Raghav Somani

Fourth year Ph.D. student at
Paul G. Allen School of Computer Science & Engineering
University of Washington
Advisor: Prof. Sewoong Oh

October 18, 2022
raghavs@cs.washington.edu
raghavsomani1995@gmail.com
raghavsomani.github.io

Research Interests

I am interested in machine learning and optimization broadly. Particularly in characterizing fundamental quantities, phenomena, and general laws that arise due to scale.

Research Interests	Machine Learning, Large-scale Optimization, Probability theory
Other Learning Interests	Geometry and Analysis, Pure Mathematics

Education

- **University of Washington** — *Sept '19 - Present*
  Ph.D. in Computer Science and Engineering
  Advisor: Prof. Sewoong Oh
  GPA: 3.90*/4
- **University of Washington** — *March '22*
  M.S. in Computer Science and Engineering
  Supervisor: Prof. Sewoong Oh
  GPA: 3.90/4
- **Indian Institute of Technology Guwahati** — *July '13 - June '17*
  Bachelor of Technology in Mathematics and Computing
  GPA: 9.10/10 (9.30/10 in major courses)

Work Experiences

- **Microsoft Research India - Research Fellow** — *July '17 - July '19*
  Advisors: Dr. Praneeth Netrapalli & Dr. Prateek Jain
  Project group - Provable Non-convex Optimization for Machine Learning Problems
- **Microsoft Research India - Research Intern** — *May '16 - July '16*
  Advisor: Dr. Sreangsu Acharyya
  Project - Recommendation systems
- **CAFRAL, Reserve Bank of India - Summer Research Intern** — *June '15 - July '15*
  Advisor: Prof. Nagpurnanand R. Prabhala
  Project - Modeling 'Economic Policy Uncertainty Index' for India

Selected Publications

Preprints

1. **Stochastic optimization on matrices and a graphon McKean-Vlasov limit** [arXiv]
   Z. Harchaoui, S. Oh, S. Pal, **R. Somani**, and R. Tripathi[#].
   Under review.
2. **Gradient flows on graphons: existence, convergence, continuity equations** [arXiv]
   S. Oh, S. Pal, **R. Somani**, and R. Tripathi[#].
   Under review.

Conference Publications

1. **Robust Meta-learning for Mixed Linear Regression with Small Batches** [paper]
   W. Kong, **R. Somani**, S. M. Kakade, and S. Oh.
   Advances in Neural Information Processing Systems (NeurIPS), December 2020.

2. **Meta-learning for mixed linear regression** [paper]
W. Kong, **R. Somani**, Z. Song, S. M. Kakade, and S. Oh.
International Conference on Machine Learning (ICML), July 2020.

3. **Support Recovery for Orthogonal Matching Pursuit: Upper and Lower bounds.** [paper]
R. Somani[*], C. Gupta[*], P. Jain, and P. Netrapalli.
Advances in Neural Information Processing Systems (NeurIPS), Montréal, Canada, December 2018. **Spotlight.**

4. **Clustered Monotone Transforms for Rating Factorization** [paper]
R. Somani[*], G. Hiranandani[*], O. Koyejo, and S. Acharyya.
ACM International Conference on Web Search and Data Mining (WSDM), Melbourne, Australia, February 2019.

Workshop Publications

1. **Gradient flows on graphons: existence, convergence, continuity equations**
S. Oh, S. Pal, **R. Somani**, and R. Tripathi[#].
OTML workshop, *Neural Information Processing Systems (NeurIPS), December 2021.*

2. **Non-Gaussianity of Stochastic Gradient Noise** [arXiv]
A. Panigrahi, **R. Somani**, N. Goyal, and P. Netrapalli.
SEDL workshop, *Neural Information Processing Systems (NeurIPS), Vancouver, Canada, December 2019.*

[*] - equal contribution, [#] - alphabetical ordering

Please visit (dblp) or (Google Scholar) for the list of all research articles.

Research Projects

Scaling laws of optimization algorithms for Deep Learning *May '21 - Present*
Advisors: Prof. Sewoong Oh, Prof. Soumik Pal & Prof. Zaid Harchaoui, University of Washington [Page]

- Neural Networks are large computational graphs. We try to understand the scaling limits of the family of first-order stochastic optimization processes on large unlabeled graphs. We use limiting graph theory, mean-field theory and the theory of gradient flows on metric spaces.
- We characterize the limiting process through an extension of the classical McKean-Vlasov limit.
- This is an attempt to generalize the Wasserstein calculus to higher-order exchangeable structures.
- Deep feedforward neural networks can be modeled as a sequence of graphs sharing vertex labels. Currently working to generalize the same theory and derive scaling limits of optimization processes on deep neural networks.
- ★ Work on gradient flows on graphons got accepted at **OTML '21** at **NeurIPS '21**. [arXiv]

Robustness and Meta Learning for Mixed Linear Regression (MLR) *Nov '19 - Mar '21*
Advisors: Prof. Sewoong Oh & Dr. Weihao Kong, University of Washington [Page]

- Analyzed conditions under which abundant tasks with small data can compensate for lack of tasks with large data in the context of MLR. [arXiv]
- Tightened conditions using the sum-of-squares method. Worked on the robust subspace estimation. [arXiv]
- Looked into Robust Subspace estimation for mixture distributions under the batch setting.
- ★ Works got accepted at **ICML '20** & **NeurIPS '20** respectively. [video 1, video 2]

Optimization and Generalization in Deep Neural Networks *July '18 - Sept '19*
Advisors: Dr. Prateek Jain, Dr. Praneeth Netrapalli & Dr. Navin Goyal, Microsoft Research [Page 1, Page 2]

- Understanding the dependence of batch-size (stochasticity), over-parameterization, and optimization on the generalization properties of a variety of neural networks on real world data distributions for classification tasks.
- Analyzing the dependence of support separation, number of hidden neurons, ambient dimension of data distribution, number of training points on optimization and generalization of shallow neural networks.
- ★ A work on distributional characterization of SGD got accepted at the workshop **SEDL '19** at **NeurIPS '19**. [arXiv]

Sparse Regression and Optimal Bounds for Orthogonal Matching Pursuit (OMP) *Sept '17 - June '18*
Advisors: Dr. Prateek Jain & Dr. Praneeth Netrapalli, Microsoft Research [Page]

- Analyzed Accelerated IHT, trying to strengthen Jain et al.'s results for better support expansion and generalization.
- Analyzed OMP for the Sparse Linear Regression problem under Restricted Strong Convexity (RSC) assumptions obtaining its support recovery and generalization guarantees. Also provide tight lower bounds for OMP. Our results are the first such matching upper and lower bounds (up to log factors) for *any* Sparse Regression algorithm under RSC assumption.
- ★ Accepted for a **Spotlight** paper presentation at **NeurIPS '18**. [paper, spotlight video]

Clustered Monotone Transforms for Rating Factorization (CMTRF) *May '16 - Aug '18*
Advisors: Dr. Sreangsu Acharyya (MSR India) & Prof. Oluwasanmi Koyejo (UIUC) [Page]

- Implemented and analyzed CMTRF for recommendation systems which performs regression under shared low-rank structure up to unknown monotonic transforms. CMTRF recovers a unique solution under mild conditions and also outperforms other state-of-the-art baselines on 7 benchmark and 2 synthetic datasets.
- ★ Accepted for an oral presentation at **WSDM '19**. [arXiv]

Selected Talks

- Scaling limits of SGD over large networks
  1. Machine Learning Foundations seminar @ Microsoft Research Redmond *Sept '22*
  2. Scaling Laws for Deep Learning Micro-workshop @ University of Washington, CSE *Aug '22*
- Gradient flows on Graphons @ The Kantorovich Initiative retreat 2022 *March '22*
- Sparse Regression and Optimal Bounds for Orthogonal Matching Pursuit (OMP) @ Machine Learning seminar, Microsoft Research India *2019*

Relevant Courses

Optimization	Theory of Optimization, Variational Analysis, Convex Optimization (Stanford MOOC)
Probability	Advanced Probability, Probability & Random Processes
Computer Science	Design and Analysis of Algorithms, Randomized Algorithms, Discrete Mathematics
Finance	Financial Engineering I & II, Monte Carlo Simulations, Stochastic Calculus for Finance
Mathematics	Linear Algebra, Real & Complex Analysis, Stochastic Calculus, Optimal Transport Matrix Computation, Scientific Computation

Professional Responsibilities

1. *Reviewer*: NeurIPS '19, NeurIPS '20, ICML '20, JMLR.
2. *Teaching Assistant*: Machine Learning CSE 446/546 Spring '21.

Scholastic Achievements

2013 Among top 1.7% of all selected candidates (126,000+) in JEE-Advanced (IIT-JEE).
2013 Among top 0.5% of all candidates (1,400,000+) in JEE-Mains.
2013 Among top 0.15% of all candidates (150,000+) in WBJEE.

References

1. *Prof. Sewoong Oh*, Associate Professor, University of Washington
2. *Prof. Soumik Pal*, Professor of Mathematics, University of Washington
3. *Dr. Praneeth Netrapalli*, Research Scientist, Google Research, India
4. *Dr. Prateek Jain*, Sr. Staff Research Scientist, Google Research, India