# Yuling Yan

*Address:* 7277 MSC, 1300 University Ave, Madison, WI 53706
*Homepage:* https://yulingy.github.io
*Email:* yuling.yan@wisc.edu

## APPOINTMENTS

**University of Wisconsin-Madison**, Madison, WI                    *Aug. 2024 - Now*

- Assistant Professor, Department of Statistics

## EDUCATION

**Massachusetts Institute of Technology**, Cambridge, MA           *July. 2023 - Aug 2024*

- Norbert Wiener Postdoctoral Associate
- Mentors: Philippe Rigollet and Martin J. Wainwright

**Princeton University**, Princeton, NJ                            *Sept. 2018 - June 2023*

- Ph.D. in Operations Research and Financial Engineering (M.A. awarded in Sept 2020)
- Advisors: Yuxin Chen and Jianqing Fan

**Peking University**, Beijing, China                             *Sept. 2014 - July 2018*

- B.S. in Statistics, School of Mathematical Sciences

## RESEARCH INTERESTS

Statistics, optimization, diffusion models, reinforcement learning, and their applications in generative AI and social sciences. A few highlights of my recent research:

- **Learning in the space of probability measures.** We developed the first efficient algorithm for solving the nonparametric MLE of Gaussian mixture models [J2.]. We established state-of-the-art convergence theory for diffusion generative models [C1.] and its adaptivity to unknown low-dimensional structures [C2., P8.].
- **Inference and uncertainty quantification for low-rank models.** We solved open problems on optimal construction of confidence intervals for noisy matrix completion [J8., P3.], and use it to evaluate the effects of *Affordable Care Act* across U.S. states [P4.]. We developed the first inferential algorithm for PCA [J3.].
- **Bridging convex and nonconvex optimization.** We settled minimax optimality of convex optimization for noisy matrix completion [J9.], robust PCA [J7.], and blind deconvolution [J6.] by establishing the striking connection that *nonconvex optimization is essentially solving convex relaxation with nuclear norm penalization*.
- **Mechanism design for peer review.** Our *isotonic mechanism* provably elicit private information from the authors to improve peer review process in machine learning conferences [J1.], and the improved efficiency is confirmed by *an experiment conducted during ICML 2023* [P2.].
- **Offline reinforcement learning.** We designed sample-optimal offline RL algorithms in both single- and multi-agent setups [J5., J4.], as well as minimax optimal reward-agnostic exploration schemes [C3.].

## AWARDS

| | | |
|---|---|---|
| **IMS Lawrence D. Brown Award**, Institute of Mathematical Statistics | | *2024* |
| **ICCM Best Thesis Award**, International Consortium of Chinese Mathematicians | | *2023* |
| — *declined due to inability to attend the award ceremony* | | |
| **Norbert Wiener Postdoctoral Fellowship**, Massachusetts Institute of Technology | | *2023* |
| IMS Hannan Graduate Student Travel Award, Institute of Mathematical Statistics | | *2023* |
| **Charlotte Elizabeth Procter Honorific Fellowship**, Princeton University | | *2022* |
| **Best Student Paper Award** (Statistical Learning & Data Science), American Statistical Association | | *2022* |
| Excellent Graduate, Peking University | | *2018* |
| Yang Xin Lotus Virtue Fellowship, Peking University | | *2017* |
| The May 4th Fellowship, Peking University | | *2016* |
| Samsung Fellowship, No. 2 High School Affiliated to East China Normal University | | *2013* |

# Journal Publications

J1. **Yuling Yan**, Weijie J. Su, Jianqing Fan, "Isotonic Mechanism for Exponential Family Estimation in Machine Learning Peer Review," *Journal of the Royal Statistical Society: Series B*, in press, 2025+.

J2. **Yuling Yan\***, Kaizheng Wang\*, Philippe Rigollet (\* = **equal contribution**), "Learning Gaussian Mixtures Using Wasserstein-Fisher-Rao Gradient Flow," *Annals of Statistics*, vol. 52, no. 4, pp. 1774-1795, 2024.

J3. **Yuling Yan**, Yuxin Chen, Jianqing Fan, "Inference for Heteroskedastic PCA with Missing Data," *Annals of Statistics*, vol. 52, no. 2, pp. 729-756, 2024.

J4. **Yuling Yan**, Gen Li, Yuxin Chen, Jianqing Fan, "Model-based Reinforcement Learning for Offline Zero-Sum Markov Games," *Operations Research*, vol. 72, no. 6, pp. 2430-2445, 2024.

J5. **Yuling Yan**, Gen Li, Yuxin Chen, Jianqing Fan, "The Efficacy of Pessimism in Asynchronous Q-Learning," *IEEE Transactions on Information Theory*, vol. 69, no. 11, pp. 7185-7219, 2023.

J6. Yuxin Chen, Jianqing Fan, Bingyan Wang, **Yuling Yan (alphabetical order)**, "Convex and Nonconvex Optimization Are Both Minimax-Optimal for Noisy Blind Deconvolution under Random Designs," *Journal of the American Statistical Association*, vol. 118, no. 542, pp. 858-868, 2023.

J7. Yuxin Chen, Jianqing Fan, Cong Ma, **Yuling Yan (alphabetical order)**, "Bridging Convex and Nonconvex Optimization in Robust PCA: Noise, Outliers, and Missing Data," *Annals of Statistics*, vol. 49, no. 5, pp. 2948-2971, 2021.

J8. Yuxin Chen, Jianqing Fan, Cong Ma, **Yuling Yan (alphabetical order)**, "Inference and Uncertainty Quantification for Noisy Matrix Completion," *Proceedings of the National Academy of Science*, vol. 116, no. 46, pp. 22931-22937, 2019.

J9. Yuxin Chen, Yuejie Chi, Jianqing Fan, Cong Ma, **Yuling Yan (alphabetical order)**, "Noisy Matrix Completion: Understanding Statistical Guarantees for Convex Relaxation via Nonconvex Optimization," *SIAM Journal on Optimization*, vol. 30, no. 4, pp. 3098-3121, 2020.

J10. **Yuling Yan**, Bret Hanlon, Sebastien Roch, Karl Rohe, "Asymptotic Seed Bias in Respondent-driven Sampling," *Electronic Journal of Statistics*, vol. 14, no. 1, pp. 1577-1610, 2020.

# Conference Publications

C1. Gen Li\*, **Yuling Yan\*** (\* = **equal contribution**), "$O(d/T)$ Convergence Theory for Diffusion Probabilistic Models under Minimal Assumptions," *International Conference on Learning Representations (ICLR)*, 2025.

C2. Gen Li\*, **Yuling Yan\*** (\* = **equal contribution**), "Adapting to Unknown Low-Dimensional Structures in Score-Based Diffusion Models," *Neural Information Processing Systems (NeurIPS)*, 2024.

C3. Gen Li, **Yuling Yan**, Yuxin Chen, Jianqing Fan, "Minimax-Optimal Reward-Agnostic Exploration in Reinforcement Learning," *Conference on Learning Theory (COLT)*, 2024.

C4. Bingyan Wang\*, **Yuling Yan\***, Jianqing Fan (\* = **equal contribution**), "Sample-Efficient Reinforcement Learning for Linearly-Parameterized MDPs with a Generative Model," *Neural Information Processing Systems (NeurIPS)*, 2021.

C5. Kaizheng Wang, **Yuling Yan**, Mateo Díaz, "Efficient Clustering for Stretched Mixtures: Landscape and Optimality," *Neural Information Processing Systems (NeurIPS)*, 2020.

# Preprints

P1. Gen Li, **Yuling Yan**, Yuxin Chen, Jianqing Fan, "Minimax-Optimal Reward-Agnostic Exploration in Reinforcement Learning," arXiv preprint arXiv:2304.07278, under major revision at *Mathematics of Operations Research*, 2025 (accepted in part to COLT 2024).

P2. Buxin Su, Jiayao Zhang, Natalie Collina, **Yuling Yan**, Didong Li, Kyunghyun Cho, Jianqing Fan, Aaron Roth, Weijie J. Su, "Analysis of the ICML 2023 Ranking Data: Can Authors' Opinions of Their Own Papers Assist Peer Review in Machine Learning?" arXiv preprint arXiv:2408.13430, under revision at *Journal of the American Statistical Association*, 2025.

P3. **Yuling Yan**, Martin J. Wainwright, "Entrywise Inference for Causal Panel Data: A Simple and Instance-Optimal Approach," arXiv preprint arXiv:2401.13665, 2024.

P4. Eric Z. Xia\*, **Yuling Yan\***, Martin J. Wainwright (\* = **equal contribution**), "Inference in staggered adoption: Effects of the Affordable Care Act," arXiv preprint arXiv:2412.09482, 2024.

P5. Gen Li\*, **Yuling Yan**\* (∗ = **equal contribution**), "$O(d/T)$ Convergence Theory for Diffusion Probabilistic Models under Minimal Assumptions," arXiv preprint arXiv:2409.18959, 2024 (accepted in part to ICLR 2025).

P6. Gen Li\*, **Yuling Yan**\* (∗ = **equal contribution**), "A Score-Based Density Formula, with Applications in Diffusion Generative Models," arXiv preprint arXiv:2408.16765, 2024.

P7. Jianqing Fan, **Yuling Yan**, Yuheng Zheng **(alphabetical order)**, "When Can Weak Latent Factors Be Statistically Inferred?" arXiv preprint arXiv:2407.03616, 2024.

P8. Jiaqi Tang, **Yuling Yan**, "Adaptivity and Convergence of Probability Flow ODEs in Diffusion Generative Models," arXiv preprint arXiv:2501.18863, 2025.

# INVITED TALKS

T1. Towards Mathematical Foundations of Score-Based Diffusion Models
- Department Seminar, Department of Statistics and Data Science, Cornell University, Nov 2024.
- Institute for Foundations of Data Science, University of Wisconsin-Madison, Oct 2024.

T2. Entrywise Inference for Causal Panel Data: A Simple and Instance-Optimal Approach.
- Department Seminar, Mitch Daniel School of Business, Purdue University, Apr 2025
- SDSCon 2024, Massachusetts Institute of Technology, Apr 2024.

T3. Learning Gaussian Mixtures Using the Wasserstein-Fisher-Rao Gradient Flow.
- Bernoulli-IMS World Congress in Probability and Statistics, Bochum, Germany, Aug 2024.
- ICORS meets DSSV 2024, George Mason University, July 2024.

T4. The Isotonic Mechanism for Exponential Family Estimation.
- CMStatistics, King's College London, Dec 2024.
- Joint Statistical Meeting, Portland, OR, Aug 2024.
- ICSA Applied Statistics Symposium, Ann Arbor, MI, Jun 2023.

T5. Inference and Uncertainty Quantification for Low-Rank Models.
- TOPS Department Seminar, NYU Stern School of Business, Feb 2023.
- Econometrics and Statistics Workshop, University of Chicago Booth School of Business, Feb 2023.
- Department Seminar, Department of Statistics, Rutgers University, Jan 2023.
- Department Seminar, Department of Statistics, University of Wisconsin-Madison, Jan 2023.
- Department Seminar, Department of Statistics, UC Davis, Jan 2023.
- Young Data Science Researcher Seminar, ETH Zurich, Dec 2022.
- Department Seminar, Department of Data Science and Operations, Marshall School of Business, University of Southern California, Dec 2022.
- CMStatistics, King's College London, Dec 2022.
- Department Colloquia, Department of Mathematics, University of Maryland, Nov 2022.

T6. Model-based Reinforcement Learning for Offline Zero-Sum Markov Games.
- INFORMS Annual Meeting, Indianapolis, IN, Oct 2022.
- ORIE Young Researchers Workshop, Cornell University, Oct 2022.

T7. Inference for Heteroskedastic PCA with Missing Data.
- INFORMS Annual Meeting, Indianapolis, IN, Oct 2022.
- Lunch Seminar, Department of Statistics, Harvard University, Oct 2022.
- Joint Statistical Meeting, Washington DC, Aug 2022.

T8. Bridging Convex and Nonconvex Optimization in Robust PCA: Noise, Outliers, and Missing Data.
- INFORMS Optimization Society Meeting, Greenville, SC, Mar 2022.
- Department Seminar, Department of Statistics and Applied Probability, UC Santa Barbara, Mar 2022.
- Two Sigma Ph.D. Research Symposium, Dec 2020.
- INFORMS Annual Meeting, Nov. 2020.
- ORFE Graduate Student Seminar, Princeton University, Oct 2020.

# TEACHING EXPERIENCE

- STAT 615: Statistical Learning, University of Wisconsin-Madison, Spring 2025, Instructor.
- STAT 615: Statistical Learning, University of Wisconsin-Madison, Fall 2024, Instructor.
- ORF 309: Probability and Stochastic Systems, Princeton University, Spring 2021, Teaching Assistant.
- ELE 522: Large-Scale Optimization for Data Science, Princeton University, Fall 2019, Teaching Assistant.

# PROFESSIONAL SERVICE

- Journal reviewer for *Annals of Statistics, Journal of the Royal Statistical Society: Series B, Journal of the American Statistical Association, Biometrika, Management Science, Operations Research, IEEE Transactions on Information Theory, Information and Inference: A Journal of the IMA, Journal of Business and Economic Statistics, IEEE Transactions on Signal Processing, Journal of Machine Learning Research, Journal of Multivariate Analysis, Journal of Computational and Graphical Statistics, IEEE Transactions on Computational Imaging, IEEE Signal Processing Letters, Journal of Nonparametric Statistics*.
- Conference reviewer for *Neural Information Processing Systems* (NeurIPS 2021-2023), *International Conference on Machine Learning* (ICML 2020-2024), *International Conference on Learning Representations* (ICLR 2021), *Conference on Learning Theory* (COLT 2025), *IEEE International Symposium on Information Theory* (ISIT 2020-2021).
- Organizing committee for *Conference on Statistical Foundations of Data Science and their Applications in celebration of Jianqing Fan's 60th Birthday*, May 2023.
- Poster session committee for *The Women in Data Science (WiDS) Cambridge Conference*, Mar 2024.

# SHORT-TIME VISITING POSITIONS

**Massachusetts Institute of Technology**, MA, USA          *Sep. 2022 - Dec. 2022*

- Visiting student at Department of Mathematics, hosted by Philippe Rigollet

**University of California, Berkeley**, CA, USA          *Aug. 2021 - Nov. 2021*

- Visiting student at Simons Institute for the Theory of Computing