Stochastic Control Theory and Applications
Théorie du contrôle stochastique et ses applications
(Org: **Ali D. Kara** (University of Michigan), **Somnath Pradhan** (Queen's University) and/et **Serdar Yuksel** (Queen's University))

PETER CAINES, McGill University
Mean Field Games on Large Sparse and Dense Networks

Mean Field Game (MFG) theory treats the existence of Nash equilibria for large population dynamical games by approximating them with infinite population games of negligible agents. The MFG equations consist of the Hamilton-Jacobi-Bellman equation for the control of a generic agent and the Fokker-Planck-Kolmogorov equation describing its state evolution; these are linked by the system's mean field, namely the state distribution of the generic agent.

Graphons are limits of node adjacency matrices (Lovasz, 2012). Graphon MFG (GMFG) systems (Caines and Huang, CDC 2018-19, SICON 2021) consist of MFGs where large nodal sub-populations interact over large dense graphs modelled in the limit by graphons. The solutions of the GMFG equations give a system's Nash equilibria and this permits an analysis of the optimality of the node dependent Nash values with respect to graphon index (Caines, Huang, Gao, Foguen-Tchuendom, CDC 2021,2022, MTNS 2022, IFAC 2023). This presentation of GMFG will use metric space embedded graph limits in the form of graphexons (Caines, CDC 2022). The formulation enables the modelling of limit networks situated in some compact space, M, by generalizing graphon functions on M x M to measures on M x M. This has the advantage of including both large sparse and dense networks; furthermore, the resulting set-up permits meaningful differentiation of variables with respect to network location which is not the case for graphon based formulations.

MARGARET CHAPMAN, University of Toronto
Risk-Aware Control Theory

Risk-aware control theory is a fascinating subfield of stochastic control theory. This subfield concerns the analysis and control of dynamical systems with respect to a vast spectrum of possibilities between the average case and the worst case. The flexibility and generality offered by risk-aware control theory has broad significance because control systems often require an awareness of rare harmful possibilities without being overly cautious and different applications have different needs and preferences about managing uncertainty. Major early players in this area include David Jacobson (1970's) and Peter Whittle (1980's-1990's). There have been many exciting developments since those times, motivated by advances in finance and operations research in the early 2000's. In this talk, I plan to present an overview of risk-aware control theory along with the recent development of risk-aware safety analysis. This is joint work with Michael Fauss (Educational Testing Services, Princeton, NJ) and Kevin Smith (Environmental and Water Resources Engineering, Tufts University).

ZITENG CHENG, University of Toronto
Mean field regret in discrete time games

We use mean field games (MFGs) to investigate approximations of N-player games with uniformly symmetrically continuous heterogeneous closed-loop actions. To incorporate agents' risk aversion (beyond the classical expected utility of total costs), we use an abstract evaluation functional for their performance criteria. Centered around the notion of regret, we conduct non-asymptotic analysis on the approximation capability of MFGs from the perspective of state-action distributions without requiring the uniqueness of equilibria. Under suitable assumptions, we first show that scenarios in the N-player games with large N and small average regrets can be well approximated by approximate solutions of MFGs with relatively small regrets. We then show that δ-mean field equilibria can be used to construct ε-equilibria in N-player games. Furthermore, in this general

setting, we prove the existence of mean field equilibria (MFEs). Our analysis above reveals an approximated refinement of N-player equilibria through MFEs. It also offers theoretical substantiation for algorithms that identify MFEs by minimizing regrets.

ASAF COHEN, University of Michigan
Deep Neural Networks Methods for Mean Field Game Master Equation

A mean field game (MFG) approximates a large-player symmetric game and its mean field equilibrium is fully characterized by the so-called master equation. The master equation is typically a non-linear, partial differential equation. While the master equation's well-posedness is known, an analytical solution is not known. What's more, classical discretization methods for solving the master equation suffer from the curse of dimensionality. In this joint work with Ethan Zell and Mathieu Lauriere, we study two algorithms to efficiently solve the master equation in many dimensions. We call one algorithm the Deep Backward Mean Field Game method (DBMFG) and the other is the Deep Galerkin Method (DGM) of Sirignano and Spiliopoulos. We provide novel proofs of the correctness of the algorithms. Due to the structure of the master equation, we cannot rely on the argument of Sirignano and Spiliopoulos for the correctness of the DGM in this application, nor can we rely on the proof of an analogous deep backward method introduced by Pham et. al. for the DBMFG. Instead, we use the structure of the MFG to overcome these difficulties. Time permitting, I will conclude with some of the numerical results.

YUNUS EMRE DEMIRCI, Queen's University
On Regularity and Ergodicity of Partially Observable Markov (Decision) Processes

In this talk, we study time-homogeneous hidden Markov models where states are not directly observable, also known as partially observable Markov processes. Instead, these states are observed through a measurement channel. The initial state is determined by an initial distribution, and as new observations are made, we update the state's conditional probability measure given the measurements, leading to a nonlinear filtering process.

The focus of our study is on the regularity properties for these nonlinear filters under the Wasserstein metric. We present conditions which lead to geometric ergodicity, implying that the filter process converges to invariance at an exponential rate (as a probability measure valued Markov chain). While unique ergodicity of such filter processes had been studied in the literature, such a geometric ergodicity result appears to be new. We also provide complementary results on unique ergodicity for such models with continuous state spaces.

As an implication of our analysis for controlled hidden Markov models, we provide new conditions for the existence of solutions to the average cost optimality equation for Partially Observable Markov Decision Processes, for which only limited results are available in the literature. We furthermore discuss implications on robustness to incorrect priors

---

DENA FIROOZI, HEC Montréal - Université de Montréal
Risk-Sensitive Control and Mean Field Games: A Variational Approach

We develop a variational approach to address risk-sensitive optimal control problems with an exponential-of-integral cost functional in a general linear-quadratic-Gaussian (LQG) single-agent setup, offering new insights into such problems. Our analysis leads to the derivation of a nonlinear necessary and sufficient condition of optimality, expressed in terms of martingale processes. Subject to specific conditions, we find an equivalent risk-neutral measure, under which a linear state feedback form can be obtained for the optimal control. It is then shown that the obtained feedback control is consistent with the imposed condition and remains optimal under the original measure. Building upon this development, we (i) propose a variational framework for general LQG risk-sensitive mean-field games (MFGs) and (ii) advance the LQG risk-sensitive MFG theory by incorporating a major agent in the framework. The major agent interacts with a large number of minor agents, and unlike the minor agents, its influence on the system remains significant even with an increasing number of minor agents. We derive the Markovian closed-loop best-response strategies of agents in the limiting case where the number of agents goes to infinity. We establish that the set of obtained best-response strategies yields a Nash equilibrium in the limiting case and an ε-Nash equilibrium in the finite-player case.

MINYI HUANG, Carleton University
Mean field social optimization: person-by-person optimality and master equations

We consider a large population optimal control problem and apply dynamic programming from the point of view of a representative agent, instead of directly treating a continuum of agents. This leads to a special HJB equation, called the master equation, for the value function of an agent. For performance analysis, we employ a two-scale master equation system to prove person-by-person optimality.

Joint work with

Shuenn-Jyi Sheu (National Central Uni. and National Chengchi Univ., Taiwan)

and Li-Hsien Sun (National Central Uni., Taiwan)

JOE JACKSON, The University of Chicago
Sharp convergence rates for mean field control on the region of strong regularity

This talk will be about the convergence of certain symmetric N-particle stochastic control problems towards their mean field limits. After a brief introduction to mean field control, we will mainly discuss the following question: how fast do the value functions V^N for the N-particle problems converge towards the value function U of the mean field problem? Or in terms of partial differential equations - how fast do the solutions of certain finite-dimensional Hamilton-Jacobi-Bellman equations converge to the solution of a corresponding Hamilton-Jacobi-Bellman equation set on the space of probability measures? If the data is smooth and convex, then U is smooth, and the rate is $O(1/N)$. When the data is not convex, U may fail to be smooth, and the answer is more subtle. On one hand, we know that the optimal global rate cannot be better than $O(1/\sqrt{N})$. On the other hand, a recent paper of Cardaliaguet and Souganidis identifies an open and dense set $\mathcal{O}$ of initial conditions (which we call the region of strong regularity, by analogy with some classical results on first order Hamilton-Jacobi equations) where U is smooth, and it is natural to wonder whether the rate of convergence might be better inside of $\mathcal{O}$. In an ongoing joint work with Cardaliaguet, Mimikos-Stamatopoulos, and Souganidis, we show that this is indeed the case: the rate is $O(1/N)$ locally uniformly inside the set $\mathcal{O}$, so the convergence is indeed faster inside $\mathcal{O}$ than it is outside.

ROLAND MALHAME, Ecole Polytechnique de Montréal
A bottom-up approach to the construction of socially optimal discrete choices under congestion

We consider the problem of N agents having a limited time to decide on a destination choice among a finite number of alternatives D. The agents attempt to minimize collective energy expenditure while favoring motion strategies which limit crowding along their paths in the state space. This can correspond to a situation of crowd evacuation or a group of micro robots distributing themselves on tasks associated to distinct geographic locations. We formulate the problem as a Min linear quadratic optimal control problem with non positive definite Q matrices accounting for negative costs accruing from decreased crowding. The solution proceeds in three stages, each one improving on the performance of the previous stage: (i) Mapping optimal paths for an arbitrary agent destination assignment; (ii) Mapping optimal paths for fixed fractions of agents assigned to each destination; (iii) Identifying the optimal fraction of agents' assignments to each destination. The cost function associated with stage (iii), as N goes to infinity, is proven to be convex, leads to simplified computations and to epsilon-optimal decentralized control policies when applied for N large.

SOMNATH PRADHAN, Queen's University
Existence and Discrete-Time Approximations of Optimal Controls for Controlled Diffusions under General Information Structures

In this talk, we present existence and discrete-time approximation results on optimal control policies for continuous-time stochastic control problems under a variety of information structures. These include fully observed models, partially observed models and multi-agent models with decentralized information structures. While there exist comprehensive existence and approximations results for the fully observed setup in the literature, few prior research exists on discrete-time approximation

results for partially observed models. For decentralized models, even existence results have not received much attention except for specialized models and approximation has been an open problem. Our existence and approximations results lead to the applicability of well-established partially observed Markov decision processes and the relatively more mature theory of discrete-time decentralized stochastic control to be applicable for computing near optimal solutions for continuous-time stochastic control.

This talk is based on joint work with Serdar Yüksel.

SINA SANJARI, Royal Military College of Canada / University of Illinois at Urbana-Champaign
Large Stochastic Exchangeable Teams, Their Mean-Field Limits, and Optimality of Symmetric Policies

Stochastic teams entail a collection of decision-makers acting together to optimize a common cost function but not necessarily sharing all the available information. We discuss a class of decentralized stochastic exchangeable teams with a finite number of decision-makers as well as their mean-field limits with infinite numbers of decision-makers.

We first consider convex teams. For this class of problems, we establish the existence of a globally optimal solution and show that it is symmetric (identical) for both the finite decision-maker regime and the infinite one. As the number of decision makers drives to infinity (that is for the mean-field limits), we establish the existence of a privately randomized globally optimal solution and show that it is symmetric among decision makers. For the class of non-convex exchangeable teams, we establish the existence of a globally optimal solution and show that it is exchangeable (the joint distribution is permutation invariant) and not necessarily symmetric for the finite decision-maker regime. For the infinite population regime, however, we show the existence of a globally optimal solution and establish that it is privately randomized and symmetric. Finally, we establish that a symmetric globally optimal solution for the mean-field problem is approximately optimal for the corresponding finite-population team with a large number of decision-makers.

BORNA SAYEDANA, McGill University
Relative Almost Sure Regret Bounds for Certainty Equivalence Control of Markov Jump Systems

In this talk, we consider the learning and control problem for unknown Markov jump linear systems (MJLS) with perfect state observations. We first establish an upper bound on regret for any learning-based algorithm. We then propose a certainty-equivalence based learning algorithm and show that this algorithm achieves a regret of $O(\sqrt{T}\log(T))$ relative to a certain subset of the sample space. As part of our analysis, we propose a switched least squares method for the identification of MJLS, show that this method is strongly consistent, and derive data-dependent and data-independent rates of convergence. These results show that certainty equivalence control along with the switched least squares method for MJLS has the same rate of convergence as the certainty equivalence control method for linear systems.

ZACHARY SELK, Queen's University
Robustness for Near-Brownian Noise via Rough Paths Theory

One frequent modelling choice made in stochastic control is assuming that the noise is a Brownian motion. However, this is only an idealization. For example, Kushner argues that "wide-band" Brownian motion is much more physical because the high frequencies are often not present. Another example is the fractional Brownian motion which can be seen as the generalization of Brownian motion to allow for correlations in the increments. There are several other "near-Brownian" but not actually Brownian driving signals which lead to questions of robustness. That is - if we assume an idealized Brownian noise apply the optimal policy to a real situation, are we near optimal? We show that the answer is yes, using rough paths theory.

Rough paths theory was in invented in the 1990s by Terry Lyons as an alternative to the standard Itô theory. One of the issues with Itô theory is a lack of continuity in the driving noise. Another issue with Itô theory is that it only allows for semimartingale noise. This disallows noise such as fractional Brownian motion. One further issue is that Itô theory is not defined pathwise - only as limits in L^2. The key insight of rough paths theory is that by enhancing the driving signal with its "iterated integrals", continuity to the solution map is restored and a wide range of signals can be integrated.

In this talk, we introduce the basics of rough paths theory and discuss a robustness result. Joint with Somnath Pradhan and Serdar Yüksel.

VIJAY SUBRAMANIAN, University of Michigan, Ann Arbor
Bayesian Learning of Optimal Policies in Markov Decision Processes with Countably Infinite State-Space

Models of many real-life applications—queuing models of communication networks—have a countably infinite state-space. Algorithmic and learning procedures that have been developed to produce optimal policies mainly focus on finite state settings, and do not apply to these models. To overcome this lacuna, we study the problem of optimal control of a family of discrete-time countable state-space Markov Decision Processes (MDPs) governed by an unknown parameter $\theta \in \Theta$, and defined on a countably-infinite state space $\mathcal{X} = \mathbb{Z}_+^d$, with finite action space $\mathcal{A}$, and an unbounded cost function. The random unknown parameter θ^* is generated via a given fixed prior distribution on Θ. To optimally control the unknown MDP, we propose an algorithm based on Thompson sampling with dynamically-sized episodes: at the beginning of each episode, the posterior distribution formed via Bayes' rule is used to produce a parameter estimate, which then decides the policy applied during the episode. To ensure the stability of the Markov chain obtained by following the policy chosen for each parameter, we impose ergodicity assumptions. From this condition and using the solution of the average cost Bellman equation, we establish an $\tilde{O}(\sqrt{|\mathcal{A}T})$ upper bound on the Bayesian regret of our algorithm, where T is the time-horizon. Finally, to elucidate the applicability of our algorithm, we consider two different queuing models with unknown dynamics, and show that our algorithm can be applied to develop approximately optimal control algorithms.

This is joint work with Saghar Adler at the University of Michigan, Ann Arbor.

JOHANNES WIESEL, Carnegie Mellon University
Martingale Schrödinger bridges

In a two-period financial market where a stock is traded dynamically and European options at maturity are traded statically, we study the so-called martingale Schrödinger bridge Q*; that is, the minimal-entropy martingale measure among all models calibrated to option prices. This minimization is shown to be in duality with an exponential utility maximization over semistatic portfolios. Under a technical condition on the physical measure P, we show that an optimal portfolio exists and provides an explicit solution for Q*. This result overcomes the remarkable issue of non-closedness of semistatic strategies discovered by Acciaio, Larsson and Schachermayer.

This talk is based on joint work with Marcel Nutz and Long Zhao.

BORA YONGACOGLU, University of Toronto
Connections between POMDPs and partially observed n-player mean-field games

In this talk, we will study a discrete-time model of mean-field games with finitely many players and partial observability of the global state, and we will describe the deep connection between such n-player mean-field games and partially observed Markov decision problems (POMDPs). We focus primarily on settings with mean-field observability, where each player privately observes its own local state as well as the complete mean-field distribution. We prove that if one's counterparts use symmetric stationary memoryless policies, then a given agent faces a fully observed, time homogenous MDP. We leverage this to prove the existence of a memoryless, stationary perfect equilibrium in the n-player game with mean-field observability. We also show that the symmetry condition cannot be relaxed without loss of generality. Under narrower observation channels, in which the mean-field information is compressed before being observed by each agent, we show that the agent faces a POMDP rather than an MDP, even when its counterparts use symmetric policies.