

# Alexander (Sasha) Rakhlin

Department of Statistics, The Wharton School  
University of Pennsylvania  
465 Jon M. Huntsman Hall, Philadelphia, PA 19104  
Email: rakhlin at wharton.upenn.edu

## ACADEMIC APPOINTMENTS

- 01/2009 - **University of Pennsylvania**  
Assistant Professor, Department of Statistics, The Wharton School  
Secondary appointment: Department of Computer & Information Science  
*Co-director*, Penn Research in Machine Learning (PRiML)
- 07/2006 - 12/2008 **University of California, Berkeley**  
Postdoctoral Scholar, Dept. of Electrical Engineering and Computer Sciences  
Supervisor: Peter L. Bartlett.

## EDUCATION

- 09/2000 - 06/2006 **Massachusetts Institute of Technology**  
Ph.D., Center for Biological and Computational Learning  
Supervisor: Tomaso Poggio. Thesis: *Applications of Empirical Processes in Learning Theory: Algorithmic Stability and Generalization Bounds.*
- 01/2000 - 06/2000 **Oxford University** – Cornell Study Abroad
- 09/1996 - 06/2000 **Cornell University**  
B.A. in Computer Science, B.A. in Mathematics. GPA 4.0/4.0

## TEACHING EXPERIENCE

- Fall 2009, 2010 **Department of Statistics, University of Pennsylvania**  
*STAT 101: Intro to Statistics*
- Spring 2009 **Department of Statistics, University of Pennsylvania**  
*STAT 991: Regularization Methods*
- Spring 2008 **Dept. of Electrical Engineering & Computer Sciences, UC Berkeley**  
*Statistical Learning Theory*, Co-Lecturer
- 2003 – 2008 **Center for Biological and Computational Learning, MIT**  
*Statistical Learning Theory and Applications*, Co-Lecturer
- Spring 2004 **Department of Brain and Cognitive Sciences, MIT**  
*Statistical Methods*, Teaching Assistant
- Fall 2002 **Department of Brain and Cognitive Sciences, MIT**  
*Laboratory in Cognitive Science*, Teaching Assistant

## AWARDS AND GRANTS

- Best Paper Award, Conference on Learning Theory, 2011
- NSF CCF-1116928, Algorithmic Foundations, Division of Computer and Comm. Foundations, 2011
- NSF CAREER Award, Division of Mathematical Sciences, 2010

- Dean's Research Fund, Wharton School, 2010
- IBM Research's 2008 Pat Goldberg Memorial Best Paper Award in CE, EE and Math

**PUBLICATIONS** ( \* denotes alphabetical ordering )

- \* D. Foster and A. Rakhlin. *No Internal Regret via Neighborhood Watch*.
- \* M. Raginsky and A. Rakhlin. *Lower Bounds for Passive and Active Learning*. NIPS 2011, to appear.
- \* A. Agarwal, D. Foster, D. Hsu, S. Kakade, and A. Rakhlin *Stochastic Convex Optimization with Bandit Feedback*. NIPS 2011, to appear.
- \* A. Rakhlin, K. Sridharan, and A. Tewari. *Online Learning: Stochastic and Constrained Adversaries*. NIPS 2011, to appear.
- \* D. Foster, A. Rakhlin, K. Sridharan, and A. Tewari. *Complexity-Based Approach to Calibration with Checking Rules*. COLT, 2011.
- \* A. Rakhlin, K. Sridharan, and A. Tewari. *Online Learning: Beyond Regret*. COLT, 2011. (**COLT'11 Best Paper Award**)
- \* J. Abernethy, E. Hazan, and A. Rakhlin. *Interior Point Methods for Full-Information and Bandit Online Learning*, submitted.
- \* M. Raginsky and A. Rakhlin. *Information-Based Complexity, Feedback and Dynamics in Convex Programming*, IEEE Transactions on Information Theory, 2011.
- \* A. Rakhlin, K. Sridharan, and A. Tewari. *Online Learning: Random Averages, Combinatorial Parameters, and Learnability*. Advances in Neural Information Processing Systems, 2010. (**Accepted for full oral presentation**)
- \* H. Narayanan and A. Rakhlin. *Random Walk Approach to Regret Minimization*. Advances in Neural Information Processing Systems, 2010.
- \* M. Raginsky, A. Rakhlin, and S. Yüksel. *Online Convex Programming in Adaptive Control*. The 49th IEEE Conference on Decision and Control, 2010.
- S. Seshia and A. Rakhlin. *Quantitative Analysis of Systems Using Game-Theoretic Learning*. ACM Transactions on Embedded Computing Systems, 2010.
- \* M. Raginsky and A. Rakhlin. *Information Complexity of Black-Box Convex Optimization: A New Look Via Feedback Information Theory*. Allerton Conference on Communication, Control, and Computing, 2009.
- \* J. Abernethy and A. Rakhlin. *Beating the Adaptive Bandit with High Probability*. COLT 2009.
- \* J. Abernethy, A. Agarwal, P. Bartlett, and A. Rakhlin. *A Stochastic View of Optimal Regret through Minimax Duality*. COLT 2009.
- S. Seshia and A. Rakhlin. *Game-Theoretic Timing Analysis*. ICCAD 2008.
- \* J. Abernethy, E. Hazan, and A. Rakhlin. *Competing in the Dark: An Efficient Algorithm for Bandit Linear Optimization*. COLT 2008. (**Machine Learning Journal Award and IBM Research's 2008 Pat Goldberg Memorial Best Paper Award in CE, EE and Math**)
- \* J. Abernethy, P. Bartlett, A. Rakhlin, and A. Tewari. *Optimal Strategies and Minimax Lower Bounds for Online Convex Games*. COLT 2008.
- \* P. Bartlett, V. Dani, T. Hayes, S. Kakade, A. Rakhlin, and A. Tewari. *High probability regret bounds for online optimization*. COLT 2008.
- \* P. L. Bartlett, E. Hazan, and A. Rakhlin. *Adaptive Online Gradient Descent*. Advances in Neural Information Processing Systems, 2007. (**Accepted for full oral presentation**)

- A. Rakhlin, J. Abernethy, and P. L. Bartlett. *Online Discovery of Similarity Mappings*. Proceedings of International Conference on Machine Learning, pages 767–774, 2007. ACM Press.
- ★ J. Abernethy, P. L. Bartlett, and A. Rakhlin. *Multitask Learning with Expert Advice*. The Twentieth Annual Conference on Learning Theory, pages 484–498, 2007.
- A. Rakhlin and A. Caponnetto. *Stability of K-means Clustering*. Advances in Neural Information Processing Systems, pages 1121–1128. MIT Press, Cambridge, MA, 2007.
- A. Rakhlin. *Applications of Empirical Processes in Learning Theory: Algorithmic Stability and Generalization Bounds*, Ph.D. Thesis, MIT, 2006.
- ★ A. Caponnetto and A. Rakhlin. *Stability Properties of Empirical Risk Minimization over Donsker Classes*. Journal of Machine Learning Research, vol. 6, 2565–2583, 2006.
- A. Rakhlin, D. Panchenko and S. Mukherjee, *Risk Bounds for Mixture Density Estimation*, ESAIM: Probability and Statistics, vol. 9, pages 220–229, 2005.
- A. Rakhlin, S. Mukherjee and T. Poggio, *Stability Results in Learning Theory*, Analysis and Applications, vol. 3, no. 4, pages 397–419, 2005.
- T. Poggio, S. Mukherjee, R. Rifkin, A. Rakhlin and A. Verri, *B. Uncertainty in Geometric Computations*, J. Winkler and M. Niranjana (eds.), Kluwer Academic Publishers, pages 131–141, 2002.
- A. Rakhlin, G. Yeo and T. Poggio, *Extra-label Information: Experiments with View-based Classification*. Proceedings of the Sixth International Conference on Knowledge-Based Intelligent Information & Engineering Systems, Crema, Italy, 2002.

## TECHNICAL REPORTS (other than above)

- A. Agarwal, A. Rakhlin, and P. L. Bartlett. *Matrix Regularization Techniques for Online Multitask Learning*. Technical Report EECS-2008-138, UC Berkeley, Oct 2008.
- A. Rakhlin, A. Tewari, and P. L. Bartlett. *Closing the Gap between Bandit and Full-Information Online Optimization: High-Probability Regret Bound*. Technical Report EECS-2007-109, UC Berkeley, Aug 2007.
- A. Rakhlin and T. Poggio, *On Stability and Concentration of Measure*. CBCL Paper 2004-239a, Massachusetts Institute of Technology, Cambridge, MA, 2004.
- T. Poggio, R. Rifkin, S. Mukherjee and A. Rakhlin, *Bagging Regularizes*. AI Memo 2002-003, Massachusetts Institute of Technology, Cambridge, MA, 2002.

## SOFTWARE

*Transfer Learning Toolkit* in Matlab (with W.C. Kao); <http://multitask.cs.berkeley.edu>

## RESEARCH INTERESTS

prediction methods • machine learning • online learning • sequential decision making • statistical learning theory • online optimization • probabilistic methods • game theory • multitask learning • algorithms • empirical process theory • concentration of measure • applied probability • information theory

## SERVICE

Conference on Learning Theory *Steering Committee Member* 2011-2014

Editorial Board: Machine Learning Journal 2011-2014

Guest Editor: Journal of Computer and System Sciences, Special Issue on Learning Theory, 2010

Program Committees: NIPS 2011 Area Chair, COLT 2011, COLT 2009, ALT 2009, AISTATS 2007

## CONFERENCE PRESENTATIONS AND INVITED TALKS

- *Learning Theory: A Minimax Analysis*, Mathematical and Computational Foundations of Learning Theory, Dagstuhl, Germany, Jul 2011; Mini-Workshop “Mathematics of Learning Theory”, Oberwolfach, Germany, Aug 2011
- *Complexity-Based Approach to Calibration with Checking Rules*, The Annual Conference on Learning Theory, July 2011.
- *From Statistical Learning to Game-Theoretic Learning*, CMU Dept. of Statistics Seminar, Oct 2010; Sparse Statistics, Optimization and Machine Learning Workshop, Banff International Research Station, Jan 2011; Information Theory and Applications, David Blackwell’s Legacy, Feb 2011; UChicago Booth Statistics and Econometrics Seminar, Apr 2011; Joint Statistical Meeting, Aug 2011
- *Online Learning: Random Averages, Combinatorial Parameters, and Learnability*, European Meeting of Statisticians, Greece, Aug 2010.
- *Regularization Methods for Sequential Prediction*, New England Statistics Symposium, Harvard, Apr 2010; Wharton-Stern Statistics Conference, May 2010.
- *Making Sequential Decisions under Limited Feedback*, Allerton Conference on Communication, Control, and Computing, Oct 2009.
- *A Stochastic View of Optimal Regret through Minimax Duality*, UPenn Stat Seminar, Apr 2009; The Annual Conference on Learning Theory, Jun 2009.
- *Statistics Meets Optimization: A New Look at the Multi-armed Bandit Problem*, Duke Statistics Seminar, Apr 2009.
- *Beating the Adaptive Bandit with High Probability*, Information Theory and Applications Workshop, Feb 2008.
- *Online Learning with Limited Feedback*, FOCM Conference, Jun 2008; UPenn CIS Seminar, Sep 2008; UPenn Stats Seminar, Mar 2008; MIT, Mar 2008.
- *Adaptive Online Gradient Descent*, NIPS Conference, full oral presentation, Dec 2007.
- *Progress in Online Gradient Descent Methods*, IBM Theory of Computation Seminar, Nov 2007.
- *Multitask Learning with Expert Advice*, The Annual Conference on Learning Theory, Jun 2007.
- *Stability and Consistency of Approximate Empirical Risk Minimization Algorithms*, WNAR/IMS (Institute of Mathematical Statistics) Conference, Jun 2006
- *Stability and Consistency of Approximate Empirical Risk Minimization Algorithms*, International Workshop on Applied Probability, May 2006.
- *Algorithmic Stability in Learning Theory*, MIT CSAIL seminar, Jan 2006.
- *Stability of Clustering Methods*, Neural Information Processing Systems workshop *Theoretical Foundations of Clustering*, Dec 2005.
- *Applications of Empirical Process Theory in Statistical Learning*, UC Berkeley CIS seminar, Sep 2005.
- *Some Properties of Empirical Risk Minimization on Donsker Classes*, The Fourth International Conference on High Dimensional Probability, Jun 2005.
- *Risk Bounds for Mixture Density Estimation*, Toyota Technological Institute Seminar, Mar 2005.
- *Stability Results in Learning Theory*, Toyota Technological Institute Seminar, Apr 2005.