

RESEARCH SUMMARY

My research focuses on “**mathematics and computation for machine learning.**” In particular, I design, develop and analyze algorithms for **machine learning, deep learning, and statistics**. My research involves convex optimization, nonconvex optimization, statistics, computer science, signal processing, differential geometry, {convex, functional, *matrix*, and *numerical*} *analysis*. I care about distributed and parallel computation, especially for **large-scale machine learning** and for **optimization under uncertainty**. I am interested in a variety of applications, including *semantic search, healthcare ML, materials science, and engineering*.

EDUCATION

Ph.D. in Computer Science

The University of Texas at Austin, Aug. 2007

Title: *Matrix Nearness Problems in Data Mining*. Advisor: I. S. Dhillon

M.S. in Computer Science

University of Texas at Austin, Aug. 2006

B. Engg. (Hons.) in Computer Science

Birla Inst. of Tech. & Science, Pilani, India, Jun. 1999

POSITIONS

Research Faculty (Principal Research Scientist) at Laboratory for Information & Decision Systems (LIDS), Massachusetts Institute of Technology (MIT); Jan. 2015–

Visiting Assistant Professor, Machine Learning Department, School of Computer Science (SCS), Carnegie Mellon University, Pittsburgh; Sep. 2013–May 2014

Visiting faculty, EECS, University of California, Berkeley; Jan.–May 2013

Senior Research Scientist in Prof. Bernhard Schölkopf’s group

Max Planck Institute for Intelligent Systems Tübingen, Germany; Jul. 2012–Aug. 2014

Research Scientist hosted by Prof. Bernhard Schölkopf

Max Planck Institute for Intelligent Systems Tübingen, Germany;
(formerly MPI for Biological Cybernetics). Oct. 2007–Jun. 2012

Visiting Researcher hosted by Prof. Jeff Bilmes

EE Department, University of Washington, Seattle. Jul.–Aug. 2011

Research Assistant to Prof. I. S. Dhillon

Univ. of Texas at Austin, Austin, TX, USA. Jan. 2002–Aug. 2007 (*several times*)

Research Intern with Arun Surendran in John Platt’s group

Microsoft Research, Redmond, WA, USA. May 2005–Aug. 2005.

HONORS & AWARDS

- **SIAM Outstanding Paper Prize**, July 2011
For the paper “*The Metric Nearness Problem*,” (with J. Brickell, I. Dhillon, and J. Tropp in *SIMAX* 2008)—**one of three papers amongst all papers published in SIAM journals within 2008–2010**.
(More information at: http://www.siam.org/prizes/sponsored/outstanding_paper.php)
- **Best Paper Runner Up Award**, Sep., 2011
European Conference on Machine Learning (ECML 2011), Athens, Greece.

-
- **Best Student Paper**, Jun. 2007
(with J. Davis, B. Kulis, P. Jain, and I. S. Dhillon)
International Conference on Machine Learning (ICML)
 - **Best of SDM Papers** Apr. 2007
(with D. Kim and I. S. Dhillon)
SIAM Data Mining Conference (SDM)
 - **Householder Symposium Attendance Award**, Jun., 2011
Award for attending the Householder Symposium XVIII, Tahoe City.
 - **MIC Postdoc Trainee Award**, Oct. 2009
IEEE Nuclear Science Symposium, Medical Imaging Conference (NSS / MIC)
 - **Best poster award**, Jul. 2009
(with S. Harmeling, M. Hirsch, and B. Schölkopf)
International Conference on Cosmology and Statistics (COSMOSTATS), Ascona, Switzerland
 - **Recipient of Microelectronics and Computer Development (MCD) Fellowship**
Univ. of Texas at Austin, Aug. 2000–Aug. 2004

TEACHING (CLASSES / SEMINARS)

- **Introduction to Machine Learning** (Fall 2017)
Course: 6.867, Graduate (MS & PhD) level course, EECS Department, MIT
(Co-taught with Devavrat Shah and David Sontag)
- **Introduction to Machine Learning** (Fall 2016)
Course: 6.867, Graduate (MS & PhD) level course, EECS Department, MIT
(Co-taught with Leslie Kaelbling)
- **OPTML++: Optimization for Machine Learning** (Fall 2016)
Research seminar, LIDS, EECS, MIT.
- **Introduction to Machine Learning** (Spring 2016)
Course: 6.036, Undergraduate level, EECS Department, MIT
(co-taught with Tommi Jaakkola and Regina Barzilay)
- **OPTML++: Optimization for Machine Learning** (Fall 2015)
Research seminar, LIDS, EECS, MIT.
- **Introduction to Machine Learning** (Spring 2015)
Course: 6.036, Undergraduate level, EECS Department, MIT
(co-taught with Tommi Jaakkola and Regina Barzilay)
- **Advanced Optimization and Randomized Methods** (Spring 2014)
Graduate course, *ML Department, Carnegie Mellon University, Pittsburgh*
(I created and co-taught this course with Prof. Alex Smola)
- **Convex optimization; EE 227A** (Spring 2013) (by invitation)
Graduate course, *EECS Department, University of California, Berkeley.*
- **Introduction to Logic.** (Univ. of Texas at Austin; Spring 2006)
Teaching Assistant for CS313K; Lecturing component: 3 hours per week
- **Elements of Java.** (Univ. of Texas at Austin; Spring 2005, Fall 2003)
Teaching Assistant for CS303E; Lecturing component: 3 hours per week
- **Graduate level numerical linear algebra.** (Univ. of Texas at Austin; Fall 2002)
Teaching Assistant for CS383C
- **Analysis of Programs.** (Univ. of Texas at Austin; Fall 2001)
Teaching Assistant for CS336

TEACHING (INVITED TUTORIALS AND LECTURES)

- **Introductory lectures on Machine Learning and Optimization**, (Jul 2017)
Invited lectures at the: Peking University (PKU), Beijing, China, 15th Annual Applied Math Summer School, PKU.
- **Introduction to Optimization in Machine Learning**, (Jun 2017)
Invited lectures at the: Machine Learning Summer School (MLSS), Tübingen (other lecturers: Michael Jordan, Bernhard Schölkopf, Zoubin Ghahramani, Jure Leskovec, and more)
- **Modern Stochastic Optimization in Machine Learning** (Dec 2016)
*Tutorial at the *Advances in Neural Information Processing Systems (NIPS)* conference (Co-taught with Francis Bach; 600+ audience (estimated).)*
- **Introduction to Machine Learning** (April 2014)
Invited short-course at the EU Regional School, 2014. RWTH Aachen, Germany.
- **Introduction to large-scale optimization**; (Jan 2013) (by invitation)
*Intensive graduate level course (20 hrs); *University Carlos III of Madrid (Spain)**
- **Optimization in Machine Learning**. (April 2011)
*Invited lecture for CS graduate students at: *Universidad Autónoma de Madrid (Spain)*.*
- **Introductory Lectures on Scientific Writing**. (Sep.–Oct. 2009)
*Three lectures in *Scientific Writing*, given at *Max-Planck Institute, Tübingen, Germany**
- **Matrix Factorization and Approximation Problems**. (April 2010)
Invited short-course at the EU Regional School, 2010. RWTH Aachen, Germany.

GRANTS / FUNDING

- **NSF-BIGDATA** (PI; #PIs: 2; \$1.02M); Oct 2017–Sep 2021.
Automating data analysis: interpretable, interactive, and scalable learning via discrete probability
- **Criteo Faculty Research Award** (PI); \$25,000;
- **Toyota Research Institute (TRI)** (Co-PI; #PIs: 5; ~\$4,000,000; 4 yrs); Mar 2017–Dec 2021
Accelerated Materials Design and Discovery
- **DARPA** (Co-PI; #PIs: 8; ~\$1,000,000); Sep 2016–Feb 2018
Fundamentals Limits of Learning (FunLoL).
- **Lincoln Labs** (PI; \$155,000); Jan 2016–Aug 2017
Statistics without Affirmed Ground Truth
- **NSF-Medium** (Co-PI; #PIs: 3; \$1,200,000); Jan 2015–Aug 2018.
CSR: Distributed Inference Algorithms for Machine Learning and Optimization.
- **Elsevier Mathematical Sciences Sponsorship Fund** \$2000.
- **Funding for OPT 2013**, 6th Int. Workshop on Optimization for Machine Learning
Value: GBP 2000 from **MSR, Cambridge**
- **OPT 2012**: €3040 from **PASCAL2**
OPT 2011: €2500 from **PASCAL2**
OPT 2010: €1550 from **PASCAL2**, \$4000 from Microsoft.
OPT 2009, €1000 from **PASCAL2**, \$2500 from **MOSEK**, \$1000 from **Microsoft Research**
OPT 2008, €4855 from **PASCAL2**, Oct. 2008
- **Funding for NUMML 2010**, NIPS Workshop on Numerical Challenges in Machine Learning
Value: €2500 from **PASCAL2**
NUMML 2009, ICML 2009 Workshop, €4100, Mar. 2009
- **Helped prepare NSF proposal** (PI: Inderjit Dhillon): *Non-Negative Matrix and Tensor Approximations: Algorithms, Software and Applications*, NSF, CCF-0728879, \$250,000, 01/01/08-12/31/10

PUBLICATIONS

1. A star (e.g., A. Student*) signifies a student who received supervision from me on that paper.
2. Google scholar: <https://scholar.google.com/citations?user=eyCw9goAAAAJ>.

Submitted work

1. Reshad Hosseini and **Suvrit Sra**. “An Alternative to EM for Gaussian Mixture Models: Batch and Stochastic Riemannian Optimization”. *Submitted*, Jun. 2017. *Submitted*
2. Zelda Mariet*, John Holodnak, Jason Matterer, and **Suvrit Sra**. “Labels as Features for Cluster-Based Evaluation of Classifiers”. *Submitted*, June 2017
3. **Suvrit Sra**. “Logarithmic inequalities under an elementary symmetric polynomial dominance order”. *arXiv:1509.01618v2*, Jun. 2017. *Submitted*
4. Chengtao Li*, David Alvarez-Melis, Keyulu Xu, Stefanie Jegelka, and **Suvrit Sra**. “Distributional Adversarial Networks”. May 2017. *Submitted*
5. Sashank Reddi*, Manzil Zaheer, **Suvrit Sra**, Francis Bach, Barnabas Poczos, Ruslan Salakhutdinov, and Alexander J. Smola. “A Generic Approach for Escaping Saddle points”. May 2017. *Submitted*
6. Chulhee Yun*, **Suvrit Sra**, and Ali Jadbabaie. “Global optimality conditions for deep neural networks”. May 2017. *Submitted*
7. Álvaro J. Barbero and **Suvrit Sra**. “Modular proximal optimization with application to total variation regularization”. *arXiv*, 2017. *Submitted*
8. Anoop Cherian, **Suvrit Sra**, and Richard Hartley. “Sequence Summarization Using Order-constrained Kernelized Feature Subspaces”. *arXiv:1705.08583*, 2017. *Submitted*

Books and Monographs

9. Suvrit Sra, Sebastian Nowozin, and Stephen J. Wright, editors. *Optimization for Machine Learning*. MIT Press, Oct. 2011. Our book distills research at the confluence of optimization and machine learning. It includes contributions from leading researchers in both fields; we aim to achieve a cogent summary of the state-of-the-art, while still remaining didactic.

Work at MIT (2015–)

10. Ke Jiang, **Suvrit Sra**, and Brian Kulis. “Combinatorial topic modling using small variance asymptotics”. In *Artificial Intelligence and Statistics (AISTATS)*, 2017
11. Chengtao Li*, Stefanie Jegelka, and **Suvrit Sra**. “Polynomial time algorithms for dual volume sampling”. In *Advances in Neural Information Processing Systems (NIPS)*, 2017
12. Zelda Mariet* and **Suvrit Sra**. “Elementary Symmetric Polynomials for Optimal Experimental Design”. In *Advances in Neural Information Processing Systems (NIPS)*, 2017
13. **Suvrit Sra**. “Directional Statistics in Machine Learning”. In C. Ley and T. Verdebout, editors, *Modern Directional Statistics*. Chapman and Hall, 2017. *Invited chapter*
14. Chengtao Li*, Stefanie Jegelka, and **Suvrit Sra**. “Fast Mixing Markov Chains for Strongly Rayleigh Measures, DPPs, and Constrained Sampling”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2016
15. Zelda Mariet* and **Suvrit Sra**. “Kronecker Determinantal Point Processes”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2016
16. Sashank Reddi*, **Suvrit Sra**, Barnabas Poczos, and Alexander J. Smola. “Fast incremental methods for smooth nonconvex optimization”. In *IEEE Conf. Decision and Control (CDC)*, December 2016

-
17. Sashank Reddi*, **Suvrit Sra**, Barnabas Poczos, and Alexander J. Smola. "Fast Stochastic Methods for Nonsmooth Nonconvex Optimization". In *Advances in Neural Information Processing Systems (NIPS)*, December 2016
 18. Hongyi Zhang*, Sashank Reddi*, and **Suvrit Sra**. "Fast stochastic optimization on Riemannian manifolds". In *Advances in Neural Information Processing Systems (NIPS)*, December 2016
 19. Sashank Reddi*, **Suvrit Sra**, Barnabas Poczos, and Alexander J. Smola. "Stochastic Frank-Wolfe Methods for Nonsmooth Nonconvex Optimization". In *54th Annual Allerton Conference on Communication, Control, Computing*, October 2016
 20. Chengtao Li*, Stefanie Jegelka, and **Suvrit Sra**. "Fast DPP Sampling for Nyström with Application to Kernel Methods". In *International Conference on Machine Learning (ICML)*, Jun. 2016
 21. Chengtao Li*, **Suvrit Sra**, and Stefanie Jegelka. "Gaussian quadrature for matrix inverse forms with applications". In *International Conference on Machine Learning (ICML)*, Jun. 2016
 22. Sashank Reddi*, Ahmed Hefny*, **Suvrit Sra**, Barnabas Poczos, and Alexander J. Smola. "Stochastic variance reduction for nonconvex optimization". In *International Conference on Machine Learning (ICML)*, Jun. 2016
 23. Yu-Xiang Wang*, Veeranjanyulu Sadhanala*, Wei Dai*, Willie Neiswanger*, **Suvrit Sra**, and Eric P. Xing. "Parallel and Distributed Block-Coordinate Frank-Wolfe Algorithms". In *International Conference on Machine Learning (ICML)*, Jun. 2016
 24. Pourya H. Zadeh*, Reshad Hosseini, and **Suvrit Sra**. "Geometric Mean Metric Learning". In *International Conference on Machine Learning (ICML)*, Jun. 2016
 25. Hongyi Zhang* and **Suvrit Sra**. "First-order methods for geodesically convex optimization". In *Conference on Learning Theory (COLT)*, Jun. 2016. arXiv:1602.06053
 26. Chengtao Li*, Stefanie Jegelka, and **Suvrit Sra**. "Efficient Sampling for K-Determinantal Point Processes". In *Artificial Intelligence and Statistics (AISTATS'16)*, May 2016
 27. Zeldia Mariet* and **Suvrit Sra**. "Diversity Networks". In *International Conference on Learning Representations (ICLR)*, May 2016. arXiv:1511.0577
 28. **Suvrit Sra**, Adams W. Yu*, Mu Li*, and Alexander Smola. "AdaDelay: Delay Adaptive Distributed Stochastic Optimization". In *Artificial Intelligence and Statistics (AISTATS'16)*, May 2016
 29. Lev Borisov, Patrizio Neff, **Suvrit Sra**, and Christian Thiel. "The sum of squared logarithms inequality in arbitrary dimensions". *Linear Algebra and its Applications (LAA)*, 2016
 30. Anoop Cherian and Suvrit Sra. "Riemannian Dictionary Learning and Sparse Coding for Positive Definite Matrices". *IEEE Tran. on Neural Networks and Learning Systems (TNNLS)*, 2016
 31. Anoop Cherian and **Suvrit Sra**. "Positive Definite Matrices: Data Representation and Applications to Computer Vision". In *Algorithmic Advances in Riemannian Geometry and Applications*. Springer, 2016
 32. Justin Solomon, Gabriel Peyré, Vladimir Kim, and **Suvrit Sra**. "Entropic Metric Alignment for Correspondence Problems". *ACM Transactions on Graphics (ToG)*, 35(4), 2016
 33. **Suvrit Sra** and Reshad Hosseini. "Geometric Optimization in Machine Learning". In *Algorithmic Advances in Riemannian Geometry and Applications*. Springer, 2016
 34. **Suvrit Sra**. "On inequalities for normalized Schur functions". *European J. Combinatorics*, Volume 51:492–494, 2016
 35. **Suvrit Sra**. "On the matrix square root and geometric optimization". *Electronic Journal on Linear Algebra (ELA)*, 2016
 36. **Suvrit Sra**. "Positive definite matrices and the S-divergence". *Proceedings American Mathematical Society (PAMS)*, 2016. published online Oct 2015
 37. Anoop Cherian and **Suvrit Sra**. "Positive Definite Matrices: Data Representation and Applications to Computer Vision". In M. H. Quang and V. Murino, editors, *Riemannian geometry in machine learning, statistics, optimization, and computer vision*. Springer, December 2015. Invited chapter

-
38. Reshad Hosseini and **Suvrit Sra**. “Matrix Manifold Optimization for Gaussian Mixtures”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2015
 39. Sashank Reddi*, Ahmed Hefny*, **Suvrit Sra**, Barnabas Poczos, and Alexander Smola. “Asynchronous variance reduced stochastic gradient descent”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2015
 40. **Suvrit Sra** and Reshad Hosseini. “Geometric optimization in machine learning”. In M. H. Quang and V. Murino, editors, *Riemannian geometry in machine learning, statistics, optimization, and computer vision*. Springer, December 2015. *Invited chapter*
 41. Zelda Mariet* and **Suvrit Sra**. “Fixed-point algorithms for learning determinantal point processes”. In *International Conference on Machine Learning (ICML)*, Jun. 2015
 42. Wolfgang Berndt and **Suvrit Sra**. “Hlawka–Popoviciu inequalities on positive definite tensors”. *Linear Algebra and its Applications*, 486:317–327, 2015
 43. Reshad Hosseini, **Suvrit Sra**, Lucas Theis, and Matthias Bethge. “Statistical inference with the Elliptical Gamma Distribution”. In *Artificial Intelligence and Statistics (AISTATS)*, 2015
 44. Minghua Lin and **Suvrit Sra**. “A proof of Thompson’s determinantal inequality”. *Mathematical Notes*, 2015
 45. Sashank Reddi*, Ahmed Hefny*, Carlton Downey*, Avinava Dubey*, and **Suvrit Sra**. “Large-scale randomized-coordinate descent methods with non-separable linear constraints”. In *Uncertainty in Artificial Intelligence (UAI)*, 2015
 46. **Suvrit Sra** and Reshad Hosseini. “Conic geometric optimisation on the manifold of positive definite matrices”. *SIAM J. Optimization (SIOPT)*, 2015

Pre MIT (2014 and earlier)

47. Adams W. Yu*, Wanli Ma*, Yaoliang Yu, Jaime G. Carbonell, and **Suvrit Sra**. “Efficient Structured Matrix Rank Minimization”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2014
48. Samaneh Azadi* and **Suvrit Sra**. “Towards optimal stochastic alternating direction method of multipliers”. In *Int. Conf. on Mach. Learning (ICML)*, Jun 2014
49. Anoop Cherian and **Suvrit Sra**. “Riemannian Sparse Coding of Positive Definite Matrices”. In *European Conf. Computer Vision (ECCV)*, Jun 2014
50. David Lopez-Paz*, **Suvrit Sra**, Alexander J. Smola, Zoubin Ghahramani, and Bernhard Schölkopf. “Randomized nonlinear component analysis”. In *Int. Conf. Machine Learning (ICML)*, Jun 2014
51. Matt Wytock, **Suvrit Sra**, and Zico Kolter. “Fast Newton methods for the group fused lasso”. In *Uncertainty in Artificial Intelligence (UAI)*, Jun 2014
52. **Suvrit Sra**. “Nonconvex proximal splitting: batch and incremental algorithms”. In J.A.K. Suykens, M. Signoretto, and A. Argyriou, editors, *Regularization, Optimization, Kernels, and Support Vector Machines*. Cambridge University Press, March 2014
53. Anoop Cherian, **Suvrit Sra**, V. Morellas, and N. Papanikolopoulos. “Efficient nearest neighbors via robust sparse hashing”. *IEEE Transactions Image Processing*, 2014
54. Reshad Hosseini, **Suvrit Sra**, Lucas Theis, and M. Bethge. “Statistical inference with the Elliptical Gamma Distribution”. *Computational Statistics and Data Analysis (CSDA)*, 2014
55. Stefanie Jegelka, Francis Bach, and **Suvrit Sra**. “Reflection methods for user-friendly submodular optimization”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2013
56. **Suvrit Sra** and Reshad Hosseini. “Geometric optimization on positive definite matrices with application to elliptically contoured distributions”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2013
57. **Suvrit Sra**. “Tractable Large-Scale Optimization in Machine Learning”. In L. Bordeaux, Y. Hamadi,

-
- P. Kohli, and R. Mateescu, editors, *Advances in Tractability*. Cambridge University Press, December 2013. 29 pages
58. Carlos M. Alaíz*, Francesco Dinuzzo, and **Suvrit Sra**. “Correlation matrix nearness and completion under observation uncertainty”. *IMA Journal of Numerical Analysis*, Oct. 2013. 16 pages
 59. **Suvrit Sra**. “Explicit eigenvalues of certain scaled trigonometric matrices”. *Linear Algebra and its Applications (LAA)*, 438:173–181, 2013. 9 pages
 60. **Suvrit Sra** and Dmitrii Karp. “The multivariate Watson distribution: Maximum-likelihood estimation and other aspects”. *Journal of Multivariate Analysis (JMVA)*, 114:256–269, 2013
 61. Anoop Cherian*, **Suvrit Sra**, A. Banerjee, and N. Papanikolopoulos. “Jensen-Bregman LogDet Divergence with Application to Efficient Similarity Search for Covariance Matrices”. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, Dec. 2012. 14 pages
 62. **Suvrit Sra**. “A new metric on the manifold of kernel matrices with application to matrix geometric means”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2012
 63. **Suvrit Sra**. “Scalable nonconvex inexact proximal splitting”. In *Advances in Neural Information Processing Systems (NIPS)*, December 2012
 64. **Suvrit Sra**. “Fast projections onto mixed-norm balls with applications”. *Data Mining and Knowledge Discovery (DMKD)*, 25(2), Nov. 2012. 20 pages
 65. Dongmin Kim, **Suvrit Sra**, and Inderjit S. Dhillon. “A non-monotonic method for large-scale non-negative least squares”. *Optimization Methods and Software (OMS)*, Dec. 2011. 28 pages
 66. Anoop Cherian*, **Suvrit Sra**, A. Banerjee, and N. Papanikolopoulos. “Efficient Similarity Search for Covariance Matrices via the Jensen-Bregman LogDet Divergence”. In *International Conference on Computer Vision (ICCV)*, Nov. 2011
 67. Mark Schmidt, Dongmin Kim*, and **Suvrit Sra**. “Projected Newton-type methods in machine learning”. In S. Sra, S. Nowozin, and S. J. Wright, editors, *Optimization for Machine Learning*. MIT Press, Oct. 2011. 25 pages
 68. **Suvrit Sra** and Anoop Cherian*. “Generalized Dictionary Learning for Symmetric Positive Definite Matrices with Application to Nearest Neighbor Retrieval”. In *European Conf. Machine Learning (ECML)*, Sep. 2011
 69. **Suvrit Sra**. “Fast projections onto $\ell_{1,q}$ -norm balls for grouped feature selection”. In *European Conf. Machine Learning (ECML)*, Sep. 2011. **Best paper runner up award**
 70. Álvaro J. Barbero* and **Suvrit Sra**. “Fast Newton-type Methods for Total-Variation with Applications”. In *Proceedings of the International Conference on Machine Learning (ICML)*, Jun. 2011
 71. Anoop Cherian*, **Suvrit Sra**, and Nikos Papanikolopoulos. “Denoising sparse noise via online dictionary learning”. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011
 72. Michael Hirsch*, Stefan Harmeling, **Suvrit Sra**, and Bernhard Schölkopf. “Online Multi-frame Blind Deconvolution with Super-resolution and Saturation Correction”. *Astronomy & Astrophysics (AA)*, Feb. 2011. 11 pages
 73. **Suvrit Sra**. “A short note on parameter approximation for von Mises-Fisher distributions: and a fast implementation of $I_s(x)$ ”. *Computational Statistics*, Jan. 2011. 14 pages
 74. Dongmin Kim*, **Suvrit Sra**, and Inderjit S. Dhillon. “Tackling box-constrained convex optimization via a new projected quasi-Newton approach”. *SIAM J. Scientific Computing (SISC)*, 32(6):3548–3563, Dec. 2010. 16 pages
 75. Michael Hirsch*, **Suvrit Sra**, Bernhard Schölkopf, and Stefan Harmeling. “Efficient Filter Flow for Space-Variant Multiframe Blind Deconvolution”. In *IEEE Conf. Computer Vision & Pattern Recognition (CVPR)*, Jun. 2010
 76. Stefan Harmeling, **Suvrit Sra**, Michael Hirsch*, and Bernhard Schölkopf. “Multiframe Blind Deconvolution, Super-Resolution, and Saturation Correction via Incremental EM”. In *IEEE International*

Conference on Image Processing (ICIP), 2010

77. Dongmin Kim*, **Suvrit Sra**, and Inderjit S. Dhillon. "A scalable trust-region algorithm with application to mixed-norm regression". In *Int. Conf. Machine Learning (ICML)*, 2010
78. **Suvrit Sra**, Dongmin Kim, Inderjit S. Dhillon, and Bernhard Schölkopf. "A new non-monotonic algorithm for PET image reconstruction". In *IEEE Nuclear Science Symp. / Medical Imaging Conf. (NSS/MIC)*, Oct. 2009
79. Stefanie Jegelka, **Suvrit Sra**, and Arindam Banerjee. "Approximation Algorithms for Tensor clustering". In *Algorithmic Learning Theory (ALT)*, Jun. 2009. also arXiv: cs.DS/0812.0389
80. Stefan Harmeling, Michael Hirsch*, **Suvrit Sra**, and Bernhard Schölkopf. "Online Blind Deconvolution for Astronomy". In *IEEE Int. Conf. on Computational Photography (ICCP)*, Apr. 2009
81. Brian Kulis, **Suvrit Sra**, and Inderjit S. Dhillon. "Convex Perturbations for Scalable Semidefinite Programming". In D. van Dyk and M. Welling, editors, *Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, volume 5 of *JMLR W & CP*, pages 296–303, Apr. 2009
82. Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and **Suvrit Sra**. *Text Mining: Theory, Applications, and Visualization A. Sriovastava and M. Sahami (eds.)*, chapter Text clustering with mixture of von Mises-Fisher distributions. Data Mining & Knowledge Discovery Series. Chapman-Hall / CRC Press, 2009
83. **Suvrit Sra**. "Block-Iterative Algorithms for Non-negative Matrix Approximation". In *IEEE Int. Conf. Data Mining (ICDM)*, pages 1037–1042, Dec. 2008
84. Justin Brickell, Inderjit S. Dhillon, **Suvrit Sra**, and Joel A. Tropp. "The Metric Nearness Problem". *SIAM J. Matrix Analysis and Applications (SIMAX)*, 30(1):375–396, 2008. **SIAM Outstanding Paper Prize, 2011**. (one out of three papers selected from papers across all SIAM Journals in the years 2008–2010)
85. Jason V. Davis, Brian Kulis, Prateek Jain, **Suvrit Sra**, and Inderjit S. Dhillon. "Information-theoretic Metric Learning". In *Int. Conf. Machine Learning (ICML)*, Jun. 2007. **Best Student Paper**
86. Dongmin Kim, **Suvrit Sra**, and Inderjit S. Dhillon. "Fast Newton-type Methods for the Least Squares Nonnegative Matrix Approximation Problem". In *SIAM Int. Conf. Data Mining (SDM)*, Apr. 2007. **Best of SDM papers**
87. Dongmin Kim, **Suvrit Sra**, and Inderjit S. Dhillon. "Fast Projection-Based Methods for the Least Squares Nonnegative Matrix Approximation Problem". *Statistical Analysis and Data Mining*, 1:38–51, 2007
88. Arun Surendran and **Suvrit Sra**. "Incremental Aspect Models for Mining Document Streams". In *Principles and Practice of Knowledge Discovery in Databases (PKDD)*, Sep. 2006
89. **Suvrit Sra**. "Efficient Large Scale Linear Programming Support Vector Machines". In *Euro. Conf. Machine Learning (ECML)*, pages 767–774, Sep. 2006
90. **Suvrit Sra** and Joel A. Tropp. "Row-action Methods for Compressed Sensing". In *Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, May 2006
91. Inderjit S. Dhillon and **Suvrit Sra**. "Generalized Nonnegative Matrix Approximations with Bregman Divergences". In *Advances Neural Information Processing Systems (NIPS)*, Dec. 2005
92. Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and **Suvrit Sra**. "Clustering on the Unit Hypersphere using von Mises-Fisher Distributions". *J. Mach. Learning Research (JMLR)*, 6:1345–1382, Sep 2005
93. Inderjit S. Dhillon, **Suvrit Sra**, and Joel A. Tropp. "Triangle Fixing Algorithms for the Metric Nearness Problem". In *Advances in Neural Information Processing Systems (NIPS)*, Dec. 2004
94. Hyuk Cho, Inderjit S. Dhillon, Yuqiang Guan, and **Suvrit Sra**. "Minimum Sum Squared Residue based Co-clustering of Gene Expression data". In *SIAM Conf. on Data Mining (SDM)*, Apr. 2004
95. Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and **Suvrit Sra**. "Generative Model-Based Clustering of Directional Data". In *ACM Int. Conf. Knowledge Discovery and Data Mining (KDD)*, Aug. 2003

Selected Technical Reports (with ≥ 10 citations)

96. Rashish Tandon and **Suvrit Sra**. “Sparse nonnegative matrix approximation: new formulations and algorithms”. Technical Report 193, Max Planck Institute for Intelligent Systems, 2010. **(13 citations)**
97. **Suvrit Sra**, Stefanie Jegelka, and Arindam Banerjee. “Approximation algorithms for Bregman clustering, co-clustering and tensor clustering”. Technical Report 177, MPI for Biological Cybernetics, Jun. 2008. **(9 citations)**
98. **Suvrit Sra**, Dongmin Kim, and Bernhard Schölkopf. “Non-monotonic Poisson likelihood maximization”. Technical Report 170, Jun. 2008. **(12 citations)**
99. **Suvrit Sra** and Inderjit S. Dhillon. “Nonnegative Matrix Approximation: Algorithms and Applications”. Technical Report TR-06-27, Jun. 2006. **(53 citations)**
100. Dongmin Kim, **Suvrit Sra**, and Inderjit S. Dhillon. “A New Projected Quasi-Newton Approach for Nonnegative Least Squares Problem”. Technical Report TR-06-54, May 2006. **(30 citations)**
101. Inderjit S. Dhillon, **Suvrit Sra**, and Joel A. Tropp. “Triangle fixing algorithms for the Metric Nearness problem”. Technical Report TR-04-22, Jun. 2004. **(8 citations)**
102. Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and **Suvrit Sra**. “Clustering on Hyperspheres using Expectation Maximization”. Technical Report TR-03-07, Feb. 2003. **(12 citations)**
103. Inderjit S. Dhillon and **Suvrit Sra**. “Modeling data using directional distributions”. Technical Report TR-03-06, Jan. 2003. **(61 citations)**

Miscellaneous

104. **S. Sra**. “A new metric on the manifold of kernel matrices”. In *NIPS Workshop on Algebraic Topology and Machine Learning*, Dec. 2012
105. M. Langovoy and **S. Sra**. “Statistical estimation for optimization problems on graphs”. In *NIPS Workshop on Discrete Optimization for Machine Learning*, Dec. 2011
106. **S. Sra**. *Matrix Nearness Problems in Data Mining*. PhD thesis, Univ. of Texas at Austin, Aug. 2007
107. J. V. Davis, B. Kulis, **S. Sra**, and I. S. Dhillon. “Information-theoretic Metric Learning”. In *NIPS Workshop on Learning to Compare Examples*, Dec. 2006

INVITED TALKS AND LECTURES

- ▶ *Simons Institute Berkeley: Workshop on Uncertainty, Optimization, Statistics*, Dec 01 2017
- ▶ *Simons Institute Berkeley: Workshop on Fast Iterative Methods in Optimization*, Oct 05 2017
- ▶ *Applied Math Summer School, Peking Univ., Beijing, China*, Jul 10-15 2017
- ▶ *Machine Learning Summer School (MLSS'15), Tübingen, Germany*, Jun 25-27 2017
- ▶ *Bayopt Conference, Data Science and Optimization, UC Davis, Davis, CA*, May 12 2017
- ▶ *BLISS Seminar, UC Berkeley*, Apr 24 2017
- ▶ *Optimization and Statistical Learning (OSL) Workshop, Les Houches, France*, Apr 11 2017
- ▶ *NIPS Workshop on Nonconvex Optimization*, Dec 09 2016
- ▶ *Data Science Seminar, Northeastern University*, Nov 04 2016
- ▶ *LIDS Seminar Series, MIT: Geometric optimization*, Sep 13 2016
- ▶ *SIAM Annual Meeting '16, Boston: Geometric (nonconvex) optimization*, Jul 11 2016
- ▶ *Advances in Non-convex analysis and Optimization, Workshop, ICML 2016*, Jun 23 2016
- ▶ *Google Machine Learning Workshop, New York City, NY*, Mar 03 2016
- ▶ *MIT Lincoln Laboratory, PED Seminar, Lexington, MA*, Feb 09 2016
- ▶ *Lectures on large-scale optimization, Hausdorff Institute of Mathematics, Bonn*, Jan 2016
- ▶ *Computer Vision Seminar, Boston University (BU), Boston, MA*, Sep 2015

▶ <i>Machine Learning Summer School Lecturer (Microsoft Research India)</i>	June 2015
▶ <i>New England Machine Learning Conference (NEML)</i>	May 18 2015
▶ <i>Massachusetts Institute of Technology (MIT), Math Dept., Combinatorics Seminar</i>	Mar 13 2015
▶ <i>Invited talk: "Riemannian geometry in machine learning", Workshop@NIPS 2014, Montreal</i>	Dec 05 2014
▶ <i>University of California San Diego, CSE Dept., San Diego, CA</i>	May 29 2014
▶ <i>Massachusetts Institute of Technology (MIT), LIDS, Cambridge, MA</i>	May 8 2014
▶ <i>Cornell University, ORIE Seminar, Ithaca, NY</i>	May 5 2014
▶ <i>Invited short-course, EU Regional School 2014, RWTH Aachen, Germany</i>	Apr 11 2014
▶ <i>Microsoft Research, Cambridge, UK</i>	Mar 17 2014
▶ <i>University of Cambridge, Engineering Dept., Cambridge, UK</i>	Mar 17 2014
▶ <i>University of British Columbia, Computer Science Dept.; Vancouver</i>	Feb 26 2014
▶ <i>Carnegie Mellon University, Tepper School of Business; Pittsburgh</i>	Feb 21 2014
▶ <i>Carnegie Mellon University, Machine Learning Dept.; Pittsburgh, PA</i>	Feb 19 2014
▶ <i>Carnegie Mellon University, Statistics Department Seminar, Pittsburgh, PA</i>	Nov 11 2013
▶ <i>Massachusetts Institute of Technology (MIT), LIDS, Cambridge, MA</i>	Oct 18 2013
▶ <i>Washington University at St. Louis, CSE Dept. Colloquium, St. Louis, MO</i>	Oct 15 2013
▶ <i>École polytechnique fédérale de Lausanne, Math Department, Lausanne</i>	Sep 18 2013
▶ <i>Technische Universität, München, Math Department.</i>	Sep 17, 2013
▶ <i>École polytechnique fédérale de Lausanne, School of Computer Sci., Lausanne</i>	Apr 15 2013
▶ <i>University of California Davis, Algebra and Discrete Math Seminar, Davis, CA</i>	Mar 14 2013
▶ <i>Carnegie Mellon University, Machine Learning Department, Pittsburgh</i>	Feb 28 2013
▶ <i>University College London, Gatsby Computational Neuroscience Unit, London</i>	Feb 18 2013
▶ <i>Cornell University, Computer Science Colloquium, Ithaca</i>	Feb 14 2013
▶ <i>Duke University, ECE Department Colloquium</i>	Feb 11 2013
▶ <i>INRIA Rocquencourt (ENS/CNRS/ENS joint laboratory), Paris</i>	Jan 21, 2013
▶ <i>Oxford University, The Mathematical Institute, Oxford</i>	Jan 09, 2013
▶ <i>University of British Columbia (UBC), ECE Dept., Vancouver.</i>	Oct 10, 2012
▶ <i>Kyungpook National Univ., Dept. of Mathematics, South Korea (5 day workshop)</i>	Jun 26–30, 2012
▶ <i>Google Research, Mountain View, CA</i>	May 24, 2012
▶ <i>IBM Research, New York, NY</i>	May 8, 2012
▶ <i>University of Massachusetts Amherst, ECE Dept., Amherst, MA</i>	Apr 10, 2012
▶ <i>Yahoo! Research Tech Talk, Sunnyvale, CA.</i>	Mar 9, 2012
▶ <i>LinkedIn Tech Talk, Mountain View, CA.</i>	Mar 6, 2012
▶ <i>Microsoft Research Tech Talk, Redmond, WA</i>	Aug 29, 2011
▶ <i>University of Washington, Seattle, EE, CS, Math Seminar, Seattle</i>	Aug 25, 2011
▶ <i>Yahoo Research, Sunnyvale, CA</i>	Aug 18, 2011
▶ <i>Toyota Technological Institute (TTI) Chicago.</i>	Aug 16, 2011
▶ <i>University of Chicago, Statistics Colloquium, Chicago</i>	Aug 15, 2011
▶ <i>Householder Symposium XVIII, Tahoe</i>	Jun 2011
▶ <i>ETH Zürich, Informatik Seminar, Zürich</i>	Apr 2011
▶ <i>Universidad Autónoma de Madrid, CS Dept, Madrid</i>	Apr 2011
▶ <i>Institute for Pure and Applied Mathematics, UCLA, Applications of Optimization in Science and Engineering workshop</i>	Dec 2010

▶ <i>First I.S.T. Austria Symposium on Computer Vision and Machine Learning, Klosterneuburg</i>	Oct 2010
▶ <i>Invited short-course at the EU Regional School 2010, RWTH Aachen, Germany.</i>	Apr 2010
▶ <i>Microsoft Research Tech Talk, Redmond.</i>	Dec 2008
▶ <i>Google, Mountain View, CA.</i>	Mar 2007
▶ <i>Yahoo Inc., Sunnyvale, CA.</i>	Mar 2007
▶ <i>SAGE Days 3, IPAM, UCLA</i>	Feb 2007

PROFESSIONAL SERVICE

ASSOCIATE EDITOR: Optimization Methods and Software (OMS); 2014–
SENIOR PROGRAM COMMITTEE / AREA CHAIR

1. Neural Information Processing Systems (NIPS), 2017, 2016, 2014, 2013
2. International Conference on Machine Learning (ICML) 2016, 2015, 2014
3. Association for the Advancement of Artificial Intelligence (AAAI) 2018
4. Artificial Intelligence & Statistics (AISTATS) 2016, 2014

SPONSORSHIP CHAIR: Artificial Intelligence and Statistics (AISTATS) 2012

MINISYMPOSIUM ORGANIZER: “Advances in large-scale optimization”

Invited minisymposium at *SIAM Annual Meeting, 2016*

WORKSHOP ORGANIZER AND CO-CHAIR: OPT 2008—OPT 2017

International workshops held at the *Neural Information Processing Systems (NIPS)* Conference

PROGRAM COMMITTEE MEMBER / REVIEWER

- *Conference on Learning Theory (COLT)* (2018, 2016, 2011)
- *ACM SIGMOD* 2016
- *SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD)* (2016, 2015, 2011, 2003–2007)
- *Uncertainty in Artificial Intelligence (UAI)* 2016, 2009
- *International Joint Conf. Artificial Intelligence (IJCAI)* 2013
- *Artificial Intelligence and Statistics (AISTATS)* 2011, 2013
- *Neural Information Processing Systems (NIPS)* (2005–2010)
- *International Conference on Machine Learning (ICML)* (2006–2010)
- *Snowbird Abstracts* (2008–2011)
- *IEEE Symposium on Foundations of Computer Science (FOCS)* (2009)
- *IEEE Int. Conference on Semantic Computing (ICSC)* (2008)
- *IEEE Int. Conference Data Mining (ICDM)* (2003–2007)
- *ACM Conf. on Information and Knowledge Management (CIKM)* (2005)
- *SIAM Int. Conference on Data Mining (SDM)* (2003–2008)

JOURNAL REVIEWING

1. *SIAM Review (SIREV)*
2. *SIAM J. on Optimization (SIOPT)*
3. *SIAM J. of Scientific Computing (SISC)*
4. *SIAM J. Control and Optimization (SICON)*
5. *SIAM J. Matrix Analysis and Applications (SIMAX)*
6. *SIAM J. on Imaging Sciences (SIIMS)*
7. *Machine Learning Journal (Springer)*
8. *Journal of Machine Learning Research (JMLR)*
9. *Annals of Statistics (AoS)*
10. *Annals of Applied Statistics (AoAS)*
11. *Mathematical Programming (MAPR)*

-
12. *IEEE Special Topics in Signal Processing (JTSP)*
 13. *IEEE Signal Processing Letters (SPL)*
 14. *IEEE Transactions on Information Theory (IT)*
 15. *IEEE Transactions on Signal Processing (TSP)*
 16. *IEEE Journal Selected Topics in Signal Processing (JTSP)*
 17. *IEEE Transactions on Image Processing (TIP)*
 18. *IEEE Transactions Knowledge and Data Engineering (TKDE)*
 19. *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*
 20. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*
 21. *Statistics and Computing (STCO)*
 22. *Computational Statistics and Data Analysis (CSDA)*
 23. *Computational Statistics (COST)*
 24. *Entropy*
 25. *Journal of Multivariate Analysis (JMVA)*
 26. *Journal of Statistical Computation and Simulation (JSCS)*
 27. *Linear Algebra and its Applications (LAA)*
 28. *Optimization (Taylor & Francis)*
 29. *Optimization and Engineering (OPTE)*
 30. *Optimization Letters (Springer)*
 31. *Signal Processing (Elsevier)*
 32. *Numerical Algorithms (Springer)*
 33. *Pattern Recognition (PR)*
 34. *Reviewer for [MATHSCI](#)NET*
 35. *American Mathematical Monthly*

OTHER SERVICE

Graduate admissions, EECS department, MIT
Reviewing Postdoc, PhD, and Master's applications at Max-Planck Institute
Interviewing Postdoc and PhD candidates at Max Planck Institute
ACM and SCS Doctoral Dissertation Committee, Carnegie Mellon University

STUDENTS

- ▶ *Current graduate students being advised*
 - Zelda Mariet (PhD student, CSAIL, MIT)
 - Hongyi Zhang (PhD student, BCS, MIT)
 - Chengtao Li (PhD student CSAIL, MIT); co-supervised with Stefanie Jegelka
 - Chulhee Yun (PhD student, EECS); co-supervised with Ali Jadbabaie
 - Jingzhao Zhang (PhD student, EECS); co-supervised with Ali Jadbabaie
- ▶ *PhD Committee member for*
 - Álvaro Barbero (UAM Madrid; July 2011)
 - Aaron Defazio (external reviewer), 2015 (Australian National University, ANU)
 - Qinxun Bai (CS Dept., Boston University), 2016
 - Sashank Reddi (ML Dept., Carnegie Mellon Univ.), 2017
 - Rasul Tutunov (CISE Dept., Univ. of Pennsylvania), 2017
 - Haihao Lu (Operations Research Center (ORC), MIT), 2017
- ▶ *Graduate students; supervised for ≥ 1 research projects:*
 - Sashank Reddi (2013–17, PhD student CMU; now Google Research NY)

- Pourya Habib'zadeh (2016, MS student, Univ. Tehran, Iran)
- Yu-Xiang Wang (2013–14, PhD student CMU; now Assitant Prof. UC Santa Barbara)
- Adams Wei Yu (2013-14, MS student CMU, now PhD student CMU)
- Ahmed Hefny (2013-15, PhD student CMU)
- Veeranjaneyulu Sadhanala (2013-14, PhD student CMU)
- Anoop Cherian (2011–2013, PhD student UMN, now Researcher ANU, Australia)
- Álvaro Barbero (2010–2011, PhD student UAM Madrid, Chief Data Scientist, IIC, UAM Madrid)
- Namhyoung Kim (2010, PhD student POSTECH, Korea, now Assistant Professor)
- Carlos María Gudín (2011–2012, PhD student UAM Madrid, now Postdoc KU Leuven)
- Radha Chitta (2013, Research Scientist PARC)
- Michael Hirsch (2008–2010) (PhD student Max-Planck, Tübingen, now Group Leader Max-Planck)

► *Undergraduates supervised for research projects:*

- Haochen Zhang (Spring 2018; Intern@MIT; Tsinghua University)
- Rajeev Parvathala (Fall 2017–Spring 2018, M.Eng; EECS MIT)
- Flora Tan (Fall 2015–Spring 2016, SuperUROP; EECS MIT)
- Lingfu Zhang (Fall 2015–Spring 2016, SuperUROP; EECS/Math MIT)
- Hayley Song (Fall 2015–Spring 2016, SuperUROP; EECS MIT)
- Alexander Amini (Summer 2015, UROP; EECS MIT)
- Rashish Tandon (2010, Max-Planck; now PhD student in CS, UT Austin)
- Samaneh Azadi (2013, UC Berkeley, now PhD student in EECS, UC Berkeley)

PATENTS

Method and device for recovering a digital image from a sequence of observed digital images

S. Harmeling, M. Hirsch, S. Sra, B. Schölkopf, and C. Schuler

International Patent: WO/2012/041492; 5th April, 2012.

Combining spectral and probabilistic clustering

A. C. Surendran and S. Sra

United States Patent: US 7,809,704 B2

SOFTWARE WRITTEN

MYSVD, EIGS: Sparse singular and eigenvector decomposition in C++

NMA: Optimized implementations of various nonnegative matrix factorization algorithms

SSLIB: Sparse matrix manipulation library (C/C++)

FSOLVER: Optimization software for large-scale linear and quadratic programs

Other software for clustering, co-clustering, Bessel functions, etc., available from my webpage

PERSONAL

Native to near-native proficiency: English, Hindi, Punjabi

Fluent: German; Intermediate: Italian, Urdu; Beginner: French, Russian.