People 's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

Mohamed Khider University Biskra

Faculty of Exact Sciences and Nature and Life Sciences

THIRD CYCLE LMD FORMATION

A Thesis submitted in partial execution of the requirements of

DOCTORATE DEGREE IN MATHEMATICS

Option

Probability

Presented by

Abdelhakim NINOUH

Titled

Some results on the stochastic control of backward doubly stochastic differential equations

---

Supersivor: **Dr. Boulakhras GHERBAL**

Examination Committee :

Mr. Mokhtar HAFAYED	Professor	University of Biskra	President
Mr. Boulakhras GHERBAL	Professor	University of Biskra	Supervisor
Mr. Amrane HAOUES	Doctor	University of Biskra	Examiner
Mr. Khalil SAADI	Professor	University of M'sila	Examiner
Mr. Abdelmoumen TIAIBA	Professor	University of M'sila	Examiner

January 2021

Acknowledgements

In the name of Allah, and may blessings and peace be upon the Messenger of Allah, Muhammad al-Sadiq al-Amin, and all his family and companions.

Firstly and foremost, I thank Allah almighty for the strength and patience to finish this Ph.D thesis. I would also like to express my sincere gratitude to my supervisor Dr. Boulakhras GHERBAL, for his patience, care and encouragement, and his guidance and for helped me in all the time, which we allowed us to orient my work in a relevant way, and which I asked Allah The Almighty to reward him. Many thanks to the members of the discussion committee : **Mr. Mokhtar HAFAYED**, **Mr. Amrane HAOUES**, **Mr. Khalil SAADI**, **Mr. Abdelmoumen TIAIBA** for their corrections and time they spent in reading and evaluating this work.

A special thanks to my family: my parents my brothers and sister for supporting me, words also cannot express how grateful I am to my mother, and father for all their sacrifices. Finally, I ask Allah Subhanah Wa Ta'ala to protect you both.

Would also like to thank all our professors from University of Biskra and to the members of the Applied Mathematics Laboratory and the Laboratory of Mathematical Analysis, Probability and Optimization, colleagues and friends, without exception.

Contents

Abstract

The objective of this thesis is to proof the existence of optimal relaxed controls as well as optimal stricts controls for systems governed by non linear forward–backward stochastic differential equations (FBSDEs). In the first part, we study an singular control problem for systems of forward-backward stochastic differential equations of mean-field type (MF-FBSDEs) in which the control variable consists of two components: an absolutely continuous control and a singular one. The coefficients depend on the states of the solution processes as well as their distribution via the expectation of some function. Moreover the cost functional is also of mean-field type. Our approach is based on weak convergence techniques in a space equipped with a suitable topological setting. We prove in first, the existence of optimal relaxed-singular controls,which are a couple of measure-valued processes and a singular control. Then, by using a convexity assumption and measurable selection arguments, the optimal regular (strict)-singular control are constructed from the optimal relaxed-singular one.

In the second part of this thesis, we concentrate on the study of a class of optimal controls for problems governed by forward-backward doubly stochastic differential equations (FBDSDEs). We prove the existence of an optimal control in the class of relaxed controls, which are measure-valued processes, generalizing the usual strict controls. The proof is based on some tightness properties and weak convergence on the space of càdlàg functions, endowed with the

Jakubowsky S-topology. Furthermore, under some convexity assumptions, we show that the optimal relaxed control is realized by a strict control.

Résumé

L'objectif de cette thèse est de prouver l'existence des contrôles relaxés optimaux ainsi que l'existence des contrôles stricts optimaux pour des systèmes gouvernés par des équations différentielles stochastiques progressives-rétrogrades non linéaires (EDSPR). Dans la première partie, nous étudions un problème de contrôle singulier pour des systèmes d'équations différentielles stochastiques progressives et rétrogrades de type champ moyen (MF-FBSDE) dans lequel la variable de contrôle est constituée de deux composants: un contrôle absolument continu et un singulier. Les coefficients dépendent des processus d'état de résolution ainsi que de leur distribution via l'espérance d'une fonction. De plus, la fonctionnelle de coût est également de type champ moyen. Notre approche est basée sur des techniques de convergence faible dans un espace muni par une topologie suitable. Nous prouvons premièrement l'existence des contrôles relaxés-singuliers optimaux, qui sont des couples du processus valorisés par des mesures et des contrôles singuliers. Ensuite, en utilisant une hypothèse de convexité et des arguments de sélection mesurables, le contrôle régulier-singulier optimal est construit à partir du contrôle relaxé-singulier optimal.

Dans la deuxième partie de cette thèse, nous concentrons sur l'étude d'une classe de contrôles optimaux pour des problèmes gouvernés par des équations différentielles doublement stochastique progressives-rétrogrades (FBDSDE). Nous prouvons l'existence d'un contrôle optimal dans la classe des contrôles relaxés, qui sont des processus a valeur mesure, généralisant les contrôles stricts usual. La preuve est basée sur quelques propriétés des tensions et des convergences faibles sur l'espace des fonctions càdlàg, muni par la S-topologie de Jakubowsky. De plus, sous certaines hypothèses de convexité, nous montrons que le contrôle relaxé optimal est réalisé par un contrôle strict.

Symbols and Abbreviations

The different symbols and abbreviations used in this thesis.

Symbols

$(\Omega, \mathcal{F}, \mathbb{P})$	:	Probability space.
$(W_t)_{t\geq 0}$	:	Brownian motion.
$\mathcal{F}^{\theta}_{s,t}$	:	σ-fields generated by $\sigma\left(\theta_r - \theta_s, s \leq r \leq t\right)$.
$\mathcal{F}^{\theta}_{t}$	:	σ-algebre generated by θ.
$\mathcal{F}_t := \mathcal{F}^B_{t,T} \vee \mathcal{F}^W_t$	:	σ-fields generated by $\mathcal{F}^B_{t,T} \cup \mathcal{F}^W_t$.
$\mathcal{G}_t = \mathcal{F}^W_t \vee \mathcal{F}^B_T$	:	The collection $(\mathcal{G}_t)_{t\geq 0}$ is a filtration.
$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$	:	A filtered probability space.
$(\mathcal{F}_t)_{t\geq 0}$	:	Filtration.
$\mathbb{R}^{n\times d}$	:	The set of all $(n \times d)$ real matrixes.
$\mathbb{R}^n$	:	n-dimensional real Euclidean space.
$\mathbb{R}$	:	Real numbers.
$\mathbb{E}[X\vert\mathcal{F}_t]$	:	Conditional expactation.
$\mathbb{E}[X]$	:	Expactation at X.
$\mathcal{N}$	:	The collection of class of P-null sets of $\mathcal{F}$.

ξ	:	A singular control.
q	:	A relaxed control.
$\delta(da)$	:	A Dirac measure.
U	:	The set of values taken by the strict control u.
U_2	:	The set of values taken by the singular part of the control.
$\mathcal{U}$	:	The set of admissible strict controls (or regular part of the control).
$\mathcal{U}_2$	:	The set of singular part of the control.
$\mathbb{V}$	:	The space of positive Radon measures on $[0;T] \times U$.
$\tilde{q}$	:	Optimal relaxed control.
$\tilde{u}$	:	Optimal strict control.
$\mathcal{R}_{ad}$	:	The set of relaxed control.
$CV\ (\cdot)$	:	The conditional variation.

Abbreviations

SDEs	:	Stochastic differential equations.
BSDEs	:	Backward stochastic differential equations.
BDSDEs	:	Backward doubly stochastic differential equation.
FBSDE	:	Forward-backward stochastic differential equations.
MF-FBSDE	:	Forward-backward stochastic differential equations of mean field type.
càdlàg	:	Right continuous with left limits.
a.s	:	Almost surely.

Introduction

The approach of relaxed controls is a relatively popular method of compactification of stochastic control problems to establish existence of solutions, which comes in several different flavors. Fleming [25] derived the first existence result of an optimal relaxed control for SDEs with uncontrolled diffusion coefficient by using compactification techniques. The case of stochastic differential equations with controlled diffusion coefficient has been solved by El-Karoui et al. [23], where the optimal relaxed control is shown to be Markovian, the authors reformulate the control problem as a relaxed controlled martingale problem. See also [28, 29, 35]. A similar approach is used by Lacker [36] in the context of MF Games.

Forward-backward stochastic differential equations (FBSDEs in short) were first studied by Antonelli (see [1]), where the system of such equations is driven by Brownian motion on a small time interval. The proof there relies on the fixed point theorem. There are also many other methods to study FBSDEs on an arbitrarily given time interval. For example, the four-step scheme approach of Ma et al. [42], in which the authors proved the existence and uniqueness of solutions for fully coupled FBSDEs on an arbitrarily given time interval, where the diffusion coefficients were assumed to be nondegenerate and deterministic. Their work is based on continuation method. For systems of FBSDE, the existence of optimal control has been proved by Bahlali, Gherbal and Mezerdi [10], see also Buckdahn et al [13]. Benbrahim and Gherbal in [12] proved existence of optimal controls for FBSDEs of mean-field type with controlled diffusion coefficient. Bahlali, Gherbal and Mezerdi in [11] proved existence of a strong optimal control for linear BSDEs and this result has been extended to a system of linear backward doubly SDEs by Gherbal [26]. The existence of relaxed solutions to mean field games with singular controls has been proved by Fu and Horst in [24]. The authors proved approximations of solutions for a particular class of mean field games with singular controls and relaxed controls by solutions for mean field games with purely regular controls, on the space of cádlág functions equipped with the Skorokhod M_1 topology.

A new kind of backward stochastic differential equations called mean-field BSDEs, were introduced by Buckdahn et al. [6], which were derived as a limit of some highly dimentional system of BSDEs, corresponding to a large number of particles. The existence of solution for forward-backward stochastic differential equations of mean-field type systems (MF-FBSDEs) has been proved by Carmona and Dularue [19].

The existence of approximate Nash equilibria in mean field games for large populations has been established in [19], using a representative agent approach. A relaxed solution concept to mean field games was introduced by Lacker in [37], in which the author studied in the framework of controlled martingale problems, a general existence theorems where the equilibrium control is Markovian. In [38], the author used the notation of weak solution and proved that the weak limit of ϵ-Nash equilibria for N player games as $N \to \infty$ is a weak solution to mean field games. Moreover, each weak solution to mean field games yields an ϵ-Nash equilibrium for N player game. See also Carmona and Delarue [20], they established a new version of the stochastic maximum principle for systems of SDEs of mean-field type.

The aim of the first part, of this thesis is to study of singular stochastic control problem for systems governed by the following MF-FBSDEs:

$$\begin{cases} dX_t^{u,\xi} = b(t, X_t^{u,\xi}, \mathbb{E}[\alpha(X_t^{u,\xi})], u_t)dt + \sigma(t, X_t^{u,\xi}, \mathbb{E}[\gamma(X_t^{u,\xi})])dB_t + \phi_t d\xi_t \\ \\ dY_t^{u,\xi} = -f(t, X_t^{u,\xi}, \mathbb{E}[\zeta(X_t^{u,\xi})], Y_t^{u,\xi}, \mathbb{E}[\eta(Y_t^{u,\xi})], u_t)dt + Z_t^{u,\xi} dB_t \\ \qquad\qquad + d\mathcal{M} - \varphi_t d\xi_t \\ \\ X_0^{u,\xi} = x_0, Y_T^{u,\xi} = h(X_T^{u,\xi}, \mathbb{E}[\theta(X_T^{u,\xi})]), \quad t \in [0,T], \end{cases} \tag{1}$$

where $b, \alpha, \gamma, \zeta, \eta, \theta, \phi$, and ψ are given functions, $(B_t, t \geq 0)$ is a standard Brownian motion, defined on some filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, satisfying the usual conditions. $\mathcal{M}$ a square integrable martingale that is orthogonal to B. The control variable is a suitable process (u, ξ) such that $u : [0,T] \times \Omega \to U_1 \subset \mathbb{R}^n, \xi : [0,T] \times \Omega \to U_2 = ([0,\infty))^l$ are $\mathcal{B}[0,T] \otimes \mathcal{F}$-measurable, $\mathcal{F}_t$-adapted, and ξ is an increasing process and cádlág, with $\xi_0 = 0$ and $\mathbb{E}[|\xi_T|^2] < +\infty$.

We shall consider a functional cost to be minimized, over the set admissible controls, as the following:

$$\begin{aligned} J(u_., \xi_.) := \mathbb{E}\big[\Psi\left(X_T^{u,\xi}, \mathbb{E}[\lambda(X_T^{u,\xi})]\right) + \Phi\left(Y_0^{u,\xi}, \mathbb{E}[\rho(Y_0^{u,\xi})]\right) \\ + \int_0^T g\left(t, X_t^{u,\xi}, \mathbb{E}[\pi(X_t^{u,\xi})], Y_t^{u,\xi}, \mathbb{E}[\varpi(Y_t^{u,\xi})], u_t\right) dt + \int_0^T \psi_t d\xi_t\big], \end{aligned} \tag{2}$$

where $\Phi, \lambda, \Psi, \rho, g, \pi, \psi$ and ϖ are appropriate functions.

Our objective is to minimize the functional cost J over the set of admissible controls $\mathcal{U}_1 \times \mathcal{U}_2$. An admissible control $(u_.^*, \xi_.^*)$ is called optimal if it satisfies $J(u_.^*, \xi_.^*) = inf\{J(u_., \xi_.); u_. \in \mathcal{U}_1, \xi_. \in \mathcal{U}_2\}$.

The considered system and the cost functional, depend not only on the state of the system, but also on the distribution of the state process, via the expectation of some function of the state. The mean-field FBSDEs (1) called McKean-Vlasov systems are obtained as the mean square limit of an interacting particle system of the form

$$\begin{cases} dX_t^{u,\xi,i,n} = b(t, X_t^{u,\xi,i,n}, \frac{1}{n}\sum\limits_{j=1}^{n}\alpha(X_t^{u,\xi,j,n}), u_t)dt \\ \qquad\qquad\qquad\qquad +\sigma(t, X_t^{u,\xi,i,n}, \frac{1}{n}\sum\limits_{j=1}^{n}\gamma(X_t^{u,\xi,j,n}))dB_t + \phi_t^i d\xi_t^i \\ \\ dY_t^{u,\xi,i,n} = -f(t, X_t^{u,\xi,i,n}, \frac{1}{n}\sum\limits_{j=1}^{n}\zeta(X_t^{u,\xi,j,n}), Y_t^{u,\xi,i,n}, \frac{1}{n}\sum\limits_{j=1}^{n}\eta(Y_t^{u,\xi,j,n}), u_t)dt \\ \qquad\qquad\qquad\qquad +Z_t^{u,\xi,i,n}dB_t^i + d\mathcal{M}_t^i - \varphi_t^i d\xi_t^i, \end{cases}$$

where $\left(B_\cdot^i, i \geq 0\right)$ is a collection of independent Brownian motion and $\frac{1}{n}\sum\limits_{j=1}^{n}\zeta(X_t^{u,\xi,j,n})$ denotes the empirical distribution of the individual players' state at time $t \in [0,T]$. Our system MF-FBSDEs (1) occur naturally in the probabilistic analysis of financial optimization and control problems of the McKean-Vlasov type.

One our main aims in first part of this thesis is to prove existence of optimal control for systems of FBSDEs of mean-field type (1) with singular controls. We prove in first, the existence of optimal relaxed controls, for MF-FBSDEs systems and when the Roxin convexity condition is fulfilled, we prove that the optimal relaxed control is in fact strict. This result is a generalization of the result given in [27], to the mean-field context and in a different topological setting.

A new class of stochastic differential equations with terminal condition, called Backward doubly stochastic differential equation (BDSDE) was introduced in 1994 by Pardoux and Peng in [39]. The authors show existence and uniqueness for this kind of stochastic equation and produce a probabilistic representation of certain quasi-linear stochastic partial differential equations (SPDE) extending the Feynman-Kac formula for linear SPDEs.

In this subject, Gherbal in [26] proved for the first time the existence of a strong optimal strict control for systems of linear backward doubly SDEs and established necessary as well as sufficient optimality conditions in the form of a stochastic maximum principle for this kind of systems.

To our best knowledge, no results in the literature have studied the existence of optimal control for systems of non linear forward-backward doubly SDEs. Which is in indeed a very difficult interesting research problem.

Our main goal in the second part of this thesis is to prove existence of optimal relaxed control as well as existence of optimal strict control for systems of forward-backward doubly SDEs.

In particular, we shall prove the existence of weak optimal solution (more precisely, a solution defined on an extended probability space) of a control problem driven by the following FBDSDE

$$\begin{cases} X_t = x + \int_0^t B\left(s, X_s, u_s\right) ds + \int_0^t \Sigma\left(s, X_s\right) dW_s, \\ \\ Y_t = H\left(X_T\right) + \int_t^T F\left(s, X_s, Y_s, u_s\right) ds + \int_t^T G\left(s, X_s, Y_s\right) d\overleftarrow{B_s} \\ \\ \qquad - \int_t^T Z_s dW_s - \left(M_T - M_t\right). \end{cases} \tag{3}$$

The mappings B, Σ, F, G and H are given, $(W_s)_{s\geq 0}$ and $(B_s)_{s\geq 0}$ be two mutually independent standard Brownian motions, defined on a probability space $(\Omega, \mathcal{F}, P)$, taking their values respectively in $\mathbb{R}^d$ and in $\mathbb{R}^k$, $u.$ represents the control variable. The processes $(X., Y., Z., M.)$ are valued in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \mathbb{R}^m$ and $M.$ is a square integrable martingale which is orthogonal to W, with $M_0 = 0$ and with càdlàg trajectories.

The integral with respect to B_s is a backward Itô integral, while the integral with respect to W_s is a standard forward Itô integral.

The functional cost to be minimized, over the set of admissible strict controls, is giving by:

$$J(u.) := \mathbb{E}\big[\varphi(X_T) + \psi(Y_0) + \int_0^T L(t, X_t, Y_t, u_s) dt\big]. \tag{4}$$

A weak solution for FBSDEs is given by Bahlali et al. [15], where the original probability is changed using Girsanov's theorem. See also Antonelli and Ma, [4] and Delarue and Guatteri,

[22], where the change of probability space comes from the construction of the forward component.

A weak solution for FBSDEs where the filtration is enlarged, have been studied by Buckdahn et al. [9], (see also [7] and [8]), using pseudopaths and the Meyer-Zheng topology, [43]. The success of Meyer-Zheng topology comes from a tightness criteria which is easily satisfied and ensures that all limits have their trajectories in the Skorokhod space $\mathbb{D}$. We use here the fact that Meyer-Zheng's criteria also yields tightness for the S-topology of Jakubowsky on $\mathbb{D}$, [31].

The existence of the orthogonal component $M.$ in our work comes from the fact that the optimal control and the corresponding solutions are obtained by taking weak limits of a minimizing sequence and the corresponding strong solution.

A weak solutions for FBSDEs where the Brownian filtration is enlarged, have been studied by Bouchemella and Raynaud de Fitte [17] and the system there also contains an orthogonal martingale component similar to ours. There is a good discussion there about the weak solutions.

One of our main aims in this second part of thesis is to prove in first, the existence of optimal relaxed controls, for the system of FBDSDE (3). The existence result is proved by using the tightness properties of the distributions of the processes defining the control problem and the Skorokhod's selection theorem on the space of Skorokhod $\mathbb{D}$, endowed with the Jakubowski S-topology.

Secondly, when the Roxin convexity condition is fulfilled, we prove that the optimal relaxed control is in fact strict and the set of admissible strict controls coincides with that of relaxed controls.

As a motivation to study a stochastic relaxed control problem for systems of FBDSDEs: is the fact that the FBDSDEs in their nature include FBSDEs, BSDEs and SDEs, which are widely used in mathematical economics and finance. FBSDEs, in particular, are encountered in stochastic recursive utility optimization problems, see [21], [34] and [45] for more details and applications.

Concerning the implementation of relaxed controls in such controlled systems, Anderson and

Djehiche [3], gives a financial application of relaxed control problem of linear SDE by studying an optimal bond portfolio problem in a market where there exists a continuum of bonds and the portfolio weights are modeled as measure-valued processes on the set of times to maturity. We emphasize also that dealing with relaxed controls is always a good choice when not having a convex domain of controls. There is a good discussion in Ahmed and Charalambous, [2], about this issue particularly and also the extension of the stochastic maximum principle to relaxed controls. We organize the dissertation as follows. Our contribution in this thesis touch on a very important aspect of optimal stochastic control which is the existence of optimal relaxed controls as well as the existence of optimal strict controls. We will present in what follows a brief description of the main results we have achieved in this work.

Chapitre (1).

The aim of this chapter is to introduce the concept of SDE and BSDE and to present the terminology used in this context. We also present the necessary conditions for the strong existence and the weak existence of the solutions. We give also a brief summary of equipping the space of càdlàg functions by the Skorokhod M1 and S-topology and their associated effects, and we will also discuss the relationship between them in the convergence and tightness. In fact the main reason for including this material here is to introduce some specific tools which will be used systematically in later chapters. It also unifies terminology and notation that are to be used in this chapters.

Chapitre (2). (The results of this chapter were the subject of a paper submitted to an international journal).

In the second chapter we deal with the problem of existence of optimal control for a singular control problem for systems of forward-backward stochastic differential equations of mean-field type (MF-FBSDEs) in which the control variable consists of two components: an absolutely continuous control and a singular one. The coefficients of the system depend on the states of the solution processes as well as their distribution via the expectation of some function. Moreover the cost functional is also of mean-field type. We prove in particular, the existence of optimal relaxed-singular controls, which are a couple of measure-valued processes and singular control as well as the existence optimal strict controls.

Chapitre (3). (The results of this chapter were the subject of a paper published in Random Operators & Stochastic Equations, 2020).

In the third chapter we study a class of optimal controls for problems governed by forward-backward doubly stochastic differential equations (FBDSDEs). Firstly, we prove existence of optimal relaxed controls, which are measure-valued processes for nonlinear FBDSDEs, by using some tightness properties and weak convergence techniques on the space of Skorokhod $\mathbb{D}$ equipped with the S-topology of Jakubowsky. Moreover, when the Roxin-type convexity condition is fulfilled, we prove that the optimal relaxed control is in fact strict.

CHAPTER 1

General Introduction of Stochastic Analysis.

CHAPTER 1

General Introduction of Stochastic Analysis.

Stochastic calculus serves as a fundamental tool throughout this thesis. The goal of this chapter is to introduce several tools from the theory of probability theory and stochastic processes, which will be useful in later chapters. It also unifies terminology and notation that are to be used in this chapters.

This Chapter is organized to into six sections in which we present the basic concepts and results fundamental of Stochastic calculus as follow:

- **In section one**, we present a some *introductory probability* will be briefly reviewed.
- **In section two**, we recall some results on *stochastic processes*.
- **In section three**, we are going to define the integral of type *Itô integral*.
- **In section four and five,** we study *stochastic differential equations* (SDEs) and (BSDEs).
- **In section six,** we give a brief summary of equipping the space of càdlàg functions by some topologies.

1.1 Introductions of Probability Theory

Let Ω be ordinary nonempty set and we denote by $\Gamma(\Omega)$ the set of parts of the set Ω.

Definition 1.1.1 *We say that $\mathcal{F} \subset\Gamma(\Omega)$ is a σ-algebra (or field) if*

1) Ω (or ϕ) belong to $\mathcal{F}$.

2) $\forall B \in \mathcal{F} \Rightarrow B^c \in \mathcal{F}$.

3) $(B_k)_{k\in\mathbb{N}} \subset \mathcal{F} \Rightarrow \cup_{k\in\mathbb{N}} B_k \in \mathcal{F}$.

Example 1.1.2 $\Gamma(\Omega)$ *the set of all collections of Ω and $\{\phi, \Omega\}$ are σ-algebras.*

A set Ω endowed with a σ-algebra $\mathcal{F} \subset\Gamma(\Omega)$ is called a measurable space. Members of $\mathcal{F}$ are called measurable sets.

Proposition 1.1.3 *Let $C\subset\Gamma(\Omega)$. Then there exists a unique σ-algebra, $\sigma(C)$ generated by C, which is the unique smallest σ-algebra containing C.*

Example 1.1.4 *The σ-algebra generated by $C = \{B\} \subset \Omega$ is $\sigma(C) = \{\phi, \Omega, B, B^c\}$.*

The σ-algebra generated by all open sets in $\mathbb{R}^d$ is called the Borel σ-algebra of $\mathbb{R}^d$ and is denoted by $\mathcal{B}(\mathbb{R}^d)$. It is defined to be the smallest σ-algebra containing all such open subsets of $\mathbb{R}^d$.

Definition 1.1.5 *we call measure μ on the measurable space $(\Omega, \mathcal{F})$ an application of $\mathcal{F}$ in $[0, \infty]$ which verifies*

1) $\mu(\phi) = 0$.

2) Let the family $(B_k)_{k\in\mathbb{N}} \subset \mathcal{F}$ is disjoint (i.e: $B_i \cap B_j = \phi$ if $i \neq j$) , then

$$\mu\left(\cup_{k\in\mathbb{N}} B_k\right) = \sum_{k\geq 0} \mu\left(B_k\right).$$

The triplet $(\Omega, \mathcal{F}, \mu)$ is called a measure space.

Remark 1.1.6 *If $\mu(\Omega) = 1$, we say that μ is a probability measure and $(\Omega, \mathcal{F}, \mu)$ is a probability space.*

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let

$$\mathcal{N} = \{A \subset \Omega / \exists B \in \mathcal{F},\ P(B) = 0 \text{ and } A \subset B\},$$

be the collection of null probability sets. We then say that the probability space $(\Omega, \mathcal{F}, P)$ is complete if $\mathcal{N} \subset \mathcal{F}$. In this thesis the probability space $(\Omega, \mathcal{F}, P)$ will always be considered as completed by the collection of null probability sets.

Now, if E is a topological space, let $\mathcal{B}(E)$ denote its Borel σ-field over E, generated by the family of open subsets of E. The mapping $X : \Omega \to E$ is an random variable E-valued. If $E = \mathbb{R}$ then X will be called a real random variable. If $E = \mathbb{R}^d$ then X will be called a d-dimensional random variable.

The probability measure $P_X : \mathcal{B}(E) \to [0,1]$ is called the law of X, defined by $P_X(B) = P(X \in B)$.

Let $X : (\Omega, \mathcal{F}, P) \to E$ be a random variable, the

$$\sigma(X) = \left\{X^{-1}(B) / B \in \mathcal{B}(E)\right\},$$

is called the σ-algebra generated by X, is the smallest -algebra which makes X measurable. Also $\sigma(X)$ represents the information carried by X. Let $\{X_k,\ k \in I\}$ family of random variables, we can associate the σ-algebra $\sigma(X_k,\ k \in I)$, which is the smallest σ-algebra containing $\cup_{k\in I}\sigma(X_k)$.

Let $\mathcal{F}$, $\mathcal{G}$ be two σ-algebras. We say that $\mathcal{F}$ and $\mathcal{G}$ are independent if

$$P(A \cap B) = P(A).P(B), \forall A \in \mathcal{F},\ B \in \mathcal{G}.$$

Let $X, Y : \Omega \to E$ be two random variables and $\mathcal{F}$ a σ-field on Ω. Then X is said to be independent of $\mathcal{F}$ if $\sigma(X)$ is independent of $\mathcal{F}$, and X is said to be independent of Y if $\sigma(X)$ and $\sigma(Y)$ are independent.

Let $\{\mathcal{F}_k, k \in I\}$ be a family of σ-fields on Ω. We define

$$\vee_{k\in I}\mathcal{F}_k = \sigma\left(\cup_{k\in I}\mathcal{F}_k\right),$$

and

$$\wedge_{k\in I}\mathcal{F}_k = \cap_{k\in I}\mathcal{F}_k.$$

It is easy to show that $\vee_{k\in I}\mathcal{F}_k$ and $\wedge_{k\in I}\mathcal{F}_k$, are both σ-fields and that they are the smallest σ-field containing all $\mathcal{F}_k$ and the largest σ-field contained in all $\mathcal{F}_k$ respectively.

Remark 1.1.7 *Note that if* $(E, \mathcal{B}(E)) = \left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$, *then*

$$\sigma\left(X_k,\ k \in I\right) = \vee_{k\in I} X^{-1}\left(\mathcal{B}\left(\mathbb{R}^d\right)\right)$$

Now, we call the expectation (or mean) of a random variable is the integral

$$\mathbb{E}[X] = \int_\Omega X(\omega)\, dP(\omega),$$

and the variance of X by

$$Var[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2.$$

We can use the notation $P(d\omega)$ for $dP(\omega)$. For all $B \in \mathcal{F}$ we define

$$\mathbb{E}[X1_B] = \int_B X(\omega)\, dP(\omega),$$

where

$$1_B(\omega) = \begin{cases} 1 \text{ if } \omega \in B \\ 0 \text{ if } \omega \notin B. \end{cases}$$

Example 1.1.8 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $X, Y : \Omega \to \mathbb{R}$ be two independent ramdom variables. Then*

$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y].$$

Let us show that if $X = 1_A$ and $Y = 1_B$, for some $A, B \in \mathcal{F}$, then

$$\begin{aligned} \mathbb{E}[XY] = \mathbb{E}[1_A 1_B] &= \int_\Omega 1_{A\cap B}(\omega)\, dP(\omega) \\ &= P(A\cap B) = P(A).P(B) \\ &= \mathbb{E}[X]\,\mathbb{E}[Y]. \end{aligned}$$

Let $X^n, X : (\Omega, \mathcal{F}, P) \to \mathbb{R}^d$, $n \in \mathbb{N}$, be two random variables. The following types of convergence will be used in this thesis:

- $X^n \to X$ $a.s$ if there exists an $A \in \mathcal{F}$, $P(A) = 1$ where

$$\lim_{n\to\infty} X^n(\omega) = X(\omega), \text{ for all } \omega \in A.$$

- $X^n \to X$ in probability if for all $\varepsilon > 0$, we have

$$\lim_{n\to\infty} P(|X^n - X| \geq \varepsilon) = 0.$$

For all $p > 0$, we denote by $L^p\left(\Omega, \mathcal{F}, P, \mathbb{R}^d\right)$, the linear space of random variables $X : \Omega \to \mathbb{R}^d$ such that $\mathbb{E}\left(|X_t|^p\right) < \infty$.

- $X^n \to X$ in $L^p\left(\Omega, \mathcal{F}, P, \mathbb{R}^d\right)$, if $X^n, X \in L^p\left(\Omega, \mathcal{F}, P, \mathbb{R}^d\right)$ and

$$\lim_{n\to\infty} \mathbb{E}\left[|X^n - X|^p\right] = 0.$$

- $X^n \to X$ in law if for any bounded continuous function $\varphi : \mathbb{R}^d \to \mathbb{R}$

$$\lim_{n\to\infty} \mathbb{E}\left[\varphi\left(X^n\right)\right] = \mathbb{E}\left[\varphi\left(X\right)\right].$$

It is well-known that if $X^n \to X$ in probability, then $\varphi\left(X^n\right) \to \varphi\left(X\right)$, for every $\varphi \in \mathcal{C}\left(\mathbb{R}^d\right)$, and there exists a subsequence $X^{n_k} \to X$ $a.s.$ Also, if $X^n \to X$ in probability then $X^n \to X$ in law.

Remark 1.1.9 *The various concepts of convergence of random variables with values in a metric space are defined analogously.*

Theorem 1.1.10 (Monotone Convergence) *Let X^n, X be two random variables, $n \in \mathbb{N}$. If*

i) $0 \leq X^1 \leq X^2 ... \leq X^n \leq ... \leq X$ $P - a.s.$

ii) $X^n \to X$ $a.s$, *then*

$$\lim_{n\to\infty} \mathbb{E}\left[X^n\right] = \mathbb{E}\left[X\right].$$

Theorem 1.1.11 (Fatou's Lemma) *Let X^n, X be two random variables and X^n positive $a.s$, for all $n \in \mathbb{N}^*$. Then*

$$\mathbb{E}\left[\liminf_{n\to\infty} X^n\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[X^n\right].$$

If moreover $X^n \to X$ in law, then X positive $a.s$ and

$$\mathbb{E}\left[X\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[X^n\right].$$

Theorem 1.1.12 (Dominated Convergence Theorem) *Let X^n, X be two d-dimensional random variables, $n \in \mathbb{N}^*$. If $X^n \to X$ in law and there exists a positive random variable $Y \in L^1(\Omega, \mathcal{F}, P)$ such that for all $n \in \mathbb{N}^* : |X^n| \leq Y \ a.s$, then*

$$\mathbb{E}[|X|] < \infty, \ \textit{and} \ \lim_{n\to\infty} \mathbb{E}[X^n] = \mathbb{E}[X].$$

1.1.1 Conditional expectation

In this Subsection, we present the notion of conditional expectation, i'ts main properties, which will be needed in the next sections and chapters.

All random variables will be assumed to be defined on a probability space $(\Omega, \mathcal{F}, P)$. We denote by $\mathbf{G}$, $\mathbf{H}$ two sub-σ-algebras of $\mathcal{F}$, and we will assume that each of them contains the collection $\mathcal{N}$ of all P-null sets of $\mathcal{F}$.

The space $L^2(\Omega, \mathbf{G}, P, \mathbb{R}^d)$ is a sub-Hilbert space of $L^2(\Omega, \mathcal{F}, P, \mathbb{R}^d)$, with respect to the scalar product $\mathbb{E}[\langle X, Y\rangle]$. Therefore, we can define the conditional expectation as follows:

Definition 1.1.13 *We call the conditional expectation with respect to $\mathbf{G}$ the orthogonal projection operator from $L^2(\Omega, \mathcal{F}, P, \mathbb{R}^d)$ into $L^2(\Omega, \mathbf{G}, P, \mathbb{R}^d)$.*

So be it X random variable square integrable ($X \in L^2(\Omega, \mathcal{F}, P, \mathbb{R}^d)$), the orthogonal projection on $L^2(\Omega, \mathbf{G}, P, \mathbb{R}^d)$ is the conditional expectation, denote by

$$\mathbb{E}[X \mid \mathbf{G}].$$

In fact, $\mathbb{E}[X \mid \mathbf{G}]$ is the unique $\mathbf{G}$-random variable satisfying

$$\begin{cases} \mathbb{E}[X \mid \mathbf{G}] \text{ is } \mathbf{G}\text{-measurable} \\ \mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[X \mid \mathbf{G}] Y], \ \forall Y \in L^2(\Omega, \mathbf{G}, P, \mathbb{R}^d). \end{cases}$$

In particular, if $Y = 1_\Omega$, we get

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \mathbf{G}]].$$

Remark 1.1.14 *If* $\mathbf{G} = \{\phi, \Omega\}$, *then* $\mathbb{E}[X] = \mathbb{E}[X \mid \mathbf{G}]$.

Now, we can established that the conditional expectation is continuous with respect to the norm in the space $L^1(\Omega, \mathcal{F}, P, \mathbb{R}^d)$, for any square integrable X, we have

$$\begin{cases} X = X^+ - X^- \\ |X| = X^+ + X^-, \end{cases}$$

then by linearity of the orthogonal projection

$$\begin{aligned} \mathbb{E}[X \mid \mathbf{G}] &= \mathbb{E}\left[X^+ \mid \mathbf{G}\right] - \mathbb{E}\left[X^- \mid \mathbf{G}\right] \\ |\mathbb{E}[X \mid \mathbf{G}]| &\leq \mathbb{E}\left[X^+ \mid \mathbf{G}\right] + \mathbb{E}\left[X^- \mid \mathbf{G}\right] \\ &\leq \mathbb{E}[|X| \mid \mathbf{G}], \end{aligned}$$

and hence

$$\mathbb{E}[|\mathbb{E}[X \mid \mathbf{G}]|] \leq \mathbb{E}[|X|].$$

Now, we us collect main properties of the conditional expectation in the following proposition.

Proposition 1.1.15 *Let* $X, Y \in L^1(\Omega, \mathcal{F}, P, \mathbb{R}^d)$, *and* $(\alpha, \beta) \in \mathbb{R}$, *then:*

1)-$X \geq 0 \ a.s \Longrightarrow \mathbb{E}[X \mid \mathbf{G}] \geq 0 \ a.s.$

2)-$\mathbb{E}\left[\alpha X+\beta Y\mid \mathbf{G}\right]=\alpha\mathbb{E}\left[X\mid \mathbf{G}\right]+\beta\mathbb{E}\left[Y\mid \mathbf{G}\right].$

3)-If X is $\mathbf{G}$-measurable and $\langle X,Y\rangle$ is integrable, then

$$\mathbb{E}\left[\langle X,Y\rangle\mid \mathbf{G}\right]=\langle X,\mathbb{E}\left[Y\mid \mathbf{G}\right]\rangle.$$

4)-If $\mathbf{G}\subset\mathbf{H}$, then

$$\mathbb{E}\left[\mathbb{E}\left[X\mid \mathbf{H}\right]\mid \mathbf{G}\right]=\mathbb{E}\left[X\mid \mathbf{G}\right]$$

5)-If $\mathbf{H}$ and $\mathbf{G}\vee\sigma\left(X\right)$ are independent, then

$$\mathbb{E}\left[X\mid \mathbf{G}\vee\mathbf{H}\right]=\mathbb{E}\left[X\mid \mathbf{G}\right].$$

In particular, if $\mathbf{G}$ and X are independent, then

$$\mathbb{E}\left[X\mid \mathbf{G}\right]=\mathbb{E}\left[X\right].$$

6)-(Jensen's inequality) Let $\varphi:\mathbb{R}^d\to\mathbb{R}$ is a convex function such that $\varphi\left(X\right)\in L^1\left(\Omega,\mathcal{F},P,\mathbb{R}\right)$. Then

$$\varphi\left(\mathbb{E}\left[X\mid \mathbf{G}\right]\right)\leq\mathbb{E}\left[\varphi\left(X\right)\mid \mathbf{G}\right]\ \ P\ a.s.$$

Corollary 1.1.16 *If $p\geq 1$, then $\mathbb{E}\left[\cdot\mid \mathbf{G}\right]$ is a linear continuous operator from $L^p\left(\Omega,\mathcal{F},P,\mathbb{R}^d\right)$ into $L^p\left(\Omega,\mathbf{G},P,\mathbb{R}^d\right)$.*

Proof. From the property (6) with $\varphi\left(X\right)=\left|X\right|^p$, we obtain

$$\mathbb{E}\left[\left|\mathbb{E}\left[X\mid \mathbf{G}\right]\right|^p\right]\leq\mathbb{E}\left[\left|X\right|^p\right].$$

■

Lemma 1.1.17 *Let $Y : (\Omega, \mathcal{F}, P) \to \mathbb{R}^d$ be a random variable, and $X \in L^1\left(\Omega, \mathcal{F}, P, \mathbb{R}^d\right)$. If $\mathbb{E}\left[X\psi\left(Y\right)\right] = 0$ for any $\psi \in C_b\left(\mathbb{R}^d\right)$, then we have*

$$\mathbb{E}\left[X 1_A\right] = 0, \ \forall A \in \sigma\left(Y\right).$$

Proposition 1.1.18 *Let $Y_k : (\Omega, \mathcal{F}, P) \to \mathbb{R}^{m_k}$ $\forall k \geq 1$, be a sequence of random variables. Let $\mathbf{G} = \vee_{k\geq 1} \sigma\left(Y_k\right)$ and $X \in L^1\left(\Omega, \mathcal{F}, P, \mathbb{R}^m\right)$. Then $\mathbb{E}\left[X \mid \mathbf{G}\right] = 0$ if and only if $\mathbb{E}\left[X\psi\left(Y_1, Y_2, ..., Y_k\right)\right] = 0$ for any $k \geq 1$ and $\forall \psi \in \mathcal{C}_b\left(\mathbb{R}^n\right)$, $n = \sum_{i=1}^k m_i$.*

Now, we turn to the conditional probability. We define

$$P\left(B \mid \mathbf{G}\right) = \mathbb{E}\left[1_B \mid \mathbf{G}\right], \ \forall B \in \mathcal{F},$$

which is called the conditional probability of the event A given the condition $\mathbf{G}$. Note that, if

$$\begin{cases} P\left(\phi \mid \mathbf{G}\right) = 0, \ P\left(\Omega \mid \mathbf{G}\right) = 1 \\ P\left(\cup_{k\geq 0} B_k \mid \mathbf{G}\right) = \sum_{k\geq 1} P\left(B_k \mid \mathbf{G}\right), \ \forall B_k \in \mathcal{F}, \text{and } \forall i \neq j \ B_i \cap B_j = \phi. \end{cases}$$

Remark 1.1.19 *For a given $\omega \in \Omega$, $P\left(\cdot \mid \mathbf{G}\right)(\omega)$ is not necessarily a probability measure on $\mathcal{F}$.*

Definition 1.1.20 *A separable complete metric space is called a Polish space.*

Proposition 1.1.21 *Let $(\Omega, \mathcal{F}, P)$ a probability space. If Ω is a Polish space and $\mathbf{G} \subset \mathcal{F}$, then there is a map $p : \Omega \times \mathcal{F} \to [0, 1]$, called a regular conditional probability, such that $p\left(\omega, \cdot\right)$ is a probability measure on $(\Omega, \mathcal{F})$ for any $\omega \in \Omega$, and $p\left(\cdot, B\right)$ is $\mathbf{G}$-measurable for any $B \in \mathcal{F}$,*

$$\mathbb{E}\left[1_B \mid \mathbf{G}\right](\omega) = P\left(B \mid \mathbf{G}\right)(\omega) = p\left(\omega, B\right), \ \forall B \in \mathcal{F}.$$

Remark 1.1.22 *If $\acute{p}$ is another regular conditional probability* $\mathbf{G}$*-measurable, then there exists a* P*-null set* $A \in \mathbf{G}$ *such that for any* $\omega \notin A$,

$$p(\omega, B) = \acute{p}(\omega, B), \ \forall B \in \mathcal{F},$$

in the sense that it is unique.

1.1.2 Convergence of probabilities

To prove some existence results in stochastic analysis and stochastic controls, it inevitably involves convergence of certain probabilities. We summarize in this subsection some important results concerning the convergence of probability measures.

Let (U, ρ) be a separable metric space and $\mathcal{B}(U)$ the Borel σ-field. We denote by $P(U)$ the set of all probability measures on the measurable space $(U, \mathcal{B}(U))$.

Definition 1.1.23 *i)-Let* $(Q^n)_{n\geq 0} \subset P(U)$, *we say that* Q^n *converges weakly to* $Q \in P(U)$, *if*

$$\lim_{n\to\infty} \int_{\mathbb{U}} \psi(u)\, Q^n(u) = \int_{\mathbb{U}} \psi(u)\, Q(u).$$

for all, bounded continuous function ψ *on* U, $(\forall \psi \in \mathcal{C}_b(U))$.

ii)-Let $(Q^n)_{n\geq 0} \subset P(U)$, *we say that* Q^n *converges vaguely to* Q, *if*

$$\lim_{n\to\infty} \int_{\mathbb{U}} \psi(u)\, Q^n(u) = \int_{\mathbb{U}} \psi(u)\, Q(u).$$

for all, continuous function with compact support ψ *on* U, $(\forall \psi \in \mathcal{C}_c(U))$.

Proposition 1.1.24 *There is a metric* d *on* $P(U)$ *such that* Q^n*converges weakly to* Q, *is equivalent to* $d(Q^n, Q) \to 0 \ a.s \ n \to \infty$.

Definition 1.1.25 *Let K be a subset of $P(U)$.*

i)- A set K is said to be relatively compact if any sequence $(Q^n)_{n\geq 0} \subset K$ contains a weakly convergent subsequence.

ii)-A set K is said to compact, if K is relatively compact and closed.

ii)-A set K is said to tight, if for any $\varepsilon > 0$ there is a compact set $\Sigma \subset U$ such that

$$Q(\Sigma) \geq 1 - \varepsilon, \quad \forall Q \in K.$$

Proposition 1.1.26 *Let K be a subset of $P(U)$. Then:*

i)-K is relatively compact if it is tight.

ii)-If (U, ρ) is complete (i.e: it is a polish space), then K is tight if it is relatively compact.

Corollary 1.1.27 *If (U, ρ) is compact, then any $K \subset P(U)$ is tight and relatively compact. In particular, $P(U)$ is compact.*

Proof. Straightforward by Definition (18) and Proposition (19). ∎

Theorem 1.1.28 (Prokhorov's) *Let (U, ρ) be a complete separable metric space, and let K be a subset of $P(U)$. Then the following two statements are equivalent:*

a)-K is compact in $P(U)$.

b)-K is tight.

Now, if X is a random variable definite from $(\Omega, \mathcal{F}, P)$ into (U, ρ), we denote by $Q = P_X \in P(U)$ the probability induced by X. We say that a family of random variables $(X^n)_{n\geq 0}$ is tight if $(P_{X^n})_{n\geq 0}$ is tight. As we say that X^n, converges to X in law if P_{X^n} converges weakly to P_X as $n \to \infty$.

1.2 Stochastic Processes

In this section we recall some results on stochastic processes.

Definition 1.2.1 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A family $(X_t)_{t\in I}$, $I \subset \mathbb{R}$ of functions from $\Omega \times I$ into $\mathbb{R}^d$ is called a stochastic process. For all $t \in I$, $X(\cdot, t)$ is an $\mathbb{R}^d$-valued random variable. The mappings $X(\omega, \cdot)$, $\omega \in \Omega$ are called the trajectories of the stochastic process X.*

Now, we give a stochastic process $(Y_t)_{t\in[0,T]}$, we say that Y_t is a modification of X_t if for all $t \in [0, T]$, we have

$$Y_t = X_t \; P - a.s, \; \forall t \in [0, T]$$

In this case, one is called a modification of the other. In the same context we say that Y_t is indistinguishable from X_t if their paths coincide $a.s$, then for any $t \in [0, T]$, there exists a P-null set $B_t \in \mathcal{F}$ such that

$$Y_t = X_t \; \forall \omega \in \Omega \setminus B_t.$$

Definition 1.2.2 *The stochastic process X_t is said to be càdlàg if for each $\omega \in \Omega$, the path X_t is right continuous and admits a left-limit i.e:*

$$\begin{cases} \textit{for } 0 \leq t < T, \; X_{t+} \equiv \lim_{s \searrow t} X_s \textit{ exists with } X_{t+} = X_t, \\ \textit{for } 0 \leq t \leq T, \; X_{t-} \equiv \lim_{s \nearrow t} X_s \textit{ exists.} \end{cases}$$

Obviously, the notion of indistinguishability is stronger than the one of modification, but if the two processes X_t and Y_t are càdlàg, and if Y_t is a modification of X_t, then X_t and Y_t are indistinguishable.

Definition 1.2.3 *We say that X_t is stochastically continuous at $t \in [0,T]$ if for any $\varepsilon > 0$*

$$\lim_{s \to t} P\left(|X_s - X_t| > \varepsilon\right) = 0.$$

Moreover, X_t is said to be continuous if there exists a P-null set $B \in \mathcal{F}$ such that $\forall \omega \in \Omega \diagdown B$ the trajectories X_t is continuous.

Definition 1.2.4 *A filtration on $(\Omega, \mathcal{F}, P)$ is an increasing family $(\mathcal{F}_t)_{0 \leq t \leq T}$ of σ-fields of $\mathcal{F}$ such that $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $0 \leq s \leq t \leq T$.*

$\mathcal{F}_t$ is interpreted as the information known at time t, and increases as time elapses. Set $\mathcal{F}_{t+} = \cap_{s>t} \mathcal{F}_s$ for any $t \in [0,T)$, and $\mathcal{F}_{t-} = \cup_{s<t} \mathcal{F}_s$ for any $t \in (0,T]$. If $\mathcal{F}_{t+} = \mathcal{F}_t$ (resp.$\mathcal{F}_{t-} = \mathcal{F}_t$), we say that $(\mathcal{F}_t)_{0 \leq t \leq T}$ is right (resp. left) continuous. We call $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ a filtered probability space. Then we say that the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ satisfies the usual condition if $(\Omega, \mathcal{F}, P)$ is complete (i.e: $\mathcal{F}_0$ contains all the P-null sets in $\mathcal{F}$), and $(\mathcal{F}_t)_{0 \leq t \leq T}$ is right continuous. The completion of the filtration means that if an event is impossible, this impossibility is already known at time 0. The natural filtration (or canonical) of X_t is the smallest σ-field under which X_s is measurable for all $0 \leq s \leq t$ such that

$$\mathcal{F}_t^X = \sigma\left(X_s; 0 \leq s \leq t\right), \ \forall t \in [0,T].$$

$\mathcal{F}_t^X$ is called the history of the process X until time $t \geq 0$.

Definition 1.2.5 *A process $(X_t)_{0 \leq t \leq T}$ is adapted with respect to $\mathcal{F}$, if for all $t \in [0,T]$, X_t is $\mathcal{F}_t$-measurable.*

When one wants to be precise with respect to which filtration the process is adapted, we write $\mathcal{F}$-adapted. Thus, an adapted process is a process whose value at any time t is revealed by the information $\mathcal{F}_t$.

Definition 1.2.6 *The process X_t is progressively measurable with respect to $(\mathcal{F}_t)_{0\leq t\leq T}$, if for all $t \in [0,T]$, the map $(\omega, t) \to X_t(\omega)$ is $\mathcal{B}([0,T]) \times \mathcal{F}_t$-measurable.*

It is clear that if X_t is $(\mathcal{F}_t)_{0\leq t\leq T}$-progressive measurable, it must be measurable and $(\mathcal{F}_t)_{0\leq t\leq T}$-adapted. Conversely, on a filtered probability space, we have the following result.

Proposition 1.2.7 *Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ be a filtered probability space, and let X_t be measurable and $(\mathcal{F}_t)_{0\leq t\leq T}$-adapted. Then there exists an $(\mathcal{F}_t)_{0\leq t\leq T}$-progressively measurable process Y_t modification to X_t.*

Remark 1.2.8 *If X_t is left (or right) continuous, then X_t itself is $(\mathcal{F}_t)_{0\leq t\leq T}$-progressively measurable.*

Let $E_T^m = \mathcal{C}([0,T], \mathbb{R}^m)$ the space of continuous functions from $[0,T]$ to $\mathbb{R}^m$. We called a Borel cylinder $B \subset E_T^m$ if there exist $0 \leq t_1 \leq t_2....t_k \leq T$ and $A \in \mathcal{B}(\mathbb{R}^m)$ such that

$$B = \{Y \in E_T^m \setminus (Y_{t_1}, Y_{t_2}, ..., Y_{t_k}) \in A\}.$$

Now, let C_s be the set of all Borel cylinders in E_s^m with $t_1, t_2...., t_k \in [0,s]$.

Lemma 1.2.9 *The σ-field generated by C_T coincides with the Borel σ-field $\mathcal{B}(E_T^m)$ of E_T^m.*

Lemma 1.2.10 *Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $Y : [0,T] \times \Omega \to \mathbb{R}^m$ a continuous process. Then there exists an $\Omega_0 \in \mathcal{F}$ with $P(\Omega_0) = 1$ such that $Y : \Omega_0 \to E_T^m$ and for any $s \in [0,T]$,*

$$\mathcal{F}_t^Y = Y^{-1}(\sigma(\mathcal{B}(E_t^m))).$$

We now recall the very well known Kolmogorov's criterion for the existence of a continuous version of a process.

Theorem 1.2.11 *Let X_t be an m-dimensional stochastic process over $[0,T]$ such that*

$$\mathbb{E}\left[|X_t - X_s|^{\alpha}\right] < K\,|t-s|^{\beta+1}\,,\ \forall t,s\in[0,T]\,,$$

for some constants α, $\beta > 0$. Then there exists an m-dimensional continuous process that is modification to X_t.

1.2.1 Tightness of the laws of processes

In this subsection we are interested in studying of limit behavior of stochastic processes one often needs to know when a sequence of random variables is convergent in distribution or, at least, has a subsequence that converges in distribution or rather of the relatively compact. Which play an important role in appropriately formulating stochastic optimal control problems studied in the subsequent chapters.

Let $(\Omega,\mathcal{F},P)$ be a probability space and $X:\Omega\to E$ be a random variable, then the probability measure $P_X:\mathcal{B}(E)\to[0,1]$ defined by

$$P_X(B) = P(X\in B)\,,$$

is called the law of X.

Denote by $P(E)$ the space of probability measures, for all $Q\in P(E)$ there exists a probability space $(\Omega,\mathcal{F},P)$ and a random variable $X:\Omega\to E$ such that $P_X = Q$ where

$$\int_{\Omega} H(X(\omega))\,dP(\omega) = \int_E H(x)\,dQ(x)\,.$$

$X^n \to X, as\ n \to \infty$ in law, or weakly if $Q^n \to Q$ or equivalently

$$\lim_{n\to\infty} \mathbb{E}\left[\varphi\left(X^n\right)\right] = \mathbb{E}\left[\varphi\left(X\right)\right],$$

for all $\varphi \in \mathcal{C}_b\left(\mathbb{R}^d\right)$. The topology induced by weak convergence on $P\left(E\right)$, is called the weak topology.

Definition 1.2.12 *i) The family $\left(X^n\right)_{n\geq 0}$ is relatively compact in law, if every subsequence $\left(X^{n_k}\right)_{k>0}$ convergent in law, i.e: there exists a probability space $\left(\Omega, \mathcal{F}, P\right)$ and a random variable $X : \Omega \to E$ such that $X^{n_k} \to X, as\ n \to \infty$ in law.*

ii) A family $\left(X^n\right)_{n\geq 0}$, of random variables is tight, if for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset E$, such that for all $n \geq 0$

$$P^n\left(X^n \in K_\varepsilon\right) \geq 1 - \varepsilon.$$

Theorem 1.2.13 *Let $\left(E, \rho\right)$ be a metric space and $\left(X^n\right)_{n\geq 0}$ be a family of E-valued random variables.*

i) If $\left(X^n\right)_{n\geq 0}$ is tight then it is relatively compact in law (or weakly).

ii) Suppose that $\left(E, \rho\right)$ is a Polish space. If $\left(X^n\right)_{n\geq 0}$ is relatively compact in law (or weakly) then it is tight.

Remark 1.2.14 *A separable complete metric space is called a Polish space.*

1.2.1.1 The space of relaxed controls

We know that in stochastic control theory, in the absence of additional convexity conditions, an optimal control may fail to exist in the set U. There are many methods used to prove existence of an optimal control of which maximum principle or the minimizing sequence and compactification (classical method). The classical method is then to introduce generalized controls, i.e: to use a bigger new class its role is to compensate strict control set. This new class of processes is suitable for the needs of the theory of limit for random functions and have a richer structure of compacity and convexity. To be convinced on the fact that strict optimal controls may not exist even in the simplest cases let us consider adeterministic example

Exemple The problem is to minimize the following cost function

$$J(u) = \int_{[0,T]} (x_t^u)^2 \, dt,$$

over the set $\mathcal{U}_{ad}$ of measurable functions $u : [0,1] \to U = \{-1, 1\}$.

Let x_t^u be the unique solution of a system that evolves according to the following (SDE)

$$x_t^u = u_t dt, \; x_0^u = 0.$$

So be it u^n be a sequence of controls, that is $u_t^n = (-1)^k$ if $\frac{k}{n} \leq t \leq \frac{k+1}{n}$, $0 \leq k \leq n-1$. Then, $\forall t \in [0,1]$, $\left|x_t^{u^n}\right| \leq \frac{1}{n}$, and therefore $0 \leq J(u) \leq \frac{1}{n^2}$ which implies that

$$\inf_{u \in \mathcal{U}_{ad}} J(u) = 0.$$

In this case, there is not u which fulfilled $J(u) = 0$, it is obvious then that it implies that $x_t^u = 0$, $\forall t \in [0,1]$ if and only if $u_t = 0$ which is impossible. Accordingly, the problem is the fact that the sequence (u^n) lacks a limit in the control space $\mathcal{U}_{ad}$, the limit that must be the natural candidate for optimality. We identify (u_t^n) with the Dirac measure on U, to put a sequence of measures $q^n(dt, du) = \delta_{u_t^n}(du)\, dt$ over the space $[0,1] \times U$, which converges weakly to $q(dt, du) = \frac{1}{2}(\delta_{-1}(du) + \delta_1(du))\, dt$. For any bounded continuous function φ on $[0,1] \times U$ we have

$$\int_{[0,1]\times U} \varphi(t,u)\, q^n(dt, du) = \sum_{k=0}^{n-1} \int_{\left[\frac{k}{n}, \frac{k+1}{n}\right]} \varphi\left(t, (-1)^k\right) dt.$$

We assume that n is pair $(n = 2N)$, let it be $\varphi(t, -1)$, and $\varphi(t, 1)$ are continuous on $[0, 1]$, therefore uniformly continuous, so $\forall\varepsilon > 0$ there is an $M > 0$ such that $\forall N \geq M$

$$|\varphi(t, u) - \varphi(s, u)| < \varepsilon, \text{ if } |t - s| < \frac{1}{N},$$

then for every $k = 0, ..., 2N - 1$,

$$\left| \int_{\left[\frac{2k}{n}, \frac{2k+1}{n}\right]} \varphi(t, 1)\, dt - \int_{\left[\frac{2k+1}{n}, \frac{2k+2}{n}\right]} \varphi(t, -1)\, dt - 0 \right| < \frac{\varepsilon}{2N}$$

so

$$\left| \sum_{k=0}^{n-1} \int_{\left[\frac{2k}{n}, \frac{2k+1}{n}\right]} \varphi(t, 1)\, dt - \sum_{k=0}^{n-1} \int_{\left[\frac{2k+1}{n}, \frac{2k+2}{n}\right]} \varphi(t, -1)\, dt - 0 \right| < \frac{\varepsilon}{2}$$

on the other hand, we have

$$\sum_{k=0}^{n-1} \int_{\left[\frac{2k}{n}, \frac{2k+1}{n}\right]} \varphi(t, 1)\, dt + \sum_{k=0}^{n-1} \int_{\left[\frac{2k+1}{n}, \frac{2k+2}{n}\right]} \varphi(t, -1)\, dt = \int_{[0,1]} \varphi\left(t, (-1)^k\right) dt$$

and hence

$$\begin{cases} \left| \sum_{k=0}^{n-1} \int_{\left[\frac{2k}{n}, \frac{2k+1}{n}\right]} \varphi(t, 1)\, dt - \frac{1}{2} \int_{[0,1]} \varphi(t, 1)\, dt \right| < \frac{\varepsilon}{2} \\ \left| \sum_{k=0}^{n-1} \int_{\left[\frac{2k+1}{n}, \frac{2k+2}{n}\right]} \varphi(t, -1)\, dt - \frac{1}{2} \int_{[0,1]} \varphi(t, -1)\, dt \right| < \frac{\varepsilon}{2}, \end{cases}$$

then

$$\left| \int_{[0,1]} \varphi\left(t, (-1)^k\right) dt - \frac{1}{2} \int_{[0,1]} \varphi(t, 1) + \varphi(t, -1)\, dt \right| < \varepsilon.$$

Thus, we find the following

$$\left| \int_{[0,1] \times U} \varphi(t, u)\, q^n(dt, du) - \frac{1}{2} \int_{[0,1] \times U} \varphi(t, u) \left(\delta_{-1}(du) + \delta_1(du)\right) dt \right| < \varepsilon.$$

This indicates that the $\mathcal{U}_{ad}$-set of strict controls is narrow and should be embedded in a broader class with a richer topological structure in which the control problem becomes solvable.

And from it us introduce the concept of relaxed controls which gives a more suitable topological structure. Let $P(U)$ denote the space of probability measures equipped with the topology of weak convergence, then $P(U)$ is also compact space. In a relaxed control problem, the U-valued process u_t is replaced by an $P(U)$-valued processq_t.

Let $\mathbb{V}$ the space of probability measures on the set $[0,T]\times U$, such that its first marginal is coincide with the Lebesgue measure, and its second marginal is a probability distribution over U. The space $\mathbb{V}$ equipped with the boreal tribe, which is the smallest tribe such that the $q \to \int \varphi(t,u)\, q(dt,du)$ application is measurable for any measurable, bounded and continuous function φ.

Definition 1.2.15 *We call a relaxed control on the probability space $(\Omega, \mathcal{F}, P)$, is a random variable $\omega \to q(\omega, dt, du)$ with values in $\mathbb{V}$, such that is $\mathcal{F}_t$-progressively measurable and such that for each t, $1_{]0,t]}q$ is $\mathcal{F}_t$- measurable.*

Proposition 1.2.16 *Let q be a relaxed control with values in $\mathbb{V}$. Then $\forall t \in [0,T]$, there exists a probability measure q_t on U such that: $q(dt,du) = q_t(du)\, dt$.*

Often, one needs to know when a sequence of random variables converge in the distribution, or at least have a convergence subsequence in the distribution. In this context, we provide a good description of the sequences in $P(U)$ that have a convergent subsequence through the concept relatively compact sets of $P(U)$.

Proposition 1.2.17 *Let U is a compact set, then $P(U)$ is a compact space for the topology of weak convergence.*

Proof. U being compact, then by Prokhorov's theorem, the space $P(U)$ is compact for the topology of weak convergence. ■

Definition 1.2.18 *Let $(q^n)_{n\geq 0} \subset \mathbb{V}$, we say that q^n converges weakly to $q \in \mathbb{V}$, if*

$$\lim_{n\to\infty} \int_{[0,T]\times\mathbb{U}} \varphi(t,u)\, q^n(dt,du) = \int_{[0,T]\times\mathbb{U}} \varphi(t,u)\, q(dt,du),$$

for all continuous and measurable function φ.

Since all the elements of $\mathbb{V}$ have the same marginal on $[0,T]$, which is Lebesgue's measure, it is possible to considerably modify the weak convergence and to weaken the hypotheses on χ by obtaining another type of convergence, which is the stable convergence, as follows

Proposition 1.2.19 *We assume that $(q^n)_{n\geq 0}$ converges to q in the set $\mathbb{V}$. Then for all continuous and measurable function φ on $[0,T]\times U$ we have*

$$\lim_{n\to\infty} \int_{[0,T]\times\mathbb{U}} \varphi(t,u)\, q^n(dt,du) = \int_{[0,T]\times\mathbb{U}} \varphi(t,u)\, q(dt,du),$$

this type of convergence is called stable convergence.

1.2.1.2 The functions space $\mathcal{C}([0,T],\mathbb{R}^m)$

Let $\mathcal{C}([0,T],\mathbb{R}^m)$ denote the space of continuous $\mathbb{R}^m$-valued functions on the interval $[0,T]$, with the sup norm. If $X \in \mathcal{C}([0,T],\mathbb{R}^m)$, equipped with uniform convergence on compact sets, $\mathcal{C}([0,T],\mathbb{R}^m)$ is a Polish space. We assume

$$\begin{cases} \|X\|_\infty = \sup_{0\leq t\leq T} |X_t| \text{ and for} \\ 0 \leq \theta \leq 1,\ \Pi_X(\theta) \equiv \sup_{|t-s|\leq\theta} |X_t - X_s|, \end{cases}$$

a necessary and sufficient condition for X to be uniformly continuous is

$$\lim_{\theta \to 0} \Pi_X(\theta) \equiv 0.$$

Let $(X^n)_{n \geq 0}$, is a family of continuous stochastic processes such that Q^n is the law of the random variable X_t^n for $0 \leq t \leq T$, then the following result is a consequence of the Arzelà–Ascoli theorem

Theorem 1.2.20 *Let $X^n : (\Omega, \mathcal{F}, P) \to \mathcal{C}([0,T], \mathbb{R}^m)$, $(X^n)_{n \geq 0}$ is tight on $\mathcal{C}([0,T], \mathbb{R}^m)$ if and only if for every $T \geq 0$:*

$$\begin{cases} \textit{i)}\ \lim_{N\to\infty} \sup_{n\geq 1} P\left(|X_0^n| \geq N\right) = 0, \\ \textit{ii)}\ \lim_{\theta\to 0} \sup_{n\geq 1} P\left(\Pi_{X^n}(\theta) \geq \alpha\right) = 0,\ \forall \alpha > 0. \end{cases}$$

Condition ii), can be put in a more compact form: for all positive N,

$$\lim_{N\to\infty} \sup_{n\geq 1} P\left(\|X^n\| \geq N\right) = 0.$$

Theorem 1.2.21 *Let $(X^n)_{n \geq 0}$ be a sequence of m-dimensional continuous processes satisfying the following two conditions:*

i) there exist positive constants C and γ such that $E[|X_0^n|^\gamma] \leq C, n \geq 0,$

ii) there exist positive constants α, β, M such that

$$\mathbb{E}\left[|X_t^n - X_s^n|^\alpha\right] < M\left|t-s\right|^{\beta+1}, \forall t, s \in [0,T],$$

then (X^n) is tight as $\mathcal{C}([0,T], \mathbb{R}^m)$-valued random variables. As a consequence, there exists a subsequenc $(X^{n_k})_{k>0}$, on a probability space $\left(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}\right)$, and m-dimensional continuous processes $\left(\hat{X}^{n_k}\right)_{k>0}$ and $\hat{X}$, defined on it such that

1) the laws of $\hat{X}^{n_k}$ and X^{n_k} coincide,

2) $\hat{X}_t^{n_k}$ converges to $\hat{X}_t$ uniformly on every finite time interval, $P-a.s.$

A sequence of processes $(X^n)_{n\geq 0}$ with paths in $\mathcal{C}([0,\infty),\mathbb{R}^m)$ is tight if and only if its restriction to the interval $[0,T]$ is tight for each $T<\infty$.

1.2.1.3 The functions space $\mathbb{D}([0,T],\mathbb{R}^m)$

Let $\mathbb{D}([0,T],\mathbb{R}^m)$ denote the space of càdlàg (ie: continuous from the right with left hand limits), $\mathbb{R}^m$-valued functions on the interval $[0,T]$.

Now, let Λ denot the class of strictly increasing continuous mappings of $[0,T]$ onto itself. If $\lambda\in\Lambda$, then $\lambda(0)=0$ and $\lambda(T)=T$. For $X,Y\in\mathbb{D}([0,T],\mathbb{R}^m)$, we define $d(x,y)$ to be the infimum of those for which there exists in Λ a λ satisfying

$$\begin{cases} \sup_{t\in[0,T]}|\lambda(t)-t| = \sup_{t\in[0,T]}\left|t-\lambda^{-1}(t)\right| \\ \sup_{t\in[0,T]}\left|X_t-Y_{\lambda(t)}\right| = \sup_{t\in[0,T]}\left|X_{\lambda^{-1}(t)}-Y_t\right|, \end{cases}$$

then the definition of $d(x,y)$ becomes

$$d(x,y)=\inf_{\lambda\in\Lambda}\left\{\sup_{t\in[0,T]}|\lambda(t)-t|\vee\sup_{t\in[0,T]}\left|X_t-Y_{\lambda(t)}\right|\right\},$$

this metric defines the Skorokhod topology.

Theorem 1.2.22 *$d(x,y)$ is a metric on $\mathbb{D}([0,T],\mathbb{R}^m)$.*

Theorem 1.2.23 *The space $\mathbb{D}([0,T],\mathbb{R}^m)$ is separable under $d(x,y)$, but is not complete.*

The sequence X^n is converge to limit X in the space $\mathbb{D}([0,T],\mathbb{R}^m)$ endowed with the Skorokhod topology if and only if there exist sequence of functions λ^n in Λ such that $\lim_{n\to\infty}X^n_{\lambda^n(t)}=X_t$ uniformly in t. If X^n converging uniformly to X, then there is convergence in the Skorokhod topology $(\lambda^n(t)=t)$. The Skorokhod convergence does imply that $\lim_{n\to\infty}X^n_t=X_t$ hold for continuity points t of X_t and hence for all but countably many t. A consequently, if X_t uniformly continuous on all $t \in [0, T]$, then the Skorokhod convergence implies uniform convergence.

Corollary 1.2.24 *The Skorokhod topology coincides with uniform topology on* $\mathcal{C}([0, T], \mathbb{R}^m)$.

For $t \in [0, T]$, we define the projection $\pi_t : X \in \mathbb{D}([0, T], \mathbb{R}) \to X_t$, for each $\lambda \in \Lambda$, $\pi_{\lambda(0)} = \pi_0$ and $\pi_{\lambda(T)} = \pi_T$ are continuous. If X^n convergence to X in the Skorokhod topology and X continuous at t, then by

$$|X_t^n - X_t| \leq \left|X_t^n - X_{\lambda^n(t)}\right| + \left|X_{\lambda^n(t)} - X_t\right|.$$

X^n convergence to X. We assume that, X is discontinuous at t, if $\lambda^n(t) = t + \frac{1}{n}$, element of Λ for $t \in [0, T]$ and if $X^n_{\lambda^n(t)} \equiv X^n_t$, then X^n_t converges to X_t in the Skorokhod topology but X^n_t does not converge to X_t. Therefore, for each $0 \leq t \leq T$, π_t is continuous at X if and only if X is continuous at t.

Corollary 1.2.25 *If* $d(X^n, X) \to 0 \; as \; n \to \infty$, *then*

$$\sup_{t \in [0,T]} |X_t^n - X_t| \to 0 \; as \; n \to \infty,$$

for each $t \notin Disc(X) = \{t \in [0, T]/X_t \neq X_{t-}\}$.

Remark 1.2.26 *If* X^n *convergence to* X *in the Skorokhod topology, then* X^n *convergence to* X *for continuty points* t *of* X *and hence for points* t *outside aset of Lebesgue measure* 0.

We can develop a simple criterion for Skorokhod convergence for monotone functions

Theorem 1.2.27 *Let* X^n *is monotone for* $\forall n \in \mathbb{N}$, *then* $X^n_t \to X_t$ *for all* t *in a dense subset of* $[0, T]$ *including* 0 *and* T *implies* X^n *convergence to* X *in the Skorokhod topology, for* $X \in \mathbb{D}([0, T], \mathbb{R}^m)$.

Under the Skorokhod topology, addition of functions $\mathbb{D}([0,T],\mathbb{R}^m)$ is not topological group. Since addition is not continuous everywhere, for this we show that it is continuous almost everywhere, andt it is measurable. It is thus important to know more about the Borel σ-fields associated with the Skorohod topologies. The σ-field of Borel subsets for the Skorokhod topology coincides with the Kolmogorov σ-field, which is generated by the projection π_t.

We can define in $\mathbb{D}([0,T],\mathbb{R}^m)$ another metric give the Skorohod topology, and under this metric $\mathbb{D}([0,T],\mathbb{R}^m)$ is complete. If λ is non decreasing function on $[0,T]$ satisfying $\lambda(0)=0$ and $\lambda(T)=T$, we put

$$\|\lambda\| = \sup_{s<t}\left|\log\left(\frac{\lambda(t)-\lambda(s)}{t-s}\right)\right|$$

Let $\hat{d}(x,y)$ to be the infimum of those for which there Λ contains some λ satisfying $(\|\lambda\|<\infty)$:

$$\hat{d}(x,y) = \inf_{\lambda\in\Lambda}\left\{\|\lambda\| \vee \sup_{t\in[0,T]}\left|X_t - Y_{\lambda(t)}\right|\right\}.$$

Theorem 1.2.28 *i) The metric $\hat{d}(x,y)$ on $\mathbb{D}([0,T],\mathbb{R}^m)$ is topologically equivalent to $d(x,y)$.*

ii) The space $\mathbb{D}([0,T],\mathbb{R}^m)$ is separable and complete under $\hat{d}(x,y)$.

Compactness characterizations on $\mathbb{D}([0,T],\mathbb{R}^m)$ translate into tightness characterizations for sets of probability measures on $\mathbb{D}([0,T],\mathbb{R}^m)$. We recall that a set B of probability measures is said to be tight if for all $\varepsilon>0$ there exists a compact subset K_ε such that

$$P(K_\varepsilon)\geq 1-\varepsilon,\ \forall P\in B.$$

Given a subdivision $0=t_0<t_1<...<t_n=T$, we define

$$\hat{\Pi}_X(\theta) = \inf_{t_k}\max_{0\leq k\leq n}\Pi_X(\theta).$$

Let $(X^n)_{n\geq 0} \subset \mathbb{D}([0,T],\mathbb{R}^m)$, to be a family of càdlàg stochastic processes such that Q^n is the law of the random variable X_t^n for $0 \leq t \leq T$, then we present the following result.

Theorem 1.2.29 *Let $X^n : (\Omega, \mathcal{F}, P) \to \mathcal{C}([0,T], \mathbb{R}^m)$, $(X^n)_{n\geq 0}$ is tight on $\mathcal{C}([0,T], \mathbb{R}^m)$ if and only if for every $T \geq 0$:*

$$\begin{cases} \textit{i)}\ \lim_{N\to\infty} \sup_{n\geq 1} P\left(\|X^n\|_\infty \geq N\right) = 0, \\ \textit{ii)}\ \lim_{\theta\to 0} \sup_{n\geq 1} P\left(\hat{\Pi}_{X^n}(\theta) \geq \alpha\right) = 0,\ \forall \alpha > 0. \end{cases}$$

1.2.2 Brownian motion

The Brownian motion, is a name given by the botanist Robert Brown in 1827 to describe the irregular motion of pollen particles in a fluid. The context of applications of Brownian motion goes far beyond the study of microscopical particles, and is now largely used in finance for modelling stock prices, historically since Bachelier in 1900.

1.2.2.1 Definition and Main Properties

Definition 1.2.30 *Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ be a filtered probability space. An $(\mathcal{F}_t)_{0\leq t\leq T}$-adapted $\mathbb{R}^m$-valued process W is called an m-dimensional $(\mathcal{F}_t)_{0\leq t\leq T}$-Brownian motion if for all $0 < s < t$, the increment $W_t - W_s$ is independent of $\mathcal{F}_s$ and is normally distributed with mean 0 and variance-covariance matrix $(t-s)\, I_m$. For any $0 < s < t$*

$$\mathbb{E}\left[W_t - W_s \mid \mathcal{F}_s\right] = 0.$$

In addition, if $W_0 = 0$, then W_t is called an m-dimensional standard $(\mathcal{F}_t)_{0\leq t\leq T}$-Brownian motion.

In the definition of a standard Brownian motion, the independence of the increments is with respect to the natural filtration of W_t. The natural filtration of W_t is sometimes called Brownian filtration. It is often to work with a larger fltration than the natural filtration. In this case, the definition remains correct for this fltration. We now give some elementary properties of Brownian motion.

Proposition 1.2.31 *Let $(W_t)_{0\le t\le T}$ be a Brownian motion with respect to $(\mathcal{F}_t)_{0\le t\le T}$.*

i) Scaling invariance: for all $\lambda > 0$, the process $\left(\lambda W_{\frac{t}{\lambda^2}}\right)_{0\le t\le T}$ is also a Brownian motion.

ii) Invariance by translation: for all $s > 0$, the process $(W_{t+s} - W_s)_{0\le t\le T}$ is a standard Brownian motion independent of $\mathcal{F}_s$.

Remark 1.2.32 *The scaling invariance property (with $\lambda = -1$) implies that standard Brownian motion is symmetric about 0. In other words, if $(W_t)_{0\le t\le T}$ is a standard Brownian motion then W_t has the same distribution as $-W_t$.*

Proposition 1.2.33 *Let $(W_t)_{0\le t\le T}$ be a Brownian motion with respect to $(\mathcal{F}_t)_{0\le t\le T}$. Then*

i) for all $t > s$, $W_t - W_s$ is independent of $\mathcal{F}_s = \sigma\left(W_r|0 \le r \le s\right)$,

ii) $\mathbb{E}\left[W_t W_s\right] = t \wedge s$.

Proof. i) Since W has independent increments, then $W_t - W_s$ is independent of $W_s - W_0$. Also $W_t - W_s$ is independent of $W_s - W_r$ for all $0 \le r \le s$. On the other hand, note that $W_r = -\left(W_s - W_r\right) + W_s$. Thus W_r is $\sigma\left(W_s\right) \vee \sigma\left(W_s - W_r\right)$-measurable for all $0 \le r \le s$. Consequently from above, we deduce that $W_t - W_s$ is independent of $\sigma\left(W_r\right)$ for all $0 \le r \le s$, which means that $W_t - W_s$ is independent of $\mathcal{F}_s = \sigma\left(W_r|0 \le r \le s\right)$ for all $0 \le s \le t$.

ii) If $t \geq s$ by using the independence of $W_t - W_s$ and W_s, we get

$$\mathbb{E}\left[\left(W_t - W_s\right)W_s + W_s^2\right] = \mathbb{E}\left[\left(W_t - W_s\right)W_s\right] + \mathbb{E}\left[W_s^2\right]$$
$$= s.$$

On the other hand, if $t \leq s$ we deduce that $\mathbb{E}\left[W_t W_s\right] = t$. ■

Definition 1.2.34 *we call infinitesimal variation of order (p) of an associated process $(X_t)_{0\leq t\leq T}$ of a subdivision $\Delta_n = (t_1^n < ..... < t_n^n)$ of $[0,T]$*

$$V_T^p\left(X_t\right) = \sum_{k=1}^{n}\left|X_{t_k} - X_{t_{k-1}}\right|^p,$$

if $V_T^p\left(X_t\right)$ admits a limit when $\|\Delta_n\| \to 0$ as $n \to \infty$ and the limit does not depend on a subdivided proportion, we call of order variation (p) on $[0,T]$.

If $p = 1$ the limit is called totol variation of X.

If $p = 2$ the limit is called quadratic variation and we denote by $\langle X, X\rangle_T$.

Proposition 1.2.35 *i) The quadratic variation of a Brownian motion on $[t,T]$ converges as a quadratic mean to $T - t$, $\forall t, T \in \mathbb{R}$ and if $(\Delta_n)_{n\geq 0}$ is a sequence of subdivisions of $[t,T]$, where $\|\Delta_n\| \to 0$ as $n \to \infty$.*

ii) If the subdivision Δ_n on $[0,T]$ verify $\sum_{n\geq 0}\|\Delta_n\| < \infty$ then

$$V_T^2\left(X_t\right) \to T.$$

Proof. Using the fact that for a random variable W_t Brownian motion is normally distributed with mean 0 and variance t, with $E\left(W_t^4\right) = 3t^2$ we have

$$\mathbb{E}\left[\left(V_{t,T}^2\left(W_t\right) - (T-t)\right)^2\right] = \mathbb{E}\left[\left(\sum_{k=1}^{n}\left(\left|W_{t_k} - W_{t_{k-1}}\right|^2 - \left(t_k - t_{k-1}\right)\right)\right)^2\right],$$

where $\Delta_n = (t_1^n < ..... < t_n^n)$ is subdivision of $[t, T]$, with $\sum_{k=1}^n (t_k - t_{k-1}) = T - t$.

Since the increment $W_{t_k} - W_{t_{k-1}}$ is independent so we get

$$\mathbb{E}\left[\left(V_{t,T}^2(W_t) - (T-t)\right)^2\right] = \sum_{k=1}^n \mathbb{E}\left[\left(\left|W_{t_k} - W_{t_{k-1}}\right|^2 - (t_k - t_{k-1})\right)^2\right].$$

On the other hand,

$$\mathbb{E}\left[\left(\left|W_{t_k} - W_{t_{k-1}}\right|^2 - (t_k - t_{k-1})\right)^2\right] = 3(t_k - t_{k-1})^2 + (t_k - t_{k-1})^2$$
$$-2(t_k - t_{k-1})\,\mathbb{E}\left(\left|W_{t_k} - W_{t_{k-1}}\right|^2\right) = 2(t_k - t_{k-1})^2.$$

Then

$$\mathbb{E}\left[\left(V_{t,T}^2(W_t) - (T-t)\right)^2\right] \leq 2\sup|t_k - t_{k-1}| \textstyle\sum_{k=1}^n (t_k - t_{k-1})$$
$$\leq 2(T-t)\sup|t_k - t_{k-1}| \to 0,\ as\ n \to \infty.$$

■

Proposition 1.2.36 *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function, then*

$$\mathbb{E}[f(W_t)] = \frac{1}{(2\pi t)^{\frac{1}{2}}} \int_{\mathbb{R}} |y| \exp\left(\frac{-y^2}{2t}\right) dy.$$

Example 1.2.37 *Let $f : R \to R$ be $f(x) = |x|$. By applying Proposition previous we find*

$$\mathbb{E}[|W_t|] = \frac{1}{(2\pi t)^{\frac{1}{2}}} \int_{\mathbb{R}} |y| \exp\left(\frac{-y^2}{2t}\right) dy$$
$$= \frac{2}{(2\pi t)^{\frac{1}{2}}} \int_0^\infty y \exp\left(\frac{-y^2}{2t}\right) dy,$$

using the variable change $u = \frac{y^2}{2t}$ we find

$$\mathbb{E}[|W_t|] = \frac{(2t)^{\frac{1}{2}}}{(\pi)^{\frac{1}{2}}} \int_0^\infty \exp(-u)\, du$$
$$= \left(\frac{2t}{\pi}\right)^{\frac{1}{2}}.$$

Proposition 1.2.38 *Let $(W_t)_{0\leq t\leq T}$ be a Brownian motion. Then $\left(W_t^2-t\right)_{0\leq t\leq T}$ is $\left(\mathcal{F}_t^W\right)_{0\leq t\leq T}$-martingales.*

Proof. Since $\mathcal{F}_t^W=\sigma\left(W_s,0\leq s\leq t\right)$ it follows that W_t-W_s is independent of $\mathcal{F}_s^W$ and hence

$$
\begin{aligned}
\mathbb{E}\left[W_t^2-t\mid\mathcal{F}_s^W\right] &= \mathbb{E}\left[W_t^2-W_s^2\mid\mathcal{F}_s^W\right]-t+W_s^2\\
&= \mathbb{E}\left[\left(W_t-W_s\right)^2-2W_s\left(W_t-W_s\right)\mid\mathcal{F}_s^W\right]-t+W_s^2\\
&= \mathbb{E}\left[\left(W_t-W_s\right)^2\right]-2W_s\mathbb{E}\left[W_t-W_s\right]-t+W_s^2\\
&= (t-s)-t+W_s^2=W_s^2-s.
\end{aligned}
$$

■

Proposition 1.2.39 *For each $\alpha>0$, there exists a constant M such that*

$$
\mathbb{E}\left[\sup_{r\in[s,t]}\left|W_t-W_s\right|^{\alpha}\right]<M\left|t-s\right|^{\frac{\alpha}{2}},\forall 0\leq s\leq t.
$$

We also recall that, the augmentation of $\left(\mathcal{F}_t^W\right)_{0\leq t\leq T}$ by all the P-null sets ($\hat{\mathcal{F}}_t=\mathcal{F}_t^W\vee\mathcal{N}$, where $\mathcal{N}$ is the set of negligible events), and W_t is still a Brownian motion on the (augmented) filtered probability space $\left(\Omega,\mathcal{F},\left(\hat{\mathcal{F}}_t\right)_{0\leq t\leq T},P\right)$. In what follows, by saying that $\left(\mathcal{F}_t\right)_{0\leq t\leq T}$ is the natural filtration generated by the Brownian motion W, we mean that $\left(\mathcal{F}_t\right)_{0\leq t\leq T}$ is generated by W and augmented by all the P-null sets in $\mathcal{F}$.

1.3 Itô's stochastic integral

In this section we give the definition of the Itô integral as well as some basic properties of such an integral. We shall describe the basic idea of defining the Itô integral. Those results will be useful later in this thesis.

1.3.1 Definition of Itô's Stochastic Integral

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ be a fixed filtered probability space satisfying the usual condition. Let $T > 0$ and recall that $L^2(\Omega, \mathcal{F}, P, \mathbb{R}^m)$ is the set of all $\mathcal{F}_t-$measurable processes Z_t defined from $[0,T]\times\Omega$ into $\mathbb{R}^m$, such that

$$\mathbb{E}\left[\int_0^T \|Z_t\|^2 \, dt\right]^{\frac{1}{2}} < \infty,$$

Now we define the Itô integral for a given Brownian motion W_t defined on the filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$. Let $\Sigma_{\mathcal{F}}([0,T]) \subset L^2(\Omega, \mathcal{F}, P, \mathbb{R}^m)$ the subset of all real simple processes Z_t of the following form

$$Z_t = \sum_{k=0}^{n-1} Z_k 1_{[t_k, t_{k+1}]},\ 0 \leq t \leq T,$$

with $n \in \mathbb{N}$, $t_0 < t_1 < ..... < t_n$ and for $0 \leq k \leq n-1$, $Z_k : \Omega \to \mathbb{R}^m$ is an $\mathcal{F}_{t_k}$-measurable bounded random variable.

Remark 1.3.1 *The set $\Sigma_{\mathcal{F}}([0,T])$ is dense in $L^2(\Omega, \mathcal{F}, P, \mathbb{R}^m)$.*

Now, for any simple process $Z_t \in \Sigma_{\mathcal{F}}([0,T])$ we define the following linear operator

$$I_t(Z) = \sum_{k=0}^{n-1} Z_k \left(W_{t_{k+1}} - W_{t_k}\right),\ 0 \leq t \leq T.$$

Lemma 1.3.2 *Let $I_t(Z)$ be a linear random variable with mean $\mathbb{E}[I_t(Z)] = 0$, and variance $\mathbb{E}\left[I_t(Z)^2\right] = \sum_{k=0}^{n-1}\mathbb{E}\left[Z_k^2\right](t_{k+1} - t_k)$.*

Proof. For each $0 \leq k \leq n-1$, we have

$$\begin{aligned}\mathbb{E}\left[Z_k\left(W_{t_{k+1}} - W_{t_k}\right)\right] &= \mathbb{E}\left[Z_k \mathbb{E}\left[\left(W_{t_{k+1}} - W_{t_k}\right) \mid \mathcal{F}_{t_k}\right]\right] \\ &= \mathbb{E}\left(Z_k \mathbb{E}\left[W_{t_{k+1}} - W_{t_k}\right]\right) = 0.\end{aligned}$$

Hence $\mathbb{E}\left(I_t\left(Z\right)\right)=0$. On the other hand, we have

$$I_t\left(Z\right)^2=\sum_{r,k=0}^{n-1}Z_rZ_k\left(W_{t_{r+1}}-W_{t_r}\right)\left(W_{t_{k+1}}-W_{t_k}\right).$$

If $r\neq k$, where $r<k$

$$\begin{aligned}\mathbb{E}\left[Z_rZ_k\left(W_{t_{r+1}}-W_{t_r}\right)\left(W_{t_{k+1}}-W_{t_k}\right)\right]&=\mathbb{E}\left[Z_rZ_k\left(W_{t_{r+1}}-W_{t_r}\right)\mathbb{E}\left[W_{t_{k+1}}-W_{t_k}\mid\mathcal{F}_{t_k}\right]\right]\\&=\mathbb{E}\left(Z_rZ_k\left(W_{t_{r+1}}-W_{t_r}\right)\mathbb{E}\left[W_{t_{k+1}}-W_{t_k}\right]\right)=0.\end{aligned}$$

Moreover, for $r=k$ we have

$$\begin{aligned}\mathbb{E}\left[Z_k^2\left(W_{t_{k+1}}-W_{t_k}\right)^2\right]&=\mathbb{E}\left[Z_k^2\mathbb{E}\left[\left(W_{t_{k+1}}-W_{t_k}\right)^2\mid\mathcal{F}_{t_k}\right]\right]\\&=\mathbb{E}\left[Z_k^2\mathbb{E}\left(W_{t_{k+1}}-W_{t_k}\right)^2\right]\\&=\mathbb{E}\left[Z_k^2\right]\left(t_{k+1}-t_k\right).\end{aligned}$$

■

For any $Z_t\in L^2\left(\Omega,\mathcal{F},P,\mathbb{R}^m\right)$, there is $\left(Z_k\right)_{k\geq0}\subset\Sigma_{\mathcal{F}}\left([0,T]\right)$ such that

$$\lim_{k\to\infty}\mathbb{E}\left[\int_0^T\left\|Z_k-Z_t\right\|^2dt\right]^{\frac{1}{2}}=0.$$

$\left(I_t\left(Z_k\right)\right)_{k\geq0}$ is Cauchy in $L^2\left(\Omega,\mathcal{F},P,\mathbb{R}^m\right)$ ie: has a unique limit in $L^2\left(\Omega,\mathcal{F},P,\mathbb{R}^m\right)$, denoted by

$I_t\left(Z\right)$ is called the Itô integral, denoted by

$$I_t\left(Z\right)=\int_0^TZ_tdW_t,\ 0\leq t\leq T.$$

Proposition 1.3.3 *For any $Z,Y\in L^2\left(\Omega,\mathcal{F},P,\mathbb{R}^m\right)$ and $0\leq s\leq t$. Then*

a) $\mathbb{E}\left[\int_s^tZ_rdW_r\mid\mathcal{F}_s^W\right]=0.$

b) $\mathbb{E}\left[\left|\int_s^tZ_rdW_r\right|^2\mid\mathcal{F}_s^W\right]=\mathbb{E}\left[\int_s^t\left|Z_r\right|^2dr\mid\mathcal{F}_s^W\right]=0.$

c) $\mathbb{E}\left[\int_s^tZ_rdW_r\int_s^tY_rdW_r\mid\mathcal{F}_s^W\right]=\mathbb{E}\left[\int_s^tZ_rY_rdW_r\mid\mathcal{F}_s^W\right]=0.$

Proposition 1.3.4 (Product Property) *For any $Z_t, Y_t \in L^2(\Omega, \mathcal{F}, P, \mathbb{R}^m)$, the following equality holds*

$$\mathbb{E}\left[\int_0^T Z_t dW_t \int_0^T Y_t dW_t\right] = \int_0^T \mathbb{E}\left[Y_t Z_t\right] dt.$$

Theorem 1.3.5 (Martingale Property) *Let $Z_t \in L^2(\Omega, \mathcal{F}, P, \mathbb{R}^m)$. Then the stochastic process*

$$X_t = \int_0^t Z_s dW_s,\ 0 \leq t \leq T,$$

is a martingale with respect to the filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$.

Example 1.3.6 *Using the definition of Itô integral, we prove that*

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds.$$

Given a subdivision $0 = t_0 < t_1 < ... < t_n = t$, we put

$$tW_t = \sum_{k=0}^{n-1} \left(t_{k+1} W_{t_{k+1}} - t_k W_{t_k}\right).$$

We note that

$$\left(t_{k+1} W_{t_{k+1}} - t_k W_{t_k}\right) = W_{t_{k+1}}\left(t_{k+1} - t_k\right) + t_k\left(W_{t_{k+1}} - W_{t_k}\right),$$

so

$$tW_t = \sum_{k=0}^{n-1} W_{t_{k+1}}\left(t_{k+1} - t_k\right) + \sum_{k=0}^{n-1} t_k\left(W_{t_{k+1}} - W_{t_k}\right),$$

and, when we let $\Delta t_k \to 0$, we get

$$tW_t = \int_0^t s dW_s + \int_0^t W_s ds,$$

such that $\Delta t_k = (t_{k+1} - t_k)$.

Hence

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds.$$

Now, we present the backward Itô integral, and the forward Itô integral with respect to Brownian motion, those two types of integrals are particular cases of the Itô-Skorohod integral. This stochastic integral, introduced for the first time by A. Skorohod in 1975, may be regarded as an extension of the Itô integral to integrands that are not necessarily $\mathcal{F}_t$-adapted. Let us now recall the definition of backward stochastic integrals, we denote by $(\pi^n)_{n\geq 0}$ any sequence of subdivisions: $\pi^n = \{t = t_0^n < t_1^n < ... < t_n^n = T\}$. Such that $|\pi^n| = \sup_{0\leq k\leq n-1}\left(W_{t_{k+1}} - W_{t_k}\right) \to 0$ as $n \to \infty$. Then the backward Itô integral can be defined as

$$\int_t^T Z_s d\overleftarrow{W_s} = \lim_{n\to 0}\sum_{k=0}^{n-1} Z_{t_{k+1}}\left(W_{t_{k+1}} - W_{t_k}\right).$$

In fact, the backward Itô integral of Z_s with respect to W_s may be understood as the forward integral of $\tilde{Z}_s = Z_{T-s}$ with respect to $\tilde{W}_s = W_{T-s} - W_T$ such that

$$\begin{aligned}\int_0^{T-t} \tilde{Z}_s d\tilde{W}_s &= \lim_{n\to 0}\sum_{k=0}^{n-1} \tilde{Z}_{t_{k+1}}\left(\tilde{W}_{t_{k+1}} - \tilde{W}_{t_k}\right)\\ &= \lim_{n\to 0}\sum_{k=0}^{n-1} Z_{T-t_k}\left(W_{T-t_{k+1}} - W_{T-t_k}\right),\end{aligned}$$

note that $r_{k+1} = T - t_k < r_k = T - t_{k+1}$ is a subdivision of $[t, T]$ then we get

$$\begin{aligned}\int_0^{T-t} \tilde{Z}_s d\tilde{W}_s &= -\lim_{n\to 0}\sum_{k=0}^{n-1} Z_{r_{k+1}}\left(W_{r_{k+1}} - W_{r_k}\right)\\ &= -\int_t^T Z_s d\overleftarrow{W_s}.\end{aligned}$$

Remark 1.3.7 *If Z_t is $\mathcal{F}_t$-adapted then, the Skorohod integral is coincides with the Itô integral.*

1.3.2 Itô's Formula

In this Subsubsection we present a stochastic version of the chain rule, or change-of-variable formula, called Itô's formula lemma which applies to the Itô integral. It is one of the most powerful and frequently used theorems in stochastic calculus.

In this context, In finance, we often use Itô processes for modeling the dynamics of asset prices. An Itô process is defined to be an adapted stochastic process that can be expressed as the sum of an integral with respect to Brownian motion and an integral with respect to time.

Definition 1.3.8 (Itô process) *Let W_t be a m-dimensional Brownian motion on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$. We define an Itô process as a process $(X_t)_{0\leq t\leq T}$ valued in $\mathbb{R}^n$ such that $a.s.$*

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s,\ 0 \leq t \leq T,$$

where X_0 is $\mathcal{F}_0$-measurable, b and σ are progressively measurable processes valued respectively in $\mathbb{R}^n$ and $\mathbb{R}^{n\times d}$ such that

$$\int_0^t |b_s|\, ds + \int_0^t |\sigma_s|^2\, ds < \infty,\ a.s., \forall t \in [0, T].$$

Theorem 1.3.9 (Itô's formula) *Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ be a filtered probability space satisfying the usual condition, W_t an m-dimensional $\mathcal{F}_t$-Brownian motion, and let X_t be a Itô process. Let $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ be a $\mathcal{C}^1(\mathbb{R})$ function with respect to t, and class $\mathcal{C}^2(\mathbb{R})$ with respect to X. Then*

$$\begin{aligned} f(t, X_t) = f(0, X_0) + \textstyle\int_0^t \frac{\partial f}{\partial t}(s, X_s)\, ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, dX_s \\ + \tfrac{1}{2}\textstyle\int_0^t \frac{\partial^2 f}{\partial x \partial x}(s, X_s) \times \sigma_s^2 ds,\ \forall t \in [0, T]. \end{aligned}$$

Example 1.3.10 *Our goal is calculation*

$$\int_0^t s dW_s = ?$$

Using Itô's formula, we put $f(t,x) = tx, \forall (t,x) \in [0,T] \times \mathbb{R}$.

Then we get

$$f(t, W_t) = tW_t = \int_0^t W_s ds + \int_0^t s dW_s + 0.$$

And therefore

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds.$$

Proposition 1.3.11 (Integration by parts formula) *Let X and Y be Itô processes in $\mathbb{R}$. Then*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t, \ t \geq 0.$$

1.3.3 Martingale representation theorem

In this Subsection, we state some martingale representation theorems by means of stochastic integrals. In probability theory, the martingale representation theorem states that a random variable that is measurable with respect to the filtration generated by a Brownian motion can be written in terms of an Itô integral with respect to this Brownian motion. The theorem only asserts the existence of the representation and does not help to find it explicitly. Similar theorems also exist for martingales on filtrations expandeds.

The first result is the Brownian martingale representation theorem.

Theorem 1.3.12 (Representation of Brownian martingales) *Assume that $(\mathcal{F}_t)_{0 \leq t \leq T}$ is the natural (augmented) filtration of a standard m-dimensional Brownian motion W. Let $(M_t)_{0 \leq t \leq T}$ be a càdlàg local martingale. Then there exists $\mathcal{F}_t$-progressively measurable process Z_t with values in $\mathbb{R}^{m \times d}$ such that*

$$\mathbb{E}\left[\int_0^T \|Z_t\|^2 dt\right] < \infty,$$

$$M_t = M_0 + \int_0^t Z_s dW_s,\ 0 \le t \le T.$$

This result shows in particular that any martingale with respect to a Brownian filtration is continuous. The second result is the projection theorem on the space of stochastic integrals with respect to a continuous local martingale.

Theorem 1.3.13 (Galtchouk-Kunita-Watanabe decomposition) *Let N_t be a square integrable $\mathcal{F}_t$-martingale. Then for any $M_t \in L^2(\Omega, \mathcal{F}, P, \mathbb{R}^m)$ we have*

$$M_t = M_0 + \int_0^t Z_s dN_s + R_t,\ 0 \le t \le T,$$

where Z_t is predictable, with

$$\mathbb{E}\left[\int_0^T \|Z_s\|^2 d\langle N, N\rangle_s\right] < \infty,$$

and R_t is an L^2-martingale with $R_0 = 0$. Further, the martingales R_t and N_t are orthogonal in the sense that$\langle R_t, N_t\rangle = 0$, $0 \le t \le T$.

Example 1.3.14 *Let W be a Brownian motion in $\mathbb{R}^n$, with $\theta(t,\omega) \in L^2(\Omega \times [0,T], \mathbb{R}^n)$, where $T < \infty$. We define*

$$Z_t = \exp\left(\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right), 0 \le t \le T.$$

By Itô's formula we get

$$dZ_t = Z_t \theta_t dW_t.$$

Consequently we get

$$Z_t = 1 + \int_0^t Z_s \theta_s dW_s.$$

1.4 Stochastic Differential Equations

In this sections we are going to study stochastic differential equations (SDEs, for short), which can be regarded as a generalization of ordinary differential equations (ODEs, for short). A stochastic differential equation SDE is a differential equation in which Including of the Itô integral will be involved, resulting in a solution which is also a stochastic process. SDEs are used to model various phenomena such as unstable stock prices or physical systems subject to thermal fluctuations. Typically, SDEs contain a variable which represents random white noise calculated as the derivative of Brownian motion or the Wiener process.

We fix a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ satisfying the usual conditions and a m-dimensional Brownian motion $(W_t)_{0\leq t\leq T}$ with respect to $(\mathcal{F}_t)_{0\leq t\leq T}$. We are given functions $(b(\omega, t, x))_{0\leq t\leq T}$ and $(\sigma(\omega, t, x))_{0\leq t\leq T}$ defined on $\Omega \times [0, T] \times \mathbb{R}^n$, and valued respectively in $\mathbb{R}^n$ and $\mathbb{R}^{n\times d}$. We assume that for all ω, the functions $b(\omega, ., .)$ and $\sigma(\omega, ., .)$ are Borelian on $[0, T] \times \mathbb{R}^n$ and for all $x \in \mathbb{R}^n$, the processes $b(., ., x)$ and $\sigma(., ., x)$, are progressively measurable. We consider the following (SDE) with valued in $\mathbb{R}^n$

$$(E_1) \begin{cases} dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t, \\ X_0 = \zeta. \end{cases}$$

In the above equation, we need to explain what it means that a stochastic process X_t is a solution of the (SDE). There are different notions of solutions to (SDE) depending on different roles that the underlying filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ and the Brownian motion W_t are playing. Let us introduce them in the following Subsubsections.

1.4.1 Strong solutions

Definition 1.4.1 *Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, P\right)$ be a filtered probability space, and W_t be a given m-dimensional standard $(\mathcal{F}_t)_{0\le t\le T}$-Brownian motion, and X_0, $\mathcal{F}_0$-measurable. An $(\mathcal{F}_t)_{0\le t\le T}$-adapted continuous process X_t, $0 \le t \le T$, is called a strong solution of (E_1) if*

$$\begin{cases} \int_0^t |b(s, X_s)|\, ds + \int_0^t |\sigma(s, X_s)|^2\, ds < \infty,\ \forall t \in [0, T],\ P - a.s, & (C_1) \\ X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s,\ t \in [0, T],\ P - a.s. & (C_2) \end{cases}$$

For any two strong solutions X and Y of (E_1) defined on $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, P\right)$ along with given standard $(W_t)_{0\le t\le T}$-Brownian motion, we have

$$P(X_t = Y_t, \forall t \in [0, T]) = 1,$$

then we say that strong uniqueness holds.

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, P\right)$ be a filtered probability space satisfying the usual condition, W_t an m-dimensional standard $(\mathcal{F}_t)_{0\le t\le T}$-Brownian motion, and an $\mathcal{F}_0$-measurable random variable. We make the following assumption for the coefficients of (E_1).

Existence and uniqueness of a strong solution to the (SDE) (E_1) is ensured by the following Lipschitz and linear growth conditions:

(H) There exists a constant K such that for all $\omega \in \Omega$, $t \in [0, T]$, $x, y \in \mathbb{R}^n$

$$|b(\omega, t, x) - b(\omega, t, y)| + |\sigma(\omega, t, x) - \sigma(\omega, t, y)| \le K(|x - y|),$$

$$|b(\omega, t, x)| + |\sigma(\omega, t, x)| \le K(1 + |x|).$$

Theorem 1.4.2 *Let (H) hold. Then, for any $t \in [0,T]$, equation (E_1) admits a unique strong solution X such that for any $T > 0$. Moreover, for any ζ $\mathcal{F}_0$-measurable random variable valued in $\mathbb{R}^n$, such that*

$$\mathbb{E}\left[|\zeta|^p\right] < \infty, \text{ for some } p > 1,$$

there is uniqueness of a strong solution X starting from ζ such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t|^2\right] < \infty.$$

1.4.2 Weak solutions

Definition 1.4.3 *A 6-tuple $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P, W, X\right)$ is called a weak solution of (E_1) if*

i) $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ is a filtered probability space satisfying the usual condition,

ii) W is an m-dimensional standard $(\mathcal{F}_t)_{0\leq t\leq T}$-Brownian motion and X is $(\mathcal{F}_t)_{0\leq t\leq T}$-adapted and continuous,

iii) X_0 and ζ have the same distribution,

iv) (C_1), and (C_2) hold.

The main difference between weak and strong solutions is indeed that for strong solutions we are given a Brownian motion W on a given probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ whereas for weak solutions $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ and W are parts of the solution. In other words, we are free to choose to choose the Brownian motion and the probability space.

Definition 1.4.4 *If for any two weak solutions $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P, W, X\right)$ and $\left(\hat{\Omega}, \hat{\mathcal{F}}, \left(\hat{\mathcal{F}}_t\right)_{0\leq t\leq T}, \hat{P}, \hat{W}, \hat{X}\right)$ of (E_1) with*

$$P\left(X_0 \in B\right) = \hat{P}\left(\hat{X}_0 \in B\right), \ \forall B \in \mathcal{B}\left(\mathbb{R}^n\right),$$

we have

$$P(X \in B) = \hat{P}\left(\hat{X} \in B\right),\ \forall B \in \mathcal{B}(\mathbb{R}^n),$$

then we say that the weak solution of (E_1) *is unique (in the sense of probability law).*

Definition 1.4.5 *If*

$$P\left(X_t = \hat{X}_t,\ 0 \le t \le T\right) = 1,$$

for any two weak solutions $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, P, W, X\right)$ *and* $\left(\Omega, \mathcal{F}, \left(\hat{\mathcal{F}}_t\right)_{0\le t\le T}, P, W, \hat{X}\right)$ *of* (E) *with*

$$P\left(X_0 = \hat{X}_0\right) = 1,$$

then we say that the weak solutions have pathwise uniqueness.

Existence of weak solutions does not imply that of strong solutions, and weak uniqueness does not imply pathwise uniqueness nor strong uniqueness.

Example 1.4.6 *Let* $(W_t)_{0\le t\le T}$ *be a Brownian motion. Then* $X_t = W_t$ *and* $\hat{X}_t = -W_t$ *are weak solutions to the (SDE)*

$$dX_t = dB_t,\ X_0 = 0.$$

Clearly

$$P\left(X_t = \hat{X}_t\right) = P(W_t = 0) = 0.$$

The notion of uniqueness for weak solutions is weak uniqueness, i.e: uniqueness in distribution.

Relations between the strong and weak solutions are presented in the following two theorems.

Theorem 1.4.7 *Let the processes* $(b(\omega, t, x))_{0 \leq t \leq T}$ *and* $(\sigma(\omega, t, x))_{0 \leq t \leq T}$ *are progressively measurable, and its values respectively in* $\mathbb{R}^n$ *and* $\mathbb{R}^{n \times d}$*. Then* (E_1) *admits a unique strong solution if and only if for any probability measure* μ *on* $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$, (E_1) *admits a weak solution with the initial distribution and pathwise uniqueness holds for* (E_1)*.*

By and large, Theorem previous tells that strong existence and uniqueness is equivalent to weak existence plus pathwise uniqueness.

Theorem 1.4.8 *Pathwise uniqueness implies weak uniqueness.*

The following is a general existence result of weak solutions for equations with only continuous (not necessarily Lipschitz continuous) coefficients.

Theorem 1.4.9 *Let the processes* $(b(\omega, t, x))_{0 \leq t \leq T}$ *and* $(\sigma(\omega, t, x))_{0 \leq t \leq T}$ *are progressively measurable, be bounded and continuous. Then there exists a weak solution of* (E_1)*.*

1.5 Backward Stochastic Differential Equations

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P\right)$ be a complete probability space, where $(\mathcal{F}_t)_{0 \leq t \leq T}$ is the natural filtration of a standard Brownian motion $(W_t)_{0 \leq t \leq T}$ on $\mathbb{R}^m$. Consider the backward stochastic differential equation (BSDE)

$$(E_2) \begin{cases} dY_t = -f(t, Y_t, Z_t)dt + Z_t dW_t, \ 0 \leq t \leq T, \\ Y_T = \xi, \end{cases}$$

where the random variable ξ is a $\mathcal{F}_T$-measurable takes values in $\mathbb{R}^n$, and the coefficient or driver which is a function

$$f : \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \to \mathbb{R}^n,$$

and the processes $(f(\omega,t,y,z))_{0\leq t\leq T}$ is progessively measurable with respect to the completed Brownian filtration $(\mathcal{F}_t)_{0\leq t\leq T}$.

We now introduce the following spaces of processes

$\mathfrak{M}^2_{\mathcal{F}_t}\left(0,T,\mathbb{R}^{n\times d}\right)$:the space of $Y:[0,T]\times\Omega\longrightarrow\mathbb{R}^{n\times d}$, $\mathcal{F}_t$-progressively measurable such that

$$\mathbb{E}\left[\int_0^T |Y_t|^2\,dt\right]<\infty.$$

$\mathcal{S}^2_{\mathcal{F}_t}(0,T,\mathbb{R}^n)$:the space of $Y:[0,T]\times\Omega\longrightarrow\mathbb{R}^n$, $\mathcal{F}_t$-progressively measurable such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|Y_t|^2\right]<\infty.$$

Definition 1.5.1 *A solution of the (BSDE) (E_2) with driver f and terminal condition ξ is a pair $(Y,Z)\in \mathcal{S}^2_{\mathcal{F}_t}(0,T,\mathbb{R}^n)\times\mathfrak{M}^2_{\mathcal{F}_t}\left(0,T,\mathbb{R}^{n\times d}\right)$, satisfying*

1)

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|Y_t|^2+\int_0^T\|Z_t\|^2\,dt\right]<\infty,$$

2) (Y,Z) verifies (E_2) $P-a.s.$

Example 1.5.2 *Let $(W_t)_{0\leq t\leq T}$ be a Brownian motion, and for any $\xi\in L^2(\Omega,\mathcal{F}_T,P,\mathbb{R}^n)$, we put*

$$Y_t=\mathbb{E}\left[\xi\mid\mathcal{F}_t\right],$$

*so by **Martingale Representation Theorem** there exists a unique $Z_t\in\mathfrak{M}^2_{\mathcal{F}_t}\left(0,T,\mathbb{R}^{n\times d}\right)$ such that*

$$Y_t=\mathbb{E}\left[\xi\right]+\int_0^t Z_s dW_s,$$

and therefore

$$Y_t=\xi+\int_t^T Z_s dW_s.$$

The first existence and uniqueness result for non linear BSDEs has been proved by Pardoux and Peng [40]. This important paper has given rise to a huge literature on BSDEs and has become a powerful tool in many fields such as financial mathematics, optimal control, stochastic games etc.

(H_1) There exists a constant $K > 0$ such that for all $\omega \in \Omega$, $t \in [0, T]$, $(y_i, z_i) \in \mathbb{R}^n \times \mathbb{R}^{n\times d}$, $i = 1.2$

$$|f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)| \leq K(|y_1 - y_2| + \|z_1 - z_2\|),$$

(H_2) The process $(f(\omega, t, 0, 0))_{0\leq t\leq T}$ is in $\mathfrak{M}^2_{\mathcal{F}_t}\left(0, T, \mathbb{R}^{n\times d}\right)$.

Theorem 1.5.3 *Under the hypothesis $(H_1) - (H_2)$. Then BSDE (E_2) admits a unique solution (Y, Z) .*

1.5.1 Weak solutions

Definition 1.5.4 *A standard set-up $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P, W\right)$ along with a couple of processes (Y, Z) defined on this set-up is called a weak solution of (E_2) if*

1) W is a standard Brownian motion with respect to the filtration $(\mathcal{F}_t)_{0\leq t\leq T}$,

2) The processes Y is $\mathbb{R}^n$-valued and càdlàg, and Z is $\mathbb{R}^{n\times d}$-valued, with

$$\mathbb{E}\left[|Y_T|^2 + \int_0^T \|Z_t\|^2 dt\right] < \infty,$$

3) (Y, Z) verifies (E_2) $P - a.s.$

Remark 1.5.5 *The solution of (E_2) takes the form of a triplet (Y, Z, M) of processes defined on an extended probability space and satisfying*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s - (M_T - M_t), \qquad (E_3)$$

where M is a square integrable martingale which is orthogonal to W. In this case, a strong solution to (E_3) coincides with that of a strong solution to (E_2), because then M would be an $\mathcal{F}_t^W$- martingale, hence $M=0$.

As in the case of stochastic differential equations, one might expect that BSDEs with continuous generator always admit at least a weak solution, that is, a solution defined on a different probability space, generally with a larger filtration than the original one. Now, we prove the existence of a weak solution to a backward stochastic differential equation (BSDE)

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s - (M_T - M_t), \qquad (E_3)$$

where the coefficient or driver which is a function

$$f : \Omega \times [0,T] \times \mathbb{R}^n \times L(\mathbb{R}^n) \to \mathbb{R}^n,$$

which satisfies the following assumptions

(H_3) There exists a constant $C \geq 0$ such that $\forall (t, y, z) \in [0,T] \times \mathbb{R}^n \times L(\mathbb{R}^n)$,

$$|f(t, y, z)| \leq C(1 + \|z\|).$$

(H_4) $f(t, y, z)$ is continuous with respect to y and affine with respect to z.

We denote by $L(\mathbb{R}^n)$ the space of linear mappings from $\mathbb{R}^d$ to $\mathbb{R}^n$.

Theorem 1.5.6 *Under the hypothesis $(H_3) - (H_4)$. Then Eq. (E_3) admits a weak solution.*

Proof. See Bouchemella & Raynaud [17]. ■

Now we review the relationship between the strong and weak solutions by the following Proposition.

Proposition 1.5.7 *Existence of a weak solution and pathwise uniqueness for BSDE* (E_3) *imply existence of a strong solution. Conversely, if every solution to* (E_3) *is strong, (see Remark* $(1.5.5)$*), then pathwise uniqueness holds for Eq.* (E_3) *.*

CHAPTER 2

A mixed Relaxed-Singular Optimal Controls For Systems of MF-FBSDE Type.

CHAPTER 2

A mixed Relaxed-Singular Optimal Controls For Systems of MF-FBSDEs Type.

Forward-backward stochastic differential equations (FBSDEs in short) were first studied by Antonelli (see [1]), where the system of such equations is driven by Brownian motion on a small time interval. The proof there relies on the fixed point theorem. A weak solution for FBSDEs is given by Bahlali et al [15], where the original probability is changed using Girsanov's theorem. See also Antonelli and Ma [4] and Delarue and Guatteri [22], where the change of probability space comes from the construction of the forward component. A weak solution for FBSDEs where the filtration is enlarged, have been studied by Buckdahn et al. [9], (see also [7] and [8]), using pseudopaths and the Meyer-Zheng topology, [43]. In this Chapter, we consider a singular control problem for systems of forward-backward stochastic differential equations of mean-field type (MF-FBSDEs) in which the control variable consists of two components: an absolutely continuous control and a singular one. The coefficients depend on the states of the solution processes as well as their distribution via the expectation of some function. Moreover the cost functional is also of mean-field type. We prove in particular, the weak existence of optimal relaxed controls, which are measure-valued processes as well as the existencet optimal strict controls.

2.1 Formulation of the problems and assumptions

Let $(B_t, t \geq 0)$ is a d-dimensional Brownian motion defined on some complete filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P\right)$ and $\mathcal{M}$ is a square integrable martingale that is orthogonal to B, with $(\mathcal{F}_t)_{t\geq 0}$ being its natural filtration, augmented by all the P-null sets.

2.1.1 Regular-singular control problem

We consider the following sets U_1 is a non empty subset of $\mathbb{R}^k$ and $U_2 = ([0;\infty))^l$. The control variable is a suitable process (u, ξ) such that $u : [0,T] \times \Omega \to U_1, \xi : [0,T] \times \Omega \to U_2$ are $\mathcal{B}[0,T] \otimes \mathcal{F}$-measurable, $\mathcal{F}_t$-adapted, such that ξ is nondecreasing, left-continuous with right limits and $\xi_0 = 0, \mathbb{E}[|\xi_T|^2] < +\infty$. ξ is called singular control. Let $\mathcal{U}_1$ the set of regular control or absolutely continuous part of the control and $\mathcal{U}_2$ the set of singular part of the control.

Property of Singular control. Let for all functions $\xi : [0,T] \times \Omega \to U_2$ that are right limit with left continuous. We define $\Delta\xi_s = \xi_{s+} - \xi_s$ and set $\{s \in [0,T] \diagdown \Delta\xi_s = 0\}$. Then, the pure jump part of ξ is defined by $\xi_t^j = \sum_{0s \leq t} \Delta\xi_s$, and the continuous part is $\xi_t^c = \xi_t - \xi_t^j$. Note that ξ_t^c is bounded variation and differentiable almost everywhere, and we have by Lebesgue decomposition Theorem that $\xi_t^c = \xi_t^{ac} + \xi_t^{sc}$, $t \in [0,T]$, where ξ_t^{ac} is called the absolutely continuous part of ξ, and ξ_t^{sc} the singularly continuous part of ξ. Thus, we obtain that

$$\xi_t = \xi_t^{ac} + \xi_t^{sc} + \xi_t^{j},\ t \in [0,T],\ \text{unique.}$$

Remark 2.1.1 *i) If we assume that $\xi_t^j = 0$, $t \in [0,T]$, then the singular control reduces to a standard control problem , since we take $\xi_t^{ac} + \xi_t^{sc}$ as a new control variable.*

ii) If we assume that $\xi_t^{ac} + \xi_t^{sc} \equiv 0$, $t \in [0,T]$, then the singular control performs a special form of a pure jump process, so called impulse control.

Now, let us consider a regular-singular control problem governed by the following controlled MF-FBSDE

$$\left(E_1^{u,\xi}\right) \begin{cases} X_t^{u,\xi} = x + \int_0^t b\left(s, X_s^{u,\xi}, \mathbb{E}[\alpha(X_s^{u,\xi})], u_s\right) ds \\ \qquad\qquad + \int_0^t \sigma\left(s, X_s^{u,\xi}, \mathbb{E}[\gamma(X_s^{u,\xi})]\right) dB_s + \int_0^t \phi_t d\xi_s \\ \\ Y_t^{u,\xi} = h(X_T, \mathbb{E}[\theta(X_T^{u,\xi})]) + \int_t^T f\left(s, X_s^{u,\xi}, \mathbb{E}[\zeta(X_s^{u,\xi})], Y_s^{u,\xi}, \mathbb{E}[\eta(Y_s^{u,\xi})], u_s\right) ds \\ \qquad\qquad - \int_t^T Z_s^{u,\xi} dB_s - (\mathcal{M}_T - \mathcal{M}_t) + \int_t^T \varphi_t d\xi_s, \\ \\ X_0^{u,\xi} = x,\ Y_T^{u,\xi} = h(X_T^{u,\xi}, \mathbb{E}[\theta(X_T^{u,\xi})]). \end{cases}$$

where u_t is the absolutely continuous part of the control and ξ_t the singular part of the control, $\mathcal{M}_t$ is a càdlàg square integrable martingale which is orthogonal to B_t with $\mathcal{M}_0 = 0$, and the mappings

$$\begin{aligned}
b &: [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times U_1 \to \mathbb{R}^n, \\
\sigma &: [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n \times d}, \\
f &: [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \times U_1 \to \mathbb{R}^m, \\
h &: [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m, \\
\alpha, \gamma, \zeta, \theta &: [0,T] \times \mathbb{R}^n \to \mathbb{R}^n, \\
\eta &: [0,T] \times \mathbb{R}^m \to \mathbb{R}^m, \\
\phi &: [0,T] \to \mathbb{R}^{n \times l}, \\
\varphi &: [0,T] \to \mathbb{R}^{m \times l},
\end{aligned}$$

are measurable and attain some other properties to be introduced below.

It should be noted that the probability space and the Brownian motion may be change with the control. Therefore, we need to defined the admissible weak control, as follows:

Definition 2.1.2 *A 7-tuple $\vartheta. = \left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P, B., u., \xi.\right)$ is called admissible regular-singular control, and $\left(X_t^{u,\xi}, Y_t^{u,\xi}, Z_t^{u,\xi}, \mathcal{M}_t^{u,\xi}\right)$ a admissible triple if:*

i)-$\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P\right)$ is a filtered probability space satisfying the usual conditions and B_t is an d-dimensional standard Brownian motion defined on this space;

ii)-(u_t, ξ_t) is an $\mathcal{F}_t$-adapted process valued in the action space $U_1 \times U_2$. ξ_t is nondecreasing, left-continuous with right limits taking values in U_2 with $\xi_0 = 0, \mathbb{E}[|\xi_T|^2] < +\infty$.

iii)-$\left(X_t^{u,\xi}, Y_t^{u,\xi}, Z_t^{u,\xi}, \mathcal{M}_t^{u,\xi}\right)$ is the solution of the MF-FBSDE $\left(E_1^{u,\xi}\right)$ under (u_t, ξ_t).

iv)$\mathcal{M}_t^{u,\xi}$ is a square integrable $\mathcal{F}_t$-martingale, orthogonal to Brownian motion B_t.

The set of all admissible regular-singular controls is denoted by $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2$.

Consider a *cost functional*

$$\begin{aligned} J(\vartheta_\cdot) := \mathbb{E}\big[\Phi\left(X_T^{u,\xi}, \mathbb{E}[\lambda(X_T^{u,\xi})]\right) + \Psi\left(Y_0^{u,\xi}, \mathbb{E}[\rho(Y_0^{u,\xi})]\right) \\ + \int_0^T g\big(t, X_t^{u,\xi}, \mathbb{E}[\pi(X_t^{u,\xi})], Y_t^{u,\xi}, \mathbb{E}[\varpi(Y_t^{u,\xi})], u_t\big)dt + \int_t^T \psi_t d\xi_t\big], \end{aligned} \tag{2.1}$$

for measurable functions

$$\begin{aligned} \Phi &: [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \\ \Psi &: [0,T] \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}, \\ g &: [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \times U_1 \to \mathbb{R}, \\ \lambda, \pi &: [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \\ \varpi, \rho &: [0,T] \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m, \\ \psi &: [0,T] \to \mathbb{R}^l. \end{aligned}$$

The control problem is to minimize the functional $J(\cdot)$ over $\mathcal{U}$. We say that the admissible control ϑ^* is optimal control, if it satisfies

$$J\left(\vartheta^*_\cdot\right) = \inf_{\vartheta_\cdot \in \mathcal{U}} J\left(\vartheta_\cdot\right). \tag{2.2}$$

2.1.2 Relaxed-singular control problem

Due to the fact the existence of optimal solution of the regular control problem may fail to exist one typically seeks a certain compactness structure. The idea is then to extended the absolutely continuous part of the control u_t from the set U_1 to the set $P(U_1)$ of probability measures (q_t). These measure valued control are called relaxed control. If $q_t(du) = \delta_{u_t}(du)$ is a Dirac measure charging u_t for each t, then we get that the set of the absolutely continuous part of the control is a subset of the set of relaxed controls.

We denote by $\mathbb{V}(0,T;U_1)$ the space of positive Radon measure-valued processes $d\upsilon_t(u) = q_t(du)\,dt$, whose projections on $[0,T]$ coincide with Lebesgue measure dt. Equipped with the topology of stable convergence of measures, $\mathbb{V}(0,T;U_1)$ is a compact metrizable space, (see Jacod and Mémin [33]).

In this case, the state equation is defined by the following MF-FBSDE

$$\left(E_2^{q,\xi}\right)\begin{cases} X_t^{q,\xi} = x + \int_0^t \int_{U_1} b(s, X_s^{q,\xi}, \mathbb{E}[\alpha\left(X_s^{q,\xi}\right)], u) q_s(du)\, ds \\ \qquad\qquad + \int_0^t \sigma(s, X_s^{q,\xi}, \mathbb{E}[\gamma\left(X_s^{q,\xi}\right)]) dB_s + \int_0^t \phi_t d\xi_s \\ \\ Y_t^{q,\xi} = h(X_T^{q,\xi}, \mathbb{E}[\theta(X_T^{q,\xi})]) \\ \qquad\qquad + \int_t^T \int_{U_1} f(s, X_s^{q,\xi}, \mathbb{E}[\zeta\left(X_s^{q,\xi}\right)], Y_s^{q,\xi}, \mathbb{E}[\eta\left(Y_s^{q,\xi}\right)], u) q_s(du)\, ds \\ \qquad\qquad - \int_t^T Z_s^{q,\xi} dB_s - (\mathcal{M}_T - \mathcal{M}_t) + \int_t^T \varphi_t d\xi_s, \\ \\ X_0^{q,\xi} = x,\ Y_T^{q,\xi} = h(X_T^{q,\xi}, \mathbb{E}[\theta(X_T^{q,\xi})]). \end{cases}$$

The definition of admissible relaxed-singular control is given by:

Definition 2.1.3 *A 7-tuple $\mu_{\cdot} = \left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P, B_{\cdot}, q_{\cdot}, \xi_{\cdot}\right)$ is called admissible relaxed-singular control, and $\left(X_t^{q,\xi}, Y_t^{q,\xi}, Z_t^{q,\xi}, \mathcal{M}_t^{q,\xi}\right)$ a admissible triple if:*

a)-$\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P\right)$ is a filtered probability space satisfying the usual conditions and B_t is an d-dimensional standard Brownian motion defined on this space;

b)-q_t is $\mathcal{F}_t$-progressively measurable and such that for each $t, 1_{]0,t]} \cdot q$ is $\mathcal{F}_t$- measurable, taking values in $P(U_1)$. ξ_t is nondecreasing, left-continuous with right limits taking values in U_2 with $\xi_0 = 0, \mathbb{E}[|\xi_T|^2] < +\infty$.

c)-$\left(X_t^{q,\xi}, Y_t^{q,\xi}, Z_t^{q,\xi}, \mathcal{M}_t^{q,\xi}\right)$ is the solution of the MF-FBSDE $\left(E_2^{q,\xi}\right)$ under (q_t, ξ_t).

d)-$\mathcal{M}_t^{q,\xi}$ is a square integrable $\mathcal{F}_t$-martingale, orthogonal to Brownian motion B_t.

The set of all admissible relaxed-singular controls is denoted by $\mathcal{R}$.

Accordingly, the cost functional to be minimized over the set $\mathcal{R}$ of admissible relaxed control, well be given by:

$$\begin{aligned} J(\mu_\cdot) := \mathbb{E}\big[\Phi\left(X_T^{q,\xi}, \mathbb{E}[\lambda(X_T^{q,\xi})]\right) + \Psi\left(Y_0^{q,\xi}, \mathbb{E}[\rho(Y_0^{q,\xi})]\right) \\ + \int_0^T \int_{U_1} g\left(t, X_t^{q,\xi}, \mathbb{E}[\pi(X_t^{q,\xi})], Y_t^{q,\xi}, \mathbb{E}[\varpi(Y_t^{q,\xi})], u\right) q_t(du)dt + \int_0^T \psi_t d\xi_t\big]. \end{aligned} \tag{2.3}$$

A relaxed-singular control $\mu_\cdot^*$ is called optimal if it satisfies

$$J\left(\mu_\cdot^*\right) = \inf_{\mu_\cdot \in \mathcal{R}} J\left(\mu_\cdot\right). \tag{2.4}$$

2.1.3 Notation and assumptions

We now introduce the following spaces of processes:

$\mathfrak{M}^2(0,T;\mathbb{R}^m)$: the set of $\mathcal{F}_t$-measurable processes $\{Y_t, t \in [0,T]\}$ with values in $\mathbb{R}^m$ such that

$$\mathbb{E}\Big[\int_0^T |Y_t|^2 \, dt\Big] < \infty.$$

Let $\mathfrak{S}^2(0,T;\mathbb{R}^n)$: the set of $\mathcal{F}_t$-measurable processes $\{X_t, t \in [0,T]\}$ with values in $\mathbb{R}^n$ such that

$$\mathbb{E}\left[\sup_{0 \le t \le T} |X_t|^2\right] < \infty.$$

$\mathfrak{C}(0, T; \mathbb{R}^n)$: the space of continuous functions from $[0, T]$ to $\mathbb{R}^n$, equipped with the topology of uniform convergence.

$\mathbb{D}(0, T; \mathbb{R}^m)$: the Skorokhod space of càdlàg functions from $[0, T]$ to $\mathbb{R}^m$, that is functions which are continuous from the right with left hand limits, equipped with the S-topology of Jakubowski (see [31]).

Let us assume the following conditions regarding the state equation and the cost function

$(\mathbf{H1})$ Assume that the functions $b, \sigma, f, h, \alpha, \gamma, \zeta, \eta$, are bounded and continuous. Moreover assume that there exist a constant $K > 0$ such that for every $(x_1, x_2, x_1', x_2') \in \mathbb{R}^{4n}$, $(y_1, y_2, y_1', y_2')\, \mathbb{R}^{4m}$,

$$\begin{aligned}
&|f(t, x_1, x_2, y_1, y_2, u) - f(t, x_1', x_2', y_1', y_2', u)| \\
&\qquad\qquad \leq K\ (|x_1 - x_1'| + |x_2 - x_2'| + |y_1 - y_1'| + |y_2 - y_2'|)\,, \\
&|b(t, x_1, x_2, u) - b(t, x_1', x_2', u)| \\
&\qquad\qquad \leq K\ (|x_1 - x_1'| + |x_2 - x_2'|)\,, \\
&|\sigma(t, x_1, x_2) - \sigma(t, x_1', x_2')| \\
&\qquad\qquad \leq K\ (|x_1 - x_1'| + |x_2 - x_2'|)\,.
\end{aligned}$$

Also, the functions α, γ, ζ, are uniformly Lipschitz in x and η are uniformly Lipschitz in y. The functions ϕ, φ are positive, continuous and bounded.

$(\mathbf{H2})$ Assume that the functions $\Phi, \Psi, g, \lambda, \rho, \pi, \varpi$ are bounded and continuous and $h(t, \cdot, \cdot, \cdot, \cdot, u)$ is Lipschitz continuous uniformly in (t, u). The function ψ is continuous and bounded.

Befor we proceed with the definition of the optimization problem let us take abrief look at the existence and uniqueness of the solutions for the system. We observd that forward and backward equations are decoupled in the sens that $X_t^{u,\xi}$ does not depend on $Y_t^{u,\xi}$, we can solve the system separately. Now, we presente the following result.

Proposition 2.1.4 *Under assumptions* $(\mathbf{H1})$ *and for each admissible control* (u_t, ξ_t), *the system* $\left(E_1^{u,\xi}\right)$ *has a unique solution* $\left(X_t^{u,\xi}, Y_t^{u,\xi}, Z_t^{u,\xi}\right) \in \mathfrak{S}^2(0,T;\mathbb{R}^n)^2 \times \mathfrak{M}^2(0,T;\mathbb{R}^m)$.

Proof. The proof idea, is based on the representation of forward and backward equations respectively as continuous process plus integral with respect to singular control as follow:

$$\left(E_3^{u,\xi}\right) \begin{cases} X_t^{u,\xi} = \mathcal{X}_t^{u,\xi} + \int_0^t \phi_s d\xi_s \\ Y_t^{u,\xi} = \mathcal{Y}_t^{u,\xi} + \int_t^T \varphi_s d\xi_s, \end{cases}$$

such that

$$\left(E_4^{u,\xi}\right) \begin{cases} \mathcal{X}_t^{u,\xi} = x + \int_0^t b\left(s, X_s^{u,\xi}, \mathbb{E}[\alpha(X_s^{u,\xi})], u_s\right) ds + \int_0^t \sigma\left(s, X_s^{u,\xi}, \mathbb{E}[\gamma(X_s^{u,\xi})]\right) dB_s \\ \mathcal{Y}_t^{u,\xi} = h(X_T, \mathbb{E}[\theta(X_T^{u,\xi})]) + \int_t^T f\left(s, X_s^{u,\xi}, \mathbb{E}[\zeta(X_s^{u,\xi})], Y_s^{u,\xi}, \mathbb{E}[\eta(Y_s^{u,\xi})], u_s\right) ds - \int_t^T Z_s^{u,\xi} dB_s. \end{cases}$$

On the other hand, by use the Picard's iteration method on system $\left(E_4^{u,\xi}\right)$, we conclude the existence and uniqueness of solutions for systems $\left(E_3^{u,\xi}\right)$, for each admissible control (u_t, ξ_t). ∎

Remark 2.1.5 *The strong solution of system* $\left(E_1^{u,\xi}\right)$ *is the quartet* $\left(X_t^{u,\xi}, Y_t^{u,\xi}, Z_t^{u,\xi}, \mathcal{M}_t^{u,\xi}\right)$ *defined on the natural filtration of the Brownian motion* (B_t), *such that* $\mathcal{M}_t^{u,\xi}$ *is a cádlág martingale orthogonal to* B_t *with* $\mathcal{M}_0^{u,\xi} = 0$, *is coincides with the strong solution of* $\left(E_3^{u,\xi}\right)$, *because then* $\mathcal{M}_t^{u,\xi}$ *would be an* $\mathcal{F}_t^B$*-martingale, hence disappears, due to the uniqueness of solutions.*

2.2 Existence of an optimal control

2.2.1 Existence of an optimal relaxed-singular control

Our first results in this paper extends those of [10], and [14] to a systems governed by FBSDE of mean-field type.

Let us give some results on the tightness of the distributions of the processes defining the control problem.

Let $\mu_{\cdot}^{n} = \left(\Omega^{n}, \mathcal{F}^{n}, (\mathcal{F}_{t}^{n})_{t\geq 0}, P^{n}, B_{\cdot}^{n}, q_{\cdot}^{n}, \xi_{\cdot}^{n}\right)$ be a minimizing sequence, that is $\lim\limits_{n\rightarrow\infty} J(\mu_{\cdot}^{n}) = \inf\limits_{\mu_{\cdot}\in\mathcal{R}} J(\mu_{\cdot})$. Let $(X_{\cdot}^{q^{n},\xi^{n}}, Y_{\cdot}^{q^{n},\xi^{n}}, Z_{\cdot}^{q^{n},\xi^{n}})$ be the unique solution of the following MF-FBSDE

$$
\begin{cases}
X_{t}^{q^{n},\xi^{n}} = x + \int_{0}^{t}\int_{U_{1}} b\left(s, X_{s}^{q^{n},\xi^{n}}, \mathbb{E}\left[\alpha\left(X_{s}^{q^{n},\xi^{n}}\right)\right], u\right) q_{s}^{n}(du)\, ds \\
\qquad + \int_{0}^{t}\sigma\left(s, X_{s}^{q^{n},\xi^{n}}, \mathbb{E}\left[\gamma\left(X_{s}^{q^{n},\xi^{n}}\right)\right]\right) dB_{s}^{n} + \int_{0}^{t}\phi_{t} d\xi_{t}^{n}, \\
\\
Y_{t}^{q^{n},\xi^{n}} = h\left(X_{T}^{q^{n},\xi^{n}}, \mathbb{E}[\theta(X_{T}^{q^{n},\xi^{n}})]\right) \\
\qquad + \int_{t}^{T}\int_{U_{1}} f\left(s, X_{s}^{q^{n},\xi^{n}}, \mathbb{E}\left[\zeta\left(X_{s}^{q^{n},\xi^{n}}\right)\right], Y_{s}^{q^{n},\xi^{n}}, \mathbb{E}\left[\eta\left(Y_{s}^{q^{n},\xi^{n}}\right)\right], u\right) q_{s}^{n}(du)\, ds \\
\qquad - \int_{t}^{T} Z_{s}^{q^{n},\xi^{n}} dB_{s}^{n} + \int_{t}^{T}\varphi_{t} d\xi_{t}^{n}, \\
\\
X_{0}^{q^{n},\xi^{n}} = x,\ Y_{t}^{q^{n},\xi^{n}} = h\left(X_{T}^{q^{n},\xi^{n}}, \mathbb{E}[\theta(X_{T}^{q^{n},\xi^{n}})]\right).
\end{cases}
$$

The space $\mathcal{U}_{2}$ of singular control is equipped with the topology of weak convergence (see Haussmann and Suo [27]). The weak convergence is a convergence in measure in the sense, for $\xi^{n}, \xi \in \mathcal{U}_{2}, \xi^{n} \rightarrow \xi$ if and only if

$$\int_{0}^{T}\Gamma_{t} d\xi_{t}^{n} \rightarrow \int_{0}^{T}\Gamma_{t} d\xi_{t},$$

for any $\Gamma \in \mathfrak{C}\left(0,T;\mathbb{R}^n\right)$.

Let us defined the following processes

$$\mathcal{X}_t^{q^n,\xi^n} = X_t^{q^n,\xi^n} - \int_0^t \phi_s d\xi_s^n,$$

$$\mathcal{Y}_t^{q^n,\xi^n} = Y_t^{q^n,\xi^n} - \int_0^t \varphi_s d\xi_s^n.$$

Proposition 2.2.1 *There exists a positive constant C such that:*

$$\sup_n \mathbb{E}\big[\sup_{0\le t\le T} |\mathcal{X}_t^{q^n,\xi^n}|^2 + \sup_{0\le t\le T} \left|\mathcal{Y}_t^{q^n,\xi^n}\right|^2 + \int_0^T \left\|Z_s^{q^n,\xi^n}\right\|^2 ds\big] \le C, \tag{2.5}$$

where $(X_\cdot^{q^n,\xi^n}, Y_\cdot^{q^n,\xi^n}, Z_\cdot^{q^n,\xi^n})$ is the unique strong solution of the system (3.1).

Proof. By using the boundedness of b and σ, it is easy to prove that

For every $0 \le t \le T$ and $n > 1$ we have

$$\begin{aligned}\left|\mathcal{X}_t^{q^n,\xi^n}\right|^2 = \Big|x + &\textstyle\int_0^t \int_{U_1} b\left(s, X_s^{q^n,\xi^n}, \mathbb{E}\left[\alpha\left(X_s^{q^n,\xi^n}\right)\right], u\right) q_s^n(du)\, ds \\ &+ \textstyle\int_0^t \sigma\left(s, X_s^{q^n,\xi^n}, \mathbb{E}\left[\gamma\left(X_s^{q^n,\xi^n}\right)\right]\right) dB_s^n\Big|^2.\end{aligned}$$

Applying the inequality $(a+b+c) \le 3\left(a^2+b^2+c^2\right)$, we obtain

$$\begin{aligned}\left|\mathcal{X}_t^{q^n,\xi^n}\right|^2 \le 3\Big[x^2 + &\left|\textstyle\int_0^t \int_{U_1} b\left(s, X_s^{q^n,\xi^n}, \mathbb{E}\left[\alpha\left(X_s^{q^n,\xi^n}\right)\right], u\right) q_s^n(du)\, ds\right|^2 \\ &+ \left|\textstyle\int_0^t \sigma\left(s, X_s^{q^n,\xi^n}, \mathbb{E}\left[\gamma\left(X_s^{q^n,\xi^n}\right)\right]\right) dB_s^n\right|^2\Big].\end{aligned}$$

Passing to the expectation, we get

$$\begin{aligned}\mathbb{E}\left[\left|\mathcal{X}_t^{q^n,\xi^n}\right|^2\right] \le 3\Big[\mathbb{E}\left[x^2\right] + &\mathbb{E}\left[\left|\textstyle\int_0^t \int_{U_1} b\left(s, X_s^{q^n,\xi^n}, \mathbb{E}\left[\alpha\left(X_s^{q^n,\xi^n}\right)\right], u\right) q_s^n(du)\, ds\right|^2\right] \\ &+ \mathbb{E}\left[\left|\textstyle\int_0^t \sigma\left(s, X_s^{q^n,\xi^n}, \mathbb{E}\left[\gamma\left(X_s^{q^n,\xi^n}\right)\right]\right) dB_s^n\right|^2\right]\Big].\end{aligned}$$

By applying Holder inequality we have

$$\mathbb{E}\left[\left|\textstyle\int_0^t \int_U B\left(s, X_s^n, u\right) q_s^n(du)\, ds\right|^2\right] \le T \times \mathbb{E}\left[\textstyle\int_0^t \int_U \left|B\left(s, X_s^n, u\right)\right|^2 q_s^n(d\alpha)\, ds\right].$$

By the Burkholder–Davis–Gundy and Holder inequality provid that is inequality

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t \Sigma\left(s,X_s^n\right)dW_s^n\right|^2\right]\leq C\mathbb{E}\left[\int_0^T\left|\Sigma\left(s,X_s^n\right)\right|^2 ds\right].$$

Hence, we have

$$\mathbb{E}\left[\left|\mathcal{X}_t^{q^n,\xi^n}\right|^2\right]\leq C\left(\mathbb{E}\left[x^2\right]+\acute{C}\mathbb{E}\left[\int_0^T\int_A\left(\left|B\left(s,X_s^n,u\right)\right|^2+\left|\Sigma\left(s,X_s^n\right)\right|^2\right)q_s^n\left(d\alpha\right)ds\right]\right).$$

Now by using the boundedness of b and σ, it is easy to prove that

$$\sup_n\mathbb{E}\left[\sup_{0\leq t\leq T}|\mathcal{X}_t^{q^n,\xi^n}|^2\right]<+\infty.$$

Applying Itô's formula to $|\mathcal{Y}_t^{q^n,\xi^n}|^2$, we get

$$\begin{aligned}d|\mathcal{Y}_t^{q^n,\xi^n}|^2&=2|\mathcal{Y}_t^{q^n,\xi^n}|d\mathcal{Y}_t^{q^n,\xi^n}+d\langle\mathcal{Y}^{q^n,\xi^n},\mathcal{Y}^{q^n,\xi^n}\rangle_t\\&=-2\langle|\mathcal{Y}_t^{q^n,\xi^n}|,f\left(t,X_t^{q^n,\xi^n},\mathbb{E}[\zeta(X_t^{q^n,\xi^n})],Y_t^{q^n,\xi^n},\mathbb{E}[\eta(Y_t^{q^n,\xi^n})],u\right)\rangle dt\\&+2\langle|\mathcal{Y}_t^{q^n,\xi^n}|,Z_t^{q^n,\xi^n}\rangle dB_t^n+\left\|Z_t^{q^n,\xi^n}\right\|^2dt.\end{aligned}$$

Passing to the integral between t and T, we obtain

$$\begin{aligned}|\mathcal{Y}_T^{q^n,\xi^n}|^2-|\mathcal{Y}_t^{q^n,\xi^n}|^2=-2\int_t^T\langle|\mathcal{Y}_s^{q^n,\xi^n}|,f\left(s,X_s^{q^n,\xi^n},\mathbb{E}[\zeta(X_s^{q^n,\xi^n})],Y_s^{q^n,\xi^n},\mathbb{E}[\eta(Y_s^{q^n,\xi^n})],u\right)\rangle dt\\+2\int_t^T\langle|\mathcal{Y}_s^{q^n,\xi^n}|,Z_s^{q^n,\xi^n}\rangle dB_s^n+\int_t^T\left\|Z_s^{q^n,\xi^n}\right\|^2ds.\end{aligned}$$

And then,

$$\begin{aligned}|\mathcal{Y}_t^{q^n,\xi^n}|^2+\int_t^T\left\|Z_s^{q^n,\xi^n}\right\|^2ds=|\mathcal{Y}_T^{q^n,\xi^n}|^2\\+2\int_t^T\langle|\mathcal{Y}_s^{q^n,\xi^n}|,f\left(s,X_s^{q^n,\xi^n},\mathbb{E}[\zeta(X_s^{q^n,\xi^n})],Y_s^{q^n,\xi^n},\mathbb{E}[\eta(Y_s^{q^n,\xi^n})],u\right)\rangle dt\\-2\int_t^T\langle|\mathcal{Y}_s^{q^n,\xi^n}|,Z_s^{q^n,\xi^n}\rangle dB_s^n.\end{aligned}$$

Taking expectation, we get

$$\begin{aligned}
&\mathbb{E}\left[|\mathcal{Y}_t^{q^n,\xi^n}|^2\right]+\mathbb{E}\left[\int_t^T|Z_s^{q^n,\xi^n}|^2ds\right]=\mathbb{E}\left[|h\left(X_T^{q^n,\xi^n},\mathbb{E}[\theta(X_T^{q^n,\xi^n})]\right)|^2\right]\\
&+2\mathbb{E}\left[\int_t^T\int_{U_1}\langle\left|\mathcal{Y}_s^{q^n,\xi^n}\right|,f\left(s,X_s^{q^n,\xi^n},\mathbb{E}[\zeta(X_s^{q^n,\xi^n})],Y_s^{q^n,\xi^n},\mathbb{E}[\eta(Y_s^{q^n,\xi^n})],u\right)\rangle q_s^n(du)ds\right]\\
&\leq\mathbb{E}\left[|h\left(X_T^{q^n,\xi^n},\mathbb{E}[\theta(X_T^{q^n,\xi^n})]\right)|^2\right]+\mathbb{E}\left[\int_t^T|\mathcal{Y}_s^{q^n,\xi^n}|^2ds\right]\\
&+\mathbb{E}\left[\int_t^T\int_{U_1}\left|f\left(s,X_s^{q^n,\xi^n},\mathbb{E}[\zeta(X_s^{q^n,\xi^n})],Y_s^{q^n,\xi^n},\mathbb{E}[\eta(Y_s^{q^n,\xi^n})],u\right)\right|^2q_s^n(du)ds\right].
\end{aligned}$$

From the fact that g and f are bounded, applying Gronwall's lemma, we obtain

$$\sup_n\mathbb{E}\left[\sup_{0\leq t\leq T}|\mathcal{Y}_t^{q^n,\xi^n}|^2+\int_0^T\left\|Z_s^{q^n,\xi^n}\right\|^2ds\right]<+\infty.$$

∎

Proposition 2.2.2 *The sequence of distributions of processes* $(\mathcal{X}_\cdot^{q^n,\xi^n},B_\cdot^n,\mathcal{Y}_\cdot^{q^n,\xi^n},\int_0^\cdot Z_s^{q^n,\xi^n}dB_s^n,q_\cdot^n,\xi_\cdot^n)$ *is tight on the space* $\Lambda:=\mathfrak{C}\left(0,T;\mathbb{R}^n\right)\times\mathfrak{C}\left(0,T;\mathbb{R}^d\right)\times\mathbb{D}\left(0,T;\mathbb{R}^m\right)\times\mathbb{D}\left(0,T;\mathbb{R}^{m\times d}\right)\times\mathbb{V}(0,T;U)\times\mathcal{U}_2$ *endowed with the topology of uniform convergence for the first and second factor, endowed with the S-topology of Jakubowski (see[31]) for the third and forth factor, equipped with the topology of stable convergence for the fifth factor and equipped with the topology of weak convergence for the sixth factor.*

Proof. Using the boundedness of b and σ, we have

$$
\begin{aligned}
\mathbb{E}\left[\left|\mathcal{X}_t^{q^n,\xi^n}-\mathcal{X}_s^{q^n,\xi^n}\right|^4\right] &= \mathbb{E}\left[\left|\left(X_t^{q^n,\xi^n}-\int_0^t \varphi_r d\xi_r^n\right)-\left(X_s^{q^n,\xi^n}-\int_0^s \varphi_r d\xi_r^n\right)\right|^4\right] \\
&= \mathbb{E}[|\int_s^t \int_{U_1} b\left(r, X_r^{q^n,\xi^n}, \mathbb{E}\left[\alpha\left(X_r^{q^n,\xi^n}\right)\right], u\right) q_r^n(du)\, dr \\
&\qquad + \int_s^t \sigma\left(r, X_r^{q^n,\xi^n}, \mathbb{E}\left[\gamma\left(X_r^{q^n,\xi^n}\right)\right]\right) dB_r^n|^4] \\
&\leq C\mathbb{E}\left[\left(\int_s^t \int_{U_1}\left|b\left(r, X_r^{q^n,\xi^n}, \mathbb{E}\left[\alpha\left(X_r^{q^n,\xi^n}\right)\right], u\right)\right|^2 q_r^n(du)\, dr\right)^2\right] \\
&\quad + C\mathbb{E}\left[\left(\int_s^t \left|\sigma\left(r, X_r^{q^n,\xi^n}, \mathbb{E}\left[\gamma\left(X_r^{q^n,\xi^n}\right)\right]\right)\right|^2 dr\right)^2\right] \\
&\leq K_1 |t-s|^2 .
\end{aligned}
$$

By the same method, it is readily seen that there exists a constants K_2 independent from n such that

$$
\mathbb{E}\left[\left|B_t^n - B_s^n\right|^4\right] \leq K_2 |t-s|^2 ,
$$

for each $s,t \in [0,T]$. Hence the Kolmogorov tightness criteria is fulfilled (see Ikeda and Watanabe [30] page 18), then the sequence $(\mathcal{X}^{q^n,\xi^n}, B^n)$ is tight.

Let us prove that $(\mathcal{Y}_{\cdot}^{q^n,\xi^n}, \int_0^{\cdot} Z_s^{q^n,\xi^n} dB_s^n)$ is tight.

Let $0 = t_0 < t_1 < ... < t_n = T$. We have

$$
\begin{aligned}
CV\left(\mathcal{Y}_{\cdot}^{q^n,\xi^n}\right) &:= \sup \mathbb{E}\left[\sum_i \left|\mathbb{E}\left(\mathcal{Y}_{t_{i+1}}^{q^n,\xi^n} - \mathcal{Y}_{t_i}^{q^n,\xi^n}\right) \mid \mathcal{F}_{t_i}^{B^n}\right|\right] \\
&\leq C\mathbb{E}\left[\int_0^T \int_{U_1} \left|f\left(s, X_s^{q^n,\xi^n}, \mathbb{E}\left[\zeta\left(X_s^{q^n,\xi^n}\right)\right], Y_s^{q^n,\xi^n}, \mathbb{E}\left[\eta\left(Y_s^{q^n,\xi^n}\right)\right], u\right)\right| q_s^n(du)\, ds\right],
\end{aligned}
$$

where $CV\left(\mathcal{Y}_{\cdot}^{q^n,\xi^n}\right)$ is the conditional variation, the supremum is taken over all partitions of the interval $[0,T]$ and C is a constant depending only on t. By combining conditions $(\mathbf{H1})$ and

Proposition 3.1, we deduce that

$$\sup_{n}\left[CV\left(\mathcal{Y}_{\cdot}^{q^n,\xi^n}\right)+\sup_{0\leq t\leq T}\mathbb{E}\left[\left|\mathcal{Y}_{t}^{q^n,\xi^n}\right|\right]+\sup_{0\leq t\leq T}\mathbb{E}\left[\left|\int_{0}^{T}Z_{s}^{q^n,\xi^n}dB_{s}^{n}\right|\right]\right]<+\infty.$$

Thus the Meyer-Zheng tightness criteria is fulfilled (see [43]), then the sequences $\mathcal{Y}_{\cdot}^{q^n,\xi^n}$ and $\int_0^{\cdot} Z_s^{q^n,\xi^n}dB_s^n$ are tight.

Also the family of distributions associated to $\left(q_{\cdot}^{n}\right)_n$ is tight, from the fact that the space $\mathbb{V}(0,T;U_1)$ of probability measures on $[0,T]\times U_1$ is compact (Prokhorov's theorem).

For the tightness of ξ_t^n, we use the same technique as is Haussmann and Suo [27]. Define the set

$$V_M=\{\xi\in\mathcal{U}_2^n:|\xi_T|\leq M\},$$

V_M is then compact for any constant $M>0$. Further, define the set

$$\mathcal{R}^{\beta}=\{\mu\in\mathcal{R}:J(\mu)\leq\beta,\}$$

where β is chosen so that $\mathcal{R}^{\beta}$ is nonempty. Clearly, we can restrict the minimizing sequence to $\mathcal{R}^{\beta}$. It also holds that

$$\lim_{M\to\infty}\inf_{\mu_{\cdot}\in\mathcal{R}^{\beta}}P(|\xi_T|\leq M)=1.$$

Thus, for any $\varepsilon>0$ there exists a compact set V_M such that for all $P^n,\xi_{\cdot}^n\in\mathcal{R}^{\beta}$, we have

$$P^n(\xi_{\cdot}^n\in V_M)\geq 1-\varepsilon.$$

Which mean that the sequence ξ_t^n is tight. ∎

Theorem 2.2.3 *Under conditions* $(\mathbf{H1})$ *and* $(\mathbf{H2})$ *, the relaxed control problem* $\{(2.4),(2.5),(2.6)\}$ *has an optimal solution.*

Proof. From the Proposition 3.2, the sequence of processes

$\Theta^n_\cdot := \left(\mathcal{X}^{q^n,\xi^n}_\cdot, B^n_\cdot, \mathcal{Y}^{q^n,\xi^n}_\cdot, \int_0^\cdot Z^{q^n,\xi^n}_s dB^n_s, q^n_\cdot, \xi^n_\cdot\right)$ is tight on the space Λ. Thus by the Skorokhod's selection theorem, there exists a probability space $(\tilde{\Omega}, \mathcal{F}, \tilde{P})$, on which is defined a sequence $\tilde{\Theta}^n_\cdot := (\tilde{\mathcal{X}}^{\tilde{q}^n,\tilde{\xi}^n}_\cdot, \tilde{B}^n_\cdot, \tilde{\mathcal{Y}}^{\tilde{q}^n,\tilde{\xi}^n}_\cdot, \int_0^\cdot \tilde{Z}^{\tilde{q}^n,\tilde{\xi}^n}_s d\tilde{B}^n_s, \tilde{q}^n_\cdot, \tilde{\xi}^n_\cdot)$ identical in law to $\Theta^n_\cdot$ and converging on this space to $\tilde{\Theta}_\cdot = (\tilde{\mathcal{X}}, \tilde{B}_\cdot, \tilde{\mathcal{Y}}_\cdot, \tilde{\mathcal{N}}_\cdot, \tilde{q}_\cdot, \tilde{\xi}_\cdot)$ in the sense, there exist a countable subset D of $[0, T]$ such that

(i) on D^c, the sequence $(\tilde{\mathcal{Y}}^{\tilde{q}^n,\tilde{\xi}^n}_\cdot, \int_0^\cdot \tilde{Z}^{\tilde{q}^n,\tilde{\xi}^n}_s d\tilde{B}^n_s)$ converges to the càdlàg processes $(\tilde{\mathcal{Y}}_\cdot, \tilde{\mathcal{N}}_\cdot)$, $dt\times\tilde{P}$-a.s and $(\tilde{\mathcal{Y}}^{\tilde{q}^n,\tilde{\xi}^n}_T, \int_0^T \tilde{Z}^{\tilde{q}^n,\tilde{\xi}^n}_s d\tilde{B}^n_s)$ converges to $(\tilde{\mathcal{Y}}_T, \tilde{\mathcal{N}}_T)$ as $n \to \infty$, $\tilde{P}$-a.s.

(ii) $\sup\limits_{0\le t\le T} \left|\tilde{\mathcal{X}}^{q^n,\xi^n}_t - \tilde{\mathcal{X}}_t\right| \to 0, \quad \tilde{P}$-a.s.

(iii) $(\tilde{q}^n_\cdot)$ converges in the stable topology to $\tilde{q}_\cdot$, $\tilde{P}$-a.s.

(iv) $\left(\tilde{\xi}^n_\cdot\right)$ converges in the topology of weak convergence to $\tilde{\xi}$, $\tilde{P}$-a.s.

From the fact that, law$\Theta^n \equiv$ law$\tilde{\Theta}^n$, we have

$$\begin{cases} \tilde{\mathcal{X}}^{\tilde{q}^n,\tilde{\xi}^n}_t = x + \int_0^t \int_{U_1} b\left(s, \tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_s, \mathbb{E}[\alpha(\tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_s)], u\right) \tilde{q}^n_s(du)ds \\ \qquad\qquad + \int_0^t \sigma\left(s, \tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_s, \mathbb{E}[\gamma(\tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_s)]\right) d\tilde{B}^n_s, \\ \\ \tilde{\mathcal{Y}}^{\tilde{q}^n,\tilde{\xi}^n}_t = h\left(\tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_T, \mathbb{E}[\theta(\tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_T)]\right) \\ \qquad\qquad + \int_t^T \int_{U_1} f\left(s, \tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_s, \mathbb{E}[\zeta(\tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_s)], \tilde{Y}^{\tilde{q}^n,\tilde{\xi}^n}_s, \mathbb{E}[\eta(\tilde{Y}^{\tilde{q}^n,\tilde{\xi}^n}_s)], u\right) \tilde{q}^n_s(du)ds \\ \qquad\qquad - \left(\tilde{\mathcal{N}}^n_T - \tilde{\mathcal{N}}^n_t\right), \\ \\ \tilde{\mathcal{Y}}^{\tilde{q}^n,\tilde{\xi}^n}_t = h\left(\tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_T, \mathbb{E}[\theta(\tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_T)]\right), \end{cases}$$

where

$$\tilde{\mathcal{N}}^n_t = \int_0^t \tilde{Z}^{q^n,\xi^n}_s d\tilde{B}^n_s,\ \tilde{\mathcal{X}}^{\tilde{q}^n,\tilde{\xi}^n}_t = \tilde{X}^{\tilde{q}^n,\tilde{\xi}^n}_t - \int_0^t \phi_s d\tilde{\xi}^n_s,\ \tilde{\mathcal{Y}}^{\tilde{q}^n,\tilde{\xi}^n}_t = \tilde{Y}^{\tilde{q}^n,\tilde{\xi}^n}_t - \int_0^t \varphi_s d\tilde{\xi}^n_s.$$

Using properties (i), (i), (iii), under (**H1**)-(**H2**) and passing to the limit in the MF-FBSDE (3.3), we obtain

$$\begin{cases} \tilde{\mathcal{X}}_t^{\tilde{q},\tilde{\xi}} = x + \int_0^t \int_{U_1} b\left(s, \tilde{X}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\alpha(\tilde{X}_s^{\tilde{q},\tilde{\xi}})], u\right) \tilde{q}_s(du)ds \\ \qquad\qquad + \int_0^t \sigma\left(s, \tilde{X}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\gamma(\tilde{X}_s^{\tilde{q},\tilde{\xi}})]\right) d\tilde{B}_s, \\ \\ \tilde{\mathcal{Y}}_t^{\tilde{q},\tilde{\xi}} = h\left(\tilde{X}_T^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\theta(\tilde{X}_T^{\tilde{q},\tilde{\xi}})]\right) \\ \qquad\qquad + \int_t^T \int_{U_1} f\left(s, \tilde{X}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\zeta(\tilde{X}_s^{\tilde{q},\tilde{\xi}})], \tilde{Y}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\eta(\tilde{Y}_s^{\tilde{q},\tilde{\xi}})], u\right) \tilde{q}_s(du)ds \\ \qquad\qquad - \left(\tilde{\mathcal{N}}_T - \tilde{\mathcal{N}}_t\right), \\ \\ \tilde{\mathcal{Y}}_t^{\tilde{q},\tilde{\xi}} = h\left(\tilde{X}_T^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\theta(\tilde{X}_T^{\tilde{q},\tilde{\xi}})]\right), \end{cases} \tag{2.6}$$

with

$$\tilde{\mathcal{X}}_t^{\tilde{q},\tilde{\xi}} = \tilde{X}_t^{\tilde{q},\tilde{\xi}} - \int_0^t \phi_s d\tilde{\xi}_s,\ \tilde{\mathcal{Y}}_t^{\tilde{q},\tilde{\xi}} = \tilde{Y}_t^{\tilde{q},\tilde{\xi}} - \int_0^t \varphi_s d\tilde{\xi}_s. \tag{2.7}$$

Let $\tilde{\mathcal{F}}_s := \mathcal{F}_s^{\tilde{X}^{\tilde{q},\tilde{\xi}}, \tilde{Y}^{\tilde{q},\tilde{\xi}}, \tilde{q}, \tilde{\xi}}$, the minimal admissible and complete filtration generated by $(\tilde{X}_r^{\tilde{q},\tilde{\xi}},\ \tilde{Y}_r^{\tilde{q},\tilde{\xi}}, \tilde{q}_r, \tilde{\xi}_r,\ r \leq s)$. We can get easily that $\tilde{\mathcal{N}}$ is a $\tilde{\mathcal{F}}_s$-martingale. Therefore by the martingale decomposition theorem, there exist a process $\tilde{Z} \in \mathfrak{M}^2(0,T;\mathbb{R}^{m\times d})$ such that

$$\tilde{\mathcal{N}}_t = \int_0^t \tilde{Z}_s d\tilde{B}_s + \tilde{\mathcal{M}}_t,\ \text{and}\ \left\langle \tilde{\mathcal{M}}, \tilde{B} \right\rangle_t = 0,$$

thus, the backward part in (2.6) becomes

$$\begin{aligned} \tilde{\mathcal{Y}}_t^{\tilde{q},\tilde{\xi}} &= h\left(\tilde{X}_T^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\theta(\tilde{X}_T^{\tilde{q},\tilde{\xi}})]\right) \\ &+ \int_t^T \int_{U_1} f\left(s, \tilde{X}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\zeta(\tilde{X}_s^{\tilde{q},\tilde{\xi}})], \tilde{Y}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\eta(\tilde{Y}_s^{\tilde{q},\tilde{\xi}})], u\right) \tilde{q}_s(du)ds \\ &- \int_t^T \tilde{Z}_s d\tilde{B}_s - \left(\tilde{\mathcal{M}}_T - \tilde{\mathcal{M}}_t\right), \end{aligned} \tag{2.8}$$

replacing (2.7) and (2.8) in (2.6), we obtain

$$\begin{cases} \tilde{X}_t^{\tilde{q},\tilde{\xi}} = x + \int_0^t \int_{U_1} b\left(s, \tilde{X}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\alpha(\tilde{X}_s^{\tilde{q},\tilde{\xi}})], u\right) \tilde{q}_s(du)ds \\ \qquad\qquad + \int_0^t \sigma\left(s, \tilde{X}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\gamma(\tilde{X}_s^{\tilde{q},\tilde{\xi}})]\right) d\tilde{B}_s + \int_0^t \phi_s d\tilde{\xi}_s, \\ \\ \tilde{Y}_t^{\tilde{q},\tilde{\xi}} = h\left(\tilde{X}_T^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\theta(\tilde{X}_T^{\tilde{q},\tilde{\xi}})]\right) \\ \qquad\qquad + \int_t^T \int_{U_1} f\left(s, \tilde{X}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\zeta(\tilde{X}_s^{\tilde{q},\tilde{\xi}})], \tilde{Y}_s^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\eta(\tilde{Y}_s^{\tilde{q},\tilde{\xi}})], u\right) \tilde{q}_s(du)ds \\ \qquad\qquad - \int_t^T \tilde{Z}_s d\tilde{B}_s - \left(\tilde{\mathcal{M}}_T - \tilde{\mathcal{M}}_t\right) + \int_0^t \varphi_s d\tilde{\xi}_s, \\ \\ \tilde{Y}_t^{\tilde{q},\tilde{\xi}} = h\left(\tilde{X}_T^{\tilde{q},\tilde{\xi}}, \mathbb{E}[\theta(\tilde{X}_T^{\tilde{q},\tilde{\xi}})]\right). \end{cases} \tag{2.9}$$

To finish the proof of our first result (Theorem 3.3), it remains to prove that $\tilde{\mu}$ minimize the cost functional J over the set $\mathcal{R}$ of admissible relaxed-singular control.

Using properties (i)-(iii), and under assumptions (**H1**) and (**H2**), we have,

$$\begin{aligned} \inf_{\mu.\in\mathcal{R}} J(\mu.) &= \lim_{n\to\infty} J(\mu^n_\cdot) = \lim_{n\to\infty} J(\tilde{\mu}^n_\cdot), \\ &= \lim_{n\to\infty} \mathbb{E}\Big[\Phi\left(X_T^{q^n,\xi^n}, \mathbb{E}[\lambda(X_T^{q^n,\xi^n})]\right) + \Psi\left(Y_0^{q^n,\xi^n}, \mathbb{E}[\rho(Y_0^{q^n,\xi^n})]\right) \\ &\quad + \int_0^T \int_{U_1} g\left(t, X_t^{q^n,\xi^n}, \mathbb{E}[\pi(X_t^{q^n,\xi^n})], Y_t^{q^n,\xi^n}, \mathbb{E}[\varpi(Y_t^{q^n,\xi^n})], u\right) q_t^n(du)dt \\ &\qquad + \int_0^T \psi_t d\xi_t^n\Big], \\ &= \lim_{n\to\infty} \widetilde{\mathbb{E}}\Big[\Phi\left(\tilde{X}_T^{\tilde{q}^n,\tilde{\xi}^n}, \mathbb{E}[\lambda(\tilde{X}_T^{\tilde{q}^n,\tilde{\xi}^n})]\right) + \Psi\left(\tilde{Y}_0^{\tilde{q}^n,\tilde{\xi}^n}, \mathbb{E}[\rho(\tilde{Y}_0^{\tilde{q}^n,\tilde{\xi}^n})]\right) \\ &\quad + \int_0^T \int_{U_1} g\left(t, \tilde{X}_t^{\tilde{q}^n,\tilde{\xi}^n}, \mathbb{E}[\pi(\tilde{X}_t^{\tilde{q}^n,\tilde{\xi}^n})], \tilde{Y}_t^{\tilde{q}^n,\tilde{\xi}^n}, \mathbb{E}[\varpi(\tilde{Y}_t^{\tilde{q}^n,\tilde{\xi}^n})], u\right) \tilde{q}_t^n(du)dt \\ &\qquad + \int_0^T \psi_t d\tilde{\xi}_t^n\Big]. \end{aligned}$$

And hence

$$
\begin{aligned}
\inf_{\mu.\in\mathcal{R}} J(\mu.) = \widetilde{\mathbb{E}}\Big[&\Phi\left(\tilde{X}_T^{\widetilde{q},\widetilde{\xi}}, \mathbb{E}[\lambda(\tilde{X}_T^{\widetilde{q},\widetilde{\xi}})]\right) + \Psi\left(\tilde{Y}_0^{\widetilde{q},\widetilde{\xi}}, \mathbb{E}[\rho(\tilde{Y}_0^{\widetilde{q},\widetilde{\xi}})]\right) \\
&+ \int_0^T \int_{U_1} g\left(t, \tilde{X}_t^{\widetilde{q},\widetilde{\xi}}, \mathbb{E}[\pi(\tilde{X}_t^{\widetilde{q},\widetilde{\xi}})], \tilde{Y}_t^{\widetilde{q},\widetilde{\xi}}, \mathbb{E}[\varpi(\tilde{Y}_t^{\widetilde{q},\widetilde{\xi}})], u\right) \tilde{q}_t(du)dt \\
&+ \int_0^T \psi_t d\widetilde{\xi}_t\Big], \\
= J(\tilde{\mu}.),
\end{aligned}
$$

thus $\widetilde{\mu}. = \left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \left(\widetilde{\mathcal{F}}_t\right)_{t\geq 0}, \widetilde{P}, \widetilde{B}., \widetilde{q}., \widetilde{\xi}.\right)$ is an relaxed-singular optimal control, then theorem (3.3) is proved. ■

2.2.2 Existence of an optimal strict-singular control

We prove in this subsection the existence of optimal solution to the control problem $\{(2.1), (2.2), (2.3)\}$. In this end, we need the following Roxin's condition:

(**H3**) : (Roxin-type convexity condition): The set

$$(b, f, g)(t, x, x', y, y', U_1) := \{b_i(t, x, x', u),$$

$$f_j(t, x, x', y, y', u), g(t, x, x', y, y', u) \diagdown u \in U_1, i = 1, \cdots, n, j = 1, \cdots, m\},$$

is convex and closed in $\mathbb{R}^{n+m+1}$.

Proposition 2.2.4 *Assume that* (**H1**)-(**H3**) *hold. Then, the strict-singular control problem* $\{(2.1), (2.2), (2.3)\}$, has an optimal solution.

Proof. The proof is inspired from that given in Yong and Zhou [49] (proof of theorem 5.3, page 71).

From (2.9), we put

$$\int_{U_1} b\left(t, \tilde{X}_t, \mathbb{E}[\alpha(\tilde{X}_t)], u\right) \tilde{q}_t(du) := \tilde{b}\,(t, w) \in b\,(t, x, x', U_1)\,,$$

$$\int_{U_1} f\left(t, \tilde{X}_t, \mathbb{E}[\zeta(\tilde{X}_t)], \tilde{Y}_t, \mathbb{E}[\eta(\tilde{Y}_t)], u\right) \tilde{q}_t(du)$$

$$:= \tilde{f}\,(t, w) \in f\,(t, x, x', y, y', U_1)\,,$$

$$\int_{U_1} g\left(t, \tilde{X}_t, \mathbb{E}[\pi(\tilde{X}_t)], \tilde{Y}_t, \mathbb{E}[\varpi(\tilde{Y}_t)], u\right) \tilde{q}_t(du)$$

$$:= \tilde{h}\,(t, w) \in h\,(t, x, x', y, y', U_1)\,.$$

Under (**H3**) and the measurable selection theorem (see Li-Yong [41] p. 102, Corollary 2.26), there is a U_1-valued, $\widetilde{\mathcal{F}}_t$-adapted process $\tilde{u}$, such that for every $t \in [0, T]$ and $w \in \tilde{\Omega}$,

$$\left(\tilde{f}, \tilde{g}\right)(t, w) = (f, h)\left(t, \tilde{X}\,(t, w)\,, \tilde{X}'\,(t, w)\,, \tilde{Y}\,(t, w)\,, \tilde{Y}'\,(t, w)\,, \tilde{u}\,(t, w)\right),$$

and

$$\tilde{b}\,(t, w) = b\left(t, \tilde{X}\,(t, w)\,, \tilde{X}'\,(t, w)\,, \tilde{u}\,(t, w)\right).$$

Thus, we have

$$\int_{U_1} b\left(t, \tilde{X}_t, \mathbb{E}[\alpha(\tilde{X}_t)], u\right) \tilde{q}_t(du) = b\left(t, \tilde{X}_t, \mathbb{E}[\alpha(\tilde{X}_t)], \tilde{u}_t\right),$$

$$\int_{U_1} f\left(t, \tilde{X}_t, \mathbb{E}[\gamma(\tilde{X}_t)], \tilde{Y}_t, \mathbb{E}[\delta(\tilde{Y}_t)], u\right) \tilde{q}_t(du)$$

$$= f\left(t, \tilde{X}_t, \mathbb{E}[\gamma(\tilde{X}_t)], \tilde{Y}_t, \mathbb{E}[\delta(\tilde{Y}_t)], \tilde{u}_t\right),$$

$$\int_{U_1} g\left(t, \tilde{X}_t, \mathbb{E}[\pi(\tilde{X}_t)], \tilde{Y}_t, \mathbb{E}[\varpi(\tilde{Y}_t)], u\right) \tilde{q}_t(du)$$

$$= g\left(t, \tilde{X}_t, \mathbb{E}[\pi(\tilde{X}_t)], \tilde{Y}_t, \mathbb{E}[\varpi(\tilde{Y}_t)], \tilde{u}_t\right).$$

Hence (2.9) becomes

$$\begin{cases} \tilde{X}_t = x + \int_0^t b\left(s, \tilde{X}_s, \mathbb{E}[\alpha(\tilde{X}_s)], \tilde{u}\right) ds \\ \qquad\qquad + \int_0^t \sigma\left(s, \tilde{X}_s, \mathbb{E}[\gamma(\tilde{X}_s)]\right) d\tilde{W}_s + \int_0^t \phi_s d\widetilde{\xi}_s, \\ \\ \tilde{Y}_t = h\left(\tilde{X}_T, \mathbb{E}[\theta(\tilde{X}_T)]\right) + \int_t^T f\left(s, \tilde{X}_s, \mathbb{E}[\zeta(\tilde{X}_s)], \tilde{Y}_s, \mathbb{E}[\eta(\tilde{Y}_s)], \tilde{u}\right) ds \\ \qquad\qquad - \int_t^T \tilde{Z}_s d\tilde{W}_s - \left(\tilde{M}_T - \tilde{M}_t\right) + \int_t^T \varphi_s d\widetilde{\xi}_s. \\ \\ \tilde{Y}_t = h\left(\tilde{X}_T, \mathbb{E}[\theta(\tilde{X}_T)]\right). \end{cases}$$

Moreover,

$$\begin{aligned} J(\tilde{q}) &= \widetilde{\mathbb{E}}\Big[\Phi\left(\tilde{X}_T, \mathbb{E}[\lambda(\tilde{X}_T)]\right) + \Psi\left(\tilde{Y}_0, \mathbb{E}[\rho(\tilde{Y}_0)]\right) \\ &\qquad + \int_0^T \int_U g\left(t, \tilde{X}_t, \mathbb{E}[\pi(\tilde{X}_t)], \tilde{Y}_t, \mathbb{E}[\varpi(\tilde{Y}_t)], u\right) \tilde{q}_t(du) dt + \int_0^T \psi_s d\widetilde{\xi}_s\Big] \\ &= \widetilde{\mathbb{E}}\Big[\Phi\left(\tilde{X}_T, \mathbb{E}[\lambda(\tilde{X}_T)]\right) + \varpi\left(\tilde{Y}_0, \mathbb{E}[\rho(\tilde{Y}_0)]\right) \\ &\qquad + \int_0^T g\left(t, \tilde{X}_t, \mathbb{E}[\pi(\tilde{X}_t)], \tilde{Y}_t, \mathbb{E}[\varpi(\tilde{Y}_t)], \tilde{u}_t\right) dt + \int_0^T \psi_s d\widetilde{\xi}_s\Big] \\ &= J(\tilde{\vartheta}), \end{aligned}$$

where $\tilde{\vartheta} = \left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left(\tilde{\mathcal{F}}_t\right)_{t\geq 0}, \tilde{P}, \tilde{W}, \tilde{u}\right)$. Which ends the proof. ■

CHAPTER 3

Existence of Optimal Controls For Systems of Controlled Forward-Backward Doubly SDEs.

CHAPTER 3

Existence of Optimal Controls For Systems of Controlled Forward-Backward Doubly SDEs.

We wish to study a class of optimal controls for problems governed by forward-backward doubly stochastic differential equations (FBDSDEs). Firstly, we prove existence of optimal relaxed controls, which are measure-valued processes for nonlinear FBDSDEs, by using some tightness properties and weak convergence techniques on the space of Skorokhod of càdlàg functions equipped with the S-topology of Jakubowsky. Moreover, when the Roxin-type convexity condition is fulfilled, we prove that the optimal relaxed control is in fact strict.

3.1 Introduction to Backward Doubly Stochastic Differential Equations BDSDEs

Pardoux and Peng in [39], introduced a new class of BSDEs with two different directions of stochastic integrals, called backward doubly stochastic differential equations (BDSDEs) with the form

$$Y_t = Y_T + \int_t^T F(s, Y_s, Z_s)\, ds + \int_t^T G(s, Y_s, Z_s)\, d\overleftarrow{B_s} - \int_t^T Z_s dW_s, \text{ for all } t \in [0, T], \tag{3.1}$$

where terminal value $Y_T = \xi$ and the functions F, G are given, the integral with respect to B_t is a backward Itô integral and the integral with respect to W_t is a standard forward Itô integral. Those two types of integrals are particular cases of the Ito-Skorohod integral, see Nualart and Pardoux [44]. Let us now recall the definition of backward stochastic integrals, we denote by $(\pi^n)_{n\geq 0}$ any sequence of subdivisions:

$\pi^n = \{t = t_0^n < t_1^n < ... < t_n^n = T\}$. Such that $|\pi^n| = \sup_{0\leq k\leq n-1}\left(B_{t_{k+1}} - B_{t_k}\right) \to 0$ as $n \to \infty$. Then the backward Itô integral can be defined as

$$\int_t^T H_s d\overleftarrow{B_s} = \lim_{n\to 0} \sum_{k=0}^{n-1} H_{t_{k+1}}\left(B_{t_{k+1}} - B_{t_k}\right),$$

is $\mathcal{F}_{t,T}^B = \sigma(B_r - B_T, t \leq r \leq T)$ backward martingale. In fact, the backward Itô integral of H_s with respect to B_s may be understood as the forward integral of $\tilde{H}_s = H_{T-s}$ with respect to

$\tilde{B}_s = B_{T-s} - B_T$ such that

$$\begin{aligned}\int_0^{T-t} \tilde{H}_s d\tilde{B}_s &= \lim_{n\to 0} \sum_{k=0}^{n-1} \tilde{H}_{t_k} \left(\tilde{B}_{t_{k+1}} - \tilde{B}_{t_k}\right) \\ &= \lim_{n\to 0} \sum_{k=0}^{n-1} H_{T-t_k} \left(B_{T-t_{k+1}} - B_{T-t_k}\right),\end{aligned}$$

note that $r_{k+1} = T - t_k < r_k = T - t_{k+1}$ is a subdivision of $[t, T]$ then we get

$$\begin{aligned}\int_0^{T-t} \tilde{H}_s d\tilde{B}_s &= -\lim_{n\to 0} \sum_{k=0}^{n-1} H_{r_{k+1}} \left(B_{r_{k+1}} - B_{r_k}\right) \\ &= -\int_t^T H_s d\overleftarrow{B_s}.\end{aligned}$$

On the other hand we have

$$\begin{aligned}\mathcal{F}^B_{T-t,T} &= \sigma\left(B_r - B_T, T - t \le r \le T\right) \\ &= \sigma\left(B_{T-s} - B_T, 0 \le s \le t\right) \\ &= \mathcal{F}^{\tilde{B}}_t.\end{aligned}$$

The same if $\tilde{W}_s = W_{T-s} - W_T$ we have $\mathcal{F}^{\tilde{W}}_{T-t,T} = \mathcal{F}^W_t \ \forall 0 \le t \le T$.

In [39], the existence and uniqueness of solution are established under uniformly Lipschitz condition on the coefficients. It is worth noting that the definition of solution of this type of equations is slightly different from that of classical (BSDEs). The BDSDEs (3.1) can be related to semilinear and quasilinear stochastic partial differential equations (SPDEs). This link was developed in many papers and has motivated many efforts to establish the existence and uniqueness of solutions under more general conditions than the global Lipschitz see for exmple ([5] [48] [50])

3.1.1 Existence and Uniqueness of BDSDEs

In this subsection, we will discuss the existence and uniqueness of adapted solution for (BDSDEs) (3.1). The basic ideas used to prove existence and unique are the use of the martingale representation property of Brownian motion and a Picard iteration scheme are used. Now we cite the some results.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $(W_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$ be two mutually independent standard Brownian motions, with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^k$. Let $\mathcal{N}$ denote the class of $P-$null sets of $\mathcal{F}$. For each $t \in [0;T]$, we define $\mathcal{F}_t := \mathcal{F}^B_{t,T} \vee \mathcal{F}^W_t$, where for any process θ_t, $\mathcal{F}^\theta_{s,t} = \sigma\left(\theta_r - \theta_s, s \leq r \leq t\right) \vee \mathcal{N}$, $\mathcal{F}^\theta_t = \mathcal{F}^\theta_{0,t}$. Note that the collection $(\mathcal{F}_t)_{0\leq t\leq T}$ is neither increasing nor decreasing, and it does not constitute a filtration.

We introduce the following notation:

$L^2_T(\mathbb{R}^m)$: is the space of $\mathbb{R}^m-$valued $\mathcal{F}_T-$measurable random variables ξ satisfying

$$\mathbb{E}\big[\, |\xi|^2 \,\big] < \infty,$$

$\mathcal{M}^2\left(0, T, \mathbb{R}^{m\times d}\right)$: the space of $\mathcal{F}_t-$measurable processes Z_t defined from $[0,T]\times\Omega$ into $\mathbb{R}^{m\times d}$, such that

$$\mathbb{E}\big[\int_0^T \|Z_t\|^2\, dt\big] < \infty,$$

and $\mathcal{S}^2\left(0, T, \mathbb{R}^m\right)$: the space of $\mathcal{F}_t-$measurable processes Y_t defined from $[0,T]\times\Omega$ into $\mathbb{R}^m$, such that

$$\mathbb{E}\big[\sup_{0\leq t\leq T} |Y_t|^2\,\big] < \infty.$$

The equation we want to solve is

$$\begin{cases} -dY_t = F(s, Y_s, Z_s)\,ds + G(s, Y_s, Z_s)\,d\overleftarrow{B_s} - Z_s dW_s, \text{ for all } t \in [0, T], \\ Y_T = \xi. \end{cases} \tag{3.2}$$

Definition 3.1.1 *A solution of equation (3.2) is a couple* (Y, Z) *which belongs to the space* $\mathcal{S}^2(0, T, \mathbb{R}^n) \times \mathcal{M}^2(0, T, \mathbb{R}^{m\times d})$ *and satisfes (3.2).*

Next we consider the following assumptions

$$F : [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \longrightarrow \mathbb{R}^m$$

$$G : [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \longrightarrow \mathbb{R}^{m\times d},$$

are measurable for each $(y, z) \in \mathbb{R}^m \times \mathbb{R}^{m\times l}$, with $\xi \in L^2_T(\mathbb{R}^m)$ and $F(., y, z) \in \mathcal{M}^2(0, T, \mathbb{R}^m)$, $G(., y, z) \in \mathcal{M}^2(0, T, \mathbb{R}^{m\times d})$ respectively.

We assume moreover that there exist a constant $K > 0$ and $0 < \alpha < 1$, such that for every $(y_1, y_2) \in \mathbb{R}^{2m}$, $(z_1, z_2) \in \mathbb{R}^{2m\times l}$

$$|F(t, y_1, z_1) - F(t, y_2, z_2)|^2 \leq K\left(|y_1 - y_2|^2 + \|z_1 - z_2\|^2\right)$$

$$\|G(t, x_1, y_1) - G(t, x_2, y_2)\|^2 \leq K\left(|y_1 - y_2|^2 + \alpha\|z_1 - z_2\|^2\right).$$

The main result of this subsection is the following

Theorem 3.1.2 *Under the above conditions, the (BDSDEs) (3.2) has a unique solution* $(Y, Z) \in \mathcal{S}^2(0, T, \mathbb{R}^n) \times \mathcal{M}^2(0, T, \mathbb{R}^{m\times d})$.

Proof. Can be found in Pardoux and Peng in [39]. ■

Now we introduce a type of forward doubly stochastic differential equation as follows

$$Y_t = \eta + \int_0^t F\left(s, Y_s, Z_s\right) ds + \int_0^t G\left(s, Y_s, Z_s\right) dB_s - \int_0^t Z_s d\overleftarrow{W}, \text{ for all } t \in [0, T], \tag{3.3}$$

where $Y_0 = \eta$ is $\mathcal{F}_0-$measurable. we recall that this type of equations appears in stochastic control problems. For example, in the stochastic formula of the principle of Pontryagin Maximum for controlled systems whose dynamics are subject to a backward doubly stochastic differential equations (BDSDEs). Now under what was mentioned in the previous subsection about the backward Itô integral, we will transform the equation (3.3) into equation of type (3.1), so we have $\tilde{B}_s = B_{T-s} - B_T, \tilde{W}_s = W_{T-s} - W_T$ and $\tilde{\mathcal{F}}_t = \mathcal{F}_{t,T}^{\tilde{B}} \vee \mathcal{F}_t^{\tilde{W}}$ $\forall 0 \leq t \leq T$ and hence

$$\tilde{Y}_t = \eta + \int_t^T F\left(s, \tilde{Y}_s, \tilde{Z}_s\right) ds + \int_t^T G\left(s, \tilde{Y}_s, Z_s\right) d\overleftarrow{\tilde{B}}_s - \int_t^T Z_s d\tilde{W}_s, \tag{3.4}$$

where $\tilde{Y}_t = Y_{T-t}, \tilde{Z}_t = Z_{T-t}$ for all $t \in [0, T]$, and η is $\tilde{\mathcal{F}}_T-$measurable. Note that (3.4) have the similar form as (3.1), then by Theorem previous there exists a unique pair of $\left(\tilde{Y}_t, \tilde{Z}_t\right)$ solving (3.4), and we have

Theorem 3.1.3 *Under the same previous conditions, the (FDSDEs) (3.3) has a unique solution* $\left(\tilde{Y}, \tilde{Z}\right) \in \mathcal{S}^2\left(0, T, \mathbb{R}^n\right) \times \mathcal{M}^2\left(0, T, \mathbb{R}^{m \times d}\right)$.

3.2 Existence of optimal controls for nonlinear FBDSDEs

3.2.1 Statement of the problems and assumptions

3.2.1.1 Strict control problem

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $(W_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$ be two mutually independent standard Brownian motions, with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^k$. Let $\mathcal{N}$ denote the class of P-null sets of $\mathcal{F}$. For each $t \in [0;T]$, we define $\mathcal{F}_t := \mathcal{F}_{t,T}^B \vee \mathcal{F}_t^W$, where for any process θ_t, $\mathcal{F}_{s,t}^\theta = \sigma(\theta_r - \theta_s, s \leq r \leq t) \vee \mathcal{N}$, $\mathcal{F}_t^\theta = \mathcal{F}_{0,t}^\theta$.

Note that the collection $(\mathcal{F}_t)_{0\leq t\leq T}$ is neither increasing nor decreasing, and it does not constitute a filtration.

We want to prove the existence of optimal strict controls for a control problem driven by the following FBDSDE

$$\begin{cases} X_t = x + \int_0^t B(s, X_s, u_s)\, ds + \int_0^t \Sigma(s, X_s)\, dW_s, \\ \\ Y_t = H(X_T) + \int_t^T F(s, X_s, Y_s, u_s)\, ds + \int_t^T G(s, X_s, Y_s)\, d\overleftarrow{B_s} - \int_t^T Z_s dW_s \\ \\ \qquad - (M_T - M_t), \end{cases} \tag{3.5}$$

where $M.$ is a square integrable martingale which is orthogonal to $W.$. With $M_0 = 0$ and with càdlàg trajectories.

Remark 3.2.1 *1)-Such a weak solution to the FBDSDE (3.5) can be considered as a generalized weak* *solution to the more classical controlled system*

$$\begin{cases} X_t = x + \int_0^t B\left(s, X_s, u_s\right) ds + \int_0^t \Sigma\left(s, X_s\right) dW_s, \\ \\ Y_t = H\left(X_T\right) + \int_t^T F\left(s, X_s, Y_s, u_s\right) ds + \int_t^T G\left(s, X_s, Y_s\right) d\overleftarrow{B_s} - \int_t^T Z_s dW_s, \end{cases} \tag{3.6}$$

2)-Existence and uniqueness of strong solution (X, Y, Z) *of the backward doubly component of (3.6) can be proved by using the same method given in [39], where* $u.$ *is the control variable and* $X.$ *is the solution of a forward SDE, if we assume that*

- *(H4)* $F(\cdot, x, y, u)$ *and* $G(\cdot, x, y)$ *are square integrable with respect to the associate norm, there exist constants* $K > 0$ *such that for any* $(w, t) \in \Omega \times [0, T], (x, y), (x', y') \in \mathbb{R}^n \times \mathbb{R}^m$,

$$|F(t, x, y, u) - F(t, x', y', u')|^2 \leq K(|x - x'|^2 + |y - y'|^2),$$

$$\|G(t, x, y) - G(t, x', y')\|^2 \leq K(|x - x'|^2 + |y - y'|^2),$$

and $H(X_T)$ *is square integrable and* $\mathcal{F}_T$*-measurable.*

Similarly, a strong solution to the backward doubly component of (3.5) should be a quadruple (X, Y, Z, M) *defined on* $\Omega \times [0, T]$, *such that* $(H4)$ *is satisfied and* $M.$ *is a càdlàg square integrable martingale which is orthogonal to* $W.$. *With* $M_0 = 0$, *but this notation coincides with that of strong solution to the backward doubly component of (3.6), because then* $M.$ *would be an* $\mathcal{G}_t = \mathcal{F}_t^W \vee \mathcal{F}_T^B$*-martingale and* $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_{t,T}^B$*-measurable, hence* $M. = 0$.

We introduce the concept of admissible controls to the FBDSDE (3.5).

Definition 3.2.2 *A 7-tuple* $\pi_{\cdot} = \left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P, W_{\cdot}, B_{\cdot}, u_{\cdot}\right)$ *is called admissible strict control, and* (X_t, Y_t, Z_t, M_t) *an admissible quadruple if:*

i)-$(\Omega, \mathcal{F}, P)$ *is a probability space;*

ii)-W_t *and* B_t *be two mutually independent standard Brownian motion defined on* $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$, *with* $\mathcal{F}_t := \mathcal{F}_{t,T}^{B} \vee \mathcal{F}_t^{W}$;

iii)-u_t *is an* $\mathcal{F}_t$*-measurable process on* $(\Omega, \mathcal{F}, P)$ *valued in* U *which is a nonempty Borel compact subset of* $\mathbb{R}^r$;

iv)-(X_t, Y_t, Z_t, M_t) *is the solution of the FBDSDE (3.5) on* $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ *under* u_t.

v)-M_t *is a square integrable martingale, orthogonal to Brownian motion* W_t.

The set of all admissible strict controls $\pi_{\cdot}$ *is denoted by* $\mathcal{U}_{ad}$.

The functional cost to be minimized, over the set $\mathcal{U}_{ad}$ of strict control, is given by

$$J(u_{\cdot}) := \mathbb{E}\big[\varphi\left(X_T\right) + \psi(Y_0) + \textstyle\int_0^T L(t, X_t, Y_t, u_t)dt\big].$$

An admissible strict control $\pi_{\cdot}^{*}$, is called optimal if it satisfy

$$J\left(\pi_{\cdot}^{*}\right) = \inf_{\pi_{\cdot} \in \mathcal{U}_{ad}} J\left(\pi_{\cdot}\right). \tag{3.7}$$

3.2.1.2 Relaxed control problem

Let us introduce the concept of relaxed controls which gives a more suitable topological structure. The weak formulation enables us to find the compactness of the image measure of some processes involved on a certain functional space, then an optimal control may fail to exist in the set U of strict controls.

To be convinced on the fact that strict optimal controls may not exist even in the simplest cases, let us consider a deterministic example.

Example 3.2.3 *Let* $U = \{-1, 1\}, \mathcal{U}_{ad} := \{u. : [0, 1] \to \{-1, 1\}/u.\text{measurable}\}$, *and* $J(u.) = \int_0^1 X_t^2 dt$, *where* X_t *denotes the solution of*

$$\begin{cases} dX_t = u_t dt, \\ X_0 = 0, t \in [0, 1]. \end{cases} \tag{3.8}$$

The optimal control problem is to find a pair $(\widetilde{X}, \widetilde{u})$ *such that*

$$J(\widetilde{u}) = \inf_{u. \in \mathcal{U}_{ad}} J(u.),$$

and $\widetilde{X}$ *satisfies (3.8). We will show that:*

$$J(\widetilde{u}) = \inf_{u. \in \mathcal{U}_{ad}} J(u.) = 0.$$

Indeed, for any $n > 0$, *we consider the sequence*

$$U_t^n = (-1)^k, \frac{k}{n} \leq t \leq \frac{k+1}{n}, 0 \leq k \leq n-1.$$

Then immediately, we have $|X_t^n| \leq \frac{1}{n}$, *where* X_t^n *is the solution of (3.8) associated with* U^n, *thus* $|J(u^n)| \leq \frac{1}{n^2}$, *which implies that* $\inf_{u. \in \mathcal{U}_{ad}} J(u.) = 0$.

On the other hand, for any $u. \in \mathcal{U}_{ad}$,

$$J(u) = \int_0^1 X_t^2 dt \geq 0.$$

Consequently, we get $J(\widetilde{u}) = 0$.

If this would have been the case, then for every $t, \widetilde{X}_t = 0$. *This implies that* $\widetilde{u}_t = 0$, *which is impossible. The problem is that the sequence* $U^n_\cdot$ *has no limit in the space of strict controls. This limit, if it exists, will be the naturale candidate for optimality. As a matter of fact, let* δ_u *denote the atomic measure concentrated at a single point* $u.$, *and we identify the sequence* U^n_t *with* $\delta_{u^n_t}(du)$, *we get*

$$dt\delta_{u^n_t}(du) \to \frac{1}{2}dt(\delta_{-1} + \delta_1)du = q_t(du), t \in [0, 1],$$

which means that $q_t(du)$ *is an optimal relaxed control. This suggests that the set of strict controls is too narrow and should be embedded into the class of relaxed control.*

Let $P(U)$ denote the space of probability measures on $\mathcal{B}(U)$ equipped with the topology of weak convergence, where U is a nonempty Borel compact subset of $\mathbb{R}^r$, then $P(U)$ is also compact metrizable space. In a relaxed control problem, the U-valued process v_t is replaced by an $P(U)$-valued process q_t.

If $\chi : U \to \mathbb{R}$ is a bounded measurable function, then we extend χ to $P(U)$ by letting $\chi(q) := \int_U \chi(u)q(du)$.

Moreover, if $q_t(du) = \delta_{v_t}(du)$ is a Dirac measure charging v_t for each t, then we get a strict control problem as a special case of the relaxed one, with the following property: For any bounded and uniformly continuous function $\varrho(t, x, u)$ defined on $[0, T] \times \mathbb{R}^n \times U$,

$$\varrho(t, x, v_t) = \int_U \varrho(t, x, u)\delta_{v.}(t, du) := \widehat{\varrho}(t, x, \delta_{v.}).$$

Denote by $\mathbb{V}$ the set of Radon measures $\mu.$ on $\mathcal{B}(U \times [0, T])$ such that $\mu(U \times \Upsilon) = Leb(\Upsilon)$ for all $\Upsilon \in \mathcal{B}([0, T])$, where Leb is the Lebesgue measure on $[0, T]$. Elements of $\mathbb{V}$ are called Young measures on U in deterministic theory. $\mathbb{V}$ is endowed with the stable topology, for which the mappings:

$$\mathbb{V} \ni \mu \mapsto \int_{\Upsilon} \int_{U} h(u)\mu(du, ds) \in \mathbb{R},$$

are continuous, $\forall \Upsilon \in \mathcal{B}([0,T])$ and $h \in C_b(U)$.

It is clear that, any element $\mu \in \mathbb{V}$ can be decomposed as

$$\mu(du, dt) = \mu_t(du)dt,$$

where $\mu_t(du) \in P(U), t \in [0,T]$, see [18]. Equipped with the topology of stable convergence of measures, $\mathbb{V}$ is a compact metrizable space, (see Jacod and Mémin [33]).

The system in this case, is then governed by the following FBDSDE

$$\begin{cases} X_t = x + \int_0^t \int_U B\left(s, X_s, u\right) q_s\left(du\right) ds + \int_0^t \Sigma\left(s, X_s\right) dW_s, \\ \\ Y_t = H\left(X_T\right) + \int_t^T \int_U F\left(s, X_s, Y_s, u\right) q_s\left(du\right) ds + \int_t^T G\left(s, X_s, Y_s\right) d\overleftarrow{B_s} \\ \\ \qquad - \int_t^T Z_s dW_s - \left(M_T - M_t\right), \end{cases} \tag{3.9}$$

where $M.$ is a square integrable martingale which is orthogonal to W. With $M_0 = 0$ and with càdlàg trajectories.

Definition 3.2.4 *A 7-tuple* $\varpi. := \left(\Omega, \mathcal{F}, \left(\mathcal{F}_t\right)_{0\le t\le T}, P, W., B., q.\right)$ *is called admissible relaxed control, and* (X_t, Y_t, Z_t, M_t) *an admissible quadruple if:*

i)-$(\Omega, \mathcal{F}, P)$ *is a probability space;*

ii)-W_t *and* B_t *be two mutually independent standard Brownian motion defined on* $\left(\Omega, \mathcal{F}, \left(\mathcal{F}_t\right)_{0\le t\le T}, P\right)$;

iii)-q_t *is* $\mathcal{F}_t$*-progressively measurable and such that for each* t, $1_{]0,t]} \cdot q$ *is* $\mathcal{F}_t$*- measurable, taking values in* $\mathbb{V}$*;*

iv)-(X_t, Y_t, Z_t, M_t) is the solution of the FBDSDE (3.9) on $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, P\right)$ under q_t.

v)-M_t is a square integrable martingale, orthogonal to Brownian motion W_t.

The set of all admissible relaxed controls is denoted by $\mathcal{R}_{ad}$.

The cost to be minimized, over the set $\mathcal{R}_{ad}$ of relaxed controls, has the form

$$J(q.) := \mathbb{E}\big[\varphi(X_T) + \psi(Y_0) + \int_0^T \int_U L(s, X_s, Y_s, u)\, q_s(du)\, ds\big].$$

A relaxed control $\varpi_.^*$ is called optimal if it satisfy

$$J(\varpi_.^*) = \inf_{\varpi. \in \mathcal{R}_{ad}} J(\varpi.)\,. \tag{3.10}$$

Let us denote by $\mathbb{D}(0, T; \mathbb{R}^m)$: the Skorokhod space of càdlàg functions from $[0, T]$ to $\mathbb{R}^m$, that is functions which are continuous from the right with left hand limits, equipped with the S-topology of Jakubowski (see [31]),

$\mathcal{C}(0, T, \mathbb{R}^n)$: the space of continuous functions from $[0, T]$ into $\mathbb{R}^n$, endowed with the topology of uniform convergence,

$\mathcal{M}^2_{\mathcal{F}_t}\left(0, T, \mathbb{R}^{m\times d}\right)$: the space of $(\mathcal{F}_t)_{0\leq t\leq T}$-measurable processes Y_t defined from $[0, T] \times \Omega$ into $\mathbb{R}^{m\times d}$, such that

$$\mathbb{E}\big[\int_0^T |Y_t|^2\, dt\big] < \infty,$$

and $\mathcal{S}^2_{\mathcal{F}_t}(0, T, \mathbb{R}^n)$: the space of $(\mathcal{F}_t)_{0\leq t\leq T}$-measurable processes X_t defined from $[0, T] \times \Omega$ into $\mathbb{R}^n$, such that

$$\mathbb{E}\big[\sup_{0\leq t\leq T} |X_t|^2\big] < \infty.$$

We shall consider in the first part of this section the following assumptions.

(**H5**). Assume that,

$$B : [0, T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}^m,$$

$$\Sigma : [0, T] \times \mathbb{R}^n \longrightarrow \mathbb{R}^{n \times d},$$

$$F : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times U \longrightarrow \mathbb{R}^m,$$

$$G : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^{m \times k},$$

$$H : \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$

are bounded and continuous. Moreover, assume that there exist a constant $K > 0$, such that for every $(x, x') \in \mathbb{R}^{2n}, (y, y') \, \mathbb{R}^{2m}$,

$$|F (t, x, y, u) - F (t, x', y', u)| \leq K \ (|x - x'| + |y - y'|) ,$$

$$|G (t, x, y) - G (t, x', y')| \leq K \ (|x - x'| + |y - y'|) ,$$

$$|B (t, x, u) - B (t, x', u)| \leq K \ |x - x'| ,$$

$$|\Sigma (t, x) - \Sigma (t, x')| \leq K \ |x - x'| .$$

(**H6**). Assume that the functions

$$L : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times U \to \mathbb{R},$$

$$\varphi : \mathbb{R}^n \to \mathbb{R},$$

$$\psi : \mathbb{R}^m \to \mathbb{R},$$

are bounded and continuous and there exist a constant $K > 0$, such that for every $(x, x', y, y') \in \mathbb{R}^{2n} \times \mathbb{R}^{2m}$,

$$|L (t, x, y, u) - L (t, x', y', u)| \leq K \ (|x - x'| + |y - y'|) .$$

3.2.2 Existence of optimal relaxed controls

The first main result in this paper is to prove existence of optimal relaxed control.

Theorem 3.2.5 *If the assumptions* $(\mathbf{H5})$ *and* $(\mathbf{H6})$ *hold. Then, the relaxed control problem* $\{(3.9),(3.10)\}$ *has an optimal solution.*

To prove Theorem 3.2.5, we need some auxiliary results on the tightness of the processes under consideration.

Let $\varpi_{\cdot}^{n} := (\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{0\leq t\leq T}, P^n, W_{\cdot}^n, B_{\cdot}^n, q_{\cdot}^n)$ be a minimizing sequence, that is

$$\lim_{n\longrightarrow\infty} J(\varpi_{\cdot}^{n}) = \inf_{\varpi_{\cdot}\in\mathcal{R}_{ad}} J(\varpi_{\cdot}).$$

Let $(X_{\cdot}^n, Y_{\cdot}^n, Z_{\cdot}^n)$ be the unique strong solution of the following FBDSDE, corresponding to $\varpi_{\cdot}^{n}$,

$$\begin{cases} X_t^n = x + \int_0^t \int_U B(s, X_s^n, u)\, q_s^n(du)\, ds + \int_0^t \Sigma(s, X_s^n)\, dW_s^n, \\ Y_t^n = H(X_T^n) + \int_t^T \int_U F(s, X_s^n, Y_s^n, u)\, q_s^n(du)\, ds + \int_t^T G(s, X_s^n, Y_s^n)\, d\overleftarrow{B_s^n} \\ \qquad - \int_t^T Z_s^n dW_s^n. \end{cases} \tag{3.11}$$

Lemma 3.2.6 *The family of distributions associated to the relaxed controls* $(q_{\cdot}^n)_{n\geq 0}$ *is tight in* $\mathbb{V}$.

Proof. From the fact that $[0,T]\times U$ is compact, then by Prokhorov's theorem, we get that the family of distributions associated to $q_{\cdot}^n$ $n\geq 0$ which valued in the compact space $\mathbb{V}$, is tight. ∎

Lemma 3.2.7 *Let* X_t^n *be the forward component of (3.11). Then, the sequence of processes* $(X_{\cdot}^n, B_{\cdot}^n, W_{\cdot}^n)$ *is tight on the space* $\mathcal{C}(0,T,\mathbb{R}^{n+k+d})$ *endowed with the topology of uniform convergence.*

Proof. By standard arguments, one can get easily that the Kolmogorov tightness criteria (see Ikeda and Watanabe [30] page 18), is fulfilled by $(X^n_\cdot, B^n_\cdot, W^n_\cdot)$, i.e., there exist a constant K such that

$$\mathbb{E}^n\left[\,|X^n_t - X^n_s|^4\,\right] \leq K\,|t-s|^2\,, \forall t, s \in [0,T],$$

where $\mathbb{E}^n$ is the expectation under P^n, and by the same method for $(B^n_\cdot, W^n_\cdot)$. Then the sequence of processes $(X^n_\cdot, B^n_\cdot, W^n_\cdot)$ is tight. ■

Lemma 3.2.8 *Let $(X^n_\cdot, Y^n_\cdot, Z^n_\cdot)$ be the unique solution of (3.11). There exists a positive constant $\beta \geq 0$ such that*

$$\sup_n \mathbb{E}^n\big[\sup_{0\leq t\leq T} |X^n_t|^2 + \sup_{0\leq t\leq T} |Y^n_t|^2 + \int_0^T \|Z^n_t\|^2\,dt\big] \leq \beta. \tag{3.12}$$

Proof. For every $0 \leq t \leq T$ and $n > 1$, we want to show that

$$\mathbb{E}^n\big[\sup_{0\leq t\leq T} |X^n_t|^2\big] < \infty.$$

We have

$$|X^n_t|^2 = \left|x + \textstyle\int_0^t \int_U B\left(s, X^n_s, u\right) q^n_s\left(du\right) ds + \int_0^t \Sigma\left(s, X^n_s\right) dW^n_s\right|^2.$$

Applying the inequality $(a+b+c) \leq 3\left(a^2+b^2+c^2\right)$, we obtain

$$|X^n_t|^2 \leq 3\left(x^2 + \left|\textstyle\int_0^t \int_U B\left(s, X^n_s, u\right) q^n_s\left(du\right) ds\right|^2 + \left|\int_0^t \Sigma\left(s, X^n_s\right) dW^n_s\right|^2\right).$$

Passing to the expectation, we get

$$\mathbb{E}\left[|X^n_t|^2\right] \leq 3\left(\mathbb{E}\left[x^2\right] + \mathbb{E}\left[\left|\textstyle\int_0^t \int_U B\left(s, X^n_s, u\right) q^n_s\left(du\right) ds\right|^2\right] + \mathbb{E}\left[\left|\int_0^t \Sigma\left(s, X^n_s\right) dW^n_s\right|^2\right]\right).$$

By applying Holder inequality we have

$$\mathbb{E}\left[\left|\textstyle\int_0^t \int_U B\left(s, X^n_s, u\right) q^n_s\left(du\right) ds\right|^2\right] \leq T \times \mathbb{E}\left[\int_0^t \int_U \left|B\left(s, X^n_s, u\right)\right|^2 q^n_s\left(d\alpha\right) ds\right].$$

By the Burkholder–Davis–Gundy and Holder inequality provid that is inequality

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t \Sigma\left(s,X_s^n\right)dW_s^n\right|^2\right]\leq C\mathbb{E}\left[\int_0^T\left|\Sigma\left(s,X_s^n\right)\right|^2 ds\right].$$

Hence, by using assumption ($\mathbf{H1}$) with a linear growth, we get

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t^n\right|^2\right]\leq C\left(\left|x\right|^2+1\right)+C\int_0^T\int_U\mathbb{E}\left[\sup_{0\leq s\leq T}\left|X_s^n\right|^2\right]q_s^n\left(d\alpha\right)ds.$$

So by applying Gronwall's lemma, we have

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t^n\right|^2\right]\leq C\left(\left|x\right|^2+1\right)\exp\left(CT\right)<\infty.$$

We now pass to prove the second inequality of Y_t^n. It then follows by using the generalized Itô's formula (See Pardoux and Peng [[39], Lemma 1.3, page 213],) we get

$$\begin{aligned}
&\mathbb{E}^n\big[\left|Y_t^n\right|^2\big]+\mathbb{E}^n\big[\int_t^T\left\|Z_s^n\right\|^2ds\big]\\
&\qquad\leq\mathbb{E}^n\big[\left|H\left(X_T^n\right)\right|^2\big]+2\mathbb{E}^n\big[\int_t^T\int_U\left|\left\langle Y_s^n,F\left(s,X_s^n,Y_s^n,u\right)\right\rangle\right|q_s^n\left(du\right)ds\big]\\
&\qquad\quad+\mathbb{E}^n\big[\int_t^T\left|G\left(s,X_s^n,Y_s^n\right)\right|^2ds\big]\\
&\qquad\leq\mathbb{E}^n\big[\left|H\left(X_T^n\right)\right|^2\big]+\mathbb{E}^n\big[\int_t^T\left|Y_s^n\right|^2ds\big]\\
&\qquad\quad+\mathbb{E}^n\big[\int_t^T\int_U\left|F\left(s,X_s^n,Y_s^n,u\right)\right|^2q_s^n\left(du\right)ds\big]+\mathbb{E}^n\big[\int_t^T\left|G\left(s,X_s^n,Y_s^n\right)\right|^2ds\big].
\end{aligned}$$

So by using ($\mathbf{H1}$), applying Gronwall's lemma and the Burkholder-Davis-Gundy inequality, we get

$$\sup_n\mathbb{E}^n\big[\sup_{0\leq t\leq T}\left|Y_t^n\right|^2+\int_0^T\left\|Z_t^n\right\|^2dt\big]<\infty.$$

■

Lemma 3.2.9 *Let $(X_\cdot^n,Y_\cdot^n,Z_\cdot^n)$ be the unique solution of (3.11). The sequence of processes $\left(Y_\cdot^n,\int_0^\cdot Z_s^n dW_s^n\right)$ is tight on the space $\mathbb{D}(0,T,\mathbb{R}^{m+m+d})$ endowed with the S-topology of Jakubowski.*

Proof. To prove this Lemma, we borrow the idea from the method using in the paper of Tang and Wu, [46]. Therefore, to prove that the sequence $\left(Y_\cdot^n, \mathcal{N}_t^n = \int_0^\cdot Z_s^n dW_s^n\right)$ satisfies the Meyer-Zheng tightness criteria, let us defined on the space $\mathbb{D}(0,T,\mathbb{R}^{m+m+d})$, the filtration $\mathcal{G}_t^n$ by

$$\mathcal{G}_t^n = \mathcal{F}_t^{X^n, Y^n, q^n, W^n} \vee \mathcal{F}_T^{B^n}.$$

Given a subdivision $0 = t_0 < t_1 < ... < t_n = T$. We define the conditional variation by

$$VC\left(Y_\cdot^n\right) := \sup \mathbb{E}^n \big[\sum_{k=0}^{n-1} \Big| \mathbb{E}^n \left(Y_{t_{k+1}}^n - Y_{t_k}^n \mid \mathcal{G}_{t_k}^n \right) \Big| \big],$$

where the supremum is taken over all partitions of the interval $[0,T]$. Since $\mathcal{N}_t^n$ is $\mathcal{G}_t^n$-martingale and $\int_t^T G\left(s, X_s^n, Y_s^n\right) d\overleftarrow{B_s^n}$ is $\mathcal{G}_t^n$-measurable, we obtain

$$VC\left(Y_\cdot^n\right) \leq \mathbb{E}^n \big[\int_0^T \int_U \left| F\left(s, X_s^n, Y_s^n, u\right) \right| q_s^n\left(du\right) ds \big].$$

Hence combining assumptions (**H1**) and Lemma 3.2.8, we have

$$\sup_n \left(VC\left(Y^n\right) + \sup_{0 \leq t \leq T} \mathbb{E}^n \big[\left| Y_t^n \right| \big] + \sup_{0 \leq t \leq T} \mathbb{E}^n \big[\left| \mathcal{N}_t^n \right| \big] \right) < \infty.$$

Thus the Meyer-Zheng tightness criteria is fulfilled (see [43]), then the sequence $(Y_\cdot^n, \mathcal{N}_\cdot^n)$ is tight in $\mathbb{D}(0,T,\mathbb{R}^{m+m+d})$ equipped with the Jakubowski S-topology. ■

3.2.2.1 Proof of existence of optimal relaxed control (Theorem 3.2.5)

Let $\varpi_\cdot^n := (\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{0 \leq t \leq T}, P^n, W_\cdot^n, B_\cdot^n, q_\cdot^n)$ be a minimizing sequence, and $(X_\cdot^n, Y_\cdot^n, Z_\cdot^n)$ be the unique solution of the FBDSDE (3.11).

From Lemma 3.2.6, Lemma 3.2.7 and Lemma 3.2.9, it follows that the sequence of processes $\vartheta^n_\cdot = (X^n_\cdot, B^n_\cdot, W^n_\cdot, q^n_\cdot, Y^n_\cdot, \mathcal{N}^n_\cdot)$ is tight on the space $\Pi = \mathcal{C}(0,T,\mathbb{R}^{n+k+d}) \times \mathbb{V} \times \mathbb{D}(0,T,\mathbb{R}^{m+m+d})$. By Skorokhod's representation theorem, one can choose a subsequence (still labeled as $\vartheta^n_\cdot$) and there exists a sequence $\widehat{\vartheta}^n_\cdot = \left(\widehat{X}^n_\cdot, \widehat{B}^n_\cdot, \widehat{W}^n_\cdot, \widehat{q}^n_\cdot, \widehat{Y}^n_\cdot, \widehat{\mathcal{M}}^n_\cdot\right)$, $n \geq 0$ and $\widehat{\vartheta}_\cdot = \left(\widehat{X}_\cdot, \widehat{B}_\cdot, \widehat{W}_\cdot, \widehat{q}_\cdot, \widehat{Y}_\cdot, \widehat{\mathcal{M}}_\cdot\right)$ defined on a suitable probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P})$, such that

$$\text{law}\,(\vartheta^n) = \text{law}(\widehat{\vartheta}^n), \forall n \geq 1, \tag{3.13}$$

$$\begin{aligned} &\text{there exists a subsequence } (\widehat{\vartheta}^{n_k}) \text{ of } (\widehat{\vartheta}^n),\ \text{still denoted } (\widehat{\vartheta}^n), \\ &\text{which converges to } \widehat{\vartheta}, \widehat{P} - a.s.,\ \text{on the space } \Pi, \end{aligned} \tag{3.14}$$

$$\begin{aligned} &(\widehat{Y}^n_\cdot, \widehat{\mathcal{M}}^n_\cdot = \textstyle\int_0^\cdot \widehat{Z}^n_s d\widehat{W}^n_s) \text{ converges to the càdlàg processes } (\widehat{Y}_\cdot, \widehat{\mathcal{M}}_\cdot), \\ &dt \times \widehat{P} - a.s,\ \text{and } (\widehat{Y}^n_T, \widehat{\mathcal{M}}^n_T) \text{ converges to } \ (\widehat{Y}_T, \widehat{\mathcal{M}}_T), \widehat{P} - a.s.,\ \text{as } n \to \infty, \end{aligned} \tag{3.15}$$

$$\sup_{0 \leq t \leq T} \left| \widehat{X}^n_t - \widehat{X}_t \right| \longrightarrow 0, \widehat{P} - a.s,\ \text{as } n \to \infty, \tag{3.16}$$

and

$$\widehat{q}^n_\cdot \text{ converges weakly to } \widehat{q}_\cdot, \widehat{P} - a.s,\ \text{as } n \to \infty \text{ on } \mathbb{V}. \tag{3.17}$$

Set that

$$\begin{cases} \widehat{\mathcal{F}}^n_t := \mathcal{F}_t^{\widehat{X}^n, \widehat{Y}^n, \widehat{q}^n, \widehat{W}^n} \vee \mathcal{F}^{\widehat{B}^n}_{t,T}, \\ \\ \widehat{\mathcal{F}}_t := \mathcal{F}_t^{\widehat{X}, \widehat{Y}, \widehat{q}, \widehat{W}} \vee \mathcal{F}^{\widehat{B}}_{t,T}, \end{cases}$$

where $\widehat{\mathcal{F}}^n_t$ is the $\sigma-$field generated by

$$\widehat{X}^n_{t_1}, \cdots, \widehat{X}^n_{t_\ell}, \widehat{Y}^n_{t_1}, \cdots, \widehat{Y}^n_{t_\ell}, \widehat{q}^n_{t_1}, \cdots, \widehat{q}^n_{t_\ell}, \widehat{W}^n_{t_1}, \cdots, \widehat{W}^n_{t_\ell}, \widehat{B}^n_r - \widehat{B}^n_{t_1}, \cdots, \widehat{B}^n_r - \widehat{B}^n_{t_j},$$

where $0 \leq t_1 \leq t_2 \leq \cdots \leq t_\ell \leq t \leq T$ and $t_\ell \leq r \leq t, \ell = 1, 2, \cdots$. A similar statement can be made for $\widehat{\mathcal{F}}_t$.

We need to show that $\widehat{W}^n_\cdot$ and $\widehat{B}^n_\cdot$ are two $\widehat{\mathcal{F}}^n_t-$Brownian motions. Note that $W^n_\cdot$ and $B^n_\cdot$ are two $\mathcal{F}^{X^n,Y^n,q^n,W^n}_t \vee \mathcal{F}^{B^n}_{t,T}-$Brownian motions. Thus for any $0 \le s \le t \le T$ and for any bounded continuous function Θ_s on $\mathcal{C}\left(0,s;\mathbb{R}^n\right) \times \mathbb{D}\left(0,s;\mathbb{R}^m\right) \times \mathbb{V}^s \times \mathcal{C}\left(0,s;\mathbb{R}^{d+k}\right)$, we have

$$\begin{aligned}
&\mathbb{E}^n\big[\Theta_s(X^n_\cdot,Y^n_\cdot,q^n_\cdot,W^n_\cdot,B^n_\cdot)\cdot(W^n_t-W^n_s)\big]=0,\\
&\mathbb{E}^n\big[\Theta_s(X^n_\cdot,Y^n_\cdot,q^n_\cdot,W^n_\cdot,B^n_\cdot)\cdot\big((W^n_t-W^n_s)(W^n_t-W^n_s)^T\big]\big]=(t-s)I,
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}^n\big[\Theta_s(X^n_\cdot,Y^n_\cdot,q^n_\cdot,W^n_\cdot,B^n_\cdot)\cdot(B^n_t-B^n_s)\big]=0,\\
&\mathbb{E}^n\big[\Theta_s(X^n_\cdot,Y^n_\cdot,q^n_\cdot,W^n_\cdot,B^n_\cdot)\cdot\big((B^n_t-B^n_s)(B^n_t-B^n_s)^T\big)\big]=(t-s)I,
\end{aligned}$$

where $\mathbb{V}^s$ denote the restriction of probability measures to the set $[0,s]\times U$.

In view of (3.13), we obtain

$$\begin{aligned}
&\widehat{\mathbb{E}}^n\big[\Theta_s(\widehat{X}^n_\cdot,\widehat{Y}^n_\cdot,\widehat{q}^n_\cdot,\widehat{W}^n_\cdot,\widehat{B}^n_\cdot)\cdot(\widehat{W}^n_t-\widehat{W}^n_s)\big]=0,\\
&\widehat{\mathbb{E}}^n\big[\Theta_s(\widehat{X}^n_\cdot,\widehat{Y}^n_\cdot,\widehat{q}^n_\cdot,\widehat{W}^n_\cdot,\widehat{B}^n_\cdot)\cdot\big((\widehat{W}^n_t-\widehat{W}^n_s)(\widehat{W}^n_t-\widehat{W}^n_s)^T\big)\big]=(t-s)I,
\end{aligned}\tag{3.18}$$

and

$$\begin{aligned}
&\widehat{\mathbb{E}}^n\big[\Theta_s(\widehat{X}^n_\cdot,\widehat{Y}^n_\cdot,\widehat{q}^n_\cdot,\widehat{W}^n_\cdot,\widehat{B}^n_\cdot)\cdot(\widehat{B}^n_t-\widehat{B}^n_s)\big]=0,\\
&\widehat{\mathbb{E}}^n\big[\Theta_s(\widehat{X}^n_\cdot,\widehat{Y}^n_\cdot,\widehat{q}^n_\cdot,\widehat{W}^n_\cdot,\widehat{B}^n_\cdot)\cdot\big((\widehat{B}^n_t-\widehat{B}^n_s)(\widehat{B}^n_t-\widehat{B}^n_s)^T\big)\big]=(t-s)I,
\end{aligned}\tag{3.19}$$

where $\widehat{\mathbb{E}}^n$ is the expectation under $\widehat{P}^n$.

According to (3.13), we have the following FBDSDE on the space $(\widehat{\Omega},\widehat{\mathcal{F}},(\widehat{\mathcal{F}}^n_t)_{0\le t\le T},\widehat{P})$,

$$\begin{cases}
\widehat{X}^n_t = x+\int_0^t\int_U B(s,\widehat{X}^n_s,\widehat{Y}^n_s,u)\widehat{q}^n_s(du)ds+\int_0^t\Sigma(s,\widehat{X}^n_s)d\widehat{W}^n_s,\\
\\
\widehat{Y}^n_t = H(\widehat{X}^n_T)+\int_t^T\int_U F(s,\widehat{X}^n_s,\widehat{Y}^n_s,u)\widehat{q}^n_s(du)ds+\int_t^T G(s,\widehat{X}^n_s,\widehat{Y}^n_s)d\overleftarrow{\widehat{B}^n_s}\\
\\
\qquad -(\widehat{\mathcal{M}}^n_T-\widehat{\mathcal{M}}^n_t),
\end{cases}\tag{3.20}$$

where $\widehat{\mathcal{M}}^n_t=\int_0^t\widehat{Z}^n_s d\widehat{W}^n_s$.

Using (3.13)-(3.17), assumption ($\mathbf{H2}$) and passing to the limit in the FBDSDE (3.20), one can show that there exists a countable set $Q \subset [0,T)$ such that

$$\begin{cases} \widehat{X}_t = x + \int_0^t \int_U B(s, \widehat{X}_s, \widehat{Y}_s, u)\widehat{q}_s(du)ds + \int_0^t \Sigma(s, \widehat{X}_s)d\widehat{W}_s, t \geq 0, \\ \\ \widehat{Y}_t = H(\widehat{X}_T) + \int_t^T \int_U F(s, \widehat{X}_s, \widehat{Y}_s, u)\widehat{q}_s(du)ds + \int_t^T G(s, \widehat{X}_s, \widehat{Y}_s)d\overleftarrow{\widehat{B}_s} \\ \\ \qquad -(\widehat{\mathcal{M}}_T - \widehat{\mathcal{M}}_t), t \in [0,T] \diagdown Q. \end{cases} \tag{3.21}$$

Since $\widehat{Y}$ et $\widehat{\mathcal{M}}$ are continuous, then are càdlàg, it follows that the doubly backward component of (3.21) is satisfied for every $t \in [0,T]$, i.e.,

$$\widehat{Y}_t = H(\widehat{X}_T) + \int_t^T \int_U F(s, \widehat{X}_s, \widehat{Y}_s, u)\widehat{q}_s(du)ds + \int_t^T G(s, \widehat{X}_s, \widehat{Y}_s)d\overleftarrow{\widehat{B}_s} \\ -(\widehat{\mathcal{M}}_T - \widehat{\mathcal{M}}_t). \tag{3.22}$$

Now, let us defined the filtration $\widehat{\mathcal{G}}_t = \mathcal{F}_t^{\widehat{X},\widehat{Y},\widehat{q},\widehat{W}} \vee \mathcal{F}_T^{\widehat{B}}$.

We need to show that $\widehat{\mathcal{M}}_t$ is $\widehat{\mathcal{G}}_t$-martingale.

In this end, for any $s,t : \ 0 \leq s \leq t \leq T$ and Θ_s a bounded continuous function defined on $\mathcal{C}(0,s;\mathbb{R}^n) \times \mathbb{D}(0,s;\mathbb{R}^m) \times \mathbb{V}^s \times \mathcal{C}\left(0,s;\mathbb{R}^{d+k}\right)$, and from the fact that $\mathcal{N}_t^n$ is $\mathcal{G}_t^n := \mathcal{F}_t^{X^n,Y^n,q^n,W^n} \vee \mathcal{F}_T^{B^n}$-martingale we have

$$\mathbb{E}^n\big[\Theta_s\left(X_\cdot^n, Y_\cdot^n, q_\cdot^n, W_\cdot^n, B_\cdot^n\right)\big(\int_0^\varepsilon (\mathcal{N}_{t+\rho}^n - \mathcal{N}_{s+\rho}^n)d\rho)\big] = 0,$$

and from (3.13) we get

$$\widehat{\mathbb{E}}^n\big[\Theta_s\left(\widehat{X}_\cdot^n, \widehat{Y}_\cdot^n, \widehat{q}_\cdot^n, \widehat{W}_\cdot^n, \widehat{B}_\cdot^n\right)\big(\int_0^\varepsilon (\widehat{\mathcal{M}}_{t+\rho}^n - \widehat{\mathcal{M}}_{s+\rho}^n)d\rho)\big] = 0.$$

Since $\left(\widehat{X}^n_\cdot, \widehat{Y}^n_\cdot, \widehat{q}^n_\cdot, \widehat{W}^n_\cdot, \widehat{B}^n_\cdot\right)$ converges weakly to $\left(\widehat{X}_\cdot, \widehat{Y}_\cdot, \widehat{q}_\cdot, \widehat{W}_\cdot, \widehat{B}_\cdot\right)$ and $\widehat{\mathbb{E}}^n\left[\sup_{0\le t\le T}\left|\widehat{\mathcal{M}}^n_t\right|^2\right] < \infty$, we have

$$\widehat{\mathbb{E}}\left[\Theta_s\left(\widehat{X}_\cdot, \widehat{Y}_\cdot, \widehat{q}_\cdot, \widehat{W}_\cdot, \widehat{B}_\cdot\right)\left(\int_0^\varepsilon (\widehat{\mathcal{M}}_{t+\rho} - \widehat{\mathcal{M}}_{s+\rho})d\rho\right)\right] = 0,$$

where $\widehat{\mathbb{E}}$ is the expectation under $\widehat{P}$, dividing by ε, sending it to 0 and using the fact that $\widehat{\mathcal{M}}_t$ is continuous from the right, we obtain

$$\widehat{\mathbb{E}}\left[\Theta_s\left(\widehat{X}_\cdot, \widehat{Y}_\cdot, \widehat{q}_\cdot, \widehat{W}_\cdot, \widehat{B}_\cdot\right)(\widehat{\mathcal{M}}_t - \widehat{\mathcal{M}}_s)\right] = 0,$$

which gives $\widehat{\mathcal{M}}_t$ is $\widehat{\mathcal{G}}_t$-martingale.

Now, from the martingale decomposition theorem, there exist a process $\widehat{Z} \in \mathcal{M}^2_{\widehat{\mathcal{G}}_t}\left(0, T, \mathbb{R}^{m\times d}\right)$ such that

$$\widehat{\mathcal{M}}_t = \int_0^t \widehat{Z}_s d\widehat{W}_s + \widehat{\aleph}_t, \text{ with } \langle \widehat{W}, \widehat{\aleph}\rangle_t = 0,$$

then (3.22) becomes

$$\begin{aligned} \widehat{Y}_t = H(\widehat{X}_T) + \textstyle\int_t^T \int_U F(s, \widehat{X}_s, \widehat{Y}_s, u)\widehat{q}_s(du)ds + \int_t^T G(s, \widehat{X}_s, \widehat{Y}_s)d\overleftarrow{\widehat{B}}_s \\ - \textstyle\int_t^T \widehat{Z}_s d\widehat{W}_s - (\widehat{\aleph}_T - \widehat{\aleph}_t), \end{aligned} \tag{3.23}$$

and by the same technique using in, [39], one can prove that $\widehat{Y}_\cdot$, $\widehat{Z}_\cdot$ and $\widehat{\aleph}_\cdot$ are in fact $\widehat{\mathcal{F}}_t$-measurable.

Finally, it remains check that $(\widehat{q}_\cdot)$ minimize the cost functional J over the set $\mathcal{R}_{ad}$.

According to properties (3.13)-(3.17) and assumption (**H6**), we have

$$
\begin{aligned}
\inf_{\varpi.\in\mathcal{R}_{ad}} J(\varpi.) &= \lim_{n\to\infty} J(\varpi.^n) = \lim_{n\to\infty} J(\widehat{\varpi}.^n),\\
&= \lim_{n\to\infty}\mathbb{E}\big[\varphi(X_T^n)+\psi(Y_0^n)+\int_0^T\int_U L(t,X_t^n,Y_t^n,u)\,q_t^n(du)\,dt\big],\\
&= \lim_{n\to\infty}\widehat{\mathbb{E}}\big[\varphi(\widehat{X}_T^n)+\psi(\widehat{Y}_0^n)+\int_0^T\int_U L(t,\widehat{X}_t^n,\widehat{Y}_t^n,u)\widehat{q}_t^n(du)dt\big],\\
&= \widehat{\mathbb{E}}\big[\varphi(\widehat{X}_T)+\psi(\widehat{Y}_0)+\int_0^T\int_U L(t,\widehat{X}_t,\widehat{Y}_t,u)\widehat{q}_t(du)dt\big],\\
&= J(\widehat{\varpi}.).
\end{aligned}
$$

Thus $\widehat{\varpi}. := (\widehat{\Omega},\widehat{\mathcal{F}},(\widehat{\mathcal{F}}_t)_{0\le t\le T},\widehat{P},\widehat{W}.,\widehat{B}.,\widehat{q}.)$ is optimal relaxed control, then theorem 3.2.5 is proved.

3.2.3 Existence of optimal strict control

The second main result in this section is to prove existence of optimal strict control. In this end we need to assume the Roxin's condition as follows:

(**H7**) : (Roxin-type convexity condition): The set

$$
\begin{aligned}
(B,F,L)(t,x,y,U) := \{&B_i(t,x,u), F_j(t,x,y,u)\\
&, L(t,x,y,u)\diagdown u\in U, i=1,\cdots,n, j=1,\cdots,m\},
\end{aligned}
$$

is convex and closed in $\mathbb{R}^{n+m+1}$.

Theorem 3.2.10 *Under* (**H5**)-(**H7**) *. The optimal relaxed control* $\widehat{q}_t$ *has the form*

$$
\widehat{q}_t(du) = \delta_{\widehat{u}_t}(du),
$$

where $\delta_{\widehat{u}_t}$ *is the Dirac measure charging a strict control* $\widehat{u}_t$*, which is optimal.*

Proof. We put

$$\int_U F\left(s,\widehat{X}_s,\widehat{Y}_s,u\right)\widehat{q}_s(du)=\widehat{F}(s,\omega)\in F(s,x,y,U)$$

$$\int_U B\left(s,\widehat{X}_s,u\right)\widehat{q}_s(du)=\widehat{B}(s,\omega)\in B(s,x,U)$$

$$\int_U L\left(s,\widehat{X}_s,\widehat{Y}_s,u\right)\widehat{q}_s(du)=\widehat{L}(s,\omega)\in L(s,x,y,U).$$

From (**H7**) and the measurable selection theorem (see Li-Yong [41] p. 102, Corollary 2.26), there is a U-valued and $\widehat{\mathcal{F}}_t$-measurable process $\widehat{u}.$, such that for every $s\in[0,T]$ and $\omega\in\widehat{\Omega}$, we have

$$\left(\widehat{F},\widehat{L}\right)(s,\omega)=(F,L)\left(s,\widehat{X}_s(\omega),\widehat{Y}_s(\omega),\widehat{u}_s(\omega)\right),$$

and for every $t\geq 0$,

$$B(t,\omega)=B\left(t,\widehat{X}_t(t),\widehat{u}_t(\omega)\right).$$

Therefore, for every $s\in[0,T], t\geq 0$, and $\omega\in\widehat{\Omega}$, we get

$$\int_U F\left(s,\widehat{X}_s,\widehat{Y}_s,u\right)\widehat{q}_s(du)=F\left(s,\widehat{X}_s(\omega),\widehat{Y}_s(\omega),\widehat{u}_s(\omega)\right)$$

$$\int_U L\left(s,\widehat{X}_s,\widehat{Y}_s,u\right)\widehat{q}_s(du)=L\left(s,\widehat{X}_s(\omega),\widehat{Y}_s(\omega),\widehat{u}_s(\omega)\right),$$

and

$$\int_U B\left(t,\widehat{X}_t,u\right)\widehat{q}_t(du)=B\left(t,\widehat{X}_t(\omega),\widehat{u}_t(\omega)\right).$$

Then the processes $\left(\widehat{X}_t,\widehat{Y}_t,\widehat{Z}_t\right)$ satisfies, for each $t\in[0,T]$ the following FBDSDE

$$\begin{cases}\widehat{X}_t=x+\int_0^t B\left(s,\widehat{X}_s,\widehat{Y}_s,\widehat{u}_t\right)ds+\int_0^t\Sigma\left(s,\widehat{X}_s\right)d\widehat{W}_s,\\ \widehat{Y}_t=H\left(\widehat{X}_T\right)+\int_t^T F\left(s,\widehat{X}_s,\widehat{Y}_s,\widehat{u}_t\right)ds+\int_t^T G\left(s,\widehat{X}_s,\widehat{Y}_s\right)d\overleftarrow{\widehat{B}}_s\\ \qquad -\int_t^T\widehat{Z}_s d\widehat{W}_s-\left(\widehat{\aleph}_T-\widehat{\aleph}_t\right).\end{cases}\tag{3.24}$$

Moreover, we have

$$
\begin{aligned}
J(\widehat{\varpi}.) &= \widehat{\mathbb{E}}\big[\varphi\left(\widehat{X}_T\right) + \psi\left(\widehat{Y}_0\right) + \int_0^T \int_U L\left(t, \widehat{X}_t, \widehat{Y}_t, u\right) \widehat{q}_t\,(du)\,dt\big] \\
&= \widehat{\mathbb{E}}\big[\varphi(\widehat{X}_T) + \psi(\widehat{Y}_0) + \int_0^T L(t, \widehat{X}_t, \widehat{Y}_t, \widehat{u}_t)dt\big], \\
&= J\left(\widehat{\pi}.\right).
\end{aligned}
$$

where $\widehat{\pi} = \left(\widehat{\Omega}, \widehat{\mathcal{F}}, \left(\widehat{\mathcal{F}}_t\right)_{0\leq t\leq T}, \widehat{P}, \widehat{W}., \widehat{B}., \widehat{u}.\right)$. Which achieves the proof. ■

Appendix: Topologies on the Skorokhod space

In this appendix we summarize related to some results by equipping the space of càdlàg functions by the Skorokhod M1 and S-topology and their associated effects, and we will also discuss the relationship between them in the convergence and tightness.

A.1 The Skorokhod M1 topology

Has been introduced a first time in 1956 on the hands of Skorohod for more details see the original paper as we refer to the book Whitt in [47], we denote for space of càdlàg functions (functions which are continuous from the right with left hand limits) from $[0, T]$ into $\mathbb{R}^m$ by $\mathbb{D}([0, T], \mathbb{R}^m)$ ie $X_t \in \mathbb{D}([0, T], \mathbb{R}^m)$ if

$$X_t = \lim_{s \searrow t} X_s = X_{t+} \text{ and } \lim_{s \nearrow t} X_s = X_{t-} \text{ exist in } \mathbb{R}^m.$$

For define the M1 metric in $\mathbb{D}([0, T], \mathbb{R}^m)$, we need to define the completed graphs of the càdlàg function X by

$$G_X = \{(z, t) \in \mathbb{R}^m \times [0, T] \setminus z \in [X_{t-}, X_t]\},$$

where denote by $[X_{t-}, X_t]$ the segments between points $X_{t-}, X_t \in \mathbb{R}^m$ i.e.,

$$[X_{t-}, X_t] = \{z \in \mathbb{R}^m / z = \alpha X_{t-} + (1-\alpha)\, X_t \text{ for } \alpha \in [0,1]\}$$

In fact X_{t-} could be bigger than X_t. Hence the segments $[X_{t-}, X_t]$ is the column that connects X_{t-} and X_t for all discontinuity points t, we denote by

$$Disc\,(X) = \{t \in [0,T] \setminus X_t \neq X_{t-}\}\,,$$

the set of discontinuous points of X,.this set is either finite or countably infinite (see the Corollary 12.2.1 In [47]).

Remark A.1.1 *In Whitt (rf: [47]): for $X_t \in \mathbb{D}([0,T], \mathbb{R}^m)$ that are left-continuous at time T, and there in, the piecewise constant functions use for approximating càdlàg functions on $[0,T]$ are precisely assumed to be continuous at terminal time T (see the Theorem 12.2.2).*

We now define order relations on the complete graph, for any $(z_1, t_1)\,, (z_2, t_2) \in G_X$ we say that $(z_1, t_1) \leq (z_2, t_2)$ if either (1) or (2) where:

(1) $t_2 > t_1$,

(2) $t_2 = t_1$ and $|X_{t_1-} - z_1| \leq |X_{t_1-} - z_2|\,.$

In other words, this relationship ensures that the normal order of G_X is from left to right (ascending order). We now define parametric representation on the complete graph. A parametric representation of X is a continuous non decreasing function (r, e)

mapping $[0, T]$ into G_X i.e.

$$(r, e) : t \in [0, T] \longrightarrow (r(t), e(t)) \in G_X,$$

where $(r, e) \in \mathcal{C}([0, T], \mathbb{R}^m)^2$. We define R_X as the set of all parametric representations of G_X.

Now we define the M1 distance between $X_1, X_2 \in \mathbb{D}([0, T], \mathbb{R}^m)$ by

$$d_{M1}(X_1, X_2) = \inf_{(r_i, e_i) \in R_X \ i=1,2} \max \left[\|e_1 - e_2\|, \|r_1 - r_2\| \right],$$

where $\|.\|$ is the usual uniform norm on $\mathcal{C}([0, T], \mathbb{R}^m)$. We now define oscillation functions that we will use with the M1 topologies, for give characterizes of M1 convergence where for $X \in \mathbb{D}([0, T], \mathbb{R}^m)$, $t \in [0, T]$ and $\delta > 0$ we have

$$w_T(X, t, \delta) = \sup_{\min(t+\delta, T) \geq t_3 > t_2 > t_1 \geq \max(t-\delta, 0)} \|X_{t_2} - [X_{t_1}, X_{t_3}]\|, \tag{1.1}$$

where $\|X_{t_2} - [X_{t_1}, X_{t_3}]\|$ is distance between a point X_{t_2} and a subset $[X_{t_1}, X_{t_3}]$ in $\mathbb{R}^m$ defined dy

$$\begin{aligned} \|X_{t_2} - [X_{t_1}, X_{t_3}]\| &= \inf_{z \in [X_{t_1}, X_{t_3}]} |X_{t_2} - z| \\ &= \inf_{\alpha \in [0,1]} |X_{t_2} - \alpha X_{t_1} + (1-\alpha) X_{t_3}|. \end{aligned}$$

Now we review some of the results of sequences convergence of what and tightness in $\mathbb{D}([0, T], \mathbb{R}^m)$ that endowed with the Skorokhod M1 topology, for more information (see [47], Chapter (12)), Now we start with the following result:

Theorem A.1.2 *The family of the functions* $(X^n)_{n\geq 0} \subset \mathbb{D}([0,T],\mathbb{R}^m)$ *converges to* $X \in \mathbb{D}([0,T],\mathbb{R}^m)$ *in the M1 topology if and only if* X_t^n *converges to* X_t *for each* $t \in Disc\,(X)^c$ *a dense subset of* $[0,T]$ *that includes* 0 *and* T*, and*

$$\lim_{\delta\longrightarrow 0} \limsup_{n\longrightarrow\infty} \sup_{0\leq t\leq T} w_T\,(X,t,\delta) = 0.$$

Theorem A.1.3 *If the sequence* $(X^n)_{n\geq 0} \subset \mathbb{D}([0,T],\mathbb{R}^m)$ *is converges to* X *in* $\mathbb{D}([0,T],\mathbb{R}^m)$ *endowed with the M1 topology.then for all* $t \in [0,T]$ *points continuation of* X *it holds that*

$$\lim_{\delta\longrightarrow 0} \limsup_{n\longrightarrow\infty} \sup_{\min(t+t,T)\geq t\geq \max(t-\delta,0)} |X_t^n - X_t|\,.$$

Remark A.1.4 *i) If* X *is continuous on the interval* $[0,T]$ *then the convergence of* $(X^n)_{n\geq 0}$ *by the M1 topology is equivalent of the uniform convergence (That's true for each* $t \notin Disc\,(X)$*).*
ii) If $\mathbb{D}([0,\infty)\,,\mathbb{R}^m)$ *so the sequence* $(X^n)_{n\geq 0}$ *converges to* X *in* $\mathbb{D}([0,\infty)\,,\mathbb{R}^m)$ *if and only if the restrictions of* X^n *in* $[0,T]$ *converge to the restriction of* X *to* $[0,T]$ *in* $\mathbb{D}([0,T],\mathbb{R}^m)$ *for all* $T > 0$ *that are continuity points of* X *(see [47], Chapter* (3)*).*

Now we define

$$\upsilon_T\,(X,t,\delta) = \sup_{\min(t+\delta,T)\geq t_2\geq t_1\geq \max(t-\delta,0)} |X_{t_1} - X_{t_2}|\,,$$

for the following result:

Theorem A.1.5 *Let subset* $K \subset \mathbb{D}([0,T],\mathbb{R}^m)$ *is relatively compact in* $\mathbb{D}([0,T],\mathbb{R}^m)$ *endowed with the M1 topology if and only if*

$$\sup_{X_t\in K} \|X_t\| \leq C,$$

such that $C \geq 0$ *finite strictly and*

$$\lim_{\delta \longrightarrow 0} \sup_{X_t \in K} \max \left[\left(\sup_{0 \leq t \leq T} w_T\left(X, t, \delta\right) \right), \upsilon_T\left(X, 0, \delta\right), \upsilon_T\left(X, T, \delta\right) \right] = 0. \qquad (1.2)$$

Remark A.1.6 *If* X *have jumps at* 0 *and* T *then we replace the condition* $(1.6.2)$ *by*

$$\lim_{\delta \longrightarrow 0} \sup_{X_t \in K} \left[\sup_{0 \leq t \leq T} w_T\left(X, t, \delta\right) + w_T\left(X, 0, \delta\right) + w_T\left(X, T, \delta\right) \right] = 0,$$

where

$$\begin{cases} w_T\left(X, 0, \delta\right) = \sup_{\delta \geq t > s \geq 0} \left\| X_s - [X_0, X_t] \right\|, \\ w_T\left(X, T, \delta\right) = \sup_{T+\delta \geq t > s \geq T} \left\| X_s - [X_T, X_t] \right\|. \end{cases}$$

We have known that $(\mathbb{D}([0,T], \mathbb{R}^m), +)$ is not topological group because addition is not continuous everywhere in $[0,T]$, and for this we show that it is continuous almost everywhere, and it is measurable so we offer the following results:

Lemma A.1.7 *The Borel* σ*-field on* $\mathbb{D}([0,T], \mathbb{R}^m)$ *coincides with the Kolmogorov* σ*-field.*

Remark A.1.8 *This result allows that the law of a process in* $\mathbb{D}([0,T], \mathbb{R}^m)$ *endowed with the M1 topology, is characterized by its finite dimensional distributions, where the Kolmorogov* σ*-field, is generated by the projection.*

Now we offer a condition for addition to be continuous

Corollary A.1.9 *Let* $(X^n)_{n \geq 0}$ *and* $(Y^n)_{n \geq 0}$ *two sequences, if* $d_{M1}\left(X^n, X\right)$ *converge to* 0 *and* $d_{M1}\left(Y^n, Y\right)$ *converge to* 0 *in* $\mathbb{D}([0,T], \mathbb{R}^m)$ *and*

$$Disc\left(X\right) \cap Disc\left(Y\right) = \varnothing,$$

then $d_{M1}\left(X^n + Y^n, X + Y\right)$ *converge to* 0 *in* $\mathbb{D}\left([0,T], \mathbb{R}^m\right)$.

The proposition shows that the M1 convergence for two sequences by the addition operation hold if two limits do not jump at the same time.

A.2 The Meyer-Zheng topology

The Meyer-Zheng topology, introduced in [43] is a the image measures on graphs $(t, X_t)_{0\leq t\leq T}$ of trajectories X_t under the measure $\lambda(t) = \exp(-t)\,dt$ called pseudo-paths, let $\mathcal{X}_t : t \in \mathbb{R}_+ \to (t, X_t) \in \mathbb{R}_+ \times \mathbb{R}^m$

$$\int_{\mathbb{R}_+} f\left(\mathcal{X}_t\right)\lambda\left(dt\right) = \int_{\mathbb{R}_+\times\mathbb{R}^m} f\left(\mathcal{X}_t\right)\lambda_{\mathcal{X}_t}\left(dt\right),$$

where $\forall A \in \mathcal{B}\left(\mathbb{R}_+ \times \mathbb{R}^m\right), \exists B \in \mathcal{B}\left(\mathbb{R}_+\right)$ such that $\lambda_{\mathcal{X}_t}\left(B\right) = \lambda\left(\mathcal{X}_t^{-1}\left(A\right)\right)$, induced by the weak topology on probability laws on compactified space $[0,\infty] \times \bar{\mathbb{R}}^m$. Thus it provides us with an imbedding of $\mathbb{D}([0,T], \mathbb{R}^m)$ into the compact Polish space $P\left([0,\infty] \times \bar{\mathbb{R}}^m\right)$ of all the probabilities on $[0,\infty] \times \bar{\mathbb{R}}^m$ (with the topology of weak convergence). The associated Borel σ-algebra on $\mathbb{D}([0,T], \mathbb{R}^m)$ is the same one that we get from the Skorohod topology. The Meyer-Zheng topology on the Skorokhod space $\mathbb{D}([0,T], \mathbb{R}^m)$ generated by the convergence in measure.

Definition A.2.1 *The topology MZ on* $\mathbb{D}(I, \mathbb{R}^m)$, *where* $I = [0,T]$ or $\mathbb{R}_+$, *is the topology* *generated by the convergence*

$$\int_I f\left(t, X_t^n\right) \lambda\left(dt\right) \to \int_I f\left(t, X_t\right) \lambda\left(dt\right), \ \forall f \in C_b\left(I \times \mathbb{R}^m\right)$$

The space $\mathbb{D}\left([0,T], \mathbb{R}^m\right)$ with the pseudo-path topology, is not a Polish space. But from the definition we know that it is homeomorphic to a subspace of the Polish space $P\left([0,\infty] \times \bar{\mathbb{R}}^m\right)$, and hence is a separable metric space. Let $\left(\Omega, \mathcal{F}, \left(\mathcal{F}_t\right)_{t\geq 0}, P\right)$ be a filtered probability space, we define the **conditional variation** of $Y \in \mathbb{D}([0,T], \mathbb{R}^k)$ over the interval $[0,T]$ as follows

$$VC\left(Y\right) = \sup \mathbb{E}\left[\sum_k \left|E\left(Y_{t_{k+1}} - Y_{t_k} \diagup \mathcal{F}_{t_k}\right)\right|\right],$$

for all the subdivision $0 = t_0 < t_1 < ... < t_n = T$, $n \in \mathbb{N}$. The process Y is call *quasimartingale* if $CV(Y) < +\infty$.

Corollary A.2.2 *Let subset $K \subset \mathbb{D}([0,T], \mathbb{R}^m)$ is relatively compact in $\mathbb{D}([0,T], \mathbb{R}^m)$ endowed with the pseudo-path topology then*

$$\sup_{X \in K} \|X\| < +\infty,$$

and for each $a < b$

$$\sup_{X \in K} N^{a,b}\left(X\right) < +\infty,$$

where $N^{a,b}\left(X\right)$ is the number of up-crossings for each $a < b$ (see p 358, in [43]).

Convergence of sequences in this topology is just the convergence in Lebesgue measure. This topology is not suitable for the "functional convergence" of stochastic processes. However, as a weak topology it can be and it is used in existence problems, due to easy-to-check compactness criteria. Now we offer Meyer-Zheng convergence naturally in the context of the following criteria of compactness

Lemma A.2.3 *Let $(Y^n)_{n\geq 0}$ be a sequence of processes on the space $\mathbb{D}([0,T],\mathbb{R}^k)$, such that:*

$$\sup_n VC\left(Y^n\right) < \infty,$$

then $(P_{Y^n})_{n\geq 0}$ is tight in $P\left(\mathbb{D}([0,T],\mathbb{R}^k)\right)$.

A.3 The Jakubowski topology

Has been introduced a first time in 1997 on the hands of Jakubowski [31], in fact this topology is weaker than the Skorokhod M1 topology as will be explained in this section and his advantage lies in ease the tightness criteria, which presented by both of the Meyer & Zheng [43].

In this section, we will provide some results about S-topology and associated effects in $\mathbb{D}([0,T],\mathbb{R}^m)$ finally we conclude with the relationship between M1 and S-topology, and we refer to original paper see.[31] for detailed of the definition and properties of S-convergence. Now we offer S-topology naturally in the context of the following criteria of compactness.

Definition A.3.1 *A subset $K \subset \mathbb{D}([0,T],\mathbb{R}^m)$, we say that K is S-Closed (Closed in S-topology) if for any $(X^n)_{n\geq 0} \subset K$ we have X^n S-convergence to X with $X \in K$.*

Definition A.3.2 *The family $(X^n)_{n\geq 0} \subset \mathbb{D}([0,T],\mathbb{R}^m)$, we say that is convergent in distribution if for any subsequence $(X^{n_k})_{k\geq 0}$ we can find a further subsequence $(X^{n_{k_i}})_{i\geq 0}$ admits a strong a.s. Skorohod representation.*

Lemma A.3.3[The a.s. Skorokhod Representation] *Let (D,S) be a topological space on which there exists a countable family of S-continuous functions separating points in X. Let $(X^n)_{n\geq 0}$ be a uniformly tight sequence of laws on D. In every subsequence (X^{n_k}) one can find a further subsequence $(X^{n_{k_i}})$ and stochastic processes $\left(Y^k\right)$ defined on $([0,T],B_{[0,T]},l)$ such that*

$$\text{the laws of } X^{n_{k_i}} \text{ and } Y^k \text{ coincide for } k=0,1,2,..., \tag{1.3}$$

$$Y^k(\omega) \longrightarrow Y^0(\omega), \qquad \text{as } k\to\infty, \tag{1.4}$$

and for each $\varepsilon > 0$ there exists an S-compact subset $K_\varepsilon \subset D$ such that

$$P\left(\omega\in[0,T] : Y^k(\omega)\in K_\varepsilon,\ k=0,1,2,...,\right) > 1-\varepsilon. \tag{1.5}$$

One can say that (1.4) and (1.5) describe “the almost sure convergence in compacts” and that $(1.3)-(1.5)$ define the strong a.s. Skorokhod representation for subsequences (“strong” because of condition (1.5)).

Proposition A.3.4 *Let subset $K \subset \mathbb{D}([0,T],\mathbb{R}^m)$ is relatively compact in $\mathbb{D}([0,T],\mathbb{R}^m)$ endowed with the S-topology then*

$$\sup_{X\in K}\|X\| < +\infty,$$

and for each $a<b$

$$\sup_{X\in K} N^{a,b}(X) < +\infty,$$

where $N^{a,b}(X)$ is the number of up-crossings for each $a<b$ see [31] and [43].

Remark A.3.5 *X^n converge to X in S-topology if, and only if, in each subsequence X^{n_k} one can find a further subsequence $X^{n_{k_l}}$ S-convergence to X.*

Lemma A.3.6 *Let the family $(X^n)_{n\geq 0}$in $\mathbb{D}([0,T],\mathbb{R}^m)$ we say that X^n S-convergence to X If we find a subsequence $(X^{n_k})_{k\geq 0}$ and countable set $D\subset[0,T]$ such that*

$$\begin{cases} X^{n_k} \longrightarrow X \text{ for all } t\in D^c = [0,T]\diagdown D, \\ \text{and } X_T^n \longrightarrow X_T \text{ for } t=T. \end{cases}$$

Remark A.3.7 *i) The set D it is a union of the discontinuity points of the signed measure on $([0,T], B([0,T]))$ see [Jakubowski p 8].*

ii) The projection $\pi_T : y \in (\mathbb{D}([0,T],\mathbb{R}^m), S) \longrightarrow \pi_T(y) = y(T) \in \mathbb{R}^m$ is continuous (see Remark (2.4), p. 8 in [31]), but $y \longrightarrow y(t)$ is not continuous for each $0\leq t\leq T$.

This topology is weaker than the Skorokhod topology and the tightness criteria are easier to establish. This criteria is the same as that of the Meyer and Zheng topology. We recall some facts about the S-topology.

Proposition A.3.8[A criteria for S-tight] *A sequence $(Y^n)_{n>0}$ is* S*-tight if and only if it is relatively compact on the* S*-topology.*

Proposition A.3.9 *Let $(Y^n)_{n>0}$ be a family of stochastic processes in $\mathbb{D}([0;T];R^m)$. Then this family is tight for the* S*-topology if and only if $(\|Y^n\|)_n$ and $\left(N^{a,b}(Y^n)\right)_n$ are tight for each $a<b$.*

We recall (see Meyer & Zheng [43] and Jakubowski [31],[32]) that for a familly $(Y^n)_n$ of quasi-martingales on the probability space $\left(\Omega,\{\mathcal{F}_t\}_{0\leq t\leq T},P\right)$, the following condition insures the tightness of the familly $(Y^n)_n$ on the space $\mathbb{D}\left([0,T];\mathbb{R}^m\right)$ endowed with the S-topology

$$\sup_n\left(CV\left(Y^n\right)\right)<+\infty,$$

where the conditional variation of Y on $[0,T]$, and is defined by

$$CV\left(Y^n\right)=\sup\mathbb{E}\left[\sum_i\left|\mathbb{E}\left(Y^n_{t_i+1}-Y^n_{t_i}\mid\mathcal{F}^n_{t_i}\right)\right|\right].$$

Proposition A.3.10 *Let $(Y^n,\ M^n)$ be a multidimensional process in $\mathbb{D}([0,\ T];\ \mathbb{R}^m)$ converging to $(Y,\ M)$ in the S-topology. Let $(\mathcal{F}^{Y^n}_t)_{t\geq0}$ (resp. $(\mathcal{F}^Y_t)_{t\geq0}$) be the minimal complete admissible filtration for Y^n (resp.Y). We assume that $\sup\limits_n\mathbb{E}\left[\sup_{0\leq t\leq T}|M^n_t|^2\right]<C_T\ \forall T>0,\ M^n$ is a $\mathcal{F}^{Y^n}$-martingale and M is a $\mathcal{F}^Y$-adapted. Then M is a $\mathcal{F}^Y$-martingale.*

Proposition A.3.11 *Let $(Y^n)_{n>0}$ be a sequence of processes converging weakly in $\mathbb{D}([0,\ T];\ R^m)$ to Y. We assume that $\sup\limits_n\mathbb{E}\left[\sup_{0\leq t\leq T}|Y^n_t|^2\right]<+\infty$. Hence, for any $t\geq0$, $\mathbb{E}\left[\sup_{0\leq t\leq T}|Y_t|^2\right]<+\infty$.*

Now the relationsbhip between M1 and S-topology come in context of the following result

Theorem A.3.12 *The M1 topology is stronger than the S-topology hence a set* $A \subset \mathbb{D}([0,T],\mathbb{R}^m)$, *which is relatively M1 compact is also relatively S compact.*

Lemma A.3.13 *Let* $X \in \mathbb{D}([0,T],\mathbb{R}^m)$, *then for each* $0 \leq a < b \leq T$ *we have if* $(b-a) > 2d\,[a,b]$

$$N^{a,b}\left(X_t\right) \leq N_{b-a}\left(X_t\right) \leq \frac{|X_a - X_b| + d\,[a,b]}{(b-a) - d\,[a,b]}$$

where $d\,[a,b] = \sup_{b \geq t_3 > t_2 > t_1 \geq a} \|X_{t_2} - [X_{t_1}, X_{t_3}]\|$, and $N_{b-a}\left(X_t\right)$ *is the number of* $(b-a)$*-oscillations of* X_t *in the interval* $[a,b]$.

That the topology M1 and S is suitable for the needs of the theory of limit for random functions in $\mathbb{D}([0,T],\mathbb{R}^m)$. He admits even such effective tools as Prohorov theories are valid. There is a Skorohod representation and finite dimensional convergence outside a countable set see original paper for Jakubowski and Billingsley (see [31] and [16] Chapter (1)).

Conclusion

In this thesis, we have investigated about two results on the existence of optimal relaxed controls as well as strict optimal controls of an stochastic optimal control problems, by using probabilistic effective tools. In the first part, we treating a singular control problem for systems of forward-backward stochastic differential equations of mean-field type (MF-FBSDEs). The ingredients used in the proof of this result is based on tightness results of the distributions of the processes defining the singular control problem and the Skorokhod representation theorem on the space of càdlàg functions, endowed with the Jakubowski S-topology. The assumptions on the coefficients which depend on the stats of the solution processes as well as their distribution via the expectation of some function, are made to ensure weak convergence of the processes under consideration and the corresponding cost functional which is also of mean-field type.

In the same context, we can reformulate the control problem by using the Jakubowski S-topology for the space of $Y^{q,\xi}$ and we replaced the topology of uniform convergence for the space of $X^{q,\xi}$ by M1 topology of the Skorokhod induced by the metric d_{M1}, which considered stronger than S-topology, which is defined using the concept of local uniform convergence at all points of continuity. Under the same assumptions one can conclude that our main result remains correct, despite the various effects associated with the M1 topology in $\mathbb{D}(0,T;\mathbb{R}^m)$.

In this part we assume that the initial and the terminal time are continuous points of $X^{q,\xi}$, $Y^{q,\xi} \in \mathbb{D}(0,T;\mathbb{R}^m)$. But it's natural to accepted jumps at initial and terminal time, this discontinuity of $X^{q,\xi}$, $Y^{q,\xi}$ at the terminal time T and initial time 0 $\Big($ie: $X_T^{q,\xi} \neq X_{T-}^{q,\xi}$, and $Y_0^{q,\xi} \neq Y_{0-}^{q,\xi}\Big)$, it is the reason for defining the cost function in this case without initial and terminal cost because, the convergence cannot be guaranteed in the M1 topology at terminal time where he is equivalent of the uniform convergence for each $t \notin Disc(X)$, on the other hand the convergence of $Y_0^{q^n,\xi^n}$ to $Y_0^{q,\xi}$ is not ensured with respect to the S-topology. But, we can calculate the value of $Y_0^{q,\xi}$ from the fact that

$$Y_0^{q^n,\xi^n} = \mathcal{Y}_0^{q^n,\xi^n} + \int_0^T \varphi_s d\xi_s^n \text{ converges to}$$

$$Y_0^{q,\xi} = \mathcal{Y}_0^{q,\xi} + \int_0^T \varphi_s d\xi_s \text{ as } n \text{ tends to } \infty.$$

This result assured by sequential continuity of the addition, with respect to the S-topology (see [31] Remark (3.12)), despite that the space of càdlàg functions endowed with S-topology is not a linear topological space. So by Skorokhod's representation theorem the sequence $Y_0^{q^n,\xi^n}$ converges in distribution to $Y_0^{q,\xi}$.

In the second part, we deal with a forward-backward doubly stochastic differential equations (FBDSDEs), where the coefficients depend on X and Y, but not on the second variable Z, with an uncontrolled diffusion coefficient. We use Jakubowski S-topology and a suitable version of the Skorokhod theorem to prove the main result. Under some additional convexity assumption, we show that the relaxed optimal control, is in fact as a strict control.

Bibliography

[1] F. Antonelli, *Backward-forward stochastic differential equations*, Ann. Appl. Probab, 3, (1993), 777–793.

[2] N.U. Ahmed and C.D. Charalambous, *Stochastic minimum principle for partially observed systems subject to continuous and jump diffusion processes and driven by relaxed controls*, SIAM J. Control Optim. 51(2013), no. 4, 3235–3257.

[3] D. Anderson and B. Djehiche, *A maximum principle for relaxed stochastic control of linear SDEs with application to bond portfolio optimization*, Mathematical methods in operations research. 72(2010), no. 2, 273–310.

[4] F. Antonelli, J. Ma, *Weak solutions of forward–backward SDE's,* Stoch. Anal. Appl. 21 (3) (2003) 493–514.

[5] B. Boufoussi, J. Van Casteren, and N. Mrhardy, *Generalized backward doubly stochasticdifferential equations and SPDEs with nonlinear Neumann boundary conditions*,

Bernoulli13(2), (2007), 423-446

[6] R. Buckdahn, B. Djehiche, J. Li, and S. Peng, *Mean-Field backward stochastic differential equations: a limit approach*. Ann. Prob., 37(4), (2009) 1524–1565.

[7] R. Buckdahn, H.-J. Engelbert, *A backward stochastic differential equation without strong solution,* Teor. Veroyatn. Primen. 50 (2) (2005) 390–396.

[8] R. Buckdahn, H.-J. Engelbert, *On the continuity of weak solutions of backward stochastic differential equations,* Teor. Veroyatn. Primen. 52 (1) (2007), 190–199.

[9] R. Buckdahn, H.-J. Engelbert, A. Rascanu, *On weak solutions of backward stochastic differential equations,* Teor. Veroyatn. Primen. 49 (1) (2004), 70–108.

[10] K. Bahlali, B. Gherbal, B. Mezerdi, *Existence of optimal controls for systems driven by FBSDEs,* Systems and Control Letters, Vol. 60 (2011), 344-349.

[11] K. Bahlali, B. Gherbal, B. Mezerdi, *Existence and optimality conditions in stochastic control of linear BSDEs,* Random Oper. Stoch. Equ., Vol. 18, No 3 (2010), 185-197.

[12] B. Radhia and B. Gherbal, *Existence of Optimal Controls for Forward-Backward Stochastic Differential Equations of Mean-Field Type*, J. of Numerical Mathematics and Stochastics, 9 (1) (2017), 33-47.

[13] R. Buckdahn, B. Labed, C. Rainer, L. Tamer, *Existence of an optimal control for stochastic systems with nonlinear cost functional,* Stochastics, Vol. 82, No 3 (2010), 241-256.

[14] K. Bahlali, M. Mezerdi and B. Mezerdi, *Existence of optimal controls for systems governed by mean-field stochastic differential equations,* Afrika Statistika, Vol. 9, (2014), 627–645.

[15] K. Bahlali, B. Mezerdi, M. N'zi, Y. Ouknine, *Weak solutions and a Yamada–Watanabe theorem for FBSDEs,* Random Oper. Stoch. Equ. 15 (3) (2007), 271–285.

[16] P. Billingsley, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, second ed., 1999. A Wiley-Interscience Publication.

[17] N. Bouchemella, P. Raynaud de Fitte, *Weak solutions of backward stochastic differential equations with continuous generator*, Stochastic Processes and their Applications. 124 (2014), 927–960.

[18] Z. Brzezniak and R. Serrano, *Optimal relaxed control of dissipative stochastic partial differential equations in Banach spaces,* SIAM Journal on Control and Optimization 51, 3 (2013), 2664–2703.

[19] R. Carmona, F. Delarue, *Probabilistic analysis of mean-field games*. SIAM J. Control Optim. 51(4), (2013), 2705–2734.

[20] R. Carmona, F. Delarue, *Forward-Backward Stochastic Differential Equations and Controlled McKean-Vlasov Dynamics*. Ann. Probab Volume 43, Number 5 (2015), 2647-2700.

[21] J. Cvitanic X.H. Wan and J.F. Zhang, *Optimal contracts in continuous-time models*, Journal of applied mathematics and stochastic analysis. (2006), Article ID 95203, 1–27.

[22] F. Delarue, G. Guatteri, *Weak existence and uniqueness for forward–backward SDEs,* Stochastic Process. Appl. 116 (12) (2006), 1712–1742.

[23] N. El Karoui, D. H. Nguyen, and M. Jeanblanc-Picqué, *Compactification methods in the control of degenerate diffusions: existence of an optimal control*, Stochastics, **20**(3), (1987), 169–219.

[24] G. Fu and U. Horst, *Mean Field Games with Singular Controls*, SIAM Journal on Control and Optimization, 55(6), (2017), 3833-3868.

[25] W.H. Fleming, *Generalized solutions in optimal stochastic control,* in: Differential Games and Control theory II, Proceedings of 2nd Conference, University of Rhode

Island, Kingston, RI, (1976), in: Lecture Notes in Pure and Appl. Math., vol. 30 (1977), Marcel Dekker, New York, 147–165

[26] B. Gherbal, *Optimal control problems for linear backward doubly stochastic differential equations*. Random. Oper. Stoc. Equ., Vol. 22, No 3, (2014), 129–138.

[27] U. Haussmann and W. Suo, *Singular optimal stochastic controls I: Existence*, SIAM. J. Control Optim., 33 (1995), pp. 916–936, https://doi.org/10.1137/S0363012993250256.

[28] U.G. Haussmann, *Existence of optimal Markovian controls for degenerate diffusions,* in: Stochastic Differential Systems, in: Lecture Notes in Control and Inform. Sci., vol. 78 (1986), Springer, Berlin, 171–186.

[29] U.G. Haussmann and J.-P. Lepeltier, *On the existence of optimal controls*. SIAM J. Cont. Optim. 28(4) (1990), 851–902.

[30] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Mathematical Library, Vol. 24, (1981). North-Holland, Amsterdam.

[31] A. Jakubowski, A., *A non-Skorokhod topology on the Skorokhod space*. *Electron. J. Probab.* **2** Paper No.4, (1997), pp.1–21.

[32] A. Jakubowski, *Convergence in various topologies for stochastic integrals driven by semimartingales*, The Annals of Probab., **24** (1996), 2141–2153.

[33] J. Jacod and J. Mémin, *Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité,* Séminaire de Probabilités, XV, Lect. Notes in Math. 850, Springer, Berlin, (1981), 529–546.

[34] S.L. Ji and X.Y. Zhou, *A maximum principle for stochastic optimal control with terminal state constraints, and its applications*, Communications in information and systems. 6(2006), no. 4, 321–337.

[35] H.J. Kushner, *Existence results for optimal stochastic controls,* J. Optim. Theory Appl. 15 (1975), 347–359.

[36] D. Lacker, *Mean field games via controlled martingale problems: existence of markovian equilibria*, Stochastic Processes and their Applications. 125 (2015), no. 7, 2856–2894.

[37] D. Lacker, *Mean Field Games via Controlled Martingale Problems: Existence of Markovian Equi- libria*, Stoch. Proc. Appl. 125 (2015), 2856-2894.

[38] D. Lacker, *A general Characterization of the Mean Field Limits for Stochastic Differential Games*, Probab. Theory Related Fields, 165 (2016), 581-648..

[39] E. Pardoux, S, Peng, *Backward doubly stochastic differential equations and system of quasilinear SPDEs.* Probab. Theory Related Fields 98, No.2 (1994), 209–227.

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

[41] X. Li, and J. Yong, *Necessary conditions of optimal control for distributed parameter systems*, SIAM J. Control & Optim., 29 (1991), 895–908.

[42] J. Ma, J. Yong, *Forward backward stochastic differential equations and their applications,* in: Lecture notes in Mathematics, vol. 1702, Springer, Berlin, Heidelberg (1999).

[43] P.A. Meyer, W.A. Zheng, *Tightness criteria for laws of semimartingales,* Ann. Inst. H. Poincaré, Probab. Statist. 20 (4) (1984), 217–248.

[44] D. Nualart, and E. Pardoux, *Stochastic calculus with anticipating integrands*, Probability Theory and Related Fields volume 78, (1988), 535–581.

[45] J.T. Shi and Z. Wu, *Maximum principle for partially-observed optimal control of fully-coupled forward-backward stochastic systems*, J. Optim. Theory Appl. 145 (2010), 543-578

[46] H. Tang and Z. Wu, *Backward stochastic differential equations with Markov chains and related asymptotic properties,* Advances in Difference Equations (2013), 2013:285.

[47] W. Whitt, *Stochastic Process Limits: an Introduction to Stochastic Process Limits and Their Application to Queues*, Springer-Verlag, New York, 2002.

[48] Wu, Z. F. Zhang, *BDSDEs with locally monotone coefficients and Sobolev solutions for SPDEs*, J. Differential Equations, 251 (2011), 759-784.

[49] J. Yong, X.Y Zhou, *Stochastic controls, in: Hamiltonian Systems and HJB Equations*, Springer, New York (1999).

[50] B. Zhu, and B. Han, *Backward doubly stochastic differential equations with infinite time-horizon*, Applications of Mathematics 6, (2012), 641-653.