Topics in Mean-Field Control and Games for Pure Jump Processes

SALAH EDDINE CHOUTRI

Doctoral Thesis
KTH Royal Institute of Technology
School of Engineering Sciences
Department of Mathematics
Division of Mathematical Statistics
Stockholm, Sweden 2018

Department of Mathematics
KTH Royal Institute of Technology
100 44 Stockholm, Sweden

TRITA-SCI-FOU 2018:55
ISBN 978-91-7873-061-2

Akademisk avhandling som med tillstånd av Kungliga Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen i tillämpad matematik och beräkningsmatematik, fredagen den 1:a februari 2019 klockan 10.00 i Kollegiesalen, Brinellvägen 8, Kungliga Tekniska högskolan, Stockholm.

Print: Universitetsservice US AB, Stockholm, 2018

To my parents and siblings +1

Abstract

This thesis is the collection of four papers addressing topics in stochastic optimal control, zero-sum games, backward stochastic differential equations, Pontryagin stochastic maximum principle and relaxed stochastic optimal control.

In the first two papers we establish existence of Markov chains of mean-field type, with countable state space and unbounded jump intensities. We further show existence of nearly-optimal controls and, using a Markov chain backward SDE approach, we derive conditions for existence of an optimal control and a saddle-point for a zero-sum differential game associated with risk-neutral and risk-sensitive payoff functionals of mean-field type, under dynamics driven by Markov chains of mean-field type. Our formulation of the control problems is of weak-type, where the dynamics are given in terms of a family of probability measures, i.e. under which the coordinate process is a pure jump process with controlled jump intensities.

In the third paper we characterize the optimal controls obtained in the first paper by deriving sufficient and necessary optimality conditions in terms of a stochastic maximum principle (SMP). Finally, within a completely different setup, in the fourth paper we establish existence of an optimal stochastic relaxed control for stochastic differential equations driven by a G-Brownian motion.

Sammanfattning

Denna avhandling består av fyra artiklar som behandlar ämnen i stokastisk optimal styrning, nollsummespel, BSDE:er, Pontryagins stokastiska maximumprincip och relaxerad stokastisk optimal styrning.

I de två första artiklarna påvisar vi existens av Markovkedjor av medelfälts-typ med uppräkneligt tillståndsrum och obegränsade hoppintensiteter. Vi visar vidare att det finns nästan-optimala kontroller och vi använder en BSDE-metod för att härleda villkor för existens av en optimal styrning till ett styrningsproblem och en sadelpunkt för ett nollsummespel i differentialform. I den första artikeln är problemen associerade med riskneutrala kostnadsfunktionaler av medelfälts-typ, medan vi behandlar riskkänsliga kostnadsfunktionaler medelfälts-typ i den andra artikeln. I både artiklarna drivs dynamiken av Markovkedjor av medelfälts-typ. Vår formulering av styrningsproblemen brukar kallas för svag typformulering, där dynamiken ges i form av en uppsättning sannolikhetsmått.

I den tredje artikeln karaktäriserar vi de optimala kontroller som erhålls i den första artikeln genom att härleda nödvändiga och tillräckliga villkor i form av en stokastisk maximumprincip. I den fjärde artikeln etablerar vi existensen av en optimal stokastisk relaxerad kontroll för stokastiska differentialekvationer drivna av G-Brownsk rörelse.

Acknowledgments

First of all, I am greatly indebted to my supervisor Professor Boualem Djehiche for his so appreciated support, patience, guidance, understanding, practical discussions and the continuous follow-up that he had to carry out during my time as a Ph.D. student. He was a real source of inspiration for me. Thank you for everything, dear Boualem.

I would like to thank my co-supervisor Professor Henrik Hult for the valuable knowledge that I acquired especially in his two reading courses.

I wish to thank Professor Hamidou Tembine for the interesting discussions and the enlightenment about certain aspects of mean-field games.

A special thanks goes to my colleague and dear friend Alexander Aurell for all the help he provided during my four years at KTH, and for all the good discussions (about mathematics, music, history, languages, football $\cdots$ etc), the chess games and Swedish tutorials.

I wish to express my gratitude to all my Ph.D. fellows in the department of mathematics, for the nice discussions we had, especially Johan Westerborn, Felix Rios and David Karagolian. Morevover, I am thankful to Nacira Agram for enriching my knowledge about topics in stochastic optimal control during my visit to University of Oslo.

I would like to thank my friend/brother Othmane Mezhar for all kind of advice and all the conversations we had together. Thanks to Anna Uth, Pierre Wikby and Olle Blomqvist for proofreading my thesis. A Special mention goes to Professor Abdelhafid Mokrane for the valuable recommendations.

Finally, I would like to show my deepest appreciation to my parents, sister, brothers and uncle Abdelaziz, for all kind of backing and encouragement.

Stockholm, 2018

Salah Choutri

Table of Contents

Introduction and Background

A Markov process is a stochastic process for which the future state depends only on the present one. Markov chains are Markov processes which take values in a discrete state space. They were introduced in 1906 by Andrei Andreyevich Markov and were named in his honor. Due to their modeling power in a variety of applied fields such as biology, economics, physics, communication networks and epidemics, etc. (see e.g. [12, 37, 47]), Markov chains are perhaps the most important class of Markov processes. We distinguish two significant classes, continuous-time and discrete-time Markov chains. A stochastic process $(x(t),\ t \geq 0)$ is called a continuous-time Markov chain (CTMC) if it has a discrete state space I and satisfies the Markov property, i.e. for all $t \geq s \geq 0$, $i, j \in I$, it holds that

$$P(x(t) = j | x(s) = i, \{x(u) : 0 \leq u < s\}) = P(x(t) = j | x(s) = i) =: \mathbf{p}_{ij}(t).$$

If, in addition, the conditional probabilities $\mathbf{p}_{ij}(t)$ depend only on $t - s$, i.e. $P[x(t) = j|x(s) = i] = P[x(t-s) = j|x(0) = i]$, then the CTMC is said to have time-homogeneous (or stationary) transition probabilities and the process is called time-homogeneous Markov chain. A continuous-time Markov chain is characterized by an initial distribution $\xi := P \circ x^{-1}(0)$ and an intensity matrix, also called Q-matrix or infinitesimal generator, $Q(t) = (q_{ij}(t),\, i, j \in I)$, $0 \leq t \leq T$, satisfying

$$q_{ij}(t) = \lim_{h \to 0} \tfrac{1}{h} P(x(t+h) = j | x(t) = i), \quad i, j \in I,\ i \neq j,$$

$$q_{ii}(t) = -\lim_{h \to 0} \tfrac{1}{h} (1 - P(x(t+h) = i | x(t) = i)), \quad i \in I,$$

$$\text{with} \quad \sum_{j:\, j \neq i} q_{ij}(t) = -q_{ii}(t) < \infty, \qquad \sum_{i,j:\, j \neq i} \int_{(0,T]} q_{ij}(t)\, dt < +\infty,$$

where $q_{ij}(t),\ i, j \in I,\ 0 \leq t \leq T$, are called the *local characteristics* of the underlying CTMC. Under some suitable assumptions, the transition matrix $\mathbf{p}(t) = (\mathbf{p}_{ij}(t),\, i, j \in I)$, $0 \leq t \leq T$, satisfies the so called *Kolmogorov backward equation*:

$$\frac{d}{dt}\mathbf{p}(t) = -Q(t)\mathbf{p}(t),$$

and the so called *Kolmogorov forward equation:*

$$\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(t)Q(t).$$

These differential systems describe the time-evolution of the transition probabilities. Moreover, the probability distribution of the holding time T_i i.e the length of time spent in state i, if the initial state is i, is given by

$$P(T_i \geq t) = \exp\left\{\int_0^t q_{ii}(s)ds\right\}.$$

Real-world problems in the fields mentioned above, increasingly suggest more complex mathematical modeling. For instance, standard continuous-time Markov chains are not suitable to model phenomena that take into account some sort of weak interaction, such as large systems of interacting queues on a network or interacting molecules in a chemical reaction. This motivated the study of a family of non-conventional Markov chains called Markov chains of mean-field type (MCMF) or simply nonlinear Markov chains. A process x with discrete state space (such as e.g. $I = \{0, 1, 2, 3, \ldots\}$ or $I = \mathbb{Z}^d$, $d \geq 1$) is called a nonlinear Markov chain if its jump intensities $q_{ij}(t)$, at time t, depend on the 'marginal law' $P \circ x^{-1}(t)$ of the chain:

$$q_{ij}(t) = \lambda_{ij}(t, P \circ x^{-1}(t)), \quad i, j \in I.$$

A typical situation is when the jump intensities depend on the mean of the underlying process $E[x(t)]$. This type of Markov chains is obtained as the limit of a system of pure jump processes with mean-field interaction, when the system size tends to infinity. The states of the finite systems are represented by a sequence of empirical distribution functions whose deterministic limit is the marginal law of the Markov chain. This marginal law satisfies a 'nonlinear' Fokker-Planck (a.k.a McKean-Vlasov equation) and represents a typical behavior of a pure jump process in the underlying system. For existence and uniqueness results for nonlinear Markov chains with bounded and unbounded jump intensities we refer to [38] and [34].

The modeling power of Markov chains of mean-field type in chemistry, physics, biology and economics etc. is well documented in the literature. Nicolis and Prigogine [37] proposed such non-conventional Markov chains as a mean-field model for a chemical reaction with spatial diffusion. Mean-field models for the so-called first and second Schlögl models [46] and the auto-catalytic processes (see e.g. [12]), which are widely used to model chemical reactions, represent interesting examples of the nonlinear Markov chains with unbounded jump intensities. These processes are obtained as limits of systems of birth and death processes with mean field interaction and have been thoroughly studied in [16], [23] and [22]. For application in the spread of epidemics see e.g. Léonard [33], Djehiche and Kaj [20], Djehiche and Schied [21].

Adding costs and a decision mechanism, to a Markov chain, with the objective to minimize them, we obtain the so-called Markov decision chain (MDC) or optimal control problem for the Markov chain. Loosely speaking, the problem is how to control a Markov chain to achieve an economic goal by minimizing a given cost functional. Control is exercised by taking a sequence of actions, each of which may depend on the currently observed state and may influence both the running cost and the next state transition. This amounts to controlling the intensity or Q-matrix of the underlying chain. We refer to [3, 14, 26, 45, 52] and the references therein for an introduction to optimal control of Markov chains.

As it is documented in the above mentioned references, the first results of optimal control of Markov chains were based on the Dynamic Programming approach which consists of reducing the optimal control problem to a system of nonlinear ODEs, the Hamilton-Jacobi-Belman equation, then solve it numerically. Bismut [5], Boel and Varaiya [7], Davis and Elliott in [15] and Wan and Davis [51], among many others, developed a general approach based on the martingale description of the evolution of the Markov chain. The main idea in the Martingale and Dynamic Programming approaches in these works is to control the probability measure describing the controlled chain through its Radon-Nikodym density (if it exists) with respect to a reference probability measure.

In their recent paper, Guo and Hernández-Lerma [25] derive conditions for existence of a value for a zero-sum game for a continuous-time Markov chain, with unbounded jump intensities, using dynamic programming. The considered game is associated to a long term average payoff criterion. Following the framework suggested in Cohen and Elliott (see e.g. [13]), Carmona and Wang [11] develop a probabilistic approach to finite state mean-field games under dynamics driven by a nonlinear Markov chain with bounded intensities. Their approach is based on a semimartingale description of the chain and the weak formulation of stochastic optimal control.

1 Weak-type formulation to optimal control of Markov chains

Stochastic optimal control problems, in general, can be studied using either the so-called strong or weak formulation. The strong formulation is based on 'fixing' the probability space, while the weak formulation consists of varying the probability space and, therefore, considering it as part of the control mechanism. This formulation is reasonable because the goal of optimal control problems is to optimize a functional that depends solely on the distribution of the underlying process.

In this thesis, we consider a weak-type formulation for optimal control problems and zero-sum games for pure jump processes. Our approach can be summarized as follows. We take $\Omega := \mathcal{D}([0,T], I)$ to be the space of functions from $[0,T]$ to I that are right continuous with left limits at each $t \in [0,T)$ and are left continuous at time T. Given $t \in [0,T]$ and $\omega \in \Omega$, we define the coordinate process $x(t,\omega) = \omega(t)$. Denote by $\mathbb{F}^0 := (\mathcal{F}^0_t)_{0\le t\le T}$, the filtration generated by x and by $\mathbb{F} := (\mathcal{F}_t)_{0\le t\le T}$ its completion with the P-null sets of Ω, where P is the probability measure under which x is time-inhomogeneous Markov chain with intensity matrix $Q(t) = (q_{ij}(t), \qquad, i,j \in I)$, $0 \le t \le T$.

We control the coordinate process through its jump intensities $Q^u(t) = (q^u_{ij}(t))$, where u belongs to the set of admissible controls $\mathcal{U}$, which in turn induces a family of probability measures $(P^u, u \in \mathcal{U})$. The construction is done through a change of measure involving the Doleans-Dade exponential martingale. Therefore, by using the density process $L^u := \frac{dP^u}{dP}$ as state dynamics, we obtain a standard optimal control problem whose dynamics are given by the controlled process L^u.

A jump process is called of mean-field type or nonlinear if its jump intensity is of the form

$$q^u_{ij}(t) = \lambda_{ij}(t, P^u \circ x^{-1}(t), u(t)), \quad i,j \in I, \quad u \in \mathcal{U},$$

where $P^u \circ x^{-1}(t)$ is the marginal distribution of $x(t)$ under P^u. Hence, the probability measure P^u should be a fixed point.

2 Performance criteria

The economic objective of an optimal control problem is to optimize a performance functional. In the case of minimization problems, risk neutral cost functionals are widely used as optimization criteria. Linear-quadratic cost functionals are among the most common risk neutral cost functionals for which one can characterize optimal controls explicitly. A cost functional is called of mean-field type, if the running and terminal costs depend on the marginal law of the chain, in which case it is of the form

$$J(u) := E^u\left[\int_0^T f(t, x(t), P^u \circ x^{-1}(t), u(t))dt + h(x(T), P^u \circ x^{-1}(T))\right]. \quad (1)$$

However, risk neutral cost functionals do not take into account attitudes toward risks. Jacobson [32] is first to show that a risk-sensitive cost functional is a reasonable way to capture risk-averse (risk-sensitive) and risk-seeking behaviors. A risk-sensitive cost functional is of the form

$$J_\theta(u) := \frac{1}{\theta}\ln E^u\Big[\exp\Big(\theta\int_0^T f(t,x(t),P^u \circ x^{-1}(t),u(t))dt + h(x(T),P^u \circ x^{-1}(T))\Big)\Big], \quad (2)$$

where $\theta > 0$ is called the risk sensitive parameter. Since the control problem is stochastic the two cost functionals (1) and (2) differ. When θ is small the cost functional (2) can be expended as

$$E^u[\mathcal{L}_T] + \frac{\theta}{2}\ Var^u(\mathcal{L}_T) + O(\theta^2),$$

where $\mathcal{L}_T = \int_0^T f(t, x(t), P^u \circ x^{-1}(t), u(t))dt + h(x(T), P^u \circ x^{-1}(T))$ and $Var^u(\cdot)$ denotes the variance under P^u. If $\theta > 0$, the decision maker is called risk-averse since the variance worsens the optimization criterion J_θ, while the opposite happens when $\theta < 0$ and the controller is called, then, risk-seeker. Furthermore, when θ goes to 0, the sequence of θ-indexed risk sensitive cost functionals converges to the risk-neutral cost functional.

Cost functionals can be characterized in terms of backward stochastic differential equations (BSDEs). Furthermore, existence of optimal control can be obtained using comparison theorems for BSDEs (see e.g. [27, 28]). BSDEs play a crucial role in stochastic maximum principle since the adjoint equations are of this type [6].

In contrast to a deterministic differential equation, one cannot simply reverse time to obtain an adapted solution to a BSDE since in practise we only have information about the past but not about the future. Therefore, another adapted (or predictable) stochastic process needs to be added to the solution to correct the 'non-adaptiveness' that may occur. Thus, an adapted solution of a BSDE is a pair of processes, where the first represents the mean evolution of the dynamics and the second represents the risk/uncertainty [53]. We further explain this fact for Markov chains in the next section.

3 Markov chain terminal value problems

A Markov chain terminal value problem is a system of a differential equation involving the martingale associated to the underlying Markov chain, and a terminal value which must be satisfied by any solution (if it exists). In contrast to a deterministic terminal value problem, a solution to a stochastic terminal value problem is defined on a filtered probability space, where the given filtration shrinks as we go backward in time. This makes the filtration play a crucial role in determining the solution, i.e. any solution must be adapted to the filtration. This adaptiveness requirement makes this kind of problems, in particular BSDEs, difficult to solve.

In this section we give an illustrative example to show the issues that might occur when formulating a Markov chain terminal value problem and how to overcome that. Given the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, consider the following differential equation

$$dY(t) = f(t, Y(t^-))dM(t), \quad Y(T) = \zeta, \quad (3)$$

where ζ is a square-integrable $\mathcal{F}_T$-measurable random variable and M is the martingale associated with the underlying Markov chain. The goal is to find an adapted solution (assuming existence) to (3), but obviously this is not possible since integrating (3), under suitable conditions, from t to T yields a candidate solution that is not $\mathbb{F}$-adapted. For instance, if we take $f = 0$, the equation (3) becomes

$$dY(t) = 0, \quad Y(T) = \zeta, \tag{4}$$

where the only candidate solution is $Y(t) = \zeta$ for all t, which is obviously $\mathcal{F}_T$-measurable, and thus not $\mathbb{F}$-adapted. Bismut [6] was first to solve a similar problem (in the diffusion case) by introducing the conditional expectation to obtain an adapted solution from the martingale representation theorem. If we apply Bismut's idea to (4) we obtain

$$Y(t) = E\left[\zeta|\mathcal{F}_t\right], \quad t \in [0, T]. \tag{5}$$

This shows that the process Y is in fact a martingale and, furthermore, it is a $\mathbb{F}$-adapted process satisfying the terminal condition. However, (5) no longer satisfies (4), thus, a reformulation of (4) is needed to have (5) as one of its $\mathbb{F}$-adapted solutions. The martingale representation theorem can be used to carry out such reformulation which eventually leads to a linear Markov chain BSDE. As a consequence, the term $Z(t)dM(t)$ is added to the original differential equation, and the process Z is a unique and $\mathbb{F}$-predictable. The latter is considered as a part of the solution, due to the fact that it makes the process Y adapted. Therefore, a solution to (4) is the pair of processes (Y, Z) satisfying the linear Markov chain BSDE

$$dY(t) = Z(t)dM(t), \quad Y(T) = \zeta. \tag{6}$$

Backward SDEs driven by Markov chains have attracted a lot of attention recently, their well-posedness has been studied, based on the martingale representation theorem, in a series of papers by Cohen and Elliott (see e.g. [13] and the references therein). Their approach is analogous to the one that deals with BSDEs driven by Brownian motions (see e.g. [40]). We refer to [35, 55] for an extensive study of Brownian motions driven BSDEs.

4 Two-person zero-sum games

Game theory is the field of mathematics that studies the interaction of rational agents (players) following strategic decision rules. Decision-makers are assumed to be rational in the sense that they minimize/maximize a performance criterion. Furthermore, they need to follow strategies (controls) taking into account other players' actions (responses). Initially, game theory addressed what is called two-person zero-sum games, in which a player loses if the other player wins. Its name comes from the fact that if we sum up the total gains and subtract the total losses, we obtain zero. In contrast, a non-zero-sum game refers to a situation where it is possible for all the players to lose/win together and clearly the sum of the total gains minus the total losses is not zero.

The so-called 'matching pennies' is a basic example of two-person zero-sum game. It involves two players (A and B) who simultaneously place a penny on the table, with the payoff depending on whether the pennies match. If both pennies are heads or tails, player A wins and takes player B's penny; if they do not match, player B earns player A's penny.

The two-person zero-sum games were first introduced by John von Neumann in 1928. Later in 1944, John von Neumann and Oscar Morgenstern extended Neumann's works and develop a general theory for zero-sum games (see e.g. [50]). The study of deterministic differential games, including zero-sum games, was initiated by Rufus Isaacs in 1965 (see e.g. [31]). Extensions to stochastic two-person zero-sum differential games are well studied and documented (see e.g. [28] and the references therein). In [28], the authors derived conditions for existence of a saddle-point (or minimax) for a stochastic two-person zero-sum game under the so called Issac's condition, where the dynamics are given by weak solutions of a Brownian motion driven SDE. The obtained result is extended to the mean-field case in [19].

Markov chain zero-sum games are stochastic games where the dynamics are given by a Markov chain. The paper [25] is first to study two-person zero-sum games for denumerable continuous-time Markov chains determined by given unbounded jump intensities, with an average performance criterion. In this thesis we deal with two-person zero-sum game where the dynamics are driven by a Markov chain whose jump intensities are of the mean-field form

$$\lambda^{u,v}(t) := \lambda_{ij}(t, x, P^u \circ x^{-1}(t), u(t), v(t)), \;\; i, j \in I, \; 0 \leq t \leq T,$$

where u and v are the controls of players who try to minimize and maximize, respectively, a common performance criterion.

5 Stochastic maximum principle

Optimal control problems in the weak-type formulation can be regarded as optimization problems in infinite-dimensional spaces, which are generally hard to solve [53]. The maximum principle is an effective mathematical tool that provides a set of necessary conditions that must be satisfied by any optimal solution. These necessary conditions can be sufficient if we assume certain convexity conditions on the involved drivers. The maximum principle originates from the work of Pontryagin and his group in the 1950s. It basically states that any optimal control-state pair must satisfy a maximum condition on a function called the Hamiltonian, and solves a forward-backward differential equation called Hamiltonian system. The power of Pontryagin's maximum principle lies within the maximization of the Hamiltonian which is an easier task to carry out and which, moreover, allows for explicit solutions for certain classes of optimal control problems such as the linear-quadratic case.

The original version of Pontryagin's maximum principle is for deterministic systems. The optimality conditions are derived by perturbing an optimal control on a small time interval, the used perturbation is known as the 'spike variation'. Then, by performing a first order Taylor expansion with respect to the perturbation and by sending the perturbation to zero, one obtains a variational inequality. Duality is then applied to obtain the maximum principle. However, one encounters a major difficulty when trying to mimic the same steps to obtain a stochastic maximum principle (SMP). The difficulty is that the backward differential equation, known as the adjoint equation, in the Hamiltonian system becomes a BSDE which, as mentioned before, requires some adaptiveness. Furthermore, Peng [41] observed that controlling the volatility term in the diffusion case poses an extra problem since the usual first-order perturbation does not suffice to obtain a maximum principle. One needs to take into account both first-order and second-order terms in the Taylor expansion, which leads to a stochastic maximum principle with two adjoint equations, both in form of linear BSDEs. Andersson and Djehiche [1] and Buckdahn, Djehiche and Li [8] were first to extend the Stochastic Maximum Principle to mean-field couplings of the form $E[\phi(x(t))]$. Hosking [29] extends the SMP to mean-field couplings of the form $E[\phi(x(t), u(t))]$ which includes the law of the control itself. Extensions to mean-field couplings of the form $\phi(t, x(t), P \circ x^{-1}(t), u(t))$, are considered in [9] and [10].

In the third paper of this thesis, we derive a stochastic maximum principle for pure jump processes of mean-field type based on the weak-type formulation considered in the first two papers, where we control the density process L^u that serves as dynamics and solves a linear SDE driven by some accompanying martingale. The adjoint process associated to the SMP solves a Markov chain backward stochastic differential equation (BSDE) driven by the accompanying martingale.

6 Relaxed optimal controls

Pontryagin's maximum principle characterizes optimal controls assuming they exist and belong to the space $\mathcal{U}$ of control processes with values in some compact metric space (U, δ). However, a control problem may fail to have a solution, simply because the set $\mathcal{U}$ might be too small to contain an optimum. As a remedy, one can enlarge the space of controls to include probability measures, called 'relaxed controls', on the set $\mathcal{U}$. In fact, we compactify the space of admissible controls $\mathcal{U}$ to have a richer topological structure for which the control problem can be solved.

This notion of relaxation was first introduced for deterministic optimal control problems in [54]. Then in [24], it was generalized to stochastic control problems for diffusion processes where the drift and the volatility are controlled.

In [30], [4] and [36], an optimal control problem of systems governed by stochastic differential equation driven by G-Brownian motion (see an account below) was studied. These authors derive sufficient and necessary optimality conditions in terms of maximum principle and dynamic programming principle. The fourth paper of the thesis investigates existence of a relaxed optimal control for SDEs driven by the G-Brownian motion in the absence of convexity assumptions on the coefficients.

7 G-Brownian motion and G-expectation

In [42, 43, 44] Peng introduces an abstract sublinear space with a process called G-Brownian motion whose definition is based on a sublinear expectation (or Choquè capacity) called G-expectation, which is in turn defined through the nonlinear heat equation in the following sense. A d-dimensional random vector X is said to be G-normally distributed under the G-expectation $\widehat{\mathbb{E}}[\cdot]$ if for each bounded and Lipschitz continuous function φ on $\mathbb{R}^d$, $\varphi \in \mathrm{Lip}(\mathbb{R}^d)$, the function u defined by

$$u(t,x) := \widehat{\mathbb{E}}[\varphi(x + \sqrt{t}X)], \quad t \geq 0, \ x \in \mathbb{R}^d$$

is the unique, bounded Lipschitz continuous viscosity solution of the following parabolic equation

$$\frac{\partial u}{\partial t} - G(D^2 u) = 0 \quad \text{on } (t,x) \in (0, +\infty) \times \mathbb{R}^d \quad \text{and} \quad u(0,x) = \varphi(x),$$

where $D^2 u = (\partial^2_{x_i x_j} u)_{1 \leq i,j \leq d}$ is the Hessian matrix of u and the nonlinear operator G is defined by

$$G(A) := \frac{1}{2} \sup_{\gamma \in \Gamma} \{\mathrm{tr}\,(\gamma\gamma^* A)\}, \quad \gamma \in \mathbb{R}^{d \times d}. \tag{7}$$

where A is a $d \times d$ symmetric matrix and Γ is a given non empty, bounded and closed subset of $\mathbb{R}^{d \times d}$. Here, v^* denotes the transpose of the vector v. This G-normal distribution is denoted by $N(0, \Sigma)$, where $\Sigma := \{\gamma\gamma^*, \ \gamma \in \Gamma\}$.

Peng ([42, 43]) shows that the G-expectation $\widehat{\mathbb{E}} : \mathcal{H} := \mathrm{Lip}(\mathbb{R}^d) \longrightarrow \mathbb{R}$ is a consistent sublinear expectation on the lattice $\mathcal{H}$ of real functions. That is, it satisfies sub-additivity, monotonicity, constant preserving and positive homogeneity (i.e. $\forall \lambda \geq 0,\ X \in \mathcal{H},\ \widehat{\mathbb{E}}[\lambda X] = \lambda \widehat{\mathbb{E}}[X]$). In [18], the authors provide a dual representation of G-expectation as a supremum of ordinary expectations over a weakly compact (tight) family $\mathcal{P}$ of possibly mutually singular probability measures. This duality expresses the G-expectation as a robust expectation with respect to $\mathcal{P}$. We refer to [17] and [18] for explicit constructions of $\mathcal{P}$. Soner, Touzi and Zhang [48, 49] perform an in-depth analysis of such a construction and its consequences on the G-stochastic analysis and in particular the question of aggregation of processes.

Summary of the Papers

Paper A: Optimal control and zero-sum games for Markov chains of mean-field type

In this paper we establish existence of Markov chains of mean-field type with unbounded jump intensities by means of a fixed point argument using the Total Variation distance. We further show existence of nearly optimal controls and we suggest conditions for existence of an optimal control $\widehat{u}$ satisfying

$$\min_{u\in\mathcal{U}} J(u) = J(\widehat{u}), \tag{1}$$

with

$$J(u) := E^u\left[\int_0^T f(t,x,P^u\circ x^{-1}(t),u(t))dt + h(x(T),P^u\circ x^{-1}(T))\right], \tag{2}$$

where f and h satisfy standard progressive measurability, linear growth and uniform boundedness assumptions. The corresponding optimal dynamics is given by the probability measure $\widehat{P}$ on $(\Omega,\mathcal{F})$ defined by

$$d\widehat{P} = L^{\widehat{u}}(T)dP, \tag{3}$$

under which the coordinate process x is a jump process with intensities

$$\lambda_{ij}^{\widehat{u}}(t) := \lambda_{ij}(t,x,P^{\widehat{u}}\circ x^{-1}(t),\widehat{u}(t)),\;\; i,j\in\{0,1,2,\dots\},\; 0\le t\le T,$$

where the jump intensity is of functional type w.r.t. the process x i.e. it depends on the whole path of the chain as long as it is predictable. We also consider a zero-sum game between two players, where the first player (with control $u\in\mathcal{U}$) wants to minimize a performance functional $J(u,v)$, while the second player (with control $v\in\mathcal{V}$) wants to maximize it. More precisely, we want to show existence of a pair $(\widehat{u},\widehat{v})$ of strategies such that

$$J(\widehat{u},v)\le J(\widehat{u},\widehat{v})\le J(u,\widehat{v}), \qquad \forall (u,v)\in\mathcal{U}\times\mathcal{V}, \tag{4}$$

where

$$J(u,v) := E^{u,v}\left[\int_0^T f(t,x,P^{u,v}\circ x^{-1}(t),u(t),v(t))dt + h(x(T),P^{u,v}\circ x^{-1}(T))\right]. \tag{5}$$

The corresponding optimal dynamics is given by the probability measure $\widehat{P}$ on $(\Omega,\mathcal{F})$ defined by

$$d\widehat{P} = L_T^{\widehat{u},\widehat{v}}dP \tag{6}$$

under which the chain has intensity

$$\lambda^{\widehat{u},\widehat{v}}(t) := \lambda_{ij}(t, x, P^{\widehat{u}} \circ x^{-1}(t), \widehat{u}(t), \widehat{v}(t)), \;\; i, j \in I, \; 0 \leq t \leq T.$$

The proof of existence of the family of Markov chains of mean-field type is based on a Girsanov-type change of measure and the Csiszár-Kullback-Pinsker inequality. The full use of the total variation distance requires L^2-boundedness of the Girsanov density, which is insured by imposing an extra regularity condition of the intensity matrix of the chain compared with what should be natural if the Wasserstein distance is used.

The main results on optimal control and zero-sum games are derived using techniques involving Markov chain backward stochastic differential equations (BSDE), where existence of an optimal control and a saddle-point strategy of the game boil down to finding a minimizer and a min-max of an underlying Hamiltonian H. Since the mean-field coupling through the marginal law of the controlled chain makes the Hamiltonian H, evaluated at time t, depend on the whole path of the control process over the time interval $[0, t]$, we cannot perform a *deterministic* minimization of H over the set of actions U and then apply a Beneš-type progressively measurable selection theorem to produce an optimal control. We should rather take the essential infimum of H over the set $\mathcal{U}$ of progressively measurable controls. This nonlocal feature of the dependence of H on the control does not seem covered by the existing powerful measurable selection theorems. Therefore, our main results are formulated *by assuming* existence of an essential minimum $\widehat{u} \in \mathcal{U}$ of H and use suitable comparison results of Markov chain BSDEs to show that $\widehat{u}$ is in fact an optimal control.

Paper B: Mean-field risk sensitive control and zero-sum games for markov chains

In this paper we generalize the results obtained in Paper A by deriving conditions for existence of an optimal control and a saddle-point for respectively the control problem and the zero-sum game considered in the first paper but when, instead, associated with performance functionals of risk-sensitive type. These performance functionals are obtained by exponentiating the risk-neutral performance functionals (2) and (5), before expectation. The considered risk-sensitive performance functional associated with the controlled nonlinear Markov chain is

$$J(u) := E^u \left[\exp\left(\int_0^T f(t, x, P^u \circ x_t^{-1}, u_t)dt + h(x_T, P^u \circ x_T^{-1}) \right) \right],$$

while the one associated with the zero-sum differential game is given by

$$J(u, v) := E^{u,v} \left[\exp\left(\int_0^T f(t, x, u, v)dt + h(x(T), P^{u,v} \circ x^{-1}(T)) \right) \right]. \tag{7}$$

In Paper A, existence and uniqueness results were derived with the help of the Total Variation (TV) distance, which required that the jump intensities are bounded from below by a strict positive constant. However, TV does not guarantee existence of finite moments, mean-field couplings of the type $E^u[X_t^u]$ or $E^u[\varphi(X_t^u)]$ where φ is a Lipschitz function, were excluded. To consider this type of couplings, the results obtained in this paper are derived using the Wasserstein distance as it is designed to guarantee finite moments. But, then we can no longer use the approach of Paper A, based on the Girsanov transform because, in general, there is no relation between the Wasserstein distance and the Hellinger (Entropy) distance, unless the nonlinear Markov chain satisfies a log-Sobolev inequality [39].

Under mild integrability and growth conditions on the unbounded jump intensities, using the Wasserstein distance, we show that by applying the Skorohod selection (or embedding) theorem and L^2-estimates, a fixed-point argument is still valid to derive existence and uniqueness of P^u. This turns out possible, thanks to Ekeland's distance on the set of admissible controls which makes it complete (or Polish space). Existence of an optimal control and a saddle-point of the game are derived using techniques involving Markov chain entropic backward stochastic differential equations (BSDE) which boils down to finding a minimizer and a min-max of an underlying Hamiltonian H. We formulate our main results as in Paper A since the mean-field coupling through the marginal law of the controlled chain makes the Hamiltonian H, evaluated at time t, depend on the whole path of the control process over the time interval $[0, t]$.

Paper C: A Stochastic Maximum Principle for Markov Chains of Mean-Field Type

This paper deals with the characterization of the optimal control of the nonlinear Markov chain obtained in Paper A. That is, we derive sufficient optimality conditions in terms of a stochastic mean-field maximum principle for optimal controls associated with cost functionals of mean-field type, under a dynamics driven by a class of Markov chains of mean-field type, when the mean-field coupling $P^u \circ x^{-1}(t)$ is given in terms of expectations $E^u[\kappa(x(t))]$, for some measurable functions κ. The fact that the cost functional is of mean-field type leads to time inconsistent control problem which means that the Bellman optimality principle does not hold since the law of iterated expectations cannot be applied on the cost functional. We illustrate the result by two examples of optimal control problems that involve two-state Markov chains and linear quadratic cost functionals. For instance, we consider a Markov chain whose state space $\mathcal{X} = \{a, b\} (0 \leq a < b)$ and mean-field jump intensity matrix,

$$\lambda^u(t) = \begin{bmatrix} -\alpha & \alpha \\ u(t) + E^u[x(t^-)] & -u(t) - E^u[x(t^-)] \end{bmatrix},$$

where $\alpha > 0$ and $u(t) + E^u[x(t^-)] \geq 0$.

The goal is to minimize the cost functional

$$J(u) = E^u\left[\frac{1}{2}\int_0^T u^2(t)dt\right] + Var^u(x(T)), \tag{8}$$

where $Var^u(x(T))$ denotes the variance of $x(T)$ under the probability P^u defined by

$$Var^u(x(T)) = E^u\left[(x(T) - E^u[x(T)])^2\right].$$

The last part of the paper is dedicated to some real-world applications such as optimal control of mean-field version of schlögl model for chemical reactions. For simplicity and to compute the optimal control explicitly, we consider linear quadratic cost functionals in all these cases.

The candidate's contribution The co-authors of paper A, B and C suggested the topic of each of the papers and helped with the formulation of the solved problems. The candidate solved the problems, performed the proofs of all the results and wrote the manuscripts.

Paper D: On Relaxed Stochastic Optimal Control for Stochastic Differential Equations Driven by G-Brownian Motion

The purpose of this paper is to study optimal control of systems subject to model uncertainty or ambiguity due to incomplete or inaccurate information, or vague concepts and principles. Climate or weather and financial markets are typical fields where information is subject to uncertainty.

Within the G-Brownian motion framework as presented in [42, 44], this paper deals with optimal control of systems governed by stochastic differential equation driven by a G-Brownian motion. More precisely, we show that there exists a stochastic optimal control $\widehat{u} \in \mathcal{U}$ with values in an action set U such that

$$J(\widehat{u}) = \inf_{u \in \mathcal{U}} J(u), \tag{9}$$

with

$$J(u) = \sup_{P \in \mathcal{P}} J^P(u) := \sup_{P \in \mathcal{P}} \mathbb{E}^P \left[\int_0^T f(t, x^u(t), u(t))dt + h(x^u(T)) \right], \tag{10}$$

where x^u is a G-SDE given by

$$\begin{cases} dx^u(t) = \sigma(t, x^u(t))dB_t + b(t, x^u(t), u(t))dt + \gamma(t, x^u(t), u(t))d\langle B \rangle_t, \\ x^u(0) = x. \end{cases} \tag{11}$$

with

$$b : [0, T] \times \mathbb{R}^d \times U \to \mathbb{R}^d, \qquad \sigma, \gamma : [0, T] \times \mathbb{R}^d \times U \to \mathbb{R}^{d \times d},$$
$$f : [0, T] \times \mathbb{R}^d \times U \to \mathbb{R}, \qquad h : \mathbb{R}^d \to \mathbb{R}.$$

Furthermore, in this work we do not assume any convexity assumptions, which may lead to non-existence of solution to the control problem (9), because $\mathcal{U}$ is too small to contain a minimizer. Therefore, to solve the problem we use an embedding technique that consists of finding a set $\mathcal{R}$ of controls that 'contains' $\mathcal{U}$ and has a richer topological structure for which the control problem becomes solvable. This embedding is often called a relaxation of the control problem and $\mathcal{R}$ is the set of relaxed controls, while $\mathcal{U}$ is called the set of strict controls. The main result of this paper is to construct the set $\mathcal{R}$ of relaxed controls as a subset of the set of probability measures on the action set U and show that

$$\inf_{u \in \mathcal{U}} J(u) = \inf_{\mu \in \mathcal{R}} J(\mu) = J(\widehat{\mu}), \tag{12}$$

where

$$\widehat{\mu} = arg \min_{\mu \in \mathcal{R}} J(\mu). \tag{13}$$

In order to show (12), we prove and use an extended version of the so called 'Chattering Lemma', which states that each relaxed control in $\mathcal{R}$ can be approximated with a sequence of strict controls from $\mathcal{U}([0, T])$, i.e. given a relaxed control $\mu \in \mathcal{R}$, there exists a sequence $(u^n)_n \in \mathcal{U}([0, T])$ of strict controls such that $\delta_{u^n(t)}(da)dt$ converges weakly to $\mu_t(da)dt$ p-a.s., for all $P \in \mathcal{P}$. The proof of (13) is based on existence of an optimal relaxed control for each $P \in \mathcal{P}$ derived in [2], together with Skorohod's embedding theorem, and a tightness argument. Stability results for the relaxed version of the G-SDE (11) are derived to complete the proofs.

The candidate's contribution This paper was co-authored with a PhD student who suggested the studied topic. Both authors equally contributed in the mathematical formulation of the problem and to the proofs of the obtained results.

References

[1] D. Andersson and B. Djehiche. A maximum principle for SDEs of mean-field type. *Applied Mathematics & Optimization*, 63(3):341–356, 2011.

[2] S. Bahlali, B. Mezerdi, and B. Djehiche. Approximation and optimality necessary conditions in relaxed stochastic control problems. *International Journal of Stochastic Analysis*, 2006:1–23, 2006.

[3] D. P. Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena scientific Belmont, MA, 2005.

[4] F. Biagini, T. Meyer-Brandis, B. Øksendal, and K. Paczka. Optimal control with delayed information flow of systems driven by G-Brownian motion. *Probability, Uncertainty and Quantitative Risk*, 2018.

[5] J. M. Bismut. Control of jump processes and applications. *Bulletin des Sciences Mathématiques*, 106(1):25–60, 1978.

[6] J. M. Bismut. An introductory approach to duality in optimal stochastic control. *SIAM Review*, 20(1):62–78, 1978.

[7] R. Boel and P. Varaiya. Optimal control of jump processes. *SIAM Journal on Control and Optimization*, 15(1):92–119, 1977.

[8] R. Buckdahn, B. Djehiche, and J. Li. A general stochastic maximum principle for SDEs of mean-field type. *Applied Mathematics & Optimization*, 64(2):197–216, 2011.

[9] R. Buckdahn, J. Li, and J. Ma. A stochastic maximum principle for general mean-field systems. *Applied Mathematics & Optimization*, 74(3):507–534, 2016.

[10] R. Carmona and F. Delarue. Forward-backward stochastic differential equations and controlled mckean–vlasov dynamics. *The Annals of Probability*, 43(05):2647–2700, 2015.

[11] R. Carmona and P. Wang. Finite state mean-field games with major and minor players. *arXiv preprint arXiv:1610.05408 [math.PR]*, 2016.

[12] M. F. Chen. *From Markov chains to non-equilibrium particle systems*. World Scientific, 2004.

[13] S. N. Cohen and R. J. Elliott. Existence, uniqueness and comparisons for BSDEs in general spaces. *The Annals of Probability*, 40(5):2264–2297, 2012.

[14] M. H. A. Davis. *Markov models & optimization*. Routledge, 2018.

[15] M. H. A. Davis and R. Elliott. Optimal control of a jump process. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 40:183–202, 1977.

[16] D. Dawson and X. Zheng. Law of large numbers and central limit theorem for unbounded jump mean-field models. *Advances in Applied Mathematics*, 12(3):293–326, 1991.

[17] L. Denis and C. Martini. A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *The Annals of Applied Probability*, 16(2):827–852, 2006.

[18] L. Denis, M. Hu, and S. Peng. Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. *Potential Analysis*, 34(2):139–161, 2011.

[19] B. Djehiche and S. Hamadène. Optimal control and zero-sum stochastic differential game problems of mean-field type. *Applied Mathematics & Optimization*, https://doi.org/10.1007/s00245-018-9525-6, 03 2018.

[20] B. Djehiche and I. Kaj. The rate function for some measure-valued jump processes. *The Annals of Probability*, 23(2):1414–1438, 1995.

[21] B. Djehiche and A. Schied. Large deviations for hierarchical systems of interacting jump processes. *Journal of Theoretical Probability*, 11(1):1–24, 1998.

[22] S. Feng. Large deviations for empirical process of mean-field interacting particle system with unbounded jumps. *The Annals of Probability*, 22(4):2122–2151, 1994.

[23] S. Feng and X. Zheng. Solutions of a class of nonlinear master equations. *Stochastic processes and their applications*, 43(1):65–84, 1992.

[24] W. Fleming. Generalized solutions in optimal stochastic control. Technical report, Lefschetz Center for Dynamical Systems at Brown University, 1976.

[25] X. Guo and O. Hernández-Lerma. Zero-sum games for continuous-time Markov chains with unbounded transition and average payoff rates. *Journal of Applied Probability*, 40 (2):327–345, 2003.

[26] X. Guo and O. Hernández-Lerma. *Continuous-time Markov decision processes*. Springer, 2009.

[27] S. Hamadène and J. Lepeltier. Backward equations, stochastic control and zero-sum stochastic differential games. *Stochastics: An International Journal of Probability and Stochastic Processes*, 54(3-4):221–231, 1995.

[28] S. Hamadène and J. P. Lepeltier. Zero-sum stochastic differential games and backward equations. *Systems & Control Letters*, 24(4):259–263, 1995.

[29] J. J. A. Hosking. A stochastic maximum principle for a stochastic differential game of a mean-field type. *Applied Mathematics and Optimization*, 66(03):415–454, 2012.

[30] M. Hu, S. Ji, and S. Yang. A stochastic recursive optimal control problem under the G-expectation framework. *Applied Mathematics & Optimization*, 70(2):253–278, 2014.

[31] R. Isaacs. *Differential Games*. John Wiley & Sons, 1965.

[32] D. Jacobson. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE transactions on automatic control*, 18(2):124–131, 1973.

[33] C. Léonard. Some epidemic systems are long range interacting particle systems. *Stochastic processes in epidemic systems (eds. J.P. Gabriel et al.), Lecture Notes in Biomathematics*, 86, 1990.

[34] C. Léonard. Large deviations for long range interacting particle systems with jumps. *Annales de l'IHP Probabilités et statistiques*, 31(2):289–323, 1995.

[35] J. Ma and J. Yong. *Forward-backward stochastic differential equations and their applications - Introduction*. Springer, 1999.

[36] A. Matoussi, D. Possamai, and C. Zhou. Robust utility maximization in non-dominated models with 2BSDEs. *Mathematical Finance*, 25(2):258–287, 2015.

[37] G. Nicolis and I. Prigogine. *Self organization in non-equilibrium systems*. Wiley-Interscience, New York, 1977.

[38] K. Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *The Annals of Probability*, 12(2):458–479, 1984.

[39] F. Otto and C. Villani. Generalization of an inequality by talagrand, and links with the logarithmic sobolev inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000.

[40] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems and Control Letters*, 14:55–61, 1990.

[41] S. Peng. A general stochastic maximum principle for optimal control problems. *SIAM Journal on control and optimization*, 28(4):966–979, 1990.

[42] S. Peng. G-expectation, G-Brownian motion and related stochastic calculus of itô type. In *Stochastic analysis and applications*, pages 541–567. Springer, 2007.

[43] S. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Processes and their Applications*, 118(12):2223–2253, 2008.

[44] S. Peng. Nonlinear expectations and stochastic calculus under uncertainty. *arXiv preprint arXiv:1002.4546*, 2010.

[45] M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[46] F. Schlögl. Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für Physik*, 253(2):147–161, 1972.

[47] L. I. Sennott. *Stochastic dynamic programming and the control of queueing systems*, volume 504. John Wiley & Sons, 2009.

[48] H. M. Soner, N. Touzi, and J. Zhang. Martingale representation theorem for the G-expectation. *Stochastic Processes and their Applications*, 121(2):265–287, 2011.

[49] H. M. Soner, N. Touzi, and J. Zhang. Quasi-sure stochastic analysis through aggregation. *Electronic Journal of Probability*, 16(67):1844–1879, 2011.

[50] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 1944.

[51] C. Wan and M. H. A. Davis. Existence of optimal controls for stochastic jump processes. *SIAM Journal on Control and Optimization*, 17(4):511–524, 1979.

[52] P. Whittle. *Optimization Over Time: Dynamic Programming and Optimal Control, Vol. II.* John Wiley and Sons, Ltd., 1983.

[53] J. Yong and X. Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.

[54] L. C. Young. Lectures on the calculus of variations and optimal control theory. *Stochastic processes and their applications*, 1969.

[55] J. Zhang. *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory*, volume 86. Springer, 2017.