THE UNIVERSITY OF CHICAGO | THE COMMITTEE ON COMPUTATIONAL AND APPLIED MATHEMATICS

Dissertation Defense:

Rong Jiang

Computational and Applied Mathematics
University of Chicago

"Nonparametric Statistics Meets Sequential Decision-Making: Minimax Optimality And Adaptivity"

Friday, April 10, 2026, at 9:30 AM
Location: Jones 111, 5747 S. Ellis

ABSTRACT

Sequential decision-making lies at the core of modern statistics and machine learning, with applications ranging from clinical trials to online recommendation systems. Classical theory typically assumes that learning algorithms can update their policies after every observation. In many contemporary experiments, however, decisions can only be revised at a small number of predetermined times due to operational constraints. Such batching restrictions limit adaptivity and inevitably affect statistical performance, raising a fundamental question: how much efficiency is lost, and how many policy updates are required to achieve optimal learning?

This thesis investigates this question in the framework of nonparametric contextual bandits with smooth reward functions. We first establish a minimax regret lower bound under batching and propose a novel batched learning algorithm that achieves the optimal regret rate up to logarithmic factors. Our results show that a nearly constant number of policy updates suffices to attain the optimal regret achievable in the fully online setting.

We then study the additional challenge of adapting to an unknown margin parameter. While adaptation is cost-free in the fully online regime, batching introduces a genuine statistical barrier. To quantify this effect, we introduce the regret inflation criterion—the ratio between the regret of an adaptive algorithm and that of an oracle that knows the margin parameter. We show that the optimal regret inflation grows polynomially with the time horizon T, with an exponent characterized by the solution of a convex optimization problem. The optimizer of this problem also directly determines the batch schedule and exploration strategy of a rate-optimal algorithm.

Together, these results provide a complete characterization of the statistical cost of limited adaptivity in nonparametric bandits learning. While batching alone can degrade performance and unknown complexity parameters can further inflate regret, this thesis shows that a modest degree of adaptivity—on the order of log log T updates—suffices to recover the optimal learning rates achievable in fully online settings.