ESHAAN NICHANI

eshnich@princeton.edu • http://eshaannichani.com

EDUCATION

Princeton University • Princeton, NJ   September 2021 – Present
PhD Candidate in Electrical & Computer Engineering
Advisors: Jason D. Lee, Yuxin Chen
GPA: 4.0/4.0

Massachusetts Institute of Technology • Cambridge, MA   June 2020 – May 2021
Masters of Engineering in EECS
Thesis: *An Empirical and Theoretical Analysis of the Role of Depth in Convolutional Neural Networks.*
Advisor: Caroline Uhler

Massachusetts Institute of Technology • Cambridge, MA   September 2016 – May 2020
Bachelors of Science in Mathematics
Bachelors of Science in Computer Science and Engineering
GPA: 5.0/5.0

PUBLICATIONS & PREPRINTS

(* denotes equal contribution or alphabetical authorship)

- **How Transformers Learn Causal Structure with Gradient Descent.**
  Eshaan Nichani, Alex Damian, Jason D. Lee.
  International Conference on Machine Learning (ICML), 2024.

- **Learning Hierarchical Polynomials with Three-Layer Neural Networks.**
  Zihao Wang, **Eshaan Nichani**, Jason D. Lee.
  International Conference on Learning Representations (ICLR), 2024.

- **Metastable Mixing of Markov Chains: Efficiently Sampling Low Temperature Exponential Random Graphs.**
  Guy Bresler*, Dheeraj Nagaraj*, **Eshaan Nichani***.
  Annals of Applied Probability, 2024.

- **Provable Guarantees for Nonlinear Feature Learning in Three-Layer Neural Networks.**
  Eshaan Nichani, Alex Damian, Jason D. Lee.
  Advances in Neural Information Processing Systems (NeurIPS), 2023 (Spotlight).

- **Smoothing the Landscape Boosts the Signal for SGD: Optimal Sample Complexity for Learning Single Index Models.**
  Alex Damian, **Eshaan Nichani**, Rong Ge, Jason D. Lee.
  Advances in Neural Information Processing Systems (NeurIPS), 2023 (Oral).

- **Fine-Tuning Language Models with Just Forward Passes.**
  Sadhika Malladi, Tianyu Gao, **Eshaan Nichani**, Alex Damian, Jason D. Lee, Danqi Chen, Sanjeev Arora.
  Advances in Neural Information Processing Systems (NeurIPS), 2023 (Oral).

- **Self-Stabilization: The Implicit Bias of Gradient Descent at the Edge of Stability.**
  Alex Damian*, **Eshaan Nichani***, Jason D. Lee.
  International Conference on Learning Representations (ICLR), 2023.

- **Identifying good directions to escape the NTK regime and efficiently learn low-degree plus sparse polynomials.**
  Eshaan Nichani, Yu Bai, Jason D. Lee.
  Advances in Neural Information Processing Systems (NeurIPS), 2022.

- **Causal Structure Discovery between Clusters of Nodes Induced by Latent Factors.**
  Chandler Squires, Annie Yun, **Eshaan Nichani**, Raj Agrawal, Caroline Uhler.
  Conference on Causal Learning and Reasoning (CLeaR), 2022.

- **Increasing Depth Leads to U-Shaped Test Risk in Over-parameterized Convolutional Networks.**
  Eshaan Nichani*, Adityanarayanan Radhakrishnan*, Caroline Uhler.
  Workshop on Overparameterization: Pitfalls and Opportunities, ICML, 2021.

- **On Alignment in Deep Linear Neural Networks.**
  Adityanarayanan Radhakrishnan*, **Eshaan Nichani***, Daniel Bernstein, Caroline Uhler.
  Workshop on Overparameterization: Pitfalls and Opportunities, ICML, 2021.

- **Adaptive diagonal curvature: a quasi-newton method for stochastic optimization**
  David Saxton, **Eshaan Nichani**.
  Workshop on Beyond First Order Methods in Machine Learning Systems, ICML, 2020.

- **Assessment of circulant copy number variant detection for cancer screening**
  Bhuvan Molparia, **Eshaan Nichani**, Ali Torkamani.
  PLoS ONE, 2017.

Honors and Awards

- Finalist for Two-Sigma PhD Fellowship — 2024
- DoD National Defense Science and Engineering Graduate Fellowship — 2021 – 2024
- Putnam Mathematics Competition – Rank N2 (top 25 in US) — 2016, 2019
- Romanian Master of Mathematics International Competition – 1st place individual worldwide (Gold medal) — 2016
- USA Mathematical Olympiad – Honorable Mention — 2016

Presentations

- *How Transformers Learn Causal Structure with Gradient Descent.*
  - Poster presentation at ICML 2024, Vienna, July 2024.
  - Citadel PhD Colloquium, New York, May 2024.
  - Flatiron Institute, New York, March 2024.
  - Siemens Research, Princeton, February 2024.
  - Broad Institute, virtual, February 2024.
  - Two Sigma PhD Fellowship Finalist Reception, virtual, February 2024.

- *Provable Guarantees for Nonlinear Feature Learning in Three Layer Neural Networks.*
  - Spotlight Poster presentation at NeurIPS 2023, New Orleans, December 2023.
  - NDSEG Fellows Conference, July 2023.
  - Princeton ML Theory Summer School, Poster presentation, June 2023.

- *Towards a Theory of Shallow Learning: End-to-end sample complexity and feature learning guarantees for gradient descent. – ECE General Exam, April 2023.*
- *Self-Stabilization: The Implicit Bias of Gradient Descent at the Edge of Stability.*
  - Poster presentation at ICLR 2023, Kigali, May 2023.
  - Alg-ML seminar, Princeton, December 2022.
  - Optimization for Machine Learning workshop at NeurIPS 2022, New Orleans, December 2022.

- *Identifying good directions to escape the NTK regime and efficiently learn low-degree plus sparse polynomials.*
  - Poster presentation at NeurIPS 2022, New Orleans, December 2022.
  - NYU, June 2022.

- *Increasing depth leads to U-shaped test risk in over-parameterized convolutional networks. – Poster presentation at ICML 2021 Workshop on Overparameterization: Pitfalls and Opportunities, virtual, July 2021.*
- *On alignment in deep linear neural networks. – Poster presentation at ICML 2021 Workshop on Overparameterization: Pitfalls and Opportunities, virtual, July 2021.*
- *When does gradient descent preserve alignment of layers of linear neural networks? – Poster presentation at MIT Foundations of Data Science Workshop, January 2020.*

Work Experience

Flatiron Institute CCM – Summer Research Associate — June 2024 – August 2024
New York, NY

- Designed and studied a theoretical model for factual recall in large language models. Advised by Dr. Alberto Bietti.

DeepMind – Research Engineering Intern — June 2019 – August 2019
London, United Kingdom

- Researched and developed a novel general-purpose optimization algorithm for training deep neural networks.

Asana – Software Engineering Intern — June 2018 – August 2018
San Francisco, CA

Teaching Experience

Teaching Assistant – Algorithms for Inference (MIT 6.438) — September 2020 – December 2020
Teaching Assistant – Design and Analysis of Algorithms (MIT 6.046) — September 2019 – May 2020

Mathematics Coach – MIT Global Teaching Labs & MISE Foundation   January 2020
Accra, Ghana
- Coached team of students who represented Ghana at the 2020 International Mathematical Olympiad.
- Taught competition mathematics & problem solving skills at 10+ high schools and middle schools in Accra.

Academic Service

Reviewing: ICML: 2022, 2023, 2024; NeurIPS: 2022, 2023 (Outstanding Reviewer Award), 2024; JMLR; OPT 2023 workshop (NeurIPS 2023); M3L workshop (NeurIPS 2023); TF2M workshop (ICML 2024); HiLD workshop (ICML 2024)

Mentoring:
- Zihao Wang (Peking University undergrad), 2023.

Leadership:
- Princeton Alg-ML Seminar, co-organizer (2023-24 academic year).

Programming Languages/Frameworks

- Python, JAX, PyTorch, TensorFlow, numpy.

Graduate Coursework

Princeton: Optimization for Machine Learning, Deep Learning Theory, Foundations of Reinforcement Learning, Mathematical Understanding of Deep Learning, Theoretical Machine Learning, Probability in High Dimension, Statistical Theory and Methods,Theory of Deep Weakly Supervised Learning.

MIT: Theoretical Foundations for Deep Learning, Mathematical Statistics: A Non-Asymptotic Approach, Graph Theory and Additive Combinatorics, Statistical Learning Theory, Bayesian Modeling and Inference, Randomized Algorithms, Theory of Probability, Inference and Information, Machine Learning, Advanced Algorithms.