LICONG LIN

PERSONAL INFORMATION

Email: liconglin@berkeley.edu
Website: https://licong-lin.github.io
Address: 437 Evans Hall, UC Berkeley, Berkeley, CA 94720

EDUCATION

University of California, Berkeley August 2021 - Present
Ph.D. candidate in Statistics
Advisors: Song Mei and Peter Bartlett

Peking University September 2017 - July 2021
B.S. in Statistics; graduated with honors
GPA: 3.82/4.00; rank: 2/45

RESEARCH INTERESTS

- Theory of AI, AI alignment, 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 the statistical foundations of architectures, algorithms and phenomena in modern AI (e.g., Transformers, in-context learning, scaling laws, contrastive learning and multimodal generative AI), and to build mathematically motivated algorithms for AI alignment (e.g., LLM unlearning).

WORK EXPERIENCE

- Applied scientist intern, **Amazon Web Services**; Santa Clara, CA 05/2024 - 08/2024
- Student researcher, **Google Research**; New York, NY 01/2025 - 05/2025
- Quantitative research intern, **The Voleon Group**; Berkeley, CA 06/2025 - 08/2025

TEACHING EXPERIENCE

Graduate Student Instructor (TA), **UC Berkeley**; Berkeley, CA 01/2022 - 12/2023

- 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 UC Berkeley 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).

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

INVITED TALKS

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

HONORS AND AWARDS

- Outstanding Graduate Student Instructor, UC Berkeley 2023
- Huaixin Bachelor, Peking University 2021
- Honor graduate of Applied Mathematics and Statistics Program, Peking University 2021
- Academic Excellence Award, Peking University 2018, 19, 20
- Peking University Scholarship, Peking University 2019, 20
- Gold Medal in Probability & Statistics, S-T. Yau College Student Mathematics Contest, placed 1st nationally, China 2020
- 1st Prize, Beijing College Student Mathematics Competition, China 2018
- 2nd Prize, China National Mathematical Olympiad, China 2016

PREPRINTS

($\star$ for co-first author or alphabetical order)

- **Licong Lin**, Jingfeng Wu, and Peter L. Bartlett. "Improved scaling laws in linear regression via data reuse." *arXiv preprint:2506.08415* (2025). arXiv.
- **Licong Lin** and Song Mei. "A statistical theory of contrastive learning via approximate sufficient statistics." *arXiv preprint:2503.17538* (2025). arXiv.
- Kazusato Oko*, **Licong Lin***, Yuhang Cai*, and Song Mei*. "A statistical theory of contrastive pre-training and multimodal generative AI." *arXiv preprint:2501.04641* (2025). arXiv.
- Ruiqi Zhang, Jingfeng Wu, **Licong Lin**, and Peter L. Bartlett. "Minimax optimal convergence of gradient descent in logistic regression via large and adaptive stepsizes." *arXiv preprint:2504.04105* (2025), an earlier version was accepted to ICML 2025. arXiv.
- Chongyu Fan*, Jiancheng Liu*, **Licong Lin***, Jinghan Jia, Ruiqi Zhang, Song Mei, and Sijia Liu. "Simplicity prevails: Rethinking negative preference optimization for LLM unlearning." *arXiv preprint:2410.07163* (2024). arXiv.
- **Licong Lin***, Fangzhou Su*, Wenlong Mou, Peng Ding, and Martin J. Wainwright. "When is it worthwhile to jackknife? Breaking the quadratic barrier for Z-estimators." *arXiv preprint:2411.02909* (2024). arXiv.

- Michael Celentano*, Zhou Fan*, **Licong Lin***, and Song Mei*. "Mean-field variational inference with the TAP free energy: Geometric and statistical properties in linear models." *arXiv preprint:2311.08442* (2023). arXiv.
- Taejoo Ahn*, **Licong Lin***, and Song Mei*. "Near-optimal multiple testing in Bayesian linear models with finite-sample FDR control." *arXiv preprint:2211.02778* (2022). arXiv.

PUBLICATIONS

- **Licong Lin***, Koulik Khamaru*, and Martin J. Wainwright. "Semiparametric inference based on adaptively collected data." *The Annals of Statistics* (2025). Journal, arXiv.
- Xuandong Zhao, Will Cai, Tianneng Shi, David Huang, **Licong Lin**, Song Mei, and Dawn Song. "Improving LLM safety alignment with dual-objective optimization." *International Conference on Machine Learning (ICML)* (2025). arXiv.
- **Licong Lin**, Jingfeng Wu, Sham M. Kakade, Peter L. Bartlett, and Jason D. Lee. "Scaling laws in linear regression: Compute, parameters, and data." *Advances in Neural Information Processing Systems (NeurIPS)* (2024). arXiv.
- Ruiqi Zhang*, **Licong Lin***, Yu Bai, Song Mei. "Negative preference optimization: From catastrophic collapse to effective unlearning." *Conference on Language Modeling (COLM)* (2024). arXiv.
- **Licong Lin**, Yu Bai, and Song Mei. "Transformers as decision makers: Provable in-context reinforcement learning via supervised pretraining." *International Conference on Learning Representations (ICLR)* (2024). arXiv.
- **Licong Lin** and Tijana Zrnic. "Plug-in performative optimization." *International Conference on Machine Learning (ICML)* (2024). arXiv.
- **Licong Lin**, Mufang Ying, Suvrojit Ghosh, Koulik Khamaru, and Cun-Hui Zhang. "Statistical limits of adaptive linear models: Low-dimensional estimation and inference." *Advances in Neural Information Processing Systems (NeurIPS)* (2023). arXiv.
- **Licong Lin** and Edgar Dobriban. "What causes the test error? Going beyond bias-variance via ANOVA." *Journal of Machine Learning Research (JMLR)* (2021). Journal, arXiv.