Alexander (Sasha) Rakhlin

MIT, E17-477
Cambridge, MA 02139

ACADEMIC APPOINTMENTS

01/2018 – **Massachusetts Institute of Technology**
Associate Professor (with tenure), Statistics and Data Science Center, IDSS
Department of Brain & Cognitive Sciences
Laboratory for Information & Decision Systems
Center for Brains, Minds, and Machines

01/2016 - 12/2016 **Massachusetts Institute of Technology**
Visiting Professor, Statistics and Data Science Center, IDSS

07/2015 - 01/2018 **University of Pennsylvania**
Associate Professor (with tenure), Department of Statistics, The Wharton School
Secondary appointment: Department of Computer & Information Science

01/2009 - 06/2015 **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.*

09/1996 - 06/2000 **Cornell University**
B.A. in Computer Science, B.A. in Mathematics. GPA 4.0/4.0

TEACHING

Winter 2018 **Department of Statistics, University of Pennsylvania**
STAT 405/705: Statistical Computing with R

Spring 2017 **Department of Statistics, University of Pennsylvania**
STAT 991: Online Methods in Machine Learning: Theory and Applications

Spring 2016 **EECS, Massachusetts Institute of Technology**
6.883: Online Methods in Machine Learning: Theory and Applications

Spring 2015 **Department of Statistics, University of Pennsylvania**
STAT 991: Concentration Inequalities

Fall 2009-2015 **Department of Statistics, University of Pennsylvania**
STAT 101: Intro to Statistics

Spring 2015	**Spring School "Structural Inference in Statistics", Germany** https://www.mathematik.hu-berlin.de/for1735/spring-school-2015
Spring 2012, 2014	**Department of Statistics, University of Pennsylvania** *STAT 928: Statistical Learning Theory and Sequential Prediction*
Summer 2012	**The 21st Machine Learning Summer School (Kyoto, Japan)** *Introduction to Statistical Learning Theory*
Spring 2012	**ENSAE, Paris Graduate School of Economics, Statistics and Finance** *From Statistical Learning to Sequential Prediction: A Minimax Approach (minicourse)*
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

AWARDS AND GRANTS

- NSF CAREER Award, Division of Mathematical Sciences, 2011–2015, *Statistical and Computational Complexities of Modern Learning Problems.*
- NSF CCF-1116928, Algorithmic Foundations, Division of Computer and Comm. Foundations, 2011–2014, *From Statistical to Worst-Case Learning: A Unified Framework.*
- NSF DMS-1521529, Computational and Data-Enabled Science and Engineering in Mathematical and Statistical Sciences, 2015-2017, *Collaborative Research: Novel Computational and Statistical Approaches to Prediction and Estimation* (co-PI with K. Sridharan)
- DARPA program on Fundamental Limits of Learning, *Foundations of Scalable Statistical Learning*, (with G. Bresler, A. Jadbabaie, A. Ozdaglar, P. Rigollet, D. Shah, S. Sra, and C. Uhler)
- Office of Naval Research Grant. *New Paradigms for Scalable Online Decentralized Optimization* (co-PI with A. Jadbabaie, A. Ribeiro, and A. Ozdaglar)
- NSF DMS-1342739, CDS&E-MSS, Division of Mathematics Sciences, 2013–2014.
- Dorinda and Mark Winkelman Distinguished Scholar Award, 2012
- Best Paper Award, Conference on Learning Theory, 2011
- Dean's Research Fund, Wharton School, 2010, 2012, 2014
- IBM Research's 2008 Pat Goldberg Memorial Best Paper Award in CE, EE and Math; Machine Learning Journal Award

EDITORIAL BOARDS, COMMITTEES

- Associate Editor, *The Annals of Statistics*, 2016–2018
- Program Chair (with V. Feldman), *Conference on Learning Theory*, 2016
- Action Editor, *Journal of Machine Learning*, 2013–present

- Editorial Board: *Machine Learning Journal*, 2011–2014
- Guest Editor: Journal of Computer and System Sciences, Special Issue on Learning Theory, 2010
- Conference on Learning Theory *Steering Committee Member*, 2011–2014
- Association for Computational Learning, Secretary, 2014–present
- Penn Research in Machine Learning, co-director, 2010–
- Penn Applied Mathematics and Computational Science, Executive Committee, 2013–
- Program Committees: COLT, NIPS, ALT, AISTATS (multiple years)

BOOK DRAFT

⋆ A. Rakhlin and K. Sridharan. *Statistical Learning and Sequential Prediction*, in preparation.

PREPRINTS

⋆ N. Golowich, A. Rakhlin, and O. Shamir. *Size-Independent Sample Complexity of Neural Networks*, 2017. arxiv.org/abs/1712.06541

⋆ T. Liang, T. Poggio, A. Rakhlin, and J. Stokes. *Fisher-Rao Metric, Geometry, and Complexity of Neural Networks*, 2017. arxiv.org/abs/1711.01530

⋆ A. Rakhlin and K. Sridharan. *A Tutorial on Online Supervised Learning with Applications to Node Classification in Social Networks*, 2016. arxiv.org/abs/1608.09014

⋆ T. Cai, T. Liang, and A. Rakhlin. *Weighted Message Passing and Minimum Energy Flow for Heterogeneous Stochastic Block Models with Side Information.* Submitted, 2017. arxiv.org/abs/1709.03907

⋆ T. Cai, T. Liang, and A. Rakhlin. *Inference via Message Passing on Partially Labeled Stochastic Block Models.* Submitted, 2016. arxiv.org/abs/1603.06923

⋆ A. Rakhlin and K. Sridharan. *Online Nonparametric Regression with General Loss Functions.* arxiv.org/abs/1501.06598

⋆ A. Rakhlin and K. Sridharan. *Sequential Probability Assignment with Binary Alphabets and Large Classes of Experts.* arxiv.org/abs/1501.07340

⋆ T. Liang, H. Narayanan, and A. Rakhlin. *On Zeroth-Order Stochastic Convex Optimization via Random Walks.* arxiv.org/abs/1402.2667

PUBLICATIONS (⋆ `denotes alphabetical ordering`)

⋆ D. Foster, A. Rakhlin, and K. Sridharan. *ZIGZAG: A new approach to adaptive online learning.* The 30th Annual Conference on Learning Theory (COLT 2017). arxiv.org/abs/1704.04010

⋆ M. Raginsky, A. Rakhlin, and M. Telgarsky. *Non-Convex Learning via Stochastic Gradient Langevin Dynamics: A Nonasymptotic Analysis.* The 30th Annual Conference on Learning Theory (COLT 2017). arxiv.org/abs/1702.03849

⋆ A. Rakhlin and K. Sridharan. *On Equivalence of Martingale Tail Bounds and Deterministic Regret Inequalities.* The 30th Annual Conference on Learning Theory (COLT 2017). arxiv.org/abs/1510.03925.

⋆ A. Rakhlin and K. Sridharan. *Efficient Multiclass Prediction on Graphs via Surrogate Losses*. The 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 2017.

⋆ T. Cai, T. Liang, and A. Rakhlin. *On Detection and Structural Reconstruction of Small-World Random Networks*. IEEE Transactions on Network Science and Engineering. Accepted, 2016. arxiv.org/abs/1604.06474

⋆ T. Cai, T. Liang, and A. Rakhlin. *Computational and Statistical Boundaries for Submatrix Localization*. The Annals of Statistics. Accepted, 2016. arxiv.org/abs/1502.01988

⋆ A. Rakhlin and K. Sridharan. *BISTRO: An Efficient Relaxation-Based Method for Contextual Bandits*. The 33rd International Conference on Machine Learning (ICML 2016). arxiv.org/abs/1602.02196

⋆ T. Cai, T. Liang, and A. Rakhlin. *Geometric Inference for General High-Dimensional Linear Inverse Problems*. The Annals of Statistics, 2016. arxiv.org/abs/1404.4408

⋆ M. Raginsky, A. Rakhlin, M. Tsao, Y. Wu, and A. Xu. *Information-Theoretic Analysis of Stability and Bias of Learning Algorithms*. The IEEE Information Theory Workshop, 2016.

• S. Shahrampour, A. Rakhlin, and A. Jadbabaie. *Multi-Armed Bandits in Multi-Agent Networks*. The 42nd IEEE International Conference on Acoustics, Speech and Signal Processing, 2016.

⋆ D. Foster, A. Rakhlin, and K. Sridharan. *Adaptive Online Learning*. Advances in Neural Information Processing Systems (NIPS 2015). arxiv.org/abs/1508.05170

⋆ A. Rakhlin and K. Sridharan. *Hierarchies of Relaxations for Online Prediction Problems with Evolving Constraints*. The 28th Annual Conference on Learning Theory (COLT 2015). arxiv.org/abs/1503.01212

⋆ T. Liang, A. Rakhlin, and K. Sridharan. *Learning with Square Loss: Localization through Offset Rademacher Complexity*. The 28th Annual Conference on Learning Theory (COLT 2015). arxiv.org/abs/1502.06134

⋆ A. Belloni, T. Liang, H. Narayanan, and A. Rakhlin. *Escaping the Local Minima via Simulated Annealing: Optimization of Approximately Convex Functions*. The 28th Annual Conference on Learning Theory (COLT 2015). arxiv.org/abs/1501.07242

⋆ H. Narayanan and A. Rakhlin. *Efficient Sampling from Time-Varying Distributions*. Journal of Machine Learning Research. Accepted, 2015. arxiv.org/abs/1309.5977

⋆ A. Rakhlin, K. Sridharan, and A. Tewari. *Sequential Complexities and Uniform Martingale Laws of Large Numbers*. Probability Theory and Related Fields. vol. 161, no. 1-2, pp. 111–153, 2015.

⋆ A. Rakhlin, K. Sridharan, and A. Tewari. *Online Learning via Sequential Complexities*. Journal of Machine Learning Research, vol 16, pp. 155–186, 2015.

• S. Shahrampour, A. Rakhlin, and A. Jadbabaie. *Distributed Detection: Finite-time Analysis and Impact of Network Topology*. IEEE Transactions on Automatic Control. Accepted, 2015. arxiv.org/abs/1409.8606

⋆ A. Jadbabaie, A. Rakhlin, S. Shahrampour, and K. Sridharan. *Online Optimization: Competing with Dynamic Comparators*. The 18th International Conference on Artificial Intelligence and Statistics (AISTATS), 2015.

⋆ A. Rakhlin, K. Sridharan, and A. Tsybakov. *Empirical Entropy, Minimax Regret and Minimax Risk*. Bernoulli Journal. Forthcoming, 2014. (accepted 09/2014)

⋆ A. Rakhlin and K. Sridharan. *Online Nonparametric Regression.* The 27th Annual Conference on Learning Theory (COLT 2014).

⋆ A. Rakhlin and K. Sridharan. *On Martingale Extensions of Vapnik-Chervonenkis Theory with Applications to Online Learning.* Book chapter: in A. Chervonenkis Festschrift. To appear, 2014.

⋆ G. Bartók, D. Foster, D. Pál, A. Rakhlin, and C. Szepesvári. *Partial Monitoring — Classification, Regret Bounds, and Algorithms.* Mathematics of Operations Research, vol. 39, no. 4, pp. 967-997, 2014.

⋆ A. Rakhlin and K. Sridharan. *On Semi-Probabilistic Universal Prediction.* IEEE Information Theory Workshop, 2013. Invited paper.

⋆ A. Rakhlin and K. Sridharan. *Optimization, Learning, and Games with Predictable Sequences.* Advances in Neural Information Processing Systems (NIPS 2013).

• S. Shahrampour, A. Rakhlin, and A. Jadbabaie. *Online Learning of Dynamic Parameters in Social Networks.* Advances in Neural Information Processing Systems (NIPS 2013).

⋆ W. Han, A. Rakhlin, and K. Sridharan. *Competing with Strategies.* The 26th Annual Conference on Learning Theory (COLT 2013).

⋆ A. Rakhlin and K. Sridharan. *Online Learning with Predictable Sequences.* The 26th Annual Conference on Learning Theory (COLT 2013).

⋆ A. Rakhlin, O. Shamir, and K. Sridharan. *Localization and Adaptation in Online Learning.* The 16th International Conference on Artificial Intelligence and Statistics (AISTATS 2013).

⋆ A. Agarwal, D. Foster, D. Hsu, S. Kakade, and A. Rakhlin. *Stochastic Convex Optimization with Bandit Feedback.* SIAM Journal on Optimization, vol. 23, no. 1, pp. 213-240, 2013. (Conference version: NIPS'11).

⋆ A. Rakhlin, O. Shamir, and K. Sridharan. *Relax and Randomize: From Value to Algorithms.* Advances in Neural Information Processing Systems (NIPS 2012). (`Accepted for full oral presentation`) .

⋆ A. Rakhlin, O. Shamir, and K. Sridharan. *Making Gradient Descent Optimal for Strongly Convex Stochastic Optimization.* The 29th International Conference on Machine Learning (ICML 2012).

⋆ D. Foster and A. Rakhlin. *No Internal Regret via Neighborhood Watch.* The 15th International Conference on Artificial Intelligence and Statistics (AISTATS 2012).

⋆ J. Abernethy, E. Hazan, and A. Rakhlin. *Interior-Point Methods for Full-Information and Bandit Online Learning.* IEEE Transactions on Information Theory, vol. 58, no. 7, pp. 4164–4175, 2012.

• S. Seshia and A. Rakhlin. *Quantitative Analysis of Systems Using Game-Theoretic Learning.* ACM Transactions on Embedded Computing Systems, vol. 11, no. S2, pp. 1-27, 2012.

⋆ M. Raginsky and A. Rakhlin. *Lower Bounds for Passive and Active Learning.* Advances in Neural Information Processing Systems (NIPS 2011).

⋆ A. Rakhlin, K. Sridharan, and A. Tewari. *Online Learning: Stochastic, Constrained, and Smoothed Adversaries.* Advances in Neural Information Processing Systems (NIPS 2011).

⋆ D. Foster, A. Rakhlin, K. Sridharan, and A. Tewari. *Complexity-Based Approach to Calibration with Checking Rules.* The 24th Annual Conference on Learning Theory (COLT 2011).

⋆ A. Rakhlin, K. Sridharan, and A. Tewari. *Online Learning: Beyond Regret.* The 24th Annual Conference on Learning Theory (COLT 2011). (**COLT'11 Best Paper Award**)

⋆ M. Raginsky and A. Rakhlin. *Information-Based Complexity, Feedback and Dynamics in Convex Programming.* IEEE Transactions on Information Theory, vol. 57, no. 10, pp. 7036–7056, 2011.

⋆ A. Rakhlin, K. Sridharan, and A. Tewari. *Online Learning: Random Averages, Combinatorial Parameters, and Learnability.* Advances in Neural Information Processing Systems (NIPS 2010). (`Accepted for full oral presentation`)

⋆ H. Narayanan and A. Rakhlin. *Random Walk Approach to Regret Minimization.* Advances in Neural Information Processing Systems (NIPS 2010).

⋆ M. Raginsky, A. Rakhlin, and S. Yüksel. *Online Convex Programming and Regularization in Adaptive Control.* The 49th IEEE Conference on Decision and Control (CDC 2010).

⋆ J. Abernethy and A. Rakhlin. *Beating the Adaptive Bandit with High Probability.* The 22nd Annual Conference on Learning Theory (COLT 2009).

⋆ J. Abernethy, A. Agarwal, P. Bartlett, and A. Rakhlin. *A Stochastic View of Optimal Regret through Minimax Duality.* The 22nd Annual Conference on Learning Theory (COLT 2009).

• S. Seshia and A. Rakhlin. *Game-Theoretic Timing Analysis.* International Conference on Computer-Aided Design (ICCAD 2008).

⋆ J. Abernethy, E. Hazan, and A. Rakhlin. *Competing in the Dark: An Efficient Algorithm for Bandit Linear Optimization.* The 21st Annual Conference on Learning Theory (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.* The 21st Annual Conference on Learning Theory (COLT 2008).

⋆ P. Bartlett, V. Dani, T. Hayes, S. Kakade, A. Rakhlin, and A. Tewari. *High-Probability Regret Bounds for Bandit Online Linear Optimization.* The 21st Annual Conference on Learning Theory (COLT 2008).

⋆ P. Bartlett, E. Hazan, and A. Rakhlin. *Adaptive Online Gradient Descent.* Advances in Neural Information Processing Systems (NIPS 2007). (`Accepted for full oral presentation`)

• A. Rakhlin, J. Abernethy, and P. Bartlett. *Online Discovery of Similarity Mappings.* The 24th International Conference on Machine Learning (ICML 2007) pp. 767–774.

⋆ J. Abernethy, P. Bartlett, and A. Rakhlin. *Multitask Learning with Expert Advice.* The 20th Annual Conference on Learning Theory (COLT 2007), pp. 484–498.

• A. Rakhlin and A. Caponnetto. *Stability of K-Means Clustering.* Advances in Neural Information Processing Systems (NIPS 2007), pp. 1121–1128.

• 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. 7, pp. 2565–2583, 2006.

• A. Rakhlin, D. Panchenko, and S. Mukherjee. *Risk Bounds for Mixture Density Estimation.* ESAIM: Probability and Statistics, vol. 9, pp. 220–229, 2005.

- A. Rakhlin, S. Mukherjee, and T. Poggio. *Stability Results in Learning Theory.* Analysis and Applications (Special Issue on Learning Theory), vol. 3, no. 4, pp. 397–417, 2005.
- T. Poggio, S. Mukherjee, R. Rifkin, A. Rakhlin and A. Verri, *B.* Uncertainty in Geometric Computations, J. Winkler and M. Niranjan (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.

RESEARCH INTERESTS

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