BIJAN MAZAHERI

bijanmazaheri.com bijan.h.mazaheri@dartmouth.edu github.com/honeybijan
Hanover, NH (781)-985-0881

I study topics related to causality, distribution shift, decision fusion, and mixture models, particularly concerning combining data and knowledge from multiple places, topics, and modalities. I approach problems through a bifocal lens, with theory guiding data-realistic approaches to application and application giving rise to new fundamental discoveries.

EDUCATION

California Institute of Technology - Pasadena, CA *Oct 2017 - Aug 2023*
Ph.D. Candidate
Department of Computing and Mathematical Sciences, GPA: 3.9/4.0
Awarded NSF Graduate Research Fellowship and Amazon AI4Science Research Fellowship
Thesis Title: Combining Sources and Leveraging Contexts

Cambridge University (Emmanuel College) - Cambridge, UK *Oct 2016 - Jun 2017*
Mathematics Part 1B
Supported by a Herchel Smith Fellowship
Additional classes in Computer Science and Mathematics Part II

Williams College - Williamstown, MA *Sep 2012 - Jun 2016*
Bachelor of Arts
Physics and Computer Science, GPA: 3.92/4.00
Highest Honors (Physics), Phi Beta Kappa, Sigma Xi, Magna Cum Laude
Thesis Title: RNA Macrostates and Macrokinetics

WORK EXPERIENCE

Thayer School of Engineering, Dartmouth College - Hanover, NH *Jan 2025 - Now*
Assistant Professor

Broad Institute of MIT and Harvard - Cambridge, MA
Eric and Wendy Schmidt Postdoctoral Associate *Oct 2023 - Dec 2024*
Visiting Research Scientist *Jan 2025 - Now*
Current projects focus on batch-effect correction, causal inference, missing data, and efficient experimental design. Primarily working with Prof. Caroline Uhler's group.

Amazon Research Causality Lab - Tübingen, Germany *Oct 2022 - Feb 2023*
Applied Scientist Intern (L5)
Worked with Dr. Michaela Hardt, Dr. Atalanti Mastakouri, and Dr. Dominik Janzing
Lead-authored a paper accepted to UAI 2023.

BioDiscovery - El Segundo, CA *Jun 2017 - Sep 2017*
Intern
I developed and implemented methods for clustering cancers based on genome aberrations. My work has now been integrated into BioDiscovery's software and presented at a conference.

IBM T.J. Watson Research Center - Yorktown Heights, NY *Jun 2016 - Sep 2016*
Intern
Worked with Dr. Victor Kravets (mentor) and Dr. Andrew Sullivan (manager).
Projects included non-greedy and map-reduce algorithms for factoring sum of products representations, with the goal of developing more efficient circuits.

TEACHING

Principles of Causality *Spring 2025*
I developed a new class on causality and causal inference for Dartmouth Engineering, first taught in Spring 2025.

Markov Chain Monte Carlo *Spring 2022*
Head TA for new class on MCMC methods in theoretical computer science. Developed solutions and grading rubrics for problem sets.

Physics and Mathematics *Sep 2013-Jun 2016*
TAed for undergraduate classes in Electricity and Magnetism, Classical Mechanics, Mathematical Methods for Scientists, Premed Physics, Discrete Mathematics.

AWARDS AND GRANTS

Eric and Wendy Schmidt Postdoctoral Fellowship *Awarded Summer 2023*
Funding for 1-3 years of research at the Broad Institute of MIT and Harvard.
Amazon AI4Science Research Fellowship *Awarded Spring 2022*
Funding for research with the potential to aid scientific discovery.
National Science Foundation Graduate Research Fellowship *Awarded Spring 2019*
3 year Ph.D. fellowship awarded for a proposal to research confounding influence in causal networks.
Herchel Smith Fellowship *Awarded Spring 2016*
Funding for 1-3 years of study at Cambridge University.

PUBLICATIONS

Bijan Mazaheri, Chandler Squires, Caroline Uhler. Synthetic Potential Outcomes and Causal Mixture Identifiability. In *The 28th International Conference on Artificial Intelligence and Statistics*, 2025.

Bijan Mazaheri, Siddharth Jain, Matthew Cook, Jehoshua Bruck. Omitted Labels Induce Nontransitive Paradoxes in Causality. In *4th Conference on Causal Learning and Reasoning*, 2025.

Spencer Gordon, Eric Jahn, ***Bijan Mazaheri**, Yuval Rabani, and Leonard J Schulman. Identification of Mixtures of Discrete Product Distributions in Near-Optimal Sample and Time Complexity. *arXiv:2309.13993*, 2023. In *The Thirty Seventh Annual Conference on Learning Theory*, pp. 2071-2091. PMLR, 2024.

Bijan Mazaheri, Atalanti Mastakouri, Dominik Janzing, and Michaela Hardt. Causal Information Splitting: Engineering Proxy Features for Robustness to Distribution Shifts. In *The 39th Conference on Uncertainty in Artificial Intelligence*, 2023.

Spencer Gordon, ***Bijan Mazaheri**, Yuval Rabani, and Leonard J Schulman. Causal Inference Despite Limited Global Confounding via Mixture Models. In *2nd Conference on Causal Learning and Reasoning*, 2023.

Siddharth Jain, **Bijan Mazaheri**, Netanel Raviv, and Jehoshua Bruck. Glioblastoma signature in the DNA of blood-derived cells. *PLOS ONE* 16(9): e0256831. 2021.

Bijan Mazaheri, Siddharth Jain, and Jehoshua Bruck. Expert Graphs: Synthesizing New Expertise via Collaboration. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pages 2447–2452, 2021.

Spencer Gordon, ***Bijan Mazaheri**, Yuval Rabani, and Leonard Schulman. Source Identification for Mixtures of Product Distributions. In *The Thirty Fourth Annual Conference on Learning Theory*, pages 2193–2216. PMLR, 2021.

Bijan Mazaheri, Siddharth Jain, and Jehoshua Bruck. Robust Correction of Sampling Bias using Cumulative Distribution Functions. *Advances in Neural Information Processing Systems*, volume 33, pages 3546–3556. Curran Associates, Inc., 2020.

* = Authorship order is alphabetical.

PREPRINTS

Bijan Mazaheri, Jiaqi Zhang, and Caroline Uhler. Meta-Dependence in Conditional Independence Testing. *arXiv:2504.12594*, 2025.

Bijan Mazaheri, Spencer Gordon, Yuval Rabani, and Leonard Schulman. Causal Discovery under Latent Class Confounding. *arXiv:2311.07454*, 2023.

Spencer Gordon, ***Bijan Mazaheri**, Yuval Rabani, and Leonard J Schulman. The sparse Hausdorff moment problem, with application to topic models. *arXiv:2007.08101*, 2020.

Siddharth Jain, **Bijan Mazaheri**, Netanel Raviv, and Jehoshua Bruck. Cancer Classification from Healthy DNA using Machine Learning. *BioRxiv*, page 517839, 2019.

Siddharth Jain, **Bijan Mazaheri**, Netanel Raviv, and Jehoshua Bruck. Short Tandem Repeats Information in TCGA is Statistically Biased by Amplification. *BioRxiv*, page 518878, 2019.

* = Authorship order is alphabetical.

PATENTS

Siddharth Jain, **Bijan Mazaheri**, Netanel Raviv, and Jehoshua Bruck. Mutation profile and related labeled genomic components, methods and systems. 2019.

WORKSHOPS

Simon's Institute for Theory of Computing: Causality *Spring 2022*
4 week workshop on Causal inference methods.

INVITED TALKS

Simon's Institute for Theory of Computing *May 2023*
Title: "Causal Discovery under Limited Global Confounding"

Jones Seminar, Thayer School of Engineering at Dartmouth *May 2024*
Title: "Latency and Heterogeneity in Data and What to do About it"

Stanford Online Causal Inference Seminar *Oct 2024*
Title: "Synthetic Potential Outcomes and the Hierarchy of Causal Identifiability"

Boston University Machine Learning Symposium *Nov 2024*
Title: "Synthetic Potential Outcomes and Causal Mixture Identifiability"

PROJECTS

LACCTiC *Sep 2021 - present*
I maintain a website for collegiate cross country with 10,000 regular users that applies concepts from batch-effect correction to ranking performances on differing terrain. The backend runs on Python and Django and the frontend uses React, and the database is hosted on AWS. I have helped advise over 20 student projects using this data.

MENTORSHIP

MIT UROP *Summer 2024*
Supervised two undergraduate students on projects related to causality.
Caltech Cross Country Team *Sep 2018 - Sep 2022*
Mentoring and supporting undergraduate students at Caltech as an assistant coach.
Data Science Projects *Sep 2019 - present*
I have supported over 20 projects with undergraduate students using data on my website, and have advised some of these students in applying to graduate school.

SERVICE

Reviewer for ICML, NeurIPS, AIStats, CLeaR, The American Statistician, Nature Machine Intelligence.
Sports statistics outreach for Cross Country and correspondent at D3 Glory Days.
Caltech Community Associate - volunteer position for building community in Caltech housing.