Paxton Turner

Curriculum Vitae

Department of Statistics
Harvard University
Science Center 400 Suite
One Oxford Street
Cambridge, MA 02138-2901 ✉ *paxtonturner@g.harvard.edu*

Academic positions

2021–2023. **Postdoctoral researcher**, *Harvard University*, Advisor: Zheng Tracy Ke.

Education

2016–2021 **Ph.D. in Applied Mathematics**, *Massachusetts Institute of Technology (MIT)*, Thesis: *Combinatorial methods in statistics*, Advisor: Philippe Rigollet.

2011–2015 **B.S. in Mathematics**, *Louisiana State University (LSU)* .

Research Interests

High dimensional statistics, minimax theory, network inference, coresets

Current projects

- Estimating heterogeneity in multinomial data with applications to text analysis . With T. Tony Cai and Zheng Tracy Ke.
  - We develop an estimator for the between-group variance of a collection of samples of multinomial distributions and apply our results to two real-world datasets: (i) Amazon reviews and (ii) abstracts from popular statistics journals.
- Phase transition for detecting a small community in a large network. With Jiashun Jin, Zheng Tracy Ke, and Anru Zhang. (Preprint available on request.)
  - We show that the signed quadrilateral test is optimal among computationally efficient estimators for detecting a planted community in a popular model for heterogeneous networks.
- Near-optimal ellipsoid fitting on random points. With Prayaag Venkat and Alex Wein.
  - When is it possible to interpolate a set of n Gaussian vectors in dimension d with an ellipsoid? This question has connections to many problems in machine learning such as independent component analysis. We show ellipsoid fitting is possible when $n \leq d^{5/3}$, improving on prior work that achieved $n \leq d^{3/2}$.
- Estimating the unseen revisited: correlated species. With Morgane Austern and Zheng Tracy Ke.
  - A biologist records the different species observed over a period of N days. If they stay an extra N days, how many new species will be observed? We study this classical statistical question in a new setting where the observations are correlated.

Publications

- Sinho Chewi, Patrik Gerber, Philippe Rigollet, and Paxton Turner. Gaussian discrepancy: a probabilistic relaxation of vector balancing. https://arxiv.org/abs/2109.08280. 2021.
  - We describe a new relaxation for vector balancing and establish an optimal online algorithm.
- Paxton Turner, Jingbo Liu, and Philippe Rigollet. A Statistical Perspective on Coreset Density Estimation. http://proceedings.mlr.press/v130/turner21b.html. *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*. (AISTATS 2021).
  - Coresets are a powerful new framework for data compression. We establish optimal statistical guarantees for coreset-based estimators in a canonical density estimation problem.
- Paxton Turner, Jingbo Liu, and Philippe Rigollet. Efficient Interpolation of Density Estimators. http://proceedings.mlr.press/v130/turner21a.html. *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*. (AISTATS 2021).
  - Naively implemented kernel density estimators are computationally expensive. We demonstrate how interpolation on a well-chosen set of query points can reduce the evaluation to sublinear time and space.
- Paxton Turner, Raghu Meka, and Philippe Rigollet. Balancing Gaussian Vectors in High Dimension. *Proceedings of Thirty Third Conference on Learning Theory* (COLT 2020).
  - We study optimal procedures for balancing collections of random vectors and establish the best-known efficient algorithms.

- Paxton Turner and Yuhuai Wu. Ehrhart Theory and Discrete Equidecomposablility of Polygons. *Discrete and Computational Geometry*. 2020.
  - We develop an invariant that detects when one polygon can be transformed into another via a piecewise linear map with rational coefficients.
- Megan Leoni, Gregg Musiker, Seth Neel, and Paxton Turner. Aztec Castles and the dP3 Quiver. *Journal of Physics A: Mathematical and Theoretical*. 2014.
  - Inspired by algorithms from high-energy physics, we give exact formulas for perfect matchings of a family subgraphs of a certain lattice in terms of mutation sequences of the del-Pezzo 3 quiver.

Manuscripts

- Younhun Kim, Elchanan Mossel, Govind Ramnarayan, and Paxton Turner. Efficient Reconstruction of Stochastic Pedigrees. https://arxiv.org/abs/2005.03810. 2020.
  - We introduce a new algorithm for reconstructing the genealogy of an extant population directly from its genetic data.

Presentations

Aug. 2022 **Estimating heterogeneity in multinomial data**, *SNAB 2022*, New York University.

June 2022 **Detecting a small community in a large network**, *SIAM Discrete Math*, Carnegie Mellon University.

Oct. 2021 **A statistical perspective on coresets**, *Stat 300 Seminar*, Harvard University.

April 2021 **A statistical perspective on coreset density estimation** , *AISTATS 2021 (virtual)*.

April 2021 **Efficient interpolation of density estimators** , *AISTATS 2021 (virtual)*.

Dec. 2020 **A statistical perspective on coresets**, *YINS Seminar*, Yale University.

July 2020 **Balancing Gaussian vectors in high dimensions**, *COLT 2020 (virtual)*, Talk available at https://www.youtube.com/watch?v=q6R0FNTM8wM .

Teaching

Fall 2020 **Teaching Assistant**, *Fundamental of Statistics (online)*, MIT Math Department.

Summer 2020 **Teaching Assistant**, *Fundamentals of Statistics (online)*, MIT edX Micromasters.

Spring 2019 **Teaching Assistant**, *Introduction to Probability and Statistics*, MIT Math Department.

Summer, Fall 2018 **Teaching Staff**, *Fundamentals of Statistics (online)*, MIT edX Micromasters.

Spring 2017 **Grader**, *Introduction to Stochastic Processes*, MIT Math Department.

Fall 2017 **Grader**, *Extremal Graph Theory and Additive Combinatorics*, MIT Math Department.

Service

Reviewer, *ICALP 2020, AISTATS 2021, UAI 2021 (top reviewer), Theory of Computing Systems, Neurips 2021 (top reviewer), ICLR 2022, AISTATS 2022, FOCS 2022, Neurips 2022*.

2016–2020 **Organizer**, *MIT Integration Bee*.

2017–2018 **Co-organizer**, *MIT Graduate Student Applied Math Seminar (SPAMS)*.

Summer 2017 **Mentor**, *Summer Program for Undergraduate Research (SPUR)*, Student: Alonso Espinosa Dominguez. Project: "On Kakeya-type problems for hyperplanes in $\mathbb{R}^d$".

January 2017 **Mentor**, *MIT Directed Reading Program*, Student: Jessy Lin. Text: *The Probabilistic method* by Noga Alon and Joel Spencer.

2016–2017 **Tutor**, *English as a Second Language (ESL) Program for MIT Facilities Department Employees*, Student: Gabriel Castrillon.

Summer 2015, 2016 **Instructor**, *LSU Math Circle*, Enrichment program for mathematically gifted high school students.

Honors and Awards

2017 Levinson Fellow - MIT Math Department

2015 Betti and Robert Giles Senior Mathematics Award

2014 Barry M. Goldwater Scholarship Recipient

2014 LSU College of Science Outstanding Junior

2014 LSU Pasquale Porcelli Senior Scholarship

2013 Barry M. Goldwater Scholarship Honorable Mention

2011–2015 LSU Chancellor's Alumni Scholarship

Computer Skills

Python (numpy, scipy, matplotlib), R, LaTeX, Matlab, Mathematica, C++,