NOAH AMSEL

noah.amsel@nyu.edu ⋄ https://NoahAmsel.github.io/

Last updated June 18, 2025

EDUCATION

Courant Institute, New York University August 2022 – Present
Ph.D. in Computer Science New York, NY
Advised by Joan Bruna and Chris Musco.

Yale University August 2016 – May 2020
B.S. in Computer Science & Mathematics, with distinction. New Haven, CT
Magna cum laude. Thesis: "Online Vector Balancing in Practice," advised by Dan Spielman.

EXPERIENCE

Adobe, Inc. June 2025 – August 2025
Research Scientist Intern *San Jose, CA*

· Led research project on auxiliary memory systems for agentic AI systems.

Polymathic AI June 2024 – August 2024
Summer Research Intern *New York, NY*

· Contributed to deep learning research on large foundation models for scientific data from heterogeneous domains.
· Led project studying transfer capabilities of models for fluid dynamics problems described by differential equations.

Qualcomm Technologies, Inc. November 2021 – July 2022
Engineer, Corporate Research & Development *New York, NY*

· Stayed on after Reservoir Labs was acquired by Qualcomm and turned into a new R&D division.
· Worked on software package for global non-convex optimization.

Reservoir Labs June 2020 – November 2021
Research Engineer *New York, NY*

· Created algorithm for simulating network performance that achieved order-of-magnitude speed up.
· Deployed our network modeling tool to DOE's Energy Sciences Network.
· Developed framework for designing provably efficient data center networks. Published it at ACM SIGCOMM.

Weizmann Institute of Science June – August 2019
Kupcinet-Getz International Summer School, Research Fellow *Rehovot, Israel*

· Developed novel spectral method for fitting latent tree graphical models of DNA data.
· Proved sufficient condition for the algorithm to succeed and finite sample guarantee.
· Wrote open source implementation that scales to to very large problems with high accuracy.

Facebook, Inc. May – August 2018
Software Engineering Intern *New York, NY*

· Built Bandwidth Estimation model for Adaptive Bitrate Streaming to improve mobile video quality.
· Implemented and tested the model in C++ and Java; refactored ABR code.
· Validated new model in production; found it reduced error by more than half (RMSE).
· Received return offer.

Off the Hook, LLC May – July 2017
Data Analyst / Software Developer *New York, NY*

· Developed computer gambling software in Python based on statistical analysis of horse racing data.
· Evaluated wager opportunities using machine learning trained on data from hundreds of thousands of past races.

PUBLICATIONS

Noah Amsel, David Persson, Christopher Musco, and Robert M. Gower. The polar express: Optimal matrix sign methods and their application to the muon algorithm, 2025a. URL https://arxiv.org/abs/2505.16932.

Noah Amsel, Tyler Chen, Feyza Duman Keles, Diana Halikias, Cameron Musco, Christopher Musco, and David Persson. Quasi-optimal hierarchically semi-separable matrix approximation, 2025b. URL https://arxiv.org/abs/2505.16937.

Gilad Yehudai, Noah Amsel, and Joan Bruna. Compositional reasoning with transformers, rnns, and chain of thought, 2025. URL https://arxiv.org/abs/2503.01544.

Yilun Kuang, Noah Amsel, Sanae Lotfi, Shikai Qiu, Andres Potapczynski, and Andrew Gordon Wilson. Customizing the inductive biases of softmax attention using structured matrices. *International Conference on Machine Learning (ICML), 2025*, 2025. URL https://icml.cc/virtual/2025/poster/45261.

Noah Amsel, Gilad Yehudai, and Joan Bruna. Quality over quantity in attention layers: When adding more heads hurts. In *The Thirteenth International Conference on Learning Representations*, 2025c. URL https://iclr.cc/virtual/2025/poster/27747.

Noah Amsel, Tyler Chen, Feyza Duman Keles, Diana Halikias, Cameron Musco, and Christopher Musco. Fixed-sparsity matrix approximation from matrix-vector products, 2024a. URL https://arxiv.org/abs/2402.09379.

Noah Amsel, Tyler Chen, Anne Greenbaum, Cameron N Musco, and Christopher Musco. Nearly optimal approximation of matrix functions by the lanczos method. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024b. URL https://openreview.net/forum?id=3s8V8QP9XV.

Yariv Aizenbud, Ariel Jaffe, Meng Wang, Amber Hu, Noah Amsel, Boaz Nadler, Joseph T Chang, and Yuval Kluger. Spectral top-down recovery of latent tree models. *Information and Inference: A Journal of the IMA*, 12(3):iaad032, 08 2023. ISSN 2049-8772. doi: 10.1093/imaiai/iaad032. URL https://doi.org/10.1093/imaiai/iaad032.

Jordi Ros-Giralt, Noah Amsel, Sruthi Yellamraju, James Ezick, Richard Lethin, Yuang Jiang, Aosong Feng, Leandros Tassiulas, Zhenguo Wu, Min Yee Teh, and Keren Bergman. Designing data center networks using bottleneck structures. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*, SIGCOMM '21, pages 319–348, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383837. doi: 10.1145/3452296.3472898. URL https://doi.org/10.1145/3452296.3472898.

Ariel Jaffe, Noah Amsel, Yariv Aizenbud, Boaz Nadler, Joseph T. Chang, and Yuval Kluger. Spectral neighbor joining for reconstruction of latent tree models. *SIAM Journal on Mathematics of Data Science (SIMODS)*, 3 (1):113–141, 2021. doi: 10.1137/20M1365715. URL https://doi.org/10.1137/20M1365715.

Noah Amsel, Jordi Ros-Giralt, Sruthi Yellamraju, James Ezick, Brendan von Hofe, Alison Ryan, and Richard Lethin. Computing bottleneck structures at scale for high-precision network performance analysis. *2020 IEEE/ACM Innovating the Network for Data-Intensive Science (INDIS)*, 11 2020. doi: 10.1109/indis51933.2020.00012. URL https://ieeexplore.ieee.org/document/9307175.

William Merrill, Lenny Khazan, Noah Amsel, Yiding Hao, Simon Mendelsohn, and Robert Frank. Finding hierarchical structure in neural stacks using unsupervised parsing. In *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 224–232, Florence, Italy, August 2019. Association for Computational Linguistics. doi: 10.18653/v1/W19-4823. URL https://aclanthology.org/W19-4823.

Yiding Hao, William Merrill, Dana Angluin, Robert Frank, Noah Amsel, Andrew Benz, and Simon Mendelsohn. Context-free transductions with neural stacks. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 306–315, Brussels, Belgium, November 2018. Association for Computational Linguistics. doi: 10.18653/v1/W18-5433. URL https://aclanthology.org/W18-5433.

TALKS

NYC Student Theory Day	The Low-Rank Bottleneck in Attention	2025
Flatiron CCM ML Seminar		2025
NYAS ML Symposium		2024
SIAM Linear Algebra	Fixed-Sparsity Matrix Approximation from Matrix-Vector Products	2024
Temple Num. Anal. Day		2024
Rutgers Theory Seminar	Near-Optimal Approx. of Matrix Functions by the Lanczos Method	2025
SIAM-NNP		2023
ACM SIGCOMM	Designing Data Center Networks Using Bottleneck Structures	2021
IEEE/ACM INDIS at SC20	Computing Bottleneck Structures at Scale	2020

POSTER PRESENTATIONS

ICML	Quality over Quantity in Attention Layers	2025
NeurIPS – *Spotlight!*	Lanczos-FA is Nearly Krylov-Optimal for Rational Functions	2024
ICERM		2023
ASE60 (MIT)		2023
Mihalis Fest (Columbia)		2023
BlackboxNLP @ ACL	Finding Hierarchical Structure in Neural Stacks Using Unsupervised Parsing	2019

TEACHING EXPERIENCE

Teaching Assistant	Algorithms for Machine Learning	*NYU*	2025
Undergraduate Learning Assistant	Algorithms	*Yale University*	2019
Volunteer Math Tutor		*Top Honors*	2013 – 2016

AWARDS

NSF Graduate Research Fellowship	2024
Phi Beta Kappa	2020
Intel Science Talent Search Semifinalist	2016

SERVICE

Organizing NYC Machine Learning Speaker Series at the Flatiron Institute (2023 – 2025).

Reviewing JMLR (2024).

Outreach NYU AI Summer School Instructor (2024).