



THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY

Department of Mathematics

**SEMINAR ON DATA SCIENCE
AND APPLIED MATHEMATICS**

**Sampling Complexity of Temporal-Difference and
Proximal Policy Optimization in RKHS**

By

Prof. Liang Ding

School of Data Science, Fudan University

Abstract

We revisit Proximal Policy Optimization (PPO) from a function-space perspective. Our analysis decouples policy evaluation and improvement in a reproducing kernel Hilbert space (RKHS): (i) A kernelized temporal-difference (TD) critic performs efficient RKHS-gradient updates using only one-step state-action transition samples; (ii) a KL-regularized, natural-gradient policy step exponentiates the evaluated action-value, recovering a PPO/TRPO-style proximal update in continuous state-action spaces. We provide non-asymptotic, instance-adaptive guarantees whose rates depend on RKHS entropy, unifying tabular, linear, Sobolev, Gaussian, and Neural Tangent Kernel (NTK) regimes, and we derive a sampling rule for the proximal update that ensures the optimal $k^{-1/2}$ convergence rate for stochastic optimization. Empirically, the theory-aligned schedule improves stability and sample efficiency on common control tasks (e.g., CartPole, Acrobot), while our TD-based critic attains favorable throughput versus a GAE baseline. Altogether, our results place PPO on a firmer theoretical footing beyond finite-dimensional assumptions and clarify when RKHS-proximal updates with kernel-TD critics yield global policy improvement with practical efficiency.

Keywords: Kernel method, Temporal-Difference, Proximal Policy

Bio: *Liang Ding is an Assistant Professor in the School of Data Science at Fudan University. Before joining Fudan, he was a Postdoctoral Research Associate jointly appointed by the Department of Industrial and Systems Engineering and the Texas A&M Institute of Data Science at Texas A&M University. He received his Ph.D. in Operations Research from the Hong Kong University of Science and Technology, supported by the Hong Kong PhD Fellowship and the Overseas Scholarship. His research interests include simulation optimization, kernel methods, and statistical learning.*

Date : 13 November 2025 (Thursday)
Time : 2:00p.m.-3:00p.m.
Venue : Room 4621 (near Lift 31/32)

All are welcome