People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

Mohamed Khider University of Biskra

Faculty of Exact Sciences and Sciences of Nature and Life

Department of Mathematics

MASTER DISSERTATION

Applied Mathematics

Submitted and defended by

Allaoui Chahrazed

Title:

On Stochastic Optimal Control Problems

Members of the Examination Committee:

Pr. Chala Adel . *University of Biskra,* ________ **President**

Dr. Lakhdari Imad Eddine. *University of Biskra,* ________ **Supervisor**

Dr. Korichi Fatiha. *University of Biskra,* ________ **Examiner**

2023

Dedication

I dedicate this work to my beloved brother, Taher Allaoui, who recently passed away. He greatly supported me in my educational journey, motivated me, and instilled a love for learning and life in me. I also dedicate it to my family, who helped me reach this level of education.

To my dear parents, my mother, with her constant prayers, staying up late with me, and her efforts, and my father, who supported me throughout my educational journey.

To my dear brother, who led me to success.

To my sisters, who have encouraged me throughout my life, their children, and their spouses.

To my dear friends, who have supported me and helped me emotionally and intellectually.

I present this work to you with the hope that it pleases and satisfies.

Acknowledgments

I would like to express my special thanks to my God, who has taught me and provided me with the opportunities to learn and succeed. Then, I would like to thank my supervisors **Dr. LAKHDARI Imad Eddine** *for their continuous support and for allowing me to benefit from their knowledge and expertise.*

I am also grateful to **Phd. ABBA Khadidja** *and* **Phd. DOUBBAKH Salima** *who provided me with valuable advice and guidance. My heartfelt appreciation goes to all of them.*

I would like to thank the members of the examination committee Pr. **CHALA Adel** *and* **Dr. KORICHI Fatiha**.

To everyone who contributed to the realization of this achievement.

Thank you.

Contents

Dedication **i**

Acknowledgments **ii**

Table of contents **iii**

Introduction **1**

1 Introduction to stochastic calculus **5**

1.1 Processus stochastic 5

1.2 Conditional expectation 7

1.3 Martingale 8

1.4 Brownian Motion 9

1.5 Quadratic variation 10

1.5.1 Stochastic integrals 11

1.5.2 Itô's process and stochastic differential equations 12

1.6 Some classes of stochastic controls 14

1.7 Useful results 14

2 Existence and uniqueness solution of SDEs and BSDEs with Lipschitz condition **18**

2.1 Stochastic differential equations (**SDE**) 18

2.1.1 Existence and uniqueness result for **SDE** 19
2.2 Backward stochastic differential equation (BSDE) 25
2.2.1 Existence and uniqueness result for **BSDE** with Lipschitz coefficients 26

3 Stochastic maximum principle **29**
3.1 Notations 30
3.2 Lions differentiability 32
3.3 Necessary optimality conditions 39

Conclusion **44**

Appendix A: Abbreviations and Notations **47**

Introduction

Introduction

In this master dissetation, we focus on a control problem within a specific filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$ that involves the following equations:

$$\begin{cases} dx^u(t) = b(t,\ x^u(t),\ P_{x^u(t)},\ u(t))dt + \sigma(t, x^u(t), P_{x^u(t)}, u(t))dB(t), \\ dy^u(t) = -g(t,\ x^u(t),\ P_{x^u(t)}, y^u(t), P_{y^u(t)}, z^u(t), P_{z^u(t)}, u(t))dt + z^u(t)dB(t), \\ x^u(0) = x_0,\ y^u(T) = h(x^u(T),\ P_{x^u(T)}), \end{cases} \tag{1}$$

where $B(\cdot)$ represents a d-dimensional $\mathcal{F}_t$-Brownian motion. Here, $P_{X(t)} = P \circ X^{-1}$ denotes the law of the random variable X. The control $u(\cdot) = (u(t))_{t\geq 0}$ is required to take values in a subset of $\mathbb{R}$ and be adapted to a subfiltration $(\mathcal{G}_t)_{t\geq 0}$ within $(\mathcal{F}_t)$. The maps

$$b : [0,T] \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{A} \to \mathbb{R},$$

$$\sigma : [0,T] \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{A} \to \mathbb{R},$$

$$g : [0,T] \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{A} \to l^2(\mathbb{R}),$$

and

$$h : \mathbb{R} \times Q_2(\mathbb{R}) \to \mathbb{R},$$

are given deterministic functions, where $Q_2(\mathbb{R})$ represents the space of probability measures μ on $(\mathbb{R} \times \beta(\mathbb{R}))$ with finite second moment, i.e., $\int_{\mathbb{R}} |x|^2 \mu(dx) < \infty$, equipped with the

2-Wasserstein metric given by: for $\mu_1, \mu_2 \in Q_2(\mathbb{R})$,

$$D_2(\mu_1, \mu_2) = \inf \left\{ \left(\int_{\mathbb{R}} |x-y|^2 \rho(dx, dy) \right)^{\frac{1}{2}}, \rho \in Q_2(\mathbb{R}), \rho(\cdot, \mathbb{R}) = \mu_1, \rho(\mathbb{R}, \cdot) = \mu_2 \right\}.$$

The McKean-Vlasov forward-backward stochastic differential equation is a general equation where the coefficients depend on the probability law of the solution. It is used in the analysis of financial optimization problems and optimal control in various fields such as economics, finance, physics, chemistry, and game theory. In the context of interacting particle systems, the equation describes the movement of particles, and the empirical measures of these particles converge to a deterministic measure as the number of particles increases. This measure corresponds to the probability distribution of the processes governed by the McKean-Vlasov system.

The cost function to be minimized, known as the McKean-Vlasov-type cost, is given by:

$$\begin{aligned} J(u(\cdot)) = \mathbb{E} \Big\{ \int_0^T & f(t, x^u(t), P_{x^u(t)}, y^u(t), P_{y^u(t)}, z^u(t), P_{z^u(t)}, u(t)) dt \\ & + \phi(x^u(T), P_{x^u(T)}) + \varphi(y^u(0), P_{y^u(0)}) \Big\}, \end{aligned} \tag{2}$$

where

$$f : [0, T] \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{A} \to \mathbb{R},$$
$$\phi : \mathbb{R} \times Q_2(\mathbb{R}) \to \mathbb{R}, \varphi : \mathbb{R} \times Q_2(\mathbb{R}) \to \mathbb{R},$$

are deterministic functions. We remark that the cost functional (2) involves the law of the solution in a nonlinear way.

The general maximum principle for stochastic differential equations (SDEs) has been explored by Meng and Tang ([12]). Mean-field-type SDEs, also known as McKean-Vlasov systems, have a historical background dating back to the works of Kac and McKean ([8], [10]), who investigated stochastic systems with a large number of interacting particles. Optimal control problems for McKean-Vlasov-type SDEs have been extensively studied by many authors ([13, 14]). Hafayed et al. ([6]) have studied the maximum principle for McKean-

Vlasov forward-backward SDEs of mean-field-type. The mean-field maximum principle for SDEs has been established by Buckdahn et al ([2]). Lasry and Lions ([9]) have introduced a mathematical modeling approach for high-dimensional systems involving a large number of particles. Buckdahn et al. ([3]) proved a Peng's type maximum principle for mean-field-type SDEs using second-order derivatives with respect to measures. Hafayed et al. ([7]) investigated the singular optimal control problem for general controlled nonlinear SDEs, where the coefficients depend on the solution process and its law.

Partial information refers to cases where the available information to the controller is limited. This situation arises in real-world applications, such as mathematical finance and economics, where obtaining full information may not be feasible. The partial information maximum principle for SDEs has been established by Wang et al ([15]).

This study is based on part of the work of Meherrem and Hafayed ([16]), where the authors proved the stochastic maximum principle for optimal control of McKean-Vlasov FBSDEs with Lévy process via the differentiability with respect to probability law.

This work has three chapters.

• The first chapter is preliminary and serves to introduce the essential tools for the second and third chapters.

• In the second chapter, we will show the existence and uniqueness of the solution to an (SDE), (BSDE) in the case where both the drift term and the diffusion term satisfy the Lipschitz condition in x, and linear growth and the generator satisfies the Lipschitz condition in y, z, and the terminal value is square integrable respectively.

• In the third chapter, we prove the necessary optimality conditions, which are our main results. The differential system is governed by general McKean–Vlasov forward-backward stochastic differential equation (FBSDE).

Chapter §.1
Introduction to stochastic calculus

Chapter 1

Introduction to stochastic calculus

This chapter serves mainly as an introduction, aimed at highlighting the tools used in our study, and provides some basic reminders about stochastic calculus.

1.1 Processus stochastic

Definition 1.1.1 ***(Random variable):*** *A random variable X is a real-valued measurable function defined as:*

$$X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})),$$

such that for any Borel set, B, its preimage is measurable:

$$\{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\} \in \mathcal{F}, \forall B \in \mathcal{B}(\mathbb{R}),$$

where $\mathcal{F}$ is the sigma-algebra of measurable sets on the sample space Ω and $\mathcal{B}(\mathbb{R})$ is the Borel sigma-algebra on the real line.

Definition 1.1.2 ***(Generated sigma-algebra):*** *The sigma-algebra generated by a random variable X defined on $(\Omega, \mathcal{F})$ is denoted by $\sigma(X)$, and it is defined as the set of all* *pre-images of Borel sets under X, i.e.,*

$$\sigma(X) = \left\{X^{-1}(A) : A \in \varepsilon\right\},$$

where $X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\}$ for any set A in a generator ε of the Borel sigma-algebra on the real line. This sigma-algebra is the smallest sigma-algebra on Ω that contains all sets of the form $X^{-1}(B)$ for Borel sets B, and makes X measurable.

Definition 1.1.3 ***(Stochastic process):*** *A stochastic process $X = (X_t)_{t \in T}$ is a family of random variables X_t indexed by a set T. Typically, $T = \mathbb{R}^+$ and the process is assumed to be indexed by time t.*

- *If T is a finite set, the process is a random vector.*
- *If $T = \mathbb{N}$, the process is a sequence of random variables.*
- *If $T \subset \mathbb{Z}$, the process is said to be discrete.*
- *If $T \subset \mathbb{R}^d$, the process is called a random field.*
- For a fixed $t \in T$, $\omega \in \Omega \to X_t(\omega)$ is a random variable on the probability space $(\Omega, \mathcal{F}, P)$.
- For a fixed $\omega \in \Omega$, $t \in T \to X_t(\omega)$ is a real-valued function, called the trajectory of the process.

Definition 1.1.4 ***(Filtration):*** *A filtration $(\mathcal{F}_t)$ on a probability space $(\Omega, \mathcal{F}, P)$ is an increasing family of sub-sigma-algebras of $\mathcal{F}$, i.e., $\mathcal{F}_s \subset \mathcal{F}_t$, $\forall s \leq t$.*

1. The space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ is called a filtered space.
2. A filtration is P-complete for a probability measure P if $\mathcal{F}_0$ contains all events of measure zero, i.e., $\mathcal{N} = \{N \in \mathcal{F} \text{ such that } P(N) = 0\} \subset \mathcal{F}_0$

3. We say that a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ satisfies the usual conditions if:

- Negligible sets are contained in $\mathcal{F}_0$, i.e., $\mathcal{N} \subset \mathcal{F}_0$,
- The filtration is right-continuous, i.e., $\mathcal{F}_t = \bigcap_{s \geq t} \mathcal{F}_s \forall t$.

Definition 1.1.5 ***(Adapted-measurable-progressively measurable)***

- *A process X is measurable if the application $(t, w) \to X_t(w)$ from $\mathbb{R}_+ \times \Omega$ to $\mathbb{R}^d$ is measurable with respect to the tribu $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ and $\mathcal{B}(\mathbb{R}^d)$.*
- *A process X is adapted with respect to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if, for all $t \geq 0$, X_t is $\mathcal{F}_t$-measurable.*
- *A process X is progressively measurable with respect to $\{\mathcal{F}_t\}_{t \geq 0}$ if for all $t \geq 0$, the application $(s, w) \to X_s(w)$ from $[0, t] \times \Omega$ to $\mathbb{R}^d$ is measurable with respect to $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ and $\mathcal{B}(\mathbb{R}^d)$.*

Remark 1.1.1 *A progressively measurable process is measurable and adapted*

Proposition 1.1.1 *If X is a stochastic process whose trajectories are right-continuous (or left-continuous), then X is measurable and X is progressively measurable if it is also adapted.*

1.2 Conditional expectation

Definition 1.2.1 ***(Conditional expectation with respect to a σ-algebra):*** *Let X be a real-valued random variable (integrable , i.e., $X \in L^1$) defined on $(\Omega, \mathcal{F}, P)$,and let $\mathcal{G}$ be a sub-sigma-algebra of $\mathcal{F}$.*

The conditional expectation $\mathbb{E}[X | \mathcal{G}]$ of X quand $\mathcal{G}$ st l'unique variable aléatoire :

- *$\mathcal{G}$-measurable.*

- $\int_A \mathbb{E}[X|\mathcal{G}]\,dP = \int_A X\ dP,\ \forall A \in \mathcal{G}$.

It is also the unique (up to almost sure equality) $\mathcal{G}$*-measurable random variable such that:*

$$\mathbb{E}[Y\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[XY],$$

for any bounded Y*,* $\mathcal{G}$*-measurable variable.*

Property 1.2.1 ***(Properties of conditional expectation):***

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space, and let $\mathcal{G}$ be a sub-sigma-algebra of $\mathcal{F}$. Let X and Y be two random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$

- Linearity: For any constants a and b, we have $\mathbb{E}(aX + bY|\mathcal{G}) = a\mathbb{E}(X|\mathcal{G}) + b\mathbb{E}(Y|\mathcal{G})$.

- Monotonicity: If X and Y are random variables such that $X \leq Y$, then $\mathbb{E}(X|\mathcal{G}) \leq \mathbb{E}(Y|\mathcal{G})$.

- If X is $\mathcal{G}$-measurable, then $\mathbb{E}(X|\mathcal{G}) = X$.

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

1.3 Martingale

Definition 1.3.1 *A stochastic process,* $\{Y_t : 0 \leq t \leq \infty\}$*, is a martingale with respect to the filtration,* $\mathcal{F}_t$ *and probability measure* P*, if*

- $\mathbb{E}^P[|Yt|] < \infty$ for all $t \geq 0$.

- $\mathbb{E}^P[Y_{t+s}|\mathcal{F}_t] = Y_t$ for all $t, s \geq 0$.

Example 1.3.1 *Let* B_t *be a Brownian motion.*
Then $B_t^2 - t, B_t^3 - 3tB_t$ *and* $\exp(-\lambda^2 \frac{t}{2}) \exp \lambda B_t$*, are all martingales.*

Definition 1.3.2 ***(Continuous-time martingale):*** *A process $(X_t)_{t\geq 0}$ adapted with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and such that for all $t \geq 0$, $X_t \in L^1$ is called:*

- A martingale if for $s \leq t : \mathbb{E}\left[X_t \left| \mathcal{F}_s\right.\right] = X_s$,
- A supermartingale if for $s \leq t : \mathbb{E}\left[X_t \left| \mathcal{F}_s\right.\right] \leq X_s$,
- A submartingale if for $s \leq t : \mathbb{E}\left[X_t \left| \mathcal{F}_s\right.\right] \geq X_s$,
- If X est une martingale $\mathbb{E}\left[X_t\right] = \mathbb{E}\left[X_0\right], \forall t$.
- If $(X_t,\ t \leq T)$ is a martingale, the process is completely determined by its terminal value: $X_t = \mathbb{E}\left[X_T \left| \mathcal{F}_t\right.\right]$. This latter property is very frequently used in finance.

1.4 Brownian Motion

Definition 1.4.1 ***(Brownian motion):*** *A stoshastic process $\{B_t : 0 \leq t \leq \infty\}$ is a standard Brownian motion:*

- $B_0 = 0$.
- *With probability* 1*, the function* $t \rightarrow B_t$ *is continuous in* t.
- *The process* $\{B_t\}_{t\geq 0}$ *has stationary, independent increments.*
- $B_t \sim N(0,t)$.

Definition 1.4.2 *(d**-dimensional Brownian motion**) An d-dimensional Wiener process is a vector-valued stochastic process, $B_t = (B_t^{(1)}, ..., B_t^{(d)})$ is a standard d-dimensional Brownian motion if each $B_t^{(i)}$ it is a standard Brownian motion and the whose components $B_t^{(i)}$'s are independent of each other.*

1.5 Quadratic variation

Suppose that B_t is a real-valued stochastic process defined on a probability space $(\Omega, \mathcal{F}, P)$ and with time index t ranging over the non-negative real numbers, consider a partition ofthe time interval, $[0; T]$ given by

$$0 = t_0 < t_1 < t_2 < ... < t_n = T.$$

Let Y_t be a Brownian motion and consider the sum of squared changes

$$Q_n(T) := \sum_{i=1}^{n} \left[B_{t_i} - B_{t_{i-1}}\right]^2 \tag{1.1}$$

Definition 1.5.1 ***(Quadratic variation):*** *The quadratic variation of a stochastic process, Y_t, is the process, written as $[Y]_t$ is equal to the limit of*

$$Q_n(T) \;\; as \;\; \Delta t := \max_i (t_i - t_{i-1}) \to 0$$

Remark 1.5.1 *The functions with which you are normally familiar, e.g. continuous differentiable functions, have quadratic variation equal to zero. Note that any continuous stochastic process or function that has non-zero quadratic variation must have infinite* ***total variation*** *where the total variation of a process, Y_t, on $[0; T]$ is defined as*

$$Total\ variation := \lim_{\Delta t \to 0} \sum_{k=1}^{n} \left|Y_{t_k} - Y_{t_{k-1}}\right|.$$

This follows by observing that

$$\sum_{k=1}^{n} \left(Y_{t_k} - Y_{t_{k-1}}\right)^2 \leq \left(\sum_{k=1}^{n} \left|Y_{t_k} - Y_{t_{k-1}}\right|\right) \max_{1 \leq k \leq n} \left|Y_{t_k} - Y_{t_{k-1}}\right|. \tag{1.2}$$

If we now let $n \to \infty$ in (1.2) then the continuity of Y_t implies the impossibility of the process having finite total variation and non-zero quadratic variation.

1.5.1 Stochastic integrals

We now discuss the concept of a stochastic integral, ignoring the various technical conditions that are required to make our definitions rigorous. In this section, we write $X_t(\omega)$ instead of the usual X_t to emphasize that the quantities in question are stochastic.

Definition 1.5.2 ***(Stopping time):*** *A stopping time of the filtration* $\mathcal{F}_t$ *is a random time* τ*, such that the event* $\{\tau \leq t\} \in \mathcal{F}_t$ *for all* $t > 0$*.*

In non-mathematical terms, we see that a stopping time is a random time whose value is part of the information accumulated by that time.

Definition 1.5.3 ***(Piece-wise):*** *We say a process* $h_t(\omega)$*, is elementary if it is* ***piece-wise*** *constant so that there exists a sequence of stopping times* $0 = t_0 < t_1 < t_2 < ... < t_n = T$*, and a set of* $\mathcal{F}_{t_i}$*-measurable functions,* $e_i(\omega)$*, such that*

$$h_t(\omega) = \sum_i e_i(\omega) I_{[t_i, t_{i+1})}(t),$$

where $I_{[t_i,t_{i+1})}(t) = 1$ *if* $t \in [t_i, t_{i+1})$ *and* 0 *otherwise.*

Definition 1.5.4 ***(Stochastic integral):*** *A stochastic integral of an elementary function,* $h_t(\omega)$*, with respect to a Brownian motion,* B_t *is defined as*

$$\int_0^T h_t(\omega)\, dB_t(\omega) := \sum_{i=0}^{n-1} e_i(\omega) \left(B_{t_{i+1}}(\omega) - B_{t_i}(\omega) \right). \tag{1.3}$$

Note that, if we interpret $h_t(\omega)$ as a trading strategy and the stochastic integral as the gains or losses from this trading strategy, then evaluating $h_t(\omega)$ at the left-hand point is equivalent to imposing the **non-anticipativity** of the trading strategy, a property that we always wish to impose.

For a more general process, $Y_t(\omega)$, we have

$$\int_0^T Y_t(\omega)\, dB_t(\omega) := \lim_{n\to 0} \int_0^T Y_t^{(n)}(\omega)\, dB_t(\omega),$$

where $Y_t^{(n)}$ is a sequence of elementary processes that converges (in an appropriate manner) to Y_t.

Definition 1.5.5 *We define the space $L^2[0,T]$ to be the space of processes, $Y_t(\omega)$ shuch that:*

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

Theorem 1.5.1 ***(Martingale property of stochastic integrals):***
The stochastic integral, $X_t := \int_0^T Y_t(\omega)\, dB_t$, is a martingale for any $Y_t(\omega) \in L^2[0,T]$.

1.5.2 Itô's process and stochastic differential equations

Definition 1.5.6 ***(Itô process):*** *An n-dimensional Itô process, Y_t, is a process of the form*

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s, \tag{1.4}$$

where B is an m-dimensional standard Brownian motion, and b and σ are n-dimensional and $n \times m$-dimensional $\mathcal{F}_t$-adapted processes, respectively.
We often use the notation

$$dX_t = b_t dt + \sigma_t dB_t,$$

as shorthand for (1.4).

Definition 1.5.7 ***(Stochastic diferential equation):*** *An n-dimensional stochastic diferential equation (SDEs) has the form*

$$dX_t = b_t(X_t, t)\, dt + \sigma_t(X_t, t)\, dB_t; \qquad X_0 = 0, \tag{1.5}$$

where as before, B_t is an m- dimensional standard Brownian motion, and b and σ are n-dimensional and $n \times m$-dimensional adapted processes, respectively. Once again, (1.5) is shorthand for

$$X_t = x + \int_0^t b_s(X_s, s)\, ds + \int_0^t \sigma_s(X_s, s)\, dB_s, \tag{1.6}$$

while we do not discuss the issue here, various conditions exist to guarantee existence and uniqueness of solutions to (1.6). A useful tool for solving SDEs is Itô's Lemma which we now discuss.

Theorem 1.5.2 ***(Itô's formula):***

Let X_t be 1-dimensional Itô process satisfying the SDEs

$$dX_t = \mu_t dt + \sigma_t dB_t.$$

If $f(t.x) : [0, \infty) \times \mathbb{R} \times \mathbb{R}$ is a $C^{1,2}$ function and $Y_t := f(t, X_t)$ then

$$\begin{aligned} dY_t &= \frac{\partial f}{\partial t}(t, X_t)\, dt + \frac{\partial f}{\partial x}(t, X_t)\, dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t)\,(dX_t)^2 \\ &= \left(\frac{\partial f}{\partial t}(t, X_t) + \frac{\partial f}{\partial x}(t, X_t)\, \mu_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t)\, \sigma_t^2\right) dt + \frac{\partial f}{\partial x}(t, X_t)\, \sigma_t dB_t. \end{aligned}$$

The "Box" calculus

In the statement of Itô's Lemma, we implicitly assumed that $(dX_t)^2 = \sigma_t^2 dt$. The box calculus is a series of simple rules for calculating such quantities. In particular, we use the rules

$$dt \times dt = dt \times dB_t = 0,$$
$$\text{and } dB_t \times dB_t = dt,$$

when determining quantities such as $(dB_t)^2$ in the statement of Itô's Lemma above. Note that these rules are consistent with Theorem (1.2.1). When we have two correlated Brownian motions, $B_t^{(1)}$ and $B_t^{(2)}$, with correlation coeficient, ρ_t, then we easily obtain that $dB_t^{(1)} \times dB_t^{(2)} = \rho_t dt$. We use the **box calculus** for computing the quadratic variation of Itô processes.

1.6 Some classes of stochastic controls

let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P\right)$ be a complete filtred probability space.

Definition 1.6.1 ***(Optimal control):*** *The goal of the optimal control problem is to minimize a cost function $J(v)$ over the set of admissible control $\mathcal{U}$. We say that the control $u(\cdot)$ is an optimal control if*

$$J(u(t)) \leq J(v(t)), \text{ for all } v(\cdot) \in \mathcal{U}.$$

Definition 1.6.2 ***(Admissible control):*** *$\mathcal{F}_t$-adapted processes $v(t)$ with values in a borelian $A \subset \mathbb{R}^n$ is An admissible control adapted processes*

$$\mathcal{U} := \{v(\cdot) : [0, T] \times \Omega \to A : v(t) \text{ is } \mathcal{F}_t - \text{adapted}\}.$$

Definition 1.6.3 ***(Feedback control):*** *We say that $v(\cdot)$ is a feedback control if the control $v(\cdot)$ depends on the state variable $X(\cdot)$. If $\mathcal{F}_t^X$ the natural filtration generated by the process X, then $v(\cdot)$ is a feedback control if $v(\cdot)$ is $\mathcal{F}_t^X$-adapted.*

1.7 Useful results

Theorem 1.7.1 ***(Brownian martingale representation theorem):*** *Let $(\mathcal{F}_t)_{0\leq t\leq T}$ be the natural filtration of the Brownian motion $(B_t)_{0\leq t\leq T}$. Let M be a square integrable continuous martingale with respect to $(\mathcal{F}_t)_{0\leq t\leq T}$. Then there exists a unique predictable process H such that:*

$$\mathbb{E}\left(\int_t^T H_s^2 ds\right) < +\infty,$$

for all $\forall t \in [0, T]$ *and*:

$$M_t = M_0 + \int_0^t H_s dB_s, \qquad P - p.s.$$

Theorem 1.7.2 ***(Hôlder inequality):*** *Let* p *and* q *be exponents in the range* $[1, +\infty]$ *such that they are conjugate, i.e.,* $1/p + 1/q = 1$*. If* f *and* g *are measurable functions, then:*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q .$$

In particular, the Hôlder inequality (in the case of $p = 2$*) yields* ***the Cauchy-Schwarz inequality****:*

$$|(f \mid g)| \leq \|f\|_2 \|g\|_2 .$$

Lemma 1.7.1 ***(Gronwall Lemma):*** *Let* f *be an integrable and non-negative function defined for* $t \geq 0$ *and satisfying*

$$f(t) \leq \beta + c \int_0^t f(s) ds,$$

where c *is a positive constant. Then we have:*

$$f(t) \leq \beta \int_0^t \exp(cs) ds.$$

Lemma 1.7.2 ***(Burkholder-Davis-Gundy inequality):*** *For any* $p > 0$ *there exists positive constants* c_p *and* C_p*, such that for any continuous local martingale* $X = (X_t)$ *with* $X(0) = 0$*,*

$$c_p \mathbb{E}\left[\langle X,\, X\rangle_\infty^{\frac{p}{2}}\right] \leq \mathbb{E}\left[\sup_{t\geq 0} |X_t|^p\right] \leq C_p \mathbb{E}\left[\langle X,\, X\rangle_\infty^{\frac{p}{2}}\right].$$

Theorem 1.7.3 ***(Itô's isometry):*** *For any* $Y_t(\omega) \in L^2[0,T]$ *we have*

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

Theorem 1.7.4 ***(Fixed-Point Theorem)****: Let* (E,d) *be a complete metric space and* $\varphi : E \to E$ *be a une application contractante, i.e. Lipschitz with a contraction factor* $k < 1$*. Then,* φ *has a unique fixed point* $x \in E$ *such that* $\varphi(x) = x$*.*

Chapter §.2
Existence and uniqueness solution of SDEs and BSDEs with Lipschitz condition

Chapter 2

Existence and uniqueness solution of SDEs and BSDEs with Lipschitz condition

In this chapter, we study the existence and uniqueness of solutions for two types of equations. Firstly, we consider Stochastic Differential Equations (SDEs) where both the drift term and the diffusion term satisfy the Lipschitz condition and linear growth. Secondly, we investigate Backward Stochastic Differential Equations (BSDEs) where the generator satisfies the Lipschitz condition and the terminal value is square integrable.

2.1 Stochastic differential equations (SDE)

A stochastic differential equation (SDE) is a perturbation of the ordinary differential equation (ODE) with a random term modeling noise around a deterministic phenomenon. The simplest perturbation is the addition of a Brownian motion.

Consider the following stochastic differential equation.

$$\begin{cases} dx_t = b(t,x)dt + \sigma(t,x_t)dB_t, \\ x_0 = \xi, \end{cases} \tag{2.1}$$

check the following conditions

$$P(x_0 = \xi) = 1,$$

$$P(\int_0^t |b(s,x_s)| + \sigma^2(s,x_s)ds < \infty) = 1,$$

$$x_t = \xi + \int_0^t b(s,x_s)ds + \int_0^t \sigma(s,x_s)dB_s.$$

The problem is since stochastic differential equations are a generalization of ordinary differential equations, to show that under certain conditions on the coefficients b, σ, the differential equation has a unique solution. We assume that

$$|b(t,x) - b(t,y)|^2 + |\sigma(t,x) - \sigma(t,y)|^2 \leq k|x-y|^2, \tag{2.2}$$

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \leq k(1 + |x_t|^2). \tag{2.3}$$

2.1.1 Existence and uniqueness result for SDE

Theorem 2.1.1 *If the coefficients b and σ satisfy the conditions (2.2) and (2.3), then the equation (2.1) has a unique strong solution $X = (X_t)_{t\in[0,T]}$, $\mathcal{F}_t$-adapted, and continuous with initial condition $X_0 = \xi$. Furthermore, this solution is Markovian and satisfies*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^p\right] \leq M, \forall p > 1.$$

Where M is a constant that depends on k, p, T and ξ.

Remark 2.1.1 *Noticing that the Lipschitz condition (2.2) ensures the existence and uniqueness of the solution to the equation (2.1).*

Remark 2.1.2 *The growth condition (2.3) prevents the solution from exploding, and if we do not have this condition, the equation (2.1) will have a unique solution only up to the time of explosion.*

Proof. Uniqueness

Let $X = (X_t)_{t\in T}$ and $Y = (Y_t)_{t\in T}$ be two solutions (2.1) such that $X_0 = Y_0 = \xi$. By applying the inequality:

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

and by using the formulas for X_t and Y_t, we obtain:

$$X_t - Y_t \leq 2\left|\int_0^t b(s,X_s) - b(s,Y_s)\,ds\right|^2 + 2\left|\int_0^t \sigma(s,X_s) - \sigma(s,Y_s)\,dB_s\right|^2.$$

Taking the mathematical expectation, we obtain:

$$\mathbb{E}\left(|X_t - Y_t|^2\right) \leq 2\mathbb{E}\left(\left|\int_0^t b(s,X_s) - b(s,Y_s)\,ds\right|^2\right) + 2\mathbb{E}\left(\left|\int_0^t \sigma(s,X_s) - \sigma(s,Y_s)\,dB_s\right|^2\right).$$

By using the Cauchy-Schwarz and Hölder-Davis-Gundy inequalities, we obtain:

$$\left\langle \int_0 g(s)\,dB_s \right\rangle_T = \int_0^t g^2(s)\,ds.$$

$$\mathbb{E}\left(|X_t - Y_t|^2\right) \leq 2T\mathbb{E}\int_0^T |b(s,X_s) - b(s,Y_s)|^2\,ds + 2\mathbb{E}\int_0^T |\sigma(s,X_s) - \sigma(s,Y_s)|^2\,ds.$$

By applying the Lipschitz condition, we obtain:

$$\mathbb{E}\left(|X_t - Y_t|^2\right) \leq c\int_0^T \mathbb{E}\,|X_s - Y_s|^2\,ds.$$

Where $c = \max(2Tk; 2k)$. By applying the Chebyshev inequality, we obtain:

$$\forall\, \varepsilon > 0; \left(P\,|X_t - Y_t|^2 > \varepsilon\right) \leq \frac{\mathbb{E}\left(|X_t - Y_t|^2\right)}{\varepsilon} \to 0.$$

Therefore, for any countable set D that is everywhere dense in $[0;T]$, we have:

$$P\left(\sup_{t\in[0;T]}|X_t - Y_t|^2 > 0\right) = 0.$$

Finally, the processes X and Y are continuous. We conclude that

$$P\left(\sup_{t\in[0;T]}|X_t - Y_t|^2 > 0\right) = 0.$$

This proves the strong uniqueness of the solution.

Existence:

We demonstrate the existence of a strong solution using the method of successive approximations and for this we set

$$X_t^n = \xi + \int_0^t b\left(s, X_s^{n-1}\right) ds + \int_0^t \sigma\left(s, X_s^{n-1}\right) dB_s. \tag{2.4}$$

We set:

$$X_t^{n+1} - X_t^n = \int_0^t b\left(s, X_s^n\right) - b\left(s, X_s^{n-1}\right) ds + \int_0^t \sigma\left(s, X_s^n\right) - \sigma\left(s, X_s^{n-1}\right) dB_s.$$

Using the same technique as for uniqueness, we obtain

$$\mathbb{E}\left(\left|X_s^{n+1} - X_s^n\right|^2\right) \leq c\int_0^t \mathbb{E}\left|X_s^n - X_s^{n-1}\right|^2 ds.$$

Where $c = \max(2Tk; 2k)$. According to (2.2), we have

$$\mathbb{E}\left(\left|X_s^1 - X_s^0\right|^2\right) \leq 2cT\left(1 + \mathbb{E}\left|X_s^0\right|^2\right),$$

since

$$\begin{aligned}
\mathbb{E}\left(\left|X_s^1 - X_s^0\right|^2\right) &\leq 2\mathbb{E}\left(\left|\int_0^t b\left(s, X_s^0\right) ds\right|^2\right) + 2\mathbb{E}\left(\left|\int_0^t \sigma\left(s, X_s^0\right) dB_s\right|^2\right) \\
&\leq 2T\mathbb{E}\int_0^t \left|b\left(s, X_s^0\right)\right|^2 ds + 2\mathbb{E}\int_0^t \left|\sigma\left(s, X_s^0\right)\right|^2 ds \\
&\leq 2Tk\int_0^t \left(1 + \mathbb{E}\left|X_s^0\right|^2\right) ds + 2k\int_0^t \left(1 + \mathbb{E}\left|X_s^0\right|^2\right) ds \\
&\leq (2Tk + 2k)\left(1 + \mathbb{E}\left|X_s^0\right|^2\right) T \\
&\leq 2cT\left(1 + \mathbb{E}\left|X_s^0\right|^2\right).
\end{aligned}$$

So:

$$\mathbb{E}\left(\left|X_s^1 - X_s^0\right|^2\right) \leq MT,$$

where $M = 2c\left(1 + \mathbb{E}\left|X_s^0\right|^2\right)$. By recurrence on n, it follows that

$$\mathbb{E}\left(\left|X_s^{n+1} - X_s^n\right|^2\right) \leq \frac{(MT)^{n+1}}{(n+1)!},$$

and we demonstrate that

$$\begin{aligned}
\mathbb{E}\left(\left|X_t^{n+2} - X_t^{n+1}\right|^2\right) &\leq \frac{(MT)^{n+2}}{(n+2)!} \\
&\leq c\int_0^t \mathbb{E}\left|X_s^n - X_s^{n-1}\right|^2 ds \\
&\leq c\int_0^t \frac{(Ms)^{n+1}}{(n+1)!} ds = c\frac{(MT)^{n+2}}{(n+2)!},
\end{aligned}$$

then we obtain,

$$\begin{aligned}\mathbb{E}\left(|X_t^m - X_t^n|^2\right)^{\frac{1}{2}} &= \|X_t^m - X_t^n\|_{L^2(\Omega)} \\ &= \sum_{k=n}^{m-1} \left\|X_t^{k+1} - X_t^k\right\|_{L^2(\Omega)} \\ &\leq \sum_{k=n}^{\infty} \left(\frac{(MT)^{k+1}}{(k+1)}\right)^{\frac{1}{2}} \to 0 \ \ (n \to \infty),\end{aligned}$$

When $n \to \infty$, we obtain $\mathbb{E}\left(|X_t^m - X_t^n|^2\right)^{\frac{1}{2}} \to 0$. Therefore, we have

$$\lim_{n\to+\infty} \mathbb{E}\left(|X_t^m - X_t^n|^2\right)^{\frac{1}{2}} = 0. \tag{2.5}$$

So X_t^n is a Cauchy sequence in $\mathbb{L}(\Omega)$ and therefore convergent. Let X_t be its limit, that is:

$$\lim_{n\to+\infty} X_t^n \to X_t.$$

Now, the process X_t^n defined by:

$$X_t^n = \xi + \int_0^t b\left(s, X_s^{n-1}\right) ds + \int_0^t \sigma\left(s, X_s^{n-1}\right) dB_s.$$

According to inequality (2.5) and Fatou's lemma, we obtain

$$\mathbb{E}\int_0^T |X_t^m - X_t^n|^2 \, ds \leq \lim_{m\to\infty} \sup \mathbb{E}\int_0^T |X_t^m - X_t^n|^2 \, ds \to 0.$$

$$\mathbb{E}\left(\lim_n \int_0^T |X_t^m - X_t^n|^2 \, ds\right) \leq \lim_{m\to\infty} \sup \mathbb{E}\left(\int_0^T |X_t^m - X_t^n|^2 \, ds\right) \to 0.$$

Therefore,

$$\mathbb{E}\left(\int_0^T |X_t - X_t^n|^2 \, ds\right) \to 0.$$

Using Itô's isometry,

$$\mathbb{E}\int_0^t \left|\sigma\left(s,X_s^n\right)-\sigma\left(s,X_s^{n-1}\right)dB_s\right|^2 \leq C\int_0^t \mathbb{E}\left(\left|X_s^n-X_s\right|^2\right)ds \to 0,$$

and furthermore,

$$\int_0^t \sigma\left(s,X_s^n\right)dB_s \to \int_0^t \sigma\left(s,X_s\right)dB_s.$$

We apply Hölder's inequality,

$$\mathbb{E}\int_0^t \left|b\left(s,X_s^n\right)-b\left(s,X_s\right)\right|^2 ds \leq cT\int_0^t \mathbb{E}\left(\left|X_s^n-X_s\right|^2\right)ds \to 0.$$

And by the continuity of $b\left(t,\cdot\right)$, we obtain

$$\int_0^t b\left(s,X_s^n\right)ds \to \int_0^t b\left(s,X_s\right)ds;\ \ \left(n\to\infty\right).$$

Passing to the limit in (2.4), we have

$$X_t = \xi + \int_0^t b\left(s,X_s\right)ds + \int_0^t \sigma\left(s,X_s\right)dB_s.$$

Thus, X_t is a solution of equation (2.1). Let's show that

$$\mathbb{E}\left(\sup_{t\in[0,T]}\left|X_t\right|^p\right) < M\ \forall p>1.$$

Using the inequality $(a+b+c)^2 \leq 3a^2+3b^2+3c^2$ and taking expectations, we have:

$$\mathbb{E}\left(\left|X_t\right|^2\right) \leq 3\mathbb{E}\left(\left|\xi\right|^2\right) + 3T\mathbb{E}\int_0^t \left|b\left(s,X_s\right)\right|^2 ds + 3\mathbb{E}\int_0^t \left|\sigma\left(s,X_s\right)\right|^2 ds,$$

According to (2.3), we have:

$$\mathbb{E}\left(\left|X_t\right|^2\right) \leq 3\mathbb{E}\left(\left|\xi\right|^2\right) + 3Tk\int_0^t \left(1+\mathbb{E}\left|X_s\right|^2\right)ds + 3k\int_0^t \left(1+\mathbb{E}\left|X_s\right|^2\right)ds.$$

Let $M = \max(3; 3Tk; 3k)$ and $c = \max(M; 2M)$, then we obtain

$$\mathbb{E}\left(|X_t|^2\right) \leq c\left(1 + \mathbb{E}|\xi|^2\right) + c\int_0^t \mathbb{E}\left(|X_s|^2\right) ds.$$

Applying the Grnwall lemma, we obtain:

$$\mathbb{E}\left(|X_t|^2\right) \leq c\left(1 + \mathbb{E}|\xi|^2\right)\exp(ct); \forall t \in T.$$

Since $\mathbb{E}|\xi|^2 < \infty$, we can set $M = c\left(1 + \mathbb{E}|\xi|^2\right)\exp(ct)$. This gives us

$$\mathbb{E}\left(|X_t|^2\right) \leq M; \qquad \forall t \in T.$$

This implies that

$$\mathbb{E}\left(\sup_{t\in[0,T]} |X_t|^p\right) < M.$$

Hence the result. ■

2.2 Backward stochastic differential equation (BSDE)

A backward stochastic differential equation (BSDE) is a mathematical equation that involves a process evolving backward in time. It is a generalization of stochastic differential equation (SDE) and plays a significant role in stochastic analysis, mathematical finance, and control theory. In a standard SDE, the evolution of a process is determined by its past and current values. In contrast, a BSDE involves the process evolving backward from a future time to the present. This backward evolution makes BSDE particularly useful for solving problems related to optimization, hedging, and risk management.

The general form of a BSDE is given by:

$$\begin{cases} -dY_t = f(t, Y_t, Z_t)dt - Z_t dB_t, 0 \leq t \leq T \\ Y_T = \xi, \end{cases} \tag{2.6}$$

where $Y(t)$ is the unknown process, $Z(t)$ is the adapted process called the control or stochastic process, $f(t, Y(t), Z(t))$ is a known function, and $W(t)$ is a standard Wiener process or Brownian motion.

A solution to the EDSR (2.6) is a pair of processes $\{(Y_t, Z_t)\}_{0 \leq t \leq T}$ that satisfies the following conditions:

- Y and Z are progressively measurable processes taking values in $\mathbb{R}^k$ and $\mathbb{R}^{k \times d}$, respectively.
- $P-p.s$
$$\left\{ \int_0^T |f(s, Y_s, Z_s)| ds - \int_0^T |Z_s|^2 ds \right\} < \infty;$$
- $P-p.s$, we have:
$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dB_s, 0 \leq t \leq T.$$

2.2.1 Existence and uniqueness result for BSDE with Lipschitz coefficients

Theorem 2.2.1 *Given the preceding data $\xi \in L^2(\mathcal{F}_T^B)$, we further assume that:*

- *f is k-uniformly Lipschitz with respect to Y and Z, that is:*
- $\exists k > 0, \forall (Y, Z) \in \mathbb{R}^k \times \mathbb{R}^{k \times d}$, $\forall (Y', Z') \in \mathbb{R}^k \times \mathbb{R}^{k \times d}$:
$$|f(t, B, Y, Z) - f(t, B, Y', Z')| \leq k(|Y - Y'| + |Z - Z'|).$$
- $|f(t, B, Y, Z)| \leq h_t(B) + \lambda(|Y| + |Z|)$,

- *Where*$h_t(\omega) \geq 0, \forall t, B$ *and* $\mathbb{E}(\int_0^T h_s^2 ds) < \infty$.
- $$\mathbb{E}\left[|\xi| + |\int_0^T f(t, B, 0, 0)|^2 ds\right] < \infty.$$

Then, the BDSR (f, ξ) *has a solution.*

Proof. The proof of the existence and uniqueness of solutions to backward stochastic differential equations (BSDEs) relies on two key theorems, as well as the Lipschitz condition. The first theorem is the fixed-point theorem, which provides a framework for establishing the existence and uniqueness of solutions. The second theorem used is the representation theorem for Brownian martingales. For more detailed insights and a rigorous understanding of the proof, you can see ([4]) . ■

Chapter §.3
Stochastic maximum principle

Chapter 3

Stochastic maximum principle

In this chapter, we are interested in the following control problem, on a given filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$

$$\begin{cases} dx^u(t) = b(t,\ x^u(t),\ P_{x^u(t)},\ u(t))dt + \sigma(t, x^u(t), P_{x^u(t)}, u(t))dB(t), \\ dy^u(t) = -g(t,\ x^u(t),\ P_{x^u(t)}, y^u(t), P_{y^u(t)}, z^u(t), P_{z^u(t)}, u(t))dt + z^u(t)dB(t), \\ x^u(0) = x_0,\ y^u(T) = h(x^u(T),\ P_{x^u(T)}). \end{cases} \tag{3.1}$$

In this context, let $B(\cdot)$ denote a d-dimensional $\mathcal{F}_t$-Brownian motion. Here, $P_{X(t)} = P \circ X^{-1}$ represents the probability measure of the random variable X. The control $u(\cdot) = (u(t))_{t\geq 0}$ is required to take values in a subset of $\mathbb{R}$ and be adapted to a subfiltration $(\mathcal{G}_t)_{t\geq 0}$ of $(\mathcal{F}_t)$. The maps

$$b : [0,T] \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{A} \to \mathbb{R},$$
$$\sigma : [0,T] \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{A} \to \mathbb{R},$$
$$g : [0,T] \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{A} \to l^2(\mathbb{R}),$$

and

$$h : \mathbb{R} \times Q_2(\mathbb{R}) \to \mathbb{R},$$

are given deterministic functions. $Q_2(\mathbb{R})$ represents the space of all probability measures

μ on $(\mathbb{R} \times \beta(\mathbb{R}))$ that have finite second moments, i.e., satisfying $\int_{\mathbb{R}} |x|^2 \mu(dx) < \infty$. This space is equipped with the 2-Wasserstein metric defined as follows: for any $\mu_1, \mu_2 \in Q_2(\mathbb{R})$,

$$D_2(\mu_1, \mu_2) = \inf \left\{ \left(\int_{\mathbb{R}} |x-y|^2 \rho(dx, dy) \right)^{\frac{1}{2}}, \rho \in Q_2(\mathbb{R}), \rho(\cdot, \mathbb{R}) = \mu_1, \rho(\mathbb{R}, \cdot) = \mu_2 \right\}.$$

The McKean-Vlasov forward-backward stochastic differential equation (3.1) is highly general, as the coefficients depend on the probability law of the solution $P_{x^u(t)}, P_{y^u(t)}, P_{z^u(t)}$ in a genuinely nonlinear manner within the space of probability measures.

The expected cost to be minimized over the class of admissible controls is also of McKean-Vlasov type, which has the form

$$\begin{aligned} J(u(\cdot)) = \mathbb{E}\Big\{ \int_0^T & f(t, x^u(t), P_{x^u(t)}, y^u(t), P_{y^u(t)}, z^u(t), P_{z^u(t)}, u(t))dt \\ & + \phi(x^u(T), P_{x^u(T)}) + \varphi(y^u(0), P_{y^u(0)}) \Big\}, \end{aligned} \tag{3.2}$$

where $f : [0,T] \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{R} \times Q_2(\mathbb{R}) \times \mathbb{A} \to \mathbb{R}$, $\phi : \mathbb{R} \times Q_2(\mathbb{R}) \to \mathbb{R}$, $\varphi : \mathbb{R} \times Q_2(\mathbb{R}) \to \mathbb{R}$ are deterministic functions. We remark that the cost functional (3.2) involves the law of the solution in a nonlinear way.

3.1 Notations

Let T is a fixed terminal time and $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \geq 0})$ be a fixed filtered probability space equipped with a P-completed right continuous filtration on which a d-dimensional Brownian motion $B(\cdot) = (B(t))_t$ is defined.

We assume that $\mathcal{F}_t$ is P-augmentation of the natural filtration $\left(\mathcal{F}_t^{(B)}\right)_{t \in [0,T]}$ defined as follows:

$$\mathcal{F}_t^{(B)} \triangleq \sigma \{ B(s) : 0 \leq s \leq t \},$$

$\mathcal{F}_0$ denotes the totality of P-null sets, and $\mathcal{F}_1$ denotes the σ-field generated by $\mathcal{F}_1$, let $\mathcal{G}_t$

be a subfiltration of $\mathcal{F}_t : t \in [0, T]$.

An admissible control is defined as a function $u(\cdot) : [0, T] \times \Omega \to \mathbb{A}$, which is G_t-pedictable, $E \int_0^T |u(s)|^2 \, ds < \infty$, such that the equation (3.1) has a unique solution. An admissible control $u^*(\cdot) \in A_{\mathcal{G}}([0, T])$ is called optimal if

$$J(u^*(\cdot)) \triangleq \inf_{u(.) \in \mathcal{A}_{\mathcal{G}}([0,T])} J(u(\cdot)). \tag{3.3}$$

Now, we introduce the fundamental notations.

- $\mathbb{R}$ is the dimensional Euclidean space.
- $\mathbb{L}^2(\Omega, \mathcal{F}, P, \mathbb{R})$ is the Banach space of $\mathbb{R}$-valued, square integrable random variables on $(\Omega, \mathcal{F}, P)$.
- l^2 : denotes the hilbert space of real-valued sequences $x = (x_n)_{n \geq 0}$ such that

$$\|x\| \triangleq \left[\sum_{n=1}^{\infty} x_n \right]^2 < \infty.$$

- $l^2(\mathbb{R})$ is the space of $\mathbb{R}$-valued $(g_n)_{n \geq 1}$ such that

$$\|g\|_{l^2(\mathbb{R})} \triangleq \left[\sum_{n=1}^{\infty} \|g_n\|_{\mathbb{R}}^2 \right]^{\frac{1}{2}} < \infty.$$

- $\mathbb{L}^2_{\mathcal{F}}([0, T]; \mathbb{R})$ the Banach space of $\mathcal{F}_t$-predictable processes g such that

$$\|g\|_{\mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R})} \triangleq \mathbb{E} \left(\int_0^T \sum_{n=1}^{\infty} \|g\|_{\mathbb{R}}^2 \, dt \right)^{\frac{1}{2}} < \infty,$$

  where $g = \{g_n(t, w) : (t, w) \in [0, T] \times \Omega, n = 1, ..., \infty\}$.

- $\mathbb{M}^2_{\mathcal{F}}([0,T];\mathbb{R})$ denotes the space of all $\mathbb{R}$-valued and $\mathcal{F}_t$-adapted processes such that

$$\|g\|_{\mathbb{M}^2_{\mathcal{F}}([0,T]\times\Omega)} \triangleq \mathbb{E}\left(\int_0^T \|g(t)\|^2_{\mathbb{R}}\, dt\right)^{\frac{1}{2}} < \infty,$$

  where $g = \{g(t,w) : (t,w) \in [0,T]\times\Omega\}$.

- $\mathcal{M}^{n\times m}(\mathbb{R})$ denotes the space of $n\times m$ real matrices.

- g is a differentiable function, we denote by $g_x(t) = \frac{\partial}{\partial x} g(t, X, P_X, u)$ its gradient with respect to the variable x.

- $\mathbb{E}^{\mathcal{G}_t}[X]$ the conditional expectation of X with respect to $\mathcal{G}_t$, $\mathbb{E}^{\mathcal{G}_t}(X) = \mathbb{E}(X\,|\mathcal{G}_t)$.

3.2 Lions differentiability

The differentiability with respect to probability measures is a powerful tool used in various areas of mathematics, including probability theory and mathematical statistics. Its use in proving the main result of a study is notable, and it was initially introduced by Lions. The fundamental idea of this technique is to identify a distribution $\mu \in Q_2(\mathbb{R})$ with a random variable $X \in \mathbb{L}^2(\mathcal{F},\mathbb{R})$, such that $\mu = P_X$. To apply the method, it is assumed that the probability space $(\Omega,\mathcal{F},P)$ provides sufficient richness so that for every $\mu \in Q_2(\mathbb{R})$, there exists a random variable $X \in \mathbb{L}^2(\mathcal{F},\mathbb{R})$ such that $\mu = P_X$. A sub-σ field $\mathcal{F}_0 \subset \mathcal{F}$, is required to meet the richness assumption and must be such that the Brownian motion $B(\cdot)$ is independent of $\mathcal{F}_0$. Furthermore, $\mathcal{F}_0$ must be rich enough, i.e.,

$$Q_2(\mathbb{R}) \triangleq \{P_X : X \in \mathbb{L}^2(\mathcal{F}_0,\mathbb{R})\}.$$

Here, $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$ denotes the filtration generated by $B(\cdot)$, completed and augmented by $\mathcal{F}_0$. Subsequently, for any function $g : Q_2(\mathbb{R}) \to \mathbb{R}$, we define a function $\widetilde{g} : \mathbb{L}^2(\mathcal{F},\mathbb{R}) \to$

$\mathbb{R}$ as follows:

$$\widetilde{g}\,(X) \triangleq g(P_X),\ X \in \mathbb{L}^2\,(\mathcal{F},\mathbb{R})\,. \tag{3.4}$$

It is noteworthy that the function $\widetilde{g}$, known as the lift of g, solely depends on the law of $X \in \mathbb{L}^2\,(\mathcal{F},\mathbb{R})$ and remains independent of the choice of representative X (refer to the work of Buckdahn et al. ([3])).

Definition 3.2.1 *A function $g : Q_2(\mathbb{R}) \to \mathbb{R}$ is said to be differentiable at a distribution $\mu_0 \in Q_2\,(\mathbb{R})$ if there exists $X_0 \in \mathbb{L}^2\,(\mathcal{F},\mathbb{R})$ whith $\mu_0 = P_{X_0}$ such that its lift $\widetilde{g}$ is Fréchet differentiable at X_0. More precisely, there exists a continuous linear functional $D\widetilde{g}(X_0) : \mathbb{L}^2\,(\mathcal{F},\mathbb{R}) \to \mathbb{R}$ such that*

$$\widetilde{g}\,(X_0+\xi) - \widetilde{g}(X_0) \triangleq \langle D\widetilde{g}(X_0)\cdot\xi\rangle + o\left(\|\xi\|_2\right) = D_\xi g(\mu_0) + o\left(\|\xi\|_2\right), \tag{3.5}$$

where $\langle\cdot,\cdot\rangle$ is the dual product on $\mathbb{L}^2\,(\mathcal{F},\mathbb{R})$. We called $D_\xi g(\mu_0)$ the Fréchet derivative of g at μ_0 in the direction ξ. In this case, we have

$$D_\xi g(\mu_0) = \langle D\widetilde{g}(X_0)\cdot\xi\rangle = \frac{d}{dt}\widetilde{g}\,(X_0+t\xi)\Big|_{t=0} \text{ with } \mu_0 = P_{X_0}. \tag{3.6}$$

By applying Riesz Representation theorem, there is a unique random variable $\Theta_0 \in \mathbb{L}^2\,(\mathcal{F},\mathbb{R})$ such that $\langle D\widetilde{g}(X_0)\cdot\xi\rangle = (\Theta_0\cdot\xi)_2 = \mathbb{E}\left[(\Theta_0\cdot\xi)_2\right]$ where $\xi \in \mathbb{L}^2\,(\mathcal{F},\mathbb{R})$ It was shown (see the work of Buckdahn et al. ([3])) that there exists a Boral function $\Phi\,[\mu_0]\,(\cdot) : \mathbb{R} \to \mathbb{R}$, depending only on the law $\mu_0 = P_{X_0}$ but not on the particular choice of the representative X_0 such that

$$\Theta_0 = \Phi\,[\mu_0]\,(X_0)\,. \tag{3.7}$$

Thus, we can write (3.5) as

$$g(P_X) - g(P_{X_0}) = (\Phi\,[\mu_0]\,(X_0)\cdot X - X_0)_2 + o\left(\|X - X_0\|_2\right),\ \forall X \in \mathbb{L}^2\,(\mathcal{F},\mathbb{R})\big)\,.$$

We denote

$$\partial_\mu g(P_{X_0}, x) = \Phi\left[\mu_0\right](x), x \in \mathbb{R}.$$

Moreover,stop we have the following identities:

$$D\widetilde{g}(X_0) = \Theta_0 = \Phi\left[\mu_0\right](X_0) = \partial_\mu g(P_{X_0}, X_0),$$

$$D_\xi g(P_{X_0}) = \left\langle \partial_\mu g(P_{X_0}, X_0) \cdot \xi \right\rangle, \tag{3.8}$$

where $\xi = X - X_0$.

For each $\mu \in Q_2(\mathbb{R})$, $\partial_\mu g(P_X, \cdot) = \Phi\left[P_X\right](\cdot)$ is only defined in a $P_X(dx) - a.e.$sense where $\mu = P_X$.

Among the different notions of differentiability of a function g defined over $Q_2(\mathbb{R})$ we apply for our control problem that introduced by Lions in his lectures at Collège de France in Paris and revised in the notes by Cardaliaguet, ([5]) we refer the reader to the work of Buckdahn et al. ([3]) .

Definition 3.2.2 (*Space of differentiable functions in* $Q_2(\mathbb{R})$).

We say that the function $g \in \mathbb{C}_b^{1,1}(Q_2(\mathbb{R}))$ if for all $X \in \mathbb{L}^2(\mathcal{F}, \mathbb{R})$ there exists a P_X-modification of $\partial_\mu g(P_X, \cdot)$ (denoted by $\partial_\mu\ g$) such that $\partial_\mu g : Q_2(\mathbb{R}) \times \mathbb{R} \rightarrow \mathbb{R}$ is bounded and Lipchitz continuous. That is, for some $C > 0$, it holds that

(i) $|\partial_\mu g(\mu, x)| \leq C, \forall \mu \in Q_2(\mathbb{R}), \forall x \in \mathbb{R};$

(ii) $|\partial_\mu g(\mu, x) - \partial_\mu g(\mu', x')| \leq C\left[\mathbb{D}_2(\mu, \mu') + |x - x'|\right], \forall \mu, \mu' \in Q_2(\mathbb{R}), \forall x, x' \in \mathbb{R}.$

Assumptions

We shall make use of the following standing assumptions on the coefficients.

Assumption (A1)

(i) For any $t \in [0,T]$, ths functions b, σ are continuously differentiable in (x,μ) and g, f are continuously differentiable in (x,y,z,u), bounded by $C(1+|x|+|y|+|z|+|u|)$.

The function ϕ, h is continuously differentiable in x, and the function φ is continuously differentiable in y.

(ii) The derivatives $b_x, b_u, \sigma_x, \sigma_u$ are bounded.The derivatives of f with respect to (x,y,z,u) are bounded by $C(1+|x|^2+|y|^2+|z|^2+|u|^2)$. The derivatives ϕ_x bounded by $C(1+|x|^2)$, and φ_y is dominated by $C(1+|y|^2)$. The terminal value $y(T) \in l^2_{\mathcal{F}}([0,T];\mathbb{R})$.

(iii) For all $t \in [0,T]$, $b(\cdot,0,0,0) \in \mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R})$, $g(\cdot,0,0,0,0,0,0,0) \in \mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R})$,

$\sigma(\cdot,0,0,0) \in \mathbb{M}^2_{\mathcal{F}}([0,T];\mathbb{R})$.

Assumption (A2)

The functions $b,\sigma,g,f,h,\phi,\varphi \in \mathbb{C}^{1,1}_b(Q_2(\mathbb{R}))$, and the derivatives $\partial_\mu^{P_x} b, \partial_\mu^{P_x}\sigma$, $(\partial_\mu^{P_x}, \partial_\mu^{P_y}, \partial_\mu^{P_z})(g,f), \partial_\mu^{P_x} h, \partial_\mu^{P_x}\phi, \partial_\mu^{P_y}\varphi$ are bounded and Lipchitz continuous, such that, for some $C>0$, it holds that

(i) $\left|\partial_\mu^{P_x}\rho(t,x,\mu)\right| \le C$, and $\left|\partial_\mu^{P_x}\rho(t,x,\mu) - \partial_\mu^{P_x}\rho(t,x',\mu')\right| \le C(\mathbb{D}_2(\mu,\mu') + |x-x'|)$,

$\forall \mu,\mu' \in Q_2(\mathbb{R}), \forall x,x' \in \mathbb{R}$ for $\rho = b,\sigma$;

(ii) $\left|\partial_\mu^{P_x}\rho(x,\mu)\right| \le C, \forall \mu \in Q_2(\mathbb{R}), \forall x \in \mathbb{R}$, and $\left|\partial_\mu^{P_x}\rho(x,\mu) - \partial_\mu^{P_x}\rho(x',\mu')\right| \le C\,(D_2(\mu,\mu') + |x-x'|), \forall \mu,\mu' \in Q_2(\mathbb{R}^d), \forall x,x' \in \mathbb{R}$, for $\rho = \Phi, h$;

(iii) $\left|\left(\partial_\mu^{P_x}, \partial_\mu^{P_y}, \partial_\mu^{P_z}\right)\rho(t,x,\mu_1,y,\mu_2,z,\mu_3)\right| \le C$,and

$$\begin{aligned}&\left|\left(\partial_\mu^{P_x}, \partial_\mu^{P_y}, \partial_\mu^{P_z}\right)\rho(t,x,\mu_1,y,\mu_2,z,\mu_3) - \left(\partial_\mu^{P_x}, \partial_\mu^{P_y}, \partial_\mu^{P_z}\right)\rho(t,x',\mu_1',y',\mu_2',z',\mu_3')\right| \\ &\le C(|x-x'|+|y-y'|+|z-z'|+\mathbb{D}_2(\mu_1,\mu_1')+\mathbb{D}_2(\mu_2,\mu_2')+\mathbb{D}_2(\mu_3,\mu_3')),\end{aligned}$$

$\forall\mu_1,\mu_1'\in Q_2(\mathbb{R}),\forall\mu_2,\mu_2'\in Q_2(\mathbb{R})$, and$\forall\mu_3,\mu_3'\in Q_2(\mathbb{R})$. In addition, $\forall x,x'\in\mathbb{R},\forall y,y'\in\mathbb{R},\forall z,z'\in\mathbb{R}$, and, for $\rho=g,f$.

Clearly, under the above Assumptions $(\mathbf{A1})-(\mathbf{A2})$, for each $u(\cdot)\in A_{\mathcal{G}}([0,T])$, the McKean-Vlasov–type FBSDE (3.1) admits a unique strong solution $(x^u(\cdot),y^u(\cdot),z^u(\cdot))\in\mathbb{R}\times\mathbb{R}\times\mathcal{M}^{m\times d}(\mathbb{R})\times l^2(\mathbb{R})$ such that

$$
\begin{aligned}
x^u(t) &= x_0+\int_0^t b(s,x^u(s),P_{x^u(s)},u(s))ds+\int_0^t\sigma(s,x^u(s),P_{x^u(s)},u(s))dB(s),\\
y^u(t) &= y^u(T)-\int_t^T g(s,x^u(s),P_{x^u(s)},y^u(s),P_{y^u(s)},z^u(s),P_{z^u}(s),u(s))ds\\
&\quad+\int_t^T z^u(s)dB(s).
\end{aligned}
$$

Let $\left(\widehat{\Omega},\mathcal{F},\widehat{P}\right)$ represent a copy of the probability space $(\Omega,\mathcal{F},P)$. For any pair of random variables $(Z,\xi)\in\mathbb{L}^2(\mathcal{F},\mathbb{R})\times\mathbb{L}^2(\mathcal{F},\mathbb{R})$, we define $\left(\widehat{Z},\widehat{\xi}\right)$ as an independent copy of (Z,ξ) on $\left(\widehat{\Omega},\mathcal{F},\widehat{P}\right)$. We consider the product probability space $\left(\Omega\times\widehat{\Omega},\mathcal{F}\otimes\mathcal{F},P\otimes\widehat{P}\right)$, and we define $(\widehat{Z},\widehat{\xi})(w,\widehat{w})=(Z(\widehat{w}),\xi(\widehat{w}))$ for any $(w,\widehat{w})\in\Omega\times\widehat{\Omega}$. Similarly, we let $(\widehat{u}^*(t),\widehat{x}^*(t),\widehat{y}^*(t),\widehat{z}^*(t))$ be an independent copy of $(u^*(t),x^*(t),y^*(t),z^*(t))$, ensuring that $P_{x^*(t)}=\widehat{P}_{\widehat{x}^*(t)}$, $P_{y^*(t)}=\widehat{P}_{\widehat{y}^*(t)}$, $P_{z^*(t)}=\widehat{P}_{\hat{z}^*(t)}$. Expectations under the probability measure P are denoted by $\widehat{\mathbb{E}}[\cdot]$. Let us represent by

$$
\begin{aligned}
\widehat{b}_\mu^{P_x}(t) &= \partial_\mu^{P_x}b(t,x^*(t),P_{x^*(t)},u^*(t);\widehat{x}^*(t)),\\
\widehat{b}_\mu^{*}(t) &= \partial_\mu^{P_x}b(t,\hat{x}^*(t),P_{x^*(t)},\hat{u}^*(t);x^*(t)),\\
\widehat{\sigma}_\mu^{P_x}(t) &= \partial_\mu^{P_x}\sigma(t,x^*(t),P_{x^*(t)},u^*(t);\widehat{x}^*(t)),\\
\widehat{\sigma}_\mu^{*}(t) &= \partial_\mu^{P_x}\sigma(t,\hat{x}^*(t),P_{x^*(t)},\hat{u}^*(t);x^*(t)),
\end{aligned}
$$

for g, f and $\xi = x, y, z$

$$\begin{aligned}
\widehat{g}_\mu^{P_\xi}(t) &= \partial_\mu^{P_\xi} g(t, x^*(t), P_{x^*(t)}, y^*(t), P_{y^*(t)}, z^*(t), P_{z^*(t)}, u^*(t); \widehat{\xi}^*(t)),\\
\widehat{g}_\mu^{P_\xi,*}(t) &= \partial_\mu^{P_\xi} g(t, \hat{x}^*(t), P_{x^*(t)}, \hat{y}^*(t), P_{y^*(t)}, \hat{z}^*(t), P_{z^*(t)}, u^*(t); \xi^*(t)),\\
\widehat{f}_\mu^{P_\xi}(t) &= \partial_\mu^{P_\xi} f(t, x^*(t), P_{x^*(t)}, y^*(t), P_{y^*(t)}, z^*(t), P_{z^*(t)}, u^*(t); \widehat{\xi}^*(t)),\\
\widehat{f}_\mu^{P_\xi,*}(t) &= \partial_\mu^{P_\xi} f(t, \hat{x}^*(t), P_{x^*(t)}, \hat{y}^*(t), P_{y^*(t)}, \hat{z}^*(t), P_{z^*(t)}, u^*(t); \xi^*(t)),
\end{aligned}$$

and

$$\begin{aligned}
\widehat{h}_\mu(t) &= \partial_\mu^{P_x} h\left(x^*(T), P_{x^*(T)}; \widehat{x}^*(T)\right),\\
\widehat{h}_\mu^*(t) &= \partial_\mu^{P_x} h\left(\hat{x}^*(T), P_{x^*(T)}; x^*(T)\right),\\
\widehat{\phi}_\mu(t) &= \partial_\mu^{P_x} A\left(x^*(T), P_{x^*(T)}; \widehat{x}^*(T)\right),\\
\widehat{\phi}_\mu^*(t) &= \partial_\mu^{P_x} A\left(\hat{x}^*(T), P_{x^*(T)}; x^*(T)\right).
\end{aligned}$$

We denote

$$\begin{aligned}
\widehat{\varphi}_\mu(t) &= \partial_\mu^{P_y} \varphi(y^*(0), P_{y^*(0)}; \widehat{y}^*(0)),\\
\widehat{\varphi}_\mu^*(T) &= \partial_\mu^{P_y} \varphi(\hat{y}^*(0), P_{y^*(0)}; y^*(0)).
\end{aligned}$$

The adjoint equations associated with the stochastic maximum principle for the control problem (3.1) − (3.2) defined by:

$$\begin{cases}
d\Phi(t) = -\{b_x(t)\Phi(t) + \widehat{\mathbb{E}}\left[\partial_\mu^{P_x} b(t)\Phi(t)\right] + \sigma_x(t)Q(t) + \widehat{\mathbb{E}}\left[\partial_\mu^{P_x}\sigma(t)Q(t)\right]\\
\quad -g_x(t)K(t) - \widehat{\mathbb{E}}\left[\partial_\mu^{P_x} g(t)K(t)\right] + f_x(t) + \widehat{\mathbb{E}}\left[\partial_\mu^{P_x} f(t)\right]\}dt + Q(t)dB(t),\\
\quad \Phi(T) = -\left[h_x\left(x(T), P_{x(T)}\right)\mathcal{K}(T) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x} h(x(T), P_{x(T)})\,\mathcal{K}(T)\right)\right]\\
\qquad\qquad + \Phi_x(x(T), P_{x(T)}) + \widehat{\mathbb{E}}[\partial_\mu^{P_x}\Phi(x(T), P_{x(T)})].\\
d\mathcal{K}(t) = \left[g_y(t)\mathcal{K}(t) + \widehat{\mathbb{E}}\left[\partial_\mu^{P_y} g(t)\mathcal{K}(t)\right] - f_y(t) - \widehat{\mathbb{E}}(\partial_\mu^{P_y} f(t))\right]dt\\
\qquad + \left[g_z(t)\mathcal{K}(t) + \widehat{\mathbb{E}}\left[\partial_\mu^{P_z} g(t)\mathcal{K}(t)\right] - f_z(t) - \widehat{\mathbb{E}}[\partial_\mu^{P_z} f(t)]\right]dB(t),\\
\mathcal{K}(0) = -\left[\varphi_y\left(y(0), P_{y(0)}\right) + \widehat{\mathbb{E}}(\partial_\mu^{P_y}\varphi\left(y(0), P_{y(0)}\right)\right].
\end{cases} \tag{3.9}$$

The mean-field structure of equation (3.9) is apparent as it arises from the terms involving Fréchet. However, if the coefficients do not explicitly depend on the law of the solution, these terms can be simplified, reducing the equation to a standard Backward Stochastic Differential Equation (BSDE). The well-posedness of the backward stochastic equation in (3.9) is established by theorem 3.1 in the work of Buckdahn et al. ([1]). To be more precise, assuming Assumptions (**A1**) and (**A2**), the Forward-Backward Stochastic Differential Equation (FBSDE) (3.9) admits a unique $\mathcal{F}$-adapted solution $(\Phi(\cdot), Q(\cdot), \mathcal{K}(\cdot))$.
The Hamiltonian function $\mathcal{H}$ associated with the McKean-Vlasov stochastic control problem (3.1) $-$ (3.2) is defined as follows:

$$\begin{aligned}
&\mathcal{H}(t, x, P_x, y, P_y, z, P_z, u, \Phi(.), Q(.), \mathcal{K}(.))\\
&= f(t, x, P_x, y, P_y, z, P_z, u) + \Phi(t)\, b(t, x, P_x, u) + Q(t)\sigma(t, x, P_x, u)\\
&-\mathcal{K}(t) g(t, x, P_x, y, P_y, z, P_z, u),
\end{aligned}$$

where $(\Phi(\cdot), Q(\cdot), \mathcal{K}(\cdot))$ solution of (3.9). conesquently, if we define

$$\mathcal{H}(t) = \mathcal{H}(t, x, P_x, y, P_y, z, P_z, u, \Phi(\cdot), u, Q(\cdot), \mathcal{K}(\cdot)),$$

then, the adjoint equation (3.9), we can express it as follows:

$$\begin{cases}
d\Phi(t) = -[\mathcal{H}_x(t) + \widehat{\mathbb{E}}\left[\partial_\mu^{P_x} \mathcal{H}(t)\right]]dt + Q(t)dB(t),\\
\Phi(T) = -[h_x\left(x(T), P_{x(T)}\right)\mathcal{K}(T) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x} h(x(T), P_{x(T)}\right)\mathcal{K}(T))]\\
\qquad +\phi_x(x(T), P_{x(T)}) + \widehat{\mathbb{E}}[\partial_\mu^{P_x}\phi(x(T), P_{x(T)})].\\
d\mathcal{K}(t) = -\left(\mathcal{H}_y(t) + \widehat{\mathbb{E}}\left[\partial_\mu^{P_y}\mathcal{H}(t)\right]\right)dt - (\mathcal{H}_z(t) + \widehat{\mathbb{E}}\left[\partial_\mu^{P_z}\mathcal{H}(t)\right])dB(t),\\
\mathcal{K}(0) = -\left[\varphi_y\left(y(0), P_{y(0)}\right) + \widehat{\mathbb{E}}(\partial_\mu^{P_y}\varphi\left(y(0), P_{y(0)}\right))\right].
\end{cases} \tag{3.10}$$

Given the well-known fact and assumptions (**A1**) and (**A2**), we can state that the adjoint equation (3.9)or (3.10) possesses a unique solution satisfying $(\Phi(\cdot), Q(\cdot), \mathcal{K}(\cdot)) \in \mathbb{L}^2_{\mathcal{F}}([0,T];\mathbb{R}) \times l^2_{\mathcal{F}}([0,T];\mathbb{R})$.

3.3 Necessary optimality conditions

In this section, we establish the necessary conditions for optimal control in the form of the Pontryagin maximum principle, using Lions derivatives. The control system is governed by McKean-Vlasov FBSDEs. In addition to the hypotheses $(\mathbf{A1}), (\mathbf{A2})$, we introduce the following assumptions:

Assumption $(\mathbf{A3})$

For any $u(\cdot), v(\cdot) \in \mathcal{A}_{\mathcal{G}}([0,T])$ with $v(\cdot)$ bounded, there exists a positive value $\delta > 0$ such that $u(\cdot) + \theta v(\cdot) \in \mathcal{A}_{\mathcal{G}}([0,T])$ for all $\theta \in (-\delta, \delta)$. Now, given a $u(\cdot), v(\cdot) \in \mathcal{A}_{\mathcal{G}}([0,T]$ with $v(\cdot)$ bounded, we define

$$
\begin{aligned}
X_1(t) &= X_1^{u^*,v}(t) := \tfrac{d}{d\theta}\left[x^{u^*+\theta v}(t)\right]\Big|_{\theta=0},\\
Y_1(t) &= Y_1^{u^*,v}(t) := \tfrac{d}{d\theta}\left[y^{u^*+\theta v}(t)\right]\Big|_{\theta=0},\\
Z_1(t) &= Z_1^{u,v}(t) := \tfrac{d}{d\theta}\left[z^{u^*+\theta v}(t)\right]\Big|_{\theta=0},
\end{aligned}
$$

where $\left(x^{u^*+\theta v}(\cdot), y^{u^*+\theta v}(\cdot), z^{u^*+\theta v}(\cdot)\right)$ represents the solution of Equation (3.1) corresponding to $(u^* + \theta v)$. It should be noted that the process $(X_1(\cdot), Y_1(\cdot), Z_1(\cdot))$ satisfies the following linear McKean-Vlasov FBSDE, known as the variational equations, which are governed by the Brownian motion $B(t)$:

$$
\begin{aligned}
dX_1(t) &= \left[b_x(t)\,X_1(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x} b(t)\widehat{X}_1(t)\right) + b_u(t)v(t)\right]dt \\
&\quad + \left[\sigma_x X_1(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x}\sigma(t)\widehat{X}_1(t)\right) + \sigma_u(t)v(t)\right]dB(t). \\
dY_1(t) &= \left[g_x(t)\,X_1(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x} g(t)\widehat{X}_1(t)\right) + g_y(t)\,Y_1(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_y} g(t)\widehat{Y}_1(t)\right)\right. \\
&\quad \left. + g_z(t)\,Y_1(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_z} g(t)\widehat{Z}_1(t)\right) + g_u(t)v(t)\right]dt + Z_1(t)dB(t), \\
Y_1(T) &= \left[h_x\left(x(T), P_{x(T)}\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x} h(x(T), P_{x(T)})\right)\right] X_1(T).
\end{aligned}
\tag{3.11}
$$

The main result of this section is stated in the following theorem.

Let $u^*(\cdot)$ be a local minimum for the cost functional J over the class of admissible controls

$\mathcal{A}_{\mathcal{G}}([0,T])$. Specifically, for all bounded $v(\cdot) \in \mathcal{A}_{\mathcal{G}}([0,T])$, there exists a positive value $\delta > 0$ such that $((u^*(\cdot) + \theta v(\cdot)) \in \mathcal{A}_{\mathcal{G}}([0,T])$ for all $\theta \in (-\delta, \delta)$, and the functional $\mathcal{J}(\theta) = J((u^*(\cdot) + \theta v(\cdot))$ achieves its minimum at $\theta = 0$,

$$\frac{d}{d\theta}[\mathcal{J}(\theta)]\bigg|_{\theta=0} = \frac{d}{d\theta}J(u^*(t) + \theta v(t))\bigg|_{\theta=0} = 0, \text{ for all } \theta \in (-\delta, \delta), \tag{3.12}$$

Under assumptions $(\mathbf{A1}) - (\mathbf{A3})$ are satisfied, let $(x^*(\cdot), y^*(\cdot), z^*(\cdot))$ denote the solution of the McKean-Vlasov FBSDE (3.1) corresponding to $u^*(\cdot)$ on a given filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$. Then, there exists a unique adapted process $(\Phi^*(\cdot), Q^*(\cdot), \mathcal{K}^*(t))$ that solves (3.9) such that $u^*(.)$ is a stationary point for $\mathbb{E}^{\mathcal{G}_t}(\mathcal{H})$. This implies that for almost all $t \in [0,T]$, we have

$$\mathbb{E}^{\mathcal{G}_t}[\mathcal{H}_u(t, x^*, P_{x^*}, y^*, P_{y^*}, z^*, P_{z^*}, u^*, \Phi^*(\cdot), Q^*(\cdot), \mathcal{K}^*(t))] = 0, \qquad a.e. \tag{3.13}$$

In order to prove this result stated in Theorem 3.3.1, the following technical lemma will be useful in the subsequent analysis.

Lemma 3.3.1 *Applying Itô's formula to the processes $\Phi^*(t)X_1(t), \mathcal{K}^*(t)Y_1(t)$, and taking expectations, we obtain the following expression:*

$$\begin{aligned}
&(\Phi^*(T)X_1(T)) + E(\mathcal{K}^*(T)Y_1(T)) \\
&= -\mathbb{E}\left[\left(\varphi_y\left(y(0), P_{y(0)}\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{p_{y^*}}\varphi\left(y(0), P_{y(0)}\right)\right)\right)Y_1(0)\right] \\
&- \mathbb{E}\int_0^T \{X_1(t)\left[f_x(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{p_{x^*}} f(t)\right)\right] + Y_1(t)\left[f_y(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{p_{y^*}} f(t)\right)\right] \\
&+ Z_1(t)\left[f_z(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{p_{z^*}} f(t)\right)\right] + f_u(t)v(t)\}dt + \mathbb{E}\int_0^T \mathcal{H}_u(t)v(t)\,dt.
\end{aligned} \tag{3.14}$$

Proof. By applying Itô's formula to $\Phi^{*}(t)X_1(t)$ and take expectation, we obtain

$$\begin{aligned}
&\mathbb{E}(\Phi^{*}(T)X_1(T)) \\
&= \mathbb{E}\int_0^T \Phi^{*}(t)dX_1(t) + \mathbb{E}\int_0^T X_1(t)d\Phi^{*}(t) \\
&+ \mathbb{E}\int_0^T Q^{*}(t)\left[\sigma_x(t)X_1(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{p_x}\sigma(t)\widehat{X}_1(t)\right) + \sigma_x(t)v(t)\right]dt \\
&= I_{1,1} + I_{1,2} + I_{1,3},
\end{aligned} \tag{3.15}$$

where

$$\begin{aligned}
I_{1,1} &= \mathbb{E}\int_0^T \Phi^{*}(t)dX_1(t) \\
&= \mathbb{E}\int_0^T \Phi^{*}(t)\left[b_x(t)X_1(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x}b(t)\widehat{X}_1(t)\right) + b_u(t)v(t)\right]dt \\
&= \mathbb{E}\int_0^T \Phi^{*}(t)b_x(t)X_1(t)dt + \mathbb{E}\int_0^T \Phi^{*}(t)\widehat{\mathbb{E}}\left(\partial_\mu^{P_x}b(t)X_1(t)\right)dt \\
&+ \mathbb{E}\int_0^T \Phi^{*}(t)b_u(t)v(t)dt.
\end{aligned}$$

By simple computations, we have

$$\begin{aligned}
I_{1,2} &= \mathbb{E}\int_0^T X_1(t)d\Phi^{*}(t) = -\mathbb{E}\int_0^T X_1(t)\{b_x(t)\Phi^{*}(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x}b(t)\right)\Phi^{*}(t) \\
&+ \left(\sigma_x(t)Q^{*}(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x}\sigma(t)Q^{*}(t)\right)\right) \\
&- g_x(t)\mathcal{K}(t) - \widehat{\mathbb{E}}\left(\partial_\mu^{P_x}f(t)\mathcal{K}(t)\right) + f_x(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x}f(t)\right)\}dt.
\end{aligned} \tag{3.16}$$

Using the same argument as before, one can demonstrate that:

$$\begin{aligned}
I_{1,3} &= \mathbb{E}\int_0^T Q^{*}(t)\left[\sigma_x(t)X_1(t) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_{x^*}}\sigma(t)\widehat{X}_1(t)\right) + \sigma_u(t)v(t)\right]dt \\
&= \mathbb{E}\int_0^T Q^{*}(t)\sigma_x(t)X_1(t)dt + \mathbb{E}\int_0^T Q^{*}(t)\widehat{\mathbb{E}}\left(\partial_\mu^{p_{x^*}}\sigma(t)\widehat{X}_1(t)\right)dt \\
&+ \mathbb{E}\int_0^T Q^{*}(t)\sigma_u(t)v(t)dt.
\end{aligned} \tag{3.17}$$

Combining equations (3.15) to (3.17), we obtain the following expression:

$$\begin{aligned}\mathbb{E}\left(\Phi^*(T)X_1\left(T\right)\right) &= \mathbb{E}\int_0^T \Phi^*(t)b_u\left(t\right)v\left(t\right)dt + \mathbb{E}\int_0^T Q^*\left(t\right)\sigma_u\left(t\right)v\left(t\right)dt \\ &- \mathbb{E}\int_0^T X_1\left(t\right)\left(f_x\left(t\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_{x^*}} f\left(t\right)\right)\right)dt \\ &- \mathbb{E}\int_0^T X_1\left(t\right)\left[g_x\left(t\right)\mathcal{K}\left(t\right) - \widehat{\mathbb{E}}\left(\partial_\mu^{P_x} g\left(t\right)\mathcal{K}\left(t\right)\right)\right]dt.\end{aligned} \tag{3.18}$$

Similarly, by applying Itô's formula to $\mathcal{K}^*(t)Y_1\left(t\right)$ and take expectation, obtain the following expression:

$$\begin{aligned}\mathbb{E}\left(\mathcal{K}^*(T)Y_1(T)\right) &= -\mathbb{E}\left\{\left(\varphi_y\left(y\left(0\right),P_{y(0)}\right)\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{p_{x^*}}\varphi\left(y\left(0\right),P_{y(0)}\right)\right)Y_1\left(0\right)\right\} \\ &+ \mathbb{E}\int_0^T \{\mathcal{K}^*\left(t\right)g_x\left(t\right)X_1\left(t\right) + \mathcal{K}^*\left(t\right)\widehat{\mathbb{E}}\left(\partial_\mu^{P_x} g\left(t\right)\widehat{X}_1\left(t\right)\right) \\ &+ \mathcal{K}^*\left(t\right)g_u\left(t\right)v\left(t\right) - Y_1\left(t\right)\left[f_y\left(t\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_y} f\left(t\right)\right)\right] \\ &- Z_1\left(t\right)\left[f_z\left(t\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_z} f\left(t\right)\right)\right]dt.\end{aligned} \tag{3.19}$$

Finally, the desired result (3.14) is fulfilled from combining (3.18) and (3.19). ■

Proof of Theorem 3.3.1. From (3.12) and by differentiating $\mathcal{J}\left(\theta\right)$ with respect to θ at $\theta = 0$, we have

$$\begin{aligned}\frac{d}{d\theta}\left[\mathcal{J}\left(\theta\right)\right]\Big|_{\theta=0} &= \mathbb{E}\int_0^T [f_x\left(t\right)X_1\left(t\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x} f\left(t\right)\widehat{X}_1\left(t\right)\right) \\ &+ f_y\left(t\right)Y_1\left(t\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_y} f\left(t\right)\widehat{Y}_1\left(t\right)\right) \\ &+ f_z\left(t\right)Z_1\left(t\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_z} f\left(t\right)\widehat{Z}_1\left(t\right)\right) + f_u\left(t\right)v\left(t\right)]dt \\ &+ \mathbb{E}\left[\Phi_x\left(x^*\left(T\right),P_{x^*(T)}\right)X_1\left(T\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_x}\Phi\left(x^*\left(T\right),P_{x^*(T)}\right)\widehat{X}_1\left(t\right)\right)\right] \\ &+ \mathbb{E}\left[\varphi_y\left(y^*\left(0\right),P_{y^*(0)}\right)Y_1\left(0\right) + \widehat{\mathbb{E}}\left(\partial_\mu^{P_y}\varphi\left(y^*\left(0\right),P_{y^*(0)}\right)\widehat{Y}_1\left(t\right)\right)\right] \\ &= 0.\end{aligned} \tag{3.20}$$

From (3.20) and (3.14), we obtain

$$\mathbb{E}\int_0^T \left[\Phi^*(t)\, b_u(t) + Q^*(t)\,\sigma_u(t) + \mathcal{K}(t)\, g_u(t) + f_u(t)\right] v(t)\, dt = 0.$$

From (3.2), we obtain

$$\mathbb{E}\int_0^T \mathcal{H}\left(t, x^*, P_{x^*}, y^*, P_{y^*}, z^*, P_{z^*}, u^*, \Phi^*(\cdot), Q^*(\cdot), \mathcal{K}^*(t)\right) v(t)\, dt = 0, \tag{3.21}$$

Since remains true for all bounded $\mathcal{G}_t$-measurable, we have $t \in [0, T]$

$$E^{\mathcal{G}_t}\left[\mathcal{H}_u\left(t, x^*, P_{x^*}, y^*, P_{y^*}, z^*, P_{z^*}, u^*(t), \Phi^*(t), Q^*(t), \mathcal{K}^*(t)\right)\right] = 0 \quad a.e.,$$

then (3.13) is fulfilled.

Conclusion

In this master dissertation, we have discussed the necessary conditions for optimal control of general controlled McKean-Vlasov FBSDEs, via the differentiability with respect to probability law. This kind of problem, which has a lot of applications in mathematical finance and economics.

Bibliography

[1] Buckdahn, R., Li, J., & Peng, S. (2009). Mean-field backward stochastic differential equations and related partial differential equations. Stochastic processes and their Applications, 119(10), 3133-3154.

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

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

[4] Briand, P. (2001). Equations différentielles stochastiques rétrogrades. Mars.

[5] Cardaliaguet, P. (2013). Notes on mean field games (from Pierre-Louis Lions' lectures at College de France). Université Paris-Dauphine: Paris, France.

[6] Hafayed, M., Abbas, S., & Abba, A. (2015). On mean-field partial information maximum principle of optimal control for stochastic systems with Lévy processes. Journal of Optimization Theory and Applications, 167, 1051-1069.

[7] Hafayed, M., Meherrem, S., Eren, Ş., & Guçoglu, D. H. (2018). On optimal singular control problem for general Mckean-Vlasov differential equations: Necessary and sufficient optimality conditions. Optimal Control Applications and Methods, 39(3), 1202-1219.

[8] Kac, M. (1956, January). Foundations of kinetic theory. In Proceedings of The third Berkeley symposium on mathematical statistics and probability (Vol. 3, pp. 171-197).

[9] Lasry, J. M., & Lions, P. L. (2007). Mean field games. Japanese journal of mathematics, 2(1), 229-260.

[10] McKean Jr, H. P. (1966). A class of Markov processes associated with nonlinear parabolic equations. Proceedings of the National Academy of Sciences, 56(6), 1907-1911.

[11] Meng, Q., & Shen, Y. (2015). Optimal control of mean-field jump-diffusion systems with delay: A stochastic maximum principle approach. Journal of computational and applied mathematics, 279, 13-30.

[12] Meng, Q., & Tang, M. (2009). Necessary and sufficient conditions for optimal control of stochastic systems associated with Lévy processes. Science in China Series F: Information Sciences, 52(11), 1982.

[13] Lakhdari, I. E., Miloudi, H., & Hafayed, M. (2021). Stochastic maximum principle for partially observed optimal control problems of general McKean–Vlasov differential equations. Bulletin of the Iranian Mathematical Society, 47, 1021-1043.

[14] Miloudi, H., Meherrem, S., Eddine Lakhdari, I., & Hafayed, M. (2022). Necessary conditions for partially observed optimal control of general McKean–Vlasov stochastic differential equations with jumps. International Journal of Control, 95(11), 3170-3181.

[15] Wang, G., Zhang, C., & Zhang, W. (2013). Stochastic maximum principle for mean-field type optimal control under partial information. IEEE Transactions on automatic control, 59(2), 522-528.

[16] Meherrem, S., & Hafayed, M. (2019). Maximum principle for optimal control of McKean-Vlasov FBSDEs with Lévy process via the differentiability with respect to probability law. Optimal Control Applications and Methods, 40(3), 499-516.

Appendix A: Abbreviations and Notations

The different abbreviations and notations used throughout this thesis are explained below:

$(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$	Complete probability space.
$\mathcal{F}_t^B$	Filtration generated by the Brownian motion B.
B	Brownian motion.
$a.e.$	Almost everywhere.
$a.s.$	Almost surely.
$\mathcal{H}(t)$	Hamiltonian function.
SDE	Stochastic differential equations.
BSDE	Backward stochastic differential equation.
$Q_2(\mathbb{R})$	The space of all probability measures μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
$\mathbb{L}^2(\otimes, \mathcal{F}, P, \mathbb{R}^n)$	Banach space of $\mathbb{R}$-valued, square integrable random variables.
$\mathcal{A}_{\mathcal{G}}([0, T])$	Set of all admissible controls.
$\mathbb{L}^2_{\mathcal{F}}([0, T]; \mathbb{R}^n)$	Banach space of $\mathcal{F}_t$-predictable processes.
$\mathbb{M}^2_{\mathcal{F}}([0, T]; \mathbb{R}^n)$	Space of all $\mathbb{R}^n$-valued and $\mathcal{F}_t$-adapted processes.
$E^{\mathcal{G}_t}[X]$	Conditional expectation of X.
l^2	Hilbert space of real-valued sequences.
$\mathcal{G}_t$	a subfiltration of $\mathcal{F}_t$.

الملخص

ان الهدف الأساسي من هذه الاطروحة هو عرض نتيجتين مهمتين. النتيجة الأولى تتعلق بوجود ووحدانية الحل بالنسبة للمعادلات التفاضلية العشوائية والمعادلات التفاضلية العشوائية التراجعية في الحالة الليبشيزية. النتيجة الثانية تركز على تحقيق مشكلة التحكم الأمثل العشوائية التي تتضمن المعادلات التفاضلية العامة من نوع McKean-Vlasov والتي يقودها قانون معين وحركة براونية مستقلة.

Abstract

The main objective of this thesis is to present two important results. The first is the existence and uniqueness of solutions for (backward-forward) stochastic differential equations in the Lipschitz case. The second result focuses on investigating a stochastic optimal control problem involving general McKean-Vlasov-type forward-backward differential equations are driven by a given law and an independent Brownian motion.

Résumé

L'objectif principal de cette thèse est de présenter deux résultats importants. Le premier est l'existence et l'unicité des solutions pour les équations différentielles stochastiques (rétrograde) dans le cas de Lipschitz. Le deuxième résultat se concentre sur l'étude d'un problème de contrôle optimal stochastique impliquant des équations différentielles stochastique (rétrograde) générales de type McKean-Vlasov régies par une loi donnée et un mouvement brownien indépendant.