

Kaizheng Wang

<https://kw2934.github.io> | kaizheng.wang@columbia.edu

ACADEMIC POSITION

Columbia University, New York, NY, USA

Jul. 2020 -

Assistant Professor, Department of Industrial Engineering and Operations Research

Member, Data Science Institute

EDUCATION

Princeton University, Princeton, NJ, USA

Sep. 2015 - Jun. 2020

Ph.D. in Operations Research and Financial Engineering, Department of ORFE

Peking University, Beijing, China

Sep. 2011 - Jul. 2015

B.S. in Mathematics and Applied Mathematics, School of Mathematical Sciences

PUBLICATIONS AND PREPRINTS

(α - β : author names are sorted alphabetically; $^+$: student/postdoc supervised.)

Preprints under review

- Adaptive Transfer Clustering: A Unified Framework.
Yuqi Gu, Zhongyuan Lyu $^+$, Kaizheng Wang. (α - β)
arXiv:2410.21263, 2024.
- Distribution-Free Predictive Inference under Unknown Temporal Drift.
Elise Han $^+$, Chengpiao Huang $^+$, Kaizheng Wang. (α - β)
arXiv:2406.06516, 2024.
- A Stability Principle for Learning under Non-Stationarity.
Chengpiao Huang $^+$, Kaizheng Wang. (α - β)
arXiv:2310.18304, 2023.
- Pseudo-Labeling for Kernel Ridge Regression under Covariate Shift.
Kaizheng Wang.

arXiv:2302.10160, 2023.

Junior Researcher Award, 2024 ICSA China Conference.

- Variable Clustering via Distributionally Robust Nodewise Regression.

Kaizheng Wang, Xiao Xu, Xun Yu Zhou. (α - β)

arXiv:2212.07944, 2022.

- Adaptive Data Fusion for Multi-Task Non-Smooth Optimization.

Henry Lam, Kaizheng Wang, Yuhang Wu[†], Yichen Zhang. (α - β)

arXiv:2210.12334, 2022.

Journal publications

- Clustering a Mixture of Gaussians with Unknown Covariance.

Damek Davis, Mateo Díaz, Kaizheng Wang. (α - β)

Bernoulli, 2024+.

- Learning Gaussian Mixtures Using the Wasserstein-Fisher-Rao Gradient Flow

Yuling Yan*, Kaizheng Wang*, Philippe Rigollet. (* = equal contribution)

Annals of Statistics 52(4): 1774-1795, 2024.

- Adaptive and Robust Multi-Task Learning.

Yaqi Duan, Kaizheng Wang. (α - β)

Annals of Statistics 51(5): 2015-2039, 2023.

- Communication-Efficient Accurate Statistical Estimation.

Jianqing Fan, Yongyi Guo, Kaizheng Wang. (α - β)

Journal of American Statistical Association 118 (542): 1000-1010, 2023.

- An ℓ_p Theory of PCA and Spectral Clustering.

Emmanuel Abbe, Jianqing Fan, Kaizheng Wang. (α - β)

Annals of Statistics 50 (4): 2359-2385, 2022.

Frontiers of Science Award in Mathematics, 2024 International Congress of Basic Science.

Presented by Jianqing Fan at the **IMS Le Cam Lecture** at the 2021 Joint Statistical Meetings.

- Modern Data Modeling: Cross-Fertilization of the Two Cultures.

Jianqing Fan, Cong Ma, Kaizheng Wang, Ziwei Zhu. (α - β)

Observational Studies 7 (1): 65-76, 2021.

- Robust High Dimensional Factor Models with Applications to Statistical Machine Learning.
Jianqing Fan, Kaizheng Wang, Yiqiao Zhong, Ziwei Zhu. (α - β)
Statistical Science 36(2): 303-327, 2021.
- Entrywise Eigenvector Analysis of Random Matrices with Low Expected Rank.
Emmanuel Abbe, Jianqing Fan, Kaizheng Wang, Yiqiao Zhong. (α - β)
Annals of Statistics 48 (3): 1452-1474, 2020.
- Implicit Regularization in Nonconvex Statistical Estimation: Gradient Descent Converges Linearly for Phase Retrieval, Matrix Completion and Blind Deconvolution.
Cong Ma, Kaizheng Wang, Yuejie Chi, Yuxin Chen.
Foundations of Computational Mathematics 20: 451–632, 2020.
Short version accepted by **International Conference on Machine Learning (ICML)** 2018.
SIAM Activity Group on Imaging Science Best Paper Prize, 2024.
- Factor-Adjusted Regularized Model Selection.
Jianqing Fan, Yuan Ke, Kaizheng Wang. (α - β)
Journal of Econometrics 216 (1): 71-85, 2020.
- Comment on “A Tuning-Free Robust and Efficient Approach to High-Dimensional Regression”.
Jianqing Fan, Cong Ma, Kaizheng Wang. (α - β)
Journal of American Statistical Association 115 (532): 1720-1725, 2020.
- Distributed Estimation of Principal Eigenspaces.
Jianqing Fan, Dong Wang, Kaizheng Wang, Ziwei Zhu. (α - β)
Annals of Statistics 47 (6): 3009-3031, 2019.
- Spectral Method and Regularized MLE are both Optimal for Top-K Ranking.
Yuxin Chen, Jianqing Fan, Cong Ma, Kaizheng Wang. (α - β)
Annals of Statistics 47 (4): 2204-2235, 2019.
- Stochastic Representations for the Wave Equation on Graphs and Their Scaling Limits.
Kaizheng Wang
Journal of Mathematical Analysis and Applications 449 (1): 808-828, 2017.
- On the Neumann Problem for Harmonic Functions in the Upper Half Plane.
Kaizheng Wang
Journal of Mathematical Analysis and Applications 419 (2): 839-848, 2014.

Conference publications

- Model Assessment and Selection under Temporal Distribution Shift.
Elise Han[†], Chengpiao Huang[†], Kaizheng Wang. (α - β)
International Conference on Machine Learning, 2024.
- Efficient Clustering for Stretched Mixtures: Landscape and Optimality.
Kaizheng Wang, Yuling Yan, Mateo Díaz.
Neural Information Processing Systems, 2020.
- Implicit Regularization in Nonconvex Statistical Estimation: Gradient Descent Converges Linearly for Phase Retrieval and Matrix Completion.
Cong Ma, Kaizheng Wang, Yuejie Chi, Yuxin Chen.
International Conference on Machine Learning, 2018.

AWARDS

- | | |
|---|-------------|
| • ICBS Frontiers of Science Award in Mathematics | 2024 |
| • ICSA China Conference Junior Researcher Award | 2024 |
| • SIAM Activity Group on Imaging Science Best Paper Prize | 2024 |
| • Second Place Award - 2023 INFORMS Blue Summit Supplies Data Challenge | 2023 |
| • Harold W. Dodds Fellowship - Princeton University | 2019 - 2020 |
| • Gordon Y. S. Wu Fellowship - Princeton University | 2015 - 2019 |
| • SEAS Award for Excellence - Princeton University | 2018 |

GRANTS

- | | |
|---|-------------|
| • NSF Grant DMS-2210907 (\$179,999) | 2022 – 2025 |
| Statistical and Computational Tools for Analyzing High-Dimensional Heterogeneous Data | |
| Role: PI | |
| • Columbia University Data Science Institute Seed Fund (\$75,000) | 2024 – 2025 |
| Policy Evaluation with Transfer Learning: How to assess safety performance of self-driving cars in NYC? | |
| Role: Co-PI | |

PROFESSIONAL ACTIVITIES AND SERVICES

- Area chair/meta-reviewer: COLT 2024 - 2025, ICML 2023 - 2024, NeurIPS 2021 - 2022
- Session chair: INFORMS Annual Meeting 2020 - 2024

- Cluster chair, 2022 CORS-INFORMS International Conference Jun. 2022
- Co-organizer, Wilks statistics seminar, Princeton University Jul. 2018 - May. 2019
- Co-organizer, the 6th Princeton Day of Statistics Jul. 2018 - Nov. 2018
- Reviewer for the following journals: Annals of Applied Probability, Annals of Statistics, Bernoulli, Biometrika, Communications on Pure and Applied Mathematics, Foundations of Computational Mathematics, IEEE Transactions on Information Theory, Journal of Business & Economic Statistics, Journal of Econometrics, Journal of Machine Learning Research, Journal of the American Statistical Association, Journal of the Royal Statistical Society: Series B, Management Science, Mathematics of Operations Research, Operations Research, etc.
- Reviewer for the following conferences: Conference on Learning Theory (COLT), International Conference on Machine Learning (ICML), IEEE International Symposium on Information Theory (ISIT), Neural Information Processing Systems (NeurIPS), ACM-SIAM Symposium on Discrete Algorithms (SODA), etc.
- Outreach program: Mentor in the Data Science Research Program for high school students (run by The Coding School in collaboration with Columbia University and other institutes), Jun. – Aug. 2024.

TEACHING EXPERIENCES

At Columbia University:

- IEOR E8100 - High-Dimensional Probability with Applications (PhD): Spring 2021, 2023 & 2024;
- IEOR E4106 - Stochastic Models (Master): Spring 2024;
- IEOR E4102 - Stochastic Modeling for Management Science and Engineering (Master): Spring 2023;
- IEOR E4307 - Statistics and Data Analysis (Undergraduate): Fall 2020 & 2021;
- IEOR E3658 - Probability for Engineers (Undergraduate): Spring 2025;
- IEOR E3106 - Stochastic Systems and Applications (Undergraduate): Fall 2021 - 2023.

At Princeton University, as Assistants in Instruction (AIs):

- ORF 525 - Statistical Learning and Nonparametric Estimation (PhD): Spring 2019;
- ORF 363 - Computing and Optimization for Physical and Social Sciences (Undergraduate): Fall 2016;
- ORF 309 - Probability and Stochastic Systems (Undergraduate): Spring 2017 & 2018;
- ORF 245 - Fundamentals of Statistics (Undergraduate): Fall 2017 & 2018 (Head AI).

RESEARCH GROUP

Postdoctoral Research Scientist

- Zhongyuan Lyu (Data Science Institute Postdoc co-mentored with Yuqi Gu).

Ph.D. students

- Chengpiao Huang
Second Place Award in the 2023 INFORMS Blue Summit Supplies Data Challenge.
- Seok-Jin Kim

Undergraduate students

- Elise Han
INFORMS Scholarship, 2024.
Bonomi Scholarship, 2024.
- Alan Ma
- Eric Chen

Alumni

- Naomi Toft (Undergraduate)
- Geraldine Nina Montano (Undergraduate): Bonomi Scholarship in 2023.
- Rain Wei (Undergraduate): Bonomi Scholarship in 2023.
- Yuhang Wu (Undergraduate)

Now a PhD student at the Decision, Risk, and Operations (DRO) division at Columbia Business School.
Second Place Award in the 2023 INFORMS Blue Summit Supplies Data Challenge.

- Alice Chen (Master)
- Sara Zhao (Undergraduate): Stephen D. Guarino Memorial Award in 2022.
- Ethan Turok (Undergraduate)