

Denny (Yi) Wu

✉ dennywu@nyu.edu
<https://dennywu1.github.io/>

Employment

2023–
Present **New York University & Flatiron Institute**, New York, NY,
CDS-Flatiron Faculty Fellow.

Education

2018–2023 **University of Toronto & Vector Institute**, Toronto, ON,
Ph.D. in Computer Science.
Supervisor: **Jimmy Ba** and **Murat A. Erdogdu**

2013–2018 **Carnegie Mellon University**, Pittsburgh, PA,
B.Sc. in Computational Biology, with University Honor.

Research Experience

- 2024.09–
Present **xAI**, Palo Alto, CA.
○ Visiting researcher.
- 2024.09–
Present **Simons Institute**, Berkeley, CA.
○ Participant of the *Modern Paradigms in Generalization* program.
- 2021.06–
2021.09 **Microsoft Research**, Redmond, WA.
○ *Deep Learning Group*. Host: Greg Yang.
- 2019.06–
2019.08 **University of Tokyo & RIKEN AIP**, Tokyo, Japan.
○ *Deep Learning Theory Team*. Host: Taiji Suzuki.
- 2017.09–
2018.03 **Carnegie Mellon University**, Pittsburgh, PA.
○ *Undergraduate Research Assistant*. Advisor: Ruslan Salakhutdinov.
- 2017.06–
2017.08 **RIKEN AIP**, Tokyo, Japan.
○ *High-dimensional Statistical Modeling Team*. Host: Makoto Yamada.
- 2016.05–
2016.08 **University of Tokyo**, Tokyo, Japan.
○ *Research Intern (Funded by ISURF)*. Host: Keisuke Goda.
- 2015.06–
2015.08 **Kyoto University**, Kyoto, Japan.
○ *Amgen Scholar*. Host: Tomokazu Umeyama.

Publications (* denotes alphabetical ordering or equal contribution)

Conference Publications

- Jason D. Lee*, Kazusato Oko*, Taiji Suzuki*, and **Denny Wu***. "Neural network learns low-dimensional polynomials near the information-theoretic limit." Advances in Neural Information Processing Systems (**NeurIPS**), 2024.
- Kazusato Oko*, Yujin Song*, Taiji Suzuki*, and **Denny Wu***. "Pretrained transformer efficiently learns low-dimensional target functions in context." Advances in Neural Information Processing Systems (**NeurIPS**), 2024.

- Zhichao Wang, **Denny Wu**, and Zhou Fan. "Nonlinear spiked covariance matrices and signal propagation in deep neural networks." Conference on Learning Theory (**COLT**), 2024.
- Kazusato Oko*, Yujin Song*, Taiji Suzuki*, and **Denny Wu***. "Learning sum of diverse features: computational hardness and efficient gradient-based training for ridge combinations." Conference on Learning Theory (**COLT**), 2024.
- Kazusato Oko, Shunta Akiyama, **Denny Wu**, Toyama Murata, and Taiji Suzuki. "SILVER: single-loop variance reduction and application to federated learning." International Conference on Machine Learning (**ICML**), 2024.
- Atsushi Nitanda*, Kazusato Oko*, Taiji Suzuki*, **Denny Wu***. "Improved statistical and computational complexity of the mean-field Langevin dynamics under structured data." International Conference on Learning Representations (**ICLR**), 2024.
- Atsushi Nitanda, Ryuhei Kikuchi, Shugo Maeda, and **Denny Wu**. "Parameter averaging for SGD stabilizes the implicit bias towards flat regions." International Conference on Artificial Intelligence and Statistics (**AISTATS**), 2024.
- Alireza Mousavi-Hosseini, **Denny Wu**, Taiji Suzuki, and Murat A. Erdogdu. "Gradient-based feature learning under structured data." Advances in Neural Information Processing Systems (**NeurIPS**), 2023.
- Jimmy Ba*, Murat A. Erdogdu*, Taiji Suzuki*, Zhichao Wang*, and **Denny Wu***. "Learning in the presence of low-dimensional structure: a spiked random matrix perspective." Advances in Neural Information Processing Systems (**NeurIPS**), 2023.
- Taiji Suzuki, **Denny Wu**, Atsushi Nitanda, and Kazusato Oko. "Feature learning via mean-field Langevin dynamics: classifying sparse parities and beyond." Advances in Neural Information Processing Systems (**NeurIPS**), 2023.
- Taiji Suzuki, **Denny Wu**, and Atsushi Nitanda. "Convergence of mean-field Langevin dynamics: time and space discretization, stochastic gradient, and variance reduction." Advances in Neural Information Processing Systems (**NeurIPS**), 2023. (**Spotlight**)
- Atsushi Nitanda, Kazusato Oko, **Denny Wu**, Nobuhito Takenouchi, and Taiji Suzuki. "Primal and dual analysis of entropic fictitious play for finite-sum problems." International Conference on Machine Learning (**ICML**), 2023.
- Taiji Suzuki, Atsushi Nitanda, and **Denny Wu**. "Uniform-in-time propagation of chaos for the mean-field Langevin dynamics." International Conference on Learning Representations (**ICLR**), 2023.
- Jimmy Ba*, Murat A. Erdogdu*, Taiji Suzuki*, Zhichao Wang*, **Denny Wu***, and Greg Yang*. "High-dimensional asymptotics of feature learning: how one gradient step improves the representation." Advances in Neural Information Processing Systems (**NeurIPS**), 2022.
- Naoki Nishikawa, Taiji Suzuki, Atsushi Nitanda, and **Denny Wu**. "Two-layer neural network on infinite-dimensional data: global optimization guarantee in the mean-field regime." Advances in Neural Information Processing Systems (**NeurIPS**), 2022.
- Atsushi Nitanda, **Denny Wu**, and Taiji Suzuki. "Convex analysis of the mean-field Langevin dynamics." International Conference on Artificial Intelligence and Statistics (**AISTATS**), 2022.
- Kazusato Oko, Taiji Suzuki, Atsushi Nitanda, and **Denny Wu**. "Particle stochastic dual coordinate ascent: exponential convergent algorithm for mean-field neural network optimization." International Conference on Learning Representations (**ICLR**), 2022.
- Jimmy Ba*, Murat A. Erdogdu*, Marzyeh Ghassemi*, Taiji Suzuki*, Shengyang Sun*, **Denny Wu***, and Tianzong Zhang*. "Understanding the variance collapse of SVGD in high dimensions." International Conference on Learning Representations (**ICLR**), 2022.
- Atsushi Nitanda, **Denny Wu**, and Taiji Suzuki. "Particle dual averaging: optimization of mean-field neural networks with global convergence rate analysis." Advances in Neural Information Processing Systems (**NeurIPS**), 2021.

- Shun-ichi Amari*, Jimmy Ba*, Roger Grosse*, Xuechen Li*, Atsushi Nitanda*, Taiji Suzuki*, **Denny Wu***, and Ji Xu*. "When does preconditioning help or hurt generalization?" International Conference on Learning Representations (**ICLR**), 2021.
[*Optimization for Machine Learning Workshop*, NeurIPS 2020. (**Best Student Paper**)]
- **Denny Wu*** and Ji Xu*. "On the optimal weighted ℓ_2 regularization in overparameterized linear regression." Advances in Neural Information Processing Systems (**NeurIPS**), 2020.
- Jimmy Ba*, Murat A. Erdogdu*, Taiji Suzuki*, **Denny Wu***, and Tianzong Zhang*. "Generalization of two-layer neural networks: an asymptotic viewpoint." International Conference on Learning Representations (**ICLR**), 2020. (**Spotlight**)
- Xuechen Li, **Denny Wu**, Lester Mackey, and Murat A. Erdogdu. "Stochastic runge-kutta accelerates Langevin monte carlo and beyond." Advances in Neural Information Processing Systems (**NeurIPS**), 2019. (**Spotlight**)
- Makoto Yamada*, **Denny Wu***, Yao-Hung Hubert Tsai, Ichiro Takeuchi, Ruslan Salakhutdinov, and Kenji Fukumizu. "Post selection inference with incomplete maximum mean discrepancy estimator." International Conference on Learning Representations (**ICLR**), 2019.

Journal Publications

- Naoki Nishikawa, Taiji Suzuki, Atsushi Nitanda, and **Denny Wu**. "Two-layer neural network on infinite-dimensional data: global optimization guarantee in the mean-field regime." **Journal of Statistical Mechanics: Theory and Experiment (invited)**, no. 11: 114007, 2023.
- Atsushi Nitanda, **Denny Wu**, and Taiji Suzuki. "Particle dual averaging: optimization of mean-field neural networks with global convergence rate analysis." **Journal of Statistical Mechanics: Theory and Experiment (invited)**, no. 11: 114010, 2022.
- Hirofumi Kobayashi, Cheng Lei, **Yi Wu**, Chun-Jung Huang, Atsushi Yasumoto, Masahiro Jona, Wenxuan Li et al. "Intelligent whole-blood imaging flow cytometry for simple, rapid, and cost-effective drug-susceptibility testing of leukemia." **Lab on a Chip** 19, no. 16: 2688-2698, 2019.
- Cheng Lei, Hirofumi Kobayashi, **Yi Wu**, Ming Li, Akihiro Isozaki, Atsushi Yasumoto, Hideharu Mikami et al. "High-throughput imaging flow cytometry by optofluidic time-stretch microscopy." **Nature Protocols** 13, no. 7: 1603-1631, 2018.
- Baoshan Guo, Cheng Lei, **Yi Wu**, Hirofumi Kobayashi, Takuro Ito, Yaxiaer Yalikun, Sangwook Lee et al. "Optofluidic time-stretch quantitative phase microscopy." **Methods** 136: 116-125, 2018.
- Hirofumi Kobayashi, Cheng Lei, **Yi Wu**, Ailin Mao, Yiyue Jiang, Baoshan Guo, Yasuyuki Ozeki, and Keisuke Goda. "Label-free detection of cellular drug responses by high-throughput bright-field imaging and machine learning." **Scientific Reports** 7, no. 1: 1-9, 2017.
- Cheng Lei, **Yi Wu**, Aswin C. Sankaranarayanan, Shih-Min Chang, Baoshan Guo, Naoto Sasaki, Hirofumi Kobayashi, Chia-Wei Sun, Yasuyuki Ozeki, and Keisuke Goda. "GHz optical time-stretch microscopy by compressive sensing." **IEEE Photonics Journal** 9, no. 2: 1-8, 2017.
- Rachel Ferebee, Ilhem F. Hakem, Amelie Koch, Maggie Chen, **Yi Wu**, Derek Loh, David C. Wilson et al. "Light scattering analysis of mono- and multi-pegylated bovine serum albumin: Role of composition on structure and interactions." **The Journal of Physical Chemistry B** 120, no. 20: 4591-4599, 2016.

Preprints & Workshop Presentations

- Gerard Ben Arous*, Murat A. Erdogdu*, Nuri Mert Vural*, and **Denny Wu***. "Learning quadratic neural networks with SGD in high dimensions." *In preparation*.
- Wei Huang, Andi Han, Yujin Song, Yilan Chen, **Denny Wu**, Difan Zou, and Taiji Suzuki. "Label noise gradient descent improves generalization in the low SNR regime." *In preparation*.
- Alireza Mousavi-Hosseini, **Denny Wu**, and Murat A. Erdogdu. "Learning multi-index models with neural networks via mean-field Langevin dynamics." *In preparation*.

- **Denny Wu**, Hirofumi Kobayashi, Charles Ding, Lei Cheng, Keisuke Goda, and Marzyeh Ghassemi. "Modeling the biological pathology continuum with HSIC-regularized wasserstein auto-encoders." Machine Learning for Health (ML4H) Workshop, NeurIPS 2018.
- Cheng Lei, Baoshan Guo, Yiyue Jiang, **Yi Wu**, Hirofumi Kobayashi, Takuro Ito, Atsushi Yasumoto, Yutaka Yatomi, Yasuyuki Ozeki, and Keisuke Goda. "High-throughput, label-free, multivariate cell analysis with optofluidic time-stretch microscopy." Conference on Lasers and Electro-Optics, 2017.

Professional Service

Conferences

Area Chair AISTATS.

Reviewer NeurIPS, ICLR, COLT, ICML, AISTATS.

Journals

Reviewer *Annals of Statistics, Information and Inference: A Journal of the IMA, SIAM Journal on Mathematics of Data Science, Journal of Machine Learning Research, Transaction of Machine Learning Research, Journal of Computational and Graphical Statistics.*

Awards & Invited Talks

Awards & Fellowships

2023 *UChicago Rising Stars in Data Science.*

2023 *DCS Departmental Fellowship.*

2021–2023 *Borealis AI Fellowship.*

2019–2022 *Vector Institute Research Grant.*

Selected Talks

- 10/2024 *LTV Talk at the Modern Paradigms in Generalization*, Simons Institute.
- 06/2024 *DIMACS Workshop on Modeling Randomness in Neural Network Training*, Rutgers University.
- 05/2024 *Lausanne Event on Machine Learning and Neural Network Theory*, EPFL.
- 05/2024 *Youth in High Dimensions: Recent Progress in Machine Learning, High-Dimensional Statistics and Inference*, ICTP Trieste.
- 01/2024 *CD3 x Simons Foundation workshop "AI-driven discovery in physics and astrophysics"*, The University of Tokyo.
- 01/2024 *The Mathematics of Data Workshop*, The National University of Singapore.
- 12/2023 *IMA Data Science Seminar*, The University of Minnesota.
- 11/2023 *IMS Young Mathematical Scientist Forum*, The National University of Singapore.
- 10/2023 *CDS Lunch Seminar*, New York University.
- 05/2023 *CFAR Outstanding PhD Student Seminar*, The Agency for Science, Technology and Research.
- 05/2023 *Department of Statistics and Data Science Seminar*, The National University of Singapore.
- 04/2023 *CIRJE Applied Statistics Workshop*, The University of Tokyo.
- 03/2023 *Department of Statistics Seminar*, The Chinese University of Hong Kong.
- 03/2023 *AI Research Talk*, Mohamed bin Zayed University of Artificial Intelligence.
- 11/2022 *The 25th Information-Based Induction Sciences Workshop (IBIS2022)*, Tsukuba International Conference Center.
- 09/2022 *One World Seminar Series on the Mathematics of Machine Learning*, Online.
- 06/2022 *Young Data Science Researcher Seminar Zurich*, ETH Zurich.

- 04/2022 *Workshop on Advances in Stein's method and Applications in Machine Learning and Optimization*, Banff International Research Station.
- 09/2021 *MSR Machine Learning Foundations Seminar*, Microsoft Research Redmond.
- 03/2021 *Workshop on Functional Inference and Machine Intelligence (FIMI)*, Online.
- 11/2020 *Workshop on Seeking Low-dimensionality in Deep Neural Networks*, Online.

Teaching

Course Instructor

Spring 2025 *DS-GA 1010: Capstone Continuation*, New York University.

Fall 2023 *DS-GA 1014: Optimization and Computational Linear Algebra*, New York University.

Teaching Assistant

Winter 2023 *CSC413: Neural Networks and Deep Learning*, University of Toronto.

Winter 2022

Winter 2020

Fall 2022 *Machine Learning and AI – Black & Indigenous Program*, Vector Institute.

Fall 2018 *CSC411: Introduction to Machine Learning*, University of Toronto.

Fall 2017 *80-257: Nietzsche*, Carnegie Mellon University.