RESEARCH SUMMARY

My research focuses on "*Mathematics of AI, Optimization, and AI driven Decision Making*." A major part of work focuses on optimization and algorithm design for **deep learning**, as well as for problems having **geometric structure**. Current topics of interest to me include: The Theory of Transformers and LLMs, AI in real-world decision making, and quantum computation. I am currently developing foundation models for AI-driven supply chains, to target: **social good, fairness, sustainability, resource efficiency, and robustness.**

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 Institute of Technology & Science, Pilani, India, Jun. 1999

POSITIONS

Alexander von Humboldt Professor of AI and **Full Professor of Mathematics**,
Mathematics Department, Technical University of Munich, Germany.   Oct 2023–now

Esther and Harold E. Edgerton (1927) Career Development Associate Professor,
Electrical Engineering and Computer Science (EECS) and Institute for Data Systems and Society (IDSS),
Massachusetts Institute of Technology (MIT).   Jan 2019–now

Chief Scientist and Co-founder,
Pendulum Systems (OPT+ML startup)   Jan 2014–now

Assistant Professor,
Electrical Engineering and Computer Science (EECS) and Institute for Data Systems and Society (IDSS),
Massachusetts Institute of Technology (MIT).   Jan 2018–Jan 2019

Principal Research Scientist (Research Professor),
Laboratory for Information & Decision Systems (LIDS), MIT.   Jan 2015–Dec 2017

Visiting Assistant Professor,
Machine Learning Department, Carnegie Mellon University, Pittsburgh.   Sep 2013–May 2014

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

Senior Research Scientist,
(Schölkopf Group), Max Planck Institute for Intelligent Systems, Tübingen.   Jul 2012–Aug 2014

Research Scientist,
(Schölkopf group), Max Planck Institute for Intelligent Systems Tübingen   Oct. 2007–Jun. 2012

Visiting Researcher,
(host: Prof. Jeff Bilmes), EE Department, University of Washington, Seattle.   Jul.–Aug. 2011

Graduate Research Assistant,
(advisor: Prof. I. S. Dhillon), University of Texas at Austin.   Jan 2002–Aug 2007

Research Intern,
(supervisor: John Platt), Microsoft Research, Redmond.   May 2005–Aug 2005

HONORS and AWARDS

- **Alexander von Humboldt Professorship** (EUR 5M award), AvH foundation, Germany, 2022.
- **NSF-CAREER Award** (Jan 2019).
- **Amazon Research Award** (Oct 2017)
- **Criteo Faculty Research Award** (Sep 2017).
- **NSF-BIGDATA Award** (Sep 2017).
- **Invited lecturer** on *"Optimization for ML"*
  at the Machine Learning Summer School (MLSS), Tübingen, Germany, Jun. 2017.
- **Elsevier Mathematical Sciences Sponsorship Award**, 2016.
- **SIAM Outstanding Paper Prize**, Jul. 2011
  One of three best papers among all papers published in SIAM journals between 2008–2010.
- **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

INVITED LONG-TERM RESEARCH VISITS

- **Visiting Professor**, Dept. of Computer Science, ETH Zürich, Jun 2019–Aug 2019
- **Invited long-term participant**, Simons Institute, UC Berkeley, Jan 2017–May 2017
  in the program *Foundations of Machine Learning.*
- **Invited long-term participant**, Hausdorff Institute for Mathematics, Bonn, Jan 2016–Mar 2016
  in the program *Mathematics of Signal Processing.*

INVITED TUTORIALS

- **Optimization for Machine Learning**, (Mar 2024)
  Invited lectures at the: Machine Learning Summer School (MLSS), Okinawa.
- **Modern optimization for machine learning**, (Aug 2019)
  Invited tutorial at *CPS-FR 2019, 1st Computational Physics School for Fusion Research, MIT*
- **Negative dependence, stable polynomials, and all that**, (Dec 2018)
  Tutorial at the *Neural Information Processing Systems* conference
  (co-taught with Stefanie Jegelka; 500+ audience (estimated).)
- **Nonconvex optimization in Machine Learning**, (Jun 2018)
  Invited lectures at the: DS3: Data Science Summer School organized by École Polytechnique and the
  DATAIA Institute, Paris.

- **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 for Machine Learning**, (Jun 2017)
  Invited lectures at the: Machine Learning Summer School (MLSS), Tübingen
- **Modern Stochastic Optimization in Machine Learning** (Dec 2016)
  Tutorial at the *Neural Information Processing Systems* 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)
  Invited 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.

TEACHING

- **Fundamentals of Optimization for Machine Learning** (Apr–July'24)
  Graduate course, Math Department, TU Munich.
  (**new graduate course developed by me**)
- **Introduction to the theory of Transformers and LLMs** (Apr–July'24)
  Graduate seminar course, Math Department, TU Munich.
  (**new graduate course developed by me**)
- **Foundations of Mathematics for Artificial Intelligence** (June 2024)
  MIT Professional Education Course
  (**new professional course developed by me**)
- **Introduction to Data Science** (Spring 2023)
  Course IDS.131, 6.3730, Capstone course (Undergrad & Graduate level), IDSS & EECS, MIT.
  (co-taught with Stefanie Jegelka, Navid Azizan)
- **Nonlinear Programming** (Spring 2022)
  Course: 6.252, Graduate (MS & PhD) level course, EECS Department, MIT
  (foundational course on nonlinear programming; co-taught with Mardavij Roozbehani)
- **Optimization for Machine Learning** (Spring 2020, 2021)
  Course: 6.881, Graduate (MS & PhD) level course, EECS Department, MIT
  (**new graduate course developed by me**)
- **Introduction to Machine Learning** (Fall 2018, Fall 2019, Fall 2021, Fall 2022)
  Course: 6.867, 6.7900 Graduate (MS & PhD) level course, EECS Department, MIT
  (co-teaching with Devavrat Shah and David Sontag; a new version of the course)
- **Modeling and Optimization for Machine Learning** (Summer 2018, 2019, 2020)
  MIT Professional Education Course
  (co-developed and taught with Justin Solomon)
- **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*
  (created and co-taught with Alex Smola)
- **Convex optimization; EE 227A** (Spring 2013) (by invitation)
  Graduate course, *EECS Department, University of California, Berkeley*
  (created entire graduate level course from scratch)
- **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

PUBLICATIONS (Please see Google Scholar)

1. In machine learning and optimization for ML, the main conferences are highly competitive and are the main venues of publication. They are equivalent to top journal publications in other fields.
2. A star (e.g., A. Student⋆) signifies student / postdoc advised by me on that paper.
3. Google scholar: *https://scholar.google.com/citations?user=eyCw9goAAAAJ*.
4. Citations: **18402**; H-Index: **68**; I10-Index: **151**   (retrieved: 06 July 2024)

Books and Monographs

1. 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.

Publications at MIT (Jan 2015–now)

102. Kwangjun Ahn⋆, Xiang Cheng⋆, Hadi Daneshmand⋆, and **Suvrit Sra**. "Transformers learn to implement preconditioned gradient descent for in-context learning". *Advances in Neural Information Processing Systems*, 36, 2024. 37 pages

101. Kwangjun Ahn⋆, Xiang Cheng⋆, Minhak Song, Chulhee Yun, Ali Jadbabaie, and **Suvrit Sra**. "Linear attention is (maybe) all you need (to understand transformer optimization)". In *International Conference on Learning Representations (ICLR)*, 2024. 14 pages

100. Xiang Cheng⋆, Yuxin Chen, and **Suvrit Sra**. "Transformers implement functional gradient descent to learn non-linear functions in context". In *International Conference on Machine Learning*, 2024. *To appear.* 37 pages.

99. Yan Dai⋆, Kwangjun Ahn⋆, and **Suvrit Sra**. "The crucial role of normalization in sharpness-aware minimization". *Advances in Neural Information Processing Systems*, 36, 2024. 30 pages

98. Derek Lim, Joshua Robinson⋆, Lingxiao Zhao, Tess Smidt, **Suvrit Sra**, Haggai Maron, and Stefanie Jegelka. "Sign and Basis Invariant Networks for Spectral Graph Representation Learning". In *International Conference on Learning Representations (ICLR 23)*, 2023. 23 pages

97. Constantin P Niculescu and **Sra, Suvrit**. "Functions with Positive Differences on Convex Cones". *Results in Mathematics*, 78(6):217, 2023

96. Yi Tian, Kaiqing Zhang, Russ Tedrake, and **Sra, Suvrit**. "Can Direct Latent Model Learning Solve Linear Quadratic Gaussian Control?" In *Learning for Dynamics and Control Conference*, pages 51–63. PMLR, 2023

95. Yi Tian⋆, Kaiqing Zhang, Russ Tedrake, and **Suvrit Sra**. "Toward Understanding State Representation Learning in MuZero: A Case Study in Linear Quadratic Gaussian Control". In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pages 6166–6171. IEEE, 2023

94. Melanie Weber and **Suvrit Sra**. "Global optimality for Euclidean CCCP under Riemannian convexity". In *International Conference on Machine Learning*, pages 36790–36803. PMLR, 2023

93. Melanie Weber and **Suvrit Sra**. "Riemannian optimization via Frank–Wolfe methods". *Mathematical Programming*, 199:525–556, 2023

92. David Xing Wu⋆, Chulhee Yun, and **Suvrit Sra**. "On the training instability of shuffling SGD with batch normalization". In *International Conference on Machine Learning*, pages 37787–37845. PMLR, 2023

91. Peiyuan Zhang⋆, Jingzhao Zhang, and Suvrit Sra. "Sion's minimax theorem in geodesic metric spaces and a Riemannian extragradient algorithm". *SIAM Journal on Optimization*, 33(4):2885–2908, 2023

90. Kwangjun Ahn and Suvrit Sra. "Understanding Nesterov's Acceleration via Proximal Point Method". In *Symposium on Simplicity in Algorithms (SOSA)*, pages 117–130. SIAM, 2022

89. Kwangjun Ahn⋆, Jingzhao Zhang, and **Suvrit Sra**. "Understanding the unstable convergence of gradient descent". In *International Conference on Machine Learning*, pages 247–257. PMLR, 2022

88. Xiang Cheng⋆, Jingzhao Zhang, and **Suvrit Sra**. "Efficient sampling on Riemannian manifolds via Langevin MCMC". *Advances in Neural Information Processing Systems*, 35:5995–6006, 2022. 73 pages

87. Jikai Jin⋆ and **Suvrit Sra**. "Understanding Riemannian acceleration via a proximal extragradient framework". In *Conference on Learning Theory*, pages 2924–2962. PMLR, 2022

86. Horia Mania⋆, Ali Jadbabaie, Devavrat Shah, and **Suvrit Sra**. "Time varying regression with hidden linear dynamics". In *Learning for Dynamics and Control Conference*, pages 858–869. PMLR, 2022

85. Anshul Shah, **Suvrit Sra**, Rama Chellappa, and Anoop Cherian. "Max-margin contrastive learning". In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 8220–8230, 2022

84. Melanie Weber and **Suvrit Sra**. "Projection-free nonconvex stochastic optimization on Riemannian manifolds". *IMA Journal of Numerical Analysis*, 42(4):3241–3271, 2022

83. Chulhee Yun⋆, Shashank Rajput⋆, and **Suvrit Sra**. "Minibatch vs local SGD with shuffling: Tight convergence bounds and beyond". In *International Conference on Learning Representations (ICLR 22)*, 2022. 76 pages

82. Alp Yurtsever and **Suvrit Sra**. "CCCP is Frank-Wolfe in disguise". *Advances in Neural Information Processing Systems*, 35:35352–35364, 2022. 18 pages

81. Pourya Habib Zadeh⋆ and **Suvrit Sra**. "Introducing discrepancy values of matrices with application to bounding norms of commutators". *Linear Algebra and its Applications*, 651:359–384, 2022

80. Jingzhao Zhang⋆, Haochuan Li, **Suvrit Sra**, and Ali Jadbabaie. "Neural network weights do not converge to stationary points: An invariant measure perspective". In *International Conference on Machine Learning*, pages 26330–26346. PMLR, 2022

79. Jingzhao Zhang⋆, Hongzhou Lin, Subhro Das, **Suvrit Sra**, and Ali Jadbabaie. "Beyond Worst-Case Analysis in Stochastic Approximation: Moment Estimation Improves Instance Complexity". In *International Conference on Machine Learning*, pages 26347–26361. PMLR, 2022

78. Joshua Robinson⋆, Li Sun, Ke Yu, Kayhan Batmanghelich, Stefanie Jegelka, and **Suvrit Sra**. "Can contrastive learning avoid shortcut solutions?" *Advances in neural information processing systems*, 34:4974–4986, 2021

77. **Suvrit Sra**. "Metrics induced by Jensen-Shannon and related divergences on positive definite matrices". *Linear Algebra and its Applications (LAA)*, 616:125–138, 2021

76. Yi Tian⋆, Yuanhao Wang, Tiancheng Yu⋆, and **Suvrit Sra**. "Online learning in unknown markov games". In *International conference on machine learning*, pages 10279–10288. PMLR, 2021

75. Chulhee Yun, **Suvrit Sra**, and Ali Jadbabaie. "Open problem: Can single-shuffle SGD be better than reshuffling SGD and GD?" In *Conference on Learning Theory*, pages 4653–4658. PMLR, 2021

74. Alp Yurtsever⋆, Alex Gu⋆, and **Suvrit Sra**. "Three operator splitting with subgradients, stochastic gradients, and adaptive learning rates". *Advances in Neural Information Processing Systems*, 34:19743–19756, 2021

73. Alp Yurtsever⋆, Varun Mangalick⋆, and **Suvrit Sra**. "Three operator splitting with a nonconvex loss function". In *International Conference on Machine Learning*, pages 12267–12277. PMLR, 2021

72. Tiancheng Yu⋆, Yi Tian⋆, Jingzhao Zhang⋆, and **Suvrit Sra**. "Provably efficient algorithms for multi-objective competitive RL". In *International Conference on Machine Learning*, pages 12167–12176. PMLR, 2021

71. Jingzhao Zhang⋆, Aditya Krishna Menon, Andreas Veit, Srinadh Bhojanapalli, Sanjiv Kumar, and **Suvrit Sra**. "Coping with Label Shift via Distributionally Robust Optimisation". In *International Conference on Learning Representations (ICLR)*, 2021

70. Kwangjun Ahn⋆ and **Suvrit Sra**. "From Nesterov's Estimate Sequence to Riemannian Acceleration". In *Conference on Learning Theory (COLT)*, 2020

69. Kwangjun Ahn⋆, Chulhee Yun⋆, and **Suvrit Sra**. "SGD with shuffling: optimal rates without component convexity and large epoch requirements". In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020

68. R Hariharan, J Sundberg, G Gallino, A Schmidt, D Arenth, **Sra, S**, and B Fels. "An Interpretable Predictive Model of Vaccine Utilization for Tanzania". *Frontiers of Artificial Intelligence*, 3:559617, 2020

67. Reshad Hosseini and **Suvrit Sra**. "Recent Advances in Stochastic Riemannian Optimization". In *Handbook of Variational Methods for Nonlinear Geometric Data*, pages 527–554. Springer International Publishing, 2020

66. Chi Jin, Tiancheng Jin, Haipeng Luo, **Suvrit Sra**, and Tiancheng Yu⋆. "Learning Adversarial MDPs with Bandit Feedback and Unknown Transition". In *International Conference on Machine Learning (ICML)*, 2020. arXiv:1912.01192

65. Joshua Robinson⋆, Stefanie Jegelka, and **Suvrit Sra**. "Strength from Weakness: Fast Learning Using Weak Supervision". In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pages 8127–8136. PMLR, 13–18 Jul 2020

64. Yi Tian⋆, Jian Qian, and **Suvrit Sra**. "Towards Minimax Optimal Reinforcement Learning in Factored Markov Decision Processes". In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020

63. Florian Yger, Sylvain Chevallier, Quentin Barthélemy, and **Sra, Suvrit**. "Geodesically-convex optimization for averaging partially observed covariance matrices". In *Proceedings of The 12th Asian Conference on Machine Learning*, volume 129, pages 417–432. PMLR, 18–20 Nov 2020

62. Jingzhao Zhang⋆, Tianxing He, **Suvrit Sra**, and Ali Jadbabaie. "Why Gradient Clipping Accelerates Training: A Theoretical Justification for Adaptivity". In *International Conference on Learning Representations (ICLR)*, 2020. (**Oral presentation**)

61. Jingzhao Zhang⋆, Sai Praneeth Karimireddy, Andreas Veit, Seungyeon Kim, Sashank J Reddi, Sanjiv Kumar, and **Suvrit Sra**. "Why are Adaptive Methods Good for Attention Models?" In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020

60. Jingzhao Zhang★, Hongzhou Lin, Stefanie Jegelka, **Suvrit Sra**, and Ali Jadbabaie. "Complexity of Finding Stationary Points of Nonconvex Nonsmooth Functions". In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pages 11173–11182. PMLR, 13–18 Jul 2020

59. Jeffery Z. HaoChen★ and **Suvrit Sra**. "Randomshuffle beats SGD after finite epochs". In *International Conference on Machine Learning (ICML)*, 2019

58. Reshad Hosseini and **Suvrit Sra**. "An Alternative to EM for Gaussian Mixture Models: Batch and Stochastic Riemannian Optimization". *Mathematical Programming, Ser. A*, pages 1–37, 2019

57. Zelda Mariet★, Mike Gartrell, and **Suvrit Sra**. "Learning Determinantal Point Processes by Corrective Negative Sampling". In *Artificial Intelligence and Statistics (AISTATS)*, volume 89 of *Proceedings of Machine Learning Research*, pages 2251–2260. PMLR, 16–18 Apr 2019

56. Joshua Robinson★, Suvrit Sra, and Stefanie Jegelka. "Flexible modeling of diversity with strongly log-concave distributions". In *Advances in Neural Information Processing Systems (NeuRIPS)*, pages 15225–15235, 2019

55. Matthew Staib★, Sashank Reddi, Satyen Kale, Sanjiv Kumar, and **Suvrit Sra**. "Escaping Saddle Points with Adaptive Gradient Methods". In *International Conference on Machine Learning (ICML)*, pages 5956–5965, 2019

54. **Suvrit Sra**. "Logarithmic inequalities under a symmetric polynomial dominance order". *Proceedings American Mathematical Society (PAMS)*, 142(2):481–486, 2019

53. Chulhee Yun★, **Suvrit Sra**, and Ali Jadbabaie. "Are deep ResNets provably better than linear predictors?" In *Advances in Neural Information Processing Systems (NeuRIPS)*, 2019. **Spotlight presentation**

52. Chulhee Yun★, **Suvrit Sra**, and Ali Jadbabaie. "Efficiently testing local optimality and escaping saddles for ReLU networks". In *International Conference on Learning Representations (ICLR)*, 2019

51. Chulhee Yun★, **Suvrit Sra**, and Ali Jadbabaie. "Small nonlinearities in activation functions create bad local minima in neural networks". In *International Conference on Learning Representations (ICLR)*, 2019

50. Chulhee Yun★, **Suvrit Sra**, and Ali Jadbabaie. "Small ReLU networks are powerful memorizers: a tight analysis of memorization capacity". In *Advances in Neural Information Processing Systems (NeuRIPS)*, 2019. **Spotlight presentation**

49. Alp Yurtsever★, **Suvrit Sra**, and Volkan Cevher. "Conditional Gradient Methods via Stochastic Path-Integrated Differential Estimator". In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 7282–7291. PMLR, 09–15 Jun 2019

48. Jingzhao Zhang★, **Suvrit Sra**, and Ali Jadbabaie. "Acceleration in First Order Quasi-strongly Convex Optimization by ODE Discretization". In *IEEE CDC Conference*, 2019. arXiv:1905.12436

47. Jingzhao Zhang★, Aryan Mokhtari, **Suvrit Sra**, and Ali Jadbabaie. "Direct Runge-Kutta discretization achieves acceleration". In *Advances in Neural Information Processing Systems (NIPS)*, Dec. 2018. **Spotlight**

46. Hongyi Zhang★ and **Suvrit Sra**. "An Estimate Sequence for Geodesically Convex Optimization". In *Conference on Learning Theory (COLT)*, Jun. 2018

45. Sashank Reddi★, Manzil Zaheer, **Suvrit Sra**, Francis Bach, Barnabas Poczos, Ruslan Salakhutdinov, and Alexander J. Smola. "A Generic Approach for Escaping Saddle points". In *Artificial Intelligence and Statistics (AISTATS)*, May 2018. *arXiv:1709.01434*

44. Anoop Cherian, **Suvrit Sra**, Stephen Gould, and Richard Hartley. "Non-Linear Temporal Subspace Representations for Activity Recognition". In *CVPR*, Apr. 2018

43. Chulhee Yun★, **Suvrit Sra**, and Ali Jadbabaie. "Global optimality conditions for deep neural networks". In *International Conference on Learning Representations (ICLR)*, Apr 2018

42. Álvaro J. Barbero and **Suvrit Sra**. "Modular proximal optimization with application to total variation regularization". *Journal of Machine Learning Research (JMLR)*, 19(1):2232–2313, 2018

41. Zelda Mariet⋆, **Suvrit Sra**, and Stefanie Jegelka. "Exponentiated Strongly-Rayleigh distributions". In *Advances in Neural Information Processing Systems (NeuRIPS)*, 2018

40. **Suvrit Sra**. "New concavity and convexity results for symmetric polynomials and their ratios". *Linear and Multilinear Algebra*, 68(5):1031–1038, 2018

39. **Suvrit Sra**, Nisheeth K. Vishnoi, and Ozan Yıldız. "On Geodesically Convex Formulations for the Brascamp-Lieb Constant". In *APPROX*, 2018

38. Chulhee Yun⋆, **Suvrit Sra**, and Ali Jadbabaie. "Global optimality conditions for deep neural networks". In *International Conference on Representation Learning (ICLR)*, 2018

37. Chengtao Li⋆, Stefanie Jegelka, and **Suvrit Sra**. "Polynomial time algorithms for dual volume sampling". In *Advances in Neural Information Processing Systems (NIPS)*, Dec 2017

36. Zelda Mariet⋆ and **Suvrit Sra**. "Elementary Symmetric Polynomials for Optimal Experimental Design". In *Advances in Neural Information Processing Systems (NIPS)*, Dec 2017

35. Ke Jiang, **Suvrit Sra**, and Brian Kulis. "Combinatorial topic modling using small variance asymptotics". In *Artificial Intelligence and Statistics (AISTATS)*, Apr 2017

34. **Suvrit Sra**. "Directional Statistics in Machine Learning". In C. Ley and T. Verdebout, editors, *Modern Directional Statistics*. Chapman and Hall, 2017. *Invited chapter*

33. 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)*, Dec 2016

32. Zelda Mariet⋆ and **Suvrit Sra**. "Kronecker Determinantal Point Processes". In *Advances in Neural Information Processing Systems (NIPS)*, Dec 2016

31. Sashank Reddi⋆, **Suvrit Sra**, Barnabas Poczos, and Alexander J. Smola. "Fast incremental methods for smooth nonconvex optimization". In *IEEE Conf. Decision and Control (CDC)*, Dec 2016

30. 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)*, Dec 2016

29. Hongyi Zhang⋆, Sashank Reddi⋆, and **Suvrit Sra**. "Riemannian SVRG: Fast Stochastic Optimization on Riemannian Manifolds". In *Advances in Neural Information Processing Systems (NIPS)*, Dec 2016

28. 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*, Oct 2016

27. 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

26. Chengtao Li⋆, **Suvrit Sra**, and Stefanie Jegelka. "Gaussian quadrature for matrix inverse forms with applications". In *International Conference on Machine Learning (ICML)*, Jun 2016

25. 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

24. Yu-Xiang Wang⋆, Veeranjaneyulu 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

23. Pourya H. Zadeh⋆, Reshad Hosseini, and **Suvrit Sra**. "Geometric Mean Metric Learning". In *International Conference on Machine Learning (ICML)*, Jun 2016

22. Hongyi Zhang⋆ and **Suvrit Sra**. "First-order methods for geodesically convex optimization". In *Conference on Learning Theory (COLT)*, Jun 2016. arXiv:1602.06053

21. Zelda Mariet⋆ and **Suvrit Sra**. "Diversity Networks". In *International Conference on Learning Representations (ICLR)*, May 2016. arXiv:1511.0577

20. **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

19. Chengtao Li⋆, Stefanie Jegelka, and **Suvrit Sra**. "Efficient Sampling for K-Determinantal Point Processes". In *Artificial Intelligence and Statistics (AISTATS'16)*, Apr 2016

18. 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

17. 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

16. 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

15. Justin Solomon, Gabriel Peyré, Vladimir Kim, and **Suvrit Sra**. "Entropic Metric Alignment for Correspondence Problems". *ACM Transactions on Graphics (ToG)*, 35(4), 2016

14. **Suvrit Sra** and Reshad Hosseini. "Geometric Optimization in Machine Learning". In *Algorithmic Advances in Riemannian Geometry and Applications*. Springer, 2016

13. **Suvrit Sra**. "On inequalities for normalized Schur functions". *European J. Combinatorics*, Volume 51:492–494, 2016

12. **Suvrit Sra**. "On the matrix square root and geometric optimization". *Electronic Journal on Linear Algebra (ELA)*, 2016

11. **Suvrit Sra**. "Positive definite matrices and the S-divergence". *Proceedings American Mathematical Society (PAMS)*, 2016

10. 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, Dec 2015. *Invited chapter*

9. Reshad Hosseini and **Suvrit Sra**. "Matrix Manifold Optimization for Gaussian Mixtures". In *Advances in Neural Information Processing Systems (NIPS)*, Dec 2015

8. 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)*, Dec 2015

7. **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, Dec 2015. *Invited chapter*

6. Zelda Mariet⋆ and **Suvrit Sra**. "Fixd-point algorithms for learning determinantal point processes". In *International Conference on Machine Learning (ICML)*, Jun 2015

5. Wolfgang Berndt and **Suvrit Sra**. "Hlawka–Popoviciu inequalities on positive definite tensors". *Linear Algebra and its Applications*, 486:317–327, 2015

4. Reshad Hosseini, **Suvrit Sra**, Lucas Theis, and Matthias Bethge. "Statistical inference with the Elliptical Gamma Distribution". In *Artificial Intelligence and Statistics (AISTATS)*, 2015

3. Minghua Lin and **Suvrit Sra**. "A proof of Thompson's determinantal inequality". *Mathematical Notes*, 2015

2. 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

1. **Suvrit Sra** and Reshad Hosseini. "Conic geometric optimisation on the manifold of positive definite matrices". *SIAM J. Optimization (SIOPT)*, 2015

Publications prior to MIT (Year ⩽ 2014)

49. 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

47. Anoop Cherian and **Suvrit Sra**. "Riemannian Sparse Coding of Positive Definite Matrices". In *European Conf. Computer Vision (ECCV)*, Jun 2014

46. 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

45. Matt Wytock, **Suvrit Sra**, and Zico Kolter. "Fast Newton methods for the group fused lasso". In *Uncertainty in Artificial Intelligence (UAI)*, Jun 2014

44. **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

43. Anoop Cherian, **Suvrit Sra**, V. Morellas, and N. Papanikolopoulos. "Efficient nearest neighbors via robust sparse hashing". *IEEE Transactions Image Processing*, 2014

42. Reshad Hosseini, **Suvrit Sra**, Lucas Theis, and M. Bethge. "Statistical inference with the Elliptical Gamma Distribution". *Computational Statistics and Data Analysis (CSDA)*, 2014

41. Stefanie Jegelka, Francis Bach, and **Suvrit Sra**. "Reflection methods for user-friendly submodular optimization". In *Advances in Neural Information Processing Systems (NIPS)*, December 2013

40. **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

39. **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

38. 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

37. **Suvrit Sra**. "Explicit eigenvalues of certain scaled trigonometric matrices". *Linear Algebra and its Applications (LAA)*, 438:173–181, 2013. 9 pages

36. **Suvrit Sra** and Dmitrii Karp. "The multivariate Watson distribution: Maximum-likelihood estimation and other aspects". *Journal of Multivariate Analysis (JMVA)*, 114:256–269, 2013

35. 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

34. **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

33. **Suvrit Sra**. "Scalable nonconvex inexact proximal splitting". In *Advances in Neural Information Processing Systems (NIPS)*, December 2012

32. **Suvrit Sra**. "Fast projections onto mixed-norm balls with applications". *Data Mining and Knowledge Discovery (DMKD)*, 25(2), Nov. 2012. 20 pages

31. 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

30. 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

29. 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

28. **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

27. **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**

26. 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

25. 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

24. 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

23. **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

22. 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

21. 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

20. 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

19. 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

18. **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

17. 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

16. Stefan Harmeling, Michael Hirsch⋆, **Suvrit Sra**, and Bernhard Schölkopf. "Online Blind Deconvolution for Astronomy". In *IEEE Int. Conf. on Computational Photography (ICCP)*, Apr. 2009

15. 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

14. Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and **Suvrit Sra**. *Text Mining: Theory, Applications, and Visualization A. Srivastava and M. Sahami (eds.)*, chapter Text clustering with mixture of von Mises-Fisher distributions. Data Mining & Knowledge Discovery Series. Chapman-Hall / CRC Press, 2009

13. **Suvrit Sra**. "Block-Iterative Algorithms for Non-negative Matrix Approximation". In *IEEE Int. Conf. Data Mining (ICDM)*, pages 1037–1042, Dec. 2008

12. 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)*

11. 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**

10. 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**

9. 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

8. Arun Surendran and **Suvrit Sra**. "Incremental Aspect Models for Mining Document Streams". In *Principles and Practice of Knowledge Discovery in Databases (PKDD)*, Sep. 2006

7. **Suvrit Sra**. "Efficient Large Scale Linear Programming Support Vector Machines". In *Euro. Conf. Machine Learning (ECML)*, pages 767–774, Sep. 2006

6. **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

5. Inderjit S. Dhillon and **Suvrit Sra**. "Generalized Nonnegative Matrix Approximations with Bregman Divergences". In *Advances Neural Information Processing Systems (NIPS)*, Dec. 2005

4. 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

3. 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

2. 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

1. 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)

7. Rashish Tandon and **Suvrit Sra**. "Sparse nonnegative matrix approximation: new formulations and algorithms". Technical Report 193, Max Planck Institute for Intelligent Systems, 2010. **(50 citations)**

6. **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. **(14 citations)**

5. **Suvrit Sra**, Dongmin Kim, and Bernhard Schölkopf. "Non-monotonic Poisson likelihood maximization". Technical Report 170, Jun. 2008. **(20 citations)**

4. **Suvrit Sra** and Inderjit S. Dhillon. "Nonnegative Matrix Approximation: Algorithms and Applications". Technical Report TR-06-27, Jun. 2006. **(87 citations)**

3. 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. **(45 citations)**

2. Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and **Suvrit Sra**. "Clustering on Hyperspheres using Expectation Maximization". Technical Report TR-03-07, Feb. 2003. **(16 citations)**

1. Inderjit S. Dhillon and **Suvrit Sra**. "Modeling data using directional distributions". Technical Report TR-03-06, Jan. 2003. **(135 citations)**

Preprints / Other Manuscripts

11. Guy Kornowski, Swati Padmanabhan, Kai Wang, Zhe Zhang, and **Suvrit Sra**. "First-Order Methods for Linearly Constrained Bilevel Optimization". *arXiv:2406.12771*, June 2024

10. Sanchayan Dutta⋆, Xiang Cheng, and **Suvrit Sra**. "Riemannian Bilevel Optimization". *arXiv:2405.15816*, May 2024

9. Yikun Hou, **Suvrit Sra**, and Alp Yurtsever. "Implicit Regularization in Matrix Factorization and Neural Networks with Diagonal Layers". *Submitted.*, (May), 2024

8. Hooman Maskan, Yikun Hou, **Suvrit Sra**, and Alp Yurtsever. "Revisiting Frank-Wolfe for Nonconvex Problems". *Submitted.*, (May), 2024

7. Adarsh Barik, **Suvrit Sra**, and Jean Honorio. "Invex Programs: First Order Algorithms and Their Convergence". *arXiv:2307.04456*, 2023. 18 pages

6. Xiang Cheng⋆, Jingzhao Zhang, and **Suvrit Sra**. "Theory and algorithms for diffusion processes on riemannian manifolds". *arXiv:2204.13665*, 2022. 156 pages

5. Melanie Weber and Suvrit Sra. "Computing Brascamp-Lieb constants through the lens of Thompson geometry". *arXiv:2208.05013*, 2022. 28 pages

4. Horia Mania⋆ and **Suvrit Sra**. "Why do classifier accuracies show linear trends under distribution shift?" *arXiv:2012.15483*, 2021. 18 pages.

3. Suvrit Sra. "Positive definite functions of noncommuting contractions, Hua-–Bellman matrices, and a new distance metric". *arXiv:2112.00056*, 2021. 11 pages

2. Pourya Habibzadeh⋆, **Suvrit Sra**, and Reshad Hosseini. "Deep RBF networks revisited: robust classification with rejection". *arXiv:1812.03190*, 2018

1. Jingzhao Zhang⋆, Hongyi Zhang⋆, and **Suvrit Sra**. "R-SPIDER: A fast Riemannian stochastic optimization with curvature independent rate". *arXiv:1811.04194*, 2018

INVITED TALKS AND LECTURES

1. Cambridge University. 26th Int. Symp. Math. Theory Networks & Systems. Plenary.   Aug 2024
   "AI and optimization through a geometric lens"

2. Int. Symposium on Math. Prog. (ISMP). Keynote speaker.   Jul 2024
   "Geometric thinking for Transformers and Non-Convex Optimization."

3. ICLR 2024. Workshop on Bridging the theory-practice gap in deep learning.   May 2024
   "How do Transformers learn to perform in-context learning."

4. Cornell University. ORIE Colloquium; Center for Data Science Distinguished Lecture.   Mar 2024
   "AI and optimization through a geometric lens."

5. Indian Statistical Institute, Kolkata. Winter School on Deep Learning. Plenary lecture.   Mar 2024
   "How do Transformers learn to perform in-context learning."

6. Google NYC. Workshop on Efficient Machine Learning.   Nov 2023
   "CCCP is Frank-Wolfe in disguise."

7. Georgia Tech. Joint Colloquium in Math, Applied and Computational Math (ACM).   Nov 2023
   "Learning, optimization, and sampling through a geometric lens."

8. UC Berkeley. Machine Learning Group Seminar   Nov 2023
   "Learning, optimization, and sampling through a geometric lens."

9. UC Berkeley. Simons Institute: Optimization and Algorithm Design Workshop.   Nov 2023
   "Geometric optimization (theory)."

10. TILOS-NSF AI Research Institute.   July 2023
    "Foundations of Geometric Sampling."

11. SIAM Conference on Optimization. Session: recent advances in nonconvex optimization. Jun 2023
"Nonconvex Sampling Through the Lens of Riemannian Geometry and Langevin MCMC."

12. Oxford University, Mathematics Institute; Department colloquium. May 05, 2023
"Numerical thinking in the AI era."

13. MIT. MIT IDSS Celebration Event; Invited talk. Apr 2023
"From OPTML to MLOPT."

14. Univ. of Chicago. Chicago Distinguished Speaker Series, Data Science Institute. Dec 2022
"Geometric optimization: old and new."

15. KAIST, Korea. Korean AI Association's Conference. Plenary speaker. Nov 2022
"Two surprises when optimization meets machine learning."

16. UC Berkeley. C3.AI DTI Colloquium. Sep 2022
"Two surprises when optimization theory meets ML practice."

17. Princeton University. Princeton Optimization Seminar, ORFE Princeton. Sep 2022
"Geometric optimization: old and new."

18. Int. Conf. on Continuous Optimization (ICCOPT), Bethlehem. Semi-plenary speaker. Jul 2022
"Two surprises when optimization theory meets machine learning practice."

19. University of Trier. Workshop on Algorithmic Optimization and Data Science. Keynote. Jul 2022
"Some Surprising Gaps between Optimization Theory and Practice in Machine Learning."

20. UC Santa Barbara. CCDC Seminar Series. Jan 2022
"A surprising non-convergence phenomenon in training deep networks."

21. UC Berkeley. BLISS seminar, EECS Department. Nov 2021
"Some surprising gaps between optimization theory and ML practice."

22. London School of Economics. Data Science Seminar. Nov 2021
"Do we understand how to find critical points in nonsmooth optimization?."

23. Urmia University, Iran. Workshop on Nonsmooth Opt. and Applications. Keynote. Sep 2021
"Finding stationary points of nonsmooth nonconvex functions."

24. ICLR 2021. Workshop on AI for Public Health. Keynote. May 2021

25. UC Berkeley. FODSI (NSF-Institute) Seminar Series. Mar 2021
"SGD without replacement: optimal rate analysis and more."

26. Boston University. Math-Physics Seminar. Mar 2021
"Accelerated gradient methods on Riemannian manifolds."

27. Caltech. ECE Department Seminar. Feb 2021
"Accelerated Gradient Methods on Riemannian Manifolds."

28. NeuRIPS 2020. OPT20 Workshop on Optimization for ML. Dec 2020
"SGD without replacement: optimal rate analysis and more."

29. FODSI Workshop (MIT, UC Berkeley) Oct 2020
"Geometric Optimization and Open problems."

30. University of Pennsylvania. PRiML–PIFODS Seminar Series. Sep 2020.
"Geometric optimization and Riemannian Accelerated Gradient."

31. EPFL Switzerland. Workshop on Learning, Optimization and Statistics Aug 2020
"Open Problems in Geometric Optimization."

32. University of Albany. Math Dept. Colloquium. Dec 06, 2019
"ReLU nets are powerful memorizers: A tight analysis of finite sample expressive power."

33. Yale University. Yale Institute for Network Science (YINS). Nov 20, 2019
"ReLU nets are powerful memorizers: a tight analysis of finite sample expressive power."

34. Fields Institute, Toronto. Conference on Data Science and Optimization.   Nov 19, 2019
    "ReLU nets are powerful memorizers: a tight analysis of finite sample expressive power."

35. UT Austin. Machine Learning Seminar (WNC Group, ECE Dept.)   Nov 08, 2019
    " ReLU nets are powerful memorizers: a tight analysis of finite sample expressive power."

36. Harvard University. CMSA Workshop on Noncommutative Analysis, Computational Complexity
    and Quantum Information.   Oct 18, 2019
    "Some nonconvex optimization problems through a geometric lens."

37. MIT. Computational Physics School for Fusion Research. Invited lecture.   Aug 30, 2019
    "Introduction to modern ML and Optimization."

38. Microsoft Research, Redmond. AI Institute workshop on Geometry of Deep Learning. Aug 26, 2019
    "Memorization capacity of ReLU nets."

39. ICCOPT 2019. Session on geometry in nonconvex optimization. Invited speaker.   Aug 08, 2019
    "Riemannian Optimization in ML, Statistics, etc."

40. ETH Zurich. Foundations of Data Science Seminar (Computer Science)   July 22, 2019
    "Foundations of geometric non-convex optimization."

41. ETH Zurich. Foundations of Data Science Seminar (Computer Science)   July 09, 2019
    "Negative dependence, stable polynomials, and all that."

42. NeurIPS 2018. Tutorial on Negative Dependence (jointly with S. Jegelka)   Dec 2018

43. Boston University. Algorithms and Complexity Seminar   Oct 2018
    "A critical view of optimality in deep learning."

44. Lehigh University. DIMACS/TRIPODS/MOPTA Plenary speaker.   Aug 13, 2018
    "Geometric nonconvex optimization."

45. Univ. Wisconsin, Madison. TRIPODS, Nonconvex Formulations Workshop   Jul 31, 2018
    "Geometric nonconvex optimization."

46. École polytechnique, Palaiseau, France. Data Science Summer School, invited lecturer.   Jul 2018
    "Nonconvex optimization"

47. MIT. EECS Special Seminar   Dec 14 2017
    "New tools for nonconvex optimization and machine learning."

48. University of Tübingen, Germany. Computer Science Department Seminar   Oct 2017
    "Frontiers of nonconvex optimization for machine learning."

49. *Simons Institute Berkeley: Workshop on Fast Iterative Methods in Optimization*   Oct 05 2017
    "Trends in nonconvex optimization."

50. Peking University. Applied Math Summer School. Invited lecturer.   Jul 10–15 2017
    "Lectures on: Optimization for machine learning."

51. MLSS 2017, Tübingen, Germany. Machine Learning Summer School. Invited lecturer. Jun 25-27 2017
    "Lectures on: Optimization for machine learning."

52. UC Davis, CA. Bayopt Conference on Data Science and Optimization.   May 12 2017
    "Geometric optimization."

53. UC Berkeley. BLISS Seminar, EECS Department.   Apr 24 2017
    "Geometric optimization: convex and nonconvex."

54. Optimization and Statistical Learning (OSL) Workshop, Les Houches, France   Apr 11 2017
    "Geometric Optimization."

55. NIPS 2016. Workshop on Nonconvex Optimization.   Dec 09 2016
    "Geometric Optimization."

56. Northeastern University, Boston. Data Science Seminar.   Nov 04 2016
    "Geometric Optimization."

57. MIT. LIDS Seminar Series.   Sep 13 2016
    "Geometric optimization."

58. SIAM Annual Meeting.   Jul 11 2016
    "Geometric (nonconvex) optimization."

59. ILAS 2016. Matrix Inequalities and Operator Means. Plenary speaker.   Jul 2016
    "Canceled plenary on: Schur convexity, matrix integrals, and combinatorics."

60. ICML 2016. Workshop on advances in non-convex analysis and optimization.   Jun 23 2016
    "Faster nonconvex finite-sum optimization."

61. Google Research, NYC. Google Machine Learning Workshop.   Mar 03 2016
    "Geometric Optimization in ML."

62. MIT Lincoln Labs. PED Seminar.   Feb 09 2016
    "Geometric convex (and nonconvex) optimization."

63. Hausdorff Center for Mathematics, Bonn.   Jan 2016
    "Lectures on large-scale optimization."

64. Boston University. Computer Vision Seminar   Sep 2015
    "Geometric optimization for GMMs (and more broadly for ML)."

65. Microsoft Research India. Machine Learning Summer School Lecturer   June 2015
    "Introduction to large-scale optimization."

66. Cambridge, MA. New England Machine Learning Conference (NEML)   May 18 2015
    "Geometric optimization in machine learning."

67. MIT. Math Dept., Combinatorics Seminar.   Mar 13 2015
    "Solution to the Cuttler-Green-Skandera conjecture on Schur Polynomials."

68. Caltech. CMS Division seminar.   June 2014
    "Inexactness, geometry, and optimization: recurrent themes in modern data analysis."

69. UC San Diego. CSE Department seminar.   May 29 2014
    "Inexactness, geometry, and optimization: recurrent themes in modern data analysis."

70. MIT. LIDS, EECS Department.   May 8 2014
    "Inexactness, geometry, and optimization: recurrent themes in modern data analysis."

71. Cornell University. ORIE Seminar   May 5 2014
    "Inexactness, geometry, and optimization: recurrent themes in modern data analysis."

72. University of Saarland, Germany. Department of Computer Science.   May 2014.
    "Inexactness, geometry, and optimization: recurrent themes in modern data analysis."

73. RWTH Aachen, Germany. EU Regional School, Invited short-course.   Apr 11 2014
    "Introduction to Machine Learning."

74. Microsoft Research, Cambridge, UK.   Mar 17 2014
    "Inexactness, Geometry, and Optimization."

75. University of Cambridge. Engineering Department.   Mar 17 2014
    "Matrix Means, Distances, Kernels, and Geometric Optimization."

76. University of British Columbia, Vancourver. Computer Science Dept. seminar.   Feb 26 2014
    "Inexactness, geometry, and optimization: recurrent themes in modern data analysis."

77. Carnegie Mellon University. Tepper School of Business.   Feb 21 2014
    "Inexactness, geometry, and optimization: recurrent themes in modern data analysis."

78. Carnegie Mellon University. Machine Learning Department   Feb 19 2014
    "Inexactness, geometry, and optimization: recurrent themes in modern data analysis."

79. Carnegie Mellon University. Statistics Department Seminar   Nov 11 2013
    "Mathematics of data analysis: Geometric Optimisation for Data Analysis."

80. MIT. LIDS seminar   Oct 18 2013
    "Matrix Means, Distances, Kernels, and Geometric Optimization."

81. Washington University at St. Louis. CSE Dept. Colloquium   Oct 15 2013
    "Challenges in data analysis: Leveraging inexactness for scalability."

82. *École polytechnique fédérale de Lausanne, Math Department, Lausanne*   Sep 18 2013
    "Leveraging inexactness for scalability."

83. TU Munich. Math Department Seminar.   Sep 17, 2013
    "Mathematics of data analysis: Geometric optimization for data analysis."

84. EPFL. School of Computer Science   Apr 15 2013
    "Leveraging inexactness for scalability."

85. UC Davis. Algebra and Discrete Math Seminar   Mar 14 2013
    "A new metric on $\mathbb{P}^n$ and its connections."

86. Carnegie Mellon University. Machine Learning Department   Feb 28 2013
    "Leveraging inexactness for scalability."

87. University College London. Gatsby Computational Neuroscience Unit   Feb 18 2013
    "Leveraging inexactness for scalability."

88. Cornell University. Computer Science Colloquium.   Feb 14 2013
    "Leveraging inexactness for scalability."

89. Duke University. ECE Department Colloquium.   Feb 11 2013
    "Leveraging inexactness for scalability."

90. INRIA Rocquencourt, (ENS/CNRS/ENS joint laboratory).   Jan 21, 2013
    "A new metric on $\mathbb{P}^n$ and matrix means."

91. Oxford University. The Mathematical Institute Seminar.   Jan 09, 2013
    "Leveraging inexactness for scalable optimisation."

92. University of British Columbia (UBC), Vancouver. ECE Department seminar   Oct 10, 2012
    "Leveraging inexactness for scalability."

93. Oxford University. Dept. of Engineering Science.   Sep 2012
    "Challenges in Big-Data: A framework for scalable nonconvex optimisation."

94. Kyungpook National Univ., Dept. of Mathematics, South Korea. Hot Topics Workshop on: "Positive
    matrices and operators: Recent developments and Advances."   Jun 26–30, 2012
    "A computationally simple distance for positive definite matrices."

95. Google Research, Mountain View, CA.   May 24, 2012
    "A framework for scalable nonsmooth optimization."

96. IBM Research, New York.   May 8, 2012
    "A framework for scalable nonsmooth optimization."

97. UMass Amherst. ECE Dept. Seminar   Apr 10, 2012
    "A framework for scalable nonsmooth optimization."

98. Yahoo! Research, Sunnyvale, CA. Invited Tech Talk.   Mar 9, 2012
    "A framework for scalable nonsmooth optimization."

99. LinkedIn, Mountain View, CA. Invited Tech Talk.   Mar 6, 2012
    "A framework for scalable nonsmooth optimization."

100. Microsoft Research, Redmond, WA. Invited Tech Talk   Aug 29, 2011
    "Positive definite matrices and the Symmetrized Stein Divergence."

101. University of Washington, Seattle. EE, CS, Math Seminar.   Aug 25, 2011
    "Positive definite matrices and the Symmetrized Stein Divergence."

102.	Yahoo Research, Santa Clara, CA. Tech Talk.	Aug 18, 2011
	"Positive definite matrices and the Symmetrized Stein Divergence."	
103.	Toyota Technological Institute (TTI), Chicago.	Aug 16, 2011
	"Positive definite matrices and the Symmetrized Stein Divergence."	
104.	University of Chicago. Statistics Colloquium	Aug 15, 2011
	"The metric nearness problem."	
105.	Householder Symposium XVIII, Tahoe.	Jun 2011
	"Solving large-scale nonnegative least-squares."	
106.	ETH Zürich. Informatik Seminar	Apr 2011
	"Scalable algorithms for nonsmooth composite convex minimization."	
107.	Universidad Autónoma de Madrid. Computer Science Dept	Apr 2011
	"Lecture on Optimization in Machine Learning."	
108.	University of Buffalo. Computer Sciences Colloquium	Mar 2011
	"arge-scale Optimization for Machine Learning and Scientific Computing."	
109.	UCLA, Institute for Pure and Applied Mathematics (IPAM). Applications of Optimization Workshop Dec 2010	
	"Multiframe blind deblurring"	
110.	IST Austria. First Symposium on Computer Vision and Machine Learning.	Oct 2010
	"Kernels for covariance matrices."	
111.	RWTH Aachen, Germany. Invited short-course at the EU Regional School	Apr 2010
	"Lecture on Matrix Factorization and Approximation Problems."	
112.	Microsoft Research, Redmond. Tech Talk	Dec 2008
	"New projected quasi-Newton methods with applications."	
113.	Google, Mountain View, CA.	Mar 2007
	"Matrix nearness problems in data mining: nonnegative matrix approximation."	
114.	Yahoo Inc., Sunnyvale, CA.	Mar 2007
	"Matrix Nearness Problems in Data Mining."	
115.	UCLA, IPAM. SAGE Days, 3rd Conference, IPAM.	Feb 2007
	"Sparse matrix computations with SAGE."	

PROFESSIONAL SERVICE

ASSOCIATE EDITOR: SIAM J. Matrix Analysis and Applications (SIMAX); Oct 2024–

ASSOCIATE EDITOR: SIAM J. Mathematics of Data Science (SIMODS); 2024–now

ASSOCIATE EDITOR: IMA Journal on Information & Inference; 2023–now

ASSOCIATE EDITOR: Journal of Machine Learning Research (JMLR); 2018–now

ASSOCIATE EDITOR: Optimization Methods and Software (OMS); 2014–2020

ASSOCIATE EDITOR: IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI); 2020

SCIENTIFIC COMMITTEE: "Machine learning and optimization" track at ISMP 2018.

WORKSHOP ORGANIZER AND CO-CHAIR: OPT 2008—OPT 2019
International workshops help at the *Neural Information Processing Systems* Conference (except 2019)

SENIOR PROGRAM COMMITTEE / SENIOR AREA CHAIR

1. International Conference on Learning Representations (ICLR) 2024
2. Neural Information Processing Systems (NeuRIPS) 2021–2024
3. International Conference Machine Learning (ICML) 2020–2023
4. Conference on Learning Theory (COLT) 2021, 2020

PROGRAM COMMITTEE / AREA CHAIR

1. Neural Information Processing Systems 2019–2016, 2014, 2013
2. International Conference on Machine Learning (ICML) 2020, 2018, 2016, 2015, 2014
3. Conference on Learning Theory (COLT) 2018
4. Systems and Machine Learning (SML) 2020
5. Association for the Advancement of Artificial Intelligence (AAAI) 2018
6. Artificial Intelligence & Statistics (AISTATS) 2020, 2016, 2014
7. Uncertainty in Artificial Intelligence (UAI) 2020

Sponsorship Chair: Artificial Intelligence and Statistics (AISTATS) 2012

Minisymposium organizer: "Advances in large-scale optimization"
 Invited minisymposium at *SIAM Annual Meeting, 2016*

Program Committee Member / Reviewer

– *Symposium on Theory of Computing (STOC)* 2020, 2018
– *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 (JSTSP)*
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. *European Journal of Combinatorics*
26. *Journal of Multivariate Analysis (JMVA)*
27. *Journal of Statistical Computation and Simulation (JSCS)*
28. *Linear Algebra and its Applications (LAA)*
29. *Optimization (Taylor & Francis)*
30. *Optimization and Engineering (OPTE)*
31. *Optimization Letters (Springer)*
32. *Science*
33. *Signal Processing (Elsevier)*
34. *Numerical Algorithms (Springer)*
35. *Pattern Recognition (PR)*
36. *Reviewer for* MATHSCINET
37. *American Mathematical Monthly*
38. *Journal American Mathematical Society (JAMS)*

OTHER SERVICE

Graduate admissions (ML Area), EECS department, MIT
Advising Course 6 undergraduates, MIT
Reviewing Postdoc, PhD, and Master's applications at Max-Planck Institute
ACM and SCS Doctoral Dissertation Committee, Carnegie Mellon University & MIT

STUDENTS and POSTDOCS

▶ *Graduated PhDs and Postdocs (with first appointment)*

1. Yi Tian (PhD (**PhD** Aug 2024, EECS, MIT; ML Scientist, Meta.)
2. Kwangjun Ahn* (**PhD** Jun 2024, EECS MIT; Sr. Researcher, Microsoft Research.)
3. Xiang Cheng (Postdoc 8/20–07/24, LIDS; PhD: UC Berkeley; joining Asst. Prof. Duke Univ.)
4. Hadi Daneshmand (Postdoc 11/22–11/23, LIDS; PhD: ETH; joining faculty position)
5. Zhang Zhe (Postdoc 8/23–01/24, LIDS; PhD: Georgia Tech; Asst. Prof. Purdue)
6. Kai Wang (Postdoc 8/23–01/24, LIDS, CSAIL; PhD: Harvard; Asst. Prof. Georgia Tech)
7. Joshua Robinson† (**PhD** 2023, EECS, MIT; Postdoc Stanford)
8. Tiancheng Yu (**PhD** 2023, EECS, MIT; Two-Sigma)
9. Alp Yurtsever (Postdoc 2020-21; Asst. Prof. Umea Univ. Sweden)
10. Jingzhao Zhang* (**PhD** 2021, EECS, MIT; Asst. Prof. Tsinghua Univ. China)
11. Chulhee Yun* (**PhD** 2021, EECS, MIT; Asst. Prof. KAIST, South Korea)
12. Horia Mania*★ (Postdoc 2020–22, LIDS, EECS; PhD: UC Berkeley; Citadel)
13. Zelda Mariet (**PhD** 2019, CSAIL, MIT; Research Scientist at Google Brain)
14. Hongyi Zhang (**PhD** 2019, BCS, MIT; Research Scientist at Bytedance Inc.)
15. Chengtao Li† (**PhD** 2019, CSAIL, MIT; co-founder of a Galixir)

Co-supervised with: *Ali Jadbabaie (MIT); †Stefanie Jegelka (MIT); ★Devavrat Shah (MIT).

▶ *Current Students Under Supervision*

1. Pourya Habibzadeh (PhD, EECS, MIT)
2. Swati Padmanabhan (Postdoc 7/23–now, LIDS, EECS; PhD: Univ. Washington)
3. Manish Lal (Postdoc 5/24–now; TUM, Math; PhD: UBC, Math)
4. Maria-Luiza Vladarean (Postdoc 6/24–now; TUM, Math; PhD: EPFL CS)
5. Alina Yang (MIT MISTI-Germany Student, Summer 2024)
6. Sanchayan Dutta (PhD student, Math, UC Davis; Intern Summer 2024)
7. Evan Montoya (MS student, Math, Georgia Tech; Intern Summer 2024)
8. Xuhui Zhang (PhD, Math, CIT TUM; 6/24–)

9. Sai Niranjan (PhD, Math, CIT TUM; 7/24–)
10. Pouria Fatemi (Phd, Math, CIT TUM; 10/24–)
11. Josef Ott (Masters, CIT TUM; 9/24–)
12. Anh Duc Nguyen (Bachelors, NUS Singapore; Intern; 2/24–)

▶ *PhD Committee Member*
- 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
- Qingkai Liang (Aeronautics and Astronautics Dept., MIT) 2017-2018
- Kenji Kawaguchi (EECS, MIT) 2020
- Manolis Zampetakis (EECS, MIT) 2020
- Curtis Northcutt (EECS, MIT) 2021
- Chenyang Yuan (EECS, MIT) 2022

▶ *Master's theses*
- Yi Tian (MS, EECS)
- Kwangjun Ahn (MS, EECS)
- Tiancheng Yu (MS, EECS)
- Pourya Habibzadeh (MS, EECS)
- Jingzhao Zhang (MS, EECS)
- Hongyi Zhang (MS, BCS)
- Rajeev Parvathala (M.Eng., EECS, June 2018)
- Chengtao Li (MS, EECS)
- Zelda Mariet (MS, EECS)
- Johannes Kitschke (Master, RWTH Aachen, 2013)

▶ *Graduate students; supervised for ≥ 1 research projects:*
- Foivos Alimisis (3/24–6/24, PhD Math, U. Geneva; intern TUM)
- Melanie Weber (2017-now, PhD Princeton Math; Asst. Prof. Harvard University)
- Jikai Jin (2021, PKU Math; now PhD student Stanford)
- Alp Yurtsever (2019, PhD EPFL; now Asst. Professor, Umea Univ.)
- Sashank Reddi (2013–17, PhD student CMU; now Google Research NY)
- Pourya Habib'zadeh (2016, MS student, Univ. Tehran, Iran; now PhD EECS MIT)
- Yu-Xiang Wang (2013–14, PhD student CMU; Assoc. Prof. UC Santa Barbara)
- Adams Wei Yu (2013-14, MS CMU; now PhD CMU MLD)
- Ahmed Hefny (2013-15, PhD student CMU)
- Veeranjaneyulu Sadhanala (2013-14, PhD student CMU)
- Anoop Cherian (2011–2013, PhD student UMN; Principal Researcher MERL, Cambridge)
- 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:*
- Jikai Jin (Peking University, Math; Intern)
- Alex Gu (SuperUROP 2020-21; MIT EECS)
- Brice Huang (Spring 2019; MIT Math; UROP)

- Haochen Zhang (Spring 2018; Intern@MIT; Tsinghua University)
- 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

COMPANIES FOUNDED

Pendulum Systems (previously macro-eyes). Co-founder. Founded 2014.

PERSONAL

Native to near-native: English, Hindi, Punjabi
Fluent: German;   Intermediate: Urdu;   Beginner: Italian, French