Sandeep Silwal

silwal@mit.edu | sandeepsilwal.com

EDUCATION

Massachusetts Institute of Technology

PhD in Computer Science   *September 2019 - June 2024 (Expected)*
Advisor: Piotr Indyk

B.S. in Mathematics with Computer Science   *2015-2019*
Advisor: Michel Goemans
Cumulative GPA: 5.0/5.0

Current Research Interests: I work in the intersection of machine learning and classical algorithms by designing provable algorithms in various ML settings, such as efficient algorithms for processing large datasets, as well as using ML to inspire and motivate algorithm design itself.

PUBLICATIONS

1. Anders Aamand, Alexandr Andoni, Justin Y. Chen, Piotr Indyk, Shyam Narayanan, and Sandeep Silwal. Data structures for density estimation. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 1–18, 2023

2. Nicholas Schiefer, Justin Y. Chen, Piotr Indyk, Shyam Narayanan, Sandeep Silwal, and Tal Wagner. Learned interpolation for better streaming quantile approximation with worst-case guarantees. In *SIAM Conference on Applied and Computational Discrete Algorithms (ACDA23)*, pages 87–97, 2023

3. Anders Aamand, Justin Y. Chen, Huy Lê Nguyen, and Sandeep Silwal. Improved Space Bounds for Learning with Experts. *arXiv e-prints*, page arXiv:2303.01453, poster at ACDA 2023

4. Ainesh Bakshi, Piotr Indyk, Praneeth Kacham, Sandeep Silwal, and Samson Zhou. Subquadratic algorithms for kernel matrices via kernel density estimation. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. **Spotlight Presentation**.

5. Sandeep Silwal, Sara Ahmadian, Andrew Nystrom, Andrew McCallum, Deepak Ramachandran, and Seyed Mehran Kazemi. Kwikbucks: Correlation clustering with cheap-weak and expensive-strong signals. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023

6. Yeshwanth Cherapanamjeri, Sandeep Silwal, David P. Woodruff, Fred Zhang, Qiuyi Zhang, and Samson Zhou. Robust algorithms on adaptive inputs from bounded adversaries. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023

7. Yeshwanth Cherapanamjeri, Sandeep Silwal, David P. Woodruff, and Samson Zhou. Optimal algorithms for linear algebra in the current matrix multiplication time. In Nikhil Bansal and Viswanath Nagarajan, editors, *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023*, pages 4026–4049. SIAM, 2023

8. Piotr Indyk and Sandeep Silwal. Faster linear algebra for distance matrices. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information*

Processing Systems, volume 35, pages 35576–35589. Curran Associates, Inc., 2022. **Oral Presentation**

9. Anders Aamand, Justin Chen, Piotr Indyk, Shyam Narayanan, Ronitt Rubinfeld, Nicholas Schiefer, Sandeep Silwal, and Tal Wagner. Exponentially improving the complexity of simulating the weisfeiler-lehman test with graph neural networks. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 27333–27346. Curran Associates, Inc., 2022

10. Elena Grigorescu, Young-San Lin, Sandeep Silwal, Maoyuan Song, and Samson Zhou. Learning-augmented algorithms for online linear and semidefinite programming. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 38643–38654. Curran Associates, Inc., 2022

11. Michael Kapralov, Mikhail Makarov, Sandeep Silwal, Christian Sohler, and Jakab Tardos. Motif cut sparsifiers. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*, pages 389–398. IEEE, 2022

12. Justin Chen, Sandeep Silwal, Ali Vakilian, and Fred Zhang. Faster fundamental graph algorithms via learned predictions. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 3583–3602. PMLR, 17–23 Jul 2022

13. Eric Price, Sandeep Silwal, and Samson Zhou. Hardness and algorithms for robust and sparse optimization. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 17926–17944. PMLR, 17–23 Jul 2022

14. Miklós Ajtai, Vladimir Braverman, T.S. Jayram, Sandeep Silwal, Alec Sun, David P. Woodruff, and Samson Zhou. The white-box adversarial data stream model. In *Proceedings of the 41st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, PODS '22, page 15–27, New York, NY, USA, 2022. Association for Computing Machinery

15. Justin Y. Chen, Talya Eden, Piotr Indyk, Honghao Lin, Shyam Narayanan, Ronitt Rubinfeld, Sandeep Silwal, Tal Wagner, David P. Woodruff, and Michael Zhang. Triangle and four cycle counting with predictions in graph streams. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022

16. Jon C. Ergun, Zhili Feng, Sandeep Silwal, David P. Woodruff, and Samson Zhou. Learning-augmented k-means clustering. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022. **Spotlight Presentation**.

17. Sandeep Silwal. A concentration inequality for the facility location problem. *Oper. Res. Lett.*, 50(2):213–217, 2022

18. Zachary Izzo, Sandeep Silwal, and Samson Zhou. Dimensionality reduction for wasserstein barycenter. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 15582–15594. Curran Associates, Inc., 2021

19. Vladimir Braverman, Avinatan Hassidim, Yossi Matias, Mariano Schain, Sandeep Silwal, and Samson Zhou. Adversarial robustness of streaming algorithms through importance sampling. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 3544–3557. Curran Associates, Inc.,

2021. **Silver Best Paper Award** for preliminary version at Adversarial ML workshop at ICML 2021.

20. Rikhav Shah and Sandeep Silwal. Smoothed analysis of the condition number under low-rank perturbations. In Mary Wootters and Laura Sanità, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2021, August 16-18, 2021, University of Washington, Seattle, Washington, USA (Virtual Conference)*, volume 207 of *LIPIcs*, pages 40:1–40:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021

21. Shyam Narayanan, Sandeep Silwal, Piotr Indyk, and Or Zamir. Randomized dimensionality reduction for facility location and single-linkage clustering. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 7948–7957. PMLR, 18–24 Jul 2021

22. Talya Eden, Piotr Indyk, Shyam Narayanan, Ronitt Rubinfeld, Sandeep Silwal, and Tal Wagner. Learning-based support estimation in sublinear time. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021.* OpenReview.net, 2021. **Spotlight Presentation**.

23. Vishesh Jain and Sandeep Silwal. A note on the universality of esds of inhomogeneous random matrices. *ALEA. Latin American Journal of Probability and Mathematical Statistics*, 18:1047 – 1059, 2021

24. Maryam Aliakbarpour and Sandeep Silwal. Testing properties of multiple distributions with few samples. In Thomas Vidick, editor, *11th Innovations in Theoretical Computer Science Conference, ITCS 2020, January 12-14, 2020, Seattle, Washington, USA*, volume 151 of *LIPIcs*, pages 69:1–69:41. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020

25. Rogers Epstein and Sandeep Silwal. Property testing of lp-type problems. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of *LIPIcs*, pages 98:1–98:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020

26. Rikhav Shah and Sandeep Silwal. Using dimensionality reduction to optimize t-sne. *CoRR*, abs/1912.01098, 2019, poster at OPT-ML Workshop at NeurIPS 2019

27. Jesse Michel, Sushruth Reddy, Rikhav Shah, Sandeep Silwal, and Ramis Movassagh. Directed random geometric graphs. *J. Complex Networks*, 7(5):792–816, 2019

MANUSCRIPTS

1. Ainesh Bakshi, Piotr Indyk, Rajesh Jayaram, Sandeep Silwal. A Near-Linear Time Algorithm for the Chamfer Distance. In Preparation

2. Anders Aamand, Justin Y. Chen, Allen Liu, Sandeep Silwal, Pattara Sukprasert, Ali Vakilian, Fred Zhang. A Constant-Factor Approximation for Individual Preference Stable Clustering. In Preparation

3. Anders Aamand, Justin Y. Chen, Huy Nguyen, Sandeep Silwal, Ali Vakilian. Improved Frequency Estimation Algorithms with and without Predictions. In Preparation

PROFESSIONAL EXPERIENCE

- Research Intern at Microsoft Research. Supervised by Sepideh Mahabadi, *Summer 2023* Jakub Tarnawski, and worked with the Algorithms team

- Research Intern at Google Research. Supervised by Mehran Kazemi *Summer 2022* and Sara Ahmadian. Work lead to publication *KwikBucks: Correlation Clustering with Cheap-Weak and Expensive-Strong Signals* at ICLR 2023

- Research Intern at Nokia Bell Labs. Supervised by Dan Kushnir   *Summer 2021*

- Research Intern at IBM Research Cambridge. Supervised by Ramis Movassagh.   *Winter 2018*
  Lead to my first publication *Directed Random Geometric Graphs*

- Trading Intern at TransMarket Group. Created tools to gather and analyze   *Summer 2017*
  large mortgage data sets to identify potentially profitable trading opportunities

- Data science intern at Soofa. Conducted research to determine the factors   *Winter 2016*
  that make a location attractive to pedestrians such as commercial presence. Created a web application using public data sets to visualize popular pedestrian locations in any city as a heatmap

- Modeling intern at Brigham and Women's Hospital. Created a physiologically   *Winter 2015*
  based model of human sleep dynamics based off of academic papers. Model correctly predicted statistics about human sleep such as the average number of times humans wake up during sleep.

TEACHING AND ADVISING EXPERIENCE

- Teaching assistant for 6.1400: Computability and Complexity Theory at MIT   *Spring 2023*
  Overall rating: 6.8/7 (highest among all TAs and Professors)

- Teaching assistant for 6.854: Advanced Graduate Algorithms at MIT   *Fall 2018*
  Overall rating: 6.4/7 (highest among all TAs)

- Helped organize and contributed towards diversity and inclusion discussion panels during MIT graduate EECS visit days.

- Volunteered at MIT GAAP program whose mission is to help underrepresented groups apply to graduate schools. Advised a student on their application to UC Berkeley EECS where they successfully matriculated.

- Ran MIT Algorithms Office Hours and improved communication between theory and non theory PhD students.

AWARDS AND ACHIEVEMENTS

- Spotlight Presentation for *Subquadratic Algorithms for Kernel Matrices*   *ICLR 2023*
  via Kernel Density Estimation. Top $\sim 8\%$ of submissions

- Oral Presentation for *Faster Linear Algebra for Distance Matrices*   *NeurIPS 2022*
  Top $\sim 2\%$ of submissions

- Spotlight Presentation for *Learning-augmented k-means clustering*   *ICLR 2022*
  Top $\sim 5\%$ of submissions

- Silver Best Paper Award Winner, AdvML @ICML Workshop for   *ICML 2021*
  Adversarial Robustness of Streaming Algorithms through Importance Sampling

- Spotlight Presentation for *Learning-based Support Estimation in Sublinear Time*   *ICLR 2021*
  Top $\sim 4\%$ of submissions

- NSF Graduate Research Fellowship   *2019*

- 1st Place Team and IBM Prize Winner, Hack MIT   *2017*

- Meritorious Team, Mathematical Contest in Modeling   *2017, 2016*

- Two Sigma Prize Winner ($1,000$), Hack MIT   *2016*

- 3rd Place Team, National Basketball Association Hackathon   *2015*

- 1st Place Team, Moody's Mega Math Challenge   *2015*

SELECTED TALKS

- CSAIL + Imagination in Action: AI Frontier & Implications event in celebration *2023*
  of CSAIL's 60th anniversary. Covered by Forbes magazine.

- Basic Algorithms Research Copenhagen at the University of Copenhagen *2023*
  on the paper *Data Structures for Density Estimation*

- UMass Amherst Theory Seminar on the paper *Data Structures for Density Estimation* *2023*

- Google Algorithms Seminar on the paper *Faster Linear Algebra for Distance Matrices* *2022*

- Talk at the Workshop on Algorithms under Uncertainty at FSTTCS 2022 *2022*
  on the paper *Faster Linear Algebra for Distance Matrices*

- Computer Science for Mathematicians Seminar at Harvard Center of Mathematical *2021*
  Sciences and Applications on the paper *Randomized Dimensionality Reduction for Facility Location and Single-Linkage Clustering*

- CMU CS Theory lunch seminar on the paper *Randomized Dimensionality Reduction* *2021*
  for Facility Location and Single-Linkage Clustering

- Workshop on Algorithms for Large Data (Online) on the paper *2021*
  Adversarial Robustness of Streaming Algorithms through Importance Sampling

OTHER PROFESSIONAL ACTIVITIES

- Reviewer for: NeurIPS, ICML, ICLR, STOC, FOCS, SODA, ICALP, Random, ITCS