

Stephen Zhang

Email : stephenn.zhang@mail.utoronto.ca

EDUCATION

- **Ph.D. Mathematics, Direct-Entry Stream** University of Toronto
Advisor: Vardan Papyan Fall 2020 – Present
- **Honours B.Sc Mathematics** Victoria College, University of Toronto
Mathematics Specialist, Computer Science Major, Statistics Minor; GPA: 3.94/4.00 Fall 2015 – Winter 2020

PUBLICATIONS & PREPRINTS

* denotes equal contribution

- [1] **Sparsest Models Elude Pruning: An Exposé of Pruning’s Current Capabilities** [arxiv]
Stephen Zhang, Vardan Papyan
International Conference in Machine Learning (ICML), 2024
- [2] **OATS: Outlier-Aware Pruning Through Sparse and Low Rank Decomposition** [arxiv]
Stephen Zhang, Vardan Papyan
International Conference on Learning Representations (ICLR), 2025
- [3] **Low-Rank is Required for Pruning LLMs** [OpenReview]
Stephen Zhang, Vardan Papyan
ICLR Workshop on Sparsity in LLMs (SLLM), 2025
- [4] **Attention Sinks and Outlier Features: A ‘Catch, Tag, and Release’ Mechanism for Embeddings** [arxiv]
Stephen Zhang*, Mustafa Khan*, Vardan Papyan
Preprint.

RESEARCH EXPERIENCE

- **Ph.D. Research** University of Toronto
Advisor: Vardan Papyan Fall 2020-Present
 - **Research Interests:** Sparse training for deep-learning, post-train pruning, sparse and low-rank structures, sparsity for acceleration and compression, sparse autoencoders.
- **Summer Research Project** University of Toronto
Project Supervisor: Jeremy Quastel Summer 2021
 - **Project Description:** Demonstrate that the kernel of the Kardar-Parisi-Zhang (KPZ) equation in a half-space with a wall parameter satisfied a version of the Kadomtsev-Petviashvili (KP) equation.
- **NSERC USRA Project** University of Toronto
Project Supervisor: Jeremy Quastel Summer 2019
 - **Project Description:** Focused on stochastic partial differential equations. Learned about topics such as Brownian motion, Ito calculus, and Feller processes.

TEACHING & OUTREACH

- **Course Instructor** University of Toronto
MAT292 - Ordinary Differential Equations Fall 2022
 - **Lectures:** Created, designed, and delivered lectures for a class size of 132 EngSci students.
 - **Assessments:** Assisted in creating the course assessments which included a quiz, midterm, and final exam.
- **Math Mentorship Program** University of Toronto
Volunteer Mentor Winter 2022
 - **Math Mentor:** Volunteered as a math mentor for high school students who wanted to learn about research in mathematics.
 - **Curriculum Design:** Designed a curriculum with minimal prerequisites to learn the basics of deep learning.
- **Teaching Assistant** University of Toronto
Hosted weekly tutorials and office hours. Performed exam invigilation and grading. Fall 2020 - Present
 - **MAT1510:** Deep Learning: Theory & Data Science
 - **MAT337:** Introduction to Real Analysis
 - **MAT334:** Complex Variables
 - **MAT235:** Multivariable Calculus
 - **MAT223:** Linear Algebra I
 - **MAT188:** Linear Algebra
 - **MAT137:** Calculus with Proofs

SCHOLARSHIPS AND AWARDS

Blyth Fellowship	2020
<i>University of Toronto, CAD \$10,000</i>	
NSERC USRA	Summer 2019
<i>Natural Sciences and Engineering Research Council of Canada, CAD \$6,150</i>	
Professor R K Arnold Scholarship II	2019
<i>Victoria College, University of Toronto, CAD \$1,000</i>	
Professor R K Arnold Scholarship I	2018
<i>Victoria College, University of Toronto, CAD \$1,000</i>	
Mary Isabel (Park) Hodgkinson Scholarship I	2017
<i>Victoria College, University of Toronto, CAD \$1,000</i>	
Dean's List Scholar	2016, 2017, 2018, 2019
<i>University of Toronto</i>	

RELEVANT COURSEWORK

MAT1525: Computational Inverse Problems	MAT1508: Applied Nonlinear Equations
MAT1128: Random Matrices and Random Planar Geometry	MAT1951: The Sherrington-Kirkpatrick Model
MAT1950: Integrable Probability and KPZ Universality	MAT1510: Deep Learning: Theory & Data Science
MAT1850: Linear Algebra and Optimization	MAT1752: Topics in Quantum Information Theory
MAT1000: Real Analysis I	MAT1001: Real Analysis II
MAT1600: Mathematical Probability I	MAT1601: Mathematical Probability II
CSC373: Algorithm Design, Analysis & Complexity	CSC2516: Neural Nets and Deep Learning