

Paxton Turner

Department of Statistics, Harvard University | <https://paxtonturner.github.io> | paxtonturner@fas.harvard.edu

RESEARCH INTERESTS AND SKILLS

I develop and analyze efficient statistical procedures for machine learning and data science using tools from high-dimensional statistics and discrete mathematics. My research works span statistical machine learning, algorithms, network science, natural language processing, data compression, and experimental design.

Computer skills: Python, R, Matlab, Mathematica, LaTeX, C++

ACADEMIC POSITION AND EDUCATION

Harvard University	June 2021 — Present
<ul style="list-style-type: none">• Postdoctoral fellow, Department of Statistics• <i>Supervisor:</i> Zheng Tracy Ke	
Massachusetts Institute of Technology (MIT)	Sept. 2016 — June 2021
<ul style="list-style-type: none">• Ph.D. in Applied Mathematics• <i>Advisor:</i> Philippe Rigollet	
Louisiana State University (LSU)	Aug. 2011 — May 2015
<ul style="list-style-type: none">• B.S. in Mathematics	

SELECTED RESEARCH PROJECTS

Balancing covariates in high dimension	Mar. 2018 — Present
<ul style="list-style-type: none">• Inspired by causal inference, investigated the problem of dividing datapoints into two groups with similar averaged covariates<ul style="list-style-type: none">– Established best known algorithms and computational lower bounds for random datapoints– Developed a near-optimal online algorithm for a new relaxation of this problem• <i>Related areas:</i> A/B testing, high-dimensional statistics, algorithms, combinatorial optimization	
Data compression and estimation	Sept. 2019 — April 2021
<ul style="list-style-type: none">• Exhibited trade-offs between statistical accuracy and data compression in nonparametric estimation<ul style="list-style-type: none">– Developed interpolation-based and subsampling-based algorithms for reducing the time and space complexity of kernel density estimators to sub-linear while achieving optimal accuracy• <i>Related areas:</i> nonparametric estimation, algorithms, numerical analysis, discrete geometry	
Statistical similarity analysis of text data	Aug. 2021 — Present
<ul style="list-style-type: none">• Developed and analyzed the De-biased and Length-Adjusted Variability Estimator (DELVE) to measure statistical similarity among collections of multinomial datasets, such as word counts of text corpora<ul style="list-style-type: none">– Using R & Python, applied DELVE to (i) Amazon movie reviews and (ii) statistics journal abstracts– DELVE identifies movies with polarized reviews and researchers with diverse interests• <i>Related areas:</i> natural language processing, topic modeling, high-dimensional statistics	
Community detection in social networks	Feb. 2022 — Present
<ul style="list-style-type: none">• Established a provably optimal computationally efficient method based on signed polygon counts for detecting small, tight-knit communities in large degree-heterogeneous networks<ul style="list-style-type: none">– Algorithm implemented in Matlab detects cliques of collaborators in co-authorship networks• <i>Related areas:</i> statistical network analysis, algorithms, complexity theory	

PUBLICATIONS AND PREPRINTS

(† indicates alphabetical ordering. Underlined hyperlinks lead to corresponding paper.)

<u>Testing High-dimensional Multinomials with Applications to Text Analysis</u> T. Tony Cai, Zheng Tracy Ke, and Paxton Turner†	<i>Under review</i>
<u>Near-optimal Fitting of Ellipsoids to Random Points</u> Aaron Potechin, Paxton Turner, Prayaag Venkat, and Alex Wein†	<i>Under review</i>
<u>Phase Transition for Detecting a Small Community in a Large Network</u> Jiashun Jin, Zheng Tracy Ke, Paxton Turner, and Anru Zhang†	<i>To appear at ICLR (2023)</i>
<u>Gaussian Discrepancy: a Probabilistic Relaxation of Vector Balancing</u> Sinho Chewi, Patrik Gerber, Philippe Rigollet, and Paxton Turner†	<i>Discret. Appl. Math (2022)</i>
<u>A Statistical Perspective on Coreset Density Estimation</u> Paxton Turner, Jingbo Liu, and Philippe Rigollet	<i>AISTATS (2021)</i>
<u>Efficient Interpolation of Density Estimators</u> Paxton Turner, Jingbo Liu, and Philippe Rigollet	<i>AISTATS (2021)</i>
<u>Balancing Gaussian Vectors in High Dimension</u> Paxton Turner, Raghu Meka, and Philippe Rigollet	<i>COLT (2020)</i>
<u>Discrete Equidecomposability and Ehrhart Theory of Polygons</u> Paxton Turner and Yuhuai Wu†	<i>Discrete and Computational Geometry (2020)</i>
<u>Efficient Reconstruction of Stochastic Pedigrees</u> Younhun Kim, Elchanan Mossel, Govind Ramnarayan, and Paxton Turner†	<i>arXiv preprint (2020)</i>
<u>Aztec Castles and the dP3 Quiver</u> Megan Leoni, Gregg Musiker, Seth Neel, and Paxton Turner†	<i>J. Phys. A Math Theor. (2014)</i>

PRESENTATIONS

Discrete Models and Methods: from Multinomial Testing to Data Compression Talk at <i>Stat 300 Seminar</i> , Harvard University	Nov. 2022
Testing Variability of Multinomial Data with Applications to Text Analysis Invited talk at <i>Stanford Statistics Seminar</i>	Oct. 2022
Detection of Heterogeneous Documents in a Corpus Poster at <i>Statistical Network Analysis & Beyond (SNAB 2022)</i> , New York University	Aug. 2022
Detecting a Small Community in a Large Network Invited talk at <i>SIAM Conference on Discrete Mathematics</i> , Carnegie Mellon University	June 2022
A Statistical Perspective on Coresets Talk at <i>Stat 300 Seminar</i> , Harvard University	Oct. 2021
A Statistical Perspective on Coreset Density Estimation Poster & talk at <i>Int'l Conference on Artificial Intelligence and Statistics (AISTATS)</i> , Virtual	April 2021
Efficient Interpolation of Density estimators Poster & talk at <i>Int'l Conference on Artificial Intelligence and Statistics (AISTATS)</i> , Virtual	April 2021
A Statistical Perspective on Coresets Invited talk at <i>Yale Institute for Network Science Seminar</i> , Virtual	Dec. 2020
Balancing Gaussian Vectors in High Dimension Talk at <i>Conference on Learning Theory (COLT)</i> , Virtual	July 2020

AWARDS

Levinson Fellowship — Department of Mathematics, MIT	2017
LSU University Medalist (4.0/4.0 GPA)	2015
Betti and Robert Giles Senior Mathematics Award — Department of Mathematics, LSU	2015
Barry M. Goldwater Scholarship Recipient	2014
LSU College of Science Outstanding Junior	2014
1st Prize - LSU Undergrad. Research Conference: Math & Phys. Sciences Oral Presentations	2013
Pasquale Porcelli Junior Scholarship — Department of Mathematics, LSU	2013
LSU Chancellor's Alumni Scholarship (full tuition, room and board)	2011–2015
LSU Science Honors Scholarship	2011–2015

TEACHING & SERVICE

- In person and online teaching experience for one undergraduate course (30-60 students) and three graduate courses (20+ students) in probability and statistics at MIT
- Reviewer for ICALP 2020, AISTATS 2021, UAI 2021 (top reviewer), Theory of Computing Systems, Neurips 2021 (top reviewer), ICLR 2022, AISTATS 2022, FOCS 2022, Neurips 2022 (top reviewer)

LEADERSHIP AND MENTORSHIP

<i>Co-organizer</i> of MIT <u>Integration Bee</u>	2016 — 2020
<i>Co-organizer</i> of <u>SPAMS seminar</u>	Sept. 2017 — May 2018
<i>Mentor</i> for <u>SPUR Undergraduate Research Program</u>	Summer 2017
<i>Mentor</i> for <u>MIT Directed Reading Program</u>	January 2017
<i>Tutor</i> for <u>MIT ESL Program for Facilities Workers</u>	Fall 2016 — Summer 2017
<i>Instructor</i> for <u>LSU Math Circle Summer Enrichment Program</u>	Summers 2015 & 2016