People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

MOHAMED KHIDER UNIVERSITY, BISKRA

Faculty of Exact Science and the Natural Science and Life

DEPARTMENT OF MATHEMATICS

A Thesis Presented for the Degree of

Doctorate of science in Applied Mathematics

By

Berrouis Nassima

Title

The study of optimal controls for forward backward doubly stochastic differential equations

Examination Commite Members

MERAGHNI Djamel	Pr	University of Biskra	President
GHERBAL Boulakhras	Pr	University of Biskra	Supervisor
HAFAYED Mokhtar	Pr	University of Biskra	Examiner
SAADI Khalil	Pr	University of M'sila	Examiner
TIAIBA Abdelmoumen	Pr	University of M'sila	Examiner
DJENAIHI Youcef	Dr	University of Setif	Examiner

2022

Acknowledgement

Foremost, I would like to express my sincere gratitude to my advisor Pr. GHERBAL Boulakhras for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Many thanks to the members of committee for their corrections and time they spent in reading and evaluating this work: Pr. MERAGHNI Djamel my committee chair, Pr. HAFAYED Mokhtar, Pr. SAADI Khalil, Pr. TIAIBA Abdelmoumen, and Dr. DJENAIHI Youcef.

I would like to thank Pr. HAFAYED Mokhtar for his advice, he was always ready to help me every time I asked.

I would like to thank my family and all my colleagues of the Mathematics Department.

Abstract

In this thesis, we are concerned with stochastic optimal control problems of systems governed by different types of forward-backward doubly stochastic differential equations.

In the first part, we prove existence of strong optimal control (that is adapted to the initial σ-algebra) for linear forward-backward doubly stochastic differential equations, with random coefficients and non linear functional cost. The control domain and the cost function were assumed convex. The proof is based on strong convergence techniques for the associated linear FBDSDEs and Mazur's theorem. We derive also necessary and sufficient conditions for optimality for this strict control problem. This result is based on the convex optimization principle.

In the second part of this thesis, we generelize the results of the first part to systems governed by linear forward-backward doubly stochastic differential equations of mean field type, in which the coefficients depend on the state process, and also on the distribution of the state process, via the expectation of some function of the state. In particularly, we establish the existence of strong optimal solutions of a control problem for dynamics driven by a linear forward-backward doubly stochastic differential equations of mean- field type (MF-LFBDSDEs), with random coefficients and non linear functional cost which is also of mean-field type. Moreover, we establish necessary as well as sufficient optimality conditions for this kind of control problem.

In the last part, we establish necessary as well as sufficient optimality conditions for existence of both optimal relaxed control and optimal strict control for dynamics of nonlinear forward-backward doubly SDEs of mean-field type.

Résumé

Dans cette thèse, nous nous intéressons aux problèmes de contrôle optimal stochastique de systèmes gouvernés par différents types d'équations différentielles doublement stochastique progressives-rétrogrades.

Dans la première partie, nous prouvons l'existence d'un contrôle optimal pour les équations différentielles doublement stochastique progressives-rétrogrades linéaires, avec des coefficients aléatoires et une fonction de coût non linéaire. Le domaine de contrôle et la fonction de coût sont supposés convexes. La preuve est basée sur des techniques de convergence forte pour les EDDSPRs linéaires et le théorème de Mazur. Nous établissons également les conditions nécessaires et suffisantes d'optimalité pour ce problème de contrôle strict. Ce résultat est basé sur le principe d'optimisation convexe.

Dans la deuxième partie de cette thèse, on généralise les résultats du première partie pour des systèmes gouvernés par des équations différentielles doublement stochastique progressives-rétrogrades linéaires de type champ moyen, dans lequel les coefficients dépendent du processus d'état, ainsi que de la distribution du processus d'état, via l'espérance d'une fonction de l'état. En particulier, nous établissons l'existence d'une solution optimale forte du problème de contrôle pour des équations différentielles doublement stochastique progressives-rétrogrades linéaires de type champ moyen, à coefficients aléatoires et une fonction de coût non linéaire qu'est aussi de type champ moyen. De plus, nous établissons des conditions nécessaires ainsi que des conditions suffisantes d'optimalité pour ce genre de problème de contrôle.

Dans la dernière partie, nous établissons les conditions nécessaires et suffisantes d'optimalité pour les deux problèmes de contrôle relaxé et strict pour les équations différentielles doublement stochastique progressives-rétrogrades non linéaires de type champ moyen.

Symbols and Acronyms

The different symbols and acronyms used in this thesis.

$(B_t)_{t\geq 0}$	: Brownian motion.
$\mathbb{E}\left[X\right]$	: Expectation at x.
$\mathbb{E}\left[X/\mathcal{F}_t\right]$	: Conditional expectation.
$(\mathcal{F}_t)_{t\geq 0}$	: Filtration.
$\mathcal{F}_t := \mathcal{F}_t^W \vee \mathcal{F}_{t,T}^B$	: $\sigma-$fields generated by $\mathcal{F}_t^W \cup \mathcal{F}_{t,T}^B$.
H	: The Hamiltonian.
$\mathbb{J}\left(u.\right)$	: The cost function.
$\mathcal{N}$	: The collection of class of $\mathbb{P}-$null sets of $\mathcal{F}$.
$q.$	: Optimal relaxed control.
$\mathcal{R}$	: The set of admissible relaxed controls.
$\mathbb{R}$	: Real numbers.
$\mathbb{R}^n$	: $n-$dimensional real Euclidean space.
$\mathbb{R}^{n\times d}$	: The set of all $(n \times d)$ real matrixes.
U	: The set of values taken by the strict control $u_\cdot$.
$\mathcal{U}$	: The set of admissible strict controls.
$u.$	: Admissible control.
$(\Omega, \mathcal{F}, \mathbb{P})$	: Probability space.
$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$	: A filtered probability space.

$SDEs$	: Stochastic differential equations.
$BSDEs$	: Backward stochastic differential equations.
$FBSDEs$	: Forward-backward stochastic differential equations.
$FBDSDEs$	: Forward-backward doubly stochastic differential equations.
$MF-FBDSDEs$	: Forward-backward doubly SDEs of mean field type.
$a.e.$	: almost everywhere.
$a.s.$	: Almost surely.
$r.v$	: random variable.

Contents

Introduction

The mathematical theory of stochastic differential equations was developed in the 1940s through the groundbreaking work of Japanese mathematician Kiyosi Itô, who introduced the concept of stochastic integral and initiated the study of nonlinear stochastic differential equations (SDEs). The linear backward stochastic differential equations (LBSDEs in short) related to the stochastic version of Pontryagin's maximum principle, has been studied by Bismut [10]. After that, the non linear BSDEs have been introduced by Pardoux and Peng [33]. Forward-backward stochastic differential equations (FBSDEs in short) were first studied by Antonelli (see [5]), 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 forward-backward stochastic differential equations on an arbitrarily given time interval. For example, the four-step scheme approach of Ma et al. [27], in which the authors proved the result of 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.

A new class of stochastic differential equations with terminal condition, called backward doubly stochastic differential equation (BDSDE) have been introduced by Pardoux and Peng in [34]. The authors show existence and uniqueness for this kind of stochastic differential equation and produce a probabilistic representation of certain quasi-linear stochastic partial differential equations (SPDE) extending the Feynman–Kac formula for linear SPDEs. Recently, Al-Hussein and Gherbal, [3], established the existence and uniqueness of the solutions of multidimensional forward-backward doubly SDEs with random jumps. The problem of existence of optimal controls for various control systems is a fundamental problem in stochastic optimal control theory. The existence of optimal controls for stochastic differential equations (SDEs), is guaranteed by the presence of the Roxin-type convexity condition (see [15, 22, 24]). Without this condition, a strict optimal control may fail to exist. In [6] Bahlali et al proved an existence result of strong optimal strict control for linear backward stochastic differential equations (BSDEs). They showed the existence in the strong formulation of the control problem, where the optimal control is adapted to the original filtration. In this subject, Gherbal in [19] proved for the first time the existence of optimal strict control for systems of linear backward doubly SDEs and establish necessary as well as sufficient optimality conditions in the form of a stochastic maximum principle for this kind of systems. Also, Al-Hussein and Gherbal established in [2] sufficient conditions for optimal control of fully coupled multi-dimensional FBDSDEs with Poisson jumps.

The first of our main aims in this work is to prove existence of optimal strict control and establish necessary as well as sufficient optimality conditions for a control problem governed by the following linear forward-backward doubly SDEs

$$\begin{cases} dX_t = (\alpha_t X_t + \beta_t u_t)dt + (\widehat{\alpha}_t X_t + \widehat{\beta}_t u_t)dW_t \\ \\ dY_t = -(\gamma_t X_t + \widehat{\gamma}_t Y_t + \delta_t Z_t + \widehat{\delta}_t u_t)dt \\ \qquad\qquad -(\eta_t X_t + \widehat{\eta}_t Y_t + \theta_t Z_t + \widehat{\theta}_t u_t)\overleftarrow{dB}_t + Z_t dW_t, \\ \\ X_0 = x, Y_T = \xi, \end{cases} \tag{1}$$

where $\alpha_., \widehat{\alpha}_., \beta_., \widehat{\beta}_., \gamma_., \widehat{\gamma}_., \delta_., \widehat{\delta}_., \eta_., \widehat{\eta}_., \theta_.$ and $\widehat{\theta}_.$ are matrix-valued functions of suitable sizes, x is a square integrable and $\mathcal{F}_0-$measurable process and ξ is a square integrable and $\mathcal{F}_T-$measurable process, the solution $(X_., Y_., Z_.)$ takes values in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$. $(W_t)_{t\geq 0}, (B_t)_{t\geq 0}$ are two mutually independent standard Brownian motions, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking their values respectively in $\mathbb{R}^d$ and in $\mathbb{R}^k$, and $u_.$ represents a strict control. The integral with respect to B is a backward Itô integral, while the integral with respect to W is a standard forward Itô integral.

We shall consider a functional cost to be minimized, over the set $\mathcal{U}$ of a admissible strict controls, as the following

$$\mathbb{J}(u.) := \mathbb{E}\left[\varphi\left(X_T^u\right) + \psi\left(Y_0^u\right) + \int_0^T L\left(t, X_t^u, Y_t^u, Z_t^u, u_t\right) dt\right], \tag{2}$$

where φ, ψ and L are appropriate functions.

Stochastic optimal control of mean-field type recently are extensively studied, due to their applications in economics and mathematical finance. In 2009, Buckdahn et al. [11] established the theory of mean-field backward stochastic differential equations which were derived as a limit of some highly dimensional system of FBSDEs, corresponding to a large number of particles. Since that, many authors treated the system of this kind of Mckean-Vlasov type (see [1] and [25]). As it is well-knew that the adjoint equation of a controlled SDEs of mean-field type is a backward-SDEs of mean-field type, the maximum principle for optimal control systems of mean-field type (MF-SDEs, MF-BSDEs and MF-FBSDEs) has becomes popular topic. In this regard, Carmona and Dularue proved in [13] the existence of solution for mean-field FBSDEs systems. A maximum principle for fully coupled FBSDEs of mean-field type has been established by Li and Liu [26], where the control domain is not assumed to be convex. A maximum principle for mean-field FBSDEs with jumps with uncontrolled diffusion, where the domain of control is not assumed to be convex, has been investigated by Hafayed [20], Hafayed et al. [21] established a maximum principle for MF-FBSDEJs with controlled diffusion, where the domain of control is assumed to be convex. One can refer to [[4], [12], [26] and [28]] for more result on the maximum principles for different types of mean-field systems. The existence of optimal control for systems of mean-field forward backward stochastic differential equations has been proved by Benbrahim and Gherbal [8], where the diffusion is controlled.

The second main result is to prove existence of strong optimal control and to establish necessary as well as sufficient optimality conditions for a control problem of systems governed by the following linear FBDSDEs of mean- field type:

$$\begin{cases} dy_t^u = (a_t y_t^u + \widehat{a}_t \mathbb{E}\left[y_t^u\right] + b_t u_t)dt + (c_t.y_t^u + \widehat{c}_t \mathbb{E}\left[y_t^u\right] + \widehat{b}_t u_t)dW_t \\ \\ dY_t^u = -(d_t y_t^u + \widehat{d}_t \mathbb{E}\left[y_t^u\right] + e_t Y_t^u + \widehat{e}_t \mathbb{E}\left[Y_t^u\right] + f_t Z_t^u + \widehat{f}_t \mathbb{E}\left[Z_t^u\right] + g_t u_t)dt \\ \qquad - \Big(h_t y_t^u + \widehat{h}_t \mathbb{E}\left[y_t^u\right] + k_t Y_t^u + \widehat{k}_t \mathbb{E}\left[Y_t^u\right] \\ \qquad\qquad + m_t Z_t^u + \widehat{m}_t \mathbb{E}\left[Z_t^u\right] + \widehat{g}_t u_t \Big) \overleftarrow{dB}_t + Z_t^u dW_t, \\ y_0^u = x, Y_T = \xi, \end{cases} \tag{3}$$

and a cost functional:

$$\mathbb{J}(u.) := \mathbb{E}\left[\alpha\left(y_T^u, \mathbb{E}\left[y_T^u\right]\right) + \beta\left(Y_0^u, \mathbb{E}\left[Y_0^u\right]\right) + \int_0^T \ell\left(t, y_t^u, \mathbb{E}\left[y_t^u\right], Y_t^u, \mathbb{E}\left[Y_t^u\right], Z_t^u, \mathbb{E}\left[Z_t^u\right], u_t\right) dt\right], \tag{4}$$

where $a., \widehat{a}., b., \widehat{b}., c., \widehat{c}., d., \widehat{d}., e., \widehat{e}., f., \widehat{f}., g., \widehat{g}., h., \widehat{h}., k., \widehat{k}., m.$ and $\widehat{m}.$ are matrix-valued functions of suitable sizes. The solution $(y., Y., Z.)$ takes values in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ and $u.$ is the control variable values in subset U of $\mathbb{R}^k$. α, β, ℓ are a given functions.

An admissible control $u.$ is a square integrable, $\mathcal{F}_t-$measurable process with values in some subset $U \subseteq \mathbb{R}^k$.

Note that we have an additional constraint that a control must be square-integrable just to ensure the existence of solutions of (3) under $u.$. We say that an admissible control $u^*_. \in \mathcal{U}$ is an optimal control if

$$\mathbb{J}\left(u^*_.\right) = \inf_{v. \in \mathcal{U}} \mathbb{J}\left(v.\right). \tag{5}$$

In this part 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 the state. The mean-field FBDSDEs (3) called also McKean-Vlasov systems are obtained as the mean square limit of an interacting particle system of the form

$$\begin{cases} dy_t^{u,i,n} = (a_t y_t^u + \widehat{a}_t \frac{1}{n}\sum_{j=1}^{n} y_t^{u,j,n} + b_t u_t)dt \\ \qquad\qquad + (c_t.y_t^u + \widehat{c}_t \frac{1}{n}\sum_{j=1}^{n} y_t^{u,j,n} + \widehat{b}_t u_t)dW_t^i \\ dY_t^u = -\Bigg(d_t y_t^u + \widehat{d}_t \frac{1}{n}\sum_{j=1}^{n} y_t^{u,j,n} + e_t Y_t^u + \widehat{e}_t \frac{1}{n}\sum_{j=1}^{n} Y_t^{u,j,n} \\ \qquad\qquad + f_t Z_t^u + \widehat{f}_t \frac{1}{n}\sum_{j=1}^{n} Z_t^{u,j,n} + g_t u_t\Bigg) dt \\ \qquad - \Bigg(h_t y_t^u + \widehat{h}_t \frac{1}{n}\sum_{j=1}^{n} y_t^{u,j,n} + k_t Y_t^u + \widehat{k}_t \frac{1}{n}\sum_{j=1}^{n} Y_t^{u,j,n} \\ \qquad\qquad + m_t Z_t^u + \widehat{m}_t \frac{1}{n}\sum_{j=1}^{n} Z_t^{u,j,n} + \widehat{g}_t u_t\Bigg) \overleftarrow{dB}_t^i + Z_t^u dW_t^i, \\ y_0^u = x, Y_T = \xi, \end{cases}$$

where (W^i), (B^i) are a collections of independent Brownian motions and $\frac{1}{n}\sum_{j=1}^{n} y_t^{u,j,n}$ denotes the empirical distribution of the individual players' state at time $t \in [0,T]$. Our system MF-FBDSDEs (3) occur naturally in the probabilistic analysis of financial optimization and control problems of the McKean-Vlasov type.

The subject 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 [16] derived the first existence result of an optimal relaxed control for SDEs with uncontrolled diffusion coefficient by using compactification techniques. For such systems of SDEs, a maximum principle has been established in Mezerdi and Bahlali [29]. The case of an SDE where the diffusion coefficient depends explicitly on the control variable has been solved by El-Karoui et al. [15], where the optimal relaxed control is shown to be Markovian.

Our third main goal in this thesis is to establish necessary as well as sufficient optimality conditions for both relaxed and strict control problems for systems driven by nonlinear mean-field forward-backward doubly stochastic differential equations.

Our contribution in this thesis touch on a very important aspect of optimal stochastic control which is the existence of optimal controls as well as the necessary and sufficient optimality conditions. We will present in what follows a brief description of the main results we have achieved in this thesis.

This thesis is organized as follows

Chapter 1:

This introductory chapter, we give some mathematical preliminaries, we provide the most important definitions and some specific tools to introduce the stochastic integral. In fact the main reason for including this material here is to introduce some specific tools which will be used systematically in later chapters.

Chapter 2: (The results of this chapter were a part of a paper [30] published in Random Operators & Stochastic Equations, 2020).

In this chapter, we deal with the problem of existence of optimal strict control of systems governed by linear forward-backward doubly stochastic differential equations, with random coefficients and non linear functional cost. The control domain and the cost function were assumed convex. The proof is based on strong convergence techniques for the associated linear FBDSDEs and Mazur's theorem. We derive also necessary and sufficient conditions for optimality for this strict control problem. This result is based on the convex optimization principle.

Chapter 3: (The results of this chapter were a part of a paper [9] published in Boletim da Sociedade Paranaense de Matemática 2020).

In this chapter, we prove the existence of a optimal strict control for a control problem governed by linear forward-backward doubly stochastic differential equations of mean-field type. The coefficients of the system depend on the states of the solution processes as well as their distribution via the expectation of the states. Moreover, the cost functional is also of mean-field type. We prove in particular, the existence of optimal strong control by using the strong convergence techniques for the associated linear MF-FBDSDEs and Mazur's theorem. We derive also necessary and sufficients conditions for optimality for this control problem of linear MF-FBDSDEs.

Chapter 4: (The results of this chapter were a part of a paper [9] published in Boletim da Sociedade Paranaense de Matemática 2020).

In this chapter, we establish necessary as well as sufficient optimality conditions for both relaxed and strict control problems governed by systems of nonlinear FBDSDEs of mean field type.

Chapter 1

Some Mathematical Preliminaries

Our aim in this chapter is to provide the most important definitions concerning stochastic calculus.

The organization of this chapeter is the following: in the first section, some introductory probability will be briefly reviewed, in the second section, we introduce the notion of filtration, stochastic process and Brownian motion, finally, the last section deals with the stochastic integral.

1.1 Probability

1.1.1 Probability spaces

Let Ω be a nonempty set and $\mathcal{F}$ be a collection of subsets of Ω.

Definition 1.1.1 *The sample space Ω of an experiment is the set of all possible outcomes.*

Definition 1.1.2 *We say that $\mathcal{F}$ is $\sigma-$algebra or $\sigma-$field if*

1. $\Omega \in \mathcal{F}$,
2. $(A \in \mathcal{F}) \Longrightarrow (A^c \in \mathcal{F})$,

3. $\left((A_n)_{n\in\mathbb{N}} \subset \mathcal{F}\right) \Longrightarrow \left(\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{F}\right)$.

If $\mathcal{F}$ and $\mathcal{G}$ are two σ-algebras on Ω and $\mathcal{G} \subseteq \mathcal{F}$, then $\mathcal{G}$ is called a sub σ- algebra of $\mathcal{F}$. The pair $(\Omega, \mathcal{F})$ is called a measurable space.

Exemple 1.1.1 *The following sets are always σ-fields*

1. $\mathcal{F}_0 = \{\phi, \Omega\}$ *(trivial σ-field),*

2. $\mathcal{P}(\Omega) = \{$*all subsets of* $\Omega\}$ *,(complete σ-field).*

Let $\{\mathcal{F}_n\}$ be a family of σ-fields on Ω. We have
$\underset{n}{\vee}\mathcal{F}_n = \sigma\left(\underset{n}{\cup}\mathcal{F}_n\right)$ is the smallest σ-field containing all $\mathcal{F}_n$.
$\underset{n}{\wedge}\mathcal{F}_n = \sigma\left(\underset{n}{\cap}\mathcal{F}_n\right)$ .is the largest σ-field contained in all $\mathcal{F}_n$.

Definition 1.1.3 *Let Ω be a topological space, then the smallest σ-field containing all open sets of Ω is called the Borel σ-algebra of Ω, denoted by $\mathcal{B}(\Omega)$.*

For example, if the collection of all open subsets of a topological space $\mathbb{R}^m$, then $\mathcal{B}(\mathbb{R}^m)$ is called the Borel σ-field on Ω and the elements $A \in \mathcal{B}(\mathbb{R}^m)$ are called Borel sets. $\mathcal{B}(\mathbb{R}^m)$ contains all open sets, all closed sets, all countable unions of closed sets, all countable intersections of such countable unions etc.

Definition 1.1.4 *Let $(\Omega, \mathcal{F})$ be a measurable space. A measure μ on $\mathcal{F}$ is a function*

$$\mu : \mathcal{F} \to [0; \infty],$$

with the following properties:

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

2. *if the family $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ is disjoint ($A_i \cap A_j = \phi$ if $i \neq j$), then*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\geq 0} \mu(A_n).$$

We say that μ is a probability measure, if $\mu(\Omega) = 1$, in this case we write $\mathbb{P}$ instead of μ, the triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space. The subsets A of Ω which belong to $\mathcal{F}$ are called $\mathcal{F}$-measurable sets. In a probability context these sets are called events and we use the interpretation

$$\mathbb{P}(A) = \text{"the probability that the event } A \text{ occurs"}.$$

In particular, if $\mathbb{P}(A) = 1$ we say that "A occurs with probability 1", or "almost surely (a.s.)".

Theorem 1.1.1 (Product Measure) *Let $(E, \mathcal{E}, \mu)$ and $(E', \mathcal{E}', \mu')$ be two σ-finite measure spaces. There exists a unique measure $\hat{\mu} = \mu \otimes \mu'$ on ξ such that*

$$\hat{\mu}(A \times A') = \mu(A)\mu'(A'),$$

for all $A \in \mathcal{E}, A' \in \mathcal{E}'$.

Theorem 1.1.2 (Fubini's Theorem) *Let $(E, \mathcal{E}, \mu)$ and $(E', \mathcal{E}', \mu')$ be two σ-finite measure spaces. Let f be $\mathcal{E}$-measurable and non-negative. Then*

$$\begin{aligned} \int_{E \times E'} f(x, x') d(\mu \otimes \mu') &= \int_E \left(\int_{E'} f(x, x') d\mu' \right) d\mu \\ &= \int_{E'} \left(\int_E f(x, x') d\mu \right) d\mu'. \end{aligned}$$

If f is integrable, then

1. *$x' \to f(x, x')$ is $\mu'-$integrable for $\mu-$almost all x,*
2. *$x \to \int_{E'} f(x, x') d\mu'$ is μ-integrable and the above equality holds.*

Definition 1.1.5 *A family of events* $\{A_i,\ i \in I\}$ *is independent if*

$$\mathbb{P}(\bigcap_{i\in J} A_i) = \prod_{i\in J} \mathbb{P}(A_i),$$

for all finite subsets J *of* I.

Definition 1.1.6 *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space*

1. *An event* A *is independent of a* σ*-field* $\mathcal{F}$ *if* A *is independent of any* $B \in \mathcal{F}$.

2. *Two* σ*-algebras* $\mathcal{F}_1$ *and* $\mathcal{F}_2$ *are independent if any event* $A \in \mathcal{F}_1$ *is independent of* $\mathcal{F}_2$.

Definition 1.1.7 *An event* A *is said a* $\mathbb{P}$*-null event if* $\mathbb{P}(A) = 0$.

A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be complete if for any $\mathbb{P}$-null set $A \in \mathcal{F}$, one has $B \in \mathcal{F}$ whenever $B \subseteq A$ (thus, it is necessary that B is also a $\mathbb{P}$-null set).
For any given probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we define

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

and $\hat{\mathcal{F}} := \mathcal{F} \vee \mathcal{N}$. Then for any $\hat{A} \in \hat{\mathcal{F}}$, there exist $A, B \in \mathcal{F}$ such that $\mathbb{P}(B) = 0$ and $\hat{A} \backslash A \subseteq B$. In such a case, we define $\mathbb{P}(\hat{A}) = \mathbb{P}(A)$. This extends $\mathbb{P}$ to $\hat{\mathcal{F}}$. Clearly, $(\Omega, \hat{\mathcal{F}}, \mathbb{P})$ is a complete probability space. Any probability space can be made complete by the augmentation procedure.

1.1.2 Random variable

Definition 1.1.8 *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, then a function* $X : \Omega \to \mathbb{R}^m$ *is called* $\mathcal{F}$*-measurable if*

$$X^{-1}(A) := \{\omega \in \Omega;\ X(\omega) \in A\} \in \mathcal{F},$$

for all open sets $A \subset \mathbb{R}^m$.

Definition 1.1.9 *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, a random variable* $X : \Omega \to \mathbb{R}^m$*, is an* $\mathcal{F}$*-measurable function.*

More general, if $(\Omega, \mathcal{F})$ *and* $(\Omega', \mathcal{F}')$ *are two measurable spaces and* $X : \Omega \to \Omega'$ *is an* $(\mathcal{F}/\mathcal{F}')$*-measurable map. We call* X *an* $(\mathcal{F}/\mathcal{F}')$*-random variable.*

If $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ are two measurable spaces and $X : \Omega \to \Omega'$ is a random variable, then $X^{-1}(\mathcal{F}')$ is a sub σ-algebra of $\mathcal{F}$, which is called the σ-algebra generated by X and denoted by $\sigma(X)$. This is the smallest σ-field in Ω under which X is measurable. Also, if $\{X_\eta,\ \eta \in \Xi\}$ is a family of random variables from Ω to Ω', then we denote by

$$\sigma(X_\eta,\ \eta \in \Xi) = \bigvee_{\eta \in \Xi} X_\eta^{-1}(\mathcal{F}'),$$

the smallest sub σ-field of $\mathcal{F}$ under which all X_η, $(\eta \in \Xi)$ are measurable.

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

Next, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(\Omega', \mathcal{F}')$ a measurable space and $X : \Omega \to \Omega'$ a random variable. Then X induces a probability measure M_X, defined by

$$\begin{aligned} \mathrm{M}_X(B) &:= \mathbb{P}\circ X^{-1}(B') = \mathbb{P}\left(X^{-1}(B')\right), \\ &= \mathbb{P}\{\omega \in \Omega / \ X(\omega) \in B'\} = \mathbb{P}\{\ X \in B'\},\ \forall B' \in \mathcal{F}'. \end{aligned}$$

M_X is called the distribution of X. In the case where $\Omega' = \mathbb{R}^m$, M_X can be uniquely determined by the following function

$$F(x_1, ..., x_m) := \mathbb{P}\{\omega \in \Omega / \ X_i(\omega) \le x_i,\ 1 \le i \le m\}.$$

We call $F(x)$ the distribution cumulative function of X, it is nonnegative, nondecreasing in each variable $x_i \in \mathbb{R}$ and

$$\lim_{\substack{x_i \to -\infty \\ \exists i}} F(x) = 0, \quad \lim_{\substack{x_i \to \infty \\ \forall x_i}} F(x) = 1.$$

If $\int_\Omega |X(\omega)| \, d\mathbb{P}(\omega) < \infty$, then the number

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

is called the expectation of X.

If $g : \mathbb{R}^m \to \mathbb{R}$ is Borel measurable function and if $\int_\Omega |g(X(\omega))| \, d\mathbb{P}(\omega) < \infty$, then we have

$$\mathbb{E}[g(X)] := \int_\Omega g(X(\omega)) \, d\mathbb{P}(\omega).$$

Let $L^p_{\mathcal{F}}(\Omega, \mathbb{R}^m) := L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^m)$ be the set of all random variables $X : \Omega \to \mathbb{R}^m$ with $|X|^p \in L^1_{\mathcal{F}}(\Omega, \mathbb{R}^m)$. This is a Banach space with the norm

$$(\mathbb{E}[|X|^p])^{\frac{1}{p}} = \left(\int_\Omega |X(\omega)|^p \, d\mathbb{P}(\omega) \right)^{\frac{1}{p}}.$$

In particular, if $p = 2$, then $L^2_{\mathcal{F}}(\Omega, \mathbb{R}^m)$ is a Hilbert space with the inner product

$$\langle X, Y \rangle_{L^2_{\mathcal{F}}(\Omega, \mathbb{R}^m)} := \int_\Omega \langle X(\omega), Y(\omega) \rangle \, d\mathbb{P}(\omega).$$

If two random variables X, $Y : \Omega \to \mathbb{R}$ are independent then

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

provided that $\mathbb{E}[|X|] < \infty$ and $\mathbb{E}[|Y|] < \infty$.

1.1.3 Modes of Convergence

Definition 1.1.10 *Let X, $X_1, X_2, \ldots$ be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Then*

1. *$X_n \to X$* ***almost surely****, if*

$$A = \{\omega \in \Omega : X_n(\omega) \to X(\omega) \text{ as } n \to \infty\} \in \mathcal{F} \text{ with } \mathbb{P}(A) = 1.$$

2. *$X_n \to X$* ***in probability,*** *if*

$$\mathbb{P}(|X_n - X| > \varepsilon) \to 0 \text{ as } n \to \infty \text{ for all } \varepsilon > 0.$$

3. *$X_n \to X$* ***in*** *$L^p_{\mathcal{F}}(\Omega, \mathbb{R}^m)$, if X_n, $X \in L^p_{\mathcal{F}}(\Omega, \mathbb{R}^m)$ and*

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

4. *$X_n \to X$* ***in distribution****, if*

$$\mathbb{P}(X_n \leq x) \to \mathbb{P}(X \leq x) \text{ as } n \to \infty,$$

for all x at which $F(x) = \mathbb{P}(X \leq x)$ is continuous.

Theorem 1.1.3 *Almost sure convergence $\Longrightarrow$ convergence in probability $\Longrightarrow$ convergence in distribution.*

Convergence of probabilities

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

Definition 1.1.11 *A sequence* $\{q_n\} \subseteq \mathbb{P}(U)$ *is said to be weakly convergent to* $q \in \mathbb{P}(U)$ *if for any* $f \in C_b(U)$,

$$\lim_{n\to\infty} \int_U f(x)\, dq_n(x) = \int_U f(x)\, dq(x).$$

Definition 1.1.12 *Let* $X_n : (\Omega_n, \mathcal{F}_n, \mathbb{P}_n) \to (U, d)$, $n = 1, 2, ...,$ *and* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (U, d)$ *be a random variables. We say that* X_n *converges to* X *in law if* $\mathrm{M}_{X_n} \to \mathrm{M}_X$ *weakly as* $n \to \infty$.

Definition 1.1.13 *A set* $K \subseteq \mathbb{P}(U)$ *is said to be*

i *relatively compact if any sequence* $\{q_n\} \subseteq K$ *contains a weakly convergent subsequence,*

ii *compact if* K *is relatively compact and closed.*

Corollary 1.1.1 *If* (U, d) *is compact, then any* $K \subseteq \mathbb{P}(U)$ *is relatively compact. In particular,* $\mathbb{P}(U)$ *is compact.*

1.1.4 Conditional expectation

In this subsection, we present the notion of conditional expectation and its main properties. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$ and $X \in L^1_{\mathcal{F}}(\Omega, \mathbb{R}^m)$.

Definition 1.1.14 *We say that* Y *is the conditional expectation of* X *with respect to* $\mathcal{G}$, *and denote it by* $\mathbb{E}[X \mid \mathcal{G}]$, *if the following two conditions hold:*

1. Y *is* $\mathcal{G}$*-measurable,*
2. $\int_\Lambda X d\mathbb{P} = \int_\Lambda Y d\mathbb{P}$ *for all* $\Lambda \in \mathcal{G}$.

It is worth noting that the expectation of X, denoted by $\mathbb{E}[X]$ is a number, while the conditional expectation $\mathbb{E}[X \mid \mathcal{G}]$ is a random variable.

Remark 1.1.1 *1. **Existence:** There is always a random variable Y satisfying the above properties (provided that $\mathbb{E}[X] < \infty$), i.e., conditional expectations always exist.*

*2. **Uniqueness**: There can be more than one random variable Y satisfying the above properties, but if Y' is another one, then $Y = Y'$ almost surely, i.e.,*

$$\mathbb{P}(\{\omega \in \Omega / \ Y(\omega) = Y'(\omega)\}) = 1.$$

Main properties

Let us collect some basic properties of the conditional expectation.

Proposition 1.1.1 *Let X and Y be two random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ taking their values in $\mathbb{R}$.*

1. *If X is $\mathcal{G}$-measurable, then*

$$\mathbb{E}[X \mid \mathcal{G}] = X \ a.s.$$

2. *If X and $\mathcal{G}$ are independent, then*

$$\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X] \ \ a.s.$$

3. *If Y is $\mathcal{G}$-measurable and $\mathbb{E}[XY] < \infty$, then*

$$\mathbb{E}[XY \mid \mathcal{G}] = Y\mathbb{E}[X \mid \mathcal{G}] \ \ a.s.$$

4. *If $\mathcal{H}$ is a sub-σ-algebra of $\mathcal{G}$, then*

$$\mathbb{E}[X \mid \mathcal{H}] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] \ \ a.s.$$

5. *If* $X, Y \in L^1_{\mathcal{F}}(\Omega, \mathbb{R})$ *and* $X \leq Y$*, then*

$$\mathbb{E}[X \mid \mathcal{G}] \leq \mathbb{E}[Y \mid \mathcal{G}] \ \ a.s.$$

6. *We have*

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

7. *Linearity:*

$$\mathbb{E}[\alpha X + \beta Y \mid \mathcal{G}] = \alpha \mathbb{E}[X \mid \mathcal{G}] + \beta \mathbb{E}[Y \mid \mathcal{G}], \ \forall \alpha, \ \beta \in \mathbb{R}.$$

8. *Conditional Lebesgue Dominated Convergence theorem: if* $X = \lim_{n\to\infty} X_n$ *a.s. and* $|X_n| \leq Y$ *for some integrable random variable* Y*, then*

$$\mathbb{E}[X \mid \mathcal{G}] = \lim_{n\to\infty} \mathbb{E}[X_n \mid \mathcal{G}].$$

Proposition 1.1.2 ***(Jensen's Inequality)*** *Let* $X \in L^1_{\mathcal{F}}(\Omega, \mathbb{R}^m)$ *and* $\varphi : \mathbb{R}^m \to \mathbb{R}$ *be a convex function such that* $\varphi(X) \in L^1_{\mathcal{F}}(\Omega, \mathbb{R}^m)$*. Then*

$$\varphi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\varphi(X) \mid \mathcal{G}], \ \ a.s.$$

In particular, for any $p \geq 1$*, provided that* $\mathbb{E}[|X|^p]$ *exists, we have*

$$|\mathbb{E}[X \mid \mathcal{G}]|^p \leq \mathbb{E}[|X|^p \mid \mathcal{G}], \ \ a.s.,$$

for any σ*-algebra* $\mathcal{G}$ *on* Ω *contained in* $\mathcal{F}$*.*

1.2 Stochastic Processes

In this section, we present some general notions that will be of constant use later.

Definition 1.2.1 *A filtration on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a collection* $(\mathcal{F}_t)_{t\geq 0}$ *of sub* σ *-fields of* $\mathcal{F}$ *which is increasing, such that*

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F} \text{ for every } s \leq t.$$

We also say that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ *is a filtered probability space.*

$\mathcal{F}_t$ is interpreted as the information known at time t, and increases as time elapses.

The canonical filtration of X is the smallest $\sigma-$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 $t \geq 0$.

Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$. We set, for every $t \in [0, T)$

$$\mathcal{F}_{t^+} = \bigcap_{s>t} \mathcal{F}_s.$$

We say that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous if

$$\mathcal{F}_t = \mathcal{F}_{t^+}, \quad \text{for any } t \in [0, T).$$

Definition 1.2.2 *We say that* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ *satisfies the usual condition if* $(\Omega, \mathcal{F}, \mathbb{P})$ *is complete,* $\mathcal{F}_0$ *contains all the* $\mathbb{P}-$*null sets in* $\mathcal{F}$*, and* $\{\mathcal{F}\}_{t\geq 0}$ *is right continuous.*

Let us turn to random processes, we define stochastic processes in general and give some results.

Definition 1.2.3 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A family $(X_t)_{t\in I}$, $I \subset \mathbb{R}$ of functions from $\Omega \times I$ into $\mathbb{R}^m$ is called a stochastic process. Note that for each $t \in I$ fixed we have a random variable, $\omega \to X_t(\omega)$, $\omega \in \Omega$. On the other hand, fixing $\omega \in \Omega$ we can consider the function $t \to X_t(\omega)$, $t \in I$, which is called a path of X.*

We shall interchangeably use $\{X_t,\ t \in I\}$, $(X_t)_{t\in I}$, X_t, or even X to denote a stochastic process.

For two stochastic processes X and Y, there exist different concepts of equality.

Definition 1.2.4 *Let X and Y be stochastic processes. Then X and Y are*

1. *equivalent if they have the same finite dimensional distributions,*
2. *modifications if $\mathbb{P}[X_t = Y_t] = 1$, for every $t \geq 0$,*
3. *indistinguishable if $\mathbb{P}[X_t = Y_t$, for every $t \geq 0] = 1$.*

Definition 1.2.5 *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0})$ be a filtered measurable space and X_t a process taking values in a metric space (U, d).*

i *The process X_t is said to be measurable if the map $(t,\omega) \longmapsto X_t(\omega)$ is $(\mathcal{B}[0,T] \times \mathcal{F}) / \mathcal{B}(U)$-measurable.*

ii *The process X_t is said to be $\{\mathcal{F}_t\}_{t\geq 0}$ $-$adapted if for all $t \in [0,T]$, the map $\omega \longmapsto X(t,\omega)$ is $\mathcal{F}_t / \mathcal{B}(U)$-measurable.*

iii *The process X_t is $\{\mathcal{F}_t\}_{t\geq 0}$ $-$progressively measurable if for all $t \in [0,T]$, the map $(s,\omega) \longmapsto X_s(\omega)$ is $\mathcal{B}[0,t] \times \mathcal{F} / \mathcal{B}(U)$-measurable.*

Note that a process progressively measurable is both adapted and measurable.

Definition 1.2.6 *Let X be a stochastic process.*

1. *The process X is said to be (a.s.) continuous if (almost) all its trajectories are continuous.*

2. *The process X is said to be (a.s.) càdlàg if (almost) all its trajectories are càdlàg.*

3. *The process X is said to be stochastically continuous (or continuous in probability) if*

$$\lim_{s \to t} \mathbb{P}\left[|X_t - X_s| > \varepsilon\right] = 0, \quad \text{for every } t \geq 0 \text{ and } \varepsilon > 0.$$

Remark 1.2.1 *We remark that the word càdlàg is the abbreviation of French "continue a droite, limité à gauche". This means that the paths of the process X are right-continuous and admit finite left-limit in every point, that is*

$$\lim_{s \searrow t} X_s = X_t, \quad \lim_{s \nearrow t} X_s \text{ exists and is finite for every } t > 0.$$

Definition 1.2.7 *We call infinitesimal variation of order p of an associated process X of a subdivision $\Delta_n = (t_1^n < ... < t_n^n)$ of $[0, T]$*

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

if $V_T^p(X_t)$ admits a limit when $\|\Delta_n\| \longrightarrow 0$ as $n \longrightarrow \infty$ 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 total variation of X.
If $p = 2$ the limit is called quadratic variation and we denote by $\langle X, X\rangle_T$.

We shall define next an important type of stochastic process.

Definition 1.2.8 *A process X is called a martingale with respect to the filtration $\{\mathcal{F}_t,\ t \geq 0\}$ if*

1. *X_t is integrable for each $t \geq 0$,*

2. $(X_t)_{t\geq 0}$ *is adapted to the filtration* $(\mathcal{F}_t)_{t\geq 0}$ *,*

3. $X_s = \mathbb{E}[X_t \mid \mathcal{F}_s]$*,* $\forall s \leq t$*.*

Remark 1.2.2 *The first condition states that the unconditional forecast is finite* $\mathbb{E}[|X_t|] < \infty$*. Condition* 2 *says that the value* X_t *is known, given the information set* $\mathcal{F}_t$*. This can be also stated by saying that* X_t *is* $\mathcal{F}_t$*-predictable. The third relation asserts that the best forecast of unobserved future values is the last observation on* X_t*.*

Remark 1.2.3 *If the third condition is replaced by*

3' $X_s \leq \mathbb{E}[X_t \mid \mathcal{F}_s]$*,* $\forall s \leq t$*,*

then X_t *is called a submartingale, and if it is replaced by*

3" $X_s \geq \mathbb{E}[X_t \mid \mathcal{F}_s]$*,* $\forall s \leq t$*,*

then X_t *is called a supermartingale.*

It is worth noting that X_t *is a submartingale if and only if* $(-X_t)$ *is a supermartingale.*

Proposition 1.2.1 *Let* $\{\mathcal{F}_t\}_{t\geq 0}$ *and* $\{\mathcal{G}_t\}_{t\geq 0}$ *be two families of sub* σ *-fields of* $\mathcal{F}$ *with* $\mathcal{G}_t \subset \mathcal{F}_t$*,* $\forall t \geq 0$*. If* X_t *is an* $\{\mathcal{F}_t\}_{t\geq 0}$ *−martingale (respectively submartingale, supermartingale), then* $Y_t := \mathbb{E}[X_t \mid \mathcal{G}_t]$ *is a* $\{\mathcal{G}_t\}_{t\geq 0}$ *−martingale (respectively submartingale, supermartingale). In particular, if* X_t *is* $\{\mathcal{G}_t\}_{t\geq 0}$ *−adapted, then* X_t *itself is a* $\{\mathcal{G}_t\}_{t\geq 0}$ *−martingale (respectively submartingale, supermartingale).*

1.2.1 The Brownian Motion

The observation made first by the botanist Robert Brown in 1827, that small pollen grains suspended in water have a very irregular and unpredictable state of motion, led to the definition of the Brownian motion, which is formalized in the following.

Definition 1.2.9 *A Brownian motion process is a stochastic process* B*, which satisfies*

1. *The process starts at the origin,* $B_0 = 0$,
2. B_t *has stationary, independent increments,*
3. *the process* B_t *is continuous in* t,
4. *the increments* $B_t - B_s$ *are normally distributed with mean zero and variance* $t - s$,

$$B_t - B_s \sim \mathcal{N}(0, t-s).$$

It is worth noting that even if B_t is continuous, it is nowhere differentiable. From condition 4 we get that B_t is normally distributed with mean $\mathbb{E}[B_t] = 0$ and $Var[B_t] = t$,

$$B_t \sim \mathcal{N}(0, t).$$

The process $X_t = x + B_t$ has all the properties of a Brownian motion that starts at x.

Main properties

Proposition 1.2.2 *Let* $(B_t)_{t\geq 0}$ *be a Brownian motion with respect to* $(\mathcal{F}_t)_{t\geq 0}$

1. ***Translation Invariance:*** *for fixed* $t_0 \geq 0$ *the stochastic process* $(B_{t+t_0} - B_{t_0})_{t\geq 0}$ *is a Brownian motion.*
2. ***Scaling Invariance:*** *for* $\beta > 0$. *Then the process* $(X_t)_{t\geq 0}$ *where* $X_t := \frac{1}{\beta} B_{\beta^2 t}$, $t \geq 0$ *is also a standard Brownian motion.*

Remark 1.2.4 *The scaling invariance property (with* $\beta = -1$*) implies that standard Brownian motion is symmetric about* 0. *In other words, if* $(B_t)_{t\geq 0}$ *is a standard Brownian motion and* $t \geq 0$, *then* B_t *has the same distribution as* $-B_t$.

Proposition 1.2.3 *Let* B *be a Brownian motion and* $\mathcal{F}_s = \sigma\{B_r \ / \ 0 \leq r \leq s\}$, $s \geq 0$. *Then*

1. *for all* $t > s$, $B_t - B_s$ *is independent of* $\mathcal{F}_s$,

2. $\mathbb{E}[B_s B_t] = \min(s, t)$.

Proof. 1. Since B has independent increments, then $B_t - B_s$ is independent of $B_s - B_0 = B_s$. Also $B_t - B_s$ is independent of $B_s - B_r$ for all $0 \leq r \leq s$. Therefore $B_t - B_s$ is independent of $\sigma(B_s)$ and $\sigma(B_s - B_r)$ and so of $\sigma(B_s) \vee \sigma(B_s - B_r)$ for all $r \leq s$.
On the other hand, note that $B_r = -(B_s - B_r) + B_s$. Thus B_r is $\sigma(B_s) \vee \sigma(B_s - B_r)-$measurable for all $0 \leq r \leq s$. Consequently from above, we deduce that $B_t - B_s$ is independent of $\sigma(B_r)$ for all $0 \leq r \leq s$, which means that $B_t - B_s$ is independent of $\mathcal{F}_s = \sigma\{B_r \;/\; 0 \leq r \leq s\}$ for all $0 \leq s \leq t$.
2. If $t \geq s$

$$\mathbb{E}[B_s B_t] = \mathbb{E}[(B_s - B_0)(B_t - B_0) + B_s^2] = \mathbb{E}[B_s - B_0]\,\mathbb{E}[B_t - B_0] + \mathbb{E}\left[B_s^2\right] = s,$$

by using the independence of $B_t - B_s$ and B_s. On the other hand, by symmetry we deduce that $\mathbb{E}[B_s B_t] = t$ if $s \geq t$. The proof is completed ■

Proposition 1.2.4 *If* $f : \mathbb{R} \to \mathbb{R}$ *is measurable, then*

$$\mathbb{E}[f(B_t)] = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} f(u)\, e^{\frac{-u^2}{2t}} du.$$

Exemple 1.2.1 *In this example let us compute* $\mathbb{E}[|B_t|]$ *by applying Proposition 1.2.4 with* $f(u) = |u|$. *We get*

$$\begin{aligned}\mathbb{E}[|B_t|] &= \tfrac{1}{\sqrt{2\pi t}} \textstyle\int_{\mathbb{R}} |u|\, e^{\frac{-u^2}{2t}} du \\ &= \tfrac{2}{\sqrt{2\pi t}} \left[\textstyle\int_0^\infty u e^{\frac{-u^2}{2t}} du\right] \\ &= \sqrt{\tfrac{2t}{\pi}} \textstyle\int_0^\infty e^{-v} dv \\ &= \sqrt{\tfrac{2t}{\pi}},\end{aligned}$$

using the variable change $v = \frac{u^2}{2t}$.

Proposition 1.2.5 *i) The quadratic variation of a Brownian motion on* $[t,T]$ *converges as a quadratic mean to* $T-t$, $\forall t \in \mathbb{R}$ *and if* $(\Delta_n)_{n\geq 0}$ *is a sequence of subdivisions of* $[t,T]$, *where* $\|\Delta_n\| \longrightarrow 0$ *as* $n \longrightarrow \infty$.

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

$$V_T^2(B_t) \longrightarrow T.$$

Proposition 1.2.6 *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* $(B_t)_{t\geq 0}$ *be a Brownian motion. Then*

I) $(B_t)_{t\geq 0}$ *is a martingale.*

II) $X_t = B_t^2 - t\ \forall t \geq 0$, *is a martingale.*

Proof. I) Let $\mathcal{F}_t = \sigma\left(B_s \ / \ 0 \leq s \leq t\right)$. It is obvious that B_t is $\sigma(B_t)$-measurable $t \geq 0$, and so B_t is $\mathcal{F}_t$-measurable $\forall t \geq 0$. Secondly, from Hölder's inequality we get

$$\mathbb{E}[|B_t|^2] \leq \sqrt{\mathbb{E}[B_t^2]} = \sqrt{t} < \infty,$$

for all $t \geq 0$, showing that B_t is integrable.

Let $s \leq t$ and write $B_t = B_s + (B_t - B_s)$. Then

$$\begin{aligned}
\mathbb{E}[B_t \mid \mathcal{F}_s] &= \mathbb{E}[B_s + (B_t - B_s) \mid \mathcal{F}_s] \\
&= \mathbb{E}[B_s \mid \mathcal{F}_s] + \mathbb{E}\left[(B_t - B_s) \mid \mathcal{F}_s\right] \\
&= B_s + \mathbb{E}\left[B_t - B_s\right] \\
&= B_s + 0,
\end{aligned}$$

for any $s \leq t$ and this shows that B is a martingale.

II) 1. Since $X_t = B_t^2 - t$ is a function of B_t, hence it is $\mathcal{F}_t$-measurable $\forall t \geq 0$, implying that X_t is $\{\mathcal{F}_t\}_{t\geq 0}$-adapted.

2. Since $|X_t| = |B_t^2 - t| \leq B_t^2 + t$ we can therefore write

$$\mathbb{E}\left[|X_t|\right] = \mathbb{E}\left[\left|B_t^2 - t\right|\right] \leq \mathbb{E}\left[B_t^2 + t\right] = 2t < \infty, \ \forall t \geq 0.$$

3. $(B_t - B_s)$ is independent of $\mathcal{F}_s$ for $s \leq t$, we have

$$\begin{aligned}
\mathbb{E}\left[(B_t^2 - t) \mid \mathcal{F}_s\right] &= \mathbb{E}\left[(B_t - B_s + B_t)^2 \mid \mathcal{F}_s\right] - t \\
&= \mathbb{E}\left[(B_t - B_s)^2 \mid \mathcal{F}_s\right] + 2\mathbb{E}\left[B_s\left(B_t - B_s\right) \mid \mathcal{F}_s\right] + \mathbb{E}\left[B_s^2 \mid \mathcal{F}_s\right] - t \\
&= t - s + 0 + B_s^2 - t \\
&= B_s^2 - s.
\end{aligned}$$

The proposition is proved. ■

1.3 Stochastic Integral

The Itô integral is defined in a way that is similar to the Riemann integral. The Itô integral is taken with respect to infinitesimal increments of a Brownian motion, dB_t, which are random variables, while the Riemann integral considers integration with respect to the predictable infinitesimal changes dt. It is worth noting that the Itô integral is a random variable, while the Riemann integral is just a real number.

In this section we will briefly review the definition and some properties of the stochastic integral.

1.3.1 Construction of Itô's Integral

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a fixed filtered probability space satisfying the usual condition. Let $T > 0$ and recall that $L^2_{\mathcal{F}}(0, T, \mathbb{R})$ the space of all stochastic processes $F_t(\omega)$, $0 \leq t \leq T$, $\omega \in \Omega$, satisfying the following conditions

1. F_t is adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$,

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

$L^2_{\mathcal{F}}\left(0, T, \mathbb{R}\right)$ is a Hilbert space with norm

$$\|F\| = \langle F, F\rangle^{\frac{1}{2}} = \sqrt{\mathbb{E}\left[\int_0^T F_t^2 dt\right]}.$$

We want to define the stochastic integral

$$\int_0^T F_t dB_t,$$

for elements F of $L^2_{\mathcal{F}}\left(0, T, \mathbb{R}\right)$.

We start with a definition for a simple class of functions F.

Definition of the Itô's Integral for Step Functions (step1)

Divide the interval $[0, T]$ into n subintervals using the partition points

$$0 = t_0 < t_1 < ... < t_{n-1} < t_n = T,$$

suppose F is a step stochastic process given by

$$F = \sum_{k=1}^{n} f_{k-1} 1_{[t_{k-1}, t_k]},$$

where f_{k-1} is $\mathcal{F}_{k-1}$-measurble and $\mathbb{E}\left[f_{k-1}\right]^2 < \infty$. We define the following linear operator

$$I\left(F\right) = \sum_{k=1}^{n} f_{k-1}\left(B_{t_k} - B_{t_{k-1}}\right).$$

Lemma 1.3.1 *Let $I\left(F\right)$ be a linear random variable with mean $\mathbb{E}\left[I\left(F\right)\right] = 0$, and variance*

$$\mathbb{E}\left[\left|I\left(F\right)\right|^2\right] = \int_0^T \mathbb{E}\left[\left|F\right|^2\right] dt.$$

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

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

Hence $\mathbb{E}\left[I\left(F\right)\right]=0$. Moreover, we have

$$\left|I\left(F\right)\right|^2=\sum_{k,l=1}^{n} f_{k-1}f_{l-1}\left(B_{t_k}-B_{t_{k-1}}\right)\left(B_{t_l}-B_{t_{l-1}}\right).$$

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

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

On the other hand, for $k=l$ we have from the independence of $B_{t_k}-B_{t_{k-1}}$ of $\mathcal{F}_{t_{k-1}}$,

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

So, we get

$$\mathbb{E}\left[|I(F)|^2\right] = \sum_{k=1}^{n} (t_k - t_{k-1})\, \mathbb{E}\left[f_{k-1}^2\right].$$

The proof is completed. ■

An approximation lemma (step2)

Lemma 1.3.2 *Suppose that* $F \in L^2_{\mathcal{F}}(0, T, \mathbb{R})$. *Then there exists a sequence* $\{F_n,\ n \geq 1\}$ *of step processes in* $L^2_{\mathcal{F}}(0, T, \mathbb{R})$ *such that*

$$\lim_{n\to\infty} \int_0^T \mathbb{E}\left[|F_t - F_n|^2\right] dt = 0. \tag{1.1}$$

Now we define the stochastic integral by using what we proved in (Step1) and (Step2)

$$\int_0^T F_t dB_t,$$

for $F \in L^2_{\mathcal{F}}(0, T, \mathbb{R})$. Apply first Lemma 1.3.2 to get a sequence $\{F_n,\ n \geq 1\}$ of adapted step stochastic processes such that (1.1) holds.

For each n, $I(F_n)$ is defined by (Step1). By Lemma 1.3.1 we have

$$\mathbb{E}\left[|I(F_n) - I(F_m)|^2\right] = \int_0^T \mathbb{E}\left[|F_n - F_m|^2\right] dt \to 0,\ \ as\ n, m \to \infty.$$

It follows that $\{I(F_n)\}$ is a Cauchy sequence in $L^2_{\mathcal{F}}(0, T, \mathbb{R})$. Thus $\{I(F_n)\}$ has a unique limit in $L^2_{\mathcal{F}}(0, T, \mathbb{R})$, denoted by $I(F)$, it is called the Itô integral, so

$$I(F) = \int_0^T F_t dB_t.$$

The integral is independent of the choice of the sequence $\{F_n,\ n \geq 1\}$.

Properties of Itô's Integral

1. Linearity: let $\alpha, \beta \in \mathbb{R}$ and $F, G \in L^2_{\mathcal{F}}(0,T,\mathbb{R})$. Then $\alpha F + \beta G \in L^2_{\mathcal{F}}(0,T,\mathbb{R})$ and

$$\int_0^T (\alpha F_t + \beta G_t)\, dB_t = \alpha \int_0^T F_t dB_t + \beta \int_0^T G_t dB_t.$$

2. Partition property

$$\int_0^T F_t dB_t = \int_0^c G_t dB_t + \int_c^T F_t dB_t,\ \forall\ 0 < c < T.$$

3. Zero mean

$$\mathbb{E}\left[\int_0^T F_t dB_t\right] = 0.$$

4. Isometry

$$\mathbb{E}\left[\left(\int_0^T F_t dB_t\right)^2\right] = \int_0^T \mathbb{E}\left[F_t^2\right] dt.$$

5. Product Property

$$\mathbb{E}\left[\left(\int_0^T F_t dB_t\right)\left(\int_0^T G_t dB_t\right)\right] = \mathbb{E}\left[\int_0^T F_t G_t dt\right].$$

Theorem 1.3.1 ***(Martingale Property)*** *Suppose $F \in L^2_{\mathcal{F}}(0,T,\mathbb{R})$. Then the stochastic process*

$$X_t = \int_0^t F_s dB_s,\ 0 \le t \le T,$$

is a martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\ge 0}$.

Examples of Itô integrals

Exemple 1.3.1 ***The case $F_t = c$, constant.***

In this case the partial sums can be computed explicitly

$$\begin{aligned}&\textstyle\sum_{k=1}^{n} f_{k-1}\left(B_{t_k}-B_{t_{k-1}}\right)\\&=\textstyle\sum_{k=1}^{n} c\left(B_{t_k}-B_{t_{k-1}}\right)\\&=c\left(B_T-B_0\right),\end{aligned}$$

and since the answer does not depend on n, we have

$$\int_0^T c dB_t = c\left(B_T-B_0\right).$$

In particular, taking $c=1$, we have the following formula

$$\int_0^T dB_t = B_T.$$

Exemple 1.3.2 ***The case*** $F_t = B_t$.

We shall integrate the process B_t between 0 and T. Considering an equidistant partition, we take $t_k = \frac{(k-1)T}{(n-1)}$, $k=1,...,n$. The partial sums are given by

$$I\left(F\right) = \textstyle\sum_{k=1}^{n} B_{t_{k-1}}\left(B_{t_k}-B_{t_{k-1}}\right).$$

Since

$$xy=\frac{1}{2}\left[(x+y)^2-x^2-y^2\right],$$

letting $x=B_{t_{k-1}}$ and $y=B_{t_k}-B_{t_{k-1}}$ yields

$$B_{t_{k-1}}\left(B_{t_k}-B_{t_{k-1}}\right)=\frac{1}{2}\left(B_{t_k}\right)^2-\frac{1}{2}\left(B_{t_{k-1}}\right)^2-\frac{1}{2}\left(B_{t_k}-B_{t_{k-1}}\right)^2.$$

Then after pair cancelations the sum becomes

$$I\left(F\right)=\frac{1}{2}\sum_{k=1}^{n}\left(B_{t_k}\right)^2-\frac{1}{2}\sum_{k=1}^{n}\left(B_{t_{k-1}}\right)^2-\frac{1}{2}\sum_{k=1}^{n}\left(B_{t_k}-B_{t_{k-1}}\right)^2,$$

so

$$I(F) = \frac{1}{2}B_{t_n} - \frac{1}{2}\sum_{k=1}^{n}\left(B_{t_k} - B_{t_{k-1}}\right)^2.$$

Using $t_n = T$ and the proposition 1.2.5, we get the following explicit formula of a stochastic integral

$$\int_0^T B_t dB_t = \frac{1}{2}B_T^2 - \frac{1}{2}T.$$

Next, we briefly discuss the higher-dimensional case. Let $B_t = (B_t^1, ..., B_t^m)$ be $m-$dimensional Brownian motion and $F = (F^1, ..., F^m) \in L^2_{\mathcal{F}}(0, T, \mathbb{R}^m)$. Then, for $i = 1, ...m$, $\int_0^t F_s^i dB_s^i$ is well-defined. We define

$$\int_0^t F_s dB_s := \sum_{i=1}^{m}\int_0^t F_s^i dB_s^i,\ t \geq 0.$$

1.3.2 Introduction to backward integrals

Let B be a Brownian motion in $\mathbb{R}^m$, and $\mathcal{F}_{t,T}^B := \sigma\left\{B_r - B_T /\ t \leq r \leq T\right\} \vee \mathcal{N}$, where $\mathcal{N}$ is the $\mathbb{P}-$null sets in Ω, then $\left\{\mathcal{F}_{t,T}^B /\ 0 \leq t \leq T\right\}$ is a backward filtration in the sense that $\mathcal{F}_{t,T}^B \supseteq \mathcal{F}_{s,T}^B$ if $s \leq t$. If $\{Z_t /\ 0 \leq t \leq T\}$ is a stochastic process over $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying Z_t is $\mathcal{F}_{t,T}^B-$measurable $\forall 0 \leq t \leq T$, we say that Z is $\left\{\mathcal{F}_{t,T}^B /\ t \leq T\right\}-$adapted.
the backward Itô integral of Z with respect to B is defined by

$$\int_t^T Z_s d\overleftarrow{B}_s := \lim_{n\to\infty}\sum_{k=1}^{n} z_{t_{k+1}}\left(B_{t_{k+1}} - B_{t_k}\right),$$

where $\pi = \{t = t_1, t_2, ..., t_{n+1} = T\}$ is a partition of $[t, T]$ satisfying

$$|\pi| := \sup_{1\leq k\leq n}(t_{k+1} - t_k) \to 0 \text{ as } n \to \infty.$$

On the other hand, letting $\check{B}_s := B_{T-s} - B_T$, $0 \leq s \leq T$, shows that $\check{B}$ is a Brownian motion as well, and for all $0 \leq s \leq T$, we have

$$\begin{aligned} \mathcal{F}^{B}_{T-t,T} &= \sigma\left\{B_r - B_T / \ T-t \leq r \leq T\right\} \\ &= \sigma\left\{B_{T-s} - B_T / \ 0 \leq s \leq t\right\}, \ \ (r = T-s), \\ &= \sigma\left\{\check{B}_s / \ 0 \leq s \leq t\right\} = \mathcal{F}^{\hat{B}}_t. \end{aligned}$$

If Z_s is $\mathcal{F}^{B}_{t,T}-$measurable for all $0 \leq s \leq T$, then $\check{Z}_s := Z_{T-s}$ is $\mathcal{F}^{\hat{B}}_s-$measurable for all $0 \leq s \leq T$.

So, the backward Itô integral of Z_s with respect to B_s may be understood as the forward integral of $\check{Z}_s = Z_{T-s}$ with respect to $\check{B}_s$

$$\begin{aligned} \textstyle\int_0^{T-t} \check{Z}_s \check{B}_s &= \lim_{n\to\infty} \textstyle\sum_{k=1}^{n} \check{z}_{s_k}\left(\check{B}_{s_{k+1}} - \check{B}_{s_k}\right) \\ &= \lim_{n\to\infty}\left[\check{z}_{s_1}\left(\check{B}_{s_2} - \check{B}_{s_1}\right) + ... + \check{z}_{s_n}\left(\check{B}_{s_{n+1}} - \check{B}_{s_n}\right)\right] \\ &= \lim_{n\to\infty}\left[\check{z}_0\left(\check{B}_{T-t_n} - \check{B}_0\right) + ... + \check{z}_{T-t_2}\left(\check{B}_{T-t_1} - \check{B}_{T-t_2}\right)\right] \\ &= \lim_{n\to\infty}\left[z_T\left(B_{t_n} - B_T\right) + ... + z_{t_2}\left(B_t - B_{t_2}\right)\right] \\ &= -\lim_{n\to\infty}\left[z_{t_2}\left(B_{t_2} - B_{t_1}\right) + ... + z_{t_{n+1}}\left(B_{t_{n+1}} - B_{t_n}\right)\right] \\ &= -\textstyle\int_t^T Z_s d\overleftarrow{B}_s, \end{aligned}$$

where $S = \{s_1 = 0, ..., s_n = T-t\}$ is a partition of $[0, T-t]$, where $s_k = T - t_{n+2-k}$ for all $1 \leq k \leq n+1$.

1.3.3 Itô's Formula

In general the basic definition of Itô integral is not very useful when we try to evaluate a given integral. In this context, however, we have no differentiation theory, only integration theory. Nevertheless it turns out that it is possible to establish an Itô integral version of the chain rule, called the Itô's Formula.

The Itô's Formula is central to the theory of stochastic calculus.

Definition 1.3.1 (***Itô's process***) *Let B be m-dimensional Brownian motion on a filtered* *probability space* $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$. *$X$ is called a Itô process, if X admits the representation*

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

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

$$\int_0^t |b_s|\, ds + \int_0^t |\sigma_s|^2\, ds < \infty,\ a.s.,\ 0 \leq t \leq T.$$

Theorem 1.3.2 (***Itô's Formula***) *Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ be a filtered probability space satisfying the usual condition, B_t be m-dimensional Brownian motion, and let X_t be an 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*

$$f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial f}{\partial t}(s, X_s)\, ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)\, dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x \partial x}(s, X_s)\, \sigma_s^2 ds,$$

$\forall t \in [0,T]$.

Proposition 1.3.1 *(**Integration by parts formula**) Let X and Y be Itô processes in $\mathbb{R}$. Then*

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

We shall need the following extension of the Itô formula.

Let $L^2_{\mathcal{F}}(0, T, \mathbb{R}^n)$: the set of all $\mathcal{F}_t-$adapted and $\mathbb{R}^n-$valued processes X, such that

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

$L^2_{\mathcal{F}}(\Omega; C(0, T, \mathbb{R}^n))$: the set of all $\mathcal{F}_t-$adapted and $\mathbb{R}^n-$valued continuous processes X, such that

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

Proposition 1.3.2 *Let* $a\in L^2_{\mathcal{F}}(\Omega;C(0,T,\mathbb{R}^n))$, $b\in L^2_{\mathcal{F}}(0,T,\mathbb{R}^n)$, $c\in L^2_{\mathcal{F}}\left(0,T,\mathbb{R}^{n\times d}\right)$ *and* $d\in L^2_{\mathcal{F}}(0,T,\mathbb{R}^{n\times m})$. *Assume that*

$$a_t=a_0+\int_0^t b_s ds+\int_0^t c_s d\overleftarrow{B}_s+\int_0^t d_s dW_s,\ \ t\in[0,T],$$

and

Then, for each $t\in[0,T]$,

$$\begin{aligned}|a_t|^2=\ & |a_0|^2+2\int_0^t\langle a_s,b_s\rangle\,ds+2\int_0^t\langle a_s,c_s\rangle\,d\overleftarrow{B}_s\\ &+2\int_0^t\langle a_s,d_s\rangle\,dW_s-\int_0^t\|c_s\|^2\,ds+\int_0^t\|d_s\|^2\,ds,\end{aligned}$$

$$\mathbb{E}\,|a_t|^2=\mathbb{E}\,|a_0|^2+2\mathbb{E}\int_0^t\langle a_s,b_s\rangle\,ds-\mathbb{E}\int_0^t\|c_s\|^2\,ds+\mathbb{E}\int_0^t\|d_s\|^2\,ds.$$

Proof. See Pardoux and Peng [34]. ■

Chapter 2

Optimal control problem for a linear FBDSDEs

In this chapter, we study a control problem of linear forward-backward doubly SDEs with non linear cost functional. We prove in first, the existence of a strong optimal controls which is adapted to the initial σ-algebra, under the convexity of the cost function and the domain of control U. The proof is based on strong convergence techniques for the associated linear FBDSDEs and Mazur's theorem. Secondly, we establish necessary as well as sufficient optimality conditions for this kind of control problem by using the convex optimization principle.

2.1 Formulation of the problem and assumptions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ 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 $\mathbb{P}-$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 π_t, $\mathcal{F}^\pi_{s,t} = \sigma\left(\pi_r - \pi_s, s \leq r \leq t\right) \vee \mathcal{N}$, $\mathcal{F}^\pi_t = \mathcal{F}^\pi_{0,t}$.

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

Given x a square integrable and $\mathcal{F}_0$-measurable process and ξ is a square integrable and $\mathcal{F}_T$-measurable process, and for any admissible control u, we consider an optimal control problem driven by the following controlled linear FBDSDE

$$\begin{cases} dX_t^u = (\alpha_t X_t^u + \beta_t u_t)\, dt + \left(\widehat{\alpha}_t X_t^u + \widehat{\beta}_t u_t\right) dW_t \\ \\ dY_t^u = -\left(\gamma_t X_t^u + \widehat{\gamma}_t Y_t^u + \delta_t Z_t^u + \widehat{\delta}_t u_t\right) dt \\ \qquad\qquad - \left(\eta_t X_t^u + \widehat{\eta}_t Y_t^u + \theta_t Z_t^u + \widehat{\theta}_t u_t\right) \overleftarrow{dB}_t + Z_t^u dW_t, \\ \\ X_0^u = x, Y_T^u = \xi, \end{cases} \tag{2.1}$$

and a functional cost to be minimized over the set admissible control $\mathcal{U}$, given by:

$$\mathbb{J}(u.) := \mathbb{E}\left[\varphi\left(X_T^u\right) + \psi\left(Y_0^u\right) + \int_0^T L\left(t, X_t^u, Y_t^u, Z_t^u, u_t\right) dt\right], \tag{2.2}$$

where $\alpha., \widehat{\alpha}., \beta., \widehat{\beta}., \gamma., \widehat{\gamma}., \delta., \widehat{\delta}., \eta., \widehat{\eta}., \theta.$ and $\widehat{\theta}.$ are matrix-valued functions of suitable sizes. The solution $(X., Y., Z.)$ takes values in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$. $u.$ is the control variable values in subset U of $\mathbb{R}^k$. φ, ψ, L are a given functions define by

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

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

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

Definition 2.1.1 *An admissible control $u.$ is a square integrable, $\mathcal{F}_t$-measurable process with values in some subset $U \subseteq \mathbb{R}^k$. We denote by $\mathcal{U}$ the set of all admissible controls.*

We assumed here that the control variable must be square-integrable just to ensure the existence of solutions of (2.1) under $u.$. We say that an admissible control $u^*_\cdot \in \mathcal{U}$ is an optimal control if

$$\mathbb{J}\left(u_{\cdot}^{*}\right)=\inf_{v_{\cdot}\in\mathcal{U}}\mathbb{J}\left(v_{\cdot}\right). \tag{2.3}$$

We shall consider, the following assumptions

$(\mathbf{H1})$: the set $U\subseteq\mathbb{R}^{k}$ is convex and compact and the cost functions L,φ and ψ are continuous, bounded and convex,

$(\mathbf{H2})$: $\alpha_t,\widehat{\alpha}_t,\beta_t,\widehat{\beta}_t,\gamma_t,\widehat{\gamma}_t,\delta_t,\widehat{\delta}_t,\eta_t,\widehat{\eta}_t$ and $\widehat{\theta}_t$ are bounded by $\rho>0$ and θ_t is bounded by $\sigma\in(0,1)$. That is:

$$\rho:=\sup_{t,\omega}\left|\phi_t\left(\omega\right)\right| \text{ and } \sigma:=\sup_{t,\omega}\left|\theta_t\left(\omega\right)\right|,$$

where $\phi_t\left(\omega\right)=\alpha_t,\widehat{\alpha}_t,\beta_t,\widehat{\beta}_t,\gamma_t,\widehat{\gamma}_t,\delta_t,\widehat{\delta}_t,\eta_t,\widehat{\eta}_t,\widehat{\theta}_t$.

2.2 Existence of optimal strict controls for linear FBDSDEs:

The first main result in this chapter is given by the following theorem:

Theorem 2.2.1 *Under assumptions* $(\mathbf{H1})-(\mathbf{H2})$, *the strict control problem defined by ((2.1), (2.2), (2.3)) has a strong optimal solution.*

Proof. Assume that $(\mathbf{H1})-(\mathbf{H2})$ hold. Let $(u_{\cdot}^{n})$ be a minimizing sequence, which satisfies

$$\lim_{n\to\infty}\mathbb{J}\left(u_{\cdot}^{n}\right)=\inf_{v_{\cdot}\in\mathcal{U}}\mathbb{J}\left(v_{\cdot}\right).$$

With corresponding trajectories $(X_\cdot^n, Y_\cdot^n, Z_\cdot^n)$ solution of the following linear FBDSDE:

$$\begin{cases} dX_t^n = \left(\alpha_t X_t^n + \beta_t u_t^n\right) dt + \left(\widehat{\alpha}_t X_t^n + \widehat{\beta}_t u_t^n\right) dW_t \\ \\ dY_t^n = -\left(\gamma_t X_t^n + \widehat{\gamma}_t Y_t^n + \delta_t Z_t^n + \widehat{\delta}_t u_t^n\right) dt \\ \qquad\qquad - \left(\eta_t X_t^n + \widehat{\eta}_t Y_t^n + \theta_t Z_t^n + \widehat{\theta}_t u_t^n\right) \overleftarrow{dB}_t + Z_t^n dW_t, \\ \\ X_0^n = x, Y_T^n = \xi, \end{cases}$$

Since U is a compact set, the sequence $(u_\cdot^n)_{n\geq 0}$ is relatively compact. So, there exists a subsequence (which is still labeled by $(u_\cdot^n)_{n\geq 0}$) such that

$$u_\cdot^n \longrightarrow \widetilde{u}_\cdot, \text{ weakly in } \mathcal{M}^2\left(\left[0,T\right], \mathbb{R}^k\right).$$

Applying Mazur's theorem (see Yosida [36], Theorem 2 page 120), there is a sequence of convex combinations defined by

$$\overline{u}_\cdot^n = \sum_{\jmath\geq 0} \pi_{\jmath n} u_\cdot^{\jmath+n} \quad (\text{with } \pi_{\jmath n} \geq 0, \quad \text{and } \sum_{\jmath\geq 0} \pi_{\jmath n} = 1),$$

such that

$$\overline{u}_\cdot^n \rightarrow \widetilde{u}_\cdot \text{ strongly in } \mathcal{M}^2\left(\left[0,T\right], \mathbb{R}^k\right). \tag{2.4}$$

Since the set of control (action space) $U \subseteq \mathbb{R}^k$ is convex and compact, it follows easily that $\widetilde{u}_\cdot \in \mathcal{U}$.

Let $(\overline{X}_\cdot^n, \overline{Y}_\cdot^n, \overline{Z}_\cdot^n)$ and $(X_\cdot^{\widetilde{u}_\cdot}, Y_\cdot^{\widetilde{u}_\cdot}, Z_\cdot^{\widetilde{u}_\cdot})$ be the solutions of the linear FBDSDE (2.1), associated with $\overline{u}_\cdot^n$ and $\widetilde{u}_\cdot$ respectively i.e.,

$$\begin{cases} d\overline{X}_t^n = \left(\alpha_t \overline{X}_t^n + \beta_t u_t^n\right) dt + \left(\widehat{\alpha}_t \overline{X}_t^n + \widehat{\beta}_t \overline{u}_t^n\right) dW_t \\ \\ d\overline{Y}_t^n = -\left(\gamma_t \overline{X}_t^n + \widehat{\gamma}_t \overline{Y}_t^n + \delta_t \overline{Z}_t^n + \widehat{\delta}_t \overline{u}_t^n\right) dt \\ \qquad\qquad - \left(\eta_t \overline{X}_t^n + \widehat{\eta}_t \overline{Y}_t^n + \theta_t \overline{Z}_t^n + \widehat{\theta}_t \overline{u}_t^n\right) \overleftarrow{dB}_t + \overline{Z}_t^n dW_t, \\ \\ \overline{X}_0^n = x, \overline{Y}_T^n = \xi, \end{cases} \tag{2.5}$$

and

$$\begin{cases} dX_t^{\widetilde{u}.} = \left(\alpha_t X_t^{\widetilde{u}.} + \beta_t \widetilde{u}_t\right) dt + \left(\widehat{\alpha}_t X_t^{\widetilde{u}.} + \widehat{\beta}_t \widetilde{u}_t\right) dW_t \\ \\ dY_t^{\widetilde{u}.} = -\left(\gamma_t X_t^{\widetilde{u}.} + \widehat{\gamma}_t Y_t^{\widetilde{u}.} + \delta_t Z_t^{\widetilde{u}.} + \widehat{\delta}_t \widetilde{u}_t\right) dt \\ \qquad\qquad - \left(\eta_t X_t^{\widetilde{u}.} + \widehat{\eta}_t Y_t^{\widetilde{u}.} + \theta_t Z_t^{\widetilde{u}.} + \widehat{\theta}_t \widetilde{u}_t\right) \overleftarrow{dB}_t + Z_t^{\widetilde{u}.} dW_t, \\ \\ X_0^{\widetilde{u}.} = x, Y_T^{\widetilde{u}.} = \xi. \end{cases} \tag{2.6}$$

Firstly, let us prove that:

$$(\overline{X}_t^n, \overline{Y}_t^n, \int_0^T \overline{Z}_s^n dW_s) \text{ converges strongly to } (X_t^{\widetilde{u}.}, Y_t^{\widetilde{u}.}, \int_0^T Z_s^{\widetilde{u}.} dW_s), \tag{2.7}$$

in $\mathcal{S}^2\left([0,T], \mathbb{R}^{n+m}\right) \times \mathcal{M}^2\left([0,T], \mathbb{R}^{m\times d}\right)$.

We have

$$\begin{aligned} \left|\overline{X}_t^n - X_t^{\widetilde{u}}\right| \leq & \left|\int_0^t \left(\alpha_s(\overline{X}_s^n - X_s^{\widetilde{u}}) + \beta_s(u_s^n - \widetilde{u}_s)\right) ds\right| \\ & + \left|\int_0^t \left(\widehat{\alpha}_s(\overline{X}_s^n - X_s^{\widetilde{u}}) + \widehat{\beta}_s(u_s^n - \widetilde{u}_s)\right) dW_s\right|, \end{aligned}$$

hence

$$(\sup_{0\le s\le t}|\overline{X}^n_s - X^{\widetilde{u}}_s|)^2 \le \int_0^t \left(|\alpha_s|^2(\sup_{0\le r\le s}|\overline{X}^n_r - X^{\widetilde{u}}_r|^2) + |\beta_s|^2|u^n_s - \widetilde{u}_s|^2\right) ds$$
$$+ \sup_{0\le s\le t}\left|\int_0^t \left(\widehat{\alpha}_s(\overline{X}^n_s - X^{\widetilde{u}}_s) + \widehat{\beta}_s(u^n_s - \widetilde{u}_s)\right) dW_s\right|^2,$$

using the Burkholder-Davis-Gundy inequality to the martingale part we get

$$\mathbb{E}\left[\sup_{0\le s\le T}|\overline{X}^n_s - X^{\widetilde{u}.}_s|)^2\right] \le K\int_0^t \mathbb{E}\left[\sup_{0\le r\le s}|\overline{X}^n_r - X^{\widetilde{u}.}_r|^2\right] ds + K'\mathbb{E}\left[\int_0^t |\overline{u}^n_s - \widetilde{u}_s|^2 ds\right].$$

By Gronwall's lemma and the fact that $\overline{u}^n_.$ converges strongly to $\widetilde{u}.$ in $\mathcal{M}^2([0,T],\mathbb{R}^k)$ (from (2.4)) we have

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{0\le s\le T}|\overline{X}^n_s - X^{\widetilde{u}}_s|^2\right] = 0. \tag{2.8}$$

On the other hand, applying Itô's formula to $\left|\overline{Y}^n_t - Y^{\widetilde{u}.}_t\right|^2$, we obtain

$$|\overline{Y}^n_t - Y^{\widetilde{u}.}_t|^2 + \int_t^T \|\overline{Z}^n_s - Z^{\widetilde{u}.}_s\|^2 ds = 2\int_t^T \langle \overline{Y}^n_s - Y^{\widetilde{u}.}_s, \gamma_s(\overline{X}^n_s - X^{\widetilde{u}.}_s) + \widehat{\gamma}_s(\overline{Y}^n_s - Y^{\widetilde{u}.}_s)$$
$$+ \delta_s(\overline{Z}^n_s - Z^{\widetilde{u}.}_s) + \widehat{\delta}_s(\overline{u}^n_s - \widetilde{u}_s)\rangle ds + 2\int_t^T \langle \overline{Y}^n_s - Y^{\widetilde{u}.}_s, \eta_s(\overline{X}^n_s - X^{\widetilde{u}}_s)$$
$$+ \widehat{\eta}_s(\overline{Y}^n_s - Y^{\widetilde{u}.}_s) + \theta_s(\overline{Z}^n_s - Z^{\widetilde{u}.}_s) + \widehat{\theta}_s(\overline{u}^n_s - \widetilde{u}_s)\rangle \overleftarrow{dB}_s$$
$$- 2\int_t^T \langle \overline{Y}^n_s - Y^{\widetilde{u}.}_s, \overline{Z}^n_s - Z^{\widetilde{u}.}_s\rangle dW_s$$
$$+ \int_t^T |\eta_s(\overline{X}^n_s - X^{\widetilde{u}.}_s) + \widehat{\eta}_s(\overline{Y}^n_s - Y^{\widetilde{u}.}_s) + \theta_s(\overline{Z}^n_s - Z^{\widetilde{u}.}_s) + \widehat{\theta}_s(\overline{u}^n_s - \widetilde{u}_s)|^2 ds.$$

Take expectation we get

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|\overline{Y}_t^n - Y_t^{\widetilde{u}.}|^2\right] + \mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \leq$$
$$2\mathbb{E}\left[\int_0^T \langle \overline{Y}_s^n - Y_s^{\widetilde{u}.}, \gamma_s(\overline{X}_s^n - X_s^{\widetilde{u}.}) + \widehat{\gamma}_s(\overline{Y}_s^n - Y_s^{\widetilde{u}.}) \right.$$
$$\left. +\delta_s(\overline{Z}_s^n - Z_s^{\widetilde{u}.}) + \widehat{\delta}_s(\overline{u}_s^n - \widetilde{u}_s)\rangle ds\right]$$
$$+ \mathbb{E}\left[\int_0^T |\eta_s(\overline{X}_s^n - X_s^{\widetilde{u}.}) + \widehat{\eta}_s(\overline{Y}_s^n - Y_s^{\widetilde{u}.}) \right.$$
$$\left. +\theta_s(\overline{Z}_s^n - Z_s^{\widetilde{u}.}) + \widehat{\theta}_s(\overline{u}_s^n - \widetilde{u}_s)|^2 ds\right].$$

Under the assumption (H2) and by using the Young's formula ($2ab \leq \frac{1}{\varepsilon_1}a^2 + \varepsilon_1 b^2$), we obtain

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|\overline{Y}_t^n - Y_t^{\widetilde{u}.}|^2\right] + \mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \leq \frac{1}{\varepsilon_1}\mathbb{E}\left[\int_0^T |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 ds\right]$$
$$+ 4\varepsilon_1\rho^2\mathbb{E}\left[\int_0^T \left(|\overline{X}_s^n - X_s^{\widetilde{u}.}|^2 + |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 \right.\right.$$
$$\left.\left. +\|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 + |\overline{u}_s^n - \widetilde{u}_s|^2\right) ds\right]$$
$$+ 3\rho^2\mathbb{E}\left[\int_0^T \left(|\overline{X}_s^n - X_s^{\widetilde{u}.}|^2 + |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 + |\overline{u}_s^n - \widetilde{u}_s|^2\right) ds\right]$$
$$+ \sigma\mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] + 2\rho\sigma\mathbb{E}\left[\int_0^T \langle(\overline{X}_s^n - X_s^{\widetilde{u}.}) + (\overline{Y}_s^n - Y_s^{\widetilde{u}.}) \right.$$
$$\left. + (\overline{u}_s^n - \widetilde{u}_s), (\overline{Z}_s^n - Z_s^{\widetilde{u}.})\rangle ds\right].$$

Hence

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leq t\leq T}|\overline{Y}_t^n - Y_t^{\widetilde{u}.}|^2\right] + \mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \\
&\quad \leq \frac{1}{\varepsilon_1}\mathbb{E}\left[\int_0^T |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 ds\right] + 4\varepsilon_1\rho^2\mathbb{E}\left[\int_0^T \left(|\overline{X}_s^n - X_s^{\widetilde{u}.}|^2 + |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2\right.\right. \\
&\quad \left.\left. + \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 + |\overline{u}_s^n - \widetilde{u}_s|^2\right) ds\right] + 3\rho^2\mathbb{E}\left[\int_0^T \left(|\overline{X}_s^n - X_s^{\widetilde{u}.}|^2\right.\right. \\
&\quad \left.\left. + |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 + |\overline{u}_s^n - \widetilde{u}_s|^2\right) ds\right] + \sigma\mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \\
&\quad + \frac{3\rho\sigma}{\varepsilon_2}\mathbb{E}\left[\int_0^T \left(|\overline{X}_s^n - X_s^{\widetilde{u}.}|^2 + |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 + |\overline{u}_s^n - \widetilde{u}_s|^2\right) ds\right] \\
&\quad + \varepsilon_2\rho\sigma\mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right],
\end{aligned}$$

and therefore

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leq t\leq T}|\overline{Y}_t^n - Y_t^{\widetilde{u}.}|^2\right] + \mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \\
&\quad \leq (\frac{1}{\varepsilon_1} + 4\varepsilon_1\rho^2 + 3\rho^2 + \frac{3\rho\varepsilon}{\varepsilon_2})\mathbb{E}\left[\int_0^T |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 ds\right] \\
&\quad + (4\varepsilon_1\rho^2 + \sigma + \varepsilon_2\rho\sigma)\mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \\
&\quad + (4\varepsilon_1\rho^2 + 3\rho^2 + \frac{3\rho\sigma}{\varepsilon_2})\mathbb{E}\left[\int_0^T |\overline{X}_s^n - X_s^{\widetilde{u}.}|^2 ds\right] \\
&\quad + (4\varepsilon_1\rho^2 + 3\rho^2 + \frac{3\rho\sigma}{\varepsilon_2})\mathbb{E}\left[\int_0^T |\overline{u}_s^n - \widetilde{u}_s|^2 ds\right],
\end{aligned}$$

Choosing

$$\varepsilon_1 = \frac{1-\sigma}{8\rho^2} > 0 \text{ and } \varepsilon_2 = \frac{1-\sigma}{3\rho\sigma} > 0,$$

the previous inequality becomes

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}|\overline{Y}_t^n - Y_t^{\widetilde{u}.}|^2\right] + k_1\mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \\
&\quad\le k_2\mathbb{E}\left[\int_0^T |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 ds\right] \\
&\quad + k_3\mathbb{E}\left[\int_0^T |\overline{X}_s^n - X_s^{\widetilde{u}.}|^2 ds\right] + k_3\mathbb{E}\left[\int_0^T |\overline{u}_s^n - \widetilde{u}_s|^2 ds\right],
\end{aligned} \tag{2.9}$$

where

$$\begin{aligned}
k_1 &= \frac{1-\sigma}{6} > 0, \\
k_2 &= \frac{8\rho^2}{1-\sigma} + \frac{1-\sigma}{2} + 3\rho^2 + \frac{9\left(\rho\sigma\right)^2}{1-\sigma} > 0, \\
k_3 &= \frac{1-\sigma}{2} + 3\rho^2 + \frac{9\left(\rho\sigma\right)^2}{1-\sigma} > 0.
\end{aligned}$$

We derive from (2.9) two inequalities:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}|\overline{Y}_t^n - Y_t^{\widetilde{u}.}|^2\right] \le k_2\mathbb{E}\left[\int_0^T |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 ds\right] \\
&\quad + k_3\mathbb{E}\left[\int_0^T |\overline{X}_s^n - X_s^{\widetilde{u}.}|^2 ds\right] + k_3\mathbb{E}\left[\int_0^T |\overline{u}_s^n - \widetilde{u}_s|^2 ds\right],
\end{aligned} \tag{2.10}$$

and

$$\begin{aligned}
&k_1\mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \le k_2\mathbb{E}\left[\int_0^T |\overline{Y}_s^n - Y_s^{\widetilde{u}.}|^2 ds\right] \\
&\quad + k_3\mathbb{E}\left[\int_0^T |\overline{X}_s^n - X_s^{\widetilde{u}.}|^2 ds\right] + k_3\mathbb{E}\left[\int_0^T |\overline{u}_s^n - \widetilde{u}_s|^2 ds\right],
\end{aligned} \tag{2.11}$$

Applying Gronwall's lemma to (2.10) and passing to the limit as $n \to \infty$, and using the convergence (2.4) and (2.8), we obtain

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{0\le t\le T}|\overline{Y}_t^n - Y_t^{\widetilde{u}.}|^2\right] = 0. \tag{2.12}$$

Then, one can shows directly from (2.4),(2.8) and (2.12) that

$$\mathbb{E}\left[\int_0^T \|\overline{Z}_s^n - Z_s^{\widetilde{u}.}\|^2 ds\right] \longrightarrow 0, \text{ as } n \to \infty,$$

which implies by applying the isometry of Itô that:

$$\int_0^T \overline{Z}_s^n dW_s \textit{ converges strongly to } \int_0^T Z_s^{\widetilde{u}.} dW_s,$$

in $\mathcal{M}^2([0,T], \mathbb{R}^{m\times d})$.

Finally, let us prove that $\widetilde{u}.$ is an optimal control.

According to the minimizing sequence, $(X_\cdot^n, Y_\cdot^n, Z_\cdot^n, u_\cdot^n)$ satisfies

$$\begin{aligned}\lim_{n\to\infty} \mathbb{J}(u_\cdot^n) &= \lim_{n\to\infty} \mathbb{E}\left[\varphi(X_T^n) + \psi(Y_0^n) + \textstyle\int_0^T L(t, X_t^n, Y_t^n, Z_t^n, u_t^n)\, dt\right] \\ &= \inf_{v.\in\mathcal{U}} \mathbb{J}(v.).\end{aligned}$$

Using the continuity of functions φ, ψ and L, we get

$$\begin{aligned}\mathbb{J}(\widetilde{u}.) &= \mathbb{E}\left[\varphi\left(X_T^{\widetilde{u}.}\right) + \psi\left(Y_0^{\widetilde{u}.}\right) + \textstyle\int_0^T L\left(t, X_t^{\widetilde{u}.}, Y_t^{\widetilde{u}.}, Z_t^{\widetilde{u}.}, \widetilde{u}_t\right) dt\right] \\ &= \lim_{n\to\infty} \mathbb{E}\left[\varphi\left(\overline{X}_T^n\right) + \psi\left(\overline{Y}_0^n\right) + \textstyle\int_0^T L\left(t, \overline{X}_t^n, \overline{Y}_t^n, \overline{Z}_t^n, \overline{u}_t^n\right) dt\right].\end{aligned}$$

By the convexity of φ, ψ and L, it follows that

$$\begin{aligned}\mathbb{J}(\widetilde{u}.) &\leq \lim_{n\to\infty} \sum_{j\geq 0} \pi_{jn} \mathbb{E}\left[\varphi\left(X_T^{j+n}\right) + \psi\left(Y_0^{j+n}\right)\right. \\ &\quad \left. + \int_0^T L\left(t, X_t^{j+n}, Y_t^{j+n}, Z_t^{j+n}, u_t^{j+n}\right) dt\right] \\ &= \lim_{n\to\infty} \sum_{j\geq 0} \pi_{jn} \mathbb{J}\left(u_\cdot^{j+n}\right), \\ &\leq \lim_{n\to\infty} \underset{1\leq j\leq i_n}{Max}\, \mathbb{J}\left(u_\cdot^{j+n}\right) \sum_{j\geq 1} \pi_{jn} = \inf_{v.\in\mathcal{U}} \mathbb{J}(v.).\end{aligned}$$

■

2.3 Necessary and sufficient conditions

We establish in this section, necessary as well as sufficient optimality conditions for a stochastic control problem governed by a linear forward-backward doubly SDE by using the convex optimization principle. In this end, and according to the convexity of the domain of control U, we use the convex perturbation method.

Let $(\widetilde{u}_{\cdot}, X_t^{\widetilde{u}_{\cdot}}, Y_t^{\widetilde{u}_{\cdot}}, Z_t^{\widetilde{u}_{\cdot}})$ be the optimal solution of the control problem (2.1)–(2.3) obtained in above section, which satisfies:

$$\begin{cases} X_t^{\widetilde{u}_{\cdot}} = x + \int_0^t \left(\alpha_s X_s^{\widetilde{u}_{\cdot}} + \beta_s \widetilde{u}_s\right) ds + \int_0^t \left(\widehat{\alpha}_s X_s^{\widetilde{u}_{\cdot}} + \widehat{\beta}_s \widetilde{u}_s\right) dW_s \\ \\ Y_t^{\widetilde{u}_{\cdot}} = \xi + \int_t^T \left(\gamma_s X_s^{\overline{u}_{\cdot}} + \widehat{\gamma}_s Y_s^{\overline{u}_{\cdot}} + \delta_s Z_s^{\widetilde{u}_{\cdot}} + \widehat{\delta}_s \widetilde{u}_s\right) ds \\ \qquad + \int_t^T \left(\eta_s X_s^{\widetilde{u}_{\cdot}} + \widehat{\eta}_s Y_s^{\widetilde{u}_{\cdot}} + \theta_s Z_s^{\widetilde{u}_{\cdot}} + \widehat{\theta}_s \widetilde{u}_s\right) \overleftarrow{dB}_s - \int_t^T Z_s^{\widetilde{u}_{\cdot}} dW_s. \end{cases} \tag{2.13}$$

Let us define the perturbed control as follow: for each admissible control $v.$

$$u_t^{\varepsilon} = \widetilde{u}_t + \varepsilon \left(v_t - \widetilde{u}_t\right),$$

where, $\varepsilon > 0$ is sufficiently small.

The perturbed control $u_{\cdot}^{\varepsilon}$ is admissible control with associated trajectory $(X_t^{\varepsilon}, Y_t^{\varepsilon}, Z_t^{\varepsilon})$, solution of the following FBDSDE:

$$\begin{cases} X_t^{\varepsilon} = x + \int_0^t \left(\alpha_s X_s^{\varepsilon} + \beta_s u_s^{\varepsilon}\right) ds + \int_0^t \left(\widehat{\alpha}_s X_s^{\varepsilon} + \widehat{\beta}_s u_s^{\varepsilon}\right) dW_s \\ \\ Y_t^{\varepsilon} = \xi + \int_t^T \left(\gamma_s X_s^{\varepsilon} + \widehat{\gamma}_s Y_s^{\varepsilon} + \delta_s Z_s^{\varepsilon} + \widehat{\delta}_s u_s^{\varepsilon}\right) ds \\ \qquad + \int_t^T \left(\eta_s X_s^{\varepsilon} + \widehat{\eta}_s Y_s^{\varepsilon} + \theta_s Z_s^{\varepsilon} + \widehat{\theta}_s u_s^{\varepsilon}\right) \overleftarrow{dB}_s - \int_t^T Z_s^{\varepsilon} dW_s. \end{cases}$$

According to the optimality of $\overline{u}.$ and by using the following inequality, one can establish the necessary optimality conditions,

$$
\begin{aligned}
0 &\leq \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(\mathbb{J}\left(u^{\varepsilon}_{\cdot}\right)-\mathbb{J}\left(\widetilde{u}_{\cdot}\right)\right)\\
&=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(\mathbb{J}\left(\widetilde{u}_{\cdot}+\varepsilon\left(v_{\cdot}-\widetilde{u}_{\cdot}\right)\right)-\mathbb{J}\left(\widetilde{u}_{\cdot}\right)\right)\\
&=\left\langle \mathbb{J}'\left(\widetilde{u}_{\cdot}\right), v_{\cdot}-\widetilde{u}_{\cdot}\right\rangle .
\end{aligned}
$$

Let us consider the following assumptions

(**H5**) (Regularity conditions)

$$
\begin{cases}
(i)\ \text{the function } L \text{ is continuously differentiable with respect to}\\
\quad (X,Y,Z,v),\ \text{and the mappings } \varphi \text{ and } \psi \text{ are continuously}\\
\quad \text{differentiable with respect to } X \text{ and } Y,\ \text{respectively,}\\
(ii)\ \text{the derivatives of } L,\varphi,\psi \text{ with respect to their arguments (given above)}\\
\quad \text{are bounded.}
\end{cases}
$$

The second main result in this chapter, is the following

Theorem 2.3.1 *(Necessary and sufficient conditions for optimality). Let $\widetilde{u}_{\cdot}$ be an admissible control (candidate to be optimal) with associated trajectories $(X^{\widetilde{u}_{\cdot}}_{\cdot}, Y^{\widetilde{u}_{\cdot}}_{\cdot}, Z^{\widetilde{u}_{\cdot}}_{\cdot})$. Then the control $\widetilde{u}_{\cdot}$ is optimal for the control problem (2.1)-(2.3), if and only if, there exists a unique solution $\left(Q^{\widetilde{u}_{\cdot}}_{\cdot}, K^{\widetilde{u}_{\cdot}}_{\cdot}, P^{\widetilde{u}_{\cdot}}_{\cdot}, \Gamma^{\widetilde{u}_{\cdot}}_{\cdot}\right)$ of the following adjoint equations of the linear forward-backward doubly SDE (2.1),*

$$
\begin{cases}
-dQ^{\widetilde{u}_{\cdot}}_t = H_X\left(t,\zeta^{\widetilde{u}_{\cdot}}_t,\widetilde{u}_t,\Pi^{\widetilde{u}_{\cdot}}_t\right)dt - K^{\widetilde{u}}_t dW_t,\\
dP^{\widetilde{u}_{\cdot}}_t = H_Y\left(t,\zeta^{\widetilde{u}_{\cdot}}_t,\widetilde{u}_t,\Pi^{\widetilde{u}_{\cdot}}_t\right)dt + H_Z\left(t,\zeta^{\widetilde{u}_{\cdot}}_t,\widetilde{u}_t,\Pi^{\widetilde{u}_{\cdot}}_t\right)dW_t - \Gamma^{\widetilde{u}}_t \overleftarrow{dB}_t,\\
Q^{\widetilde{u}_{\cdot}}_T = \varphi_X\left(X^{\widetilde{u}_{\cdot}}_T\right),\ P^{\widetilde{u}_{\cdot}}_0 = \psi_Y\left(Y^{\widetilde{u}_{\cdot}}_0\right),
\end{cases}
\tag{2.14}
$$

such that

$$\langle H_v(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t,\Pi_t^{\widetilde{u}.}),v_t-\widetilde{u}_t\rangle \geq 0,\ \forall v_t\in U,\ a.e,\ as, \tag{2.15}$$

where $H_\lambda(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t,\Pi_t^{\widetilde{u}.})$ *is the gradient*

$$\nabla_\lambda H(t,X_t^{\widetilde{u}.},Y_t^{\widetilde{u}.},Z_t^{\widetilde{u}.},\widetilde{u}_t,Q_t^{\widetilde{u}.},K_t^{\widetilde{u}.},P_t^{\widetilde{u}.},\Gamma_t^{\widetilde{u}.}),\lambda:=X,Y,Z,$$

$$\zeta_t^{\widetilde{u}.}:=(X_t^{\widetilde{u}},Y_t^{\widetilde{u}.},Z_t^{\widetilde{u}.}),\Pi_t^{\widetilde{u}.}:=(Q_t^{\widetilde{u}.},K_t^{\widetilde{u}.},P_t^{\widetilde{u}.},\Gamma_t^{\widetilde{u}.}),$$

and the Hamiltonian function is given by

$$\begin{aligned}H\left(t,X,Y,Z,v,Q,K,P,\Gamma\right)&=\langle Q,\alpha X+\beta v\rangle+\left\langle K,\widehat{\alpha}X+\widehat{\beta}v\right\rangle\\&+\left\langle P,\gamma X+\widehat{\gamma}Y+\delta Z+\widehat{\delta}v\right\rangle+\left\langle\Gamma,\eta X+\widehat{\eta}Y+\theta Z+\widehat{\theta}v\right\rangle+\ (t,X,Y,Z,v)\,.\end{aligned}$$

Proof. To establish a necessary and sufficient optimality conditions, we use *The convex optimization principle* (see Ekeland-Temam ([14], prop 2.1, page 35):
Since the domain of control U is convex, the functional $\mathbb{J}$ is convex in $\widetilde{u}.$, continuous and Gâteaux-differentiable with continuous derivative $\mathbb{J}'$, we have

$$(\widetilde{u}.\text{ minimize }\mathbb{J})\Leftrightarrow\langle\mathbb{J}'\left(\widetilde{u}.\right),v.-\widetilde{u}.\rangle\geq 0;\ \forall v.\in U. \tag{2.16}$$

Let us calculate the Gâteaux derivative of $\mathbb{J}$ at a point $\widetilde{u}.$ and in the direction $(v.-\widetilde{u}.)$, we obtain

$$\begin{aligned}\langle\mathbb{J}'\left(\widetilde{u}.\right),v.-\widetilde{u}.\rangle&=\mathbb{E}\left[\langle\varphi_X(X_T^{\widetilde{u}.}),X_T^{v}-X_T^{\widetilde{u}.}\rangle\right]+\mathbb{E}\left[\langle\psi_Y(Y_0^{\widetilde{u}.}),Y_0^{v.}-Y_0^{\widetilde{u}.}\rangle\right]\\&+\mathbb{E}\left[\int_0^T\langle L_X(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t),X_t^{v.}-X_t^{\widetilde{u}.}\rangle dt\right]+\mathbb{E}\left[\int_0^T\langle L_Y(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t),Y_t^{v.}-Y_t^{\widetilde{u}.}\rangle dt\right]\\&+\mathbb{E}\left[\int_0^T\langle L_Z(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t),Z_t^{v.}-Z_t^{\widetilde{u}.}\rangle dt\right]+\mathbb{E}\left[\int_0^T\langle L_v(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t),v_t-\widetilde{u}_t\rangle dt\right].\end{aligned} \tag{2.17}$$

The system of adjoint equations (2.14) can be rewritten as follows

$$\begin{cases} dQ_t^{\widetilde{u}} = -\left(Q_t^{\widetilde{u}}\alpha_t + K_t^{\widetilde{u}}\widehat{\alpha}_t + P_t^{\widetilde{u}}\gamma_t + \Gamma_t^{\widetilde{u}}\eta_t + L_X(t,\zeta_t^{\widetilde{u}},\widetilde{u}_t)\right)dt + K_t^{\widetilde{u}}dW_t \\ \\ dP_t^{\widetilde{u}} = \left(P_t^{\widetilde{u}}\widehat{\gamma}_t + \Gamma_t^{\widetilde{u}}\widehat{\eta}_t + L_Y(t,\zeta_t^{\widetilde{u}},\widetilde{u}_t)\right)dt \\ \qquad\qquad + \left(P_t^{\widetilde{u}}\delta_t + \Gamma_t^{\widetilde{u}}\theta_t + L_Z(t,\zeta_t^{\widetilde{u}},\widetilde{u}_t)\right)dW_t - \Gamma_t^{\widetilde{u}}\overleftarrow{dB}_t \\ \\ Q_T^{\widetilde{u}.} = \varphi_X\left(X_T^{\widetilde{u}.}\right), P_0^{\widetilde{u}.} = \psi_Y\left(Y_0^{\widetilde{u}.}\right). \end{cases}$$

From (2.14), the equality (2.17) becomes

$$\begin{aligned} \langle \mathbb{J}'(\widetilde{u}.), v. - \widetilde{u}.\rangle &= \mathbb{E}\left[\langle Q_T^{\widetilde{u}}, X_T^{v.} - X_T^{\widetilde{u}.}\rangle\right] + \mathbb{E}\left[\langle P_0^{\widetilde{u}}, Y_0^{v.} - Y_0^{\widetilde{u}.}\rangle\right] \\ &+ \mathbb{E}\left[\int_0^T \langle L_X(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t), X_t^{v.} - X_t^{\widetilde{u}.}\rangle dt\right] + \mathbb{E}\left[\int_0^T \langle L_Y(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t), Y_t^{v.} - Y_t^{\widetilde{u}.}\rangle dt\right] \\ &+ \mathbb{E}\left[\int_0^T \langle L_Z(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t), Z_t^{v.} - Z_t^{\widetilde{u}.}\rangle dt\right] + \mathbb{E}\left[\int_0^T \langle L_v(t,\zeta_t^{\widetilde{u}.},\widetilde{u}_t), v_t - \widetilde{u}_t\rangle dt\right]. \end{aligned} \tag{2.18}$$

Applying integration by parts to $\langle P_t^{\widetilde{u}}, Y_t^{v.} - Y_t^{\widetilde{u}.}\rangle$ and $\langle Q_t^{\widetilde{u}}, X_t^{v.} - X_t^{\widetilde{u}.}\rangle$, passing to integral on $[0,T]$ and take the expectations we obtain

$$\begin{aligned} \mathbb{E}\left[\langle Q_T^{\widetilde{u}.}, X_T^{v.} - X_T^{\widetilde{u}.}\rangle\right] &= \mathbb{E}\left[\int_0^T \langle Q_t^{\widetilde{u}.}, \alpha_t(X_t^{v.} - X_t^{\widetilde{u}.}) + \beta_t(v_t - \widetilde{u}_t)\rangle dt\right] \\ &- \mathbb{E}\left[\int_0^T \langle Q_t^{\widetilde{u}}\alpha_t + K_t^{\widetilde{u}}\widehat{\alpha}_t + P_t^{\widetilde{u}}\gamma_t + \Gamma_t^{\widetilde{u}}\eta_t + L_X(t,\zeta_t^{\widetilde{u}},\widetilde{u}_t), X_t^{v.} - X_t^{\widetilde{u}.}\rangle dt\right] \\ &+ \mathbb{E}\left[\int_0^T \langle K_t^{\widetilde{u}.}, \widehat{\alpha}_t(X_t^{v.} - X_t^{\widetilde{u}.}) + \widehat{\beta}_t(v_t - \widetilde{u}_t)\rangle dt\right], \end{aligned} \tag{2.19}$$

and

$$\mathbb{E}\left[\langle P_0^{\widetilde{u}.}, Y_0^{v.} - Y_0^{\widetilde{u}.}\rangle\right] = -\mathbb{E}\left[\int_0^T \langle P_t^{\widetilde{u}}\widehat{\gamma}_t + \Gamma_t^{\widetilde{u}}\widehat{\eta}_t + L_Y(t, \zeta_t^{\widetilde{u}}, \widetilde{u}_t, Y_t^{v.} - Y_t^{\widetilde{u}.}\rangle dt\right] \tag{2.20}$$
$$+ \mathbb{E}\left[\int_0^T \langle P_t^{\widetilde{u}.}, \gamma_t(X_t^{v.} - X_t^{\widetilde{u}.}) + \widehat{\gamma}_t(Y_t^{v.} - Y_t^{\widetilde{u}.}) + \delta_t(Z_t^{v.} - Z_t^{\widetilde{u}.})\right.$$
$$\left. + \widehat{\delta}_t(v_t - \widetilde{u}_t)\rangle dt\right] + \mathbb{E}\left[\int_0^T \langle \Gamma_t^{\widetilde{u}.}, \eta_t(X_t^{v.} - X_t^{\widetilde{u}.}) + \widehat{\eta}_t(Y_t^{v.} - Y_t^{\widetilde{u}.})\right.$$
$$\left. + \theta_t(Z_t^{v.} - Z_t^{\widetilde{u}.}) + \widehat{\theta}_t(v_t - \widetilde{u}_t)\rangle dt\right]$$
$$- \mathbb{E}\left[\int_0^T \langle P_t^{\widetilde{u}}\delta_t + \Gamma_t^{\widetilde{u}}\theta_t + L_Z(t, \zeta_t^{\widetilde{u}}, \widetilde{u}_t), Z_t^{v.} - Z_t^{\widetilde{u}.}\rangle dt\right].$$

Replacing (2.19), (2.20) in (2.18) we get

$$\langle \mathbb{J}'(\widetilde{u}), v. - \widetilde{u}.\rangle = \mathbb{E}\left[\int_0^T \langle Q_t^{\widetilde{u}.}\beta_t + K_t^{\widetilde{u}.}\widehat{\beta}_t + P_t^{\widetilde{u}.}\widehat{\delta}_t\right.$$
$$\left. + \Gamma_t^{\widetilde{u}.}\widehat{\theta}_t + L_v(t, \zeta_t^{\widetilde{u}.}, \widetilde{u}_t), v_t - \widetilde{u}_t)\rangle dt\right].$$

On the other hand, we calculate the Gâteaux derivative of H at a point $\widetilde{u}.$ and in the direction $(v. - \widetilde{u}.)$, we have

$$\mathbb{E}\left[\int_0^T \langle H_v(t, \zeta_t^{\widetilde{u}.}, \widetilde{u}_t, \Pi_t^{\widetilde{u}.}), v_t - \widetilde{u}_t\rangle dt\right] = \mathbb{E}\left[\int_0^T \langle Q_t^{\widetilde{u}.}\beta_t + K_t^{\widetilde{u}.}\widehat{\beta}_t + P_t^{\widetilde{u}.}\widehat{\delta}_t\right.$$
$$\left. + \Gamma_t^{\widetilde{u}.}\widehat{\theta}_t + L_v(t, \zeta_t^{\widetilde{u}.}, \widetilde{u}_t), v_t - \widetilde{u}_t)\rangle dt\right] \tag{2.21}$$
$$= \langle \mathbb{J}'(\widetilde{u}), v. - \widetilde{u}.\rangle$$

From (2.21) and (2.16), we have

$$(\widetilde{u} \text{ minimize } \mathbb{J}) \Leftrightarrow \mathbb{E}\left[\int_0^T \langle H_v(t, \zeta_t^{\widetilde{u}.}, \widetilde{u}_t, \Pi_t^{\widetilde{u}.}), v_t - \widetilde{u}_t\rangle dt\right] \geq 0, \ \forall v \in \mathcal{U}.$$

Which implies that

$$\mathbb{E}\left[\langle H_v(t, \zeta_t^{\widetilde{u}.}, \widetilde{u}_t, \Pi_t^{\widetilde{u}.}), v_t - \widetilde{u}_t\rangle\right] \geq 0, \ dt\text{-a.e.}$$

Now, let Π be an arbitrary element of the σ-algebra $\mathcal{F}_t$, and set

$$\pi_t = v_t 1_\Pi + \widetilde{u}_t 1_{\Omega-\Pi}.$$

It is not difficult to see that the control π is an element of $\mathcal{U}$.

Applying the above inequality with , we obtain

$$\mathbb{E}\left[\langle 1_\Pi H_v(t, \zeta_t^{\widetilde{u}.}, \widetilde{u}_t, \Pi_t^{\widetilde{u}.}), v_t - \widetilde{u}_t\rangle\right] \geq 0, \forall \Pi \in \mathcal{F}_t.$$

Which implies that

$$\mathbb{E}\left[\langle H_v(t, \zeta_t^{\widetilde{u}.}, \widetilde{u}_t, \Pi_t^{\widetilde{u}.}), v_t - \widetilde{u}_t\rangle \mid \mathcal{F}_t\right] \geq 0.$$

The quantity inside the conditional expectation is $\mathcal{F}_t-$measurable, and thus the result follows. ■

Chapter 3

Optimal control problem for a linear MF-FBDSDEs

In this chapter, we prove the existence of a strong optimal strict control which is adapted to the initial σ-algebra, under the convexity of the cost function and the action space U. Here, the systems governed by linear forward-backward doubly stochastic differential equations of mean field type, in which the coefficients depend on the state process, and also on the distribution of the state process, via the expectation of the state. Moreover, we establish necessary as well as sufficient optimality conditions for this kind of control problem.

3.1 Formulation of the problem and assumptions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Let $(W_t)_{t\in[0,T]}$ and $(B_t)_{t\in[0,T]}$ be two mutually independent standard Brownian motions, with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^l$.
Let $\mathcal{N}$ denote the class of $\mathbb{P}$-null sets of $\mathcal{F}$. For each $t \in [0,T]$, we define $\mathcal{F}_t := \mathcal{F}_t^W \vee \mathcal{F}_{t,T}^B$, where for any process $\{\delta_t\}$, we set $\mathcal{F}_{s,t}^\delta = \sigma(\delta_r - \delta_s; s \leq r \leq t) \vee \mathcal{N}, \mathcal{F}_t^\delta = \mathcal{F}_{0,t}^\delta$.
Note that the collection $\{\mathcal{F}_t, t \in [0,T]\}$ is neither increasing nor decreasing, then it does not constitute a classical filtration.

Given ξ a square integrable and $\mathcal{F}_T$-measurable process, x a square integrable and $\mathcal{F}_0$-measurable process and for any admissible control u.

We consider a control problem governed by the following controlled linear FBDSDE of mean-field type:

$$\begin{cases} dy_t^u = b(t, y_t^u, \mathbb{E}[y_t^u], u_t)dt + \sigma(t, y_t^u, \mathbb{E}[y_t^u], u_t)dW_t \\ \\ dY_t^u = -f(t, y_t^u, \mathbb{E}[y_t^u], Y_t^u, \mathbb{E}[Y_t^u], Z_t^u, \mathbb{E}[Z_t^u], u_t)dt \\ \qquad\qquad -g(t, y_t^u, \mathbb{E}[y_t^u], Y_t^u, \mathbb{E}[Y_t^u], Z_t^u, \mathbb{E}[Z_t^u], u_t)\overleftarrow{dB}_t + Z_t^u dW_t, \\ y_0^u = x, Y_T = h(y_T^u, \mathbb{E}[y_T^u]), \end{cases} \tag{3.1}$$

with

$$b(t, y_t^u, \mathbb{E}[y_t^u], u_t) = a_t y_t^u + \widehat{a}_t \mathbb{E}\left[y_t^u\right] + b_t u_t,$$

$$\sigma(t, y_t^u, \mathbb{E}[y_t^u], u_t) = c_t . y_t^u + \widehat{c}_t \mathbb{E}\left[y_t^u\right] + \widehat{b}_t u_t,$$

$$\begin{aligned} f(t, y_t^u, \mathbb{E}[y_t^u], Y_t^u, \mathbb{E}[Y_t^u], Z_t^u, \mathbb{E}[Z_t^u], u_t) &= d_t y_t^u + \widehat{d}_t \mathbb{E}\left[y_t^u\right] + e_t Y_t^u + \widehat{e}_t \mathbb{E}\left[Y_t^u\right] \\ &\quad + f_t Z_t^u + \widehat{f}_t \mathbb{E}\left[Z_t^u\right] + g_t u_t, \end{aligned}$$

$$\begin{aligned} g(t, y_t^u, \mathbb{E}[y_t^u], Y_t^u, \mathbb{E}[Y_t^u], Z_t^u, \mathbb{E}[Z_t^u], u_t) &= h_t y_t^u + \widehat{h}_t \mathbb{E}\left[y_t^u\right] + k_t Y_t^u + \widehat{k}_t \mathbb{E}\left[Y_t^u\right] \\ &\quad + m_t Z_t^u + \widehat{m}_t \mathbb{E}\left[Z_t^u\right] + \widehat{g}_t u_t, \end{aligned}$$

$$h(y_T^u, \mathbb{E}[y_T^u]) = \xi,$$

and a cost functional:

$$\mathbb{J}(u.) := \mathbb{E}\left[\alpha\left(y_T^u, \mathbb{E}\left[y_T^u\right]\right) + \beta\left(Y_0^u, \mathbb{E}\left[Y_0^u\right]\right) + \int_0^T \ell\left(t, y_t^u, \mathbb{E}\left[y_t^u\right], Y_t^u, \mathbb{E}\left[Y_t^u\right], Z_t^u, \mathbb{E}\left[Z_t^u\right], u_t\right) dt\right], \tag{3.2}$$

where $a_{\cdot}, \widehat{a}_{\cdot}, b_{\cdot}, \widehat{b}_{\cdot}, c_{\cdot}, \widehat{c}_{\cdot}, d_{\cdot}, \widehat{d}_{\cdot}, e_{\cdot}, \widehat{e}_{\cdot}, f_{\cdot}, \widehat{f}_{\cdot}, g_{\cdot}, \widehat{g}_{\cdot}, h_{\cdot}, \widehat{h}_{\cdot}, k_{\cdot}, \widehat{k}_{\cdot}, m_{\cdot}$ and $\widehat{m}_{\cdot}$ are matrix-valued functions of suitable sizes. The solution $(y_{\cdot}, Y_{\cdot}, Z_{\cdot})$ takes values in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ and $u_{\cdot}$ is the control variable values in subset U of $\mathbb{R}^k$. α, β, ℓ are a given functions define by

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

$$\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$

$$\beta : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}.$$

Definition 3.1.1 *An admissible control $u_{\cdot}$ is a square integrable, $\mathcal{F}_t$-measurable process with values in some subset $U \subseteq \mathbb{R}^k$. We denote by $\mathcal{U}$ the set of all admissible controls.*

Note that we have an additional constraint that a control must be square-integrable just to ensure the existence of solutions of (3.1) under $u_{\cdot}$. We say that an admissible control $u^*_{\cdot} \in \mathcal{U}$ is an optimal control if

$$\mathbb{J}\left(u^*_{\cdot}\right) = \inf_{v_{\cdot} \in \mathcal{U}} \mathbb{J}\left(v_{\cdot}\right). \tag{3.3}$$

The following notations are needed

$\mathcal{M}^2_{\mathcal{F}}\left(0,T;\mathbb{R}^m\right)$: the set of process $\pi_{\cdot}$, $\mathcal{F}_t$-adapted with values in $\mathbb{R}^m$ such that

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

$\mathcal{S}^2_{\mathcal{F}}\left(0,T;\mathbb{R}^n\right)$: the set of process $\eta_{\cdot}$, $\mathcal{F}_t$-adapted and $\mathbb{R}^n$-valued continuous processes such that

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

$\mathcal{U} := \left\{v_{\cdot} \in \mathcal{M}^2_{\mathcal{F}}\left(0,T;\mathbb{R}^k\right) / v_t \in U, a.e.t \in [0,T], \mathbb{P} - a.s.\right\}.$

We shall consider the following assumptions

$(\mathbf{H1})$: the set $U \subseteq \mathbb{R}^k$ is convex and compact and the functions ℓ, α and β are continuous, bounded and convex,

$(\mathbf{H2})$: $a_t, \widehat{a}_t, b_t, \widehat{b}_t, c_t, \widehat{c}_t, d_t, \widehat{d}_t, e_t, \widehat{e}_t, f_t, \widehat{f}_t, g_t, \widehat{g}_t, h_t, \widehat{h}_t, k_t$ and $\widehat{k}_t$ are bounded by $\lambda > 0$ and $m_t, \widehat{m}_t$ are bounded by $\gamma \in \left]0, \frac{1}{2}\right[$. That is:

$$\lambda := \sup_{t,\omega} |\varphi_t(\omega)| \text{ and } \gamma := \sup_{t,\omega} |\sigma_t(\omega)|,$$

where $\varphi_t(\omega) = a_t, \widehat{a}_t, b_t, \widehat{b}_t, c_t, \widehat{c}_t, d_t, \widehat{d}_t, e_t, \widehat{e}_t, f_t, \widehat{f}_t, g_t, \widehat{g}_t, h_t, \widehat{h}_t, k_t, \widehat{k}_t$ and $\sigma_t = m_t, \widehat{m}_t$.

Proposition 3.1.1 *Under assumptions* $(\mathbf{H1}) - (\mathbf{H2})$ *the system of linear FBDSDE of mean-field type (3.1), has a unique strong solution.*

Proof. The proof of this proposition is established in Zhu and Shi [37], by using a method of continuation, and the fact that our system (3.1) is a special case of the one given in the paper [37]. ■

Remark 3.1.1 *A special case is that in which both* α, β *and* l *are convex quadratic functions. The control problem* $\{(3.1), (3.2), (3.3)\}$ *is then reduced to a stochastic linear quadratic optimal control problem.*

3.2 Existence of a strong optimal control

The following theorem confirms the existence of a strong optimal solutions for the control problem $\{(3.1), (3.2), (3.3)\}$.

Theorem 3.2.1 *Under either* $(\mathbf{H1})-(\mathbf{H2})$*, if the strict control problem* $\{(3.1), (3.2), (3.3)\}$ *is finite, then it admits an optimal strong solution.*

Proof. Assume that $(\mathbf{H1})$-$(\mathbf{H2})$ holds. Let $(u^n_\cdot)$ be a minimizing sequence, i.e.,

$$\lim_{n\to\infty} \mathbb{J}(u^n_\cdot) = \inf_{v_\cdot \in \mathcal{U}} \mathbb{J}(v_\cdot).$$

With associated trajectories $\left(y_\cdot^{u_\cdot^n}, Y_\cdot^{u_\cdot^n}, Z_\cdot^{u_\cdot^n}\right)$ satisfies the linear FBDSDE of mean-field type (3.1). From the fact that U is a compact set, there exists a subsequence (which is still labeled by $(u_\cdot^n)_{n\geq 0}$) such that

$$u_\cdot^n \longrightarrow \overline{u}_\cdot, \text{ weakly in } \mathcal{M}_{\mathcal{F}}^2\left(0,T;\mathbb{R}^k\right).$$

Applying Mazur's theorem, there is a sequence of convex combinations

$$\widetilde{U}_\cdot^n = \sum_{\jmath\geq 0}\theta_{\jmath n} u_\cdot^{\jmath+n} \quad (\text{with } \theta_{\jmath n}\geq 0, \quad \text{and } \sum_{\jmath\geq 0}\theta_{\jmath n}=1),$$

such that

$$\widetilde{U}_\cdot^n \rightarrow \overline{u}_\cdot \text{ strongly in } \mathcal{M}_{\mathcal{F}}^2\left(0,T;\mathbb{R}^k\right). \tag{3.4}$$

Since the set $U\subseteq\mathbb{R}^k$ is convex and compact, it follows that $\overline{u}_\cdot\in\mathcal{U}$. Let $(y_\cdot^{\widetilde{U}_\cdot^n}, Y_\cdot^{\widetilde{U}_\cdot^n}, Z_\cdot^{\widetilde{U}_\cdot^n})$ and $(y_\cdot^{\overline{u}_\cdot}, Y_\cdot^{\overline{u}_\cdot}, Z_\cdot^{\overline{u}_\cdot})$ be the solutions of the linear MF-FBDSDE (3.1), associated with $\widetilde{U}_\cdot^n$ and $\overline{u}_\cdot$ respectively i.e.,

$$\begin{cases} dy_t^{\widetilde{U}_\cdot^n} = \left(a_t y_t^{\widetilde{U}_\cdot^n} + \widehat{a}_t\mathbb{E}[y_t^{\widetilde{U}_\cdot^n}] + b_t\widetilde{U}_t^n\right)dt + \left(c_t y_t^{\widetilde{U}_\cdot^n} + \widehat{c}_t\mathbb{E}[y_t^{\widetilde{U}_\cdot^n}] + \widehat{b}_t\widetilde{U}_t^n\right)dW_t \\ \\ dY_t^{\widetilde{U}_\cdot^n} = -\left(d_t y_t^{\widetilde{U}_\cdot^n} + \widehat{d}_t\mathbb{E}[y_t^{\widetilde{U}_\cdot^n}] + e_t Y_t^{\widetilde{U}_\cdot^n} + \widehat{e}_t\mathbb{E}[Y_t^{\widetilde{U}_\cdot^n}] + f_t Z_t^{\widetilde{U}_\cdot^n} + \widehat{f}_t\mathbb{E}[Z_t^{\widetilde{U}_\cdot^n}] + g_t\widetilde{U}_t^n\right)dt \\ \qquad -\left(h_t y_t^{\widetilde{U}_\cdot^n} + \widehat{h}_t\mathbb{E}[y_t^{\widetilde{U}_\cdot^n}] + k_t Y_t^{\widetilde{U}_\cdot^n} + \widehat{k}_t\mathbb{E}[Y_t^{\widetilde{U}_\cdot^n}] + m_t Z_t^{\widetilde{U}_\cdot^n} + \widehat{m}_t\mathbb{E}[Z_t^{\widetilde{U}_\cdot^n}] \right. \\ \qquad\qquad \left. +\widehat{g}_t\widetilde{U}_t^n\right)\overleftarrow{dB}_t + Z_t^{\widetilde{U}_\cdot^n}dW_t, \\ y_0^{\widetilde{U}_\cdot^n} = x, Y_T^{\widetilde{U}_\cdot^n} = \xi, \end{cases} \tag{3.5}$$

and

$$\begin{cases} dy_t^{\overline{u}.} = \big(a_t y_t^{\overline{u}.} + \widehat{a}_t \mathbb{E}[y_t^{\overline{u}.}] + b_t \overline{u}_t\big)dt + \big(c_t y_t^{\overline{u}.} + \widehat{c}_t \mathbb{E}[y_t^{\overline{u}.}] + \widehat{b}_t \overline{u}_t\big)dW_t \\ \\ dY_t^{\overline{u}.} = -\big(d_t y_t^{\overline{u}.} + \widehat{d}_t \mathbb{E}[y_t^{\overline{u}.}] + e_t Y_t^{\overline{u}.} + \widehat{e}_t \mathbb{E}[Y_t^{\overline{u}.}] + f_t Z_t^{\overline{u}.} + \widehat{f}_t \mathbb{E}[Z_t^{\overline{u}.}] + g_t \overline{u}_t\big)dt \\ \qquad -\big(h_t y_t^{\overline{u}.} + \widehat{h}_t \mathbb{E}[y_t^{\overline{u}.}] + k_t Y_t^{\overline{u}.} + \widehat{k}_t \mathbb{E}[Y_t^{\overline{u}.}] + m_t Z_t^{\overline{u}.} + \widehat{m}_t \mathbb{E}[Z_t^{\overline{u}.}] \\ \qquad\qquad + \widehat{g}_t \overline{u}_t\big)\overleftarrow{dB}_t + Z_t^{\overline{u}.} dW_t, \\ y_0^{\overline{u}.} = x, Y_T^{\overline{u}.} = \xi. \end{cases} \tag{3.6}$$

Then let us prove

$$(y_t^{\widetilde{U}^n_.}, Y_t^{\widetilde{U}^n_.}, \int_0^T Z_s^{\widetilde{U}^n_.} dW_s) \text{ converges strongly to } (y_t^{\overline{u}.}, Y_t^{\overline{u}.}, \int_0^T Z_s^{\overline{u}.} dW_s), \tag{3.7}$$

in $\mathcal{S}^2_{\mathcal{F}}(0, T; \mathbb{R}^{n+m}) \times \mathcal{M}^2_{\mathcal{F}}\big(0, T; \mathbb{R}^{m\times d}\big)$.

Firstly, we have

$$\begin{aligned} (\sup_{0\le s\le t} |y_s^{\widetilde{U}^n_.} - y_s^{\overline{u}.}|^2) \le \int_0^t & \big(|a_s|^2 (\sup_{0\le r\le s} |y_r^{\widetilde{U}^n_.} - y_r^{\overline{u}.}|^2) + |\widehat{a}_s|^2 \mathbb{E}[\sup_{0\le r\le s} |y_r^{\widetilde{U}^n_.} - y_r^{\overline{u}.}|^2] \\ & + |b_s|^2 |\widetilde{U}^n_s - \overline{u}_s|^2\big) ds + \sup_{0\le s\le t} (\big|\int_0^t \big(c_s (y_s^{\widetilde{U}^n_.} - y_s^{\overline{u}.}) \\ & + \overline{c}_s (\mathbb{E}[y_s^{\widetilde{U}^n_.} - y_s^{\overline{u}.}]) + \widehat{b}_s (\widetilde{U}^n_s - \overline{u}_s)\big) dW_s\big|)^2, \end{aligned}$$

using the Burkholder-Davis-Gundy inequality to the martingale part, we can show

$$\mathbb{E}\big[\sup_{0\le s\le T} |y_s^{\widetilde{U}^n_.} - y_s^{\overline{u}.}|^2\big] \le K \int_0^t \mathbb{E}\big[\sup_{0\le r\le s} |y_r^{\widetilde{U}^n_.} - y_r^{\overline{u}.}|^2\big] ds + K' \mathbb{E}\big[\int_0^t |\widetilde{U}^n_s - \overline{u}_s|^2 ds\big].$$

Applying Gronwall's lemma and using (3.4), we get

$$\lim_{n\to\infty} \mathbb{E}\big[\sup_{0\le s\le T} |y_s^{\widetilde{U}^n_.} - y_s^{\overline{u}.}|^2\big] = 0. \tag{3.8}$$

Secondly, applying Itô's formula to $\Big|Y_t^{\widetilde{U}^n_\cdot} - Y_t^{\overline{u}_\cdot}\Big|^2$ and taking expectation, we get

$$\begin{aligned}
&\mathbb{E}\big[\sup_{0\leq t\leq T}|Y_t^{\widetilde{U}^n_\cdot} - Y_t^{\overline{u}_\cdot}|^2\big] + \mathbb{E}\big[\int_0^T \|Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}\|^2 ds\big] \leq \\
&2\mathbb{E}\big[\int_t^T \langle Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}, d_s(y_s^{\widetilde{U}^n_\cdot} - y_s^{\overline{u}_\cdot}) + \widehat{d}_s\mathbb{E}[y_s^{\widetilde{U}^n_\cdot} - y_s^{\overline{u}_\cdot}] + e_s(Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}) \\
&+\widehat{e}_s\mathbb{E}[Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}] + f_s(Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}) + \widehat{f}_s\mathbb{E}[Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}] + g_s(\widetilde{U}^n_s - \overline{u}_s)\rangle ds\big] \\
&+\mathbb{E}\big[\int_0^T |h_s(y_s^{\widetilde{U}^n_\cdot} - y_s^{\overline{u}_\cdot}) + \widehat{h}_s\mathbb{E}[y_s^{\widetilde{U}^n_\cdot} - y_s^{\overline{u}_\cdot}] + k_s(Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}) \\
&+\widehat{k}_s\mathbb{E}[Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}] + m_s(Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}) + \widehat{m}_s\mathbb{E}[Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}] + \widehat{g}_s(\widetilde{U}^n_s - \overline{u}_s)|^2 ds\big].
\end{aligned}$$

According to the assumption (H2) and by using the Young's formula, we obtain

$$\begin{aligned}
&\mathbb{E}\big[\sup_{0\leq t\leq T}|Y_t^{\widetilde{U}^n_\cdot} - Y_t^{\overline{u}_\cdot}|^2\big] + \mathbb{E}\big[\int_0^T \|Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}\|^2 ds\big] \\
&\leq \frac{1}{\rho_1}\mathbb{E}\big[\int_0^T |Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}|^2 ds\big] + 14\rho_1\lambda^2\mathbb{E}\big[\int_0^T \big(|y_s^{\widetilde{U}^n_\cdot} - y_s^{\overline{u}_\cdot}|^2 + |Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}|^2 \\
&+\|Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}\|^2 + \frac{1}{2}|\widetilde{U}^n_s - \overline{u}_s|^2\big) ds\big] + 10\lambda^2\mathbb{E}\big[\int_0^T \big(|y_s^{\widetilde{U}^n_\cdot} - y_s^{\overline{u}_\cdot}|^2 \\
&+|Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}|^2 + \frac{1}{2}|\widetilde{U}^n_s - \overline{u}_s|^2\big) ds\big] + 4\gamma^2\mathbb{E}\big[\int_0^T \|Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}\|^2 ds\big] \\
&+\frac{5\lambda\gamma}{\rho_2}\mathbb{E}\big[\int_0^T \big(|y_s^{\widetilde{U}^n_\cdot} - y_s^{\overline{u}_\cdot}|^2 + \mathbb{E}[|y_s^{\widetilde{U}^n_\cdot} - y_s^{\overline{u}_\cdot}|^2] + |Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}|^2 \\
&+\mathbb{E}[|Y_s^{\widetilde{U}^n_\cdot} - Y_s^{\overline{u}_\cdot}|^2] + |\widetilde{U}^n_s - \overline{u}_s|^2\big) ds\big] \\
&+2\rho_2\lambda\gamma\mathbb{E}\big[\int_0^T \big(\|Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}\|^2 + \mathbb{E}[\|Z_s^{\widetilde{U}^n_\cdot} - Z_s^{\overline{u}_\cdot}\|^2]\big) ds\big],
\end{aligned}$$

and therefore

$$
\begin{aligned}
&\mathbb{E}\big[\sup_{0\leq t\leq T}|Y_t^{\widetilde{U}^n_\cdot}-Y_t^{\overline{u}_\cdot}|^2\big]+\mathbb{E}\big[\int_0^T\|Z_s^{\widetilde{U}^n_\cdot}-Z_s^{\overline{u}_\cdot}\|^2ds\big]\\
&\leq(\frac{1}{\rho_1}+14\rho_1\lambda^2+10\lambda^2+\frac{10\lambda\gamma}{\rho_2})\mathbb{E}\big[\int_0^T|Y_s^{\widetilde{U}^n_\cdot}-Y_s^{\overline{u}_\cdot}|^2ds\big]\\
&+(14\rho_1\lambda^2+4\gamma^2+4\rho_2\lambda\gamma)\mathbb{E}\big[\int_0^T\|Z_s^{\widetilde{U}^n_\cdot}-Z_s^{\overline{u}_\cdot}\|^2ds\big]\\
&+(14\rho_1\lambda^2+10\lambda^2+\frac{10\lambda\gamma}{\rho_2})\mathbb{E}\big[\int_0^T|y_s^{\widetilde{U}^n_\cdot}-y_s^{\overline{u}_\cdot}|^2ds\big]\\
&+(7\rho_1\lambda^2+5\lambda^2+\frac{5\lambda\gamma}{\rho_2})\mathbb{E}\big[\int_0^T|\widetilde{U}^n_s-\overline{u}_s|^2ds\big].
\end{aligned}
$$

Choosing

$$
\rho_1=\frac{1-4\gamma^2}{28\lambda^2}>0 \text{ and } \rho_2=\frac{1-4\gamma^2}{12\lambda\gamma}>0 \text{ because } 0<\gamma<\frac{1}{2},
$$

the previous inequality becomes

$$
\begin{aligned}
\mathbb{E}\big[\sup_{0\leq t\leq T}|Y_t^{\widetilde{U}^n_\cdot}-Y_t^{\overline{u}_\cdot}|^2\big]+\mu_1\mathbb{E}\big[\int_0^T\|Z_s^{\widetilde{U}^n_\cdot}-Z_s^{\overline{u}_\cdot}\|^2ds\big]&\leq\mu_2\mathbb{E}\big[\int_0^T|Y_s^{\widetilde{U}^n_\cdot}-Y_s^{\overline{u}_\cdot}|^2ds\big]\\
+\mu_3\mathbb{E}\big[\int_0^T|y_s^{\widetilde{U}^n_\cdot}-y_s^{\overline{u}_\cdot}|^2ds\big]+\mu_4\mathbb{E}\big[\int_0^T|\widetilde{U}^n_s-\overline{u}_s|^2ds\big],& \qquad (3.9)
\end{aligned}
$$

where

$$
\begin{aligned}
\mu_1&=\frac{1-4\gamma^2}{6}>0,\\
\mu_2&=\frac{28\lambda^2}{1-4\gamma^2}+\frac{1-4\gamma^2}{2}+10\lambda^2+\frac{120\,(\lambda\gamma)^2}{1-4\gamma^2}>0,\\
\mu_3&=\frac{1-4\gamma^2}{2}+10\lambda^2+\frac{120\,(\lambda\gamma)^2}{1-4\gamma^2}>0,\\
\mu_4&=\frac{1-4\gamma^2}{4}+5\lambda^2+\frac{60\,(\lambda\gamma)^2}{1-4\gamma^2}>0.
\end{aligned}
$$

We derive two inequalities from (3.9),

$$\begin{aligned}\mathbb{E}\big[\sup_{0\leq t\leq T}|Y_t^{\widetilde{U}^n_\cdot}-Y_t^{\overline{u}_\cdot}|^2\big]\leq \mu_2\mathbb{E}\big[\int_0^T|Y_s^{\widetilde{U}^n_\cdot}-Y_s^{\overline{u}_\cdot}|^2ds\big]\\+\mu_3\mathbb{E}\big[\int_0^T|y_s^{\widetilde{U}^n_\cdot}-y_s^{\overline{u}_\cdot}|^2ds\big]+\mu_4\mathbb{E}\big[\int_0^T|\widetilde{U}^n_s-\overline{u}_s|^2ds\big],\end{aligned}\tag{3.10}$$

and

$$\begin{aligned}\mu_1\mathbb{E}\big[\int_0^T\|Z_s^{\widetilde{U}^n_\cdot}-Z_s^{\overline{u}_\cdot}\|^2ds\big]\leq \mu_2\mathbb{E}\big[\int_0^T|Y_s^{\widetilde{U}^n_\cdot}-Y_s^{\overline{u}_\cdot}|^2ds\big]\\+\mu_3\mathbb{E}\big[\int_0^T|y_s^{\widetilde{U}^n_\cdot}-y_s^{\overline{u}_\cdot}|^2ds\big]+\mu_4\mathbb{E}\big[\int_0^T|\widetilde{U}^n_s-\overline{u}_s|^2ds\big].\end{aligned}\tag{3.11}$$

Using Burkholder-Davis-Gundy's inequality, applying Gronwall's lemma to (3.10) and passing to the limit as $n\to\infty$, and using the convergence (3.4) and (3.17), we obtain

$$\lim_{n\to\infty}\mathbb{E}\big[\sup_{0\leq s\leq T}|Y_s^{\widetilde{U}^n_\cdot}-Y_s^{\overline{u}_\cdot}|^2\big]=0.\tag{3.12}$$

Then, one can shows directly from (3.4),(3.8) and (3.11) that

$$\mathbb{E}\big[\int_0^T\|Z_s^{\widetilde{U}^n_\cdot}-Z_s^{\overline{u}_\cdot}\|^2ds\big]\longrightarrow 0,\ \text{as } n\to\infty,$$

which gives the result by applying the isometry of Itô. Finally, let us prove that $\overline{u}_\cdot$ is an optimal control. Using the continuity of functions α,β and ℓ, we get

$$\begin{aligned}\mathbb{J}\left(\overline{u}_\cdot\right)&=\mathbb{E}\big[\alpha\left(y_T^{\overline{u}_\cdot},\mathbb{E}\left[y_T^{\overline{u}_\cdot}\right]\right)+\beta\left(Y_0^{\overline{u}_\cdot},\mathbb{E}\left[Y_0^{\overline{u}_\cdot}\right]\right)\\&\quad+\textstyle\int_0^T\ell\big(t,y_t^{\overline{u}_\cdot},\mathbb{E}\left[y_t^{\overline{u}_\cdot}\right],Y_t^{\overline{u}_\cdot},\mathbb{E}\left[Y_t^{\overline{u}_\cdot}\right],Z_t^{\overline{u}_\cdot},\mathbb{E}\left[Z_t^{\overline{u}_\cdot}\right],\overline{u}_t\big)dt\big]\\&=\lim_{n\to\infty}\mathbb{E}\big[\alpha\left(y_T^{\widetilde{U}^n_\cdot},\mathbb{E}\left[y_T^{\widetilde{U}^n_\cdot}\right]\right)+\beta\left(Y_0^{\widetilde{U}^n_\cdot},\mathbb{E}\left[Y_0^{\widetilde{U}^n_\cdot}\right]\right)\\&\quad+\textstyle\int_0^T\ell\big(t,y_t^{\widetilde{U}^n_\cdot},\mathbb{E}\left[y_t^{\widetilde{U}^n_\cdot}\right],Y_t^{\widetilde{U}^n_\cdot},\mathbb{E}\left[Y_t^{\widetilde{U}^n_\cdot}\right],Z_t^{\widetilde{U}^n_\cdot},\mathbb{E}\left[Z_t^{\widetilde{U}^n_\cdot}\right],\widetilde{U}^n_t\big)dt\big].\end{aligned}$$

By the convexity of α, β and ℓ, it follows that

$$\begin{aligned}\mathbb{J}\left(\overline{u}.\right) \leq & \lim_{n\rightarrow\infty}\sum_{\jmath\geq 0}\theta_{\jmath n}\mathbb{E}\Big[\alpha\left(y_T^{u_\cdot^{\jmath+n}}, \mathbb{E}\left[y_T^{u_\cdot^{\jmath+n}}\right]\right) + \beta\left(Y_0^{u_\cdot^{\jmath+n}}, \mathbb{E}\left[Y_0^{u_\cdot^{\jmath+n}}\right]\right) \\ & + \int_0^T \ell\big(t, y_t^{u_\cdot^{\jmath+n}}, \mathbb{E}\left[y_t^{u_\cdot^{\jmath+n}}\right], Y_t^{u_\cdot^{\jmath+n}}, \mathbb{E}\left[Y_t^{u_\cdot^{\jmath+n}}\right], Z_t^{u_\cdot^{\jmath+n}}, \mathbb{E}\left[Z_t^{u_\cdot^{\jmath+n}}\right], u_t^{\jmath+n}\big)dt\Big] \\ & = \lim_{n\rightarrow\infty}\sum_{k\geq 0}\theta_{\jmath n}\mathbb{J}\left(u_\cdot^{\jmath+n}\right) \leq \lim_{n\rightarrow\infty}\underset{1\leq\jmath\leq i_n}{Max}\,\mathbb{J}\left(u_\cdot^{\jmath+n}\right)\sum_{\jmath\geq 1}\theta_{\jmath n} = \inf_{v.\in\mathcal{U}}\mathbb{J}\left(v.\right).\end{aligned}$$

This completes the proof. ■

3.3 Necessary and sufficient conditions for optimality

In this section, we establish necessary as well as sufficient optimality conditions for a strict control problem driven by a linear MF-FBDSDE. In this end, we use the convex perturbation method because the domain of control U is convex.

Let $(\overline{u}., y_t^{\overline{u}.}, Y_t^{\overline{u}.}, Z_t^{\overline{u}.})$ be the optimal solution of the control problem $\{(3.1), (3.2), (3.3)\}$ obtained in section (3.2). Let us define the perturbed control as follow: for each admissible control $v.$

$$u_t^{\varepsilon} = \overline{u}_t + \varepsilon\left(v_t - \overline{u}_t\right),$$

where, $\varepsilon > 0$ is sufficiently small.

It's clear that $u_\cdot^{\varepsilon}$ is admissible control and let $\left(y_t^{u_\cdot^{\varepsilon}}, Y_t^{u_\cdot^{\varepsilon}}, Z_t^{u_\cdot^{\varepsilon}}\right)$ be the solution of (3.1) corresponding to $u_\cdot^{\varepsilon}$.

The necessary conditions for optimality will be derived by using the optimality of $\overline{u}.$ and the following inequality,

$$\begin{aligned}0 &\leq \lim_{\varepsilon\rightarrow 0}\frac{1}{\varepsilon}\left(\mathbb{J}\left(u_\cdot^{\varepsilon}\right) - \mathbb{J}\left(\overline{u}.\right)\right) \\ &= \lim_{\varepsilon\rightarrow 0}\frac{1}{\varepsilon}\left(\mathbb{J}\left(\overline{u}. + \varepsilon\left(v. - \overline{u}.\right)\right) - \mathbb{J}\left(\overline{u}.\right)\right) \\ &= \left\langle \mathbb{J}'\left(\overline{u}.\right), v. - \overline{u}.\right\rangle.\end{aligned}$$

Considering in this section the following assumptions

(**H3**) (Regularity conditions)

$$\begin{cases} (i) \text{ the function } \ell \text{ is continuously differentiable with respect to} \\ \quad (y,y',Y,Y',Z,Z',v), \text{ and the mappings } \alpha \text{ and } \beta \text{ are continuously} \\ \quad \text{differentiable with respect to } (y,y') \text{ and } (Y,Y'), \text{ respectively,} \\ (ii) \text{ the derivatives of } \ell,\alpha,\beta \text{ with respect to their arguments are bounded.} \end{cases}$$

The main result in this section, is the following

Theorem 3.3.1 *(Necessary and sufficient conditions for optimality). Let $\overline{u}_{\cdot}$ be an admissible control (candidate to be optimal) with associated trajectories $(y^{\overline{u}_{\cdot}}_{\cdot}, Y^{\overline{u}_{\cdot}}_{\cdot}, Z^{\overline{u}_{\cdot}}_{\cdot})$. Then $\overline{u}_{\cdot}$ is an optimal control for the strict control problem $\{(3.1),(3.2),(3.3)\}$, if and only if, there exists a unique solution $(\Phi^{\overline{u}_{\cdot}}_{\cdot}, \Psi^{\overline{u}_{\cdot}}_{\cdot}, \Sigma^{\overline{u}_{\cdot}}_{\cdot}, \Pi^{\overline{u}_{\cdot}}_{\cdot})$ of the following adjoint equations of the MF-FBDSDE (3.1),*

$$\begin{cases} -d\Phi^{\overline{u}_{\cdot}}_t = \big(\big(H_y\left(t,\zeta^{\overline{u}_{\cdot}}_t,\overline{u}_t,\chi^{\overline{u}_{\cdot}}_t\right) + \mathbb{E}[\big(H_{y'}\left(t,\zeta^{\overline{u}_{\cdot}}_t,\overline{u}_t,\chi^{\overline{u}_{\cdot}}_t\right)]\big)dt - \Sigma^{\overline{u}}_t dW_t, \\ \\ d\Psi^{\overline{u}_{\cdot}}_t = \big(\big(H_Y\left(t,\zeta^{\overline{u}_{\cdot}}_t,\overline{u}_t,\chi^{\overline{u}_{\cdot}}_t\right) + \mathbb{E}[\big(H_{Y'}\left(t,\zeta^{\overline{u}_{\cdot}}_t,\overline{u}_t,\chi^{\overline{u}_{\cdot}}_t\right)]\big)dt \\ \qquad\qquad + \big(\big(H_Z\left(t,\zeta^{\overline{u}_{\cdot}}_t,\overline{u}_t,\chi^{\overline{u}_{\cdot}}_t\right) + \mathbb{E}[\big(H_{Z'}\left(t,\zeta^{\overline{u}_{\cdot}}_t,\overline{u}_t,\chi^{\overline{u}_{\cdot}}_t\right)]\big)dW_t - \Pi^{\overline{u}}_t \overleftarrow{dB}_t, \\ \\ \Phi^{\overline{u}_{\cdot}}_T = \alpha_y\left(y^{\overline{u}_{\cdot}}_T,\mathbb{E}[y^{\overline{u}_{\cdot}}_T]\right) + \mathbb{E}[\alpha_{y'}\left(y^{\overline{u}}_T,\mathbb{E}[y^{\overline{u}_{\cdot}}_T]\right)], \\ \Psi^{\overline{u}_{\cdot}}_0 = \beta_Y\left(Y^{\overline{u}_{\cdot}}_0,\mathbb{E}[Y^{\overline{u}_{\cdot}}_0]\right) + \mathbb{E}[\beta_{Y'}\left(Y^{\overline{u}_{\cdot}}_0,\mathbb{E}[Y^{\overline{u}_{\cdot}}_0]\right)], \end{cases} \tag{3.13}$$

such that

$$\langle\big(H_v(t,\zeta^{\overline{u}_{\cdot}}_t,\overline{u}_t,\chi^{\overline{u}_{\cdot}}_t),v_t-\overline{u}_t\rangle \geq 0,\ \forall v_{\cdot}\in\mathcal{U},\ a.e,\ as, \tag{3.14}$$

where $\big(H_\varpi(t,\zeta^{\overline{u}_{\cdot}}_t,\overline{u}_t,\chi^{\overline{u}_{\cdot}}_t)$ *with* $\varpi := y,y',Y,Y',Z,Z'$, *is the gradient*

$$\nabla_\varpi\big(H(t,y^{\overline{u}_{\cdot}}_t,\mathbb{E}[y^{\overline{u}_{\cdot}}_t],Y^{\overline{u}_{\cdot}}_t,\mathbb{E}[Y^{\overline{u}_{\cdot}}_t],Z^{\overline{u}_{\cdot}}_t,\mathbb{E}[Z^{\overline{u}_{\cdot}}_t],\overline{u}_t,\Phi^{\overline{u}_{\cdot}}_t,\Psi^{\overline{u}_{\cdot}}_t,\Sigma^{\overline{u}_{\cdot}}_t,\Pi^{\overline{u}_{\cdot}}_t),$$

$$(t,\zeta_t^{\overline{u}.},\overline{u}_t,\chi_t^{\overline{u}.}) := (t,y_t^{\overline{u}.},\mathbb{E}[y_t^{\overline{u}.}],Y_t^{\overline{u}.},\mathbb{E}[Y_t^{\overline{u}.}],Z_t^{\overline{u}.},\mathbb{E}[Z_t^{\overline{u}.}],\overline{u}_t,\Phi_t^{\overline{u}.},\Psi_t^{\overline{u}.},\Sigma_t^{\overline{u}.},\Pi_t^{\overline{u}.}),$$

and the Hamiltonian function is given by

$$\begin{aligned}\big(H\left(t,y,y',Y,Y',Z,Z',v,\Phi,\Psi,\Sigma,\Pi\right) &= \left\langle \Psi, dy+\widehat{d}y'+eY+\widehat{e}Y'+fZ+\widehat{f}Z'+gv\right\rangle \\ &+\left\langle \Phi, ay+\widehat{a}y'+bv\right\rangle+\left\langle \Pi, hy+\widehat{h}y'+kY+\widehat{k}Y'+mZ+\widehat{m}Z'+\widehat{g}v\right\rangle \\ &+\left\langle \Sigma, cy+\widehat{c}y'+\widehat{b}v\right\rangle+\ell\left(t,y,y',Y,Y',Z,Z',v\right).\end{aligned}$$

Proof. Our control problem is governed by a linear system, so to establish a necessary and sufficient optimality conditions, we use the following principle: *The convex optimization principle* (see Ekeland-Temam ([14], prop 2.1, p 35). Since the domain of control U is convex, the functional $\mathbb{J}$ is convex in $\overline{u}.$, continuous and Gâteaux-differentiable with continuous derivative $\mathbb{J}'$, thus, we have

$$(\overline{u}.\text{ minimize } \mathbb{J}) \Leftrightarrow \langle \mathbb{J}'(\overline{u}.), v.-\overline{u}.\rangle \geq 0; \forall v. \in \mathcal{U}. \tag{3.15}$$

Firstly, let us calculate the Gâteaux derivative of $\mathbb{J}$ at a point $\overline{u}.$ and in the direction $(v.-\overline{u}.)$, we obtain

$$\begin{aligned}\langle \mathbb{J}'(\overline{u}.), v.-\overline{u}.\rangle &= \mathbb{E}\big[\langle \alpha_y(y_T^{\overline{u}.},\mathbb{E}[y_T^{\overline{u}.}])+\mathbb{E}[\alpha_{y'}(y_T^{\overline{u}.},\mathbb{E}[y_T^{\overline{u}.}])], y_T^{v.}-y_T^{\overline{u}.}\rangle\big] \\ &+\mathbb{E}\big[\langle \beta_Y(Y_0^{\overline{u}.},\mathbb{E}[Y_0^{\overline{u}.}])+\mathbb{E}[\beta_{Y'}(Y_0^{\overline{u}.},\mathbb{E}[Y_0^{\overline{u}.}])], Y_0^{v.}-Y_0^{\overline{u}.}\rangle\big] \\ &+\mathbb{E}\big[\int_0^T \langle \ell_y(t,\zeta_t^{\overline{u}.},\overline{u}_t)+\mathbb{E}[\ell_{y'}(t,\zeta_t^{\overline{u}.},\overline{u}_t)], y_t^{v.}-y_t^{\overline{u}.}\rangle dt\big] \\ &+\mathbb{E}\big[\int_0^T \langle \ell_Y(t,\zeta_t^{\overline{u}.},\overline{u}_t)+\mathbb{E}[\ell_{Y'}(t,\zeta_t^{\overline{u}.},\overline{u}_t)], Y_t^{v.}-Y_t^{\overline{u}.}\rangle dt\big] \\ &+\mathbb{E}\big[\int_0^T \langle \ell_Z(t,\zeta_t^{\overline{u}.},\overline{u}_t)+\mathbb{E}[\ell_{Z'}(t,\zeta_t^{\overline{u}.},\overline{u}_t)], Z_t^{v.}-Z_t^{\overline{u}.}\rangle dt\big] \\ &+\mathbb{E}\big[\int_0^T \langle \ell_v(t,\zeta_t^{\overline{u}.},\overline{u}_t), v_t-\overline{u}_t\rangle dt\big].\end{aligned} \tag{3.16}$$

The adjoint equations (3.13) can be rewritten as follows

$$
\begin{cases}
-d\Phi_t^{\overline{u}} = \big(\Psi_t^{\overline{u}} d_t + \Phi_t^{\overline{u}} a_t + \Pi_t^{\overline{u}} h_t + \Sigma_t^{\overline{u}} c_t + \ell_y(t, \zeta_t^{\overline{u}}, \overline{u}_t) \\
\qquad + \mathbb{E}[\Psi_t^{\overline{u}} \widehat{d}_t + \Phi_t^{\overline{u}} \widehat{a}_t + \Pi_t^{\overline{u}} \widehat{h}_t + \Sigma_t^{\overline{u}} \widehat{c}_t + \ell_{y'}(t, \zeta_t^{\overline{u}}, \overline{u}_t)]\big) dt - \Sigma_t^{\overline{u}} dW_t, \\
\\
d\Psi_t^{\overline{u}} = \big(\Psi_t^{\overline{u}} e_t + \Pi_t^{\overline{u}} k_t + \ell_Y(t, \zeta_t^{\overline{u}}, \overline{u}_t) + \mathbb{E}[\big(\Psi_t^{\overline{u}} \widehat{e}_t + \Pi_t^{\overline{u}} \widehat{k}_t + \ell_{Y'}(t, \zeta_t^{\overline{u}}, \overline{u}_t)]\big) dt \\
\qquad + \big(\Psi_t^{\overline{u}} f_t + \Pi_t^{\overline{u}} m_t + \ell_Z(t, \zeta_t^{\overline{u}}, \overline{u}_t) + \mathbb{E}[\big(\Psi_t^{\overline{u}} \widehat{f}_t + \Pi_t^{\overline{u}} \widehat{m}_t + \ell_{Z'}(t, \zeta_t^{\overline{u}}, \overline{u}_t)]\big) dW_t, \\
\qquad - \Pi_t^{\overline{u}} \overleftarrow{dB}_t \\
\\
\Phi_T^{\overline{u}} = \alpha_y \left(y_T^{\overline{u}}, \mathbb{E}[y_T^{\overline{u}}]\right) + \mathbb{E}[\alpha_{y'} \left(y_T^{\overline{u}}, \mathbb{E}[y_T^{\overline{u}}]\right)], \\
\Psi_0^{\overline{u}} = \beta_Y \left(Y_0^{\overline{u}}, \mathbb{E}[Y_0^{\overline{u}}]\right) + \mathbb{E}[\beta_{Y'} \left(Y_0^{\overline{u}}, \mathbb{E}[Y_0^{\overline{u}}]\right)].
\end{cases}
$$

From (3.13), the equality (3.16) becomes

$$
\begin{aligned}
\langle \mathbb{J}'(\overline{u}_\cdot), v_\cdot - \overline{u}_\cdot \rangle &= \mathbb{E}\big[\langle \Phi_T^{\overline{u}}, y_T^{v_\cdot} - y_T^{\overline{u}_\cdot} \rangle\big] + \mathbb{E}\big[\langle \Psi_0^{\overline{u}}, Y_0^{v_\cdot} - Y_0^{\overline{u}_\cdot} \rangle\big] \\
&+ \mathbb{E}\Big[\int_0^T \langle \ell_y(t, \zeta_t^{\overline{u}_\cdot}, \overline{u}_t) + \mathbb{E}[\ell_{y'}(t, \zeta_t^{\overline{u}_\cdot}, \overline{u}_t)], y_t^{v_\cdot} - y_t^{\overline{u}_\cdot} \rangle dt\Big] \\
&+ \mathbb{E}\Big[\int_0^T \langle \ell_Y(t, \zeta_t^{\overline{u}_\cdot}, \overline{u}_t) + \mathbb{E}[\ell_{Y'}(t, \zeta_t^{\overline{u}_\cdot}, \overline{u}_t)], Y_t^{v_\cdot} - Y_t^{\overline{u}_\cdot} \rangle dt\Big] \\
&+ \mathbb{E}\Big[\int_0^T \langle \ell_Z(t, \zeta_t^{\overline{u}_\cdot}, \overline{u}_t) + \mathbb{E}[\ell_{Z'}(t, \zeta_t^{\overline{u}_\cdot}, \overline{u}_t)], Z_t^{v_\cdot} - Z_t^{\overline{u}_\cdot} \rangle dt\Big] \\
&+ \mathbb{E}\Big[\int_0^T \langle \ell_v(t, \zeta_t^{\overline{u}_\cdot}, \overline{u}_t), v_t - \overline{u}_t \rangle dt\Big].
\end{aligned} \tag{3.17}
$$

Applying integration by part to $\langle \Psi_t^{\overline{u}}, Y_t^{v_\cdot} - Y_t^{\overline{u}_\cdot} \rangle$ and $\langle \Phi_t^{\overline{u}}, y_t^{v_\cdot} - y_t^{\overline{u}_\cdot} \rangle$, passing to integral on

$[0, T]$ and taking the expectations to deduce

$$\begin{aligned}\mathbb{E}\big[\langle \Phi_T^{\overline{u}.}, y_T^{v.} - y_T^{\overline{u}.}\rangle\big] &= -\mathbb{E}\big[\int_0^T \langle \Psi_t^{\overline{u}.} d_t + \Phi_t^{\overline{u}.} a_t + \Pi_t^{\overline{u}.} h_t + \Sigma_t^{\overline{u}.} c_t + \ell_y(t, \zeta_t^{\overline{u}.}, \overline{u}_t) \\ &+\mathbb{E}[\Psi_t^{\overline{u}.} \widehat{d}_t + \Phi_t^{\overline{u}.} \widehat{a}_t + \Pi_t^{\overline{u}.} \widehat{h}_t + \Sigma_t^{\overline{u}.} \widehat{c}_t + \ell_{y'}(t, \zeta_t^{\overline{u}.}, \overline{u}_t)], y_t^{v.} - y_t^{\overline{u}.}\rangle dt\big] \\ &+\mathbb{E}\big[\int_0^T \langle \Phi_t^{\overline{u}.}, a_t(y_t^{v.} - y_t^{\overline{u}.}) + \widehat{a}_t\mathbb{E}[y_t^{v.} - y_t^{\overline{u}.}] + b_t(v_t - \overline{u}_t)\rangle dt\big] \\ &+\mathbb{E}\big[\int_0^T \langle \Sigma_t^{\overline{u}.}, c_t(y_t^{v.} - y_t^{\overline{u}.}) + \widehat{c}_t\mathbb{E}[y_t^{v.} - y_t^{\overline{u}.}] + \widehat{b}_t(v_t - \overline{u}_t)\rangle dt\big], \end{aligned} \tag{3.18}$$

and

$$\begin{aligned}\mathbb{E}\big[\langle \Psi_0^{\overline{u}.}, Y_0^{v.} - Y_0^{\overline{u}.}\rangle\big] &= -\mathbb{E}\big[\int_0^T \langle \Psi_t^{\overline{u}.} e_t + \Pi_t^{\overline{u}.} k_t + \ell_Y(t, \zeta_t^{\overline{u}.}, \overline{u}_t) \\ &+\mathbb{E}[\Psi_t^{\overline{u}.} \widehat{e}_t + \Pi_t^{\overline{u}.} \widehat{k}_t + \ell_{Y'}(t, \zeta_t^{\overline{u}.}, \overline{u}_t)], Y_t^{v.} - Y_t^{\overline{u}.}\rangle dt\big] \\ &+\mathbb{E}\big[\int_0^T \langle \Psi_t^{\overline{u}.}, d_t(y_t^{v.} - y_t^{\overline{u}.}) + \widehat{d}_t\mathbb{E}[y_t^{v.} - y_t^{\overline{u}.}] + e_t(Y_t^{v.} - Y_t^{\overline{u}.}) \\ &+\widehat{e}_t\mathbb{E}[Y_t^{v.} - Y_t^{\overline{u}.}] + f_t(Z_t^{v.} - Z_t^{\overline{u}.}) + \widehat{f}_t\mathbb{E}[Z_t^{v.} - Z_t^{\overline{u}.}] + g_t(v_t - \overline{u}_t)\rangle dt\big] \\ &+\mathbb{E}\big[\int_0^T \langle \Pi_t^{\overline{u}.}, h_t(y_t^{v.} - y_t^{\overline{u}.}) + \widehat{h}_t\mathbb{E}[y_t^{v.} - y_t^{\overline{u}.}] + k_t(Y_t^{v.} - Y_t^{\overline{u}.}) \\ &+\widehat{k}_t\mathbb{E}[Y_t^{v.} - Y_t^{\overline{u}.}] + m_t(Z_t^{v.} - Z_t^{\overline{u}.}) + \widehat{m}_t\mathbb{E}[Z_t^{v.} - Z_t^{\overline{u}.}] + \widehat{g}_t(v_t - \overline{u}_t)\rangle dt\big] \\ &-\mathbb{E}\big[\int_0^T \langle \Psi_t^{\overline{u}.} f_t + \Pi_t^{\overline{u}.} m_t + \ell_Z(t, \zeta_t^{\overline{u}.}, \overline{u}_t) \\ &+\mathbb{E}[\Psi_t^{\overline{u}.} \widehat{f}_t + \Pi_t^{\overline{u}.} \widehat{m}_t + \ell_{Z'}(t, \zeta_t^{\overline{u}.}, \overline{u}_t)], Z_t^{v.} - Z_t^{\overline{u}.}\rangle dt\big]. \end{aligned} \tag{3.19}$$

Combining (3.17), (3.18) and (3.19), we obtain

$$\langle \mathbb{J}'(\overline{u}), v. - \overline{u}.\rangle = \mathbb{E}\big[\int_0^T \langle \Phi_t^{\overline{u}.} b_t + \Sigma_t^{\overline{u}.} \widehat{b}_t + \Psi_t^{\overline{u}.} g_t + \Pi_t^{\overline{u}.} \widehat{g}_t + \ell_v(t, \zeta_t^{\overline{u}.}, \overline{u}_t), v_t - \overline{u}_t)\rangle dt\big].$$

On the other hand, we calculate the Gâteaux derivative of $\big(H$ at a point $\overline{u}.$ in the direction

$(v. - \overline{u}.)$, we have

$$\begin{aligned}\mathbb{E}\Big[\int_0^T \langle \big(H_v(t,\zeta_t^{\overline{u}.},\overline{u}_t,\chi_t^{\overline{u}.}),v_t-\overline{u}_t\rangle dt\Big] &= \mathbb{E}\Big[\int_0^T \langle \Phi_t^{\overline{u}.} b_t + \Sigma_t^{\overline{u}.}\widehat{b}_t + \Psi_t^{\overline{u}.} g_t + \Pi_t^{\overline{u}.}\widehat{g}_t \\ &\quad + \ell_v(t,\zeta_t^{\overline{u}.},\overline{u}_t), v_t - \overline{u}_t)\rangle dt\Big] \\ &= \langle \mathbb{J}'(\overline{u}), v. - \overline{u}.\rangle. \end{aligned} \tag{3.20}$$

Combines (3.15) and (3.20), we get

$$(\overline{u} \text{ minimize } \mathbb{J}) \Leftrightarrow \mathbb{E}\Big[\int_0^T \langle \big(H_v(t,\zeta_t^{\overline{u}.},\overline{u}_t,\chi_t^{\overline{u}.}),v_t-\overline{u}_t\rangle dt\Big] \geq 0, \forall v. \in \mathcal{U}.$$

By a standard argument we get the result. ■

Chapter 4

Necessary and sufficient optimality conditions for both relaxed and strict control problems for nonlinear MF-FBDSDEs

Stochastic control problems have gained a particular interest due to their broad applications in economics, finance, engineering, etc. Mean-field models are useful to characterize the asymptotic behavior when the size of the system is getting very large, in 2009, Buckdahn et al. [11] established the theory of mean-field backward stochastic differential equations which were derived as a limit of some highly dimensional system of FBSDEs, corresponding to a large number of particles. Since that, many authors treated the system of this kind of Mckean-Vlasov type (see [25] and [1]).

In the other hand, the existence of optimal relaxed controls and optimal strict controls for systems of mean-field forward backward stochastic differential equations has been proved by Benbrahim and Gherbal [8], where the diffusion is controlled. The existence of relaxed solutions to mean field games with singular controls has been proved by Fu and Horst in

[18].

In this chapter, we establish necessary as well as sufficient optimality conditions for both relaxed and strict control problems driven by systems of nonlinear MF-FBDSDEs, where the action space U is not necessary convex.

4.1 Statement of the problems

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Let $\{W_t, t \in [0,T]\}$ and $\{B_t, t \in [0,T]\}$ be two mutually independent standard Brownian motions, with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^l$.

Let $\mathcal{N}$ denote the class of $\mathbb{P}$-null sets of $\mathcal{F}$. For each $0 \leq t \leq T$, we define $\mathcal{F}_t := \mathcal{F}_t^W \vee \mathcal{F}_{t,T}^B$, where for any process $\{\delta_t\}$, we set $\mathcal{F}_{s,t}^\delta = \sigma\left(\delta_r - \delta_s; s \leq r \leq t\right) \vee \mathcal{N}, \mathcal{F}_t^\delta = \mathcal{F}_{0,t}^\delta$.

Note that the collection $\{\mathcal{F}_t, t \in [0,T]\}$ is neither increasing nor decreasing, then it does not constitute a classical filtration.

4.1.1 Strict control problem

Definition 4.1.1 *An admissible control u. is a square integrable, $\mathcal{F}_t$-measurable process with values in some subset $U \subseteq \mathbb{R}^k$. We denote by $\mathcal{U}$ the set of all admissible controls.*

For any $v \in \mathcal{U}$,we consider the following MF-FBDSDE

$$\begin{cases} y_t^v = x + \int_0^t b(s, y_s^v, \mathbb{E}[y_s^v], v_s)ds + \int_0^t \sigma(s, y_s^v, \mathbb{E}[y_s^v])dW_s \\ \\ Y_t^v = h(y_T^v, \mathbb{E}[y_T^v]) + \int_t^T f(s, y_s^v, \mathbb{E}[y_s^v], Y_s^v, \mathbb{E}[Y_s^v], Z_s^v, \mathbb{E}[Z_s^v], v_s)ds \\ \qquad + \int_t^T g(s, y_s^v, \mathbb{E}[y_s^v], Y_s^v, \mathbb{E}[Y_s^v], Z_s^v, \mathbb{E}[Z_s^v])\overleftarrow{dB}_t - \int_t^T Z_s^v dW_s, \end{cases} \tag{4.1}$$

and the functional cost to be minimize over the set of strict controls $\mathcal{U}$ is given by

$$\begin{aligned}\mathbb{J}(v.) := \mathbb{E}\big[\alpha\left(y_T^v, \mathbb{E}\left[y_T^v\right]\right) + \beta\left(Y_0^v, \mathbb{E}\left[Y_0^v\right]\right) \\ + \textstyle\int_0^T \ell\big(t, y_t^v, \mathbb{E}\left[y_t^v\right], Y_t^v, \mathbb{E}\left[Y_t^v\right], Z_t^v, \mathbb{E}\left[Z_t^v\right], v_t\big) dt\big].\end{aligned} \tag{4.2}$$

where

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

$$\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$

$$\beta : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}.$$

We say that a strict control $u.$ is an optimal control if

$$\mathbb{J}(u.) = \inf_{v.\in\mathcal{U}} \mathbb{J}(v.). \tag{4.3}$$

4.1.2 Relaxed control problem

Without the convexity condition an optimal control does not necessarily exist in U, we need to use a bigger new class its role is to compensate strict control set. The idea is then to replace the $U-$valued process u_t with $\mathbb{P}(U)-$valued process (q_t), where $\mathbb{P}(U)$ is the space of probability measures on U equipped with the topology of stable convergence. These measure valued control are called relaxed control. It turns out that this class of controls enjoys good topological properties. 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 case. Thus the set of strict controls may be identified as a subset of relaxed controls.

Let $\mathbb{V}$ the set of Radon measures q on the set $[0,T]\times U$, whose projections on $[0,T]$ coincide with the Lebesgue measure dt, and whose projection on U coincide with some probability measure $q_t(da) \in \mathbb{P}(U)$ i.e. $q(da, dt) = q_t(da)\, dt$. $\mathbb{V}$ is a compact metrizable space, see [23]. The topology of stable convergence of measures is the coarsest topology which makes the mapping

$$q \mapsto \int_0^T \int_U \psi(t, a)\, q_t(dt, da)$$

continuous, for all bounded measurable functions $\psi(t, a)$ such that for fixed t, $\psi(t, a)$ is continuous.

The definition of admissible relaxed control is given by

Definition 4.1.2 *A stochastic relaxed control (or simply a relaxed control) $\mu_.$ is a measurable $\mathbb{P}(U)$ −valued process,i.e.,*

$$\begin{aligned}[0, T] \times \Omega &\to \mathbb{P}(U) \\ (t, \omega) &\mapsto \mu_t(\omega, .),\end{aligned}$$

is measurable. We say that a relaxed control $\mu_.$ is admissible if μ_t is $\mathcal{F}_t$−progressively measurable, in the sense that, for any bounded measurable function $\Phi : [0, T] \times U \to \mathbb{P}(U)$, the process $\int_0^t \int_U \Phi(s, a)\, \mu_s(da)\, ds$ is $\mathcal{F}_t$−measurable for all $t \in [0, T]$, and if, moreover,

$$E\left[\sup_{t\in[0,T]} \int_U |a|^2 \mu_t(da)\right] < \infty.$$

Let us denote by $\mathcal{R}$ to the set of all such admissible relaxed controls.

For any $\mu \in \mathcal{R}$ we consider a relaxed control problem governed by the following MF-

FBDSDE:

$$\begin{cases} dy_t^\mu = \int_U b(t, y_t^\mu, \mathbb{E}[y_t^\mu], u)\mu_t(du)dt + \sigma(t, y_t^\mu, \mathbb{E}[y_t^\mu])dW_t \\ dY_t^\mu = -\int_U f(t, y_t^\mu, \mathbb{E}[y_t^\mu], Y_t^\mu, \mathbb{E}[Y_t^\mu], Z_t^\mu, \mathbb{E}[Z_t^\mu], u)\mu_t(du)dt \\ \qquad\qquad -g(t, y_t^\mu, \mathbb{E}[y_t^\mu], Y_t^\mu, \mathbb{E}[Y_t^\mu], Z_t^\mu, \mathbb{E}[Z_t^\mu])\overleftarrow{dB}_t + Z_t^\mu dW_t \\ y_0^\mu = x, Y_T^\mu = h(y_T^\mu, \mathbb{E}[y_T^\mu]), \quad t \in [0, T], \end{cases} \tag{4.4}$$

and the functional cost is given by

$$\begin{aligned} \mathbb{J}(\mu.) := \mathbb{E}\big[&\alpha\left(y_T^\mu, \mathbb{E}\left[y_T^\mu\right]\right) + \beta\left(Y_0^\mu, \mathbb{E}\left[Y_0^\mu\right]\right) \\ &+ \int_0^T \int_U \ell\big(t, y_t^\mu, \mathbb{E}\left[y_t^\mu\right], Y_t^\mu, \mathbb{E}\left[Y_t^\mu\right], Z_t^\mu, \mathbb{E}\left[Z_t^\mu\right], u\big)\mu_t(du)dt\big]. \end{aligned} \tag{4.5}$$

We say that a relaxed control $q.$ is an optimal control if

$$\mathbb{J}(q.) = \inf_{\mu. \in \mathcal{R}} \mathbb{J}(\mu.). \tag{4.6}$$

4.2 Necessary and sufficient optimality conditions for relaxed control problems

In this section, we study the problem {(4.4),(4.5),(4.6)} and we establish necessary condition of optimality for relaxed controls.

According to the fact that the set of relaxed controls is convex, then to establish necessary optimality condition we use the convex perturbation method. Let $q.$ be an optimal relaxed control with associated trajectories (y_t^q, Y_t^q, Z_t^q) solution of the MF-FBDSDEs (4.4). Then, we can define a perturbed relaxed control by

$$q_t^\varepsilon = q_t + \varepsilon(\mu_t - q_t),$$

where $\varepsilon > 0$ is sufficiently small and $\mu_.$ is an arbitrary element of $\mathcal{R}$. Denote by $(y_t^\varepsilon, Y_t^\varepsilon, Z_t^\varepsilon)$ the solution of the system (4.4) corresponding to $q_.^\varepsilon$.

We shall consider in this section the following assumptions.

- (H4) (Lipschitz condition)

  $\exists\, C > 0, 0 < \gamma < \frac{1}{2}$ such that $\forall\, y_1, y_1', y_2, y_2', Y_1, Y_1', Y_2, Y_2', Z_1, Z_1', Z_2, Z_2', u,$

$$|b(t, y_1, y_1', u) - b(t, y_2, y_2', u)|^2 \leq C\ (|y_1 - y_2|^2 + |y_1' - y_2'|^2),$$

$$|\sigma(t, y_1, y_1') - \sigma(t, y_2, y_2')|^2 \leq C\ (|y_1 - y_2|^2 + |y_1' - y_2'|^2),$$

$$\begin{aligned} & |f(t, y_1, y_1', Y_1, Y_1', Z_1, Z_1', u) - f(t, y_2, y_2', Y_2, Y_2', Z_2, Z_2', u)|^2 \\ \leq\ & C\ (|y_1 - y_2|^2 + |y_1' - y_2'|^2 + |Y_1 - Y_2|^2 + |Y_1' - Y_2'|^2 \\ & \qquad + \|Z_1 - Z_2\|^2 + \|Z_1' - Z_2'\|^2), \end{aligned}$$

$$\begin{aligned} & |\ell(t, y_1, y_1', Y_1, Y_1', Z_1, Z_1', u) - \ell(t, y_2, y_2', Y_2, Y_2', Z_2, Z_2', u)|^2 \\ \leq\ & C\ (|y_1 - y_2|^2 + |y_1' - y_2'|^2 + |Y_1 - Y_2|^2 + |Y_1' - Y_2'|^2 \\ & \qquad + \|Z_1 - Z_2\|^2 + \|Z_1' - Z_2'\|^2), \end{aligned}$$

$$\begin{aligned} & |g(t, y_1, y_1', Y_1, Y_1', Z_1, Z_1') - g(t, y_2, y_2', Y_2, Y_2', Z_2, Z_2')|^2 \\ \leq\ & C\ (|y_1 - y_2|^2 + |y_1' - y_2'|^2 + |Y_1 - Y_2|^2 + |Y_1' - Y_2'|^2) \\ & \qquad + \gamma(\|Z_1 - Z_2\|^2 + \|Z_1' - Z_2'\|^2). \end{aligned}$$

- (H5) (Regularity conditions)

$$\begin{cases} (i) \text{ the mappings } b, h, \sigma, \alpha \text{ are bounded and continuously differentiable with} \\ \quad \text{respect to } (y, y'), \text{ and the functions } f, g \text{ and } \beta \text{ are bounded and continuously} \\ \quad \text{differentiable with respect to } (y, y', Y, Y', Z, Z') \text{ and } (y, y'), \text{ respectively,} \\ (ii) \text{ the derivatives of } b, h, g, \sigma, f \text{ with respect to the above arguments are} \\ \quad \text{continuous and bounded,} \\ (iii) \text{ the derivatives of } \ell \text{ are bounded by } C(1 + |y| + |y'| + |Y| + |Y'| + |Z| + |Z'|), \\ (iv) \text{ the derivatives of } \alpha \text{ and } \beta \text{ are bounded by } C\left(1 + |y| + |y'|\right) \text{ and} \\ \quad C\left(1 + |Y| + |Y'|\right) \text{ respectively,} \end{cases}$$

for some positive constant C.

Remark 4.2.1 *Under the above hypothesis, for every $\mu_\cdot \in \mathcal{R}$, equation (4.4) has a unique strong solution and the functional cost $\mathbb{J}$ is well defined from $\mathcal{R}$ into $\mathbb{R}$.*

4.2.1 The variational inequality

Using the optimality of $q_\cdot$, the variational inequality will be derived from the following inequality

$$0 \leq \mathbb{J}(q^\varepsilon) - \mathbb{J}(q).$$

For this end, we need some results.

Proposition 4.2.1 *Under assumptions* $(\mathbf{H4}) - (\mathbf{H5})$, *we have*

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\sup_{0 \leq t \leq T} |y_t^\varepsilon - y_t^q|^2\right] = 0, \tag{4.7}$$

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\sup_{0 \leq t \leq T} |Y_t^\varepsilon - Y_t^q|^2\right] = 0, \tag{4.8}$$

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\int_0^T \|Z_t^\varepsilon - Z_t^q\|^2 dt\right] = 0. \tag{4.9}$$

Proof. We calculate $\mathbb{E}\left[|y_t^\varepsilon - y_t^q|^2\right]$ and using the definition of q_t^ε to get

$$\begin{aligned}\mathbb{E}\left[|y_t^\varepsilon - y_t^q|^2\right] \leq C\mathbb{E}\Big[\int_0^t \Big| &\int_U b\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], u\right) q_s(du) \\ &- \int_U b\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], u\right) q_s(du)\Big|^2 ds\Big] \\ +C\varepsilon^2\mathbb{E}\Big[\int_0^t \Big|\int_U b\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], u\right) \mu_s(du) &- \int_U b\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], u\right) q_s(du)\Big|^2 ds\Big] \\ +C\mathbb{E}\Big[\int_0^t \left|\sigma\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right]\right) - \sigma\left(s, y_s^q, \mathbb{E}\left[y_s^q\right]\right)\right|^2 ds\Big].\end{aligned}$$

Since b and σ are uniformly Lipschitz and b is bounded, we can show

$$\mathbb{E}\left[|y_t^\varepsilon - y_t^q|^2\right] \leq C\mathbb{E}\Big[\int_0^t |y_s^\varepsilon - y_s^q|^2 ds\Big] + C\varepsilon^2.$$

Applying Granwall's lemma and Burkholder-Davis-Gundy inequality, we get (4.7).

On the other hand, applying Itô's formula to $(Y_t^\varepsilon - Y_t^q)^2$, taking expectation and applying Young's inequality, to obtain

$$\begin{aligned}\mathbb{E}\left[|Y_t^\varepsilon - Y_t^q|^2\right] + \mathbb{E}\left[\int_t^T \|Z_s^\varepsilon - Z_s^q\|^2 ds\right] &\leq \mathbb{E}\left[|h(y_T^\varepsilon, \mathbb{E}[y_T^\varepsilon]) - h(y_T^q, \mathbb{E}[y_T^q])|^2\right] \\ +\frac{1}{\theta}\mathbb{E}\left[\int_t^T |Y_s^\varepsilon - Y_s^q|^2 ds\right] + \theta\mathbb{E}\Big[\int_t^T \Big|&\int_U f\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], Y_s^\varepsilon, \mathbb{E}\left[Y_s^\varepsilon\right], Z_s^\varepsilon, \mathbb{E}\left[Z_s^\varepsilon\right], u\right) q_s^\varepsilon(du) \\ &- \int_U f\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], Y_s^q, \mathbb{E}\left[Y_s^q\right], Z_s^q, \mathbb{E}\left[Z_s^q\right], u\right) q_s(du)\Big|^2 ds\Big] \\ +\mathbb{E}\Big[\int_t^T \big|g\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], Y_s^\varepsilon, \mathbb{E}\left[Y_s^\varepsilon\right], Z_s^\varepsilon, \mathbb{E}\left[Z_s^\varepsilon\right]\right)& \\ &-g\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], Y_s^q, \mathbb{E}\left[Y_s^q\right], Z_s^q, \mathbb{E}\left[Z_s^q\right]\right)\big|^2 ds\Big].\end{aligned}$$

Using the definition of q_t^ε, we obtain

$$\begin{aligned}
&\mathbb{E}\left[|Y_t^\varepsilon - Y_t^q|^2\right] + \mathbb{E}\left[\int_t^T \|Z_s^\varepsilon - Z_s^q\|^2 ds\right] \leq \mathbb{E}\left[|h(y_T^\varepsilon, \mathbb{E}[y_T^\varepsilon]) - h(y_T^q, \mathbb{E}[y_T^q])|^2\right] \\
&\quad + \frac{1}{\theta}\mathbb{E}\left[\int_t^T |Y_s^\varepsilon - Y_s^q|^2 ds\right] \\
&\quad + C\theta\varepsilon^2 \mathbb{E}\Big[\int_t^T \Big| \int_U f\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], Y_s^\varepsilon, \mathbb{E}\left[Y_s^\varepsilon\right], Z_s^\varepsilon, \mathbb{E}\left[Z_s^\varepsilon\right], u\right) \mu_s(du) \\
&\qquad\qquad - \int_U f\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], Y_s^\varepsilon, \mathbb{E}\left[Y_s^\varepsilon\right], Z_s^\varepsilon, \mathbb{E}\left[Z_s^\varepsilon\right], u\right) q_s(du) \Big|^2 ds\Big] \\
&\quad + C\theta \mathbb{E}\Big[\int_t^T \Big| \int_U f\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], Y_s^\varepsilon, \mathbb{E}\left[Y_s^\varepsilon\right], Z_s^\varepsilon, \mathbb{E}\left[Z_s^\varepsilon\right], u\right) q_s(du) \\
&\qquad\qquad - \int_U f\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], Y_s^q, \mathbb{E}\left[Y_s^q\right], Z_s^q, \mathbb{E}\left[Z_s^q\right], u\right) q_s(du) \Big|^2 ds\Big] \\
&\quad + \mathbb{E}\Big[\int_t^T |g\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], Y_s^\varepsilon, \mathbb{E}\left[Y_s^\varepsilon\right], Z_s^\varepsilon, \mathbb{E}\left[Z_s^\varepsilon\right]\right) \\
&\qquad\qquad - g\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], Y_s^q, \mathbb{E}\left[Y_s^q\right], Z_s^q, \mathbb{E}\left[Z_s^q\right]\right)|^2 ds\Big].
\end{aligned}$$

Since f and h are uniformly Lipschitz with respect to their arguments, we have

$$\begin{aligned}
\mathbb{E}\left[|Y_t^\varepsilon - Y_t^q|^2\right] + \mathbb{E}\left[\int_t^T \|Z_s^\varepsilon - Z_s^q\|^2 ds\right] &\leq (\frac{1}{\theta} + 2C\theta + 2C)\mathbb{E}\left[\int_t^T |Y_s^\varepsilon - Y_s^q|^2 ds\right] \\
&\quad + (2C\theta + 2\gamma)\mathbb{E}\left[\int_t^T \|Z_s^\varepsilon - Z_s^q\|^2 ds\right] + \phi_t^\varepsilon,
\end{aligned} \tag{4.10}$$

where

$$\phi_t^\varepsilon = 2C\mathbb{E}\left[|y_T^\varepsilon - y_T^q|^2\right] + (2C\theta + 2C)\mathbb{E}\Big[\int_t^T |y_s^\varepsilon - y_s^q|^2 ds\Big] + C\varepsilon\theta^2.$$

From (4.7) we can show that

$$\lim_{\varepsilon \to 0} \phi_t^\varepsilon = 0. \tag{4.11}$$

Choose $\theta = \frac{1-2\gamma}{4C} > 0$, thus $2C\theta + 2\gamma = \frac{1-2\gamma}{2} + 2\gamma = \frac{1+2\gamma}{2} < 1$, so the inequality (4.10) becomes

$$\mathbb{E}\left[|Y_t^\varepsilon - Y_t^q|^2\right] + \frac{1-2\gamma}{2}\mathbb{E}\left[\int_t^T \|Z_s^\varepsilon - Z_s^q\|^2 ds\right] \le C\mathbb{E}\left[\int_t^T |Y_s^\varepsilon - Y_s^q|^2 ds\right] + \phi_t^\varepsilon,$$

we derive from this inequality, two inequalities

$$\mathbb{E}\left[|Y_t^\varepsilon - Y_t^q|^2\right] \le C\mathbb{E}\left[\int_t^T |Y_s^\varepsilon - Y_s^q|^2 ds\right] + \phi_t^\varepsilon, \tag{4.12}$$

and

$$\mathbb{E}\left[\int_t^T \|Z_s^\varepsilon - Z_s^q\|^2 ds\right] \le C\mathbb{E}\left[\int_t^T |Y_s^\varepsilon - Y_s^q|^2 ds\right] + \phi_t^\varepsilon. \tag{4.13}$$

Applying Granwall's lemma and Burkholder-Davis-Gundy inequality in (4.12) and using (4.7) and (4.11) to get (4.8). Finally (4.9) derived from (4.8), (4.11) and (4.13). ■

Proposition 4.2.2 *Let* $\left(\widehat{y}_t, \widehat{Y}_t, \widehat{Z}_t\right)$ *, be the solution of the following variational equations* *of MF-FBDSDE (4.4)*

$$\begin{cases} d\widehat{y}_t = \int_U b_y\left(t, y_t^q, \mathbb{E}\left[y_t^q\right], u\right) q_t(du)\widehat{y}_t dt \\ \qquad +\mathbb{E}\left[\int_U b_{y'}\left(t, y_t^q, \mathbb{E}\left[y_t^q\right], u\right) q_t(du)\mathbb{E}[\widehat{y}_t]\right] dt \\ \quad + \left(\sigma_y\left(t, y_t^q, \mathbb{E}\left[y_t^q\right]\right)\widehat{y}_t + \mathbb{E}\left[\sigma_{y'}\left(t, y_t^q, \mathbb{E}\left[y_t^q\right]\right)\mathbb{E}[\widehat{y}_t]\right]\right) dW_t \\ \quad + \left(\int_U b\left(t, y_t^q, \mathbb{E}\left[y_t^q\right], u\right) q_t(du) - \int_U b\left(t, y_t^q, \mathbb{E}\left[y_t^q\right], u\right)\mu_t(du)\right) dt \\ \\ d\widehat{Y}_t = -(\int_U f_y(t, \pi_t^q, u)q_t(du)\widehat{y}_t + \mathbb{E}\left[\int_U f_{y'}(t, \pi_t^q, u)q_t(du)\mathbb{E}[\widehat{y}_t]\right] \\ \qquad + \int_U f_Y(t, \pi_t^q, u)q_t(du)\widehat{Y}_t + \mathbb{E}\left[\int_U f_{Y'}(t, \pi_t^q, u)q_t(du)\mathbb{E}[\widehat{Y}_t]\right] \\ \qquad + \int_U f_Z(t, \pi_t^q, u)q_t(du)\widehat{Z}_t + \mathbb{E}\left[\int_U f_{Z'}(t, \pi_t^q, u)q_t(du)\mathbb{E}[\widehat{Z}_t]\right] \\ \qquad +(\int_U f\left(t, \pi_t^q, u\right) q_t(du) - \int_U f\left(t, \pi_t^q, u\right)\mu_t(du))dt \\ \quad -(g_y(t, \pi_t^q)\widehat{y}_t + \mathbb{E}\left[g_{y'}(t, \pi_t^q)\mathbb{E}[\widehat{y}_t]\right] + g_Y(t, \pi_t^q)\widehat{Y}_t + \mathbb{E}\left[g_{Y'}(t, \pi_t^q)\mathbb{E}[\widehat{Y}_t]\right] \\ \qquad + g_Z(t, \pi_t^q)\widehat{Z}_t + \mathbb{E}\left[g_{Z'}(t, \pi_t^q)\mathbb{E}[\widehat{Z}_t]\right])\overleftarrow{dB}_t + \widehat{Z}_t dW_t, \\ \\ \widehat{y}_0 = 0, \widehat{Y}_T = h_y\left(y_T^q, \mathbb{E}[y_T^q]\right)\widehat{y}_T + \mathbb{E}\left[h_{y'}\left(y_T^q, \mathbb{E}[y_T^q]\right)\mathbb{E}[\widehat{y}_T]\right], \end{cases} \tag{4.14}$$

where $(t, \pi_t^q, u) := (t, y_t^q, \mathbb{E}\left[y_t^q\right], Y_t^q, \mathbb{E}\left[Y_t^q\right], Z_t^q, \mathbb{E}\left[Z_t^q\right], u)$. *We have the following estimates*

$$\lim_{\varepsilon \to 0}\mathbb{E}\left[\sup_{0 \le t \le T}\left|\frac{1}{\varepsilon}(y_t^\varepsilon - y_t^q) - \widehat{y}_t\right|^2\right] = 0, \tag{4.15}$$

$$\lim_{\varepsilon \to 0}\mathbb{E}\left[\sup_{0 \le t \le T}\left|\frac{1}{\varepsilon}(Y_t^\varepsilon - Y_t^q) - \widehat{Y}_t\right|^2\right] = 0, \tag{4.16}$$

$$\lim_{\varepsilon \to 0}\mathbb{E}\left[\int_0^T\left\|\frac{1}{\varepsilon}(Z_t^\varepsilon - Z_t^q) - \widehat{Z}_t\right\|^2 dt\right] = 0. \tag{4.17}$$

Proof. For simplicity, denote by

$$\Upsilon_t^\varepsilon = \frac{1}{\varepsilon}(y_t^\varepsilon - y_t^q) - \widehat{y}_t, \mathbb{Y}_t^\varepsilon = \frac{1}{\varepsilon}(Y_t^\varepsilon - Y_t^q) - \widehat{Y}_t, \mathbb{Z}_t^\varepsilon = \frac{1}{\varepsilon}(Z_t^\varepsilon - Z_t^q) - \widehat{Z}_t. \tag{4.18}$$

i) Let us prove (4.15). From (4.4), (4.14) and notations (4.18), we have

$$\begin{aligned}
\Upsilon_t^\varepsilon &= \frac{1}{\varepsilon}\int_0^t \left[\int_U b\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right], u\right) q_s^\varepsilon(du) - \int_U b\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], u\right) q_s^\varepsilon(du)\right] ds \\
&+\frac{1}{\varepsilon}\int_0^t \left[\int_U b\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], u\right) q_s^\varepsilon(du) - \int_U b\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], u\right) q_s(du)\right] ds \\
&+\frac{1}{\varepsilon}\int_0^t \left[\sigma\left(s, y_s^\varepsilon, \mathbb{E}\left[y_s^\varepsilon\right]\right) - \sigma\left(s, y_s^q, \mathbb{E}\left[y_s^q\right]\right)\right] dW_s \\
&-\int_0^t \int_U b_y\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], u\right) q_s(du)\widehat{y}_s ds \\
&\qquad\qquad -\int_0^t \mathbb{E}\left[\int_U b_{y'}\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], u\right) q_s(du)\mathbb{E}[\widehat{y}_s]\right] ds \\
&-\int_0^t \left(\sigma_y\left(s, y_s^q, \mathbb{E}\left[y_s^q\right]\right)\widehat{y}_s + \mathbb{E}\left[\sigma_{y'}\left(s, y_s^q, \mathbb{E}\left[y_s^q\right]\right)\mathbb{E}[\widehat{y}_s]\right]\right) dW_s \\
&-\int_0^t \left(\int_U b\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], u\right) q_s(du) - \int_U b\left(s, y_s^q, \mathbb{E}\left[y_s^q\right], u\right) \mu_s(du)\right) ds.
\end{aligned}$$

Using the definition of q_s^ε and taking expectation, we obtain

$$\begin{aligned}
\mathbb{E}\left[|\Upsilon_t^\varepsilon|^2\right] \leq C\mathbb{E}&\left[\int_0^t\int_0^1\int_U \left|b_y\left(s, \Lambda_s^\varepsilon, u\right)\Upsilon_t^\varepsilon\right|^2 q_s(du)d\lambda ds\right] \\
&+C\mathbb{E}\left[\int_0^t\int_0^1\int_U \left|\mathbb{E}\left[b_{y'}\left(s, \Lambda_s^\varepsilon, u\right)\mathbb{E}[\Upsilon_t^\varepsilon]\right]\right|^2 q_s(du)d\lambda ds\right] \\
&+C\mathbb{E}\left[\int_0^t\int_0^1 \left|\sigma_y\left(s, \Lambda_s^\varepsilon\right)\Upsilon_t^\varepsilon\right|^2 d\lambda ds\right] \\
&+C\mathbb{E}\left[\int_0^t\int_0^1 \left|\mathbb{E}\left[\sigma_{y'}\left(s, \Lambda_s^\varepsilon\right)\mathbb{E}[\Upsilon_t^\varepsilon]\right]\right|^2 d\lambda ds\right] + C\mathbb{E}\left[|\Gamma_t^\varepsilon|^2\right],
\end{aligned}$$

where $(s, \Lambda_s^\varepsilon, u) := (s, y_s^q + \lambda\varepsilon(\Upsilon_s^\varepsilon + \widehat{y}_s), \mathbb{E}[y_s^q + \lambda\varepsilon(\Upsilon_s^\varepsilon + \widehat{y}_s)], u)$, and

$$\begin{aligned}
\Gamma_t^\varepsilon =& \int_0^t \int_0^1 \int_U b_y\left(s, \Lambda_s^\varepsilon, u\right)\left(y_s^\varepsilon - y_s^q\right) \mu_s(du) d\lambda ds \\
&+ \int_0^t \int_0^1 \int_U \mathbb{E}\left[b_{y'}\left(s, \Lambda_s^\varepsilon, u\right) \mathbb{E}[y_s^\varepsilon - y_s^q]\right] \mu_s(du) d\lambda ds \\
&- \int_0^t \int_0^1 \int_U b_y\left(s, \Lambda_s^\varepsilon, u\right)\left(y_s^\varepsilon - y_s^q\right) q_s(du) d\lambda ds \\
&- \int_0^t \int_0^1 \int_U \mathbb{E}\left[b_{y'}\left(s, \Lambda_s^\varepsilon, u\right) \mathbb{E}[y_s^\varepsilon - y_s^q]\right] q_s(du) d\lambda ds \\
&+ \int_0^t \int_0^1 \int_U \left(b_y\left(s, \Lambda_s^\varepsilon, u\right) \widehat{y}_s + \mathbb{E}\left[b_{y'}\left(s, \Lambda_s^\varepsilon, u\right) \mathbb{E}[\widehat{y}_s]\right]\right) q_s(du) d\lambda ds \\
&\quad + \int_0^t \int_0^1 \left(\sigma_y\left(s, \Lambda_s^\varepsilon\right) \widehat{y}_t + \mathbb{E}\left[\sigma_{y'}\left(s, \Lambda_s^\varepsilon\right) \mathbb{E}[\widehat{y}_s]\right]\right) d\lambda dW_s \\
&- \int_0^t \int_U b_y\left(s, y_s^q, \mathbb{E}[y_s^q], u\right) \widehat{y}_s q_s(du) ds \\
&\quad - \int_0^t \int_U \mathbb{E}\left[b_{y'}\left(s, y_s^q, \mathbb{E}[y_s^q], u\right) \mathbb{E}[\widehat{y}_s]\right] q_s(du) ds \\
&- \int_0^t \left(\sigma_y\left(s, y_s^q, \mathbb{E}[y_s^q]\right) \widehat{y}_s + \mathbb{E}\left[\sigma_{y'}\left(s, y_s^q, \mathbb{E}[y_s^q]\right) \mathbb{E}[\widehat{y}_s]\right]\right) dW_s,
\end{aligned}$$

since $b_y, b_{y'}, \sigma_y, \sigma_{y'}$ are continuous and bounded we have

$$\mathbb{E}\left[|\Upsilon_t^\varepsilon|^2\right] \leq C\mathbb{E}\left[\int_0^t |\Upsilon_s^\varepsilon|^2 ds\right] + C\mathbb{E}\left[|\Gamma_t^\varepsilon|^2\right], \tag{4.19}$$

and

$$\lim_{\varepsilon\to 0}\mathbb{E}\left[|\Gamma_t^\varepsilon|^2\right] = 0. \tag{4.20}$$

By using (4.20), Granwall's lemma and Burkholder-Davis-Gundy inequality in (4.19), one can show (4.15).

ii) Let us prove (4.16) and (4.17). We put

$$(s,\Delta_s^\varepsilon,u):=\big(s,y_s^q+\lambda\varepsilon(\Upsilon_s^\varepsilon+\widehat{y}_s),\mathbb{E}[y_s^q+\lambda\varepsilon(\Upsilon_s^\varepsilon+\widehat{y}_s)],Y_s^q+\lambda\varepsilon(\mathbb{Y}_s^\varepsilon+\widehat{Y}_s)$$
$$,\mathbb{E}[Y_s^q+\lambda\varepsilon(\mathbb{Y}_s^\varepsilon+\widehat{Y}_s)],Z_s^q+\lambda\varepsilon(\mathbb{Z}_s^\varepsilon+\widehat{Z}_s),\mathbb{E}[Z_s^q+\lambda\varepsilon(\mathbb{Z}_s^\varepsilon+\widehat{Z}_s)],u\big).$$

From (4.4), (4.14) and (4.18) we have

$$\begin{cases} d\mathbb{Y}_t^\varepsilon=-\big(F_Y^\varepsilon\mathbb{Y}_t^\varepsilon+\mathbb{E}\left[F_{Y'}^\varepsilon\mathbb{E}[\mathbb{Y}_t^\varepsilon]\right]+F_Z^\varepsilon\mathbb{Z}_t^\varepsilon+\mathbb{E}\left[F_{Z'}^\varepsilon\mathbb{E}[\mathbb{Z}_t^\varepsilon]\right]+\Theta_t^\varepsilon\big)dt \\ \qquad -\big(g_Y\left(t,\Delta_t^\varepsilon\right)\mathbb{Y}_t^\varepsilon+\mathbb{E}\left[g_{Y'}\left(t,\Delta_t^\varepsilon\right)\mathbb{E}[\mathbb{Y}_t^\varepsilon]\right]+g_Z\left(t,\Delta_t^\varepsilon\right)\mathbb{Z}_t^\varepsilon \\ \qquad\qquad +\mathbb{E}\left[g_{Z'}\left(t,\Delta_t^\varepsilon\right)\mathbb{E}[\mathbb{Z}_t^\varepsilon]\right]+\Xi_t^\varepsilon\big)\overleftarrow{dB}_t+\mathbb{Z}_t^\varepsilon dW_t \\ \\ \mathbb{Y}_T^\varepsilon=\frac{1}{\varepsilon}\left(h\left(y_T^\varepsilon,\mathbb{E}[y_T^\varepsilon]\right)-h\left(y_T^q,\mathbb{E}[y_T^q]\right)\right) \\ \qquad\qquad -\left(h_{y'}\left(y_T^q,\mathbb{E}[y_T^q]\right)\widehat{y}_T+\mathbb{E}\left[h_{y'}\left(y_T^q,\mathbb{E}[y_T^q]\right)\mathbb{E}[\widehat{y}_T]\right]\right), \end{cases} \tag{4.21}$$

where

$$F_\varpi^{\varepsilon,q}=\int_0^1\int_U f_\varpi\big(t,\Delta_t^\varepsilon,u\big)q_t(du)d\lambda,\quad \text{for } \varpi=y,y',Y,Y',Z,Z',$$

$$
\begin{aligned}
\Theta_t^\varepsilon &= F_y^{\varepsilon,q}\Upsilon_t^\varepsilon + \mathbb{E}\left[F_{y'}^{\varepsilon,q}\mathbb{E}[\Upsilon_t^\varepsilon]\right] + F_y^{\varepsilon,q}\widehat{y}_t + \mathbb{E}\left[F_{y'}^{\varepsilon,q}\mathbb{E}[\widehat{y}_t]\right] \\
&\quad -\int_U f_y\big(t,\pi_t^q,u\big)q_t(du)\widehat{y}_t - \mathbb{E}\left[\int_U f_{y'}\big(t,\pi_t^q,u\big)q_t(du)\mathbb{E}[\widehat{y}_t]\right] + F_Y^{\varepsilon,q}\widehat{Y}_t \\
&\quad +\mathbb{E}\left[F_{Y'}^{\varepsilon,q}\mathbb{E}[\widehat{Y}_t]\right] - \int_U f_Y\big(t,\pi_t^q,u\big)q_t(du)\widehat{Y}_t - \mathbb{E}\left[\int_U f_{Y'}\big(t,\pi_t^q,u\big)q_t(du)\mathbb{E}[\widehat{Y}_t]\right] \\
&\quad +F_Z^{\varepsilon,q}\widehat{Z}_t + \mathbb{E}\left[F_{Z'}^{\varepsilon,q}\mathbb{E}[\widehat{Z}_t]\right] - \int_U f_Z\big(t,\pi_t^q,u\big)q_t(du)\widehat{Z}_t \\
&\qquad\quad -\mathbb{E}\left[\int_U f_{Z'}\big(t,\pi_t^q,u\big)q_t(du)\mathbb{E}[\widehat{Z}_t]\right] \\
&\quad +F_y^{\varepsilon,\mu}\left(y_t^\varepsilon - y_t^q\right) + \mathbb{E}\left[F_{y'}^{\varepsilon,\mu}\mathbb{E}[y_t^\varepsilon - y_t^q]\right] + F_Y^{\varepsilon,\mu}\left(Y_t^\varepsilon - Y_t^q\right) \\
&\qquad\quad +\mathbb{E}\left[F_{Y'}^{\varepsilon,\mu}\mathbb{E}[Y_t^\varepsilon - Y_t^q]\right] + F_Z^{\varepsilon,\mu}\left(Z_t^\varepsilon - Z_t^q\right) + \mathbb{E}\left[F_{Z'}^{\varepsilon,\mu}\mathbb{E}[Z_t^\varepsilon - Z_t^q]\right] \\
&\quad -\big(F_y^{\varepsilon,q}\left(y_t^\varepsilon - y_t^q\right) + \mathbb{E}\left[F_{y'}^{\varepsilon,q}\mathbb{E}[y_t^\varepsilon - y_t^q]\right] + F_Y^{\varepsilon,q}\left(Y_t^\varepsilon - Y_t^q\right) \\
&\qquad\quad +\mathbb{E}\left[F_{Y'}^{\varepsilon,q}\mathbb{E}[Y_t^\varepsilon - Y_t^q]\right] + F_Z^{\varepsilon,q}\left(Z_t^\varepsilon - Z_t^q\right) + \mathbb{E}\left[F_{Z'}^{\varepsilon,q}\mathbb{E}[Z_t^\varepsilon - Z_t^q]\right]\big),
\end{aligned}
$$

and

$$
\begin{aligned}
\Xi_t^\varepsilon &= \int_0^1 \big(g_y\big(t,\Delta_t^\varepsilon\big)\widehat{y}_t + \mathbb{E}\left[g_{y'}\big(t,\Delta_t^\varepsilon\big)\mathbb{E}[\widehat{y}_t]\right] - g_y\big(t,\pi_t^q\big)\widehat{y}_t - \mathbb{E}\left[g_{y'}\big(t,\pi_t^q\big)\mathbb{E}[\widehat{y}_t]\right]\big)d\lambda\overleftarrow{dB}_t \\
&\quad + \int_0^1 \big(g_Y\big(t,\Delta_t^\varepsilon\big)\widehat{Y}_t + \mathbb{E}\left[g_{Y'}\big(t,\Delta_t^\varepsilon\big)\mathbb{E}[\widehat{Y}_t]\right] - g_Y\big(t,\pi_t^q\big)\widehat{Y}_t - \mathbb{E}\left[g_{Y'}\big(t,\pi_t^q\big)\mathbb{E}[\widehat{Y}_t]\right]\big)d\lambda\overleftarrow{dB}_t \\
&\quad + \int_0^1 \big(g_Z\big(t,\Delta_t^\varepsilon\big)\widehat{Z}_t + \mathbb{E}\left[g_{Z'}\big(t,\Delta_t^\varepsilon\big)\mathbb{E}[\widehat{Z}_t]\right] - g_Z\big(t,\pi_t^q\big)\widehat{Z}_t - \mathbb{E}\left[g_{Z'}\big(t,\pi_t^q\big)\mathbb{E}[\widehat{Z}_t]\right]\big)d\lambda\overleftarrow{dB}_t.
\end{aligned}
$$

Using the fact that the derivatives $f_y, f_{y'}, f_Y, f_{Y'}, f_Z, f_{Z'}$ are continuous and bounded and from (4.7), (4.8), (4.9) and (4.15) we show

$$
\lim_{\varepsilon\to 0}\mathbb{E}\left[\int_t^T |\Theta_s^\varepsilon|^2 ds\right] = 0, \text{ and } \lim_{\varepsilon\to 0}\mathbb{E}\left[\int_t^T |\Xi_s^\varepsilon|^2 ds\right] = 0. \tag{4.22}
$$

Applying Itô's formula to $|\mathbb{Y}_t^\varepsilon|^2$ we obtain

$$\begin{aligned}\mathbb{E}\left[|\mathbb{Y}_t^\varepsilon|^2\right]+\mathbb{E}\left[\int_t^T\|\mathbb{Z}_s^\varepsilon\|^2ds\right]&=\mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right]+2\mathbb{E}\left[\int_t^T\langle\mathbb{Y}_s^\varepsilon,+F_s^Y\mathbb{Y}_s^\varepsilon+\mathbb{E}\left[F_s^{Y'}\mathbb{E}[\mathbb{Y}_s^\varepsilon]\right]\right.\\&\left.+F_s^Z\mathbb{Z}_s^\varepsilon+\mathbb{E}\left[F_s^{Z'}\mathbb{E}[\mathbb{Z}_s^\varepsilon]\right]+\Theta_s^\varepsilon\rangle ds\right]+\mathbb{E}\left[\int_t^T|g_Y\left(t,\Delta_t^\varepsilon\right)\mathbb{Y}_s^\varepsilon\right.\\&\left.+\mathbb{E}\left[g_{Y'}\left(t,\Delta_t^\varepsilon\right)\mathbb{E}[\mathbb{Y}_s^\varepsilon]\right]+g_Z\left(t,\Delta_t^\varepsilon\right)\mathbb{Z}_s^\varepsilon+\mathbb{E}\left[g_{Z'}\left(t,\Delta_t^\varepsilon\right)\mathbb{E}[\mathbb{Z}_s^\varepsilon]\right]+\Xi_s^\varepsilon|^2ds\right].\end{aligned}$$

Applying Young's inequality and the boundedness of the derivatives $F_s^Y, F_s^{Y'}, F_s^Z, F_s^{Z'}, g_Y, g_{Y'}, g_Z, g_{Z'}$, we obtain

$$\begin{aligned}\mathbb{E}\left[|\mathbb{Y}_t^\varepsilon|^2\right]+\mathbb{E}\left[\int_t^T\|\mathbb{Z}_s^\varepsilon\|^2ds\right]&\leq\mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right]+\frac{1}{\theta_1}\mathbb{E}\left[\int_t^T|\mathbb{Y}_s^\varepsilon|^2ds\right]\\&+5C\theta_1\mathbb{E}\left[\int_t^T\left(|\mathbb{Y}_s^\varepsilon|^2+\mathbb{E}\left[|\mathbb{Y}_s^\varepsilon|^2\right]+\|\mathbb{Z}_s^\varepsilon\|^2+\mathbb{E}\left[\|\mathbb{Z}_s^\varepsilon\|^2\right]+|\Theta_s^\varepsilon|^2\right)ds\right]\\&+3C\mathbb{E}\left[\int_t^T\left(|\mathbb{Y}_s^\varepsilon|^2+\mathbb{E}\left[|\mathbb{Y}_s^\varepsilon|^2\right]+|\Xi_s^\varepsilon|^2\right)ds\right]\\&+2\gamma^2\mathbb{E}\left[\int_t^T\left(\|\mathbb{Z}_s^\varepsilon\|^2+\mathbb{E}\left[\|\mathbb{Z}_s^\varepsilon\|^2\right]\right)ds\right]\\&+2C\gamma\mathbb{E}\left[\int_t^T\langle\mathbb{Y}_s^\varepsilon+\mathbb{E}\left[\mathbb{Y}_s^\varepsilon\right]+\Xi_s^\varepsilon,\mathbb{Z}_s^\varepsilon+\mathbb{E}\left[\mathbb{Z}_s^\varepsilon\right]\rangle ds\right].\end{aligned}$$

Applying Young's inequality again

$$\begin{aligned}\mathbb{E}\left[|\mathbb{Y}_t^\varepsilon|^2\right]+\mathbb{E}\left[\int_t^T\|\mathbb{Z}_s^\varepsilon\|^2ds\right]&\leq\mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right]+\frac{1}{\theta_1}\mathbb{E}\left[\int_t^T|\mathbb{Y}_s^\varepsilon|^2ds\right]\\&+5C\theta_1\mathbb{E}\left[\int_t^T\left(|\mathbb{Y}_s^\varepsilon|^2+\mathbb{E}\left[|\mathbb{Y}_s^\varepsilon|^2\right]+\|\mathbb{Z}_s^\varepsilon\|^2+\mathbb{E}\left[\|\mathbb{Z}_s^\varepsilon\|^2\right]+|\Theta_s^\varepsilon|^2\right)ds\right]\\&+3C\mathbb{E}\left[\int_t^T\left(|\mathbb{Y}_s^\varepsilon|^2+\mathbb{E}\left[|\mathbb{Y}_s^\varepsilon|^2\right]+|\Xi_s^\varepsilon|^2\right)ds\right]+2\gamma^2\mathbb{E}\left[\int_t^T\left(\|\mathbb{Z}_s^\varepsilon\|^2+\mathbb{E}\left[\|\mathbb{Z}_s^\varepsilon\|^2\right]\right)ds\right]\\&+\frac{6C\gamma}{\theta_2}\mathbb{E}\left[\int_t^T\left(|\mathbb{Y}_s^\varepsilon|^2+\mathbb{E}\left[|\mathbb{Y}_s^\varepsilon|^2\right]+|\Xi_s^\varepsilon|^2\right)ds\right]\\&+2C\gamma\,\theta_2\mathbb{E}\left[\int_t^T\left(\|\mathbb{Z}_s^\varepsilon\|^2+\mathbb{E}\left[\|\mathbb{Z}_s^\varepsilon\|^2\right]\right)ds\right].\end{aligned}$$

Hence

$$
\begin{aligned}
\mathbb{E}\left[|\mathbb{Y}_t^\varepsilon|^2\right] + \mathbb{E}\left[\int_t^T \|\mathbb{Z}_s^\varepsilon\|^2 ds\right] \\
&\leq \mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right] + (\frac{1}{\theta_1} + 10C\ \theta_1 + 6C + \frac{12C\gamma}{\theta_2})\mathbb{E}\left[\int_t^T |\mathbb{Y}_s^\varepsilon|^2 ds\right] \\
&+(10C\ \theta_1 + 4\gamma^2 + 8C\gamma\ \theta_2)\mathbb{E}\left[\int_t^T \|\mathbb{Z}_s^\varepsilon\|^2 ds\right] \\
&+5C\ \theta_1 \mathbb{E}\left[\int_t^T |\Theta_s^\varepsilon|^2 ds\right] + (3C + \frac{6C\gamma}{\theta_2})\mathbb{E}\left[\int_t^T |\Xi_s^\varepsilon|^2 ds\right]. \qquad (4.23)
\end{aligned}
$$

We choose

$$
\theta_1 = \frac{1-4\gamma^2}{20C} > 0, \theta_2 = \frac{1-4\gamma^2}{24C\ \gamma} > 0,
$$

thus

$$
10C\ \theta_1 + 4\gamma^2 + 8C\gamma\ \theta_2 = \frac{1-4\gamma^2}{2} + 4\gamma^2 + \frac{1-4\gamma^2}{3} = \frac{5+4\gamma^2}{6} < 1.
$$

Then the inequality (4.22) becomes

$$
\begin{aligned}
\mathbb{E}\left[|\mathbb{Y}_t^\varepsilon|^2\right] + K_1\mathbb{E}\left[\int_t^T \|\mathbb{Z}_s^\varepsilon\|^2 ds\right] \leq \mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right] + K_2\mathbb{E}\left[\int_t^T |\mathbb{Y}_s^\varepsilon|^2 ds\right] \\
+K_3\mathbb{E}\left[\int_t^T |\Theta_s^\varepsilon|^2 ds\right] + K_4\mathbb{E}\left[\int_t^T |\Xi_s^\varepsilon|^2 ds\right], \qquad (4.24)
\end{aligned}
$$

with $K_1 = \frac{1-4\gamma^2}{6} > 0, K_2 > 0, K_3 > 0, K_4 > 0$.

We derive from (4.24) two inequality

$$
\begin{aligned}
\mathbb{E}\left[|\mathbb{Y}_t^\varepsilon|^2\right] \leq \mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right] + K_2\mathbb{E}\left[\int_t^T |\mathbb{Y}_s^\varepsilon|^2 ds\right] \\
+K_3\mathbb{E}\left[\int_t^T |\Theta_s^\varepsilon|^2 ds\right] + K_4\mathbb{E}\left[\int_t^T |\Xi_s^\varepsilon|^2 ds\right], \qquad (4.25)
\end{aligned}
$$

and

$$\mathbb{E}\left[\int_t^T \|\mathbb{Z}_s^\varepsilon\|^2 ds\right] \leq \frac{1}{K_1}\mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right] + \frac{K_2}{K_1}\mathbb{E}\left[\int_t^T |\mathbb{Y}_s^\varepsilon|^2 ds\right] + \frac{K_3}{K_1}\mathbb{E}\left[\int_t^T |\Theta_s^\varepsilon|^2 ds\right] + \frac{K_4}{K_1}\mathbb{E}\left[\int_t^T |\Xi_s^\varepsilon|^2 ds\right]. \qquad (4.26)$$

On the other hand we have

$$\begin{aligned}\mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right] &= \mathbb{E}\Bigg[\Bigg|\frac{1}{\varepsilon}\left(h\left(y_T^\varepsilon, \mathbb{E}[y_T^\varepsilon]\right) - h\left(h_T^q, \mathbb{E}[y_T^q]\right)\right) \\ &\qquad - \left(h_y\left(y_T^q, \mathbb{E}[y_T^q]\right)\widehat{y}_T + \mathbb{E}\left[h_{y'}\left(y_T^q, \mathbb{E}[y_T^q]\right)\mathbb{E}[\widehat{y}_T]\right]\right)\Bigg|^2\Bigg] \\ &\leq 4\mathbb{E}\left[\left|\int_0^1 h_y\left(\Lambda_T^\varepsilon\right)d\lambda - h_y\left(y_T^q, \mathbb{E}[y_T^q]\right)\right|^2 \cdot |\widehat{y}_T|^2\right] \\ &\qquad + 4\mathbb{E}\left[\mathbb{E}\left[\left|\int_0^1 h_{y'}\left(\Lambda_T^\varepsilon\right)d\lambda - h_{y'}\left(y_T^q, \mathbb{E}[y_T^q]\right)\right|^2\right] \cdot \mathbb{E}\left[|\widehat{y}_T|^2\right]\right] \\ &\qquad + 4\mathbb{E}\left[\int_0^1 \left(|h_y\left(\Lambda_T^\varepsilon\right)|^2 \cdot |\Upsilon_T^\varepsilon|^2 + \mathbb{E}\left[|h_{y'}\left(\Lambda_T^\varepsilon\right)|^2\right] \cdot \mathbb{E}\left[|\Upsilon_T^\varepsilon|^2\right]\right)d\lambda\right].\end{aligned}$$

Since h_y, h'_y are continuous and bounded, using (4.15) to get

$$\lim_{\varepsilon\to 0}\mathbb{E}\left[|\mathbb{Y}_T^\varepsilon|^2\right] = 0. \qquad (4.27)$$

Now, applying Gronwall's lemma in (4.25) and using (4.22) and (4.27) to obtain (4.16) and from (4.16), (4.22) and (4.27) we get (4.17). ■

Proposition 4.2.3 *[Variational inequality] Let* $(\mathbf{H4}) - (\mathbf{H5})$, *hold. Let* $q_\cdot$ *be an optimal relaxed control with associated trajectories* $(y_t^{q_\cdot}, Y_t^{q_\cdot}, Z_t^{q_\cdot})$. *Then, for any element* $\mu_\cdot$ *of* $\mathcal{R}$, *we have*

$$
\begin{aligned}
0 \leq\ & \mathbb{E}\left[\alpha_y(y_T^q, \mathbb{E}[y_T^q])\widehat{y}_T + \mathbb{E}\left[\alpha_{y'}(y_T^q, \mathbb{E}\left[y_T^q\right])\mathbb{E}[\widehat{y}_T]\right]\right] \\
& + \mathbb{E}\left[\beta_Y(Y_0^q, \mathbb{E}[Y_0^q])\widehat{Y}_0 + \mathbb{E}\left[\beta_{Y'}(Y_0^q, \mathbb{E}[Y_0^q])\mathbb{E}[\widehat{Y}_0]\right]\right] \\
& + \mathbb{E}\Big[\int_0^T \int_U \big(\ell_y(t, \pi_t^q, u)\widehat{y}_t + \mathbb{E}\left[\ell_{y'}(t, \pi_t^q, u)\mathbb{E}[\widehat{y}_t]\right] \\
& \qquad + \ell_Y(t, \pi_t^q, u)\widehat{Y}_t + \mathbb{E}\left[\ell_{Y'}(t, \pi_t^q, u)\mathbb{E}[\widehat{Y}_t]\right] \\
& \qquad + \ell_Z(t, \pi_t^q, u)\widehat{Z}_t + \mathbb{E}\left[\ell_{Z'}(t, \pi_t^q, u)\mathbb{E}[\widehat{Z}_t]\right]\big) q_t(du)dt\Big] \\
& + \mathbb{E}\Big[\int_0^T \Big(\int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)\mu_t(du) \\
& \qquad - \int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)q_t(du)\Big)dt\Big]. \qquad (4.28)
\end{aligned}
$$

Proof. From the optimality of $q.$ we have

$$
\begin{aligned}
0 \leq\ & \mathbb{E}\left[\alpha(y_T^\varepsilon, \mathbb{E}[y_T^\varepsilon]) - \alpha(y_T^q, \mathbb{E}[y_T^q])\right] + \mathbb{E}\left[\beta(Y_0^\varepsilon, \mathbb{E}[Y_0^\varepsilon]) - \beta(Y_0^q, \mathbb{E}[Y_0^q])\right] \\
& + \mathbb{E}\Big[\int_0^T \Big(\int_U \ell(t, y_t^\varepsilon, \mathbb{E}[y_t^\varepsilon], Y_t^\varepsilon, \mathbb{E}[Y_t^\varepsilon], Z_t^\varepsilon, \mathbb{E}[Z_t^\varepsilon], u)q_t^\varepsilon(du) \\
& \qquad - \int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)q_t^\varepsilon(du)\Big)dt\Big] \\
& + \mathbb{E}\Big[\int_0^T \Big(\int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)q_t^\varepsilon(du) \\
& \qquad - \int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)q_t(du)\Big)dt\Big].
\end{aligned}
$$

Let us divide this inequality by ε and using the definition of q_t^ε and from the notation (4.18), we have

$$
\begin{aligned}
0 \leq \mathbb{E}&\left[\int_0^1 \left(\alpha_y(\Lambda_T^\varepsilon)\widehat{y}_T + \mathbb{E}\left[\alpha_{y'}(\Lambda_T^\varepsilon)\mathbb{E}[\widehat{y}_T]\right]\right) d\lambda\right] \\
&+\mathbb{E}\left[\int_0^1 \Big(\beta_Y(Y_0^q + \lambda\varepsilon(\mathbb{Y}_0^\varepsilon + \widehat{Y}_0), \mathbb{E}[Y_0^q + \lambda\varepsilon(\mathbb{Y}_0^\varepsilon + \widehat{Y}_s)])\widehat{Y}_0\right. \\
&\left.+\mathbb{E}\left[\beta_{Y'}(Y_0^q + \lambda\varepsilon(\mathbb{Y}_0^\varepsilon + \widehat{Y}_0), \mathbb{E}[Y_0^q + \lambda\varepsilon(\mathbb{Y}_0^\varepsilon + \widehat{Y}_s)])\mathbb{E}[\widehat{Y}_0]\right]\Big) d\lambda\right] \\
&+\mathbb{E}\Big[\int_0^T \int_0^1 \int_U \big(\ell_y(t, \Delta_t^\varepsilon, u)\widehat{y}_t + \mathbb{E}\left[\ell_{y'}(t, \Delta_t^\varepsilon, u)\mathbb{E}[\widehat{y}_t]\right] \\
&\qquad\qquad +\ell_Y(t, \Delta_t^\varepsilon, u)\widehat{Y}_t + \mathbb{E}\left[\ell_{Y'}(t, \Delta_t^\varepsilon, u)\mathbb{E}[\widehat{Y}_t]\right] \\
&\qquad\qquad +\ell_Z(t, \Delta_t^\varepsilon, u)\widehat{Y}_t + \mathbb{E}\left[\ell_{Z'}(t, \Delta_t^\varepsilon, u)\mathbb{E}[\widehat{Y}_t]\right]\big) q_t(du) d\lambda dt\Big] \\
&+\mathbb{E}\Big[\int_0^T \Big(\int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)\mu_t(du) \\
&\qquad\qquad - \int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u) q_t(du)\Big) dt\Big] + \nabla_t^\varepsilon,
\end{aligned}
\tag{4.29}
$$

where ∇_t^ε is given by

$$\begin{aligned}
\nabla_t^\varepsilon &= \mathbb{E}\left[\int_0^1 \left(\alpha_y(\Lambda_T^\varepsilon)\Upsilon_T^\varepsilon + \mathbb{E}\left[\alpha_{y'}(\Lambda_T^\varepsilon)\mathbb{E}[\Upsilon_T^\varepsilon]\right]\right) d\lambda\right] \\
&+\mathbb{E}\left[\int_0^1 \Big(\beta_Y\big(Y_0^q + \lambda\varepsilon(\mathbb{Y}_0^\varepsilon + \widehat{Y}_0), \mathbb{E}[Y_0^q + \lambda\varepsilon(\mathbb{Y}_0^\varepsilon + \widehat{Y}_s)]\big)\mathbb{Y}_0^\varepsilon\right. \\
&+\mathbb{E}\left.\Big[\beta_{Y'}\big(Y_0^q + \lambda\varepsilon(\mathbb{Y}_0^\varepsilon + \widehat{Y}_0), \mathbb{E}[Y_0^q + \lambda\varepsilon(\mathbb{Y}_0^\varepsilon + \widehat{Y}_s)]\big)\mathbb{E}[\mathbb{Y}_0^\varepsilon]\Big]\Big) d\lambda\right] \\
&+\mathbb{E}\Big[\int_0^T\int_0^1\int_U \big(\ell_y(t,\Delta_t^\varepsilon,u)(y_t^\varepsilon - y_t^q) + \mathbb{E}\left[\ell_{y'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[y_t^\varepsilon - y_t^q]\right] \\
&\qquad +\ell_Y(t,\Delta_t^\varepsilon,u)(Y_t^\varepsilon - Y_t^q) + \mathbb{E}\left[\ell_{Y'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[Y_t^\varepsilon - Y_t^q]\right] \\
&\qquad +\ell_Z(t,\Delta_t^\varepsilon,u)(Z_t^\varepsilon - Z_t^q) + \mathbb{E}\left[\ell_{Z'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[Z_t^\varepsilon - Z_t^q]\right]\big)\mu_t(du)d\lambda dt\Big] \\
&-\mathbb{E}\Big[\int_0^T\int_0^1\int_U \big(\ell_y(t,\Delta_t^\varepsilon,u)(y_t^\varepsilon - y_t^q) + \mathbb{E}\left[\ell_{y'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[y_t^\varepsilon - y_t^q]\right] \\
&\qquad +\ell_Y(t,\Delta_t^\varepsilon,u)(Y_t^\varepsilon - Y_t^q) + \mathbb{E}\left[\ell_{Y'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[Y_t^\varepsilon - Y_t^q]\right] \\
&\qquad +\ell_Z(t,\Delta_t^\varepsilon,u)(Z_t^\varepsilon - Z_t^q) + \mathbb{E}\left[\ell_{Z'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[Z_t^\varepsilon - Z_t^q]\right]\big)q_t(du)d\lambda dt\Big] \\
&+\mathbb{E}\Big[\int_0^T\int_0^1\int_U \big(\ell_y(t,\Delta_t^\varepsilon,u)\Upsilon_t^\varepsilon + \mathbb{E}\left[\ell_{y'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[\Upsilon_t^\varepsilon]\right] \\
&\qquad +\ell_Y(t,\Delta_t^\varepsilon,u)\mathbb{Y}_t^\varepsilon + \mathbb{E}\left[\ell_{Y'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[\mathbb{Y}_t^\varepsilon]\right] \\
&\qquad +\ell_Z(t,\Delta_t^\varepsilon,u)\mathbb{Z}_t^\varepsilon + \mathbb{E}\left[\ell_{Z'}(t,\Delta_t^\varepsilon,u)\mathbb{E}[\mathbb{Z}_t^\varepsilon]\right]\big)q_t(du)d\lambda dt\Big].
\end{aligned}$$

Since the derivatives $\alpha_y, \alpha_{y'}, \beta_Y, \beta_{Y'}, \ell_y, \ell_{y'}, \ell_Y, \ell_{Y'}, \ell_Z, \ell_{Z'}$ are continuous and bounded, then by using (4.7), (4.8),(4.9), (4.15), (4.16),(4.17) and the Cauchy-Schwartz inequality we show that

$$\lim_{\varepsilon\to 0}\mathbb{E}\left[|\nabla_t^\varepsilon|^2\right] = 0.$$

Then let ε go to 0 in (4.29), we get the variational inequality. ■

4.2.2 Necessary optimality conditions for relaxed control

Let us introduce the adjoint equations of the MF-FBDSDE (4.4) and then gives the maximum principle.

Define the Hamiltonian H from

$$[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d}\times U\times\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^{m\times l}\times\mathbb{R}^{n\times d},$$

to $\mathbb{R}$ by

$$\begin{aligned} H(t,y,y',Y,Y',Z,Z',\mu,\Phi,\Psi,\Sigma,\Pi) &:= \Phi\int_U b(t,y,y',u)\mu(du)+\Sigma\sigma(t,y,y') \\ &+\Psi\int_U f(t,y,y',Y,Y',Z,Z',u)\mu(du)+\Pi g(t,y,y',Y,Y',Z,Z') \\ &+\int_U \ell(t,y,y',Y,Y',Z,Z',u)\mu(du). \end{aligned} \tag{4.30}$$

Theorem 4.2.1 *(Necessary optimality conditions for relaxed control) Assume that* (**H4**)$-$(**H5**) , *hold. Let* $q. \in \mathcal{R}$ *be an optimal relaxed control. Let* (y^q,Y^q,Z^q) *be the associated solution of MF-FBDSDE (4.4). Then there exists a unique solution* $(\Phi^q,\Psi^q,\Sigma^q,\Pi^q)$ *of the following adjoint equations of MF-FBDSDE (4.4):*

$$\begin{cases} d\Phi^q_t = -\big(H_y(t,\zeta^q_t,q_t,\chi^q_t)+\mathbb{E}\big[H_{y'}(t,\zeta^q_t,q_t,\chi^q_t)\big]\big)dt+\Sigma^q_t dW_t, \\ d\Psi^q_t = \big(H_Y(t,\zeta^q_t,q_t,\chi^q_t)+\mathbb{E}\big[H_{Y'}(t,\zeta^q_t,q_t,\chi^q_t)\big]\big)dt \\ \qquad +\big(H_Z(t,\zeta^q_t,q_t,\chi^q_t)+\mathbb{E}\big[H_{Z'}(t,\zeta^q_t,q_t,\chi^q_t)\big]\big)dW_t-\Pi^q_t \overleftarrow{dB}_t, \\ \\ \Psi^q_0 = \beta_Y(Y^q_0,\mathbb{E}[Y^q_0])+\mathbb{E}\big[\beta_{Y'}(Y^q_0,\mathbb{E}[Y^q_0])\big], \\ \Phi^q_T = \alpha_y(y^q_T,\mathbb{E}[y^q_T])+\mathbb{E}\big[\alpha_{y'}(y^q_T,\mathbb{E}[y^q_T])\big] \\ \qquad +h_y(y^q_T,\mathbb{E}[y^q_T])\Psi^q_T+\ \mathbb{E}\big[h_{y'}(y^q_T,\mathbb{E}[y^q_T])\mathbb{E}[\Psi^q_T]\big], \end{cases} \tag{4.31}$$

such that

$$\begin{aligned} &H(t,y^q_t,\mathbb{E}[y^q_t],Y^q_t,\mathbb{E}[Y^q_t],Z^q_t,\mathbb{E}[Z^q_t],q_t,\Phi^q_t,\Psi^q_t,\Sigma^q_t,\Pi^q_t) \\ &\quad\le H(t,y^q_t,\mathbb{E}[y^q_t],Y^q_t,\mathbb{E}[Y^q_t],Z^q_t,\mathbb{E}[Z^q_t],\mu_t,\Phi^q_t,\Psi^q_t,\Sigma^q_t,\Pi^q_t) \\ &\qquad ,\ a.e.\, t,\mathbb{P}-a.s.,\ \forall\mu\in\mathbb{P}(U), \end{aligned} \tag{4.32}$$

where $(t, \zeta_t^q, q_t, \chi_t^q) := (t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], q_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)$.

Proof. From (4.31), the variational inequality (4.28) becomes

$$\begin{aligned}
0 \leq \; & \mathbb{E}\left[\langle \Phi_T^q, \widehat{y}_T \rangle\right] - \mathbb{E}\left[h_y(y_T^q, \mathbb{E}[y_T^q])\Psi_T^q + \mathbb{E}\big[h_{y'}(y_T^q, \mathbb{E}[y_T^q])\mathbb{E}[\Psi_T^q]\big]\right] \\
& + \mathbb{E}\left[\langle \Psi_0^q, \widehat{Y}_0 \rangle\right] + \mathbb{E}\Big[\int_0^T \int_U \big(\ell_y(t, \pi_t^q, u)\widehat{y}_t + \mathbb{E}\left[\ell_{y'}(t, \pi_t^q, u)\mathbb{E}[\widehat{y}_t]\right] \\
& \qquad + \ell_Y(t, \pi_t^q, u)\widehat{Y}_t + \mathbb{E}\left[\ell_{Y'}(t, \pi_t^q, u)\mathbb{E}[\widehat{Y}_t]\right] \\
& \qquad + \ell_Z(t, \pi_t^q, u)\widehat{Z}_t + \mathbb{E}\left[\ell_{Z'}(t, \pi_t^q, u)\mathbb{E}[\widehat{Z}_t]\right]\big) q_t(du)dt\Big] \\
& + \mathbb{E}\Big[\int_0^T \Big(\int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)\mu_t(du) \\
& \qquad - \int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)q_t(du)\Big)dt\Big]. \qquad (4.33)
\end{aligned}$$

Now applying Itô's formula to compute $\langle \Phi_t^q, \widehat{y}_t \rangle$ and $\langle \Psi_t^q, \widehat{Y}_t \rangle$ and taking the expectations we derive

$$\begin{aligned}
\mathbb{E}\big[\langle \Phi_T^q, \widehat{y}_T \rangle\big] = & - \mathbb{E}\Big[\int_0^T \langle \Psi_t^q \int_U \big(f_y(t, \pi_t^q, u) + \mathbb{E}\big[f_{y'}(t, \pi_t^q, u)\big]\big) q_t(du) \\
& + \Pi_t^q\big(g_y(t, \pi_t^q) + \mathbb{E}\big[g_{y'}(t, \pi_t^q)\big]\big) + \int_U \big(\ell_y(t, \pi_t^q, u) + \mathbb{E}\big[\ell_{y'}(t, \pi_t^q, u)\big]\big) q_t(du), \widehat{y}_t \rangle dt\Big] \\
& + \mathbb{E}\Big[\int_0^T \Phi_t^q\Big(\int_U b(t, y_t^q, \mathbb{E}[y_t^q], u)q_t(du) - \int_U b(t, y_t^q, \mathbb{E}[y_t^q], u)\mu_t(du)\Big)dt\Big],
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\big[\langle \Psi_0^q, \widehat{Y}_0\rangle\big] = & \ \mathbb{E}\big[\langle \Psi_T^q, \widehat{Y}_T\rangle\big] \\
& + \mathbb{E}\Big[\int_0^T \langle \Psi_t^q, \int_U \big(f_y(t,\pi_t^q,u)\widehat{y}_t + \mathbb{E}\big[f_{y'}(t,\pi_t^q,u)\mathbb{E}[\widehat{y}_t]\big]\big) q_t(du)\rangle dt\Big] \\
& + \mathbb{E}\Big[\int_0^T \langle \Pi_t^q, \big(g_y(t,\pi_t^q)\widehat{y}_t + \mathbb{E}\big[g_{y'}(t,\pi_t^q)\mathbb{E}[\widehat{y}_t]\big]\big)\rangle dt\Big] \\
& - \mathbb{E}\Big[\int_0^T \langle \int_U \big(\ell_Y(t,\pi_t^q,u) + \mathbb{E}\big[\ell_{Y'}(t,\pi_t^q,u)\big]\big) q_t(du), \widehat{Y}_t\rangle dt\Big] \\
& - \mathbb{E}\Big[\int_0^T \langle \int_U \big(\ell_Z(t,\pi_t^q,u) + \mathbb{E}\big[\ell_{Z'}(t,\pi_t^q,u)\big]\big) q_t(du), \widehat{Z}_t\rangle dt\Big] \\
& + \mathbb{E}\Big[\int_0^T \Psi_t^q\Big(\int_U f(t,\pi_t^q,u) q_t(du) - \int_U f(t,\pi_t^q,u)\mu_t(du)\Big) dt\Big].
\end{aligned}$$

Substitute the above equalities in inequality (4.33) to get, for every $\mu \in \mathcal{R}$,

$$\begin{aligned}
0 \leq \ \mathbb{E}\Big[\int_0^T & \big(H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], q_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q) \\
& - H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], \mu_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)\big) dt\Big].
\end{aligned}$$

Therefore inequality (4.32) follows by a standard arguments. ■

4.2.3 Sufficient optimality conditions for relaxed control

In this subsection we study when the necessary conditions for optimality in Theorem 4.2.1 become sufficient as well.

Theorem 4.2.2 *(Sufficient optimality conditions for relaxed control) Assume that (H4) holds. Given* $q_\cdot \in \mathcal{R}$*, let* (y^q, Y^q, Z^q) *and* $(\Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)$ *be the corresponding solutions of the MF-FBDSDEs (4.4) and (4.31) respectively. Suppose that* α, β, ℓ *and the function* $H(t, \cdot, \cdot, \cdot, \cdot, q_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)$ *are convex.*
Then $(y_\cdot^q, Y_\cdot^q, Z_\cdot^q, q_\cdot)$ *is an optimal solution of the control problem (4.4)–(4.6) if it satisfies (4.32).*

Proof. Let $q_{\cdot} \in \mathcal{R}$ be arbitrary (candidate to be optimal), and let $(y_{\cdot}^{q}, Y_{\cdot}^{q}, Z_{\cdot}^{q})$ denote the trajectory associated to $q_{\cdot}$. For any $\mu_{\cdot} \in \mathcal{R}$ with associated trajectory $(y_{\cdot}^{\mu}, Y_{\cdot}^{\mu}, Z_{\cdot}^{\mu})$, we have

$$\begin{aligned}\mathbb{J}(\mu_{\cdot}) - \mathbb{J}(q_{\cdot}) = {} & \mathbb{E}\big[\alpha(y_T^{\mu}, \mathbb{E}[y_T^{\mu}]) - \alpha(y_T^{q}, \mathbb{E}[y_T^{q}])\big] + \mathbb{E}\big[\beta(Y_0^{\mu}, \mathbb{E}[Y_0^{\mu}]) - \beta(Y_0^{q}, \mathbb{E}[Y_0^{q}])\big] \\ & + \mathbb{E}\Big[\int_0^T \Big(\int_U \ell(t, y_t^{\mu}, \mathbb{E}[y_t^{\mu}], Y_t^{\mu}, \mathbb{E}[Y_t^{\mu}], Z_t^{\mu}, \mathbb{E}[Z_t^{\mu}], u)\mu_t(du) \\ & \qquad - \int_U \ell(t, y_t^{q}, \mathbb{E}[y_t^{q}], Y_t^{q}, \mathbb{E}[Y_t^{q}], Z_t^{q}, \mathbb{E}[Z_t^{q}], u)q_t(du)\Big)dt\Big].\end{aligned}$$

Since α and β are convex, we get

$$\begin{aligned}\alpha(y_T^{\mu}, \mathbb{E}[y_T^{\mu}]) - \alpha(y_T^{q}, \mathbb{E}[y_T^{q}]) \geq {} & \langle \alpha_y(y_T^{q}, \mathbb{E}[y_T^{q}]), y_T^{\mu} - y_T^{q}\rangle \\ & + \mathbb{E}\big[\langle \alpha_{y'}(y_T^{q}, \mathbb{E}[y_T^{q}]), \mathbb{E}[y_T^{\mu} - y_T^{q}]\rangle\big], \\ \beta(Y_0^{\mu}, \mathbb{E}[Y_0^{\mu}]) - \beta(Y_0^{q}, \mathbb{E}[Y_0^{q}]) \geq {} & \langle \beta_Y(Y_0^{q}, \mathbb{E}[Y_0^{q}]), Y_0^{\mu} - Y_0^{q}\rangle \\ & + \mathbb{E}\big[\langle \beta_{Y'}(Y_0^{q}, \mathbb{E}[Y_0^{q}]), \mathbb{E}[Y_0^{\mu} - Y_0^{q}]\rangle\big].\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{J}(\mu_{\cdot}) - \mathbb{J}(q_{\cdot}) \geq {} & \langle \alpha_y(y_T^{q}, \mathbb{E}[y_T^{q}]), y_T^{\mu} - y_T^{q}\rangle + \mathbb{E}\big[\langle \alpha_{y'}(y_T^{q}, \mathbb{E}[y_T^{q}]), \mathbb{E}[y_T^{\mu} - y_T^{q}]\rangle\big] \\ & + \langle \beta_Y(Y_0^{q}, \mathbb{E}[Y_0^{q}]), Y_0^{\mu} - Y_0^{q}\rangle + \mathbb{E}\big[\langle \beta_{Y'}(Y_0^{q}, \mathbb{E}[Y_0^{q}]), \mathbb{E}[Y_0^{\mu} - Y_0^{q}]\rangle\big] \\ & + \mathbb{E}\Big[\int_0^T \Big(\int_U \ell(t, y_t^{\mu}, \mathbb{E}[y_t^{\mu}], Y_t^{\mu}, \mathbb{E}[Y_t^{\mu}], Z_t^{\mu}, \mathbb{E}[Z_t^{\mu}], u)\mu_t(du) \\ & \qquad - \int_U \ell(t, y_t^{q}, \mathbb{E}[y_t^{q}], Y_t^{q}, \mathbb{E}[Y_t^{q}], Z_t^{q}, \mathbb{E}[Z_t^{q}], u)q_t(du)\Big)dt\Big].\end{aligned}$$

Therefore after recalling also (4.31) one gets

$$\begin{aligned}\mathbb{J}(\mu_.) - \mathbb{J}(q_.) \geq \mathbb{E}\big[\langle \Phi_T^q, y_T^\mu - y_T^q\rangle\big] \\ -\mathbb{E}\big[\langle h_y(y_T^q, \mathbb{E}[y_T^q])\Psi_T^q + \mathbb{E}\big[h_{y'}(y_T^q, \mathbb{E}[y_T^q])\mathbb{E}[\Psi_T^q]\big], y_T^\mu - y_T^q\rangle\big] \\ + \mathbb{E}\big[\langle \Psi_0^q, Y_0^\mu - Y_0^q\rangle\big] \\ + \mathbb{E}\big[\int_0^T \big(\int_U \ell(t, y_t^\mu, \mathbb{E}[y_t^\mu], Y_t^\mu, \mathbb{E}[Y_t^\mu], Z_t^\mu, \mathbb{E}[Z_t^\mu], u)\mu_t(du) \\ - \int_U \ell(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)q_t(du)\big)dt\big]. \qquad (4.34)\end{aligned}$$

Applying Itô's formula to $\langle \Phi_t^q, y_t^\mu - y_t^q\rangle$ and $\langle \Psi_t^q, Y_t^\mu - Y_t^q\rangle$, we obtain

$$\begin{aligned}\mathbb{E}\big[\,\langle \Phi_T^q, y_T^\mu - y_T^q\rangle\,\big] = \mathbb{E}\big[\int_0^T \langle \Phi_t^q, \int_U b(t, y_t^\mu, \mathbb{E}[y_t^\mu], u)\mu_t(du) \\ - \int_U b(t, y_t^q, \mathbb{E}[y_t^q], u)q_t(du)\rangle dt\big] \\ + \mathbb{E}\big[\int_0^T \langle \Sigma_t^q, \sigma(t, y_t^\mu, \mathbb{E}[y_t^\mu]) - \sigma(t, y_t^q, \mathbb{E}[y_t^q])\rangle\, dt\big] \\ - \mathbb{E}\big[\int_0^T \langle H_y(t, \zeta_t^q, q_t, \chi_t^q) + \mathbb{E}[H_{y'}(t, \zeta_t^q, q_t, \chi_t^q)], y_t^\mu - y_t^q\rangle dt\big], \qquad (4.35)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\big[\,\langle \Psi_0^q, Y_0^\mu - Y_0^q\rangle\,\big] = \mathbb{E}\big[\,\langle \Psi_T^q, Y_T^\mu - Y_T^q\rangle\,\big] \\ - \mathbb{E}\big[\int_0^T \langle H_Y(t, \zeta_t^q, q_t, \chi_t^q) + \mathbb{E}[H_{Y'}(t, \zeta_t^q, q_t, \chi_t^q)], Y_t^\mu - Y_t^q\rangle dt\big] \\ + \mathbb{E}\big[\int_0^T \langle \Psi_t^q, \int_U f(t, y_t^\mu, \mathbb{E}[y_t^\mu], Y_t^\mu, \mathbb{E}[Y_t^\mu], Z_t^\mu, \mathbb{E}[Z_t^\mu], u)\mu_t(du) \\ - \int_U f(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], u)q_t(du)\rangle dt\big] \\ - \mathbb{E}\big[\int_0^T \langle H_Z(t, \zeta_t^q, q_t, \chi_t^q) + \mathbb{E}[H_{Z'}(t, \zeta_t^q, q_t, \chi_t^q)], Z_t^\mu - Z_t^q\rangle dt\big] \\ + \mathbb{E}\big[\int_0^T \langle \Psi_t^q, g(t, y_t^\mu, \mathbb{E}[y_t^\mu], Y_t^\mu, \mathbb{E}[Y_t^\mu], Z_t^\mu, \mathbb{E}[Z_t^\mu]) \\ - g(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q])\rangle dt\big]. \qquad (4.36)\end{aligned}$$

From the convexity of h we have

$$\begin{aligned}\mathbb{E}\big[\,\langle \Psi_T^q, Y_T^\mu - Y_T^q\rangle\,\big] &= \mathbb{E}\big[\,\langle \Psi_T^q, h(y_T^\mu, \mathbb{E}[y_T^\mu]) - h(y_T^q, \mathbb{E}[y_T^q])\rangle\,\big]\\ &\geq \mathbb{E}\big[\,\big\langle h_y(y_T^q, \mathbb{E}[y_T^q])\Psi_T^q + \mathbb{E}\big[h_{y'}(y_T^q, \mathbb{E}[y_T^q])\mathbb{E}[\Psi_T^q]\big], y_T^\mu - y_T^q\big\rangle\,\big]. \end{aligned} \tag{4.37}$$

Replacing (4.35) and (4.36) in inequality (4.34) and using (4.37), we get

$$\begin{aligned}\mathbb{J}(\mu_\cdot) - \mathbb{J}(q_\cdot) \geq \mathbb{E}\Big[\int_0^T &\big(H(t, \zeta_t^q, \mu_t, \chi_t^q) - H(t, \zeta_t^q, q_t, \chi_t^q)\big)dt\\ &- \mathbb{E}\Big[\int_0^T \langle H_y(t, \zeta_t^q, q_t, \chi_t^q) + \mathbb{E}[H_{y'}(t, \zeta_t^q, q_t, \chi_t^q)], y_t^\mu - y_t^q\rangle dt\Big]\\ &- \mathbb{E}\Big[\int_0^T \langle H_Y(t, \zeta_t^q, q_t, \chi_t^q) + \mathbb{E}[H_{Y'}(t, \zeta_t^q, q_t, \chi_t^q)], Y_t^\mu - Y_t^q\rangle dt\Big]\\ &- \mathbb{E}\Big[\int_0^T \langle H_Z(t, \zeta_t^q, q_t, \chi_t^q) + \mathbb{E}[H_{Z'}(t, \zeta_t^q, q_t, \chi_t^q)], Z_t^\mu - Z_t^q\rangle dt\Big]. \end{aligned} \tag{4.38}$$

On the other hand, by the convexity of $H(t, y, y', Y, Y', Z, Z', q, \Phi, \Psi, \Sigma, \Pi)$ in (y, y', Y, Y', Z, Z') and its linearity in q, then by using the clarke generalized gradient of H evaluated at (y, y', Y, Y', Z, Z'), we obtain

$$\begin{aligned}H(t, \zeta_t^q, \mu_t, \chi_t^q) - H(t, \zeta_t^q, q_t, \chi_t^q) \geq\, & H_y(t, \zeta_t^q, q_t, \chi_t^q)\big(y_t^\mu - y_t^q\big)\\ &+ \mathbb{E}\big[H_{y'}(t, \zeta_t^q, q_t, \chi_t^q)\mathbb{E}[y_t^\mu - y_t^q]\big] + H_Y(t, \zeta_t^q, q_t, \chi_t^q)\big(Y_t^\mu - Y_t^q\big)\\ &+\mathbb{E}\big[H_{Y'}(t, \zeta_t^q, q_t, \chi_t^q)\mathbb{E}[Y_t^\mu - Y_t^q]\big] + H_Z(t, \zeta_t^q, q_t, \chi_t^q)\big(Z_t^\mu - Z_t^q\big)\\ &+\mathbb{E}\big[H_{Z'}(t, \zeta_t^q, q_t, \chi_t^q)\mathbb{E}[Z_t^\mu - Z_t^q]\big]. \end{aligned}$$

Therefore, applying this inequality in (4.38) gives

$$\mathbb{J}(\mu_\cdot) - \mathbb{J}(q_\cdot) \geq 0, \forall \mu \in \mathcal{R}.$$

The theorem is proved. ■

4.3 Necessary and Sufficient optimality conditions for strict control

In this section, we study the strict control problem $\{(4.1), (4.2), (4.3)\}$ and from the results of section (4.2), we derive the optimality conditions for strict controls. For this end, consider the following subset of $\mathcal{R}$

$$\mathcal{R}^{\delta}=\{\mu_{\cdot}\in\mathcal{R}/\ \ \mu=\delta_v : v\in\mathcal{U}\},$$

the set of all relaxed controls in the form of Dirac measure charging a strict control. Denote by $\mathbb{P}(U^{\delta})$ the action set of all relaxed control $\mathcal{R}^{\delta}$.

4.3.1 Necessary optimality conditions for strict control

Define the Hamiltonian $\mathcal{H}$ in the strict control problem from

$$[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d}\times U\times\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^{m\times l}\times\mathbb{R}^{n\times d},$$

to $\mathbb{R}$ by

$$\begin{aligned}\mathcal{H}(t,y,y',Y,Y',Z,Z',v,\Phi,\Psi,\Sigma,\Pi) &:= +\Phi b(t,y,y',v)+\ \Sigma\sigma(t,y,y')\\ &+\Psi f(t,y,y',Y,Y',Z,Z',v)+\Pi g(t,y,y',Y,Y',Z,Z',v)\\ &+\ell(t,y,y',Y,Y',Z,Z',v). \end{aligned} \tag{4.39}$$

Theorem 4.3.1 *(Necessary optimality conditions for strict control.) Let $u_{\cdot}\in\mathcal{U}$ be an optimal strict control. Let (y^u,Y^u,Z^u) be the associated solution of MF-FBDSDE (4.1). Then there exists a unique solution $(\Phi^u,\Psi^u,\Sigma^u,\Pi^u)$ of the following adjoint equations of*

MF-FBDSDE (4.1):

$$\begin{cases} d\Phi^u_t = -\big(\mathcal{H}_y(t,\zeta^u_t,u_t,\chi^u_t) + \mathbb{E}\big[\mathcal{H}_{y'}(t,\zeta^u_t,u_t,\chi^u_t)\big]\big)dt + \Sigma^u_t dW_t, \\ d\Psi^u_t = \big(\mathcal{H}_Y(t,\zeta^u_t,u_t,\chi^u_t) + \mathbb{E}\big[\mathcal{H}_{Y'}(t,\zeta^u_t,u_t,\chi^u_t)\big]\big)dt \\ \qquad\qquad + \big(\mathcal{H}_Z(t,\zeta^u_t,u_t,\chi^u_t) + \mathbb{E}\big[\mathcal{H}_{Z'}(t,\zeta^u_t,u_t,\chi^u_t)\big]\big)dW_t - \Pi^u_t \overleftarrow{dB}_t, \\ \\ \Psi^u_0 = \beta_Y(Y^u_0,\mathbb{E}[Y^u_0]) + \mathbb{E}\big[\beta_{Y'}(Y^u_0,\mathbb{E}[Y^u_0])\big], \\ \Phi^u_T = \alpha_y(y^u_T,\mathbb{E}[y^u_T]) + \mathbb{E}\big[\alpha_{y'}(y^u_T,\mathbb{E}[y^u_T])\big] \\ \qquad\qquad + h_y(y^u_T,\mathbb{E}[y^u_T])\Psi^u_T + \mathbb{E}\big[h_{y'}(y^u_T,\mathbb{E}[y^u_T])\mathbb{E}[\Psi^u_T]\big], \end{cases} \tag{4.40}$$

such that

$$\begin{aligned} &\mathcal{H}(t,y^u_t,\mathbb{E}[y^u_t],Y^u_t,\mathbb{E}[Y^u_t],Z^u_t,\mathbb{E}[Z^u_t],u_t,\Phi^u_t,\Psi^u_t,\Sigma^u_t,\Pi^u_t) \\ &\quad \le \mathcal{H}(t,y^u_t,\mathbb{E}[y^u_t],Y^u_t,\mathbb{E}[Y^u_t],Z^u_t,\mathbb{E}[Z^u_t],v_t,\Phi^u_t,\Psi^u_t,\Sigma^u_t,\Pi^u_t), \ \ a.e.\, t, \mathbb{P}\text{-}a.s., \ \forall v \in \mathcal{U}, \end{aligned} \tag{4.41}$$

where $(t,\zeta^u_t,u_t,\chi^u_t) := (t,y^u_t,\mathbb{E}[y^u_t],Y^u_t,\mathbb{E}[Y^u_t],Z^u_t,\mathbb{E}[Z^u_t],u_t,\Phi^u_t,\Psi^u_t,\Sigma^u_t,\Pi^u_t)$.

Proof. Note that the strict $u.$ embedded into the space $\mathbb{V}$ in the sense that $u.$ is corresponding with the Dirac measure $\lambda_{u.}(dt,da) = \delta_{u.}(du)$ with the propriety: For any bounded and uniformly continuous function $\hbar(t,y,y',Y,Y',Z,Z',u)$ we have

$$\begin{aligned} \hbar(t,y,y',Y,Y',Z,Z',u_t) &= \int_U \hbar(t,y,y',Y,Y',Z,Z',u)\delta_{u_t}(du) \\ &:= \widehat{\hbar}(t,y,y',Y,Y',Z,Z',\lambda_u). \end{aligned} \tag{4.42}$$

From the necessary optimality condition for relaxed controls (Theorem 4.2.1), there exist a unique solution $(\Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)$ of (4.31) such that

$$\begin{aligned}&H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], q_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)\\ &\quad \leq H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], \mu_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q),\ a.e.\, t, \mathbb{P}\text{-}a.s.,\ \forall \mu \in \mathcal{R},\end{aligned}$$

and since $\mathcal{R}^\delta \subset \mathcal{R}$ we have

$$\begin{aligned}&H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], q_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)\\ &\quad \leq H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], \mu_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q),\ a.e.\, t, \mathbb{P}\text{-}a.s.,\ \forall \mu \in \mathcal{R}^\delta.\end{aligned} \tag{4.43}$$

Using the fact that if $\mu \in \mathcal{R}^\delta$, then there exist $v_t \in U^\delta \subset U$ such that $\mu = \delta_v$, and if the optimal relaxed control $q_t(du) = \delta_{u_t}(du)$ with u_t an optimal strict control, then we can show that

$$\begin{aligned}&(y_t^q, Y_t^q, Z_t^q) = (y_t^u, Y_t^u, Z_t^u),\ (y_t^\mu, Y_t^\mu, Z_t^\mu) = (y_t^v, Y_t^v, Z_t^v),\\ &(\Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q) = (\Phi_t^u, \Psi_t^u, \Sigma_t^u, \Pi_t^u),\ (\Phi_t^\mu, \Psi_t^\mu, \Sigma_t^\mu, \Pi_t^\mu) = (\Phi_t^v, \Psi_t^v, \Sigma_t^v, \Pi_t^v),\\ &H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], q_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)\\ &\qquad\qquad = \mathcal{H}(t, y_t^u, \mathbb{E}[y_t^u], Y_t^u, \mathbb{E}[Y_t^u], Z_t^u, \mathbb{E}[Z_t^u], u_t, \Phi_t^u, \Psi_t^u, \Sigma_t^u, \Pi_t^u),\\ &H(t, y_t^\mu, \mathbb{E}[y_t^\mu], Y_t^\mu, \mathbb{E}[Y_t^\mu], Z_t^\mu, \mathbb{E}[Z_t^\mu], \mu_t, \Phi_t^\mu, \Psi_t^\mu, \Sigma_t^\mu, \Pi_t^\mu)\\ &\qquad\qquad = \mathcal{H}(t, y_t^v, \mathbb{E}[y_t^v], Y_t^v, \mathbb{E}[Y_t^v], Z_t^v, \mathbb{E}[Z_t^v], v_t, \Phi_t^v, \Psi_t^v, \Sigma_t^v, \Pi_t^v).\end{aligned} \tag{4.44}$$

Using (4.42) and (4.43) we get (4.41). The proof is completed. ■

4.3.2 Sufficient optimality conditions for strict control

We shall try to shows if the necessary optimality conditions (4.41) for strict control problem $\{(4.1), (4.2), (4.3)\}$ becomes sufficient.

Theorem 4.3.2 *(Sufficient optimality conditions for strict control.) Assume that the functions* α, β, ℓ *and* $\mathcal{H}(t, \cdot, \cdot, \cdot, \cdot, u_t, \Phi_t^u, \Psi_t^u, \Sigma_t^u, \Pi_t^u)$ *are convex. Then* $(y_\cdot^u, Y_\cdot^u, Z_\cdot^u, u_\cdot)$ *is an optimal solution of the strict control problem* $\{(4.1), (4.2), (4.3)\}$ *if it satisfies (4.41).*

Proof. Let u_t be an arbitrary element of U^δ such that the necessary optimality conditions for strict control (4.41) hold, i.e.

$$\begin{aligned}&\mathcal{H}(t, y_t^u, \mathbb{E}[y_t^u], Y_t^u, \mathbb{E}[Y_t^u], Z_t^u, \mathbb{E}[Z_t^u], u_t, \Phi_t^u, \Psi_t^u, \Sigma_t^u, \Pi_t^u)\\ &\quad \leq \mathcal{H}(t, y_t^u, \mathbb{E}[y_t^u], Y_t^u, \mathbb{E}[Y_t^u], Z_t^u, \mathbb{E}[Z_t^u], v_t, \Phi_t^u, \Psi_t^u, \Sigma_t^u, \Pi_t^u),\ a.e.\, t, \mathbb{P}\text{-}a.s.,\ \forall v \in \mathcal{U}^\delta,\end{aligned}$$

and by applying the embedding mentioned in (4.42), one can show that

$$\begin{aligned}&H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], q_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q)\\ &\quad \leq H(t, y_t^q, \mathbb{E}[y_t^q], Y_t^q, \mathbb{E}[Y_t^q], Z_t^q, \mathbb{E}[Z_t^q], \mu_t, \Phi_t^q, \Psi_t^q, \Sigma_t^q, \Pi_t^q),\ a.e.\, t, \mathbb{P}\text{-}a.s.,\ \forall \mu \in \mathcal{R}^\delta.\end{aligned}$$

Thus by sufficient optimality conditions for relaxed control (Theorem 4.2.2) we have

$$\mathbb{J}(q_\cdot) = \inf_{\mu_\cdot \in \mathcal{R}^\delta} \mathbb{J}(\mu_\cdot),$$

and from the fact that the optimal relaxed control is a Dirac measure charging in optimal strict control ($q_t(du) = \delta_{u_t}(du)$) and by using (4.44), we can show that

$$\mathbb{J}(u_\cdot) = \inf_{v_\cdot \in \mathcal{U}^\delta} \mathbb{J}(v_\cdot).$$

The prove is completed. ∎

Conclusion

In this thesis, we have investigated some results on the existence of the optimal control as well as the necessary and sufficient conditions for optimality.

We have studied in first optimal control problems for systems governed by a linear forward-backward doubly stochastic differential equations (FBDSDEs), we have proved the existence of a strong optimal control (that is adapted to a fixed sigma algebra), by using the convexity of the cost functional and the domain of control and the Mazur's theorem. We have derived also necessary and sufficient conditions for optimality for this control problem. This result is based on the convex optimization principle.

In the second part, we have established the existence of strong optimal solutions of a control problem for dynamics driven by a linear forward-backward doubly stochastic differential equations of mean-field type (MF-FBDSDEs) in which the coefficients of the system depend on the states of the solution processes as well as their distribution via the expectation of the states. Moreover, the cost functional is also of mean-field type. Moreover, we have established necessary as well as sufficient optimality conditions satisfied by an optimal strict control of control problem of MF-FBDSDEs.

In the third part, we have established necessary as well as sufficient optimality conditions for both relaxed and strict control problems driven by systems of nonlinear MF-FBDSDEs. In the same context, we can reformulate the control problem for systems governed by a linear or nonlinear forward-backward doubly stochastic differential equations of McKean–Vlasov type (MV-FBDSDEs). The coefficients of the McKean–Vlasov systems depend on the state of the solution process as well as of its probability law. And we can prove the existence of the optimal control and establish the necessary as well as sufficient conditions for optimality for this kind of systems.

A special case is that in which both α, β and l are convex quadratic functions. The control problem $\{(3.1), (3.2), (3.3)\}$ is then reduced to a stochastic linear quadratic optimal control problem.

Bibliography

[1] N.U. Ahmed and X. Ding, *Controlled McKean-Vlasov equations*, Comm. in Appl. Analysis. 5, no. 2, 183-206, (2001).

[2] A. Al-Hussein and B. Gherbal, *Sufficient conditions of optimality for forward-backward doubly SDEs with jumps,* Statistical methods and applications in insurance and finance, Eddahbi, et al. (eds.), Springer proceedings in mathematics and statistics, 158, 173-191. (2016).

[3] A. Al-Hussein and B. Gherbal, *Existence and Uniqueness of the Solutions of Forward Backward Doubly Stochastic Differential Equations with Poisson Jumps*, Random. Oper. Stoc. Equ., Vol. 28, No 4, 253-268. (2020).

[4] D. Andersson and B. Djehiche, *A maximum principle for SDEs of mean-field type*, Appl. Math. and Optim, 63, no. 3, 341-356, (2010).

[5] F. Antonelli, *Backward-forward stochastic differential equations*, Ann. Appl. Probab, 3, 777–793, (1993).

[6] K. Bahlali, B. Gherbal and B. Mezerdi, *Existence and optimality conditions in stochastic control of linear BSDEs,* Random Oper. Stoch. Equ., Vol. 18, No 3, 185-197. (2010).

[7] Bahlali S, Necessary and sufficient optimality conditions for relaxed and strict control problems, SIAM J. Control Optim., 2008, 47(4): 2078–2095.

[8] R. Benbrahim and B. Gherbal, *Existence of Optimal Controls for Forward-Backward Stochastic Differential Equations of Mean-Field Type*, J. of Numerical Mathematics and Stochastics, 9, no. 1, 33-47, (2017).

[9] N. Berrouis, B. Gherbal and A. Ninouh, *Stochastic optimal control for dynamics of forward backward doubly SDEs of mean-field type*, Bol. Soc. Paran. Mat. c SPM – ISSN-2175-1188 on line, (2020).

[10] J.M. Bismut, *Conjugate convex functions in optimal stochastic control*, J. Math. Anal. Appl. 44, 384–404, (1973).

[11] R. Buckdahn, B. Djehiche, J. Li and S. Peng, *Mean-Field backward stochastic differential equations: a limit approach*, Ann. Prob., 37, no. 4, 1524-1565, (2009).

[12] R. Buckdahn, B. Djehiche and J. Li, *A general stochastic maximum principle for SDEs of mean-field type,* Appl. Math. Optim. 64, no. 2, 197-216, (2011).

[13] R. Carmona and F. Delarue, *Probabilistic analysis of mean-field games.* SIAM J. Control Optim. 51, no. 4, 2705-2734, (2013).

[14] I. Ekeland and R. Temam, *Analyse convexe et problème variationnel*, Dunod (1974).

[15] 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), 169–219. (1987).

[16] W. H. Fleming, *Generalized solutions in optimal stochastic control, Differential Games and Control Theory*, II (Proceedings of 2nd Conference, University of Rhode Island, Kingston, RI, (1976), Lecture Notes in Pure and Appl. Math., 30, Marcel Dekker, New York, pp. 147–165, (1977).

[17] J. François Le Gall, *Brownian Motion, Martingales, and Stochastic Calculus,* Springer International Publishing Switzerland (2016).

[18] G. Fu and U. Horst, *Mean Field Games with Singular Controls*, SIAM J. Control and Optim., 55, no. 6, 3833-3868, (2017).

[19] B. Gherbal, *Optimal control problems for linear backward doubly stochastic differential equations*. Random. Oper. Stoc. Equ., Vol. 22, No 3, 129–138. (2014).

[20] M. Hafayed, *A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes,* Int. J. Dyn. Control 1, no. 4, 300-315, (2013).

[21] M. Hafayed, M. Tabet and S. Boukaf, *Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and applications to mean-variance portfolio problem,* Comm. Math. Stat, 3(2), 163–186, (2015).

[22] 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, Springer, Berlin, pp. 171–186. (1986).

[23] 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, 529–546, (1981).

[24] H.J. Kushner, *Existence results for optimal stochastic controls,* J. Optim. Theory Appl. 15, 347–359. (1975).

[25] J.M. Lasry and P.L. Lions, *Mean-field games*. Japan. J. Math. 2, 229-260, (2007).

[26] R. Li and B. Liu, *A maximum principle for fully coupled stochastic control systems of mean-field type,* J. Math. Anal. Appl, 415, 902–930, (2014).

[27] J. Ma and J. Yong, *Forward backward stochastic differential equations and their applications*, in: Lecture notes in Mathematics, vol. 1702, Springer, Berlin, Heidelberg, (1999).

[28] T. Meyer-Brandis, B. Øksendal and X.Y. Zhou, *A mean-field stochastic maximum principle via Malliavin calculus*, Stochastics 84, 643-666, (2012).

[29] B. Mezerdi and S. Bahlali, *Necessary conditions for optimality in relaxed stochastic control problems*, Stochastics and Stochastics Reports 73, no. 3-4, 201–218, (2002).

[30] A. Ninouh, B. Gherbal and N. Berrouis, *Existence of optimal controls for systems of controlled forward-backward doubly SDEs.* Random. Oper. Stoc. Equ., Vol. 28, No 2, 93-112, (2020).

[31] M. Nisio, *Stochastic Control Theory Dynamic Programming Principle,* Springer Japan (2015).

[32] B. Øksendal, *Stochastic differential equations: An introduction with applications*, 6th edition, Springer, Berlin, (2003).

[33] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems and Control Letters. 14, no.1, 55–61, (1990).

[34] E. Pardoux and S. Peng, *Backward doubly stochastic differential equations and system of quasilinear SPDEs.* Probab. Theory Related Fields 98, No.2, pp 209-227. (1994).

[35] J. Yong and X.Y Zhou, *Stochastic controls, in: Hamiltonian Systems and HJB Equations*, Springer, New York (1999).

[36] K. Yosida, *Functional analysis*, 6th edition, Springer-Verlag, Berlin, (1980).

[37] Q. Zhu and Y. Shi, *Mean-field Forward-Backward Doubly Stochastic Differential Equations and Related Nonlocal Stochastic Partial Differential Equations*, Abstract and Applied Analysis, Hindawi, Article ID 194341, 10 pages, (2014).

Appendix

Mazur's theorem:

Let $(x_n) \to x$ weakly as $n \to \infty$ in a normed linear space $\mathcal{X}$. Then there exists, for any $\varepsilon > 0$, a convex combination $\sum_{j=1}^{n} \alpha_j x_j \left(\alpha_j \geq 0, \ \sum_{j=1}^{n} \alpha_j = 1\right)$ of x_j's such that

$$\left\| x - \sum_{j=1}^{n} \alpha_j x_j \right\| \leq \varepsilon.$$

We need some inequalities in this thesis.

A quadratic inequality:

For any real numbers a and b, we have

$$(a+b)^2 \leq 2a^2 + 2b^2.$$

Young inequality:

For a, $b \geq 0$ and $\varepsilon > 0$, we have

$$ab \leq \frac{a^2}{2\varepsilon} + \frac{\varepsilon b^2}{2}.$$

We introduce Gronwall's inequality

Gronwall's inequality:

Let $v : [0, T] \to [0, \infty)$ be a nonnegative continuous function such that

$$v(t) \le C + A \int_0^t v(s)\, ds \qquad \forall\ 0 \le t \le T.$$

for some nonnegative constants C and A. Then

$$v(t) \le C \exp(At) \qquad \forall\ 0 \le t \le T.$$

Burkholder-Davis-Gundy inequality:

Let X be a square integrable continuous, $\mathcal{F}_t-$martingale, with $X_0 = 0$ Then, for $p \in (0, \infty)$ there exist positive constants c_p and C_p, such that

$$c_p \mathbb{E}\left[\sup_{0 \le s \le t} |X_s|^{2p}\right] \le \mathbb{E}\left[\langle X_t \rangle^p\right] \le C_p \mathbb{E}\left[\sup_{0 \le s \le t} |X_s|^{2p}\right].$$