PARIKSHIT SHAH

Research Scientist • MIT PhD • 6 years experience

PERSONAL INFORMATION

email	pshah@discovery.wisc.edu
website	http://people.lids.mit.edu/pari/
phone	+1 (617) 388 2985

Area of Interest

Applying optimization, statistics, and data science to solve challenging problems in industry.

WORK EXPERIENCE

04/2018-present — Facebook

Research Scientist, Feed Science

Work on developing algorithmic fairness principles for News Feed.

- **Algorithmic fairness, tradeoffs, and ranking effects in News Feed Integrity**: This ongoing project focuses on developing a framework to reason about ranking impacts of feed integrity classifiers and their algorithmic fairness. The work also quantifies the disparity of impact of classifier based demotions and their decomposition into intrinsic and extrinsic components, which inform our fairness principles and enforcement mechanisms.

02/2015-03/2018 — Yahoo Research

Research Scientist

Research in several areas including optimization, machine learning, ranking and statistics and their applications to web-scale problems. Used insights from research to deliver several large scale ($>$ 100 million users) production systems involving computational advertising, personalization and ranking at Yahoo.

- **Online ranking with constraints**: This work addresses the problem of optimally ranking content on page in an online manner to optimally navigate the tradeoff between engagement and satisfaction of other business constraints such as revenue and clicks delivered to publishing properties. We developed a scalable, linear programming-based algorithm that navigates these tradeoffs optimally. Solution yielded 3% improvement in engagement metrics while satisfying contraints. Work presented as oral paper at KDD 2017.
- **Contextual bandits for comment-ranking**: In this project, we delivered a framework for optimally ranking comments on Yahoo homepage content using a neural network within a bandit learning framework.
- **A geometric framework for false discovery in low rank matrix estimation**: We develop a geometric framework for reasoning about false discovery in low-rank settings. We provide a method for controlling for false discovery in this setting via a novel bagging and tangent-space averaging procedure that we call subspace stability selection. We provide analytical guarantees and show that our approach provides strong empirical improvements on a number of datasets ranging from recommendation system problems to hyperspectral imaging.
- **Exploration-exploitation for cold ads**: Developed and deployed a exploration-exploitation system to full production traffic to address click-prediction for cold ads in the search advertising system. The system was based on a Upper Confidence Bound (UCB) based contextual bandit formulation, with several practical innovations that facilitated exploration of cold ads in a revenue-positive manner and achieves a discovery rate of 10%. Work presented as oral paper at KDD 2017.
- **Large-scale click-modeling**: Developed an end-to-end click-modeling pipeline for a new search advertising product at Yahoo. The model, which was deployed to full-production traffic, was trained using a parameter-server

pipeline with 200 million features and yielded improvements of 12% improvement in CTR and RPM metrics.

10/2012-02/2015 **Philips Research**

Member of Research Staff (Research Scientist)

Statistical modeling and optimization for problems in healthcare and energy services. Efforts include:

- Patient clustering: Led project on development and implementation of metric learning and clustering algorithms on large-scale patient data.
- Patient volume prediction: Developed models for demand, competition and patient choice for equipment utilization, and using these to develop new pricing models.
- Energy analytics and services: Data analytics for buildings, algorithms for data-driven building automation, occupancy modeling and visualization, demand response algorithms, intelligent maintenance.
- Other activities: Generation of intellectual property (**30 invention disclosures and patent applications**), writing grant proposals (NIH, CMS).

07/2011-08/2012 **University of Wisconsin-Madison**

Research Associate

Fundamental research in areas of covariance estimation, time series modeling, and statistical signal processing. New methodologies were created for these problems by using core ideas from optimization. Developed novel method for learning succinct dynamic models for time-series data, key idea was a new notion of regularization via *atomic norms*. Worked on covariance estimation in high-dimensional settings, created a new, reliable and lossless means of compressing statistical information via the notion of sketching.
Host: Benjamin Recht. brecht@berkeley.edu

EDUCATION

2005-2011 **Massachusetts Institute of Technology**

PhD, Electrical Engineering and Computer Science

Thesis: *A Partial Order Approach to Decentralized Decision-Making*
Description: PhD thesis asks and answers a fundamental question in optimization: when are large scale information-constrained decision-making problems tractable? Work delineates and characterizes tractable decentralized decision-making problems from intractable ones, provides first algorithmic approaches, basic architectural insights and establishes new connections with combinatorics, optimization and algebra.
Advisor: Pablo Parrilo.

2003-2005 **Stanford University**

S.M., Electrical Engineering

Description: Extensive coursework and research on optimization, signal processing, and decision-making.

1999-2003 **Indian Institute of Technology, Bombay**

B. Tech., Aerospace Engineering

Description: Ranked 2^{nd} in graduating class within major, undergraduate thesis work published in leading conference and journal.

OTHER EXPERIENCE

08/2012-present **University of Wisconsin-Madison**

Visiting Researcher

- Tensor methods for signal processing: while tensors play an important role in many signal processing and machine learning applications, tensor decompositions are computationally intractable in general. We draw connections to polynomial optimization and develop semidefinite programming approaches that are provably exact.
- Relative entropy relaxations: While optimizing sums of exponentials is important in many applications (e.g. learning mixtures of logistic models),

the resulting optimization problems (called signomial programs) are intractable. In this work, we develop the first principled approaches for such problems via the notion of *relative entropy relaxations* and optimization hierarchies.

- Algorithms for atomic norm regularization: Many important regularized regression problems can be stated via *atomic norms*, a notion of universal regularization. We establish a general optimization and algorithmic framework for such problems that is inspired by the Frank-Wolfe method: selected applications include tensor learning, deconvolving graphs, and learning sparse models.

Fall 2010 — *Visiting Researcher, UCLA*

Institute for Pure and Applied Mathematics

Semester-long research program for optimization experts, involving research exchanges, collaborations and new directions. Developed new approach for signal processing on manifolds. Attended and delivered research talks and together with other experts in the optimization community helped identify major open questions and challenges faced by the community. In the process, established connections with other experts and have become a known researcher.

05/2010-08/2010 — *Research Intern*

IBM T. J. Watson Research Center

Researched and applied state-of-the-art convex optimization and machine learning methods to speech recognition problems. Worked on two major aspects of speech recognition, language modeling and acoustic modeling. On the language modeling side, researched and implemented fast first order methods for regularized max-ent models. On the acoustic modeling side, developed new convex approaches for feature classification. Work led to multiple papers in top conferences.

2005-2011 — *Research Assistant, Laboratory for Information and Decision Systems*

Massachusetts Institute of Technology

Conducted basic research in the area of optimization, semidefinite programming and control. Developed first computational approaches for a broad class of decentralized decision-making problems as well as for dynamic games. Work establishes newt connections between tractability of decentralized decision-making, information-flow, and convex optimization. Novel insights draw on understanding the hidden algebraic and combinatorial structure.
Advisor: Pablo A. Parrilo. parrilo@mit.edu

Summer 2004 — *Summer Intern*

NASA Ames Research Center

Developed software for object tracking and visualization.

Awards

- INFORMS Optimization Society Young Researcher Prize, 2016.
- CDC general chairs recognition award, Conference on Decision and Control, 2009.
- School of Engineering Fellowship, Stanford University,2003.

Programming and Data Science Skills

Python, Scala, Spark, Bash Scripting, C++, MATLAB, SQL, Hadoop, Mapreduce.

Publications

PUBLICATIONS

1. A. Taeb, P. Shah and V. Chandrasekaran, "Latent Variable Graphical Model Selection for Ising Models and Generalized Linear Models," In preparation.
2. A. Taeb, P. Shah and V. Chandrasekaran, "A Geometric False Discovery Framework for Low-Rank Estimation," submitted, *Journal of the American Statistical Association*.

3. P. Shah, A. Soni, T. Chevalier, "Online Ranking with Constraints: A Primal-Dual Algorithm and Applications to Web Traffic-Shaping", *Knowledge Discovery in Databases*, 2017.
4. V. Chandrasekaran, P. Shah, "Relative Entropy Relaxations and Its Applications," *Mathematical Programming*, 2017.
5. V. Chandrasekaran, P. Shah, "Relative Entropy Relaxations for Signomial Optimization", *SIAM Journal on Optimization*, 2016.
6. P. Shah, M. Yang, S. Alle, A. Ratnaparkhi, B. Shahshahani, R. Chandra "A Practical Exploration System for Search Advertising", *Knowledge Discovery in Databases*, 2017.
7. G. Dasarathy, P. Shah, R. Baraniuk, "Sketched Covariance Testing: A Compression-Statistics Tradeoff", *IEEE Symposium on Information Theory*, 2017.
8. P. Shah, N. Rao, G. Tang, "Optimal Low-Rank Tensor Recovery from Separable Measurements: Four Contractions Suffice", Preprint.
9. P. Shah and V. Saligrama, "Causal Inference for Time Series: A Semidefinite Programming Approach", Preprint.
10. P. Shah, N. Rao, G. Tang, "Sparse and Low-Rank Tensor Decompositions", *Advances in Neural Information Processing Systems*, 2015
11. G. Tang and P. Shah, "Gauranteed Tensor Decomposition: A Moment Approach", *International Conference on Machine Learning*, 2015
12. N. Rao, P. Shah, S. Wright, "Forward-Backward Greedy Algorithms for Atomic-Norm Regularization," *IEEE Transactions on Signal Processing*, 2015.
13. G. Dasarathy, P. Shah, B. Bhaskar and R. Nowak, "Sketching Sparse Matrices," *IEEE Transactions on Information Theory*, 2015.
14. P. Shah, B. Bhaskar, G. Tang, and B. Recht, "Linear System Identification via atomic norm regularization," Preprint.
15. G. Tang, B. Bhaskar, P. Shah, and B. Recht, "Compressed Sensing Off the Grid," IEEE Transactions of Information Theory, Vol. 59, No. 11, 2013.
16. P. Shah and V. Chandrasekaran, "Group Symmetry and Covariance Regularization," *Electronic Journal of Statistics, Vol. 6, pp. 1600-1640, 2012*.
17. V. Chandrasekaran, P. Shah, "Relative Entropy Relaxations for Signomial Optimization," Allerton Conference on Communication, Control, and Computing, 2014.
18. N. Rao, P. Shah, S. Wright, "Forward-Backward Greedy Algorithms for Convex Demixing," Invited Paper, Asilomar Conference on Signals, Systems and Computers, 2014.
19. G. Tang, B. Bhaskar, P. Shah, B. Recht, "Robust Line Spectral Estimation," Invited Paper, Asilomar Conference on Signals, Systems and Computers, 2014.
20. N. Rao, P. Shah, S. Wright, "Conditional Gradient with Enhancement and Truncation for Atomic-Norm Regularization," Neural Information Processing Systems, Workshop on Greedy Algorithms, 2013.
21. G. Dasarathy, P. Shah, B. Bhaskar, R. Nowak, "Sketching Sparse Covariance Matrices and Graphs", Neural Information Processing Systems, workshop on Randomized Methods in Machine Learning, 2013.
22. G. Tang, B. Bhaskar, P. Shah, R. Nowak, and B. Recht, "Robust signal deconvolution," to be submitted.
23. N. Rao, P. Shah, S. Wright, and R. Nowak, "Forward-Backward Greedy Algorithms for Atomic Norm Minimization", International Conference on Acoustic, Speech, and Signal Processing, 2013.
24. P. Shah, B. Bhaskar, G. Tang, and B. Recht, "Linear System Identification via Atomic Norm Minimization," IEEE Conference on Decision and Control, Dec. 2012.
25. G. Tang, B. Bhaskar, P. Shah, and B. Recht, "Compressed Sensing Off the Grid,"Allerton Conference on Communication Control and Computing, Sep. 2012.

26. G. Dasarathy, P. Shah, B. Bhaskar and R. Nowak, "Covariance Sketching," Allerton Conference on Communication Control and Computing, Sep. 2012.
27. P. Shah and P. Parrilo, "An Optimal Controller Architecture for Poset-Causal systems," Submitted, .
28. P. Shah and P. Parrilo, "H_2-Optimal Decentralized Control over Posets: A State-Space Solution for State-Feedback," IEEE Transactions on Automatic Control, Vol. 58, No. 12, 2013.
29. P. Shah and V. Chandrasekaran, "Iterative Thresholding for Signal Identification on Manifolds", Allerton Conference on Communication Control and Computing, 2011.
30. P. Shah and P. Parrilo, "An Optimal Controller Architecture for Poset-Causal Systems", IEEE Conference on Decision and Control 2011.
31. M. Kristalny and P. Shah, "On the fully decentralized two-block H_2 model-matching problem with one-sided dynamics", Submitted American Control Conference 2011.
32. T. Sainath, D. Nahamoo, D. Kanevsky, B. Ramabhadran, P. Shah,"A Convex Hull Approach to Sparse Representations for Exemplar-Based Speech Recognition", IEEE Conference on Automatic Speech Recognition and Understanding 2011.
33. T. Sainath, D. Nahamoo, B. Ramabhadran, D. Kanevsky, V. Goel, P. Shah, "Exemplar-Based Sparse Representation Phone Identification Features", IEEE Conference on Acoustic, Speech and Signal Processing 2011.
34. P. Shah and P. Parrilo, "H_2-Optimal Decentralized Control over Posets: A State-Space Solution for State-Feedback", IEEE Conference on Decision and Control 2010.
35. P. Shah and P. Parrilo, 'A Poset Framework to Model Decentralized Control Problems," IEEE Conference on Decision and Control 2009.
36. P. Shah and P. Parrilo, "A Partial Order Approach to Spatially Invariant Systems," Allerton Conference on Communication Control and Computing 2008.
37. P. Shah and P. Parrilo, "Polynomial Stochastic Games using Sums-of-Squares Programming," IEEE Conference on Decision and Control 2007.
38. P. Shah and P. Parrilo, "A Partial Order Approach to Decentralized Control," IEEE Conference on Decision and Control 2008.
39. N. Ananthkrishnan, P. Shah and S. Unnikrishnan, "Approximate Analytical Criterion for Wing Rock Onset," AIAA Journal on Guidance, Control and Dynamics, Volume 27, No. 2, 2004.
40. P. Shah, N. Ananthkrishnan, and S. Unnikrishnan, "Analytical Criteria for Wing Rock Onset," AIAA Atmospheric Flight Mechanics Conference 2003.

SELECTED INVITED TALKS

(some slides available at: http://people.lids.mit.edu/pari/talks.html)

1. SIAM Conference on Optimization, May 2017
2. INFORMS Annual Meeting, Nashville, Oct. 2016
3. International Symposium on Mathematical Programming, Philadelphia, July 2015
4. SIAM Conference on Optimization, May 2014
5. IBM T. J. Watson Research Center, New York, September 2013
6. University of Southern California, Dept. Electrical and Computer Engineering, April 2013
7. Harvard University, Division of Engineering and Applied Science, March 2013
8. International Symposium on Mathematical Programming, Berlin, August 2012
9. Institute for Pure and Applied Mathematics (UCLA), Lake Arrowhead, June 2012

10. Conference on Information Sciences and Systems, Princeton University, March 2012
11. Automatic Control Laboratory, Lund University, May 2011
12. Automatic Control Department, KTH Stockholm, May 2011
13. Institute for Systems Theory and Automatic Control, University of Stuttgart, May 2011 (**visiting researcher position, Summer 2011.**)
14. Electrical and Systems Engineering, University of Pennsylvania, March 2011
15. Electrical and Computer Engineering, University of Texas at Austin, March 2011
16. Computer Science Department, University of Wisconsin, December 2010
17. Department of Electrical Engineering, Stanford University, September 2009
18. Mathematical Theory of Networks and Systems, VirginiaTech, July 2008
19. International Symposium of Mathematical Programming, June 2006
20. Northeast Control Workshop, University of Pennsylvania, 2006

OTHER RESEARCH PROGRAMS AND WORKSHOPS

- Reunion Conference on Optimization, Institute for Pure and Applied Mathematics, Lake Arrowhead, 2012.
- Machine Learning: Theory and Computation, Institute for Mathematics and its Applications, 2012.
- Workshop on Semidefinite Optimization: Theory, Algorithms and Applications, Oberwolfach, 2010.
- Convex algebraic geometry, optimization and applications, American Institute of Mathematics, 2009.
- Algebraic Algorithms in Optimization, Institute for Mathematics and its Applications, 2006.
- Optimization and Control, Institute for Mathematics and its Applications, 2006.

SKILLS AND PROFESSIONAL ACTIVITIES

Selected Reviewing Activities

- Journal of Machine Learning Research
- International Conference on Machine Learning (ICML)
- Neural Information Processing Systems (NIPS)
- AAAI Conference on Artificial Intelligence (AAAI)
- International Conference on Artificial Intelligence and Statistics (AISTATS)
- Knowledge Discovery in Databases
- SIAM Journal on Optimization
- Mathematical Programming
- SIAM Journal on Matrix Analysis and Applications
- IEEE Transactions on Information Theory
- IEEE Transactions on Signal Processing
- International Conference on Acoustic, Speech, and Signal Processing
- International Symposium on Information Theory
- IEEE Transactions on Automatic Control
- IEEE Conference on Decision and Control