Yijun Dong

Updated September 20, 2025

✉ *yd1319@nyu.edu*
dyjdongyijun.github.io

Research Interests

Numerical Linear Algebra, **Probability Theory**, **Machine Learning Theory**, **Statistics**.

Employment

2023-Present **New York University, Courant Institute of Mathematical Sciences**, New York, NY,
Assistant Professor/Courant Instructor (fixed-term).

Education

2018-2023 **University of Texas at Austin, Oden Institute**, Austin, TX,
Ph.D. in Computational Science, Engineering, and Mathematics (CSEM).
- Advisors: **Per-Gunnar Martinsson**, **Rachel Ward**.
- Thesis: Randomized Dimension Reduction with Statistical Guarantees.

2014-2018 **Emory University**, Atlanta, GA,
B.S. in Applied Mathematics & Engineering Science, Magna Cum Laude.
- Advisors: **Effrosyni Seitaridou**, **Eric Weeks**.
- Thesis: Crystals and Liquids in Gravitationally Confined Quasi-2D Colloidal Systems.

Publications (∗ denotes equal contribution or alphabetical order)

Conference Publications

- Yijun Dong, Yicheng Li, Yunai Li, Jason D. Lee, Qi Lei. "Discrepancies are Virtue: Weak-to-Strong Generalization through Lens of Intrinsic Dimension." *International Conference on Machine Learning (ICML)*, 2025.
- Jianwei Li, Yijun Dong, Qi Lei. "Greedy Output Approximation: Towards Efficient Structured Pruning for LLMs Without Retraining." *Conference on Parsimony and Learning (CPAL)*, 2025.
- Yijun Dong*, Hoang Phan*, Xiang Pan*, Qi Lei. "Sketchy Moment Matching: Toward Fast and Provable Data Selection for Finetuning." *Conference on Neural Information Processing Systems (NeurIPS)*, 2024.
- Yijun Dong*, Kevin Miller*, Qi Lei, Rachel Ward. "Cluster-aware Semi-supervised Learning: Relational Knowledge Distillation Provably Learns Clustering." *Conference on Neural Information Processing Systems (NeurIPS)*, 2023.
- Yijun Dong*, Yuege Xie*, Rachel Ward. "Adaptively Weighted Data Augmentation Consistency Regularization for Robust Optimization under Concept Shift." *International Conference on Machine Learning (ICML)*, 2023.
- Shuo Yang*, Yijun Dong*, Rachel Ward, Inderjit S Dhillon, Sujay Sanghavi, Qi Lei. "Sample Efficiency of Data Augmentation Consistency Regularization." *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2023.

Journal Publications

- Yijun Dong, Chao Chen, Per-Gunnar Martinsson, Katherine Pearce. "Robust Blockwise Random Pivoting: Fast and Accurate Adaptive Interpolative Decomposition." *SIAM Journal on Matrix Analysis and Applications*, 2025.
- Katherine Pearce, Chao Chen, Yijun Dong, Per-Gunnar Martinsson. "Adaptive Parallelizable Algorithms for Interpolative Decompositions via Partially Pivoted LU." *Numerical Linear Algebra with Applications*, 2025.
- Yijun Dong, Per-Gunnar Martinsson, Yuji Nakatsukasa. "Efficient Bounds and Estimates for Canonical Angles in Randomized Subspace Approximations." *SIAM Journal on Matrix Analysis and Applications*, 2024.
- Yijun Dong, Per-Gunnar Martinsson. "Simpler is better: A comparative study of randomized algorithms for computing the CUR decomposition." *Advances in Computational Mathematics*, 2023.
- Chen Cheng*, Yijun Dong*, Matthew Dorian*, Farhan Kamili*, Effrosyni Seitaridou. "Quantifying Biofilm Formation of *Sinorhizobium meliloti* Bacterial Strains in Microfluidic Platforms by Measuring the Diffusion Coefficient of Polystyrene Beads." *Open Journal of Biophysics*, 2017.

Preprints & Other Publications

- Zihan Wang, Yijun Dong, Qi Lei. "When does Chain-of-Thought Help? A Markovian Perspective." *Manuscript in preparation*, 2025.
- Chenruo Liu*, Yijun Dong*, Qi Lei. "Does Weak-to-strong Generalization Happen under Spurious Correlation?" *Manuscript in preparation*, 2025.
- Hoang Phan, Yijun Dong, Andrew Gordon Wilson, Qi Lei. "Balanced Locality-sensitive Hashing for Online Data Selection." *Manuscript in preparation*, 2025.
- Yijun Dong, Paul Schwerdtner, Benjamin Peherstorfer. "Randomized Time Stepping of Nonlinearly Parametrized Solutions of Evolution Problems." *Manuscript in preparation, submitted to FoCM*, 2025.
- Yijun Dong*, Xiang Pan*, Hoang Phan*, Qi Lei. "Randomly Pivoted V-optimal Design: Fast Data Selection under Low Intrinsic Dimension." *Workshop on Machine Learning and Compression @ NeurIPS*, 2024.

Teaching Experience

2023-Present **Instructor**, Courant Institute, New York University, New York, NY.

- Fall 2025: Probability and Statistics (undergraduate).
- Spring 2025: Introduction to Mathematical Modeling (undergraduate).
- Fall 2024: Computational Statistics (graduate).
- Spring 2024: Mathematics for Economics (undergraduate).
- Fall 2023: Discrete Mathematics (undergraduate).

2023/07 **Teaching Assistant**, Simons Laufer Mathematical Sciences Institute (SLMath) Summer Graduate School, IBM Almaden, San Jose, CA.

- Mathematics of Big Data: Sketching and (Multi-) Linear Algebra (TA for Kenneth Clarkson, Lior Horesh, Misha Kilmer, Tamara Kolda, and Shashanka Ubaru).

2020-2022 **Teaching Assistant**, UT Austin, Austin, TX.
- Fall 2022: Differential Equations with Linear Algebra (TA for Michael Novack).
- Fall 2021: Numerical Analysis: Linear Algebra (TA for Per-Gunnar Martinsson).
- Fall 2020: Differential Equations with Linear Algebra (TA for Sam Raskin).

2015-2016 **Student Tutor**, Emory University, Oxford, GA.
- Introduction to Physics, Modern Physics.

Industrial Experience

2022/06-2022/08 **Dell Technologies**, Research Intern, Austin, TX.
- Semi-supervised tabular learning with data augmentation and consistency regularization.

2021/05-2021/08 **Dell Technologies**, Research Intern, Austin, TX.
- Streaming telemetry time series compression on edge devices.

Academic Visits

Invited Talks

2025/11 SIAM New York-New Jersey-Pennsylvania (NNP) Sectional Meeting 2025 minisymposium on "Recent advances in low-rank methods and their applications", Pennsylvania State University, University Park, PA.

2025/05 Flatiron Institute Center for Computational Mathematics (CCM) Machine Learning Theory Seminar, New York, NY.

2025/04 John Hopkins University Postdoc Seminar, Virtual.

2025/03 UT Austin Data & Algebra seminar, Virtual.

2025/03 SIAM Conference on Computational Science and Engineering (CSE25) minisymposium on "Scientific Machine Learning for Stable Prediction of Dynamical Systems", Fort Worth, TX.

2025/01 Joint Mathematics Meetings (JMM 2025) ILAS Special Session on Randomness in Numerical Linear Algebra, Seattle, WA.

2024/10 University of Delaware Numerical Analysis & PDE seminar, Newark, DE.

2024/03 SIAM Conference on Parallel Processing for Scientific Computing (PP24) minisymposium on "Randomized Methods in Linear Solvers and Matrix Factorizations", Baltimore, MD.

2023/11 SIAM Texas-Louisiana Section Meeting 2023 minisymposium on "Nonlinear Algebra in Applications", Lafayette, LA.

2023/08 International Council for Industrial and Applied Mathematics (ICIAM) minisymposium on "Randomized Numerical Linear Algebra", Tokyo, Japan.

2023/04 Rising Stars Workshop in Computational and Data Sciences, Austin, TX.

2022/01 Jane Street Symposium 2022, Virtual.

Posters & Contributed Talks

2025/07 ICML 2025, Vancouver, Canada. (poster)

2025/06 Householder Symposium XXII, Ithaca, NY. (poster)

2024/12 NeurIPS 2024, Vancouver, Canada. (poster)

2024/10 SIAM Conference on Mathematics of Data Science (MDS24) minisymposium on "Efficient Computation and Learning with Randomized Sampling and Pruning", Atlanta, GA. (poster)

2023/12 NeurIPS 2023, New Orleans, LA. (poster)

2023/07 ICML 2023, Honolulu, HI. (poster)

2023/04 Institute for Foundations of Machine Learning (IFML) Workshop 2023, Seattle, WA. (contributed)

2023/03 Texas Women in Math Symposium, Austin, TX. (contributed)

2022/11 Institute for Pure & Applied Mathematics (IPAM) Workshop IV: Multi-Modal Imaging with Deep Learning and Modeling, University of California, Los Angeles, CA. (poster)

2022/10 Oden Institute CSEM Student Forum, Austin, TX. (contributed)

2022/09 SIAM Conference on Mathematics of Data Science (MDS22), San Diego, CA. (poster)

2021/05 SIAM Conference on Applied Linear Algebra (LA21), Virtual. (poster)

2018/03 American Physical Society (APS) March Meeting, Los Angeles, CA. (contributed)

Other Academic Visits

2026/02 Institute for Computational and Experimental Research in Mathematics (ICERM) Workshop on Randomized Numerical Linear Algebra, Brown University, Providence, RI.

2025/08 Institute for Pure & Applied Mathematics (IPAM) Research Collaboration Workshop on Randomized Numerical Linear Algebra, UCLA, Los Angeles, CA.

Student Advising

Undergraduate Research

2024-2025 Yansong Zhang (NYU), recipient of the DURF grant.

2024-2025 Yicheng Li (NYU Shanghai → MS at UPenn), co-advised with Qi Lei.

2024-2025 Yunai Li (SJTU → PhD at Northwestern), co-advised with Qi Lei.

Academic Service

Reviewer
- Journals: Annals of Applied Probability, BIT Numerical Mathematics, Calcolo, IEEE Transactions on Signal Processing, IMA Journal of Numerical Analysis, Journal of Computational Mathematics and Data Science, Linear Algebra and Its Applications, Numerical Algorithms, SIAM Journal on Matrix Analysis and Applications, SIAM Journal on Scientific Computing.
- Conferences: Conference on Artificial Intelligence and Statistics (AISTATS), Conference on Neural Information Processing Systems (NeurIPS), International Conference on Learning Representations (ICLR), International Conference on Machine Learning (ICML), Learning for Dynamics & Control Conference (L4DC).

Organizer
- SIAM MDS24 minisymposium on "Efficient Computation and Learning with Randomized Sampling and Pruning", with Yifan Chen and Qi Lei.

Distinctions & Fellowships

2025 AMS-Simons Travel Grant.

2023 Graduate School Summer Fellowship, UT Austin.
2023 Rising Stars in Computational and Data Sciences.
2019 NIMS Graduate Fellowship, UT Austin.
2018 Peter O'Donnell Graduate Fellowship, UT Austin.
2018 Trevor Evans Award, Emory University.
2017 Phi Beta Kappa & Sigma Pi Sigma, Emory University.
2016 SURE Summer Research Fellowship, Emory University.
2016 Dan C. Moore Mathematics Award, Emory University.
2016 Williams Baird Physics Award, Emory University.

Travel Awards (for individual conferences)

2025 Householder XXII Early Career Travel Support.
2024 SIAM Early Career Travel Award for PP24.
2022 Professional Development Award, UT Austin.

Skills

Computer
- LaTeX, Python, MATLAB, Java, C++.

Language
- Chinese (native), English (fluent), Japanese (elementary).