

# Workshop on Stochastic Control, Financial Technology, and Machine Learning

in Conjunction with

the 60th Birthday Celebration of Prof. Xunyu Zhou

The Hong Kong Polytechnic University

December 11 – 14, 2025

## Conference Venue

- Room GH201, The Hong Kong Polytechnic University (PolyU)

## Local Organizing Committee:

- Kexin Chen (PolyU), Nan Chen (CUHK), Wai Ki Ching (HKU), Min Dai (PolyU, Chair), Guanxing Fu (PolyU), Xuedong He (CUHK), Yanwei Jia (CUHK), Zhaoli Jiang (PolyU), Xun Li (PolyU), Zuoquan Xu (PolyU), Chen Yang (CUHK), Xiang Yu (PolyU)

## Sponsors

- Research Centre for Quantitative Finance & Department of Applied Mathematics, The Hong Kong Polytechnic University

## Content

- Schedule (Pages 2-5)
- Daily Detailed Program (Pages 6-9)
- Talk Information (Pages 10-27)
- Map (Page 28)

## Contact

- Min Dai (mindai@polyu.edu.hk); Guanxing Fu (guanxing.fu@polyu.edu.hk); Zhaoli Jiang (zhaoli.jiang@polyu.edu.hk); Zuoquan Xu (maxu@polyu.edu.hk)

**Reminder:** A QR entry code is required to access the campus. Your QR code has been sent to you via email. If you are unable to find it, please contact the organizers for assistance.

## Schedule

**Thursday, 11 December, Day 1**  
**Conference Venue: Room GH201**

Time	Session	Chaired by
8:30 – 17:20	Registration at GH201	
8:50 – 9:00	Opening Ceremony	Defeng Sun
9:00 – 9:30	Shige Peng	
9:30 – 10:00	George Yin	
10:00 – 10:30	Wei Lin	
10:30 – 11:00	Coffee Break <sup>1</sup>	
11:00 – 11:30	Nizar Touzi	
11:30 – 12:00	Jie Xiong	
12:00 – 12:30	Jianfeng Zhang	
12:30 – 14:00	Lunch <sup>1,2</sup>	
14:00 – 14:30	Huyên Pham	
14:30 – 15:00	Jiro Akahori	
15:00 – 15:30	Moris Strub	
15:30 – 16:10	Poster Session 1 & Coffee Break <sup>1</sup>	
16:10 – 16:30	Wai Ki Ching	Xunyu Zhou
16:30 – 16:50	Xuedong He	
16:50 – 17:10	Yilie Huang	
17:10 – 17:30	Hanqing Jin	
17:30 – 17:50	Zuoquan Xu	
18:10 – 20:00	Reception Dinner <sup>2,3</sup>	

1. Coffee Break and Lunch are provided to all invited speakers, registered attendees, organizers, and student helpers.
2. Lunch and Reception Dinner are located at Ju Yin House Seafood Restaurant, 4/F, Communal Building (VA). 聚賢樓飯店，校內文康樓4樓。
3. Reception Dinner is offered to all invited speakers, registered attendees, and organizers.

**Friday, 12 December, Day 2**  
**Conference Venue: Room GH201**

Time	Session	Chaired by
9:00 – 9:30	David Yao	
9:30 – 10:00	Ruodu Wang	
10:00 – 10:30	Wenbin Tang	
10:30 – 11:00	Coffee Break <sup>1</sup>	
11:00 – 11:30	Xin Guo	
11:30 – 12:00	Yu-Jui Huang	
12:00 – 12:40	Jin Ma	
12:40 – 14:10	Lunch <sup>1,2</sup>	
14:10 – 14:40	Jan Oblój	
14:40 – 15:10	Xiangyu Cui	
15:10 – 15:40	Yufei Zhang	
15:40 – 16:20	Poster Session 2 & Coffee Break <sup>1</sup>	
16:20 – 16:40	Chun Hung Chiu	Xunyu Zhou
16:40 – 17:00	Sang Hu	
17:00 – 17:20	Yanwei Jia	
17:20 – 17:40	Xun Li	
17:40 – 18:00	Wei Wei	
18:20 – 20:20	Banquet <sup>3</sup>	

1. Coffee Break and Lunch are provided to all registered attendees, organizers, and student helpers.
2. Lunch venue is located at Ju Yin House Seafood Restaurant, 4/F, Communal Building (VA). 聚賢樓飯店，校內文康樓4樓。
3. Banquet is provided to all invited speakers, conference organizers, as well as registered academic and industry attendees, and registered students who have paid the banquet fee. It will be held at SJ Cuisine restaurant, 2/F, Harbour Crystal Centre, 100 Granville Road, Tsim Sha Tsui. It is a five-minute walk from PolyU campus (exit from core D), linked by a footbridge. 港晶軒飯店，尖沙咀加連威老道100號港晶中心2樓。

**Saturday, 13 December, Day 3**  
**Conference Venue: Room GH201**

Time	Session	Chaired by
9:00 – 9:30	Rama Cont	
9:30 – 10:00	Andrew Lim	
10:00 – 10:30	Jin Liang	
10:30 – 11:10	Poster Session 3 & Coffee Break <sup>1</sup>	
11:10 – 11:30	Min Dai	
11:30 – 11:50	Guanxing Fu	
11:50 – 12:10	Lingfei Li	
12:10 – 12:30	Shuaijie Qian	
12:30 – 12:50	Chen Yang	
12:50 – 14:00	Lunch <sup>1,2</sup>	
14:00 – 17:30	Free Discussion	

1. Coffee Break and Lunch are provided to all registered attendees, organizers, and student helpers.
2. Lunch venue is located at Ju Yin House Seafood Restaurant, 4/F, Communal Building (VA). 聚賢樓飯店，校內文康樓4樓。

**Sunday, 14 December, Day 4**  
**Conference Venue: Room GH201**

Time	Session	Chaired by
9:00 – 9:30	Jose Blanchet	
9:30 – 10:00	Luhao Zhang	
10:00 – 10:30	Renyuan Xu	
10:30 – 11:00	Coffee Break <sup>1</sup>	
11:00 – 11:20	Nan Chen	
11:20 – 11:40	Xuefeng Gao	
11:40 – 12:00	Wei Jiang	
12:00 – 12:20	Xiang Yu	
12:20 – 14:00	Lunch <sup>1,2</sup>	
14:00 – 14:30	Ulrich Horst	
14:30 – 15:00	Carole Bernard	
15:00 – 15:30	Fausto Gozzii	
15:30 – 15:40	Address by Xunyu Zhou	

1. Coffee Break and Lunch are provided to all registered attendees, organizers, and student helpers.
2. Lunch venue is located at Ju Yin House Seafood Restaurant, 4/F, Communal Building (VA). 聚賢樓飯店，校內文康樓4樓。

# Daily Detailed Program

Thursday, 11 December, Day 1

Time	Session	Talk Title
9:00 – 9:30	Shige Peng	Nonlinear Expectations in Machine Learning
9:30 – 10:00	George Yin	Computational Nonlinear Filtering Using A Deep Learning Approach
10:00 – 10:30	Wei Lin	Recent Advance in Nonlinear Control with Input Delay: From Deterministic Systems to Stochastic Systems
11:00 – 11:30	Nizar Touzi	Bridging Schrödinger and Bass for Generative Diffusion Modelling: A Theory
11:30 – 12:00	Jie Xiong	Mean-field Stochastic Linear Quadratic Control Problem with Random Coefficients
12:00 – 12:30	Jianfeng Zhang	Reinforcement Learning for Stochastic Controls with Model Uncertainty
14:00 – 14:30	Huyên Pham	Bridging Schrödinger and Bass for Generative Diffusion Modelling: Applications
14:30 – 15:00	Jiro Akahori	Randomization of Option Price under the Binomial Model
15:00 – 15:30	Moris Strub	Optimal Investment to Reach a Financial Goal: A Stochastic Control Framework
15:30 – 16:10	Poster Session 1	
16:10 – 16:30	Wai Ki Ching	Online Portfolio Selection Problems
16:30 – 16:50	Xuedong He	Dynamic Asset Pricing with $\alpha$ -MEU Model
16:50 – 17:10	Yilie Huang	Adaptive Reparameterized Time for Diffusion Sampling: A Reinforcement Learning Approach to Timestep Allocation
17:10 – 17:30	Hanqing Jin	Optimal Stopping without Time Consistency
17:30 – 17:50	Zuoquan Xu	Comparison Theorems for Multi-Dimensional BSDEs with Jumps and Applications to Stochastic LQ Control

## Poster Session 1: 15:30-16:10

- Niccolò Bagnoli, Carlo Sala. *Recovering the Physical Measure from Options: A Non-Parametric, Constraint-Based Approach.*
- Min Dai, Hanqing Jin, Xi Yang. *Data-Driven Option Pricing.*
- Yucheng Guo, Qinxin Yan. *Particle Systems with Local Interactions via Hitting Times and Cascades on Graphs.*
- Xuefeng Gao, Mengying He, Xuedong He. *Factor-Based Conditional Diffusion Model for Portfolio Optimization.*
- Samuel N. Cohen, Jackson Hebner, Deqing Jiang, Justin Sirignano. *Neural Actor-Critic Methods for Hamilton–Jacobi–Bellman PDEs: Asymptotic Analysis and Numerical Studies.*
- Devang Sinha, Shashi Jain, Srikanth K. Iyer. *Efficient Simulation of Correlated Defaults via Sector Apportionment.*

**Friday, 12 December, Day 2**

Time	Session	Talk Title
9:00 – 9:30	David Yao	Production Planning with Risk Hedging and Clean-tech Adoption
9:30 – 10:00	Ruodu Wang	Risk Aversion and Risk-Insurance Parity
10:00 – 10:30	Wenpin Tang	Some Recent Progress on Diffusion Models: Conditioning and Projection
11:00 – 11:30	Xin Guo	From LLM to RL and Diffusion Models, via (rough) Differential Equations
11:30 – 12:00	Yu-Jui Huang	Mean-Field Langevin Diffusions with Density-dependent Temperature
12:00 – 12:40	Jin Ma	A Short XYZ of XunYu Zhou
14:10 – 14:40	Jan Oblój	Robust Pricing and Hedging of American Options in Continuous Time
14:40 – 15:10	Xiangyu Cui	Reinforcement Learning Mean-Variance Policy under Factor Model
15:10 – 15:40	Yufei Zhang	Deterministic Policy Gradient Methods in Continuous Time and Space
15:40 – 16:20	Poster Session 2	
16:20 – 16:40	Chun Hung Chiu	Buy-Online-Pick-Up-At-Store Benefits Supply Chains Considering Supplier Encroachment
16:40 – 17:00	Sang Hu	Equilibrium Policy on Dividend and Capital Injection under Time-inconsistent Preferences
17:00 – 17:20	Yanwei Jia	Accuracy of Discretely Sampled Stochastic Policies in Continuous-time Reinforcement Learning
17:20 – 17:40	Xun Li	Optimal Consumption under a Drawdown Constraint Over a Finite Horizon
17:40 – 18:00	Wei Wei	DPM: Deep Penalty Methods: A Class of Deep Learning Algorithms for Solving High Dimensional Optimal Stopping Problems

**Poster Session 2: 15:40-16:20**

- Boris Baros. *Mean-Field Generalisation Bounds for Learning Controls in Stochastic Environments.*
- [Eliza RezvANOVA](#), Chiheb Ben Hammouda, Erik von Schwerin, Raul Tempone. *A Lagrangian Relaxation for Continuous-Time Stochastic Optimal Control for Uncertainty Control of Coupled Power Systems with Time Delays.*
- [Jinye Du](#), Zhongfei Li, Moris S. Strub. *How Bad is Myopia for a Mean-Variance Investor?.*
- [Jiming Pan](#), Xiao Qiao, Moris S. Strub. *Information Acquisition, Portfolio Choice, and Diversification.*
- [Saifeddine Ben Naamia](#), Christian Bayer, Erik von Schwerin, Raul Tempone. *Stochastic Optimal Control with Partial Information.*
- [Eya Ben Amar](#), Nadhir Ben Rached, Raúl Tempone. *Hierarchical Importance Sampling for Estimating Occupation Times of SDE Solutions.*

**Saturday, 13 December, Day 3**

Time	Session	Talk Title
9:00 – 9:30	Rama Cont	Causal Transport for Continuous-time Stochastic Processes
9:30 – 10:00	Andrew Lim	Optimal Trade Execution with Learning
10:00 – 10:30	Jin Liang	Free Boundary Problems for Measuring Credit Rating Migration Risks
10:30 – 11:10	Poster Session 3	
11:10 – 11:30	Min Dai	Connecting Singular Control and Optimal Stopping in Multi-Asset Portfolio Selection with Transaction Costs
11:30 – 11:50	Guanxing Fu	Multiasset Liquidation in Dark Pools
11:50 – 12:10	Lingfei Li	Value-Aware Model Learning for Stochastic Control
12:10 – 12:30	Shuaijie Qian	Optimal Principal-agent Contract Design via HJB Equation
12:30 – 12:50	Chen Yang	Periodic Evaluation with Non-Concave Utility

**Poster Session 3: 10:30-11:10**

- Shanyu Han, Yang Liu, Xiang Yu. *Risk-sensitive Reinforcement Learning Based on Convex Scoring Functions.*
- Purba Banerjee, Srikanth Iyer, Shashi Jain. *Robust Hedging of Path-dependent Options Using A Min-Max Algorithm.*
- Wong Hoi Yat. *Optimal Switching via Randomized Formulation: A Numerical Approach.*
- Yang Feng, Jinxia Zhu, Tak Kuen Siu. *How Might Model Uncertainty and Transaction Costs Impact Cash Payout Strategies?.*
- Li Xia, Zhihui Yu. *Mean-Variance Optimization and Algorithm for Finite-Horizon Markov Decision Processes.*
- Yijie Huang, Mengge Li, Xiang Yu, Zhou Zhou. *Continuous-Time Reinforcement Learning for Optimal Switching over Multiple Regimes .*

**Sunday, 14 December, Day 4**

Time	Session	Talk Title
9:00 – 9:30	Jose Blanchet	Bayesian Distributionally Robust Merton Problem with Nonlinear Wasserstein Projections
9:30 – 10:00	Luhao Zhang	Neural Network Approach for Discrete-Time Portfolio Control with Provable Guarantees
10:00 – 10:30	Renyuan Xu	Schrödinger Bridge for Generative AI: Relaxed Formulation and Convergence Analysis
11:00 – 11:20	Nan Chen	Optimal Bonding Function in Decentralized Exchanges: A Convex Analysis Approach
11:20 – 11:40	Xuefeng Gao	Data-driven Generative Simulation of SDEs using Diffusion Models
11:40 – 12:00	Wei Jiang	Optimal Time-Consistent Fiscal and Monetary Policies: A Mathematical Macro-Finance Approach
12:00 – 12:20	Xiang Yu	Mean Field Control with Poissonian Common Noise: A Pathwise Compactification Approach
14:00 – 14:30	Ulrich Horst	Optimal Exploitation of Random Resources
14:30 – 15:00	Carole Bernard	Preference Robust Distortion Risk Measures
15:00 – 15:30	Fausto Gozzii	On Mean-Field Control and Games in Infinite Dimension
15:30 – 15:40	Address by Xunyu Zhou	

## Talk Information

### Randomization of Option Price under the Binomial Model

Jiro Akahori  
Ritsumeikan University

This talk presents a “Bayesian extension” of the classical discrete market model, specifically the Cox-Ross-Rubinstein (CRR) binomial model, aimed at incorporating parameter uncertainty and path-dependent dynamics into derivative pricing.

In the classical CRR framework, the risk-neutral probability is constant and determined solely by the up and down steps. We propose a generalized model where the state-dependent up and down factors ( $u_{r,w}$  and  $d_{r,w}$ ) are introduced, realizing a “reinforced scheme” analogous to the Pólya Urn process. This approach replaces the constant parameters with variables whose dynamics satisfy the exchangeability condition, effectively randomizing the underlying dynamics.

The core result is the complete characterization of all solutions satisfying this exchangeability. We prove that the sequence of state-dependent risk-neutral probabilities is uniquely parameterized by two probability measures on  $[0, 1]$ . Consequently, the derived option price is given by the expectation of the classical CRR price with respect to these measures, representing a randomized binomial option price.

We conclude by discussing the financial implications of this measure-parameterized solution and briefly address its continuous-time limit, which yields the Black-Scholes formula with a randomized volatility parameter studied by Takaoka (2006). This model provides a theoretically grounded method for incorporating path-dependence and uncertainty of the parameters (volatility and market price of risk) into derivative valuation.

### Preference Robust Distortion Risk Measures

Carole Bernard  
Grenoble Ecole de Management

We introduce a framework for preference-robust decision making when preferences over risk are modeled through generalized distortion risk measures. Unlike distributional robustness, our approach addresses ambiguity in the risk functional itself. We construct ambiguity sets on distortion (weight) functions using the Wasserstein distance and Bregman–Wasserstein divergences, and derive closed-form expressions for the worst- and best-case distortion risk measures. We further extend the framework to Rank-Dependent Expected Utility, yielding preference-robust behavioral models. This is joint work with Silvana Pesenti (University of Toronto).

# Bayesian Distributionally Robust Merton Problem with Nonlinear Wasserstein Projections

Jose Blanchet  
Stanford University

We revisit Merton’s continuous-time portfolio selection through a data-driven, distributionally robust lens. Our aim is to tap the benefits of frequent trading over short horizons; while acknowledging that drift is hard to pin down whereas volatility can be screened using realized or implied measures for appropriately selected assets. Rather than time-rectangular distributional robust control, which replenishes adversarial power at every instant and induces over-pessimism, we place a single ambiguity set on the drift prior within a Bayesian Merton model. This prior-level ambiguity preserves learning and tractability: a minimax swap reduces the robust control to optimizing a nonlinear functional of the prior, enabling Karatzas and Zhao’s closed-form evaluation for each candidate prior. We then characterize small-radius worst-case priors under Wasserstein uncertainty via an explicit asymptotically optimal pushforward of the nominal prior, and we calibrate the ambiguity radius through a nonlinear Wasserstein projection tailored to the Merton functional. Synthetic and real-data studies demonstrate reduced pessimism relative to DRC and improved performance over myopic DRO–Markowitz under frequent rebalancing. Joint work with Jiayi Chen, Hao Liu, and Yang Liu.

## Optimal Bonding Function in Decentralized Exchanges: A Convex Analysis Approach

Nan Chen  
The Chinese University of Hong Kong

Decentralized exchanges rely on automated market makers to provide liquidity passively. This exposes liquidity providers to both stale-price arbitrage and reverse-trade arbitrage. In this paper, we use tools from the calculus of variations and convex optimization to study the optimal design of bonding functions, the core mechanism of AMMs, in the presence of these two forms of arbitrage. We also discuss the economic implications of our findings. This is a joint work with Agostino Capponi (Columbia), Ling Qin (Shanghai Tech University), and Zhou Yang (South China Normal University).

## Online Portfolio Selection Problems

Wai Ki Ching  
The University of Hong Kong

Online portfolio selection (OLPS) is attracting a lot of attention due to its efficiency and practicability in deriving optimal investment strategies in the investment activities where the market information is constantly renewed in a short period. The first issue in OLPS is the prediction

of the future returns of risky assets given historical data. The second issue is to obtain optimal investment strategies based on the predictions. In this talk, we present some models and adaptive algorithms for the captured problems. Numerical experiments are given to validate the effectiveness of the proposed models.

## **Buy-Online-Pick-Up-At-Store Benefits Supply Chains Considering Supplier Encroachment**

Chun Hung Chiu  
Sun Yat-Sen University

Buy-online-pick-up-at-store (BOPS) is a prevalent strategy in omnichannel retailing. It is widely adopted by retailers, such as Target and Walmart, to attract customers via providing real-time information on product inventory and instant in-store pickup. This strategy is distinguished by its integration of online and offline channels, offering an exclusive service when compared to suppliers who operate solely online. Intuitively, BOPS enhances the competitive advantage of retailers, thereby countering suppliers' encroachment in retailers' online markets. However, our research provides the opposite results. We study the impact of BOPS on supplier encroachment by investigating a two-echelon supply chain involving a retailer considering BOPS implementation and a supplier contemplating online encroachment. Our research reveals that BOPS does not always prevent supplier encroachment; instead, it may trigger it. This occurs because the introduction of BOPS attracts customers who previously encountered stockout in physical stores, encouraging them to switch to the supplier's online channel. This migration prompts the retailer to increase their fulfillment rate to better compete, which subsequently elevates the wholesale profit and benefits the supplier. Moreover, we explore five extensive models, including customer return, storeonly customers, endogenous wholesale price, endogenous retail prices and cross-selling effect, with the results indicating that BOPS consistently facilitates supplier encroachment across all scenarios. In conclusion, our research suggests that BOPS might not be a prudent choice since it not only attracts supplier encroachment but also hurts retailers.

## **Causal Transport for Continuous-time Stochastic Processes**

Rama Cont  
University of Oxford

We study properties of causal couplings for probability measures on the space of continuous functions. We first provide a characterization of bicausal couplings between weak solutions of stochastic differential equations. We then provide a complete description of all such bicausal Monge couplings. In particular, we show that bicausal Monge couplings of  $d$ -dimensional Wiener measures are induced by stochastic integrals of rotation-valued integrands. As an application, we give necessary and sufficient conditions for bicausal couplings to be induced by Monge maps and show that such bicausal Monge transports are dense in the set of bicausal couplings between laws of SDEs with regular coefficients. Our result lead to a stochastic control formulation of

causal optimal transport problems for SDEs. Joint work with Fang Rui LIM (University of Michigan, Ann Arbor)

## **Reinforcement Learning Mean-Variance Policy under Factor Model**

Xiangyu Cui

Shanghai University of Finance and Economics

This paper investigates a discrete-time mean–variance portfolio optimization model under a factor structure. Unlike the classical framework, which treats factor selection, parameter estimation, and dynamic optimization as separate steps, our approach integrates factor selection, parameter estimation, and stochastic policy optimization into a single unified optimization problem. We study both deterministic and stochastic policies, propose an algorithm to compute these policies, and establish the asymptotic properties of our procedure. Simulation studies and empirical evidence together demonstrate the effectiveness and robustness of the proposed approach.

## **Connecting Singular Control and Optimal Stopping in Multi-Asset Portfolio Selection with Transaction Costs**

Min Dai

The Hong Kong Polytechnic University

It is well known that certain singular control problems can be reformulated as optimal stopping problems. For portfolio selection with transaction costs and a single risky asset, this equivalence was established by Dai and Yi (2009). In this talk, we extend the equivalence to the multi-asset case.

## **Multiasset Liquidation in Dark Pools**

Guanxing Fu

The Hong Kong Polytechnic University

We investigate multi-asset portfolio liquidation in dark pools under adverse selection by analyzing a multidimensional stochastic control problem, which leads to a matrix-valued BSDE with jumps and a singular terminal condition. Beyond establishing the existence of a solution, our key contribution lies in proving uniqueness, a result that is novel even in simplified settings.

In the two-asset case, we examine the interaction between asset correlation and adverse selection. On the one hand, we find that the protection of a well-diversified portfolio is not determined by adverse selection but rather by the spillover effects of dark-pool orders. On the other hand, we show that the primary role of adverse selection is to determine whether a poorly diversified portfolio can be rescued through dark-pool trading.

## Data-driven Generative Simulation of SDEs using Diffusion Models

Xuefeng Gao

The Chinese University of Hong Kong

This paper introduces a new approach to generating sample paths of unknown stochastic differential equations (SDEs) using diffusion models, a class of generative AI models commonly employed in image and video applications. Unlike the traditional Monte Carlo methods for simulating SDEs, which require explicit specifications of the drift and diffusion coefficients, our method takes a model-free, data-driven approach. Given a finite set of sample paths from an SDE, we utilize conditional diffusion models to generate new, synthetic paths of the same SDE. To demonstrate the effectiveness of our approach, we conduct numerical experiments to compare our method with alternative benchmark ones including neural SDEs. Furthermore, we present a theoretical error analysis of our SDE path generation approach, deriving an explicit bound on the Kullback-Leibler divergence between the distribution of the target SDE path and the distribution of the generated path. Finally, in simulation and empirical studies we leverage these synthetically generated sample paths to enhance the performance of reinforcement learning algorithms for continuous-time mean-variance portfolio selection, hinting promising applications of diffusion models in financial analysis and decision-making.

## On Mean-Field Control and Games in Infinite Dimension

Fausto Gozzi

Luiss University

In this talk I will briefly explain the motivation to study Mean-Field Control and Games in Infinite Dimension. Then, after a brief overview of the literature on the topic, we will pass to present a going-on result on the case of path dependent state equations.

## From LLM to RL and diffusion models, via (rough) differential equations

Xin Guo

UC Berkeley

Transfer learning is a machine learning technique that leverages knowledge acquired in one domain to enhance performance on a related task. It plays a central role in the success of large language models (LLMs) such as GPT and BERT, which leverage pretraining to enable broad generalization across downstream applications. In this talk, I will discuss how reinforcement learning (RL), and in particular continuous time RL, can benefit from transfer learning principles. I will present convergence results formulated through stability analysis for stochastic control systems, using rough differential equation techniques. Finally, I will show how this analysis yields a natural corollary establishing robustness guarantees for a class of score-based generative diffusion models. Based on joint work with Zijiu Lyu of UC Berkeley.

## Dynamic Asset Pricing with $\alpha$ -MEU Model

Xuedong He

The Chinese University of Hong Kong

We study a dynamic asset pricing problem in which a representative agent is ambiguous about the aggregate endowment growth rate and trades a risky stock, human capital, and a risk-free asset to maximize her preference value of consumption represented by the  $\alpha$ -maxmin expected utility model. This preference model is known to be dynamically inconsistent, so we consider intra-personal equilibrium strategies for the representative agent and define the market equilibrium as the one in which the strategy that clears the market is an intra-personal equilibrium. We prove the existence and uniqueness of the market equilibrium and show that the asset prices in the equilibrium are the same as in the case when the agent does not perceive any ambiguity but believes in a particular probabilistic model of the endowment process. We show that with reasonable parameter values, the more ambiguity the agent perceives or the more ambiguity-averse she is, the lower the risk-free rate, the higher the stock price, the higher the stock risk premium, and the lower the stock volatility.

## Optimal Exploitation of Random Resources

Ulrich Horst

Humboldt University

We consider single-player and mean-field game models of optimal exploitation of an exhaustible resource when only the distribution but not the exact resource level is known. Our model results in a new class of infinite-horizon path-dependent optimization problems. Under mild assumptions on the distribution of the resource level, we prove that a single-player's value function can be characterized as the unique viscosity solution to a certain infinite-horizon stochastic HJB equation. The particular structure of the HJB equation allows for a surprisingly simple extension of single-player models to mean-field games. The benchmark model where the distribution of the resource is the only available information can be solved in closed form. The talk is based on ongoing work with Jinniao Qiu and Yang Yang.

## Equilibrium Policy on Dividend and Capital Injection under Time-inconsistent Preferences

Sang Hu

Chinese University of Hong Kong (Shenzhen)

This paper studies the dividend and capital injection problem under a diffusion risk model with general discount functions. A proportional cost is imposed when injecting capitals. For exponential discounting as time-consistent benchmark, we obtain the closed-form solutions and show that the optimal strategies are of threshold type. Under general discount function which

leads to time-inconsistency, we adopt the definition of weak equilibrium and obtain the extended HJB equation system. An explicit solution is derived under pseudo-exponential discounting where three cases of the dividend and capital injection thresholds are obtained. Numerical examples show that large capital injection cost may lead to no capital injection at all, while larger difference in group discount rate leads to higher equilibrium value function.

## **Adaptive Reparameterized Time for Diffusion Sampling: A Reinforcement Learning Approach to Timestep Allocation**

Yilie Huang  
Columbia University

We study time discretization for score-based diffusion models, where samples are generated by integrating a learned reverse-time process on a finite time grid. Existing samplers typically use uniform or hand-crafted grids, which can allocate function evaluations inefficiently under a fixed computational budget. We propose Adaptive Reparameterized Time (ART), which treats the speed of the reverse dynamics as a control that reparameterizes time and adaptively redistributes computation along the trajectory while keeping the overall horizon fixed. To make this control problem tractable, we introduce an auxiliary randomized control problem, ART-RL, casting ART as a continuous-time reinforcement learning (RL) task with specific Gaussian policies over the progression speed of the reverse dynamics. We prove that the mean of the optimal Gaussian policy in ART-RL coincides with the optimal ART control and derive data-driven actor-critic update rules with policy improvement guarantees for the learned time reparameterization. Empirically, the resulting ART-RL schedules consistently outperform uniform grids and Elucidated Diffusion Model (EDM) schedules in a one-dimensional benchmark and achieve lower FID score than all baselines with a compact score network on MNIST. Within the official EDM sampling pipeline for high-dimensional images, the learned ART-RL schedule, together with its interpolations to other step counts, improves FID across a broad range of sampling budgets on CIFAR-10 and transfers without retraining to AFHQv2, FFHQ, and ImageNet.

## **Mean-Field Langevin Diffusions with Density-dependent Temperature**

Yu-Jui Huang  
University of Colorado Boulder

In the context of non-convex optimization, we let the temperature of a Langevin diffusion to depend on the diffusion’s own density function. The rationale is that the induced density captures to some extent the landscape imposed by the non-convex function to be minimized, such that a density-dependent temperature provides location-wise random perturbation that may better react to, for instance, the location and depth of local minimizers. As the Langevin dynamics is now self-regulated by its own density, it forms a mean-field stochastic differential equation (SDE) of the Nemytskii type, distinct from the standard McKean-Vlasov equations. Relying on Wasserstein subdifferential calculus, we first show that the corresponding (nonlinear)

Fokker-Planck equation has a unique solution. Next, a weak solution to the SDE is constructed from the solution to the Fokker-Planck equation, by Trevisan's superposition principle. As time goes to infinity, we further show that the induced density converges to an invariant distribution, which admits an explicit formula in terms of the Lambert W function. A numerical example suggests that the density-dependent temperature can simultaneously improve the accuracy of and rate of convergence to the estimate of global minimizers.

## **Accuracy of Discretely Sampled Stochastic Policies in Continuous-time Reinforcement Learning**

Yanwei Jia

The Chinese University of Hong Kong

Stochastic policies (also known as relaxed controls) are widely used in continuous-time reinforcement learning algorithms. However, executing a stochastic policy and evaluating its performance in a continuous-time environment remain open challenges. This work introduces and rigorously analyzes a policy execution framework that samples actions from a stochastic policy at discrete time points and implements them as piecewise constant controls. We prove that as the sampling mesh size tends to zero, the controlled state process converges weakly to the dynamics with coefficients aggregated according to the stochastic policy. We explicitly quantify the convergence rate based on the regularity of the coefficients and establish an optimal first-order convergence rate for sufficiently regular coefficients. Additionally, we prove a 1/2-order weak convergence rate that holds uniformly over the sampling noise with high probability, and establish a 1/2-order pathwise convergence for each realization of the system noise in the absence of volatility control. Building on these results, we analyze the bias and variance of various policy evaluation and policy gradient estimators based on discrete-time observations. Our results provide theoretical justification for the exploratory stochastic control framework in [H. Wang, T. Zariphopoulou, and X.Y. Zhou, *J. Mach. Learn. Res.*, 21 (2020), pp. 1-34].

## **Optimal Time-Consistent Fiscal and Monetary Policies: A Mathematical Macro-Finance Approach**

Wei Jiang

The Hong Kong University of Science and Technology

This talk introduces a mathematical macro-finance approach to studying optimal time-consistent fiscal and monetary policies. By extending two macro-finance models into a stochastic continuous-time setting, we examine two types of time consistency: (1) the Markov-perfect Nash equilibrium, as defined in Björk, Murgoci, and Zhou (2011), and (2) the Ramsey equilibrium that can be implemented through active management strategies, following the spirit of Lucas and Stokey (1983). Our mathematical macro-finance framework offers two key advantages: 1) it delivers sharp characterizations of both forms of time-consistent policies, and 2) it identifies the necessary commitment requirements and active management strategies under which a Markov equilibrium can sustain a Ramsey equilibrium.

## Optimal Stopping without Time Consistency

Hanqing Jin  
University of Oxford

In this work, we propose a new framework for solving a general dynamic optimal stopping problem without time consistency. A solution is proposed and is well-defined for any time setting and any form of objectives. A backward iteration is proposed to find the solution. The iteration works with some extra condition, which does not hold in general, but fortunately holds for the time inconsistency resulted by non-exponential discounting. When the iteration does not work, the equilibrium solution can be studied by a forward definition. The solution does not respect the widely-accepted backward induction solution (BIS) defined only for finite discrete time setting.

## Value-Aware Model Learning for Stochastic Control

Lingfei Li  
The Chinese University of Hong Kong

Accurate model learning in noisy environments remains a fundamental challenge for stochastic control, where traditional statistical methods often fail to perform well. We propose a value aware framework for learning continuous time systems governed by stochastic differential equations. Unlike traditional estimation methods, our approach identifies models whose value functions closely match empirical rewards observed under chosen policies. To enable efficient learning, we introduce a surrogate loss function with provable convergence guarantees. In this talk, we focus on the application to learning stock market dynamics with hidden regimes, where our method significantly outperforms established techniques such as the EM algorithm, yielding markedly improved policy predictions for optimal investment.

## Optimal Consumption under a Drawdown Constraint Over a Finite Horizon

Xun Li  
The Hong Kong Polytechnic University

This work studies a finite horizon utility maximization problem on excessive consumption under a drawdown constraint. Our control problem is an extension of the one considered in Bahman et al. (2019) to the model with a finite horizon and an extension of the one considered in Jeon and Oh (2022) to the model with zero interest rate. Contrary to Bahman et al. (2019), we encounter a parabolic nonlinear HJB variational inequality with a gradient constraint, in which some time-dependent free boundaries complicate the analysis significantly. Meanwhile, our methodology is built on technical PDE arguments, which differs from the martingale approach in Jeon and Oh (2022). Using the dual transform and considering the auxiliary variational inequality with gradient and function constraints, we establish the existence and uniqueness of

the classical solution to the HJB variational inequality after the dimension reduction, and the associated free boundaries can be characterized in analytical form. Consequently, the piecewise optimal feedback controls and the time-dependent thresholds for the ratio of wealth and historical consumption peak can be obtained.

## **Free Boundary Problems for Measuring Credit Rating Migration Risks**

Jin Liang  
Tongji University

This talk delivers a research synthesis report centered on the measurement of credit rating migration risks. Based on Merton’s structural approach, we conduct a systematic review of the developmental progress of modeling frameworks, critical theoretical outcomes, and applicable numerical techniques related to this topic—with particular emphasis on the free migration boundary. Additionally, we put forward a series of open research questions.

## **Optimal Trade Execution with Learning**

Andrew Lim  
National University of Singapore

The problem of optimal trade execution is to find the trading rate that minimizes the cost of liquidating an inventory of stock by a given time  $T$  in the presence of market impact costs. We consider the case of linear price impact when the drift and permanent price impact parameters are constant but unknown to the trader and adopt a Bayesian control formulation where the prior is updated using return data generated during the trade. Since the trading decision affects both the cost as well as the data, there is a tradeoff between exploration and exploitation which is fully characterized by the optimal policy. This is difficult, however, because the Hamilton-Jacobi-Bellman (HJB) dynamic programming equation has no explicit solution and is computationally intractable due to the curse of dimensionality. To isolate the exploration and exploitation components of the optimal trade, we consider an asymptotic expansion in the regime where the noise in the return is large. The zeroth-order (“exploitation”) term is the solution of the original (Almgren-Chriss) trade execution problem with uncertain parameters replaced by their prior means, while the first-order (“exploration”) term captures the tradeoff between the temporary price impact cost, the cost of using inventory for exploration, and the value of resolving uncertainty about the drift and the price impact. Both terms in the expansion are easy to compute: There is an explicit expression for the Almgren-Chriss solution while the exploration adjustment is the solution of a system of linear ordinary differential equations. Our analysis shows an intuitive but non-trivial dependence of optimal exploration on inventory and horizon.

## Recent Advance in Nonlinear Control with Input Delay: From Deterministic Systems to Stochastic Systems

Wei Lin

Case Western Reserve University

This talk presents recent advance in the study of when a globally asymptotically stabilizable nonlinear system by either state or output feedback is tolerant to input delay. The first part of this lecture focuses on the deterministic case when the  $C^0$  system is not stabilizable by linear or smooth feedback, and thus local exponential stabilizability is not satisfied. With the aid of the Lyapunov–Krasovskii functional method, the notion of weighted homogeneity and its properties, we prove that for a MIMO nonlinear system with homogeneous degree zero, global asymptotic stabilizability by nonsmooth but non-Lipschitz continuous state or output feedback implies global asymptotic stability of the closed-loop nonlinear system with delay, i.e., global input delay tolerance (GIDT). We then demonstrate applications of the obtained GIDT results to significant classes of time-delay nonlinear systems with uncontrollable/unobservable linearization, including but not limited to, a chain of nonlinear integrators, strongly nonlinear systems with a strict feedback or feedforward structure, and beyond.

In the second part of this talk, we discuss briefly how the deterministic results can be carried over to  $C^0$  input-delayed stochastic systems with inherent nonlinearity (e.g., without locally Lipschitz continuity (LLC) or linear growth). This is accomplished by proving the existence of weak solutions of  $C^0$  stochastic functional differential equation (SFDE), introducing the notion of almost sure forward completeness, and revealing its instrumental role in addressing input delay tolerance of stochastic nonlinear systems under nonsmooth yet Holder continuous feedback.

### s A Short XYZ of XunYu Zhou

Jin Ma

University of Southern California

Abstract is omitted.

### Robust pricing and hedging of American options in continuous time

Jan Obłój

University of Oxford

We consider the robust pricing and hedging of American options in a continuous time setting. We assume asset prices are continuous semimartingales, but we allow for general model uncertainty specification via adapted closed convex constraints on the volatility. We prove the robust pricing-hedging duality. When European options with given prices are available for static

trading, we show that duality holds against richer models where these options are traded dynamically. Our proofs rely on probabilistic treatment of randomised stopping times and suitable measure decoupling, and on optimal transport duality. In addition, similarly to the approach of Aksamit et al. (2019) in discrete time, we identify American options with European options on an enlarged space. Joint work with Ivan Guo.

## **Nonlinear Expectations in Machine Learning**

Shige Peng  
Shandong University

An important advantage of nonlinear expectation theory is its successful applications in measuring and regulating the probabilistic distributional uncertainty in our real world. In fact this is a challenging problem, theoretically and practically. In this talk we provide some typical examples in machine learning e.g. the QKV mechanism in transformer, and discuss related algorithms based on nonlinear expectation theory.

## **Bridging Schrödinger and Bass for Generative Diffusion Modelling: Applications**

Huyên Pham  
Ecole Polytechnique

Abstract is omitted.

## **Optimal Principal-agent Contract Design via HJB Equation**

Shuaijie Qian  
The Hong Kong University of Science and Technology

Recent advances reduce the Principal-agent contract design problem to a classical stochastic control problem. This transformation opens a promising avenue by characterizing the optimal Markov dynamic contracts via the corresponding Hamilton–Jacobi–Bellman (HJB) equation. In this talk, we will show by PDE techniques that this PDE solution permits Sobolev derivatives up to second order, which facilitates the derivation of the optimal contract. The main technical difficulty is from the potentially unbounded coefficients of the PDE, which prohibit the standard techniques from the Evans-Krylov theory. We overcome this difficulty by an a priori estimate derived from the stochastic control problem. Further numerical analysis is also provided.

## **Optimal Investment to Reach a Financial Goal: A Stochastic Control Framework**

Moris Strub  
Warwick University

We develop a framework for an investor who trades until she either reaches a financial goal or an exogenous deadline arrives. Analogous to utility functions over wealth, we measure satisfaction with the timing of reaching a goal by a discount function. For a continuous-time market where a stochastic factor drives the dynamics of stock prices and the financial goals, the investor maximizes the expected discount at the goal-reaching time and the expected utility of the funding ratio if the goal remains unreached by the deadline. This setup leads to a new class of stochastic control problems. We establish Bellman's principle of optimality and characterize the value function as a viscosity solution to the associated Hamilton-Jacobi-Bellman equation. When the deadline is infinite and the goal is constant, the HJB equation reduces to a form related to the backward heat equation, for which we provide a complete characterization of smooth solutions. When analytical solutions are unavailable, we develop an approach based on Howard's algorithm to obtain numerical solutions. Our analysis shows that optimal investment policies for goal reaching problems can be decreasing in the drift of risky assets and need not converge to full risk-free investment even as volatility diverges to infinity.

## **Some Recent Progress on Diffusion Models: Conditioning and Projection**

Wenpin Tang  
Columbia University

In this talk, I will discuss two aspects of diffusion models. First, I will describe a stochastic analysis paradigm for diffusion conditioning and guidance. Second, I will talk about dimension reduction techniques for diffusion models. Applications to queues and financial stress tests will be discussed. Besides, I will say a few things about how Xun Yu has influenced my research so far.

## **Bridging Schrödinger and Bass for Generative Diffusion Modelling: A Theory**

Nizar Touzi  
New York University

Abstract is omitted.

## Risk Aversion and Risk-Insurance Parity

Ruodu Wang  
University of Waterloo

Risk aversion and insurance are two prominent and interconnected concepts in economics and finance. To explore their fundamental connection, we introduce risk-insurance parity, which associates various classes of insurance contracts with different notions of risk attitudes. We characterize the Arrow-Pratt risk aversion in terms of propensity to full insurance and the stronger notion of risk aversion of Rothschild-Stiglitz in terms of propensity to deductible-limit or proportional insurance. Risk-insurance parity allows us to define two new notions of risk aversion, between weak and strong, characterized by insurance propensity to deductible-only and limit-only contracts respectively.

## DPM: Deep Penalty Methods: A Class of Deep Learning Algorithms for Solving High Dimensional Optimal Stopping Problems

Wei Wei  
Heriot-Watt University

We propose a deep learning algorithm for high dimensional optimal stopping problems. Our method is inspired by the penalty method for solving free boundary PDEs. Within our approach, the penalized PDE is approximated using the Deep BSDE framework proposed by E et al. (2017), which leads us to coin the term “Deep Penalty Method (DPM)” to refer to our algorithm. We show that the error of the DPM can be bounded by the cost function and  $O(\frac{1}{\lambda}) + O(\lambda h) + O(\sqrt{h})$ , where  $h$  is the step size in time and  $\lambda$  is the penalty parameter. This finding emphasizes the need for careful consideration when selecting the penalization parameter and suggests that the discretization error converges at a rate of order  $\frac{1}{2}$ . We validate the efficacy of the DPM through numerical tests conducted on a high-dimensional optimal stopping model in the area of American option pricing. The numerical tests confirm both the accuracy and the computational efficiency of our proposed algorithm.

## Mean-field Stochastic Linear Quadratic Control Problem with Random Coefficients

Jie Xiong  
Southern University of Science and Technology

In this talk, we first prove that the mean-field stochastic linear quadratic (MFSLQ) control problem with random coefficients has a unique optimal control and derive a preliminary stochastic maximum principle to characterize this optimal control by an optimality system. However, because of the term of the form  $\mathbb{E}[A(\cdot)X(\cdot)]$  in the adjoint equation, which cannot be represented in the form  $\mathbb{E}[A(\cdot)]\mathbb{E}[X(\cdot)]$ , we cannot solve this optimality system explicitly. To this end, we

decompose the MFSLQ control problem into two constrained SLQ control problems without the mean-field terms. These SLQ control problems can be solved explicitly by an extended LaGrange multiplier method.

## **Schrödinger Bridge for Generative AI: Relaxed Formulation and Convergence Analysis**

Renyuan Xu  
Stanford University

Generative AI can be framed as the problem of learning a model that maps simple reference measures into complex data distributions, and it has recently found a strong connection to the classical theory of the Schrödinger bridge problems (SBPs) due partly to their common nature of interpolating between prescribed marginals via entropy-regularized stochastic dynamics. However, the classical SBP enforces hard terminal constraints, which often leads to instability in practical implementations, especially in high-dimensional or data-scarce regimes. To address this challenge, we follow the idea of the so-called soft-constrained Schrödinger bridge problem (SCSBP), in which the terminal constraint is replaced by a general penalty function. This relaxation leads to a more flexible stochastic control formulation of McKean–Vlasov type.

We establish the existence of optimal solutions for all penalty levels and prove that, as the penalty grows, both the controls and value functions converge to those of the classical SBP at a linear rate. Our analysis builds on Doob’s  $h$ -transform representations, the stability results of Schrödinger potentials,  $\Gamma$ -convergence, and a novel fixed-point argument that couples an optimization problem over the space of measures with an auxiliary entropic optimal transport problem. These results not only provide the first quantitative convergence guarantees for soft-constrained bridges but also shed light on how penalty regularization enables robust generative modeling, fine-tuning, and transfer learning.

## **Comparison Theorems for Multi-Dimensional BSDEs with Jumps and Applications to Stochastic LQ Control**

Zuoquan Xu  
The Hong Kong Polytechnic University

In this paper, we, for the first time, establish two comparison theorems for multi-dimensional backward stochastic differential equations with jumps. Our approach is novel and completely different from the existing ones for one-dimensional case. Using these and other delicate tools, we then construct solutions to coupled two-dimensional stochastic Riccati equation with jumps in both standard and singular cases. In the end, these results are applied to solve a cone-constrained stochastic linear-quadratic control problem and a mean-variance portfolio selection problem with jumps. Different from no jump problems, the optimal (relative) state processes may change their signs, which is of course due to the presence of jumps. This is a joint work with Ying Hu (Univ Rennes) and Xiaomin Shi (Shandong University of Finance and Economics).

## **Periodic Evaluation with Non-Concave Utility**

Chen Yang

The Chinese University of Hong Kong

A fund manager's performance is often evaluated annually and compared with a benchmark, such as a market index. In addition, the manager may be subject to trading constraints, such as limited use of leverage, no short-selling, and a forced liquidation clause. We formulate this as a periodic evaluation problem with a non-concave utility, a stochastic reference point, and trading constraints. The value function is characterized as the unique solution to a Hamilton-Jacobi-Bellman equation with periodic terminal and boundary conditions, which must be imposed carefully due to possible discontinuities at the terminal time and/or on the liquidation boundary. We find that, at evaluation time, future investment opportunities induce a discontinuity in the value function at the liquidation boundary, leading to a substantial change in local risk aversion. More importantly, this local concavity/convexity weakens and shifts inward from the liquidation boundary to the interior region as the evaluation horizon increases. As a result, the joint effect of periodic evaluation and forced liquidation can generate highly nonlinear investment strategies, thereby helping to understand the complexity of trading strategies in the loss region.

## **Production Planning with Risk Hedging and Clean-tech Adoption**

David Yao

Columbia University

Abstract is omitted.

## **Computational Nonlinear Filtering Using A Deep Learning Approach**

George Yin

University of Connecticut

Nonlinear filtering is a fundamental problem in signal processing, information theory, communication, control and optimization, and systems theory. In the 1960s, celebrated results on nonlinear filtering were obtained. Nevertheless, the computational issues for nonlinear filtering remained a long-standing and challenging problem. In this talk, in lieu of treating the stochastic partial differential equations for obtaining the conditional distribution or conditional measure, we present our recent work to construct finite-dimensional approximations using a deep neural network for the optimal weights of a neural network. Two recursions are used in the algorithm. One of them is the approximation of the optimal weight and the other is for approximating the optimal learning rate. This is a joint work with Qing Zhang (University of Georgia), and Hongjiang Qian (University of Auburn).

# Mean Field Control with Poissonian Common Noise: A Pathwise Compactification Approach

Xiang Yu

The Hong Kong Polytechnic University

This paper contributes to the compactification approach to study mean-field control problems with Poissonian common noise. To overcome the lack of compactness and continuity issues caused by common noise, we exploit the point process representation of the Poisson random measure with finite intensity and propose a pathwise formulation in a two-step procedure by freezing a sample path of the common noise. In the first step, we establish the existence of the optimal relaxed control in the pathwise formulation as if common noise is absent, but with finite deterministic jumping times. The second step plays the key role in our approach, which is to aggregate the optimal solutions in the pathwise formulation over all sample paths of common noise and show that it yields an optimal solution in the original model. To this end, with the help of concatenation techniques, we first develop a pathwise superposition principle in the model with deterministic jumping times, drawing a relationship between the pathwise relaxed control problem and the pathwise measure-valued control problem. As a result, we can further bridge the equivalence among different problem formulations and verify that the constructed solution under aggregation is indeed optimal in the original problem. We also extend the methodology to solve mean-field games with Poissonian common noise, confirming the existence of a strong mean field equilibrium.

## Reinforcement Learning for Stochastic Controls with Model Uncertainty

Jianfeng Zhang

University of Southern California

In this talk we propose a reinforcement learning algorithm for an entropy-regularized HJB equation, arising from stochastic control problems with model uncertainty. The standard policy iteration algorithm, for example the one in our previous paper, has two drawbacks: (i) The iteration relies on the unknown parameter and thus it is not implementable; (ii) The iteration relies on the derivatives of the value function. Since an  $L^2$  (or uniform) approximation of the value function does not lead to a good approximation of its derivatives, the convergence of the algorithm is in danger. In this talk we shall propose a new algorithm which overcomes both difficulties. Our algorithm requires the volatility coefficient to be known, which is benign since the volatility is relatively easy to estimate, but allows the drift coefficient to be unknown. We obtain the rate of convergence and demonstrate the efficiency of the algorithm through several numerical examples. The talk is based on an ongoing joint work with Jin Ma, Gaozhan Wang, and Xunyu Zhou.

# Neural Network Approach for Discrete-Time Portfolio Control with Provable Guarantees

Luhao Zhang  
John Hopkins University

We consider data-driven dynamic portfolio control, where policies may depend on the entire history of observed returns rather than a Markov state. To capture this dependence, we propose a neural network parameterization that maps historical return trajectories to portfolio weights, which aligns empirical performance with rigorous theoretical guarantees. With the mean-field interpretation of single-hidden-layer neural networks and the concavity of the utility function, we obtain a global convergence guarantee for noisy stochastic gradient descent and further derive performance bounds for the resulting policy. We validate the approach through empirical experiments demonstrating competitive out-of-sample returns.

## Deterministic Policy Gradient Methods in Continuous Time and Space

Yufei Zhang  
Imperial College London

The theory of continuous-time reinforcement learning (RL) has progressed rapidly in recent years. Although the ultimate goal of continuous-time RL is to learn a deterministic policy, most existing continuous-time methods rely on stochastic policies. This reliance complicates policy execution, leads to high variance and slow convergence during learning, and makes extracting a deterministic policy nearly impossible.

To address these challenges, we introduce deterministic policy gradient methods for continuous-time RL. We derive a continuous-time policy gradient formula expressed as the expected gradient of the advantage rate function and establish its martingale characterization. This concise formulation allows deterministic policy gradients to be estimated far more efficiently than usual stochastic policy gradients. Building on this theoretical foundation, we propose the CT-DDPG algorithm, which enables stable learning with deterministic policies in continuous-time environments. Numerical experiments show that CT-DDPG achieves improved stability and faster convergence compared to both discrete-time and continuous-time stochastic-policy methods, across a wide range of control tasks with varying time discretizations and noise levels.

# Maps



- Lunch and Reception Dinner: 4/F, Communal Building (upper)
- Conference Venue: Room GH201 (middle)
- Banquet (bottom left)
- New World Millennium Hong Kong Hotel (bottom right)