THE UNIVERSITY OF HONG KONG

Institute of Mathematical Research
Department of Mathematics

Seminar

High-dimensional Optimization in Machine Learning with Applications to Scaling Limits and Compute-Optimal Neural Scaling Laws

Prof. Courtney Paquette

McGill University

Abstract

Given the massive scale of modern ML models, we now only get a single shot to train them effectively. This restricts our ability to test multiple architectures and hyper-parameter configurations. Instead, we need to understand how these models scale, allowing us to experiment with smaller problems and then apply those insights to larger-scale models. In this talk, I will present a framework for analyzing scaling laws in stochastic learning algorithms using a power-law random features model, leveraging high-dimensional probability and random matrix theory. I will then use this scaling law to address the compute-optimal question: How should we choose model size and hyper-parameters to achieve the best possible performance in the most compute-efficient manner? Additionally, I will introduce a scaling limit commonly seen in ML optimization algorithms which has origins in statistical physics and I will highlight several promising research directions in scaling laws that remain underexplored but offer significant potential.

Date:	September 18, 2025 (Thursday)
Time:	9:00 – 10:00 am
Venue:	ZOOM: https://hku.zoom.us/j/ Meeting ID: 955 3432 6925 Password: 328613

All are welcome