CURRICULUM VITAE

Name: RAHUL MAZUMDER

Sloan Group: Operations Research and Statistics (OR & Stats) Group

Place of Birth: Kolkata, India

I Education

Degree	*Subject*	*School*	*Date*
PhD	Statistics	Stanford University, Stanford, CA	2007 – 2012
Masters in Statistics	Statistics	Indian Statistical Institute, Kolkata, India	2005 – 2007
Bachelors in Statistics	Statistics	Indian Statistical Institute, Kolkata, India	2002 – 2005

II Title of Doctoral Thesis and Name of Thesis Advisor

Title: Topics in Sparse Multivariate Statistics

Advisor: Trevor Hastie, Stanford University.

III Principal Fields of Interest

- Statistics, Machine Learning, Mathematical Optimization (Convex optimization, Mixed Integer optimization), Large Scale Optimization Algorithms.
- High dimensional statistics and sparsity, combinatorial statistical modeling & computation, nonparametric function estimation (e.g., shape constrained inference).
- Applications of the above in recommender systems, computational finance, computational biology & healthcare, survey research, insurance pricing, etc. Conditional computing in neural networks.

IV Name and Rank of other Sloan Faculty in Same Field

Name	*Rank*
Arnold Barnett	Professor
Dimitris Bertsimas	Professor
Colin Fogarty	Assistant Professor
Robert Freund	Professor
David Gamarnik	Professor
Alexandre Jacquillat	Assistant Professor
Thomas Magnanti	Institute Professor
James Orlin	Professor
Georgia Perakis	Professor
Bart Van Parys	Assistant Professor
Roy Welsch	Professor

V Non-MIT Employment

Employer	*Position*	*Start-End*
Columbia University New York	Tenure Track Assistant Professor Department of Statistics	07/2013-06/2015
Yahoo! Research Santa Clara, CA	Research Intern Machine Learning & Statistics Group	07/2010 - 09/2010
Yahoo! Research Santa Clara, CA	Research Intern Machine Learning & Statistics Group	07/2009 - 09/2009

VI History of MIT Appointments

Employer	*Position*	*Start-End*
Sloan School of Management	Associate Professor, without tenure (OR & Stats group)	07/2019-
Sloan School of Management	Robert G. James Career Development Professor	02/2019-
Sloan School of Management	Assistant Professor (OR & Stats group)	07/2015-06/2019
MIT IBM Watson AI Lab	Principal Investigator	09/2019-
Center for Statistics and Data Science	Core Faculty	07/2017-
Operations Research Center	Affiliate	07/2015-
Operations Research Center	Postdoctoral Associate	10/2012-06/2013

VII MIT Activities

Committee	*Year*
Dean's Social and Ethical Responsibilities of Computing (SERC) Action Group for Teaching	2021
MIT Sloan OR & Stats faculty hiring committee (junior faculty member)	2017-2020
MIT Sloan First Reappointment Committee Member (Dr. Jackson Lu)	2020
Operations Research Center Seminar Series Faculty Coordinator	2018
Admissions Committee, Masters in Business Analytics	2017-2021
Graduate Admissions Committee, MIT Operations Research Center	2016
Coordinator, Probability Qualifiers Examination, MIT Operations Research Center	2016, 2018
Committee Member, Probability Qualifiers Examination, MIT Operations Research Center	2016-2020
Chair, Best Student Paper Award Committee, MIT Operations Research Center	2016
Best Student Paper Award Committee, MIT Operations Research Center	2019
PhD General Exam Committee Member, MIT Operations Research Center	2016-2021
PhD Thesis Committee Member, MIT Operations Research Center	2016-2021

VIII Industry, Government Collaboration and Consulting Activities

- Part-time consultant (Sr. Staff Research Scientist) at LinkedIn AI: Estimated start date Oct 25, 2021.
- I have worked with the following companies and government agencies as a research collaborator. Some of these engagements are in the form of sponsored research agreements between MIT and the company.

Name	*Location*	*Period*
LinkedIn	USA	2019-2020
Google	USA	2019-
IBM	USA	2019-
Wells Fargo (IBM Member company)	USA	2021-
Refinitiv (IBM Member company)	USA	2019-2021
Liberty Mutual Insurance	USA	2020-
HiCore Tech	China	2019-2020
Processminer Inc.	USA	2018-2019
United States Census Bureau	USA	2017-
Pandora Media Inc.	USA	2016-2019

- I have worked with the following companies via advising students on their MIT MBAN Capstone and MIT LGO thesis projects.

Name	*Location*	*Period*
MailChimp	USA	2018, 2020
Rapid7	USA	2019, 2020
Anheuser-Busch	USA	2019
Pratt & Whitney	USA	2018
GE Appliances	USA	2017

IX Awards, Honors and Research Grants

SELECTED AWARDS & HONORS

- Donald P. Gaver, Jr. Early Career Award for Excellence in Operations Research, 2021
  - To be announced during INFORMS 2021 Annual Meeting.
  - "The purpose of the award is to support creative and diverse work in operations research in the early career of the recipient. Winning entry chosen based on (1) individual's demonstrated creative and diverse contributions to Operations Research(OR)/Management Science(MS) research, (2) individual's contributions to disseminating OR/MS knowledge; and (3) evidence of successful collaboration with scholars outside individual's home institution and potential for growth through collaboration."
- INFORMS Optimization Society Young Researchers Award, 2020
  - "Established in 1998, this is awarded to one or more young researcher(s) for an outstanding paper in optimization that is published in, or submitted to and accepted by, a refereed professional journal within the four calendar years preceding the year of the award"
- Office of Naval Research Young Investigator Award (ONR YIP), 2018
  - One of 31 awardees nation-wide across all disciplines in 2018.
  - "Introduced in 1985, the ONR YIP is one of the nation's oldest and most selective science and technology basic research programs. Its purpose is to fund early-career academic researchers...whose scientific pursuits show outstanding promise for supporting the Department of Defense..."
- Annals of Statistics *Special Invited Session* Speaker at the Joint Statistical Meetings (JSM), 2017
  - Editors' choice for one of the best four papers accepted to the Annals of Statistics in the previous two years. As an awardee, I presented at a JSM session sponsored by the Institute of Mathematical Statistics.

PAPER AWARDS TO STUDENT CO-AUTHORS

- Mixed Integer Programming Workshop Student Poster Award, 2021 (most popular poster, as advisor)
- INFORMS Computing Society Student Paper Prize, 2020 (honorable mention, as advisor)
- MIT Operations Research Center Best Student Paper Award, 2020 (as advisor)
- Mixed Integer Programming Workshop Student Poster Award, 2019 (honorable mention, as advisor)
- INFORMS Optimization Society Student Paper Prize, 2015 (as co-author).

OTHER AWARDS AND HONORS

- Google Research-MIT AI Collaboration Research Award, 2021
- MIT-IBM Watson AI Initiative Research Awards: 2019, 2020
- Plenary Speaker, 14th UMBC Probability and Statistics Day, 2020. (postponed due to pandemic)
- Plenary tutorial speaker at 2016-17 SAMSI Optimization Program.
- MIT Research Support Committee Award 2017.
- Stanford Statistics Department Award for Excellence in Teaching (2010-2011).
- Travel awards from American Statistical Association (2011), Institute of Pure and Applied Mathematics UCLA (2010), Theory and Practice of Computational Learning Theory, University of Chicago (2009).
- Tata Consultancy Services (TCS) award for the best Masters thesis, Indian Statistical Institute, 2007.

Research Grants and Contracts

Note: While at MIT since 2015, as a solo-PI, I have raised a total of $3.05 million in research funding (this includes $175,000 from MIT internal research support JFRAP, RSC)

FEDERAL RESEARCH GRANTS

- **ONR:** Office of Naval Research (Division of Mathematical Data Science) grant on *"Statistical Learning with large parameter spaces: Interpretable Nonparametrics, Conditional Computing and Beyond"*. Award N00014-21-1-2841
(Sole PI: R. Mazumder, Period: 10/01/2021 - 09/31/2024, Amount: $\$420,000$)
- **ONR:** Office of Naval Research (Young Investigator Award, Division of Mathematical Data Science) grant on *"Combinatorial Statistical Inference with Mathematical Optimization"*. Award N00014-18-1-2298.
(Sole PI: R. Mazumder, Period: 06/01/2018 - 09/31/2021, Amount: $\$508,780$)
- **NSF-IIS:** National Science Foundation (Division of Information & Intelligent Systems) grant on *"A New Perspective on Grouped Variable Selection via Modern Optimization"*. Award: 1718258
(sole PI: R. Mazumder, Period: 09/01/2017 - 08/31/2022, Amount: $\$318,000$)
- **ONR:** Office of Naval Research (Division of Mathematical Data Science) on *"Rigorous Modeling and Computation for Sparse Multivariate Statistical Problems"*. Award N00014-15-1-2342.
(Sole PI: R. Mazumder, Period: 07/01/2015 - 12/31/2018, Amount: $\$360,000$)

INDUSTRY, FOUNDATIONS AND INTERNAL RESEARCH GRANTS

- **Google:** Google Research award on *"Improving Conditional Computation in Neural Networks using (Combinatorial) Optimization"*
(Sole PI: R. Mazumder, Period: 2021 - 2022, Amount: $\$160,000$)
- **Liberty-Mutual:** Liberty Mutual Insurance grant on *"Multitask gradient boosted decision trees for Insurance Pricing"*
(Sole PI: R. Mazumder, Period: 2020 - 2022, Amount: $\$314,000$)
- **IBM:** MIT-IBM Watson AI Lab Research Award on: *"Knowledge Graph Based Explainable Financial Forecasting Framework for Portfolio Optimization"*. Current IBM member company sponsors: Refinitiv and Wells Fargo
(Sole PI: R. Mazumder, Period: 2019-2022. Amount: $\$675,000$, including supplemental funding from Refinitiv and Wells Fargo)
- **HiCore:** HiCore Tech, Shanghai grant on *"Unsupervised methods for automated fraud detection"*
(Sole PI: R. Mazumder, Period: 2019 - 2020, Amount: $\$100,000$)
- **MIT-JFRAP:** MIT Sloan Junior Faculty Research Assistance Program (JFRAP)
(Sole PI: R. Mazumder. Years awarded: 2016, 2017, 2019. Amount: $\$100,000$)
- **Processminer:** Processminer Inc., USA: unrestricted gift for research
(Sole PI: R. Mazumder. Year awarded 2018. Amount: $\$10,000$)
- **MIT-RSC:** MIT Research Support Committee (RSC) Award: Solomon Buchsbaum grant.
(Sole PI: R. Mazumder. Year awarded 2017. Amount: $\$75,000$)
- **Moore-Sloan:** Gordon and Betty Moore Foundation and Alfred P. Sloan Foundation Interface Grant Award (Columbia University) on *"Mining an Ocean of Data: Applications of Modern Statistical Methods for Addressing Biological Oceanographic Questions"*
(PIs: J. Goes and R. Mazumder. Period: 2014 - 2015. Total Amount: $\$304,562$, equally shared.)

X Professional Membership and Activities

Editorial Boards of Journals

- Associate Editor, Bernoulli (2022-)
- Associate Editor, Annals of Statistics (2020 -)
- Editorial Board Member, Journal of Machine Learning Research (2020 -)

Services as a Reviewer and Panelist

- Reviewer for the following journals and conferences:

  Artificial Intelligence and Statistics (AISTATS), Annals of Applied Statistics, Annals of Statistics, Biometrika, Computational Statistics and Data Analysis, Journal of the American Statistical Association, Journal of Machine Learning Research, Journal of the Royal Statistical Society (Series B), Journal of the Royal Statistical Society (Series C), Journal of Computational and Graphical Statistics , Journal of Optimization Theory and Applications, IEEE Transactions on Signal Processing, IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Information Theory, Information and Inference, INFORMS Journal on Computing, INFORMS Journal on Optimization, Management Science, Neural and Information Processing Systems, Operations Research, Operations Research Letters, Optimization Methods and Software, SIAM Journal on Optimization.

- Panelist on the NSF panels:
  - Division of Computing and Communication Foundations 2018
  - Division of Information and Intelligent Systems 2017

- Ad-hoc reviewer of grant proposals for:
  - Office of Naval Research (2020)
  - US Department of Energy (2020)
  - Army Research Office (2018)
  - NSF Division of Computing and Communication Foundations (2018)
  - Israel Science Foundation (2015,2016)
  - National Security Agency-American Mathematical Society (2015)

Other professional activities

- Professional Activities as a Visiting Fellow/Scholar
  - Visiting Fellow at "Initiative on Data Science in Business and Economics" 2017-2018 at the University of Chicago, Booth School of Business (April - May, 2018)
  - Summer at Census Scholar, United States Census Bureau (August 2018)

- Conference, Workshop Organization and Program Committee Member
  - Founding Cluster Chair: Machine Learning Cluster within INFORMS Optimization Society, 2021
  - Program Committee Member (with Damek Davis): "Optimization for Data Science Track" at the INFORMS Optimization Society Conference, 2020. *Postponed to 2022.*

  - Co-organizer (with Roger Koenker and Guillaume Pouliot) of "Optimization-Conscious Econometrics Conference" at Becker Friedman Institute for Economics at University of Chicago, 2019.
  - Co-organizer (local committee) of "Mixed Integer Programming" Workshop at MIT, Cambridge, 2019.

- Session Chair and co-organizer
  - Session Chair/co-organizer INFORMS Annual Conference, Hybrid, USA, 2021.
  - Session Chair/co-organizer INFORMS Annual Conference, Virtual, USA, 2020.
  - Session Chair/co-organizer INFORMS Annual Conference, USA, 2019.
  - Session co-organizer Computational and Methodological Statistics, UK, 2017.
  - Session chair/co-organizer, International Indian Statistical Association Annual Conference, USA, 2016.
  - Session chair/co-organizer at INFORMS Annual Conference, USA, 2016.

- Membership in Professional Societies

  American Statistical Association (ASA), Institute of Mathematical Statistics (IMS), International Indian Statistical Association (IISA), Institute for Operations Research and the Management Sciences (INFORMS), Society for Industrial and Applied Mathematics (SIAM).

XI Subjects Taught

Teaching at MIT

Number	*Title*	*Date*
15.060	Data Models and Decisions (MBA Core)	Fall 2016, 2017, 2019, 2020, 2021
15.S15	Readings in Statistics (co-taught with Colin Fogarty)	Fall 2016
15.075	Statistical Thinking and Data Analysis (Undergraduate Course)	Spring 2016
15.097	Statistical Learning via a Modern Optimization Lens (PhD course, co-taught with Dimitris Bertsimas)	Spring 2016

Teaching prior to MIT

- Teaching at Columbia University

– W4240 "Data Mining", Spring 2015 (MS Class)
– W4105 "Introduction to Probability", Fall 2014 (MS Class)
– W4107 "Introduction to Statistical Inference", Fall 2013 (MS Class)
– G8325 "Topics in Modern Multivariate Statistics: Modeling, Computation and Theory" (Advanced Topics Course in Statistics, PhD Class)
– S4204 Data Mining, Summer 2014 (MS Class)

- Teaching at Stanford University as a Teaching Assistant

  – STATS 60: Spring 2011, Undergraduate Course in Introductory Statistics (pre-calculus).
  – STATS 110: Spring 2011, Introductory Statistics course for Engineers (post-calculus).
  – STATS 306B: Spring 2009, PhD first year course in Applied Statistics
  – STATS 300A: Fall 2008, PhD first year course in Theoretical Statistics
  – STATS 217/218/219: Winter 2008, Spring 2008, Summer 2008, Fall 2009, Winter 2009, Summer 2011, Graduate course in Stochastic Processes
  – STATS 305: Fall 2010, PhD first year course in Applied Statistics
  – STATS 315A: Winter 2011, Modern Applied Statistics

XII Student and Thesis Supervision

1. DOCTORAL THESES SUPERVISED

(i) Hussein Hazimeh, MIT Operations Research Center, 2016–2021
PhD in Operations Research
Title: *Sparse Learning using Discrete Optimization: Scalable Algorithms and Statistical Insights*
Current Position: Research Scientist, Google Research, NY.

2. MASTER'S THESES SUPERVISED

(i) Antoine Dedieu, MIT Operations Research Center, 2016–2018
MS in Operations Research
Title: *Sparse learning: Statistical and optimization perspectives*
Current Position: Researcher at Vicarious AI, CA.

(ii) Zachary Blanks, MIT Operations Research Center, 2017– 2019
MS in Operations Research (co-advisor Dr. Troy Lau, DRAPER)
Title: *A Generalized Hierarchical Approach for Data Labeling*
Current Position: PhD Candidate at UVA School of Data Science

(iii) Eamonn Shirey, MIT Leaders for Global Operations Program, 2017-2019
Joint MS and MBA in the LGO program (co-advisor David E Hardt, MIT)
Title: *Predicting Jet Engine Component Wear to Enable Proactive Fleet Maintenance*
Current Position: Associate Director, Operations Analytics at Wayfair

(iv) Suyash Gupta, Indian Statistical Institute, 2017
MStat in Statistics
Title: *Sentiment Analysis*
(Co-supervisor: Indranil Mukhopadhyay, Indian Statistical Institute)
Current Position: PhD candidate, Stanford University.

3. THESES IN PROGRESS

- Currently serving as the thesis advisor for the following students:

Name	*Program*	*Degree*	*Year*
Wenyu Chen	MIT Operations Research Center	PhD	2018–
Kayhan Behdin	MIT Operations Research Center	PhD	2019–
Shibal Ibrahim	MIT EECS	PhD	2019–
Haoyue Wang	MIT Operations Research Center	PhD	2019–
Brian Liu	MIT Operations Research Center	PhD	2021–
Xiang Meng	MIT Operations Research Center	PhD	2021–
Riade Benbanki	MIT Operations Research Center	MS	2021–

4. DOCTORAL THESIS COMMITTEE MEMBER

- PhD Thesis Committee Member inside and outside MIT (does not include students supervised by me):

  – Driss Lahlou Kitane (MIT ORC, advisor D. Bertsimas) 2021
  – Peter Cohen (MIT ORC, advisor Colin Fogarty) 2021
  – Lea Kapelevich (MIT ORC, advisor Juan Pablo Vielma) 2021
  – Chris Coey (MIT ORC, advisor Juan Pablo Vielma) 2021
  – Guanyi Wang (Georgia Institute of Technology, advisor Santanu Dey) 2021
  – Hui Li (Harvard University, advisor Xihong Lin) 2021
  – Gabriel Loewinger (Harvard University, advisor Giovanni Parmigiani) 2020-2022
  – Matthew Sobiesk (MIT ORC, advisor D. Bertsimas) 2020
  – Hari Bandi (MIT ORC, advisor D. Bertsimas) 2020
  – Yuchen Wang (MIT ORC, advisor D. Bertsimas) 2019-2020
  – Haihao Lu (MIT ORC and Math, advisor R. Freund) 2019
  – Edward Cheng (Columbia Statistics, advisor D. Madigan) 2018
  – Jack Dunn (MIT ORC, advisor D. Bertsimas) 2018
  – Andrew Li (MIT ORC, advisor Vivek Farias) 2018
  – Chris Marks (MIT ORC, advisor T. Zaman) 2017
  – Paul Grigas (MIT ORC, advisor R. Freund) 2016
  – Diego Saldana (Columbia Statistics, advisor Y. Feng) 2016

- MIT Operations Research Center, PhD Generals Committee Member (does not include students supervised by me):

  – Zhen Lin (advisor D. Bertsimas) 2021
  – Cynthia Zheng (advisor D. Bertsimas) 2021
  – Vassilis Digalakis (advisor D. Bertsimas) 2020
  – Michael Li (advisor D. Bertsimas) 2020
  – Peter Cohen (advisor Colin Fogarty) 2019
  – Lea Kapelevich (advisor Juan Pablo Vielma) 2019
  – Haihao Lu (advisor Robert Freund) 2018
  – Hari Bandi (advisor D. Bertsimas) 2018
  – Matthew Sobiesk (advisor D. Bertsimas) 2018
  – Brad Sturt (advisor D. Bertsimas) 2017
  – Julia Yan (advisor D. Bertsimas) 2017
  – Lauren Berk (advisor R. Freund) 2016
  – Jack Dunn (advisor D. Bertsimas) 2016

– Colin Pawlowski (advisor D. Bertsimas) 2016
– Daisy Zhou (advisor D. Bertsimas) 2016

5. Other Student Supervision

– Mathieu Jonathan Paul Sibue (MIT MBAN, RA) 2021-2022
– Aniruddh Hari (MIT MBAN Capstone) 2021
– Rebecca Schubertruegmer (MIT MBAN Capstone) 2021
– Peijun Xu (MIT MBAN, RA) 2020-2021
– Brian Hsu (MIT MBAN, RA) 2020-2021
– Xiaming Jin (MIT MBAN, RA) 2020-2021
– Denis Sai (MIT MBAN, RA) 2020-2021
– Zion Hadley (MIT UROP, RA) 2020
– Jingjing Piao (MIT MBAN Capstone) 2020
– Alessandro Previero (MIT MBAN Capstone) 2020
– Gabrielle Rappaport (MIT MBAN Capstone) 2020
– Raphaelle Diane Delpont (MIT MBAN Capstone) 2020
– Danying Xiao (MIT MBAN, RA) 2019-2020
– Timothy Nonet (MIT MBAN, RA) 2019-2020
– Francois Pierre Caprasse (MIT MBAN Capstone) 2019
– Alessandro Scaglia (MIT MBAN Capstone) 2019
– Jocelyn Beauchesne (MIT MBAN Capstone) 2019
– Johnny Oh (MIT MBAN Capstone) 2019
– Annelise Steele (MIT MBAN, RA) 2018-2019
– Subhashree Rengarajan (MIT MBAN Capstone) 2018
– Yingtian Yang (MIT MBAN Capstone) 2018
– Bibek K Pandit (MIT UROP, RA) 2018
– Tiffany Chen (Wellesley UROP, RA) 2018
– Yifei Huang (MIT MBAN Capstone) 2017
– Tianzhao Wu (MIT MBAN Capstone) 2017
– Zakaria El Hjouji (MIT MBAN RA) 2017
– Erica Chan (MIT UROP, RA) 2017
– Kliment Serafimov (MIT UROP, RA) 2017
– Suyash Gupta (BS/MS Indian Statistical Institute and Columbia Intern, RA), 2015-2016
– Debarghya Mukherjee (BS/MS Indian Statistical Institute), 2016
– Nabarun Deb (BS/MS Indian Statistical Institute), 2016
– Koulik Khamaru (BS/MS Indian Statistical Institute and Columbia Intern, RA), 2013-2015
– Jonathan Goetz (BA/MA Columbia University, RA), 2014-2015
– Wodan Lin (PhD student RA, Columbia University, Biostatistics, 2014-2015)

XIII Publications

Theses

Phd: "Topics in Sparse Multivariate Statistics", Stanford University, Dept. of Statistics, 2012
Advisor: Dr. Trevor J. Hastie
Committee: Emmanuel Candes, Jerome Friedman, Rob Tibshirani and Michael Saunders (external).

Masters: "Local scale-space contrasts via Gaussian mixture ensembles for speech signal segmentation".
Indian Statistical Institute, 2007.
Advisor: Dr. Debapriya Sengupta.
(awarded the *TCS award for best Masters thesis*, Indian Statistical Institute, 2005-2007.)

Refereed Journal Articles

Note: (i) Students supervised by me denoted by an asterisk *. Other student co-authors denoted by **.

(ii) Papers where all authors are alphabetically ordered are indicated with the footnote[†]. In other cases, when a subset of authors have equal contribution, the authors are indicated by the footnote[‡]. In all other cases, the first author is the primary contributor.

Methodological Publications

[J1] *"Solving L1-regularized SVMs and related linear programs: Revisiting the effectiveness of Column and Constraint Generation"*
[†]Antoine Dedieu*, Rahul Mazumder and Haoyue Wang*, 2021+
Journal of Machine Learning Research (accepted pending minor revisions)
`https://arxiv.org/abs/1901.01585` [stat.ML]

[J2] *"Sparse Regression at Scale: Branch-and-Bound rooted in First-Order Optimization"*
[†]Hussein Hazimeh*, Rahul Mazumder and Ali Saab**, 2021
Mathematical Programming (to appear)
`https://arxiv.org/abs/2004.06152` [stat.CO]
► *MIT Operations Research Center Best Student Paper Award, 2020. (Awardee: Hazimeh)*
► *INFORMS Computing Society Student Paper Prize (honorable mention), 2020. (Awardee: Hazimeh)*

[J3] *"Using ℓ_1-relaxation and integer programming to obtain dual bounds for sparse PCA"*
[†]Santanu Dey, Rahul Mazumder and Guanyi Wang**, 2021+.
Operations Research (to appear)
`https://arxiv.org/abs/1810.09062` [math.OC]

[J4] *"Learning Sparse Classifiers: Continuous and Mixed Integer Optimization Perspectives"*
[†]Antoine Dedieu*, Hussein Hazimeh* and Rahul Mazumder, 2021
Journal of Machine Learning Research

[J5] *"Randomized Gradient Boosting Machine"*
[†]Haihao Lu** and Rahul Mazumder, 2020
SIAM Journal on Optimization

[J6] *"Fast Best Subset Selection: Coordinate Descent and Local Combinatorial Optimization Algorithms"*
[†]Hussein Hazimeh* and Rahul Mazumder, 2020
Operations Research
► *Awarded the INFORMS Optimization Society Young Researchers Prize, 2020. (Awardees: Hazimeh and Mazumder)*

[J7] *"Computing the degrees of freedom of rank-regularized estimators and cousins"*
[†]Rahul Mazumder and Haolei Weng**, 2020
Electronic Journal of Statistics

[†]All authors in this paper are arranged alphabetically
[‡]These authors contributed equally

[J8] *"Matrix completion with nonconvex regularization: spectral operators and scalable algorithms"*
†Rahul Mazumder, Diego Saldana** and Haolei Weng**, 2020
Statistics and Computing

[J9] *"Computation of the Maximum Likelihood estimator in low-rank Factor Analysis"*
†Koulik Khamaru* and Rahul Mazumder, 2019
Mathematical Programming

[J10] *"A Computational Framework for Multivariate Convex Regression and its Variants"*
Rahul Mazumder, Arkopal Choudhury**, Garud Iyengar and Bodhisattva Sen, 2019
Journal of the American Statistical Association

[J11] *"Learning a Mixture of Gaussians via Mixed Integer Optimization"*
†Hari Bandi*, Dimitris Bertsimas and Rahul Mazumder, 2019
Informs Journal on Optimization

[J12] *"Flexible low-rank statistical modeling with missing data and side information"*
†William Fithian and Rahul Mazumder, 2018
Statistical Science

[J13] *"Certifiably Optimal Low Rank Factor Analysis"*
†Dimitris Bertsimas, Martin Copenhaver** and Rahul Mazumder, 2017
Journal of Machine Learning Research

[J14] *"The Discrete Dantzig Selector: Estimating Sparse Linear Models via Mixed Integer Linear Optimization"*
†Rahul Mazumder and Peter Radchenko, 2017
IEEE Transactions on Information Theory

[J15] *"An Extended Frank-Wolfe Method with "In-Face" Directions, and its Application to Low-Rank Matrix Completion"*
†Robert Freund, Paul Grigas** and Rahul Mazumder, 2017
SIAM Journal on Optimization

[J16] *"A New Perspective on Boosting in Linear Regression via Subgradient Optimization and Relatives"*
†Robert Freund, Paul Grigas** and Rahul Mazumder, 2017
Annals of Statistics

► *Editors' choice for one of the best four papers accepted to the Annals of Statistics in the previous two years. Special Invited Session presentation by Mazumder at the Joint Statistical Meetings, 2017.*
► *INFORMS Optimization Society Student Paper Award, 2015. (Awardee Grigas)*

[J17] *"Best Subset Selection via a Modern Optimization Lens"*
†Dimitris Bertsimas, Angela King** and Rahul Mazumder, 2016
Annals of Statistics

[J18] *"Matrix Completion and Low-Rank SVD via Fast Alternating Least Squares"*
Trevor Hastie‡, Rahul Mazumder‡, Jason Lee and Reza Zadeh, 2015
Journal of Machine Learning Research

[J19] *"Least Quantile of Squares Regression via Modern Optimization"*
†Dimitris Bertsimas and Rahul Mazumder, 2014
Annals of Statistics

[J20] *"The Graphical Lasso: New Insights and Alternatives"*
Rahul Mazumder and Trevor Hastie, 2012
Electronic Journal of Statistics

[J21] *"Exact covariance thresholding into connected components for large-scale Graphical Lasso"*
Rahul Mazumder and Trevor Hastie, 2012
Journal of Machine Learning Research

[J22] *"SparseNet: Coordinate Descent with Non-Convex Penalties"*
Rahul Mazumder, Jerome Friedman and Trevor Hastie, 2011
Journal of American Statistical Association, Theory and Methods

[J23] *"Spectral Regularization Algorithms for Learning Large Incomplete Matrices"*
Rahul Mazumder, Trevor Hastie and Robert Tibshirani, 2010
Journal of Machine Learning Research

Invited Discussion Articles

[J24] *"Discussion of "Best Subset, Forward Stepwise or Lasso? Analysis and Recommendations Based on Extensive Comparisons""*
Rahul Mazumder, 2020
Statistical Science

Application oriented publications

[J25] *"Analysis of correlations between local geographic atrophy growth rates and local OCT angiography-measured choriocapillaris flow deficits"*
Eric M Moult**, Yingying Shi, Qinqin Zhang, Liang Wang, Rahul Mazumder, Siyu Chen, Zhongdi Chu, William Feuer, Nadia K Waheed, Giovanni Gregori, Ruikang K Wang, Philip J Rosenfeld, James G Fujimoto, 2021
Biomedical Optics Express

[J26] *"Integration of Survival Data from Multiple Studies"*
Steffen Ventz, Rahul Mazumder[‡] and Lorenzo Trippa[‡], 2021
Biometrics

[J27] *"Mining Events with Declassified Diplomatic Documents"*
Yuanjun Gao**, Jack Goetz**, Matthew Connelly and Rahul Mazumder, 2020
Annals of Applied Statistics

[J28] *"Assessing the significance of global and local correlations under spatial autocorrelation: a nonparametric approach"*
Julia Validomat, Rahul Mazumder, Alex McInturff, Douglas McCauley and Trevor Hastie, 2014
Biometrics

[J29] *"Turbulence, suspension and downstream fining over a sand-gravel mixture bed"*
Koeli Ghoshal, Rahul Mazumder, Chandan Chakraborty and Bijoy Mazumder, 2013
International Journal of Sediment Research

[J30] *"Modeling Item-Item Similarities for Personalized Recommendations on Yahoo! Front Page"*
Deepak Agarwal, Liang Zhang and Rahul Mazumder, 2011
Annals of Applied Statistics

[J31] *"Fluid flow pattern analysis in a trough region: a nonparametric approach"*
Rahul Mazumder, 2008
Journal of Applied Statistics

[J32] *"Statistical characterization of circulation patterns and direction of turbulent flow over a waveform structure"*
Rahul Mazumder and Bijoy Mazumder, 2006
Environmetrics

[J33] *"Clustering based on geometry and interactions of turbulence bursting rate processes in a trough region"*
Rahul Mazumder, 2007
Environmetrics

Articles in Refereed Conference Proceedings

[C1] *"DSelect-k: Differentiable Selection in the Mixture of Experts with Applications to Multi-Task Learning"*
Hussein Hazimeh*, Zhe Zhao, Aakanksha Chowdhery, Maheswaran Sathiamoorthy, Yihua Chen, Rahul Mazumder, Lichan Hong, and Ed H. Chi, 2021
Neural Information Processing Systems (NeurIPS) (to appear)
`https://arxiv.org/abs/2106.03760` [cs.LG]

[C2] *"Linear Regression with Mismatched Data: a Provably Optimal Local Search Algorithm"*
†Rahul Mazumder and Haoyue Wang*, 2021
Integer Programming and Combinatorial Optimization (IPCO)
Journal version of this paper (=R4) is under review.

[C3] *"ECLIPSE: An Extreme-Scale Linear Program Solver for Web-Applications"*
†Kinjal Basu, Amol Ghoting, Rahul Mazumder and Yao Pan, 2020
Proceedings of the 37th International Conference of Machine Learning (ICML)

[C4] *"The Tree Ensemble Layer: Differentiability meets Conditional Computation"*
Hussein Hazimeh*, Natalia Ponomareva, Petros Mol, Zhenyu Tan and Rahul Mazumder, 2020
Proceedings of the 37th International Conference on Machine Learning (ICML)

[C5] *"Learning Hierarchical Interactions at Scale: A Convex Optimization Approach"*
†Hussein Hazimeh* and Rahul Mazumder, 2020
23rd International Conference on Artificial Intelligence and Statistics (AISTATS)

► *Mixed Integer Programming Workshop student poster award (honorable mention), 2019. (Awardee: Hazimeh).*

[C6] *"Hierarchical Modeling and Shrinkage for User Session Length Prediction in Media Streaming"*
Antoine Dedieu*, Rahul Mazumder, Zhen Zhu and Hossein Vahabi, 2018
Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM).

[C7] *"Non-Negative Matrix Completion for Bandwidth Extension: A Convex Optimization Approach"*
Dennis Sun and Rahul Mazumder, 2013
IEEE Machine Learning for Signal Processing

[C8] *"Projected likelihood contrasts for testing homogeneity in finite mixture models with nuisance parameters"*
Debapriya Sengupta and Rahul Mazumder, 2008
Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen (Beachwood, Ohio, USA: Institute of Mathematical Statistics), 272-281.

Articles in Progress or under Review

Articles with recently decisioned revision requests

[R1] *"Frank-Wolfe Methods with an Unbounded Feasible Region and Applications to Structured Learning"*
Haoyue Wang*, Haihao Lu‡ and Rahul Mazumder‡, 2020
First revision submitted, SIAM Journal on Optimization
`https://arxiv.org/abs/2012.15361` [math.OC]

[R2] *"Grouped Variable Selection with Discrete Optimization: Computational and Statistical Perspectives"*
†Hussein Hazimeh*, Rahul Mazumder and Peter Radchenko, 2021.
First revision submitted, Annals of Statistics
`https://arxiv.org/abs/2104.07084` [stat.ME]

[R3] *"Subset Selection with Shrinkage: Sparse Linear Modeling when the SNR is low"*
Rahul Mazumder‡, Peter Radchenko‡ and Antoine Dedieu*, 2021
Revision submitted, Operations Research
`https://arxiv.org/abs/1708.03288` [stat.ME]

[R4] *"Linear regression with partially mismatched data: local search with theoretical guarantees"*
†Rahul Mazumder and Haoyue Wang*, 2021
Revision requested, Mathematical Programming
`https://arxiv.org/abs/2106.02175` [math.OC]
An extended abstract of this paper appeared in IPCO

[R5] *"Archetypal Analysis for Sparse Nonnegative Matrix Factorization: Robustness Under Misspecification"*
†Kayhan Behdin* and Rahul Mazumder, 2021
Reject with encouragement to resubmit, Journal of Machine Learning Research
`https://arxiv.org/abs/2104.03527` [stat.ML]

[R6] *"Multivariate convex regression at scale"*
†Wenyu Chen* and Rahul Mazumder, 2020
Revision requested, SIAM Journal on Optimization
`https://arxiv.org/abs/2005.11588` [math.OC]

Other articles under review or in Progress

[R7] *"Sparse PCA: A New Scalable Estimator Based On Integer Programming"*
†Kayhan Behdin* and Rahul Mazumder, 2021
`https://arxiv.org/abs/2109.11142` [stat.ME]
Status: Submitted

► *Mixed Integer Programming Workshop student poster award (most popular poster) 2021. (Awardee: Behdin).*

[R8] *"Predicting Census Survey Response Rates via Interpretable Nonparametric Additive Models with Structured Interactions"*
Shibal Ibrahim*, Rahul Mazumder, Peter Radchenko, and Emanuel Ben-David, 2021
`https://arxiv.org/abs/2108.11328` [stat.ML]
Status: Preprint, in preparation for submission.

[R9] *"Nonparametric Finite Mixture Models with Possible Shape Constraints: A Cubic Newton Approach."*
Haoyue Wang*, Shibal Ibrahim*, and Rahul Mazumder, 2021
`https://arxiv.org/abs/2107.08535` [stat.CO]
Status: Submitted

[R10] *"A new computational framework for log-concave density estimation"*
†Wenyu Chen*, Rahul Mazumder and Richard J. Samworth, 2021
`https://arxiv.org/abs/2105.11387` [stat.CO]
Status: Submitted

[R11] *"Optimal Ensemble Construction for Multi-Study Prediction with Applications to COVID-19 Excess Mortality Estimation"*
Gabriel Loewinger**, Rolando Acosta Nunez**, Rahul Mazumder, Giovanni Parmigiani, 2021
`https://arxiv.org/abs/2109.09164` [stat.ML]
Status: Paper Submitted

► *Best Abstract award at Annual Symposium Brigham Health/Harvard Medical School Computational Neuroscience Outcomes Center, 2020.*

Technical Reports

[T1] *"A convex optimization framework for gene-level tissue network estimation with missing data and its application in understanding disease architecture"*
†Kushal Dey and Rahul Mazumder, 2020
`https://www.biorxiv.org/content/10.1101/2020.03.16.994020v1`

[T2] *"Computing Estimators of Dantzig Selector type via Column and Constraint Generation."*
†Rahul Mazumder, Stephen Wright, and Andrew Zheng*, 2019
`https://arxiv.org/abs/1908.06515`

[T3] *"Condition number analysis of logistic regression, and its implications for standard first-order solution methods"*
Robert Freund, Paul Grigas and Rahul Mazumder, 2018
`https://arxiv.org/abs/1810.08727`

[T4] *"The trimmed lasso: Sparsity and robustness."*
†Dimitris Bertsimas, Martin Copenhaver**, and Rahul Mazumder, 2017
`https://arxiv.org/abs/1708.04527`

[T5] *"Scalable computation of regularized precision matrices via stochastic optimization*
Yves Atchade‡, Rahul Mazumder‡, and Jie Chen, 2015
`https://arxiv.org/abs/1509.00426`

XIV Invited Oral Presentations

- "Multi-Task Learning with Soft Tree Ensembles with applications to Insurance Pricing"
  – Global Data Science Forum, Liberty Mutual Insurance 2021 (Oct)
- "Extreme Scale Linear Programming"
  – SIAM Optimization Conference 2021 (July)
- " Sparse Learning at Scale: Convex, Mixed Integer Programming, and Statistical Perspectives"
  – University of Minnesota, 2021 (Nov)
  – University of Michigan, 2021 (Oct)

- "Predictive and Prescriptive Analytics at Scale"
  – MIT Industrial Liaison Program and MIT Quest for Intelligence 2020 (July)
- "Solving Large Scale Linear Programs In Machine Learning Tasks"
  – INFORMS Annual Meeting 2019 (October)
- "Structured Learning at Scale: Continuous and Mixed Integer Programming Perspectives"
  – INFORMS Annual Meeting 2021 (Nov)
  – INFORMS Annual Meeting 2020 (Oct)
  – University of Chicago, 2020 (Oct)
  – Baidu Research, 2020 (April)
  – INFORMS Annual Conference, 2020 (Oct)
  – New York University (Data Science), 2019 (September)
  – Joint Statistical Meetings (JSM), 2019 (July)
  – Mixed Integer Programming Workshop, 2019 (July)
  – ICCOPT Berlin[‖], 2019 (August)
  – ONR Program Review, 2019 (October)
  – University of Chicago, 2019 (November)
  – INFORMS Annual Meeting 2019 (October)
  – George Washington University 2019 (April)
  – Indian Statistical Institute, 2019 (January)
- "Mining Events with Declassified Diplomatic Documents"
  – US Census Bureau, 2018 (August)
- "Factor Analysis via Modern Optimization"
  – INFORMS Optimization Society Conference, 2018 (March)
- "Solving Structured Nonconvex Problems in Statistics"[¶]
  – International Conference on Big Data and Information Analytics (BigDIA) 2018 (December)
  – INFORMS Annual Meeting 2018 (November)
  – Duke University, Fuqua School of Business, 2018 (November)
  – The Fields Institute for Research in Mathematical Sciences, University of Toronto, 2018 (September)
  – Pandora Media Inc, 2018 (September)
  – Adobe Research, 2018 (September)
  – International Indian Statistical Association (IISA), 2018 (May)
  – University of Chicago Booth School of Business, 2018 (May)
  – Harvard University, Statistics Department Seminar, 2018 (March)
  – University of Cambridge, Isaac Newton Institute, 2018 (March)
  – Conference on Optimisation and Machine Learning in Economics, University College of London, 2018 (March)
  – Michigan State University (Machine Learning Seminar Series, joint seminar across Statistics, Computer Science and Computational Mathematics Science and Engineering departments), 2018 (April)
  – George Mason University (Statistics and Operations Research joint seminar), 2017 (December)
  – INFORMS Annual Meetings, 2017 (October)
  – U.S. Census Bureau, 2017 (August)
  – 61st ISI World Statistics Congress (WSC) Marrakech, Morocco, 2017 (July)
  – Conference on Nonconvex Statistical Learning (CNSL), University of Southern California, 2017 (July)
  – ICSA Applied Statistics Symposium, Chicago, 2017 (June)

[‖] Cancelled talk due to unavoidable circumstances
[¶] The exact title and content varied, but the topic was the same.

– Georgia Institute of Technology, ISYE department, 2017 (May)
– Plenary talk at SAMSI, 2016 (September)

- "A New Perspective on Boosting in Linear Regression via Modern Optimization"
  – Joint Statistical Meetings, 2017 (August).
  – SIAM Conference on Optimization, British Columbia, 2017 (May)
  – INFORMS Annual Meetings, 2016 (November)
  – International Indian Statistical Association Conference, 2016 (August)
- "An Algorithmic Approach to Nonparametric Function Estimation with Shape Constraints"
  – Joint Statistical Meetings, 2020
  – IMS at National University of Singapore, 2016
- "Exact Subset Selection in Regression via Modern Optimization"
  – Joint Statistical Meetings (JSM), 2016
  – Lehigh University, 2015
  – University of Pennsylvania, 2015
  – Princeton University, 2015
  – University of Southern California, 2015
  – University of Michigan, 2015
  – 2-Sigma Investments, 2015
  – CMStatistics conference, London, UK, 2015
  – Stanford University, 2014
- "Burstiness Analysis: Detecting Events Through Traffic Analysis by Geography and Subject"
  – Famine and feast, international historical research in the digital age: Workshop, London School of Economics, 2015 (January)
- "Learning with Low Rank Matrices: Flexible Modeling and Scalable Computation"
  – Joint Statistical Meetings, 2019
  – MIT, 2014
  – Joint Statistical Meetings, 2014
  – Pandora Media, Inc., 2014
- "Factor Analysis via a Modern Optimization Lens"
  – 2014 INFORMS Optimization Society Conference in Houston, TX in March.
- "Low-rank Matrix Completion: Statistical Models and Large Scale Algorithms"
  – MIT Stochastics and Statistics Seminar, November 2013.
- "Convex Regularization Algorithms for Learning Large Incomplete Matrices"
  – University of Southern California, January 2012
  – University of Florida, January 2012
  – University of Chicago, January 2012
  – Cornell University, January 2012
  – University of Michigan, January 2012
  – Rutgers University, January 2012
  – University of Wisconsin Madision, February 2012
  – Columbia University, February 2012
  – MIT, February 2012
  – Harvard University , February 2012
  – Princeton University, February 2012
  – Carnegie Mellon University, February 2012.

- "Time-Sensitive Collaborative Filtering via Item-Item Similarities"
  – Joint Statistical Meetings, 2011

- "Spectral Regularization Algorithms for Learning Large Incomplete Matrices"
  – Eastern North American Region (ENAR) Annual Meeting, 2011
  – Joint Statistical Meetings, 2010

- "Regularization Methods for Learning Large Incomplete Matrices"
  –SIAM Conference on Mathematics for Industry: Challenges and Frontiers. October 2009, SF, CA.
  –SF Bay Area Chapter of the American Statistical Association, Hayward, CA, June 2011.
  –Industrial Affiliates Annual Conference, Stanford University, May 2010.
  –Berkeley Stanford Joint Student Colloquium, Stanford University, April 2011.
  –Stanford Statistics Students' Retreat Conference, Asilomar, CA, 2011 & 2009.