Andrew C. Miller

Jerome L. Greene Science Center
3227 Broadway
New York, NY 10027

am5171@columbia.edu
http://andymiller.github.io/

RESEARCH INTERESTS

probabilistic modeling, scalable statistical inference methods, semi-parametric models, spatiotemporal statistics; application to astronomy, sports analytics, and health care

EDUCATION

Harvard University, Cambridge, MA *2012–2018*
Ph.D. Computer Science
- Advisor: Ryan Adams (Princeton CS, Google Brain)

Brown University, Providence, RI *2005–2010*
B.A., Sc.M. Computer Science; B.A. Music
- Advisor: Erik Sudderth
- Graduated with Honors, **3.9** (of 4) GPA

PROFESSIONAL EXPERIENCE

Columbia University, New York, NY *Sept. 2018–present*
Postdoctoral Research Scientist, Data Science Institute
- Working with John Cunningham (Statistics) and Dave Blei (Statistics and Computer Science)

Google, Mountain View, CA *June–Sept. 2015*
Software Engineering Intern: Ads Quality
- Causal analysis of the effects of ad quality on long-term user behavior

Google, Mountain View, CA/London, UK *June–Sept. 2014, Jan. 2015*
Software Engineering Intern: Search
- Designed a time series prediction system for long-term forecasts
- Developed models for tracking influenza-like illness (ILI) rates by jointly modeling CDC measurements and Google search query volumes (published methods)

Computer Vision Group, Inc., Providence, RI *2010–2012*
Research Developer
- Designed and implemented real-time 3-D reconstruction system

PUBLICATIONS

Andrew C. Miller, Nicholas Foti, Alexander D'Amour, and Ryan P. Adams. Reducing Reparameterization Gradient Variance. *Advances in Neural Information Processing Systems*, 2017a.

Andrew C. Miller, Nicholas Foti, and Ryan P. Adams. Taylor Residual Estimators via Automatic Differentiation. *Advances in Approximate Bayesian Inference (NIPS Workshop)*, 2017b.

Andrew C. Miller, Sendhil Mullainathan, and Ziad Obermeyer. A Hierarchical Generative Model of Electrocardiogram Records. *Machine Learning for Health (NIPS Workshop)*, 2017c.

Andrew C. Miller, Nicholas Foti, and Ryan P. Adams. Variational Boosting: Iteratively Refining Posterior Approximations. *Proceedings of the 34th Annual International Conference on Machine Learning*, 2017d.

Scott Linderman, Matthew Johnson, **Andrew C. Miller**, Ryan Adams, David Blei, and Liam Paninski. Bayesian Learning and Inference in Recurrent Switching Linear Dynamical Systems. *Artificial Intelligence and Statistics*, page 914–922, 2017.

Andrew C. Miller and Luke Bornn. Possession sketches: Mapping NBA strategies. *11th Annual MIT Sloan Sports Analytics Conference*, 2017.

Luke Bornn, Daniel Cervone, Alexander Franks, and **Andrew C. Miller**. Studying Basketball through the Lens of Player Tracking Data. In *Handbook of Statistical Methods and Analyses in Sports*. Chapman and Hall/CRC, 2016.

Andrew C. Miller, Albert Wu, Jeff Regier, Jon McAuliffe, Dustin Lang, Mr Prabhat, David Schlegel, and Ryan P Adams. A Gaussian Process Model of Quasar Spectral Energy Distributions. *Advances in Neural Information Processing Systems*, page 2485–2493, 2015.

Jeffrey Regier, **Andrew C. Miller**, Jon McAuliffe, Ryan Adams, Matt Hoffman, Dustin Lang, David Schlegel, and Mr Prabhat. Celeste: Variational Inference for a generative model of astronomical images. *International Conference on Machine Learning*, page 2095–2103, 2015.

Vasileios Lampos, **Andrew C. Miller**, Steve Crossan, and Christian Stefansen. Advances in nowcasting influenza-like illness rates using search query logs. *Scientific Reports*, 5, 2015.

Alexander Franks, **Andrew C. Miller**, Luke Bornn, and Kirk Goldsberry. Characterizing the Spatial Structure of Defensive Skill in Professional Basketball. *The Annals of Applied Statistics*, 9(1):94–121, 2015a.

Alexander Franks, **Andrew C. Miller**, Luke Bornn, and Kirk Goldsberry. Counterpoints: Advanced defensive metrics for NBA basketball. *9th Annual MIT Sloan Sports Analytics Conference*, 2015b.

Andrew C. Miller, Luke Bornn, Ryan Adams, and Kirk Goldsberry. Factorized Point Process Intensities: A Spatial Analysis of Professional Basketball. *Proceedings of the 31st Annual International Conference on Machine Learning*, page 235–243, 2014.

Vishal Jain, **Andrew C. Miller**, and Joseph L Mundy. A Multi-sensor Fusion Framework in 3-D. *Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on*, page 314–319, 2013.

Andrew C. Miller, Vishal Jain, and Joseph L Mundy. Real-time rendering and dynamic updating of 3-d volumetric data. *Proceedings of the Fourth Workshop on General Purpose Processing on Graphics Processing Units*, page 8–16, 2011.

SELECTED HONORS

Siebel Scholar, Class of 2018 — *Fall 2017*

Third Place, MIT Sloan Sports Analytics Conference Research Paper Competition — *Spring 2017*

Certificate of Distinction in Teaching, Bok Center for Teaching and Learning — *Fall 2015*

Best Research Paper, MIT Sloan Sports Analytics Conference Research Paper Competition — *Spring 2015*

Goldman Sachs Fellowship — *Fall 2013*

Harvard SEAS Graduate Student Fellowship — *Fall 2012*

Weston Prize for Excellence in Music — *Spring 2009*

TEACHING EXPERIENCE

School of Engineering and Applied Sciences, Harvard University — *Spring 2015*
Head Teaching Fellow, CS281: Advanced Machine Learning

- Taught technical material in weekly sections and review sessions
- Held office hours and wrote exam questions
- Awarded Certificate of Distinction in Teaching by the Bok Center

Department of Computer Science, Brown University — *Sept. 2007–May 2009*
Teaching Assistant, Various Courses

- Designed and implemented artificial intelligence programming projects for upper level computer science students
- Designed and implemented programming projects for second semester computer science students, teaching algorithms and data structures
- Instructed and graded assignments for introductory computer science students, teaching mostly object oriented programming and software design

INVITED TALKS

Taylor Residual Estimators via Automatic Differentiation. Advances in Approximate Bayesian Inference, NIPS Workshop (Dec. 2017)

Advances in Monte Carlo Variational Inference. CS Colloquium, Viterbi School of Engineering, USC. (Nov. 2017)

Possession Sketches: Mapping NBA Strategies. Sloan Sports Analytics Conference. (Mar. 2017)

Stealing the Playbook: Structure discovery in NBA player-tracking data. The Cascadia Symposium on Statistics in Sports. (Sept. 2016)

Communication Panel, NFL Football Performance and Technology Symposium, Indianapolis, IN. (Feb. 2016)

Counterpoints: Advanced Defensive Metrics for NBA Basketball. MIT Sloan Sports Analytics Conference. (Feb. 2015)

Characterizing the Spatial Structure of Defensive Skill in Professional Basketball. KDD Workshop on Large-Scale Sports Analytics. (Aug. 2014)

A Spatiotemporal Analysis of Professional Basketball. Joint Statistical Meetings, 2014. Session: Bayes and Big Data. (Aug. 2014)

Characterizing The Spatial Structure of Defensive Skill in Professional Basketball (poster). International Society for Bayesian Analysis World Meeting. (July 2014)

Factorized Point Process Intensities: A Spatial Analysis of Professional Basketball. International Conference on Machine Learning, Beijing, China. (June 2014)

Quantifying Offensive Player Types in the NBA with Non-Negative Matrix Factorization. New England Symposium on Statistics in Sports, Harvard University, Cambridge, MA. (Sept. 2013)

Real-time Rendering and Dynamic Updating of Dense 3-d Volumetric Data. HPC and GPGPU Meetup, Boston, MA. (July 2011)

OpenCL Implementation of a Heterogeneous Computing System for Real-Time Rendering and Dynamic Updating of Dense 3-d Volumetric Data. AMD Fusion Developer Summit, Bellevue, WA (2011)

Real-time Rendering and Dynamic Updating of Dense 3-d Volumetric Data. Fourth Workshop on GPGPUs, ASPLOS 2011. Newport, CA. (Mar. 2011)

RESEARCH PROJECTS

Modeling Electrocardiogram Records *2016–present*

- Building accurate and interpretable probabilistic generative models of patient-specific electrocardiogram data
- Using the model to automatically annotate ECGs with common tags; predicting test outcomes (in a semi-supervised setting)

Flexible and Robust Monte Carlo Variational Inference *2016–present*

- Developing new variational inference methods for fast approximate Bayesian Inference
- Developing variance reduction technique useful for a large class of variational inference methods

Unsupervised Learning of Structured Agent Interaction *2016–present*

- Developing methods for discovering structure in collections of tracking-data
- Applying to large database of basketball possessions, discovering natural clusters useful for player analysis and sports journalism

Joint Photometric and Spectroscopic Modeling for Redshift Identification *2014–2015*

- Developed a model of two sources of information about point sources (stars and quasars)
- Used this method to make accurate photometric redshift measurements at extreme ranges

Robust methods for 'nowcasting' disease rates with search queries *June–Sept. 2014*

- Developed and experimented with existing high-dimensional regression techniques to accurately and robustly predict population disease rates from search query volumes (ongoing research at Google)
- *Topics*: high-dimensional regression, subspace identification, non-linear time series, Gaussian processes

Structured spatiotemporal point process priors for prediction *Jan. 2014–present*

- Developing dynamic models for point process data, with a focus on prediction
- Applications to crime prediction, transportation, sports analytics
- *Topics*: point processes, Gaussian processes, MCMC

New measures of basketball ability. *Nov. 2013–June 2014*

- Developed statistical methods to evaluate player effectiveness (particularly defensive ability) from dynamic player-tracking data
- Applied Bayesian spatiotemporal models to rich, large-scale data (10 player x, y coordinates, 25 data points per second, for an NBA season, 50-100GB)
- *Topics*: hierarchical modeling, hidden Markov models, point processes, nonnegative matrix factorization

Optimization and application of automatic 3-d modeling system. *July 2010–May 2012*

- Fusing RGB and IR imagery for classifying types of material and detecting foreground changes
- *Topics*: parallel computing, volumetric modeling, 3-d reconstruction

REVIEWING AND SERVICE

Reviewer, American Medical Informatics Association Informatics Summit	2019
Reviewer, International Conference on Learning Representations	2018
Reviewer, International Conference on Machine Learning (ICML)	2015–2018
Reviewer, Neural Information Processing Systems (NIPS)	2014–2018
Reviewer, Deep Learning for Physical Sciences (NIPS Workshop)	2017
Reviewer, Advances in Approximate Bayesian Inference (NIPS Workshop)	2016–2017
Reviewer, Artificial Intelligence and Statistics (AISTATS)	2017–2018
Reviewer, Journal of Quantitative Analysis of Sports (JQAS)	2014
Reviewer, Statistics and Computing (STCO)	2016
Reviewer, IEEE Transactions on Knowledge and Data Engineering	2014–2016

TECHNICAL SKILLS

Programming Languages and Frameworks: Python, Numpy/Scipy, PyTorch, Tensorflow, C/C++, OpenCL, MATLAB, Javascript, LaTeX

OTHER

Captained Brown Men's Ultimate Team (2008–2009)
Jazz piano/guitar, music composition
Eagle Scout (2003)