Jeffrey Carroll Regier

University of Michigan
Department of Statistics
250 West Hall
1085 S University Ave
Ann Arbor, MI 48109

+1 (734) 647-1670
regier@umich.edu
https://regier.stat.lsa.umich.edu/

Academic positions

2019 –
University of Michigan, Ann Arbor
Department of Statistics
Assistant Professor

2016 –
2019
University of California, Berkeley
Department of Electrical Engineering and Computer Sciences
Postdoctoral Researcher
advisor: Michael I. Jordan

Education

2016
Ph.D. Statistics
University of California, Berkeley
Designated Emphasis in Communication, Computation and Statistics
advisor: Jon McAuliffe

2013
M.A. Mathematics
University of California, Berkeley

2004
M.S. Computer Science
Columbia University

2003
B.A. Computer Science & Economics (double major)
Swarthmore College

Publications

Journal articles

[J1]   **Jeffrey Regier** and Philip B. Stark. "Mini-minimax uncertainty quantification for emulators". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 686–708. [PDF]

[J2]   Romain Lopez, **Jeffrey Regier**, Michael Cole, Michael I. Jordan, and Nir Yosef. "Deep generative modeling for single-cell transcriptomics". In: *Nature Methods* 15 (Dec. 2018), pp. 1053–1058. [PDF]

[J3]   **Jeffrey Regier**, Keno Fischer, Kiran Pamnany, Andreas Noack, Jarrett Revels, Maximilian Lam, Steve Howard, Ryan Giordano, David Schlegel, Jon McAuliffe, Rollin Thomas, and Prabhat. "Cataloging the visible universe through Bayesian inference in Julia at petascale". In: *Journal of Parallel and Distributed Computing* 127 (May 2019), pp. 89–104. [PDF]

[J4]   **Jeffrey Regier**, Andrew Miller, David Schlegel, Ryan P. Adams, Jon McAuliffe, and Prabhat. "Approximate inference for constructing astronomical catalogs from images". In: *Annals of Applied Statistics* 13.3 (2019), pp. 1884–1926. [PDF]

[J5]   Adam Gayoso, Zoë Steier, Romain Lopez, **Jeffrey Regier**, Kristopher Nazor, Aaron Streets, and Nir Yosef. "Joint probabilistic modeling of single-cell multi-omic data with totalVI". In: *Nature Methods* 18 (Mar. 2021), pp. 272–282. [PDF]

[J6]   Chenling Xu, Romain Lopez, Edouard Mehlman, **Jeffrey Regier**, Michael I. Jordan, and Nir Yosef. "Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models". In: *Molecular Systems Biology* 17.1 (2021), e9620. [PDF]

[J7]   Adam Gayoso, Romain Lopez, Galen Xing, Pierre Boyeau, Valeh Valiollah Pour Amiri, Justin Hong, Katherine Wu, Michael Jayasuriya, Edouard Mehlman, Maxime Langevin, Yining Liu, Jules Samaran, Gabriel Misrachi, Achille Nazaret, Oscar Clivio, Chenling Xu, Tal Ashuach, Mariano Gabitto, Mohammad Lotfollahi, Valentine Svensson, Eduardo da Veiga Beltrame, Vitalii Kleshchevnikov, Carlos Talavera-López, Lior Pachter, Fabian J. Theis, Aaron Streets, Michael I. Jordan, **Jeffrey Regier**, and Nir Yosef. "A Python library for probabilistic analysis of single-cell omics data". In: *Nature Biotechnology* 40 (2022), pp. 163–166. [PDF]

[J8]   Pierre Boyeau, **Jeffrey Regier**, Adam Gayoso, Michael Jordan, Romain Lopez, and Nir Yosef. "An empirical Bayes method for differential expression analysis of single cells with deep generative models". In: *Proceedings of the National Academy of Sciences* (2023). [PDF]

[J9]   Runjing Liu, Jon D. McAuliffe, and **Jeffrey Regier**. "Variational inference for deblending crowded starfields". In: *Journal of Machine Learning Research* (2023). [PDF]

ARTICLES IN HIGHLY SELECTIVE CONFERENCE PROCEEDINGS

[C1]   **Jeffrey Regier**, Andrew Miller, Jon McAuliffe, Ryan Adams, Matt Hoffman, Dustin Lang, David Schlegel, and Prabhat. "Celeste: Variational inference for a generative model of astronomical images". In: *International Conference on Machine Learning (ICML)*. 2015. [PDF]

[C2]   Andrew Miller, Albert Wu, **Jeffrey Regier**, Jon McAuliffe, Dustin Lang, Prabhat, David Schlegel, and Ryan Adams. "A Gaussian process model of quasar spectral energy distributions". In: *Neural Information Processing Systems (NIPS)*. 2015. [PDF]

[C3]   **Jeffrey Regier**, Michael I. Jordan, and Jon McAuliffe. "Fast black-box variational inference through stochastic trust-region optimization". In: *Neural Information Processing Systems (NIPS)*. 2017. Selected for spotlight presentation (4% acceptance rate). [PDF]

[C4] **Jeffrey Regier**, Kiran Pamnany, Keno Fischer, Andreas Noack, Maximilian Lam, Jarrett Revels, Steve Howard, Ryan Giordano, David Schlegel, Jon McAuliffe, Rollin Thomas, and Prabhat. "Cataloging the visible universe through Bayesian inference at petascale". In: *International Parallel and Distributed Processing Symposium (IPDPS)*. 2018. [PDF]

[C5] Romain Lopez, **Jeffrey Regier**, Michael I. Jordan, and Nir Yosef. "Information constraints on Auto-Encoding Variational Bayes". In: *Neural Information Processing Systems (NeurIPS)*. 2018. [PDF]

[C6] Nilesh Tripuraneni, Mitchell Stern, Chi Jin, **Jeffrey Regier**, and Michael I. Jordan. "Stochastic cubic regularization for fast nonconvex optimization". In: *Neural Information Processing Systems (NeurIPS)*. 2018. Selected for oral presentation (1% acceptance rate). [PDF]

[C7] Runjing Liu, **Jeffrey Regier**, Nilesh Tripuraneni, Michael I. Jordan, and Jon McAuliffe. "Rao-Blackwellized stochastic gradients for discrete distributions". In: *International Conference on Machine Learning (ICML)*. 2019. [PDF]

[C8] Romain Lopez, Pierre Boyeau, Nir Yosef, Michael I. Jordan, and **Jeffrey Regier**. "Decision-making with Auto-Encoding Variational Bayes". In: *Neural Information Processing Systems (NeurIPS)*. 2020. [PDF]

[C9] Derek Hansen, Brian Manzo, and **Jeffrey Regier**. "Normalizing flows for knockoff-free controlled feature selection". In: *Neural Information Processing Systems (NeurIPS)*. 2022. [PDF]

Manuscripts in submission

[S1] Jackson Loper, Robert Barton, Meena Subramaniam, Maxime Dhainaut, and **Jeffrey Regier**. "Model-Free Error Assessment for Breadth-First Studies, with Applications to Cell Perturbation Experiments". In: *arXiv: 2208.01745* (2023). [PDF]

[S2] Jackson Loper, Noam Solomon, and **Jeffrey Regier**. "Improving Accuracy in Cell Perturbation Experiments by Leveraging Auxiliary Information". In: *arXiv: 2307.11686* (2023). [PDF]

[S3] Yash Patel, Declan McNamara, Jackson Loper, **Jeffrey Regier**, and Ambuj Tewari. "Variational Inference with Coverage Guarantees". In: *arXiv: 2305.14275* (2023). [PDF]

Refereed workshop papers

[W1] **Jeffrey Regier**, Brenton Partridge, Jon McAuliffe, Ryan Adams, Matt Hoffman, Dustin Lang, David Schlegel, and Prabhat. "Celeste: Scalable variational inference for a generative model of astronomical images". In: *NIPS Workshop on Advances in Variational Inference*. 2014. Selected for spotlight presentation. [PDF]

[W2] **Jeffrey Regier**, Jon McAuliffe, and Prabhat. "A deep generative model for astronomical images of galaxies". In: *NIPS Workshop on Advances in Approximate Bayesian Inference*. 2015. [PDF]

[W3] **Jeffrey Regier** and Jon McAuliffe. "Second-order stochastic variational inference". In: *Bay Area Machine Learning Symposium*. 2016. [PDF]

[W4] Romain Lopez, **Jeffrey Regier**, Michael Cole, Michael I. Jordan, and Nir Yosef. "A deep generative model for single-cell RNA sequencing with application to detecting differentially expressed genes". In: *NIPS Workshop on Machine Learning in Computational Biology*. 2017. Selected for spotlight presentation. [PDF]

[W5] Romain Lopez, **Jeffrey Regier**, Michael I. Jordan, and Nir Yosef. "A deep generative model for single-cell RNA sequencing". In: *Bay Area Machine Learning Symposium*. 2017. Selected for oral presentation. [PDF]

[W6] Maxime Langevin, Edouard Mehlman, **Jeffrey Regier**, Romain Lopez, Michael I. Jordan, and Nir Yosef. "A deep generative model for semi-supervised classification with noisy labels". In: *Bay Area Machine Learning Symposium*. 2018. Selected for oral presentation. [PDF]

[W7] Pierre Boyeau, Romain Lopez, **Jeffrey Regier**, Adam Gayoso, Michael I. Jordan, and Nir Yosef. "Deep generative models for detecting differential expression in single cells". In: *Machine Learning in Computational Biology (MLCB) Meeting*. 2019. [PDF]

[W8] Oscar Clivio, Romain Lopez, **Jeffrey Regier**, Adam Gayoso, Michael I. Jordan, and Nir Yosef. "Detecting zero-inflated genes in single-cell transcriptomics data". In: *Machine Learning in Computational Biology (MLCB) Meeting*. 2019. Selected for spotlight presentation. [PDF]

[W9] Adam Gayoso, Romain Lopez, Zoë Steier, **Jeffrey Regier**, Aaron Streets, and Nir Yosef. "A joint model of RNA expression and surface protein abundance in single cells". In: *Machine Learning in Computational Biology (MLCB) Meeting*. 2019. [PDF]

[W10] Romain Lopez, Achille Nazaret, Maxime Langevin, Jules Samaran, **Jeffrey Regier**, Michael Jordan, and Nir Yosef. "A joint model of unpaired data from scRNA-seq and spatial transcriptomics for imputing missing gene expression measurements". In: *ICML Workshop on Computational Biology*. 2019. Best student poster award. [PDF]

[W11] Oscar Clivio, Drausin Wulsin, Evgeny Kiner, Noam Solomon, Luis Voloch, and **Jeffrey Regier**. "Cell-type annotation priors for scRNA-seq". In: *Machine Learning in Computational Biology (MLCB) Meeting*. 2020. [PDF]

[W12] Tianci Liu and **Jeffrey Regier**. "An empirical comparison of GANs and normalizing flows for density estimation". In: *NeurIPS Workshop on Bayesian Deep Learning* (2021). [PDF]

[W13] Prayag Chatha, Yixin Wang, Zhenke Wu, and **Jeffrey Regier**. "Dynamic survival transformers for causal inference with electronic health records". In: *NeurIPS Workshop on Learning from Time Series for Health* (2022). Selected for spotlight presentation. [PDF]

[W14] Derek Hansen, Ismael Mendoza, Runjing Liu, Ziteng Pang, Zhe Zhao, Camille Avestruz, and **Jeffrey Regier**. "Scalable Bayesian inference for detection and deblending in astronomical images". In: *ICML Workshop on Machine Learning for Astrophysics* (2022). [PDF]

[W15] Yash Patel and **Jeffrey Regier**. "Scalable Bayesian inference for detecting strong gravitational lensing systems". In: *NeurIPS Workshop on Machine Learning and the Physical Sciences* (2022). [PDF]

[W16] Mallory Wang, Ismael Mendoza, Cheng Wang, Camille Avestruz, and **Jeffrey Regier**. "Statistical inference for coadded astronomical images". In: *NeurIPS Workshop on Machine Learning and the Physical Sciences* (2022). [PDF]

[W17]   Zhiwei Xue, Yuhang Li, Yash Patel, and **Jeffrey Regier**. "Diffusion Models for Probabilistic Deconvolution of Galaxy Images". In: *ICML Workshop on Machine Learning for Astrophysics* (2023). [PDF]

PATENTS

[P1]   **Jeffrey Regier** and Uri Avissar. *System and method for retrieving and intelligently grouping definitions found in a repository of documents.* US Patent 7,747,555. 2010. [PDF]

Funding

Department of Energy
Title: Deep Generative Models for Transforming Pixels to Catalogs in Cosmology Analyses
Role: Principal Investigator (Share - 67%)
Total award: $224,144
Project duration: 12/1/2022–11/30/2025

National Science Foundation
Title: Elements: Scalable Bayesian Software for Interpreting Astronomical Images
Role: Principal Investigator (Share - 55%)
Total award: $500,000
Project duration: 9/1/2022–8/31/2025

Immunai Inc.
Title: Statistical methods for interpreting single-cell transcriptomics data
Role: Principal Investigator (Share - 100%)
Total award: $275,405
Project duration: 5/18/2021–5/17/2023

Shark Tank for Research Ideas in Data Science and Statistics,
University of Michigan
Title: Transformers for Robust Causal Inference using Electronic Health Records and Claims Data
Role: Co-Principal Investigator (Share - 50%)
Total award: $10,000
Project duration: 3/1/2021–3/1/2022

Awards & fellowships

2017	Best Reviewer Award, Neural Information Processing Systems (NIPS)
2017	Hyperion HPC Innovation Excellence Award, for the Celeste project
2013	Google Ph.D. Fellowship in Machine Learning, awarded annually to three students world-

wide, including one student from North America

2013	Citadel Ph.D. Fellowship (declined), awarded annually to one statistics Ph.D. student at UC Berkeley
2013	Elizabeth Scott Memorial Award, awarded at most annually to the statistics M.A. recipient at UC Berkeley showing the greatest promise in statistical research

Invited talks

Jun 2023	Statistical Challenges in Modern Astronomy VIII
Jun 2022	Penn State, Summer School in Astroinformatics
Dec 2021	Helmholtz Zentrum München, Computational Health Seminar
Mar 2021	University of Washington, Data Intensive Research in Astrophysics and Cosmology
Apr 2020	Brown University, Institute for Computational and Experimental Research in Mathematics
Sep 2018	Lawrence Berkeley National Laboratory, ML4Sci Workshop
Jul 2018	Lawrence Berkeley National Laboratory, Big Data Summit
Mar 2018	Voleon Capital Management, seminar series on machine learning
Mar 2018	University of Utah, From Stars to Bytes seminar series
Jul 2017	UCSF, Department of Radiology & Biomedical Imaging, seminar series
Jun 2017	JuliaCon 2017, keynote address
Jun 2017	Dark Energy Spectroscopic Instrument collaboration meeting, plenary talk
Dec 2016	NIPS Workshop on Advances in Approximate Bayesian Inference
Nov 2016	Lawrence Berkeley National Laboratory, Cosmology Group, seminar series
Mar 2016	Ohio State University, Statistics Department, seminar series
Aug 2015	Lawrence Berkeley National Laboratory, MANTISSA Day

Contributed talks

May 2018	International Parallel and Distributed Processing Symposium (IPDPS)
Dec 2017	Neural Information Processing Systems (NIPS)
Aug 2016	SAMSI ASTRO Opening Workshop
Jul 2015	International Conference on Machine Learning (ICML)
Dec 2014	NIPS Workshop on Advances in Variational Inference

Teaching experience

STATISTICS 507, UNIVERSITY OF MICHIGAN, ANN ARBOR
Graduate course on data analysis in Python.
Primary Instructor. Winter 2020, Fall 2020, Winter 2021

STATISTICS 415, UNIVERSITY OF MICHIGAN, ANN ARBOR
Advanced undergraduate course in modern statistical prediction and machine learning.
Primary Instructor. Fall 2019, Winter 2022, Winter 2023.

STATISTICS 215B, UNIVERSITY OF CALIFORNIA, BERKELEY
Second core graduate course in applied statistics.
Graduate Student Instructor. Spring 2013, Spring 2014.

STATISTICS 154, UNIVERSITY OF CALIFORNIA, BERKELEY
Advanced undergraduate course in modern statistical prediction and machine learning.
Graduate Student Instructor: Spring 2012.
Primary Instructor: Summer 2015.

Industry experience

May 2014
–Aug 2014
GOOGLE RESEARCH. *Intern.*
Hosted by Yoram Singer and Amar Subramanya. Developed a mathematical model for entity linking, derived an optimization procedure for it, and implemented the procedure first in Lua and then in C++ within a map-reduce framework.

May 2004
–Jun 2007
METAGLOSSARY.COM. *Co-Founder, Lead Developer.*
Invented and implemented a machine-learning-based system for extracting terms and definitions from web pages containing glossaries (US Patent 7,747,555). Raised seed funding. Developed and administered a website that served hundreds of thousands of unique visitors.

Aug 2006
–Mar 2007
GOOGLE. *Software Engineer, Contractor.*
Developed C++ code to parse web server logs on a computing farm and output models of latency distributions. Created a JavaScript-intensive web application in Python and C++ that allows users to explore these web server latency distributions.

May 2006
–Aug 2006
RUBICONSOFT. *Software Engineer, Contractor.*
Developed an extensible, multithreaded web-crawler in Java for extracting product data from e-commerce websites.

May 2005
–May 2006
GOLDMAN, SACHS & CO. *Analyst, Software Engineer.*
Developed a distributed system in Perl, C++, and a proprietary programming language for computing the firm's exposure to various types of financial risk.

Dec 2004
–May 2005
SKY SOLUTIONS. *Business Intelligence Consultant.*
Built business analytics web applications for clients using Java, SQL, and Cognos reporting software.

Journal & conference reviewing

Neural Information Processing Systems (NeurIPS), Area Chair, 2020, 2021

Annals of Applied Statistics (AoAS), 2022

Machine Learning in Computational Biology (MLCB), 2020

PLOS Computational Biology, 2020

Journal of Machine Learning Research (JMLR), 2018, 2020

The Astronomical Journal (ApJ), 2019

International Conference on Machine Learning (ICML), 2013–2016, 2018

Neural Information Processing Systems (NIPS), 2017

IEEE Transactions on Knowledge and Data Engineering (TKDE), 2014

International Joint Conferences on Artificial Intelligence (IJCAI), 2013

Artificial Intelligence and Statistics (AISTATS), 2013

Press

"Celeste: A new model for cataloging the universe". In: *EurekAlert* (Apr. 2015).

"Celeste enhancements create new opportunities in sky surveys". In: *National Energy Research Scientific Computing Center* (Nov. 2016).

"Parallel supercomputing for astronomy". In: *Julia Computing* (July 2017).

"Julia joins petaflop club". In: *HPCwire* (Sept. 2017).

"Julia language delivers petascale HPC performance". In: *The Next Platform* (Nov. 2017).

"Count based autoencoders and the future for scRNA-seq analysis". In: *What Do You Mean "Heterogeneity"?* (Apr. 2018).

Statistical advising

Immunai Inc., Scientific Advisory Board, 2020–

National Security Agency, Science Advisory Group, 2013–2022