Raaz Dwivedi

🏛 Harvard/MIT 🏠 rzrsk.github.io ✉ raaz@seas.harvard.edu ✉ raaz@mit.edu G ◯ in 🐦

ACADEMIC APPOINTMENTS	**FODSI Postdoctoral Fellow**	*2021—*
	Harvard University, School of Engineering, and Applied Sciences, Boston, USA Massachusetts Institute of Technology, Department of EECS, Cambridge, USA • Advisors: *Prof. Susan Murphy & Prof. Devavrat Shah*	
EDUCATION	**Ph.D., Electrical Engineering, and Computer Sciences**	*2015—2021*
	University of California, Berkeley, USA • Advisors: *Prof. Martin Wainwright & Prof. Bin Yu* • Thesis title: *Principled statistical approaches for sampling, and inference in high dimensions* • Thesis committee members: *Prof. David Aldous & Prof. Peter Bartlett*	
	B. Tech., Honors, Electrical Engineering	*2010—2014*
	Indian Institute of Technology, Bombay, India • Advisor: *Prof. Vivek Borkar & Prof. Juzer Vasi* • Graduated with Institute Rank 1, and Minors in Mathematics	
RESEARCH INTERESTS	Sequential decision making, causal inference, reinforcement learning, MCMC, and broadly high-dimensional statistical machine learning	
SELECTED ACHIEVEMENTS & AWARDS	Certificate of Distinction and Excellence in Teaching (Q Award), Harvard University	*2022*
	Institute of Mathematical Statistics (IMS) New Researcher Travel Award, London	*2022*
	Best Presentation Award, Machine Learning, and Statistics Session, Laboratory of Information, and Decision Systems (LIDS) Student Conference, MIT	*2022*
	Best Student Paper Award, Sections on Statistical Computing, and Statistical Graphics, American Statistical Association (ASA)	*2022*
	Outstanding Graduate Student Instructor Award, UC Berkeley	*2020*
	Berkeley Fellowship, the most prestigious fellowship for incoming students	*2015*
	President of India Gold Medal, IIT Bombay, for the highest GPA in the institute	*2014*
	Institute Silver Medal, IIT Bombay, for the highest Honors GPA (EE department)	*2014*
	Best B. Tech. Project (undergraduate thesis) Award, IIT Bombay	*2014*
	All India Rank 10 (amongst half a million), IIT Joint Entrance Exam (IIT-JEE)	*2010*
WORK EXPERIENCE	**Microsoft Research**, Research Intern (with Lester Mackey), New England, USA	*Summer 2019*
	Mist Systems (Juniper Networks), Data Science Intern, Cupertino, USA	*Summer 2017*
	WorldQuant Research, Senior Quantitative Researcher, Mumbai, India	*Jul 2014—Jul 2015*
	Stanford University, Research intern (with Prof. Balaji Prabhakar), USA	*Summer 2013*
	Ivy Mobility, Data Science Intern, Chennai, India	*Winter 2012*
PRE-PRINTS	P1. **Raaz Dwivedi**, Susan Murphy, and Devavrat Shah, "Counterfactual inference in sequential experimental design", *arXiv preprint 2022.*	

P2. **Raaz Dwivedi**[*], Chandan Singh[*], Bin Yu, and Martin J. Wainwright, "Revisiting minimum description length complexity in overparameterized models", *arXiv preprint 2021 (in journal revision)*.

P3. Nhat Ho[*], Koulik Khamaru[*], **Raaz Dwivedi**[*], Martin J. Wainwright, Michael I. Jordan, and Bin Yu, "Instability, computational efficiency, and statistical accuracy", *arXiv preprint 2021 (in journal submission)*.

JOURNAL PUBLICATIONS	(⋆ denotes equal contribution, and † denotes alphabetical ordering; title is hyperlinked to the online pdf of the paper)

J1. **Raaz Dwivedi**[*], Yan Shuo Tan[*], Briton Park, Mian Wei, Kevin Horgan, David Madigan, and Bin Yu, "Stable discovery of interpretable subgroups via calibration in causal studies", *International Statistical Review (ISR), 2020*.

J2. Nick Altieri[†], Rebecca L. Barter, James Duncan, **Raaz Dwivedi**, Karl Kumbier, Xiao Li, Robert Netzorg, Briton Park, Chandan Singh, Yan Shuo Tan, Tiffany Tang, Yu Wang, Chao Zhang, and Bin Yu, "Curating a COVID-19 data repository and forecasting county-level death counts in the United States", *Harvard Data Science Review (HDSR), 2020*.

J3. **Raaz Dwivedi**[*], Nhat Ho[*], Koulik Khamaru[*], Martin J. Wainwright, Michael I. Jordan, and Bin Yu, "Singularity, misspecification, and the convergence rate of EM", *Annals of Statistics (AoS), 2020*.

J4. Yuansi Chen, **Raaz Dwivedi**, Martin J. Wainwright, and Bin Yu, "Fast mixing of Metropolized Hamiltonian Monte Carlo: Benefits of multi-step gradients", *Journal of Machine Learning Research (JMLR), 2020*.

J5. **Raaz Dwivedi**[*], Yuansi Chen[*], Martin J. Wainwright, and Bin Yu, "Log-concave sampling: Metropolis-Hastings algorithms are fast", *Journal of Machine Learning Research (JMLR), 2019*.

J6. **Raaz Dwivedi**[†], Ohad N. Feldheim, Ori Gurel-Gurevich, and Aaditya Ramdas, "The power of online thinning in reducing discrepancy", *Probability Theory, and Related Fields (PTRF), 2019*.

J7. Yuansi Chen[*], **Raaz Dwivedi**[*], Martin J. Wainwright, and Bin Yu, "Fast MCMC sampling algorithms on polytopes", *Journal of Machine Learning Research (JMLR), 2018*.

J8. Vivek Borkar[†], **Raaz Dwivedi**, and Neeraja Sahasrabudhe, "Gaussian approximations in high dimensional estimation", *Systems & Control Letters, 2016*.

CONFERENCE PUBLICATIONS	

C1. **Raaz Dwivedi**, and Lester Mackey, "Generalized kernel thinning", To appear in *International Conference on Learning Representations (ICLR), 2022*.

C2. Abhishek Shetty, **Raaz Dwivedi**, and Lester Mackey, "Distribution compression in near-linear time", To appear in *International Conference on Learning Representations (ICLR), 2022*.

C3. **Raaz Dwivedi**, and Lester Mackey, "Kernel thinning", Extended abstract in *Conference on Learning Theory (COLT), 2021. Full version in journal submission*.

C4. **Raaz Dwivedi**[*], Nhat Ho[*], Koulik Khamaru[*], Martin J. Wainwright, Michael I. Jordan, and Bin Yu, "Sharp analysis of Expectation-Maximization for weakly identifiable models", *The 23rd International Conference on Artificial Intelligence, and Statistics (AISTATS), 2020*.

C5. **Raaz Dwivedi**[*], Nhat Ho[*], Koulik Khamaru[*], Martin J. Wainwright, and Michael I. Jordan, "Theoretical guarantees for EM under misspecified Gaussian mixture models", *Advances in Neural Information Processing Systems (NeurIPS), Montreál, 2018*.

C6. **Raaz Dwivedi**[*], Yuansi Chen[*], Martin J. Wainwright, and Bin Yu, "Log-concave sampling: Metropolis-Hastings algorithms are fast", Extended abstract in *Conference on Learning Theory (COLT), Stockholm, 2018*.

C7. Yuansi Chen[*], **Raaz Dwivedi**[*], Martin J. Wainwright, and Bin Yu, "Vaidya walk: A sampling algorithm based on the volumetric barrier", *Communication, Control, and Computing (Allerton), 55th Annual Allerton Conference, 2017*.

C8. **Raaz Dwivedi**, and Vivek Borkar, "Removing sampling bias in networked stochastic approximation", *International Conference on Signal Processing, and Communications (SPCOM), Bangalore, 2014.*

SOFTWARES &
METHODOLOGIES

P1. **Raaz Dwivedi**, Lester Mackey, Python package "Kernel Thinning", available at `https://github.com/microsoft/goodpoints`.

P2. Abhishek Shetty[*], **Raaz Dwivedi**[*], and Lester Mackey, Python package "Compress++", available at `https://github.com/microsoft/goodpoints`.

P3. **Raaz Dwivedi**[*], Yan Shuo Tan[*], Briton Park, Mian Wei, Kevin Horgan, David Madigan, and Bin Yu, GitHub Repo "StaDISC" available at `https://github.com/Yu-Group/stadisc`.

P4. Yuansi Chen[*], **Raaz Dwivedi**[*], Martin Wainwright, and Bin Yu, Python package (C++ implementation) "Vaidya and John walks", available at `https://github.com/yuachen/polytopewalk`.

SELECTED INVITED
TALKS

IT1. Revisiting minimum description length complexity in overparameterized models. *Algorithmic Information Theory & Machine Learning Symp., Alan Turing Institute, London.*   July 2022

IT2. Counterfactual inference in sequential experimental design. *Institute of Mathematical Statistics (IMS) Annual Meeting, Statistical Machine Learning Session, London.*   June 2022

IT3. Counterfactual inference in sequential experimental design. *Causal Inference Reading Group, University of Cambridge.*   June 2022

IT4. Near-optimal compression in near-linear time. *Sympo. on Kernel methods for numerical integration, SIAM Conference on Uncertainty Quantification, Atlanta.*   Apr 2022

IT5. Near-optimal compression in near-linear time. *Stable, Generalizable, & Transferable Statistical Learning Workshop, Mathematical Sciences Research Institute, Berkeley.*   Mar 2022

IT6. Counterfactual inference in sequential experimental design. *Learning from Interventions Workshop, Simons Institute, Berkeley.*   Feb 2022

IT7. Revisiting minimum description length complexity in overparameterized models. *Collaborations on the Theoretical Foundations of Deep Learning.*   Nov 2021

IT8. StaDISC: Stable discovery of interpretable subgroups via calibration. *Young Data Scientist Research Seminar, ETH Zurich.*   Sep 2020

IT9. Veridical data science, and the PCS framework. *ASA Annual Symposium on Data Science, and Statistics (SDSS).*   Jun 2020

IT10. Statistics meets optimization: Two vignettes on the Intersection. *Department of Mathematics, and Statistics, IIT Kanpur.*   Jan 2020

IT11. Singularity, misspecification, and the convergence rate of Expectation-Maximization. *Fall Western Sectional Meeting of the AMS, UC Riverside.*   Nov 2019

IT12. Power of gradients, and accept-reject step in MCMC algorithms. *BIDS Statistics, and Machine Learning Forum, UC Berkeley.*   Mar 2019

IT13. Theoretical guarantees for MCMC algorithms, *STCS Seminar, TIFR Bombay.*   Jan 2018

TEACHING
EXPERIENCE

Teaching Fellow, Harvard University   *2022–*

- *Sequential Decision Making*: STAT 234, Spring 2022, Taught by Prof. Susan Murphy. Besides usual duties, gave guest lectures, and mentored several half semester-long class projects.

Graduate Student Instructor, UC Berkeley   *2018—2019*

- *Modern Statistical Prediction, and Machine Learning*: STAT 154, Spring 2019, Taught by Prof. Bin Yu. Helped in *redesigning* the class along with one other TA, Yuansi Chen, for a class of 140+ students.

- *Introduction to Machine Learning*: EECS 189, Spring 2018, Taught by Prof. Anant Sahai, and Prof. Jennifer Listgarten. *Co-led* the content development (homeworks, discussions, and exams) in a team of 20+ TAs in a class of 350+ students.

Teaching Assistant, IIT Bombay, and Government of India (GoI) *2011—2014*

Responsible for weekly discussions (40 students) besides exam grading, and (optional) help sessions for larger groups (up to 200 students)

- *Linear Algebra*: MA 106, Spring 2012, Spring 2013, Spring 2014 (for IIT Bombay, and GoI)
- *Ordinary Differential Equations*: MA 108, Spring 2012, Spring 2013
- *Partial Differential Equations*: MA 207, Summer 2012
- *Calculus*: MA 105, Fall 2011, Fall 2012
- *Electromagnetism*: PH 103, Fall 2011

Guest lectures

- *STAT 234, Harvard University*: Three lectures on regret analysis of posterior sampling in episodic reinforcement learning with titles: (i) The fellowship of Bayesian regret, (ii) the two Bellman operators, and (iii) the return of the martingale *Apr 2022*
- *STAT 234, Harvard University*: Offline off-policy reinforcement learning *Feb 2022*
- *STAT 212, UC Berkeley*: Revisiting Complexity, and the Bias-Variance Tradeoff: Using Minimum Description Length *Apr 2021*
- *EECS 189, UC Berkeley*: An Introduction to Ensemble Methods *Oct 2019*
- *STAT 154, UC Berkeley*: Boosting *Apr 2019*

ACADEMIC SERVICES

Committees
- Member, Committee on Equality and Diversity, IMS *2022—*

Scientific Meetings
- Chair, New Researchers Group Session, IMS Annual Meeting *June 2022*
- Chair, Statistical Machine Learning Session, IMS Annual Meeting *June 2022*
- Mentor, Summer Institute on Just-in-Time Adaptive Interventions via MRTs *Oct 2021*

Institutional Mentoring Activities
- MIT Institute for Data, Systems, & Society (IDSS) Postdoc Mentors for *PhDs* *2022—*
- UC Berkeley Artificial Intelligence Research (BAIR) Buddies for *incoming PhDs* *2020—2021*
- UC Berkeley BAIR Undergraduate Mentoring Program for *undergraduates* *2017—2021*
- IIT Bombay Student Mentoring Program (ISMP) for *incoming undergraduates* *2013—2014*
- IIT Bombay Academic Mentoring Program (DAMP) for *sophomores & juniors* *2012—2014*
- IIT Bombay Intensive Program for Entrants (IPE) for *incoming undergraduates* *2012—2013*

Reviewing Activities
- *Journals*: JMLR, IEEE-IT, JRSSB, Bernoulli, MOR, Stats & Comp., SIAM
- *Conferences*: COLT, ICML, AISTATs, NeurIPS, FOCS, SODA, AAAI

Graduate Admissions
- EECS Graduate Admissions Committee, MIT *2021*
- EECS Graduate Admissions Committee, UC Berkeley *2018—2020*