

Soham Jana

240 Hayes-Healy Center
University of Notre Dame
Notre Dame, IN, USA

Updated on: June 21, 2025
Website: <https://janasoham.github.io>
Email: sjana2-at-nd-dot-edu
Phone: +1 574-631-5503

Research Interests

Theoretical and methodological aspects of high-dimensional statistics, robust estimation, neural networks, causal inference.

Education

PhD. in Statistics and Data Science May 2022
Yale University, New Haven, CT, USA
Thesis: Learning non-parametric and high-dimensional distributions
via information-theoretic methods
Advisor: Prof. Yihong Wu

Master of Statistics (Hons.) (First class with distinction) May 2017
Indian Statistical Institute, Kolkata, West Bengal, India
Specialization: Theoretical Statistics
Dissertation: Characterization of single-integral non-kernel divergences
Advisor: Prof. Ayanendranath Basu

Bachelor of Statistics (Hons.) (First class with Distinction) May 2015
Indian Statistical Institute, Kolkata, West Bengal, India

Work experiences

University of Notre Dame, Notre Dame, IN, USA
Assistant Professor, Department of Applied and Computational Mathematics and Statistics. August 2024 – Current

Princeton University, Princeton, NJ, USA
Postdoc, Department of Operations Research and Financial Engineering June 2022 – July 2024

Hosts: Prof. Sanjeev Kulkarni and Prof. Jianqing Fan
Researcher, The First Republic Bank Research and Lifelong Learning Program June 2022 – May 2023

Lecturer Spring 2023 and Fall 2023

Grants and Awards

Professional Development

Kaneb Center Course Design Academy, University of Notre Dame
Award Amount: USD 5000

2024 – 2025

Preprints (“*”: Authors list not in alphabetical order)

1. Chen, Xi[†], Soham Jana[†], Christopher A. Metzler, Arian Maleki, and Shirin Jalali. **Multilook Coherent Imaging: Theoretical Guarantees and Algorithms**. arXiv preprint arXiv:2505.23594 (2025). [†] Equal Contributions.
2. Shange Tang, Soham Jana, Jianqing Fan. **Factor adjusted spectral clustering for mixture models**. arXiv preprint arXiv:2408.12564 (2024).
3. Soham Jana, Kun Yang, and Sanjeev Kulkarni*. **Adversarially robust clustering with optimality guarantees**. arXiv preprint arXiv:2306.09977 (2023).
4. Soham Jana, Yury Polyanskiy, and Yihong Wu. **Optimal empirical Bayes estimation for the Poisson model via minimum-distance methods**. arXiv preprint arXiv:2209.01328 (2022). Under minor revision at **Information and Inference - a journal of IMA**.

Journal publications (“*”: Authors list not in alphabetical order)

1. Soham Jana, Jianqing Fan, Sanjeev Kulkarni*. **A provable initialization and robust clustering method for general mixture models**. arXiv preprint arXiv:2401.05574 (2024). Accepted at **IEEE Transactions on Information Theory**.
2. Soham Jana, Henry Li, Yutaro Yamada, and Ofir Lindenbaum. **Support recovery with Stochastic Gates: theory and application for linear models**. Elsevier Signal Processing (2023), 213, p.109193.
3. Yanjun Han, Soham Jana and Yihong Wu, **Optimal Prediction of Markov Chains With and Without Spectral Gap**, in IEEE Transactions on Information Theory, vol. 69, no. 6, pp. 3920-3959, June 2023, doi: 10.1109/TIT.2023.3239508. (**Extended from the NeurIPS version with analysis of higher-order Markov chains and different loss functions**)
4. Soham Jana and Ayanendranath Basu.* **A characterization of all single-integral, non-kernel divergence estimators**. IEEE Transactions on Information Theory 65.12 (2019): 7976-7984.

Conference publications (“*”: Authors list not in alphabetical order)

1. Soham Jana, Yury Polyanskiy, Anzo Teh, and Yihong Wu. **Empirical Bayes via ERM and Rademacher complexities: the Poisson model**. In Conference on Learning Theory 2023 Jul 15, PMLR 195:5199-5235.
2. Yanjun Han, Soham Jana, and Yihong Wu. **Optimal prediction of Markov chains with and without spectral gap**. NeurIPS 2021.

3. Soham Jana, Yury Polyanskiy, and Yihong Wu. [Extrapolating the profile of a finite population](#). In Conference on Learning Theory 2020 Jul 15 (pp. 2011-2033). PMLR.

Conferences and Invited Talks

| | |
|---|---------------|
| Joint Statistical Meetings Nashville, TN, USA | August 2025 |
| International Indian Statistical Association Cochin, Kerala, India | December 2024 |
| Joint Statistical Meetings Portland, OR, USA | August 2024 |
| University of Notre Dame Statistics Department Seminar Notre Dame, IN, USA | February 2024 |
| University of Wisconsin-Madison Statistics Department Seminar Madison, WI, USA | February 2024 |
| University of Texas at Dallas Statistics Department Seminar Richardson, TX, USA | January 2024 |
| Indian Statistical Institute ISRU Department Seminar Kolkata, West Bengal, India | July 2023 |
| Conference on Learning Theory (COLT) Bangalore, Karnataka, India | July 2023 |
| Neural Information Processing systems (NeurIPS) Virtual | December 2021 |
| Conference on Learning Theory (COLT) Virtual | July 2020 |

Teaching

University of Notre Dame

| | |
|--|-------------|
| Introduction to probability (ACMS 30530) | Fall 2024 |
| Modern Machine Learning Techniques with Application (ACMS 80870) | Spring 2025 |

Princeton University

| | |
|--|-------------|
| Probability and stochastic systems (ORF 309/ENG 309/MAT 380) | Spring 2023 |
| Statistical machine learning (ORF 570) | Fall 2023 |

Honors and awards

| | |
|-------------------------------------|-----------|
| INSPIRE Scholarship, Govt. of India | 2012-2017 |
|-------------------------------------|-----------|

| | |
|---|------|
| Indian National Mathematical Olympiad (INMO) merit certificate (For being among top 75 in the country) | 2012 |
|---|------|

Services

Paper reviewer

Annals of Statistics (2)
IEEE Transactions on Information Theory (7)
IEEE International Symposium on Information Theory (1)
Electronic Journal of Statistics (2)
Stat - an ISI Journal (1)
Algorithmic Learning Theory (3)
Bernoulli (1)

Organizational duties at conferences

| | |
|---|---------------|
| <i>International Joint Conference CFE-CMStatistics</i> | December 2025 |
| <i>Session organizer: Recent advances in Causal Inference</i> | |
| <i>Joint Statistical Meetings</i> | August 2024 |
| <i>Session chair: New Advances in Nonparametric Hypothesis Testing - Part I</i> | |
| <i>Session chair: New Developments in Non-Euclidean Statistics</i> | |
| <i>IEEE Conference on Information Sciences and Systems</i> | March 2024 |
| <i>Session chair: Machine learning and statistical inference</i> | |

| | |
|---|-------------|
| Community Service: Teaching at Math Circle, Notre Dame | Spring 2025 |
| <i>Promoting STEM education among school children</i> | |

| | |
|---|------|
| Yale S&DS M.A. admission committee | 2021 |
| Reviewer: one of the committee members handling over 150 applications and making admission recommendations | |

| | |
|---|------|
| Yale S&DS graduate reading group | 2020 |
| Co-organizer Scheduled talks and lead discussion sessions | |

| | |
|---|------|
| Yale Women in Data Science (WiDS) workshop | 2020 |
| Served as a mentor for Yale undergrad students participating in the WiDS Datathon Challenge 2020 | |

| | |
|--|-------------|
| South Asian Graduate and Professional Association at Yale (SAGA) | 2018 – 2021 |
| Treasurer, core committee member and cultural committee head Objective: organizing socio-cultural events to promote diversity and inclusion at Yale | |