SPENCER FREI

Assistant Professor of Statistics, UC Davis ◇ spencerfrei.github.io ◇ sfrei@ucdavis.edu

RESEARCH INTERESTS

Machine learning, statistics, and optimization, with a particular interest in understanding and improving deep learning. Areas of focus include generalization theory for neural networks, implicit regularization of optimization algorithms, and large language models.

EMPLOYMENT

Tenure-track Assistant Professor, Department of Statistics, University of California, Davis *2023–present*

Postdoctoral Fellow, Simons Institute for the Theory of Computing, UC Berkeley *2021–2023*
— Mentors: Peter Bartlett and Bin Yu.
— Part of the NSF/Simons Collaboration on the Theoretical Foundations of Deep Learning.
— Visitor at École polytechnique fédérale de Lausanne (EPFL), Sep. & Oct. 2022, hosted by Emmanuel Abbé.

EDUCATION

Ph.D, Statistics, UCLA *2015–2021*
— Co-advisors: Quanquan Gu and Ying Nian Wu; committee members: Arash Amini and Qing Zhou.

MSc., Mathematics, University of British Columbia, Vancouver *2013–2015*
— Advisor: Edwin A. Perkins.

BSc., Mathematics, McGill University, Montréal *2009–2013*
— First class honours.

PUBLICATIONS AND PREPRINTS

1. Nikhil Ghosh, **Spencer Frei**, Wooseok Ha, and Bin Yu. The effect of SGD batch size on autoencoder learning: Sparsity, sharpness and feature learning. *Preprint*, arXiv:2308.03215.
2. Ruiqi Zhang, **Spencer Frei**, and Peter L. Bartlett. Trained transformers learn linear models in-context. *Preprint*, arXiv:2306.09927.
3. **Spencer Frei***, Gal Vardi*, Peter L. Bartlett, and Nathan Srebro. Benign overfitting in linear classifiers and leaky ReLU networks from KKT conditions for margin maximization. *Conference on Learning Theory (COLT),* 2023.
4. **Spencer Frei***, Gal Vardi*, Peter L. Bartlett, and Nathan Srebro. The double-edged sword of implicit bias: Generalization vs. robustness in ReLU networks. *Preprint*, arXiv:2303.01456.
5. **Spencer Frei***, Gal Vardi*, Peter L. Bartlett, Nathan Srebro, and Wei Hu. Implicit bias in leaky ReLU networks trained on high-dimensional data. *International Conference on Learning Representations (ICLR)*, 2023. **(Spotlight)**
6. **Spencer Frei**, Niladri Chatterji, and Peter L. Bartlett. Random feature amplification: Feature learning and generalization in neural networks. *Journal of Machine Learning Research*, 2023+ (accepted pending minor revision).
7. **Spencer Frei**, Niladri Chatterji, and Peter L. Bartlett. Benign overfitting without linearity: Neural network classifiers trained by gradient descent for noisy linear data. *Conference on Learning Theory (COLT)*, 2022.

8. **Spencer Frei***, Difan Zou*, Zixiang Chen*, and Quanquan Gu. Self-training converts weak learners to strong learners in mixture models. *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2022.

9. **Spencer Frei** and Quanquan Gu. Proxy convexity: A unified framework for the analysis of neural networks trained by gradient descent. *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

10. Difan Zou*, **Spencer Frei***, and Quanquan Gu. Provable robustness of adversarial training for learning halfspaces with noise. *International Conference on Machine Learning (ICML)*, 2021.

11. **Spencer Frei**, Yuan Cao, and Quanquan Gu. Provable generalization of SGD-trained neural networks of any width in the presence of adversarial label noise. *International Conference on Machine Learning (ICML)*, 2021.

12. **Spencer Frei**, Yuan Cao, and Quanquan Gu. Agnostic learning of halfspaces with gradient descent via soft margins. *International Conference on Machine Learning (ICML)*, 2021. **(Long Talk)**

13. **Spencer Frei**, Yuan Cao, and Quanquan Gu. Agnostic learning of a single neuron with gradient descent. *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

14. Ariana E. Anderson, Mirella Diaz-Santos, **Spencer Frei** *et al.* Hemodynamic latency is associated with reduced intelligence across the lifespan: an fMRI DCM study of aging, cerebrovascular integrity, and cognitive ability. *Brain Structure and Function*, 2020.

15. **Spencer Frei**, Yuan Cao, and Quanquan Gu. Algorithm-dependent generalization bounds for overparameterized deep residual networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

16. **Spencer Frei** and Edwin Perkins. A lower bound for p_c in range-R bond percolation in two and three dimensions. *Electronic Journal of Probability* 21(56), 2016.

17. **Spencer Frei**, Kathryn Lockwood, Greg Stewart, Justin Boyer, and Burt S. Tilley, On thermal resistance in concentric residential geothermal heat exchangers. *Journal of Engineering Mathematics* 86(1), 2014.

* indicates equal contribution.

INDUSTRY EXPERIENCE

Applied Scientist Intern, Amazon Alexa AI, Cambridge, MA *Summer 2020*

— Worked on natural language understanding using Transformer-based multilingual language models.

Student Researcher, Chatterbaby/UCLA School of Medicine, Los Angeles, CA *2018–2020*

— Developed deep learning models for audio analysis of infant cries that were deployed in Chatterbaby app (100k+ downloads on Android). Employed 20 hr/week when not teaching.

Biostatistical Consultant, Ritter Pharmaceuticals, Los Angeles, CA *2017–2019*

— Consulted on the analysis of clinical trial data using linear regression and mixed effects models.

Student Researcher, Blackthorn Therapeutics/UCLA School of Medicine, Los Angeles, CA *2016–2018*

— Analyzed MRI and neuropsychiatric data using generalized linear models for a biotech company-sponsored project. Employed 20 hr/week when not teaching.

HONORS

- Rising Star in Machine Learning, University of Maryland, 2022.
- Best Reviewer Award: ICML 2020, ICML 2021, NeurIPS 2021, ICLR 2022.
- Dissertation Year Fellowship, UCLA, 2020–2021.
- Most Promising Computational Statistician, UCLA, 2016.
- Research Fellowship, Montréal Institut des sciences mathématiques, 2012.

TALKS

• "Learning linear models in-context with transformers"
— University of Basel, Aurelien Lucchi group meeting *November 2023*
— University of Oxford, Computational Statistics and Machine Learning Seminar *October 2023*
— University of Cambridge, Machine Learning Group meeting *October 2023*
— Google DeepMind, London *October 2023*
— Imperial College London, Imperial + AI Seminar *October 2023*
— Google Research *August 2023*

• "Benign overfitting from KKT conditions for margin maximization"
— Conference on Learning Theory, Bangalore *July 2023*

• "Implicit regularization and benign overfitting for neural networks in high dimensions"
— Youth in High Dimensions Workshop, International Center for Theoretical Physics, Trieste *May 2023*
— UC Berkeley, Department of Biostatistics Seminar *April 2023*
— University of British Columbia, Mathematics of Information, Learning, and Data Seminar *January 2023*
— University of Alberta, Dept. of Mathematical and Statistical Sciences, Statistics Seminar *October 2022*
— EPFL, Fundamentals of Learning and Artificial Intelligence Seminar *September 2022*

• "Statistical and computational phenomena in deep learning"
— University of Wisconsin–Madison, Department of Statistics Seminar *February 2023*
— University of British Columbia, Department of Statistics Seminar *February 2023*
— University of Illinois, Urbana-Champaign, Department of Statistics Seminar *February 2023*
— University of Michigan, Department of Statistics Seminar *January 2023*
— University of British Columbia, Department of Mathematics Colloquium *January 2023*
— Northwestern University, Department of Statistics and Data Science Seminar *January 2023*
— University of California, Davis, Department of Statistics Seminar *January 2023*
— University of California, Davis, Department of Mathematics Colloquium *January 2023*
— Boston University, Department of Mathematics and Statistics Seminar *January 2023*
— University of California, San Diego, Department of Mathematics Colloquium *December 2022*
— University of California, Irvine, Department of Statistics Seminar *December 2022*

• "Statistical learning theory, optimization, and neural networks" (two hour tutorial)
— Simons Institute, Deep Learning Theory Workshop & Summer School *August 2022*

• "Benign overfitting without linearity"
— University of Southern California, Symposium on Frontiers of Machine Learning and AI *November 2022*
— Joint Statistical Meetings, Washington, D.C. *August 2022*
— Conference on Learning Theory, London *July 2022*
— ETH Zürich, Data, Algorithms, Combinatorics, & Optimization Seminar *June 2022*
— Harvard University, Probabilitas Seminar *May 2022*
— University of Toronto, Statistics Research Day *May 2022*
— University of British Columbia, Christos Thrampoulidis group meeting *April 2022*
— Theory of Overparameterized Machine Learning Workshop *April 2022*
— Google Research, Algorithms Seminar *March 2022*
— Oxford University, Yee Whye Teh group meeting *March 2022*
— NSF/Simons Mathematics of Deep Learning Seminar *March 2022*

• "Random feature amplification: Feature learning and generalization in neural networks"
— EPFL, Nicolas Flammarion group meeting *October 2022*
— Microsoft Research, Machine Learning Foundations Seminar *April 2022*
— Columbia University, Daniel Hsu group meeting *April 2022*
— Theory of Overparameterized Machine Learning Workshop *April 2022*

• "Self-training converts weak learners to strong learners in mixture models"
— Simons Institute for the Theory of Computing, Deep Learning Theory Symposium *December 2021*

• "Proxy convexity: A unified optimization framework for neural networks trained by gradient descent"
— Simons Institute for the Theory of Computing, Meet the Fellows Welcome Event *September 2021*

• "Generalization of SGD-trained neural networks in the presence of adversarial label noise"
— ETH Zürich, Young Data Science Researchers Seminar *April 2021*
— Johns Hopkins University, Machine Learning Seminar *April 2021*
— Max-Planck-Institute MiS, Machine Learning Seminar *March 2021*
— NSF/Simons Mathematics of Deep Learning Seminar *February 2021*

PROFESSIONAL SERVICE

• Workshop co-organizer: Deep Learning Theory Workshop & Summer School, Simons Institute for the Theory of Computing, Berkeley, 2022.
• Area Chair/Senior PC for conferences: NeurIPS 2023, ALT 2024.
• Reviewer for journals: *Annals of Statistics*, *Journal of Machine Learning Research*, *SIAM Journal on Mathematics of Data Science*, *Neural Computation*, *Mathematics of Operations Research*, *Transactions on Machine Learning Research*.
• Reviewer for conferences: ICML 2020, NeurIPS 2020, AISTATS 2021, ICML 2021, NeurIPS 2021, ICLR 2022, AISTATS 2022, ICML 2022, ICLR 2023, COLT 2023.
• Reviewer for workshops: Theory of Overparameterized Machine Learning (TOPML) 2021, ICML Workshop on Overparameterization: Pitfalls & Opportunities (ICMLOPPO) 2021, TOPML 2022.
• Reviewer, ENVISION Research Competition for Women in STEM, 2022.
• Volunteer for Queer in AI.

CONFERENCE, WORKSHOP, AND PROGRAM PARTICIPATION

• Conference on Learning Theory. Bangalore, 2023.
• Youth in High Dimensions. Trieste, 2023.
• Joint Statistical Meetings. Washington, DC, 2022.
• Deep Learning Theory Summer Cluster. Simons Institute for the Theory of Computing, Berkeley, 2022.
• Conference on Learning Theory. London, 2022.
• Deep Learning Theory Symposium. Simons Institute for the Theory of Computing, Berkeley, 2021.
• Mathematical/Scientific Foundations of Deep Learning Annual Meeting. Simons Foundation, New York, 2021.
• ICML Workshop on Overparameterization: Pitfalls and Opportunities. Online, 2021.
• Theory of Overparameterized Machine Learning (TOPML) Workshop. Rice University (online), 2021; 2022.
• Theory of Deep Learning Special Quarter. TTIC/Northwestern Institute for Data, Econometrics, Algorithms, and Learning (online) 2020.
• Emerging Challenges in Deep Learning Workshop. Simons Institute, Berkeley, 2019.
• Summer School in Probability. Pacific Institute for the Mathematical Sciences, Vancouver, BC, 2014.

TEACHING EXPERIENCE

• UCLA, Department of Statistics
— TA, Stats 100C: Linear Models, Spring 2020.
— TA, Stats 102C: Monte Carlo Methods, Fall 2019.
— TA, Stats 100B: Mathematical Statistics, Winter 2016.
— TA, Stats 100A: Probability Theory, Fall 2016.
— TA, Stats 10: Intro to Statistics, Summer 2016.

PROGRAMMING LANGUAGES AND SKILLS

Python (including PyTorch & TensorFlow), R, Matlab.