

LIBIN ZHU

Phone: (+1)614-282-0789 | Email: libinzhu@uw.edu | Website: libinzhu.github.io

Academic Experience

University of Washington

Postdoctoral Fellow, Institute for Foundations of Data Science (IFDS)
Advisor: Dmitriy Drusvyatskiy and Maryam Fazel

2024–present

Seattle, WA

Education

University of California, San Diego

Ph.D., Computer Science and Engineering
Advisor: Mikhail Belkin

2024

San Diego, CA

Thesis: Toward Understanding the Dynamics of Over-parameterized Neural Networks

Zhejiang University

B.Sc., Mathematics

2018

Hangzhou, China

Research Interests

Mathematical foundations of machine learning: feature learning, kernel methods, over-parameterized neural networks; and optimization.

Selected Publications

- **L. Zhu**, D. Davis, D. Drusvyatskiy, & M. Fazel (2025). *Iteratively reweighted kernel machines efficiently learn sparse functions*. Submitted to *Foundations of Computational Mathematics*. Available at [arXiv](https://arxiv.org/abs/2501.08811).
- **L. Zhu**, C. Liu, A. Radhakrishnan, & M. Belkin (2024). *Catapults in SGD: spikes in the training loss and their impact on generalization through feature learning*. [ICML](#).
- **L. Zhu**, C. Liu, A. Radhakrishnan, & M. Belkin (2024). *Quadratic models for understanding neural network dynamics*. [ICLR](#).
- **L. Zhu**, C. Liu, & M. Belkin (2022). *Transition to linearity of general neural networks with directed acyclic graph architecture*. [NeurIPS](#).
- C. Liu, **L. Zhu**, & M. Belkin (2022). *Loss landscapes and optimization in over-parameterized non-linear systems and neural networks*. *Applied and Computational Harmonic Analysis* ([ACHA](#)).
- C. Liu, **L. Zhu**, & M. Belkin (2022). *Transition to linearity of wide neural networks is an emerging property of assembling weak models*. [ICLR \(Spotlight\)](#).
- C. Liu, **L. Zhu**, & M. Belkin (2020). *On the linearity of large non-linear models: when and why the tangent kernel is constant*. [NeurIPS \(Spotlight\)](#).

Research Programs

- Princeton Machine Learning Theory Summer School 2022

Awards & Grants

- Google Research Credits Program (\$5,000 equivalent). 2025
- Microsoft Research Credits Program (\$6,000 equivalent). 2025
- Tinker Research Credits Program (\$5,000 equivalent). 2025
- ACCESS-CIS (CIS220009) Computational Award (Co-PI). 2025
- Data Science Research Fellowship, University of Washington. 2024
- Scholarship of Outstanding Students, Zhejiang University. 2015
- Guanghua Educational Scholarship, Zhejiang University. 2014

Invited Talks

- Applied Math, UW 2025
- Pacific Northwest Section of SIAM Biennial Meeting 2025
- IFDS Workshop on “Theoretical Foundations of Applied AI”, UW 2025
- Joint Mathematics Meetings (JMM) 2025
- IFDS seminar, UW 2024
- MoDL Collaboration Workshop, UCSD 2024
- Co-PI Seminars on Optimization, Control, and Learning, UCSD 2024
- Deep Learning and Optimization Seminar, Westlake University 2023
- Andrew Stuart’s Lab, Caltech 2023
- Laboratory for Information & Decision Systems (LIDS), MIT 2022
- INFORMS: Topics in Theory of Neural Networks 2022
- Information Theory and Applications (ITA) Workshop 2022

Academic Service

Organizer

- IFDS weekly seminar, University of Washington 2024–2025
- MoDL Collaboration Workshop 2024

Reviewer

- SIMODS, JMLR, NeurIPS, ICML, ICLR, UAI

Mentorship

- Bhavesh Kumar, Graduate student at UW 2024–present
- Ethan Fang, Undergraduate student at UCSD 2023–2024

Teaching

- Teaching Assistant: DSC 206 — Algorithms in Data Science, UCSD Winter 2024
 - Guest lectures on counting distinct elements in data streaming.
- Teaching Assistant: DSC 240 — Machine Learning, UCSD Fall 2023
 - Guest lectures on Probability theory and Bayes optimal classifier.
- Teaching Assistant: MSRI-UCSD Summer School on Machine Learning Summer 2023
- Teaching Assistant: DSC 291 — Topics in Mathematics of Deep Learning, UCSD Spring 2023
 - Guest lectures on Neural Tangent Kernel and Optimization theory.
- Teaching Assistant: DSC 212 — Probability and Statistics for Data Science, UCSD Winter 2023
 - Guest lectures on Bootstrapping.
- Teaching Assistant: DSC 291 — Topics in Mathematics of Deep Learning, UCSD Spring 2022
- Teaching Assistant: DSC 140A — Probabilistic Modeling and ML, UCSD Winter 2022

All Publications (Reverse Chronological)

Conference proceedings

- N. Mallinar, D. Beaglehole, **L. Zhu**, A. Radhakrishnan, P. Pandit, & M. Belkin (2025). *Emergence in non-neural models: grokking modular arithmetic via average gradient outer product*. ICML (**Oral Presentation**).

- **L. Zhu**, C. Liu, A. Radhakrishnan, & M. Belkin (2024). *Catapults in SGD: spikes in the training loss and their impact on generalization through feature learning*. ICML.
- **L. Zhu**, C. Liu, A. Radhakrishnan, & M. Belkin (2024). *Quadratic models for understanding neural network dynamics*. ICLR.
- A. Banerjee, P. Cisneros-Velarde, **L. Zhu**, & M. Belkin (2023). *Neural tangent kernel at initialization: linear width suffices*. UAI.
- A. Banerjee, P. Cisneros-Velarde, **L. Zhu**, & M. Belkin (2023). *Restricted strong convexity of deep learning models with smooth activations*. ICLR.
- **L. Zhu**, C. Liu, & M. Belkin (2022). *Transition to linearity of general neural networks with directed acyclic graph architecture*. NeurIPS.
- C. Liu, **L. Zhu**, & M. Belkin (2022). *Transition to linearity of wide neural networks is an emerging property of assembling weak models*. ICLR (**Spotlight**).
- C. Liu, **L. Zhu**, & M. Belkin (2020). *On the linearity of large non-linear models: when and why the tangent kernel is constant*. NeurIPS (**Spotlight**).

Journal publications

- C. Liu, **L. Zhu**, & M. Belkin (2022). *Loss landscapes and optimization in over-parameterized non-linear systems and neural networks*. *Applied and Computational Harmonic Analysis*.
- C. Liu, **L. Zhu**, & M. Belkin (2025). *Assembly and iteration: transition to linearity of wide neural networks*. *Applied and Computational Harmonic Analysis*.

Preprints/Articles in Review

- **L. Zhu**, D. Davis, D. Drusvyatskiy, & M. Fazel (2025). *Iteratively reweighted kernel machines efficiently learn sparse functions*. Submitted to *Foundations of Computational Mathematics*.
- **L. Zhu**, D. Davis, D. Drusvyatskiy, & M. Fazel (2025). *Spectral norm bound for the product of random Fourier–Walsh matrices*. Submitted to *Annals of Applied Probability*.
- **L. Zhu**, P. Pandit, & M. Belkin (2022). *A note on linear bottleneck networks and their transition to multilinearity*.

Technical Strengths

- **Languages:** Python, C, C++, PyTorch, TensorFlow, Keras, SQL, MATLAB, R
- **Frameworks & Tools:** Linux, Git, large-scale distributed training

Industry Experience

Meta Platforms, Inc.

Machine Learning Engineer

Jun 2022–Sep 2022

Palo Alto, CA

- Designed and built pipelines to analyze impacts of Ads Xout bid on revenue and negative feedback for Instagram stream.
- Proposed and implemented transformations on Ads Xout bid, resulting in a decrease of over 40% in negative feedback on the Instagram stream while maintaining neutral revenue.