and Adaptive Stepsizes
Ruiqi Zhang, Jingfeng Wu, **Licong Lin**, Peter L Bartlett
arXiv preprint arXiv:2504.04105; an earlier version was accepted to ICML (2025). 2025

Simplicity Prevails: Rethinking Negative Preference Optimization for LLM Unlearning
Chongyu Fan*, Jiancheng Liu*, **Licong Lin***, Jinghan Jia, Ruiqi Zhang, Song Mei, Sijia Liu
arXiv preprint arXiv:2410.07163 (2024). 2024

When is it worthwhile to jackknife? Breaking the quadratic barrier for Z-estimators
Licong Lin*, Fangzhou Su*, Wenlong Mou, Peng Ding, Martin Wainwright
arXiv preprint arXiv:2411.02909 (2024). 2024

Mean-field variational inference with the TAP free energy: Geometric and statistical properties in linear models
Michael Celentano*, Zhou Fan*, **Licong Lin***, Song Mei*
arXiv preprint arXiv:2311.08442 (2023). 2023

Near-optimal multiple testing in Bayesian linear models with finite-sample FDR control
Taejoo Ahn*, **Licong Lin***, Song Mei*
arXiv preprint arXiv:2211.02778 (2022). 2022

Publications

Semiparametric inference based on adaptively collected data
Licong Lin*, Koulik Khamaru*, Martin J Wainwright
The Annals of Statistics 53.3 (2025) pp. 989–1014. 2025

Improving llm safety alignment with dual-objective optimization
Xuandong Zhao, Will Cai, Tianneng Shi, David Huang, **Licong Lin**, Song Mei, Dawn Song
International Conference on Machine Learning (ICML) (2025). 2025

Transformers as Decision Makers: Provable In-Context Reinforcement Learning via Supervised Pretraining
Licong Lin, Yu Bai, Song Mei
International Conference on Learning Representations (ICLR) (2024). 2024

Scaling Laws in Linear Regression: Compute, Parameters, and Data
Licong Lin, Jingfeng Wu, Sham M Kakade, Peter L Bartlett, Jason D Lee
Advances in Neural Information Processing Systems 37 (2024). 2024

Plug-in Performative Optimization
Licong Lin, Tijana Zrnic
International Conference on Machine Learning (ICML) (2024). 2024

Negative Preference Optimization: From Catastrophic Collapse to Effective Unlearning
Ruiqi Zhang*, **Licong Lin***, Yu Bai, Song Mei
Conference on Language Modeling (COLM) (2024). 2024

Statistical Limits of Adaptive Linear Models: Low-Dimensional Estimation and Inference
Licong Lin, Mufang Ying, Suvrojit Ghosh, Koulik Khamaru, Cun-Hui Zhang
Advances in Neural Information Processing Systems 36 (2023). 2023

What causes the test error? going beyond bias-variance via anova
Licong Lin, Edgar Dobriban
The Journal of Machine Learning Research 22.1 (2021) pp. 6925–7006. JMLRORG, 2021