People 's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

THIRD CYCLE LMD FORMATION

A Thesis submitted in partial execution of the requirements for the degree of

DOCTOR IN MATHEMATICS

Suggested by

Mohamed Khidher University Biskra

Presented by

Hanine AZIZI

Titled

Contribution to the Study of Backward SDEs and Their Applications to Stochastic Optimal Control

---

Supervisor: **Pr. Nabil KHELFALLAH**

Examination Committee :

Mr. Abdelhakim NECIR	Professor	University of Biskra	President
Mr. Nabil KHELFALLAH	Professor	University of Biskra	Supervisor
Mr. Salah Eddine REBIAI	Professor	University of Batna 2	Examiner
Mr. Farid CHIGHOUB	Professor	University of Biskra	Examiner
Mr. Boubakeur LABED	MCA	University of Biskra	Examiner
Mr. Imad Eddine LAKHDARI	MCA	University of Biskra	Examiner

Dedication

To my dear parents,

Who always encouraged me to learn until the last breath of my life, who sacrificed their lives to raise me, and who helped me a lot to become the person I am today.

I am proud to be my parent's daughter.

Acknowledgment

First and foremost, I would like to express my heartfelt gratitude to my Thesis Director, Mr. Khelfallah Nabil, Professor at MOHAMED KHIDER BISKRA University, who had the honour of completing my research work after Allah the Almighty. I'd like to express my heartfelt gratitude to him for the exceptional quality of his supervision, his encouragement, assistance, and invaluable advice, as well as his availability and ongoing support, which were invaluable to me.

I acknowledge, with great respect the jury members Pr. Abdelhakim NECIR, Pr. Salah Eddine REBIAI, Pr. Farid CHIGHOUB, Dr. Boubakeur LABED, and Dr. Imad Eddine LAKHDARI, for their confidence in our research work, and I would like to convey my gratitude to them for participating in the examination of this work and for properly evaluating my thesis. I would like to express my gratitude to all my teachers and members of the Mathematics laboratory team.

ملخص

في هذه الأطروحة ، تطرقنا إلى تعميم بعض النتائج الموجودة في الأبحاث التي تتعلق بالحد الأقصى العشوائي للمعادلات التفاضلية العشوائية التراجعية ذات شرط نهائي أو جملة معادلات تفاضلية تتكون من معادلتين الأولى ذات شرط ابتدائي والثانية ذات شرط نهائي. في الفصل الثاني، ناقشنا مشكلة التحكم العشوائي للمعادلات التفاضلية العشوائية التراجعية. ذات المعاملات الليبشيتزية المحلية التي تعد أول عمل وهي مسعى جديد حيث أثبتنا وجود ووحدانية حلول المعادلة المساعدة (معادلة تفاضلية عشوائية خطية ذات معاملات محددة محليا). بعد ذلك، تحصلنا على كل من الشروط اللازمة والكافية لتحقيق الأمثل.

في الفصل الثالث، قمنا بالعمل على مبدأ القيم الحدية القصوى للتحكم الأمثل في جملة المعادلات التفاضلية العشوائية ذات معاملات غير قابلة للتفاضل.

Résumé

L'objectif de cette thèse est de généraliser, dans deux directions différentes, certains résultats existants dans la littérature qui concernent le principe du maximum stochastique pour les équations différentielles stochastiques rétrogrades (EDSR) ou les équations différentielles stochastiques progressives rétrogrades (EDSPR).

La première direction est dédiée au problème de contrôle stochastique pour des EDSR ayant des générateurs localement Lipschitz, où le domaine de contrôle n'est pas nécessairement convexe. Nous établissons une condition nécessaire et suffisante d'optimalité satisfaite par tous les contrôles optimaux. Ces conditions sont décrites par une EDS linéaire localement Lipschitz et une condition maximale sur l'hamiltonien. Nous prouvons d'abord, sous certaines conditions convenable, l'existence d'une solution unique de l'équation adjointe résultante. Ensuite, à l'aide d'un argument d'approximation sur les coefficients, nous définissons une famille d'EDSRs contrôlées avec des générateurs globalement lipschitzienne. Puis, nous dérivons un principe du maximum stochastique approché de tels systèmes. Finalement, nous revenons au problème de contrôle initial en passant à la limite.

La deuxième direction est consacrée au principe du maximum stochastique pour une EDSPR avec des coefficients non différentiables et une diffusion peut être dégénérée. Nous supposons que les coefficients satisfont les conditions de Lipschitz, le domaine de contrôle est non convexe et le coefficient de diffusion n'est pas contrôlé. L'approche que nous allons utilisé est celle de Bouleau-Hirsch. Grâce à cette propriété, nous pouvons définir le processus adjoint en utilisant des dérivées au sens des distributions. Ensuite, nous prouvons la condition nécessaire d'optimalité sous forme d'un principe du maximum de Pontraygin.

Mots clés: Équations différentielles stochastiques rétrograde, équations différentielles stochastiques progressive rétrograde, Contrôle stochastique optimal, principe du maximum stochastique, coefficients localement lipschitzienne, coefficients non differentiable.

Abstract

In this thesis, we aim to generalize some existing results in the literature that concern a stochastic maximum principle for backward stochastic differential equations (BSDEs) or forward-backward stochastic differential equation (FBSDEs), with two possible directions. The first direction is concerned with the stochastic control problem for BSDEs with locally Lipschitz generators, where the domain is not necessarily convex, we establish a necessary and sufficient condition for optimality satisfied by all optimal controls. These conditions are described by a linear locally Lipschitz SDE and a maximum condition on the Hamiltonian. We first prove, under some convenient conditions, the existence of a unique solution to the resulting adjoint equation. Then, with the help of an approximation argument on the coefficients, we define a family of control problems with globally Lipschitz coefficients whereby we derive a stochastic maximum principle for near optimality to such approximated systems. Thereafter, we turn back to the original control problem by passing to the limits.

The second direction is devoted to the stochastic maximum principle in optimal control of possibly degenerate FBSDEs, with irregular coefficients. We assume that the coefficients satisfy the Lipschitz conditions, the control domain is non-convex and the control variable does not enter to the diffusion coefficient. We obtain the necessary conditions for optimality utilizing an adjoint process, which is the unique solution of a linear backward-forward stochastic differential equation and a maximal condition on the Hamiltonian. Thanks to the Bouleau-Hirsch flow property, we are able to define the adjoint process employing the derivatives of the coefficients in the sense of distributions.

Moreover, the adjoint process does not depend on the choice of the representatives of the weak derivatives.

Keys words: Backward stochastic differential equations, forward-backward stochastic differential equation, Optimal stochastic control, stochastic maximum principle, Locally Lipschitz coefficients, non-smooth coefficients.

Contents

Dedication i

Aknowledgemnt ii

Abstract in Arabic iv

Abstract in French iv

Abstract in English vi

Symbols and Abbreviations **x**

General Introduction **xi**

1 *Stochastic Maximum Principle for Optimal Control Problem of Forward-Backward System* **9**

1.1 Introduction 9

1.2 Assumptions and Statement of the Problem 9

1.3 Variational Equations and Variational Inequality 10

1.4 The Maximum Principle in Global Form 19

2 *The Maximum Principle for Optimal Control of BSDEs with Locally Lipschitz Coefficients* **22**

2.1 Introduction 22

2.2 Problem Formulation and Assumptions 22

2.2.1 Formulation of the Control Problem 22

2.2.2 Assumptions 23

2.3 Some Existence and uniqueness Results 25

2.4 A family of Control Problems 27

2.5 Maximum Principle for Optimality 33

2.5.1 Some Convergence Lemmas 33

2.5.2 Necessary Condition for Optimality 38

2.5.3 Sufficient Condition of Optimality 41

3 *A Stochastic Maximum Principle in Optimal Control of FBSDE with Irregular Coefficients* 43

3.1 Introduction 43

3.2 Problem Statement and the Main Result 43

3.2.1 Formulation of Control Problem 43

3.3 A Maximum Principle for a Family of Perturbed Control Problems 47

3.3.1 Estimation Between Two Solutions and some Technical Results . . 51

3.3.2 Maximum Principle for Optimality 57

Conclusion 61

Bibliography 62

List of Symbols and Abbreviations

The different symbols and abbreviations used in this thesis.

- $a.e$: almost everywhere.
- $a.s$: almost surely.
- $\mathbb{R}$: real numbers.
- $\mathbb{R}^n$: n-dimensional real Euclidean space.
- $\mathbb{R}^{n\times d}$: the set of all $(n \times d)$ real matrixes.
- $\bar{A}$: the closure of the set A.
- $\mathbb{1}_A$: the indicator function of the set A.
- $\sigma(A)$: $\sigma-$algebra generated by A.
- $\mathbb{E}\,(x)$: expectation at x.
- $\mathbb{E}\,(\cdot \mid \mathcal{G})$: conditional expectation.
- $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$: filtered probability space.
- $\mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$: filtration.
- $W = (W_t)_{t\in[0,T]}$: Brownian motion.
- $P \otimes dt$: the product measure of $\mathbb{P}$ with the Lebesgue measure dt.
- $S^2\left([0,T], \mathbb{R}^n\right)$: the set of continuous and $\mathbb{F}-$adapted stochastic processes $\{\rho(t); t \in [0,T]\}$, such that $\mathbb{E}\left(\sup_{0\le t\le T} |\rho\,(t)|^2\right) < \infty$.
- $\mathcal{M}^2\left([0,T], \mathbb{R}^n\right)$: the set of $\mathbb{F}-$predictable and $\mathbb{R}^n-$valued processes $\{\rho(t); t \in [0,T]\}$, such that $\mathbb{E}\int_0^T |\rho\,(r)|^2\,\mathrm{d}r < \infty$.

- $(D_t\zeta)_{0\leq t\leq T}$: The Malliavin derivative with respect to W at time t.
- $\mathbb{D}^{1,2}$: The set of all random variables which are Malliavin differentiable.
- $SDEs$: Stochastic differential equations.
- $BSDEs$: Backward stochastic differential equations.
- $FBSDEs$: Forward-backward stochastic differential equations.
- $\mathcal{J}(\cdot)$: The cost function.
- $u(\cdot)$: Optimal control.
- $\mathcal{H}$: The Hamiltonian.
- U : The set of values taken by the strict control v.
- $\mathcal{U}_{ad}$: The set of admissible strict controls.

General Introduction

The field of stochastic optimal control, though a relatively young branch in the realm of mathematics, has captivated the attention of numerous researchers. Its appeal lies not only in its theoretical intricacies but also in its diverse applications across various domains, including mathematical finance, insurance, physics, economics, and more.

Typically, the exploration of stochastic optimal control problems involves two pivotal methodologies. The first is rooted in Bellman's dynamic programming principle, while the second revolves around Pontryagin's maximum principle, also recognized as necessary optimality conditions—an area of focused investigation within this dissertation. The primary aim of the latter is to ascertain optimal controls that minimize or maximize a specified cost functional among the set of all admissible controls, elucidating the necessary and sufficient conditions met by these controls. This endeavor necessitates the introduction of an adjoint process, delineated by a linear stochastic system, and a variational inequality validated by the optimal controls.

We consider the stochastic control problem of minimizing the cost functional

$$\mathcal{J}\left(v\left(\cdot\right)\right)=\mathbb{E}\left[\gamma\left(x(T)\right)+\int_0^T l\left(t,x(t),v(t)\right)\mathrm{d}t\right], \tag{0.1}$$

subject to the following forward stochastic controlled system

$$\begin{cases} dx(t)=b(t,x(t),v(t))\mathrm{d}t+\sigma(t,x(t))\mathrm{d}W_t, \\ x(0)=x, \end{cases} \tag{0.2}$$

where b, σ, l and γ are given maps. $W = (W_t)_{t\in[0,T]}$ is a standard Brownian motion, defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$ satisfying the usual conditions. The control variable $v = (v(t))_{t\in[0,T]}$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-adapted process with values in some subset U of $\mathbb{R}^k$. We denote by $\mathcal{U}_{ad}$ the set of all admissible controls.

A control process $u(\cdot)$ that solves the problem $\{$(0.1),(0.2)$\}$ is called an optimal control that is $u(\cdot)$ satisfies

$$\mathcal{J}(u(\cdot)) = \inf_{v\in\mathcal{U}_{ad}} \mathcal{J}(v(\cdot)). \tag{0.3}$$

Under some differentiability assumptions on the data, the stochastic maximum principle states that

$$\mathcal{H}(t, x(t), u(t), P(t)) = \max_{v(\cdot)\in\mathcal{U}_{ad}} \mathcal{H}(t, x(t), v(t), P(t)); \quad \mathrm{d}t - a.e, \ \ P - a.s, \tag{0.4}$$

which represents the strong form, and the weak form is given by the following equality

$$\mathcal{H}(t, x(t), u(t), P(t)) = \mathcal{H}_v(t, x(t), u(t), P(t))\,(v - u(t))\,; \ \ \mathrm{d}t - a.e, \ \ P - a.s, \tag{0.5}$$

where the Hamiltonian function $\mathcal{H}(t, x, u, P) = Pb(t, x, u) - l(t, x, u)$ and for each $t \in [0, T]$, the adapted process $P(\cdot)$ is given by

$$P(t) = -\mathbb{E}\left[\int_t^T \phi^*(r, t) l_x(r, x(r), u(r))\mathrm{d}r + \phi^*(T, t)\gamma_x(x(T)) \mid \mathcal{F}_t\right], \tag{0.6}$$

here $\phi^*(r, t)$ denotes the transpose of $\phi(r, t)$ $(r \geq t)$ which is the fundamental solution of the linear equation

$$\begin{cases} d\phi(t) = b_x(t, x(t), u(t))\phi(t)\mathrm{d}t + \sum_{j\leq d} \sigma_x^j(t, x(t))\phi(t)\mathrm{d}W_t, \\ \phi(t, t) = I_d. \end{cases} \tag{0.7}$$

The stochastic maximum principle (SMP for short) problem for forward stochastic systems have been extensively studied since the 1970's. According to the convexity or non-convexity of the control domain and the diffusion depend or does not depend on the control, we can split up the existing studies in the literature into four categories.

The first category pertains to situations where the control domain is convex, and the diffusion coefficient σ is independent of the control variable. In this contest, Kushner [36] obtained an SMP of the type (0.4) for a class of controlled stochastic differential equations (SDEs) with smooth coefficients. Subsequently, Haussmann [33] developed a robust formulation of the stochastic maximum principle for a significant class of feedback controls. This formulation allows the control variable to depend on the current state of the system, expressed as $v = (v(x(t)))_{t\in[0,T]}$.

The second category, where the control domain is convex and the diffusion coefficient σ depends explicitly on the control variable was derived by Arkin and Saksonov [4], Bismut [13, 14, 15]. Their findings assert that the optimal control adheres to the stochastic maximum principle (SMP) as indicated by (0.5). Notably, the necessary and sufficient conditions for optimality for linear systems with random coefficients have been established by Cadellinas-Karatzas [21].

The third category represents the cases where the diffusion coefficient σ is independent of the control variable and has been thoroughly explored in [12]. For a comprehensive compilation of references on the stochastic control problem, an extensive list is available in [39, 45].

The fourth category addresses the general case where the control domain is not necessarily convex and the diffusion coefficient σ may involve the control variable, which was established by Peng [42]. He introduced two adjoint processes, the first and second-order, to obtain the second-order variational inequality.

All the previously mentioned papers addressed the maximum principle for stochastic systems with smooth coefficients. The natural question that arises is whether we can derive the necessary conditions of optimality under a set of conditions on the coefficients weaker than the differentiability condition. In this context, several attempts have been made to relax the assumptions on the coefficients to establish a stochastic maximum principle for a broad class of controlled stochastic differential equations (SDEs) in some irregular cases. Specifically, in cases where the coefficients are only globally Lipschitz (not necessarily differentiable). Based on the existing results in the literature, it can be concluded that there are three distinct methods to address the aforementioned control problems.

The first method relies on Clarke's generalized gradients. Utilizing this approach, Mezerdi in 1988 [20] addressed the control problem $\{(0.1),(0.2)\}$ in instances where the drift b is non-smooth but Lipschitz and $l = 0$. The author established a stochastic maximum principle for a controlled stochastic differential equation (SDE) by employing Clarke's generalized gradients and the stable convergence of probability measures. This result serves as a generalization of Kushner's maximum principle.

The second method predominantly relies on Krylov's inequality, necessitating the uniform ellipticity of the diffusion matrix. Employing this approach, Bahlali et al. [11] have developed optimality necessary conditions for the control problem $\{(0.1),(0.2)\}$ in scenarios where the coefficients b, σ and l are Lipschitz continuous but not necessarily differentiable and the diffusion matrix σ is non-degenerate. Utilizing Rademacher's theorem (which asserts that every Lipschitz function is differentiable almost everywhere) along with bounded Borel measurable derivatives, they derived an explicit formula for the adjoint process and established inequalities between the Hamiltonians. Subsequently, Ekeland's variational principle was applied to derive the necessary conditions satisfied by a sequence of near-optimal controls. Finally, the convergence of the scheme, aided by Krylov's inequality, led to necessary conditions for optimality.

The third method utilizes the renowned Bouleau–Hirsch flow property. Bahlali et al. [8] established a stochastic maximum principle for a general class of degenerate diffusion processes, assuming that the coefficients of the state equation b and σ are only Lipschitz continuous and those of the cost functional γ and l are continuously differentiable with respect to the space variables. They employed distributional derivatives of the coefficients and a technique introduced initially by Bouleau and Hirsch [18, 19] to define the adjoint process as the solution of a linear backward stochastic differential equation defined on an extension of the initial probability space. Chighoub et *al.* [24] extended the results of [8] to the case where the coefficients of the state equation b and σ and those of the cost functional γ and l are not differentiable. In the stochastic maximum principle, a significant challenge arises in the computation, particularly numerically, of the adjoint process as expressed by the equality (0.6) which involves a conditional expectation. To address this challenge, we leverage Ito's formula and the martingale representation theorem, demonstrating that the process outlined in (0.6) satisfies the following new equation

$$\begin{cases} -dP(t) = \left[b_x^*(t,x(t),u(t))P(t) + \sum_{j=1}^d \sigma_x^{j,*}(t,x(t))z^j(t) + l_x(t,x(t),u(t))\right] \mathrm{d}t - z(t)\mathrm{d}W_t, \\ P(T) = \gamma_x\left(x(T)\right). \end{cases} \tag{0.8}$$

The equation (0.8), referred to as a linear backward stochastic differential equation (BSDE for short), was introduced by Bismut see [15, 16]. Subsequently, the theory of BSDEs has experienced rapid development at the hands of numerous academic researchers. Notably, among these authors, Pardoux-Peng introduced the nonlinear form of BSDE.

$$\begin{cases} \mathrm{d}y(t) = -f(t,y(t),z(t))\mathrm{d}t + z(t)\mathrm{d}W_t, \\ y(T) = \xi. \end{cases} \tag{0.9}$$

This nonlinear form of BSDE has found significant applications, particularly in the realms of partial differential equations, optimal stochastic control problems, mathematical finance, and stochastic games. For an extensive exploration of its applications, we direct the reader to the seminal papers [40, 41, 26, 1] . Given its widespread utility, it becomes inherently compelling to delve into control problems associated with systems governed by such stochastic equations.

We introduce the following backward stochastic differential equations, for any $v(\cdot) \in \mathcal{U}_{ad}$

$$\begin{cases} dy(t) = & -f\left(t,y(t),z(t),v(t)\right)\mathrm{d}t + z(t)\mathrm{d}W_t, \\ y(T) = & \xi, \end{cases} \tag{0.10}$$

and the expected cost has the form

$$\mathcal{J}(v\left(\cdot\right)) = \mathbb{E}\left[g(y(0))\right], \tag{0.11}$$

where f and g are given functions with appropriate dimensions. Observing that the system (0.10) can be coupled with a controlled stochastic differential equation (SDE) in two distinct manners. The first is termed a fully coupled forward-backward stochastic differential equation (FBSDE), wherein all coefficients depend on the states of the solution processes. The second is referred to as a decoupled FBSDE, such that the forward equation does not depend on the solutions of the backward equation.

There exists an extensive body of literature dedicated to the investigation of stochastic optimal control problems for backward stochastic differential equations (BSDEs) and forward-backward stochastic differential equations (FBSDEs) within the globally Lipschitz framework, coupled with the differentiability of coefficients. For a comprehensive overview, the reader is encouraged to explore works such as those presented in [2, 3, 6, 9, 10, 17, 22, 28, 32, 34, 35, 37, 30] and the associated references.

The initial breakthrough in relaxing the smoothness conditions on the coefficients of controlled FBSDEs, assuming solely the globally Lipschitz condition for the forward part was achieved by Xinwei Feng in [27]. To be more specific, the author established a stochastic maximum principle for optimal control problems of FBSDEs of the following type:

$$\begin{cases} dx(t) = & b\left(t, x(t), v(t)\right) \mathrm{d}t + \sigma\left(t, x(t)\right) \mathrm{d}W_t, \\ X(0) = & x, \\ dy(t) = & -f\left(t, x(t), y(t), z(t), v(t)\right) \mathrm{d}t + z(t) \mathrm{d}W_t, \\ y(T) = & h\left(x(T)\right). \end{cases} \tag{0.12}$$

In this case, the coefficients of the forward part represented by b and σ, are Lipschitz continuous, and the domain of the controls is not necessarily convex. The author applied a technique akin to the one developed by Bahlali et *al.* [11].

As opposed to the globally Lipschitz case, only a limited number of papers have addressed the stochastic maximum principle (SMP) for stochastic differential equations (SDEs) and backward stochastic differential equations (BSDEs) with coefficients that satisfy conditions weaker than the globally Lipschitz condition.

Orrieri, C. in [39], proved a version of a maximum principle for the problem $\{(0.1),(0.2)\}$ where the diffusion coefficient σ depends on the control variable. In this work, he replaced the usual Lipschitz assumption on the drift b with dissipativity conditions, allowing polynomial growth.

Xu, R., & Wu, T. in [43] achieved an existence and uniqueness result for mild solutions to mean-field backward stochastic evolution equations in Hilbert spaces, relaxing the Lipschitz condition. Following this, they established a maximum principle for optimal control problems governed by backward stochastic partial differential equations of mean-field type.

More recently, Azizi, H., & Khelfallah, N in [5] studied stochastic optimal control problems of a BSDE of the type (0.10) Where the coefficients b is a given progressively measurable function which is supposed to be locally Lipschitz with respect to the state variables y and z, the terminal data ξ is bounded and $\mathcal{F}_T$-measurable random variable.

It is worth pointing out that during this work we have encountered two major difficulties. The first one is the coefficients of the following resultant locally Lipschitz linear adjoint equation,

$$\begin{cases} -\mathrm{d}x(t) = b_y(t, y(t), z(t), u(t))x(t)\mathrm{d}t + b_z(t, y(t), z(t), u(t))x(t)\mathrm{d}W_t, \\ x(0) = g_y(y(0)), \end{cases} \tag{0.13}$$

are only locally bounded, and hence, they are locally Lipschitz on x and they do not satisfy the linear growth condition. As a consequence, we can not confirm whether the adjoint Eq.(0.13) admits a unique solution or not. To get around this obstacle, we propose two suitable different sets of conditions, whereby, we can prove that SDE Eq.(0.13) has a unique solution. The second drawback is because the generator of the controlled BSDE is merely locally Lipschitz which makes it technically difficult to apply the standard duality method to derive the necessary condition for optimality. To overcome this difficulty, we propose to convert the problem into the globally Lipschitz framework, by using an approximating argument on the coefficients. Then, by using Ekeland's variational principle and limit argument we investigate the stochastic maximum principle of Pontraygin's type.

Let us briefly describe the contents of this dissertation:

Chapter 1

In the first chapter, we will introduce a stochastic maximum principle for nonlinear controlled forward-backward systems. This analysis will focus on cases where the diffusion coefficient σ does not depend on the control variable and the control domain is not necessarily convex. We formulate the problem as $\{$(0.12),(0.11)$\}$ and outline the main assumptions. We then introduce the conventional first-order variational equations and the variational inequality to derive the maximum principle in its global form. This groundbreaking work was initially explored by Xu, W. [44] and subsequently extended by [42] to encompass the scenario where *sigma* contains the control variable and the control domain need not be convex.

Chapter 2

The second chapter dealt with an optimal control problem for locally Lipschitz BSDE $\{(0.10),(0.11)\}$ which is described by [5]. As the first result of this chapter, we give a new existence and uniqueness result for one type of linear SDEs with locally bounded coefficients. Then, we obtain the necessary conditions for optimality, under two different sets of assumptions as two separate cases. Subsequently, under some further convexity assumptions, we prove that the necessary condition for optimality is in fact sufficient for optimal controls. To the best of our knowledge, this is the first paper studies the stochastic maximum principle for BSDEs under conditions weaker than globally Lipschitz one. More precisely, it extends the stochastic optimal control theory to a large class of BSDEs.

Chapter 3

The third chapter establishes a stochastic maximum principle for the optimal control problem $\{(0.12),(0.11)\}$ applied to possibly degenerate controlled forward-backward stochastic differential equations (FBSDEs). Here, the coefficients b and σ are only Lipschitz continuous concerning the state variable X. Additionally, the diffusion matrix σ does not involve the control variable, and the control domain need not be convex. The method employed to prove the main result revolves around approximating the initial control problem using a sequence of control problems with smooth coefficients, obtained through regularization of the original ones. For the approximate problem, we derive optimality necessary conditions for near optimality by employing Ekeland's variational principle. The adjoint process and the variational inequality between Hamiltonians are derived by transitioning to the limits in the approximate maximum principle, utilizing the Bouleau-Hirsch flow property. Consequently, we obtain an adjoint process that serves as the unique solution of a linear forward-backward stochastic differential equation defined on an extension of the initial probability space.

Chapter 1

Stochastic Maximum Principle for Optimal Control Problem of Forward-Backward System

1.1 Introduction

In this chapter, we will give a detailed demonstration of the maximum principle for optimal control of systems driven by forward and backward stochastic differential equations, where the control variable appears only in the drift and the control domain is not necessarily convex, this work was first investigated in 1995 by Xu, W.[44]. The authors studied the maximum principle in global form.

1.2 Assumptions and Statement of the Problem

We denote $(\Omega, \mathcal{F}, \mathbb{P})$ a filtered probability space equipped with a natural filtration $\mathcal{F}_t = \sigma(W(s), 0 \leq s \leq t)$, where $W(\cdot)$ an $\mathbb{R}^d$-valued standard Wiener process. We denote $\mathcal{U}_{ad}$ the set of all admissible controls $v(\cdot)$, such that $v(\cdot)$ is a measurable $\mathcal{F}_t-$adapted process with values in a compact subset U of $\mathbb{R}^k$.

We consider the stochastic control problem of minimizing the cost function

$$\mathcal{J}(u(\cdot)) = \mathbb{E}\left[\gamma(y_0)\right].$$

Among a set of admissible controls subject to the following forward and backward stochastic control system

$$\begin{cases} dx(t) = b(t, x(t), v(t))dt + \sigma(t, x(t))\mathrm{d}W_t, \\ x(0) = x, \\ dy(t) = f(t, x(t), y(t), z(t), v(t))dt + z(t)\mathrm{d}W_t, \\ y(T) = h(x(T)), \end{cases} \tag{1.1}$$

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

We assume that the following hypothesis holds.

($\mathbf{A}_1$) b, f, σ, h and γ are continuously differentiable with respect to (x, y, z);

($\mathbf{A}_2$) the derivatives of b, f and σ with respect to x, y, z are bounded.

($\mathbf{A}_3$) There exists a constant $C > 0$, such that

$$|h_x| \leq C(1 + |x|), \qquad |\gamma_y| \leq C(1 + |y|).$$

($\mathbf{A}_4$) There exists a constant $C > 0$, such that

$$|b(t, x, v)| \leq C(1 + |x|), \qquad |f(t, x, y, z, v)| \leq C(1 + |x| + |y|).$$

1.3 Variational Equations and Variational Inequality

Let $u(\cdot) \in \mathcal{U}_{ad}$ be optimal, $(x(\cdot), y(\cdot), z(\cdot))$ be the corresponding optimal trajectory of (1.1) be is the solution to the optimal problem. We introduce the spike variational with respect to $u(\cdot)$ as follows:

$$u^\epsilon(t) = \begin{cases} v, \text{ if } \ \tau \leq t \leq \tau + \epsilon, \\ u(t), \text{ otherwise}, \end{cases}$$

where $\epsilon > 0$ is sufficiently small, $v \in U$ is an arbitrary $\mathcal{F}_\tau$-measurable random variable, $t \in [0, T]$, and $sup_{\omega \in \Omega} |v(\omega)| < \infty$. Let $(x^\epsilon(\cdot), y^\epsilon(\cdot), z^\epsilon(\cdot))$ is the trajectory of system (1.1) corresponding to control $u^\epsilon(\cdot)$.

For simplification, we introduce the notations

$$\begin{aligned} b_x &\triangleq b_x(t, x(t), u(t)), & f_x &\triangleq f_x(t, x(t), y(t), z(t), u(t)), \\ b(u^\epsilon) &\triangleq b(t, x(t), u^\epsilon(t)), & b(t) &\triangleq b(t, x(t), u(t)). \end{aligned}$$

Now we introduce the following variational equations

$$\begin{cases} \mathrm{d}x^1(t) = [b_x x^1(t) + b(u^\epsilon) - b(u)]\mathrm{d}t + \sigma_x x^1(t)\mathrm{d}W_t, \\ x^1(0) = 0, \\ \mathrm{d}y^1(t) = [f_x x^1(t) + f_y y^1(t) + f_z z^1(t) + f(u^\epsilon) - f(u)]dt + z^1\mathrm{d}W_t, \\ y^1(T) = h_x(x(T))x^1(T). \end{cases} \tag{1.2}$$

The following lemmas are needed to establish the variational inequality which can be obtained from the fact $\mathcal{J}(u^\epsilon) - \mathcal{J}(u) \geq 0$.

Lemma 1.3.1

We assume $(\mathbf{A}_1) - (\mathbf{A}_4)$ *hold. Then we have*

$$\sup_{0\leq t\leq T} \mathbb{E}\left|x^1(t)\right|^2 \leq C\epsilon^2, \tag{1.3}$$

$$\sup_{0\leq t\leq T} \mathbb{E}\left|x^1(t)\right|^4 \leq C\epsilon^4, \tag{1.4}$$

$$\sup_{0\leq t\leq T} \mathbb{E}\left|y^1(t)\right|^2 \leq C\epsilon^2, \tag{1.5}$$

$$\sup_{0\leq t\leq T} \mathbb{E}\left|y^1(t)\right|^4 \leq C\epsilon^4, \tag{1.6}$$

$$\mathbb{E}\left(\int_0^T (z^1(r))^2\mathrm{d}r\right) \leq C\epsilon^2, \tag{1.7}$$

$$\mathbb{E}\left(\int_0^T (z^1(r))^2\mathrm{d}r\right)^2 \leq C\epsilon^4. \tag{1.8}$$

Proof: By the first equation of (1.2), we get

$$\begin{aligned} &\mathbb{E}\left|x^1(r)\right|^2 \\ &\leq 3\left[\mathbb{E}\left(\int_0^t b_x x^1(r)\mathrm{d}r\right)^2 + \mathbb{E}\left(\int_0^t \left|b(u^\epsilon(r)) - b(u(r))\right|\mathrm{d}r\right)^2 + \mathbb{E}\int_0^t (\sigma_x x^1(r))^2\mathrm{d}r\right], \end{aligned}$$

by Hölder's inequality and assumption $(\mathbf{A}_2)$, we get

$$\begin{aligned} &\mathbb{E}\left|x^1(t)\right|^2 \\ &\leq 3C^2T\mathbb{E}\left(\int_0^t (x^1(r))^2\mathrm{d}r\right) + 3\mathbb{E}\left(\int_0^t \left|b(u^\epsilon(r)) - b(u(r))\right|\mathrm{d}r\right)^2 + 3C^2T\mathbb{E}\left(\int_0^t (x^1(r))^2\mathrm{d}r\right) \\ &\leq 6C^2T\mathbb{E}\left(\int_0^t (x^1(r))^2\mathrm{d}r\right) + 3\mathbb{E}\left(\int_0^t \left|b(u^\epsilon(r)) - b(u(r))\right|\mathrm{d}r\right)^2 \\ &\leq 6C^2T\mathbb{E}\left(\int_0^t (x^1(r))^2\mathrm{d}r\right) + 3\mathbb{E}\left(\int_\tau^{\tau+\epsilon} \left|b(v(r)) - b(u(r))\right|\mathrm{d}r\right)^2 \end{aligned}$$

Using the assumption $(\mathbf{A}_4)$, we obtain

$$\begin{aligned}\mathbb{E}\left|x^1(t)\right|^2 &\leq 6C^2T\mathbb{E}\left(\int_0^t (x^1(r))^2\mathrm{d}r\right) + 3C\left(\int_\tau^{\tau+\epsilon}\mathrm{d}r\right)^2 \\ &\leq 6C^2T\mathbb{E}\left(\int_0^t (x^1(r))^2\mathrm{d}r\right) + 3C\epsilon^2.\end{aligned}$$

By Gronwall's inequality, we have

$$\mathbb{E}\left|x^1(t)\right|^2 \leq C\epsilon^2,\ t \in [0,T].$$

The proof of (1.4) can be obtained similarly.

Then we prove (1.5) and (1.7). Squaring both sides of

$$\begin{aligned}&-y^1(t) - \int_t^T z^1(r)\mathrm{d}W_r \\ &= -h_x(x(T))x^1(T) + \int_t^T \left[f_x x^1(r) + f_y y^1(r) + f_z z^1(r) + f(u^\epsilon(r)) - f(u(r))\right]\mathrm{d}r,\end{aligned}$$

we obtain

$$\mathbb{E}\left(-y^1(t) - \int_t^T z^1(r)\mathrm{d}W_r\right)^2 = \mathbb{E}\left|y^1(t)\right|^2 + \mathbb{E}\left(\int_t^T z^1(r)\mathrm{d}W_r\right)^2 + 2\mathbb{E}\left[y^1(t)\int_t^T z^1(r)\mathrm{d}W_\mathrm{r}\right],$$

and using the fact that

$$\begin{aligned}\mathbb{E}\left[y^1(t)\int_t^T z^1(r)\mathrm{d}W_r\right] &= \mathbb{E}\left[\mathbb{E}\left(y^1(t)\int_t^T z^1(r)\mathrm{d}W_r \mid \mathcal{F}_t\right)\right] \\ &= \mathbb{E}\left[y^1(t)\mathbb{E}\left(\int_t^T z^1(r)\mathrm{d}W_r \mid \mathcal{F}_t\right)\right] \\ &= \mathbb{E}\left[y^1(t)\mathbb{E}\left(\left[\int_0^T z^1(r) - \int_0^t z^1(r)\right]\mathrm{d}W_r \mid \mathcal{F}_t\right)\right] \\ &= \mathbb{E}\left[y^1(t)\mathbb{E}\left(\int_0^T z^1(r)\mathrm{d}W_r \mid \mathcal{F}_t\right) - y^1(t)\int_0^t z^1(r)\mathrm{d}W_r\right] \\ &= \mathbb{E}\left[y^1(t)\int_0^t z^1(r)\mathrm{d}W_r - y^1(t)\int_0^t z^1(r)\mathrm{d}W_r\right] = 0,\end{aligned}$$

we obtain

$$\begin{aligned}&\mathbb{E}\left(-y^1(t) - \int_t^T z^1(r)\mathrm{d}W_r\right)^2 \\ &= \mathbb{E}\left(-h_x(x(T))x^1(T) + \int_t^T \left[f_x x^1(r) + f_y y^1(r) + f_z z^1(r) + f(u^\epsilon(r)) - f(u(r))\right]\mathrm{d}r\right)^2 \\ &\leq 5\mathbb{E}(-h_x(x(T))x^1(T))^2 + 5\mathbb{E}\left(\int_t^T f_x x^1(r)\mathrm{d}r\right)^2 + 5\mathbb{E}\left(\int_t^T f_y y^1(r)\mathrm{d}r\right)^2 \\ &+ 5\mathbb{E}\left(\int_t^T f_z z^1(r)\mathrm{d}r\right)^2 + 5\mathbb{E}\left(\int_t^T \left[f(u^\epsilon(r)) - f(u(r))\right]\mathrm{d}r\right)^2.\end{aligned}$$

From Hölder's inequality and $(\mathbf{A}_2)$, we get

$$\begin{aligned}
&\mathbb{E}\left|y^1(t)\right|^2 + \mathbb{E}\left(\int_t^T (z^1(r))^2 \mathrm{d}r\right) \\
&\leq 5C^2\mathbb{E}\left((x^1(T))^2\right) + 5C^2T\mathbb{E}\left(\int_t^T (x^1(r))^2 \mathrm{d}r\right) + 5C^2T\mathbb{E}\left(\int_t^T (y^1(r))^2 \mathrm{d}r\right) \\
&+ 5C^2(T-t)\mathbb{E}\left(\int_t^T (z^1(r))^2 \mathrm{d}r\right) + 5\mathbb{E}\left(\int_t^T \left[f(u^\epsilon(r)) - f(u(r))\right] \mathrm{d}r\right)^2,
\end{aligned}$$

with $\delta = \dfrac{1}{10C^2}$, $t \in [T-\delta, T]$

$$\begin{aligned}
\mathbb{E}\left|y_t^1\right|^2 + \frac{1}{2}\mathbb{E}\left(\int_t^T (z_s^1)^2 \mathrm{d}r\right) &\leq 5C^2\mathbb{E}\left((x_T^1)^2\right) + 5C^2T\mathbb{E}\left(\int_0^T (x_s^1)^2 \mathrm{d}r\right) \\
&+ 5C^2T\mathbb{E}\left(\int_t^T (y^1(r))^2 \mathrm{d}r\right) + 5\mathbb{E}\left(\int_t^T \left[f(u^\epsilon(r)) - f(u(r))\right] \mathrm{d}r\right)^2.
\end{aligned}$$

Using the assumption $(\mathbf{A}_4)$ and Gronwall's inequality, we obtain

$$\begin{aligned}
\mathbb{E}\left|y^1(t)\right|^2 &\leq C\epsilon^2,\ t \in [T-\delta, T], \\
\mathbb{E}\left(\int_t^T (z^1(r))^2 \mathrm{d}r\right) &\leq C\epsilon^2,\ t \in [T-\delta, T].
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&-y^1(t) - \int_t^{T-\delta} z^1(r) \mathrm{d}W_r \\
&= -y^1(T-\delta) + \int_t^{T-\delta} \Big[f_x x^1(r) + f_y y^1(r) + f_z z^1(r) + f(u^\epsilon(r)) - f(u(r))\Big] \mathrm{d}r,
\end{aligned}$$

then

$$\begin{aligned}
&\mathbb{E}\left|y^1(t)\right|^2 + \mathbb{E}\left(\int_t^{T-\delta} (z^1(r))^2 \mathrm{d}r\right) \\
&\leq 5\mathbb{E}\left|y^1(T-\delta)\right|^2 + 5C^2T\mathbb{E}\left(\int_t^{T-\delta} (x^1(r))^2 \mathrm{d}r\right) + 5C^2T\mathbb{E}\left(\int_t^{T-\delta} (y^1(r))^2 \mathrm{d}r\right) \\
&+ 5C^2(T-\delta-t)\mathbb{E}\left(\int_t^{T-\delta} (z^1(r))^2 \mathrm{d}r\right) + 5\mathbb{E}\left(\int_t^{T-\delta} \left[f(u^\epsilon(r)) - f(u(r))\right] \mathrm{d}r\right)^2.
\end{aligned}$$

So

$$\begin{aligned}
\mathbb{E}\left|y^1(t)\right|^2 &\leq C\epsilon^2,\ t \in [T-2\delta, T], \\
\mathbb{E}\left(\int_t^T (z^1(r))^2 \mathrm{d}r\right) &\leq C\epsilon^2,\ t \in [T-2\delta, T].
\end{aligned}$$

(1.5) and (1.7) are obtained after a finite number of iterations. Then, by using a similar method and the inequality

$$\mathbb{E}\left(\int_t^T z^1(r)\mathrm{d}B_r\right)^4 \geq \beta\mathbb{E}\left(\int_t^T (z^1(r))^2\mathrm{d}r\right)^2, \beta>0,$$

can be proved (1.6) and (1.8). ∎

Lemma 1.3.2

Let assumptions $(\mathbf{A}_1)$ *and* $(\mathbf{A}_2)$*hold. Then we have the following estimations:*

$$\sup_{0\leq t\leq T}\mathbb{E}\left|x^\epsilon(t)-x(t)-x^1(t)\right|^2 \leq C_\epsilon\epsilon^2,\, C_\epsilon\to 0, \tag{1.9}$$

$$\sup_{0\leq t\leq T}\mathbb{E}\left|y^\epsilon(t)-y(t)-y^1(t)\right|^2 \leq C_\epsilon\epsilon^2,\;\; C_\epsilon\to 0, \tag{1.10}$$

$$\sup_{0\leq t\leq T}\mathbb{E}\int_t^T\left|z^\epsilon(r)-z(r)-z^1(r)\right|^2\mathrm{d}r \leq C_\epsilon\epsilon^2,\;\; C_\epsilon\to 0. \tag{1.11}$$

Proof: Let us first prove (1.9), It can be easily checked that

$$\begin{aligned}
&\int_0^t b(x(r)+x^1(r),u^\epsilon(r))\mathrm{d}r+\int_0^t \sigma(x(r)+x^1(r))\mathrm{d}W_r \\
&=\int_0^t\left[b(x(r)+x^1(r),u^\epsilon(r))-b(x(r),u^\epsilon(r))+b(x(r),u^\epsilon(r))\right]\mathrm{d}r \\
&+\int_0^t\left[\sigma(x(r)+x^1(r))-\sigma(x(r))+\sigma(x(r))\right]\mathrm{d}W_r \\
&=\int_0^t\left[b(x(r),u^\epsilon(r))+\int_0^1 b_x(x(r)+\lambda x^1(r),u^\epsilon(r))x^1(r)d\lambda\right]\mathrm{d}r \\
&+\int_0^t\left[\sigma(x(r))+\int_0^1 \sigma_x(x(r)+\lambda x^1(r))x^1(r)d\lambda\right]\mathrm{d}W_r \\
&=\int_0^t\left[b(x(r),u^\epsilon(r))+\int_0^1 b_x(x(r)+\lambda x^1(r),u^\epsilon(r))x^1(r)d\lambda\right]\mathrm{d}r \\
&+\int_0^t\left[b(x(r),u(r))-b(x(r),u(r))+b_x(x(r),u(r))x^1(r)-b_x(x(r),u(r))x^1(r)\right]\mathrm{d}r \\
&+\int_0^t\left[\sigma(x(r))+\sigma_x(x(r))x^1(r)-\sigma_x(x(r))x^1(r)+\int_0^1 \sigma_x(x(r)+\lambda x^1(r))x^1(r)d\lambda\right]\mathrm{d}W_r
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^t b(x(r),u(r))\mathrm{d}r + \int_0^t \sigma(x(r))\mathrm{d}W_r + \int_0^t \Big[b_x(x(r),u(r))x^1(r) + b(x(r),u^\epsilon(r)) - b(x(r),u(r))\Big]\,\mathrm{d}r \\
&+ \int_0^t \left[\int_0^1 \Big[b_x(x(r)+\lambda x^1(r),u^\epsilon(r)) - b_x(x(r),u(r))\Big]\, dx^1(r)\lambda\right]\mathrm{d}r + \int_0^t \sigma(x(r))x^1(r)\mathrm{d}W_r \\
&+ \int_0^t \left[\int_0^1 \Big[\sigma_x(x(r)+\lambda x^1(r)) - \sigma_x(x^1(r))\Big]\, x^1(r)d\lambda\right]\mathrm{d}W_r \\
&= \int_0^t b(x(r),u(r))\mathrm{d}r + \int_0^t \sigma(x(r))\mathrm{d}W_r + \int_0^t \Big[b_x(x(r),u(r))x^1_s + b(x(r),u^\epsilon(r)) - b(x(r),u(r))\Big]\,\mathrm{d}r \\
&+ \int_0^t \sigma_x(x(r))x^1(r)\mathrm{d}B_r + \int_0^t A^\epsilon(r)\mathrm{d}r + \int_0^t B^\epsilon(r)\mathrm{d}W_r \\
&= x(t) - x(0) + x^1(t) + \int_0^t A^\epsilon(r)\mathrm{d}r + \int_0^t B^\epsilon(r)\mathrm{d}W_r,
\end{aligned}
$$

where

$$
\begin{aligned}
A^\epsilon(r) &= \int_0^1 \Big[b_x(x(r)+\lambda x^1(r),u^\epsilon(r)) - b_x(x(r),u(r))\Big]\, x^1(r)d\lambda, \\
B^\epsilon(r) &= \int_0^1 \Big[\sigma_x(x(r)+\lambda x^1(r)) - \sigma_x(x(r))\Big]\, x^1(r)d\lambda.
\end{aligned}
$$

From Lemma (1.3.1) we can easily get

$$
\sup_{0\le t\le T} \mathbb{E}\left[\left(\int_0^t A^\epsilon(r)\mathrm{d}r\right)^2 \left(\int_0^t B^\epsilon(r)\mathrm{d}W_r\right)^2\right] = o(\epsilon^2). \tag{1.12}
$$

Then by

$$
x^\epsilon(t) - x(0) = \int_0^t b(x^\epsilon(r),u^\epsilon(r))\mathrm{d}r + \int_0^t \sigma(x^\epsilon(r))\mathrm{d}W_r,
$$

we have

$$
\begin{aligned}
&x^\epsilon(t) - x(t) - x^1(t) \\
&= \int_0^t b(x^\epsilon(r),u^\epsilon(r))\mathrm{d}r + \int_0^t \sigma(x^\epsilon(r))\mathrm{d}W_r - \int_0^t b(x(r),u(r))\mathrm{d}r - \int_0^t \sigma(x(r))\mathrm{d}W_r \\
&\quad - \int_0^t \Big[b_x(x(r),u(r))x^1(r) + b(x(r),u^\epsilon(r)) - b(x(r),u(r))\Big]\,\mathrm{d}r - \int_0^t \sigma_x(x(r))x^1(r)\mathrm{d}W_r \\
&= \int_0^t b(x^\epsilon(r),u^\epsilon(r))\mathrm{d}r - \int_0^t b(x(r)+x^1(r),u^\epsilon(r))\mathrm{d}r + \int_0^t b(x(r)+x^1(r),u^\epsilon(r))\mathrm{d}r \\
&\quad + \int_0^t \sigma(x^\epsilon(r))\mathrm{d}W_r - \int_0^t \sigma(x(r)+x^1(r))\mathrm{d}W_r + \int_0^t \sigma(x(r)+x^1(r))\mathrm{d}W_r \\
&\quad - \int_0^t b(x(r),u(r))\mathrm{d}r - \int_0^t \sigma(x(r))\mathrm{d}W_r \\
&\quad - \int_0^t \Big[b_x(x(r),u(r))x^1(r) + b(x(r),u^\epsilon(r)) - b(x(r),u(r))\Big]\,\mathrm{d}r - \int_0^t \sigma_x(x(r))x^1(r)\mathrm{d}W_r \\
&= \int_0^t \Big[b(x^\epsilon(r),u^\epsilon(r)) - b(x(r)+x^1(r),u^\epsilon(r))\Big]\,\mathrm{d}r + \int_0^t \Big[\sigma(x^\epsilon(r)) - \sigma(x(r)+x^1(r))\Big]\,\mathrm{d}W_r \\
&\quad + \int_0^t \Big[b(x(r)+x^1(r),u^\epsilon(r)) - b(x(r),u^\epsilon(r))\Big]\,\mathrm{d}r + \int_0^t \Big[\sigma(x(r)+x^1(r)) - \sigma(x(r))\Big]\,\mathrm{d}W_r \\
&\quad - \int_0^t b_x(x(r),u(r))x^1(r)\mathrm{d}r - \int_0^t \sigma_x(x(r))x^1(r)\mathrm{d}W_r,
\end{aligned}
$$

then

$$
\begin{aligned}
&x^{\epsilon}(r)-x(r)-x^{1}(r)\\
&=\int_0^t\left[\int_0^1 b_x(x(r)+x^1(r)+\lambda(x^{\epsilon}(r)-x(r)-x^1(r)),u^{\epsilon}(r))(x^{\epsilon}(r)-x(r)-x^1(r))d\lambda\right]\mathrm{d}r\\
&+\int_0^t\left[\int_0^1 \sigma_x(x(r)+x^1(r)+\lambda(x^{\epsilon}(r)-x(r)-x^1(r)))(x^{\epsilon}(r)-x(r)-x^1(r))d\lambda\right]\mathrm{d}W_r\\
&+\int_0^t\left[\int_0^1 \left[b_x(x(r)+\lambda x^1(r),u^{\epsilon}(r))-b_x(x(r),u(r))\right]x^1(r)d\lambda\right]\mathrm{d}r\\
&+\int_0^t\left[\int_0^1 \left[\sigma_x(x(r)+\lambda x^1(r))-\sigma_x(x(r))\right]x^1(r)d\lambda\right]\mathrm{d}W_r,
\end{aligned}
$$

we obtain

$$
\begin{aligned}
&x^{\epsilon}(r)-x(r)-x^{1}(r)\\
&=\int_0^t C^{\epsilon}(r)(x^{\epsilon}(r)-x(r)-x^1(r))\mathrm{d}r+\int_0^t D^{\epsilon}(r)(x^{\epsilon}(r)-x(r)-x^1(r))\mathrm{d}W_r\\
&+\int_0^t A^{\epsilon}(r)\mathrm{d}r+\int_0^t B^{\epsilon}(r)\mathrm{d}W_r,
\end{aligned}
$$

where

$$
\begin{aligned}
C^{\epsilon}(r)&=\int_0^1 b_x(x(r)-x^1(r)+\lambda(x^{\epsilon}(r)-x(r)-x^1(r)),u^{\epsilon}(r))d\lambda,\\
D^{\epsilon}(r)&=\int_0^1 \sigma_x(x(r)-x^1(r)+\lambda(x^{\epsilon}(r)-x(r)-x^1(r)))d\lambda.
\end{aligned}
$$

Using Gronwall's inequality, (1.9) follows from the above relation and (1.12).

Now we prove (1.10) and (1.11). it easy to see that

$$
\begin{aligned}
&\int_t^T f(x(r)+x^1(r),y(r)+y^1(r),z(r)+z^1(r),u^{\epsilon}(r))+\int_t^T(z(r)+z^1(r))\mathrm{d}W_r\\
&=\int_t^T f(x(r)+x^1(r),y(r)+y^1(r),z(r)+z^1(r),u^{\epsilon}(r))\mathrm{d}r-\int_t^T f(x(r),y(r),z(r),u^{\epsilon}(r))\mathrm{d}r\\
&+\int_t^T f(x(r),y(r),z(r),u^{\epsilon}(r))\mathrm{d}r+\int_t^T(f_x x^1(r)+f_y y^1(r)+f_z z^1(r))\mathrm{d}r\\
&-\int_t^T(f_x x^1(r)+f_y y^1(r)+f_z z^1(r))\mathrm{d}r+\int_t^T(z(r)+z^1(r))\mathrm{d}W_r\\
&+\int_t^T f(x(r),y(r),z(r),u(r))\mathrm{d}r-\int_t^T f(x(r),y(r),z(r),u(r))\mathrm{d}r,
\end{aligned}
$$

then

$$
\begin{aligned}
&\int_t^T f(x(r)+x^1(r),y(r)+y^1(r),z(r)+z^1(r),u^\epsilon(r))\mathrm{d}r+\int_t^T (z(r)+z^1(r))\mathrm{d}W_r \\
&=\int_t^T \left[\int_0^1 \left[f_x(x(r)+\lambda x^1(r),y(r)+\lambda y^1(r),z(r)+\lambda z^1(r),u^\epsilon(r))-f_x\right]d\lambda x^1(r)\right]\mathrm{d}r \\
&+\int_t^T \left[\int_0^1 \left[f_y(x(r)+\lambda x^1(r),y(r)+\lambda y^1(r),z(r)+\lambda z^1(r),u^\epsilon(r))-f_y\right]d\lambda y^1(r)\right]\mathrm{d}r \\
&\int_t^T \left[\int_0^1 \left[f_z(x(r)+\lambda x^1(r),y(r)+\lambda y^1(r),z(r)+\lambda z^1(r),u^\epsilon(r))-f_z\right]d\lambda z^1(r)\right]\mathrm{d}r \\
&+\int_t^T \left[\int_0^1 \left[f_x x^1(r)+f_y y^1(r)+f_z z^1(r)+f(x(r),y(r),z(r),u^\epsilon(r))-f(x(r),y(r),z(r),u(r))\right]\right]\mathrm{d}r \\
&+\int_t^T z^1(r)\mathrm{d}W_r+\int_t^T f(x(r),y(r),z(r),u(r))\mathrm{d}r+\int_t^T z(r)\mathrm{d}W_r \\
&=\int_t^T G^\epsilon(r)\mathrm{d}r+h_x(x(T))x^1(t)-y^1(t)+h(x(T))-y(t),
\end{aligned}
$$

where

$$
\begin{aligned}
G^\epsilon(r)&=\int_0^1 \left[f_x(x(r)+\lambda x^1(r),y(r)+\lambda y^1(r),z(r)+\lambda z^1(r),u^\epsilon(r))-f_x\right]d\lambda x^1(r) \\
&+\int_0^1 \left[f_y(x(r)+\lambda x^1(r),y(r)+\lambda y^1(r),z(r)+\lambda z^1(r),u^\epsilon(r))-f_y\right]d\lambda y^1(r) \\
&+\int_0^1 \left[f_z(x(r)+\lambda x^1(r),y(r)+\lambda y^1(r),z(r)+\lambda z^1(r),u^\epsilon(r))-f_z\right]d\lambda z^1(r).
\end{aligned}
$$

So we have

$$
\begin{aligned}
&-(y^\epsilon(t)-y(t)-y^1(t)) \\
&=-h(x^\epsilon(T))+\int_t^T f(x^\epsilon(r),y^\epsilon(r),z^\epsilon(r),u^\epsilon(r))\mathrm{d}r \\
&+\int_t^T z^\epsilon(r)\mathrm{d}W_r+h(x(T))-\int_t^T f(x(r),y(r),z(r),u(r))\mathrm{d}r \\
&-\int_t^T z(r)\mathrm{d}W_r+h_x(x(T))x^1(T)-\int_t^T z^1(r)\mathrm{d}W_r \\
&-\int_t^T \left[f_x x^1(r)+f_y y^1(r)+f_z z^1(r)+f(x(r),y(r),z(r),u^\epsilon(r))-f(x(r),y(r),z(r),u(r))\right]\mathrm{d}r \\
&-\int_t^T f(x(r)+x^1(r),y(r)+y^1(r),z(r)+z^1(r),u^\epsilon(r))\mathrm{d}r \\
&+\int_t^T f(x(r)+x^1(r),y(r)+y^1(r),z(r)+z^1(r),u^\epsilon(r))\mathrm{d}r \\
&=-(h(x^\epsilon(T))-h(x(T)))+h_x(x(T))x^1(T) \\
&+\int_t^T \left[f(x^\epsilon(r),y^\epsilon(r),z^\epsilon(r),u^\epsilon(r))-f(x(r)+x^1(r),y(r)+y^1(r),z(r)+z^1(r),u^\epsilon(r))\right]\mathrm{d}r \\
&+\int_t^T (z^\epsilon(r)-z(r)-z^1(r))\mathrm{d}W_r+\int_t^T G^\epsilon(r)\mathrm{d}r,
\end{aligned}
$$

we get

$$\mathbb{E}\left[-(y^{\epsilon}(t)-y(t)-y^{1}(t))-\int_{t}^{T}(z^{\epsilon}(r)-z(r)-z^{1}(r))\mathrm{d}W_{r}\right]^{2}$$
$$=\mathbb{E}[-h(x^{\epsilon}(T))-h(x(T))+h_{x}(x(T))x^{1}(T)+\int_{t}^{T}G^{\epsilon}(r)\mathrm{d}r$$
$$+\int_{t}^{T}\left[f(x^{\epsilon}(r),y^{\epsilon}(r),z^{\epsilon}(r),u^{\epsilon}(r))-f(x(r)+x^{1}(r),y(r)+y^{1}(r),z(r)+z^{1}(r),u^{\epsilon}(r))\right]\mathrm{d}r$$
$$+h(x(T))-h(x(T))+h(x^{1}(T))-h(x^{1}(T))]^{2},$$

then

$$\mathbb{E}\left|(y^{\epsilon}(t)-y(t)-y^{1}(t))\right|^{2}+\mathbb{E}\left(\int_{t}^{T}\left|(z^{\epsilon}(r)-z(r)-z^{1}(r))\right|^{2}\mathrm{d}r\right)$$
$$=\mathbb{E}[-h(x^{\epsilon}(T))-h(x(T)+x^{1}(T))+\int_{t}^{T}G^{\epsilon}(r)\mathrm{d}r$$
$$+\int_{t}^{T}\left[f(x^{\epsilon}(r),y^{\epsilon}(r),z^{\epsilon}(r),u^{\epsilon}(r)-f(x(r)+x^{1}(r),y(r)+y^{1}(r),z(r)+z^{1}(r),u^{\epsilon}(r))\right]\mathrm{d}r$$
$$-[h(x(T)+x^{1}(T))-h(x(T))]+h_{x}(x(T))x^{1}(T)]^{2},$$

we obtain

$$\mathbb{E}\left|(y^{\epsilon}(t)-y(t)-y^{1}(t))\right|^{2}+\mathbb{E}\left(\int_{t}^{T}\left|(z^{\epsilon}(r)-z(r)-z^{1}(r))\right|^{2}\mathrm{d}r\right)$$
$$=\mathbb{E}\Bigg\{-h(x^{\epsilon}(T))-h(x(T)+x^{1}(T))+\int_{t}^{T}G^{\epsilon}(r)\mathrm{d}r$$
$$+\int_{t}^{T}\left[f(x^{\epsilon}(r),y^{\epsilon}(r),z^{\epsilon}(r),u^{\epsilon}(r))-f(x(r)+x^{1}(r),y(r)+y^{1}(r),z(r)+z^{1}(r),u^{\epsilon}(r))\right]\mathrm{d}r$$
$$-\int_{0}^{1}\left[h_{x}(x(T)+x^{1}(T)-h_{x}(x(T))\right]x^{1}(T)d\lambda\Bigg\}^{2}.$$

By Lemma (1.3.1) and (1.9), we have

$$\sup_{0\leq t\leq T}\mathbb{E}\left(\int_{t}^{T}G^{\epsilon}(r)\mathrm{d}r\right)^{2}=o(\epsilon^{2}),$$
$$\mathbb{E}[h(x^{\epsilon}(T))-\varphi(x(T)+x^{1}(T))]^{2}=o(\epsilon^{2}).$$

We obtain (1.10) and (1.11) by using the same iteration method of Lemma (1.3.1) to the above relation. ■

Lemma 1.3.3

Under the assumptions $(\mathbf{A}_1)-(\mathbf{A}_4)$, *the following variational inequality holds*

$$\mathbb{E}\left[\gamma_y(y(0))y^1(0)\right] \geq o(\epsilon). \tag{1.13}$$

Proof: From Lemma (1.3.2), we get

$$\mathbb{E}\left[\gamma(y^\epsilon(0)) - \gamma(y(0)+y^1(0))\right] = o(\epsilon),$$

therefore

$$\begin{aligned} 0 &\leq \mathbb{E}[\gamma(y(0)+y^1(0)) - \gamma(y(0))] + o(\epsilon) \\ &= \mathbb{E}\left[\gamma_y(y(0))y^1(0)\right] + o(\epsilon). \end{aligned}$$

■

1.4 The Maximum Principle in Global Form

In this Section, we introduce the adjoint equations. Then, we present the main result concerning the maximum principle for optimal control problems. To this end, let us define the Hamiltonian $\mathcal{H}$ from $[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{n\times d}$ to $\mathbb{R}^n$ by

$$\mathcal{H}(t,x,y,z,u,p,q,k) \triangleq \langle p, b(t,x,u)\rangle + \langle q, f(t,x,y,z,u)\rangle + \langle k, \sigma(t,x)\rangle,$$

and the Hamilton function for our problem. Starting from the variational inequality obtained in Lemma **1.3.3**, the maximum principle can be proved by using Ito's formula.

The adjoint equations are

$$\begin{cases} -\mathrm{d}p(t) = (b_x^* p(t) + f_x^* q(t) + \sigma_x^* k(t))\mathrm{d}t - k(t)\mathrm{d}W_t, \\ p(T) = -h_x^*(x(T))q(T), \\ -\mathrm{d}q(t) = f_y^* q(t)\mathrm{d}t + f_z^* q(t)\mathrm{d}W_t, \\ q(0) = -\gamma_y(y(0)). \end{cases} \tag{1.14}$$

The adjoint equations (1.14) can be rewritten in terms of the derivatives of the Hamiltonian as

$$\begin{cases} -\mathrm{d}p(t) = \mathcal{H}_x dt - k(t)\mathrm{d}W_t, \\ p(T) = -h_x^*(x(T))q(T), \\ -\mathrm{d}q(t) = \mathcal{H}_y dt + \mathcal{H}_z \mathrm{d}W_t, \\ q(0) = -\gamma_y(y(0)). \end{cases} \tag{1.15}$$

From (1.15) and Lemma **1.3.3**, we have the following theorem

Theorem 1.4.1

Assume $(\mathbf{A}_1) - (\mathbf{A}_4)$ *hold. Let* $(x(\cdot),\ y(\cdot),\ z(\cdot),\ u(\cdot))$ *be an optimal control and its corresponding trajectory of (1.1),* $(p(\cdot), q(\cdot), k(\cdot))$ *be the corresponding solutions of (1.14)Then the maximum principle holds, that is*

$$\mathcal{H}(t, x(t), y(t), z(t), v(t), p(t), q(t), k(t)) \tag{1.16}$$

$$\geq \mathcal{H}(t, x(t), y(t), z(t), u(t), p(t), q(t), k(t)),\ \forall v \in U,\quad a.e,\quad a.s. \tag{1.17}$$

Proof: Using Ito's formula to $\langle p(t), x^1(t)\rangle$ and $\langle q(t), y^1(t)\rangle$, and we use the fact that $q(0) = -\gamma_y(y(0))$ and $p(T) = -h_x(x(T)q(T)$, $y^1(T) = h_x(x(T))x^1(T)$, we obtain

$$\begin{aligned} &\mathbb{E}\left[-h_x(x(T))q(T)x^1(T)\right] \\ &= -\mathbb{E}\left[\int_0^T f_x(t, x(t), y(t), z(t), u(t))q(t)x^1(t)\mathrm{d}t\right] \\ &+ \mathbb{E}\left[\int_0^T [b(t, x(t), u^\epsilon(t)) - b(t, x(t), u(t))]\right] p(t)\mathrm{d}t, \end{aligned} \tag{1.18}$$

and

$$\begin{aligned} &\mathbb{E}\left[h_x(x(T))q(T)x^1(T)\right] + \mathbb{E}\left[\gamma_y(y(0))y^1(0)\right] \\ &= \mathbb{E}\left[\int_0^T f_x(t, x(t), y(t), z(t), u(t))q(t)x^1(t)\mathrm{d}t\right] \\ &+ \mathbb{E}\left[\int_0^T [f(t, x(t), y(t), z(t), u^\epsilon(t)) - f(t, x(t), y(t), z(t), u(t))]q(t)\mathrm{d}t\right]. \end{aligned} \tag{1.19}$$

Using (1.18) and (1.19), so we get

$$\begin{aligned} \mathbb{E}\left[\gamma_y(y(0))y^1(0)\right] = \mathbb{E}[\int_0^T &(p(t)[b(t, x(t), u^\epsilon(t)) - b(t, x(t), u(t)] \\ &+ q(t)[f(t, x(t), y(t), z(t), u^\epsilon(t)) - f(t, x(t), y(t), z(t), u(t))])]\mathrm{d}t, \end{aligned} \tag{1.20}$$

by Replacing (1.20) in (1.13), we obtain

$$\begin{aligned}o(\epsilon) \leq \mathbb{E}[\int_0^T &(p(t)[b(t,x(t),u^\epsilon(t)) - b(t,x(t),u(t))]\\ &+ q(t)[f(t,x(t),y(t),z(t),u^\epsilon(t)) - f(t,x(t),y(t),z(t),u(t))])]\mathrm{d}t,\end{aligned}$$

$$\begin{aligned}o(\epsilon) \leq &\ \mathbb{E}\left[\int_0^T [p(t)b(t,x(t),u^\epsilon(t)) + q(t)f(t,x(t),y(t),z(t),u^\epsilon(t))]\mathrm{d}t\right]\\ &- \mathbb{E}\left[\int_0^T [p(t)b(t,x(t),u(t)) + q(t)f(t,x(t),y(t),z(t),u(t))]\mathrm{d}t\right],\end{aligned}$$

$$\begin{aligned}o(\epsilon) \leq &\ \mathbb{E}\left[\int_0^T [p(t)b(t,x(t),u^\epsilon(t)) + q(t)f(t,x(t),y(t),z(t),u^\epsilon(t)) + k(t)\sigma(t,x(t))]\mathrm{d}t\right]\\ &- \mathbb{E}\left[\int_0^T [p(t)b(t,x(t),u(t)) + q(t)f(t,x(t),y(t),z(t),u(t)) + k(t)\sigma(t,x(t))]\mathrm{d}t\right].\end{aligned}$$

Thus it follows that

$$\begin{aligned}&\mathbb{E}\left[\int_0^T \mathcal{H}(t,x(t),y(t),z(t),u^\epsilon(t),p(t),q(t),k(t))\mathrm{d}t\right]\\ &-\mathbb{E}\left[\int_0^T \mathcal{H}(t,x(t),y(t),z(t),u(t),p(t),q(t),k(t))\mathrm{d}t\right] \geq o(\epsilon).\end{aligned}$$

Which gives the desired result. ■

Chapter 2

The Maximum Principle for Optimal Control of BSDEs with Locally Lipschitz Coefficients

2.1 Introduction

In the second chapter, we study the Maximum Principle for Optimal Control of BSDEs with Locally Lipschitz Coefficients. For the first time, we prove a new existence result to one kind of linear SDEs with locally bounded coefficients. Then, we state the control problem along with some auxiliary results. The last topic of this chapter is devoted to the study of the necessary and sufficient conditions of optimality.

2.2 Problem Formulation and Assumptions

2.2.1 Formulation of the Control Problem

Let T be a strictly positive real number, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ be a complete probability space equipped with a filtration satisfying the usual conditions, on which a d-dimensional Brownian motion $W = (W_t)_{0\leq t\leq T}$ is defined. We assume that $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ is the P-augmentation of natural filtration generated by $(W_t)_{0\leq t\leq T}$. Throughout this chapter, we will use the following spaces:

- $S^2\left([0,T], \mathbb{R}^n\right)$: denotes the set of continuous and $\mathbb{F}$-adapted stochastic processes $\{y(t); t \in [0,T]\}$, such that $\mathbb{E}\left(\sup_{0\leq t\leq T} |y(t)|^2\right) < \infty$.
- $\mathcal{M}^2\left([0,T], \mathbb{R}^n\right)$: denotes the set of $\mathbb{F}$-predictable and $\mathbb{R}^n$-valued processes $\{z(t); t \in [0,T]\}$, such that $\mathbb{E}\int_0^T |z(r)|^2 \, \mathrm{d}r < \infty$.

We consider the following controlled backward stochastic differential equation (BSDE for short):

$$\begin{cases} \mathrm{d}y(t) = b(t, y(t), z(t), v(t))\mathrm{d}t + z(t)\mathrm{d}W_t, \\ y(T) = \xi, \end{cases} \tag{2.1}$$

where $b : [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times U \longrightarrow \mathbb{R}^n$ is a given measurable function which supposed to be locally Lipschitz with respect to the state variables y and z, the terminal datum ξ is a bounded and $\mathcal{F}_T$-adapted random variable. The process $v(\cdot)$ stands for the control variable, which is assumed to be an $\mathbb{F}$-adapted process that takes values in a given non-empty subset U of $\mathbb{R}^n$. We denote the set of all admissible controls by $\mathcal{U}_{ad}$. For a given measurable function $g : \mathbb{R}^n \to \mathbb{R}$, we introduce the cost functional of our stochastic control problem

$$\mathcal{J}(v(\cdot)) = \mathbb{E}\left[g(y(0))\right]. \tag{2.2}$$

The controller wants to minimize the cost functional Eq(2.2) among the set of all admissible controls. Now, we can formulate our control problem as the following:

Problem (A): To find $u(\cdot) \in \mathcal{U}_{ad}$ such that $u(\cdot)$ minimizes the cost functional Eq.(2.2) subject to Eq.(2.1).

2.2.2 Assumptions

Throughout this chapter, we shall work on the following two sets of assumptions as two separate cases that we will deal with.

Assumption 2.1

(**H.1**) b and g are continuously differentiable with respect to (y,z), and the derivatives b_y, b_z and g_y are continuous in y and z.

(H.2) There exist a constant, $M > 0$ such that for every y and z,

$$\langle y, b(t,y,z,v)\rangle \leq M(1 + |y|^2 + |y|\,|z|); \qquad \mathbb{P}-a.s., a.e.\ t \in [0,T].$$

(**H.3**) There exist two constants, $M > 0$, $\alpha \in (0,1)$ and a positive function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, such that

$$|b(t,y,z,v)| \leq M(1 + \varphi(|y|) + |z|^{\alpha}); \qquad \mathbb{P}-a.s., a.e.\ t \in [0,T].$$

(**H.4**) For every $N \in \mathbb{N}$, there exists a constants $L_N > 0$, such that

$$|b(t,y,z,v) - b(t,y',z,v)| \leq L_N |y - y'|; \qquad \mathbb{P}-a.s., a.e.\ t \in [0,T],$$

and $\forall y, y'$, such that $|y| \leq N$, $|y'| \leq N$.

(**H.5**) There exists a constant $L > 0$, such that

$$|b(t,y,z,u) - b(t,y,z',u)| \leq L(|z - z'|);\ \ \mathbb{P}-a.s., a.e.\ t \in [0,T].$$

(H.6) There exists a constant $M > 0$ and a positive function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, such that,

$$|b_y((t,y,z,v))| \leq M(1 + \varphi|y|); \qquad \mathbb{P}-a.s., a.e.\ t \in [0,T]. \tag{2.3}$$

We let $\mathbb{D}^{1,2}$ denote the set of all random variables which are Malliavin differentiable; $(D_t\zeta)_{0\leq t\leq T}$ the Malliavin derivative with respect to $W.$ at time t of a given random variable $\zeta \in \mathbb{D}^{1,2}$. We refer the reader to [38] for more information about Malliavin's calculus and its applications. Now, we introduce the following Assumptions,

Assumption 2.2

(**H.7**) Assume that **(H.1)** is fulfilled. Furthermore, suppose that ξ is an element of $\mathbb{D}^{1,2}$ and there exists a constant M such that

$$\left|D_t^i \xi\right| \leq M, \forall t \leq T; i = 1, p,$$

(**H.8**) There exists a constant $M > 0$ such that,

$$|b(t,y,z,v)| \leq M(1 + |y| + |z|); \qquad \mathbb{P}-a.s., a.e.\ t \in [0,T].$$

(H.9) b_z satisfies **(H.8),** there exists a constant $M > 0$ and $\alpha \in]0,1[$ such that

$$|b_y(t,y,z,v)| \leq M\left(1 + \ln\left(1 + (|y| + |z|)^{\alpha}\right)\right); \mathbb{P}-a.s., a.e.\ t \in [0,T].$$

(**H.10**) For every $N \in \mathbb{N}$, there exists a constants $L_N > 0$, such that

$$|b(t,y,z,v) - b(t,y',z',v)| \leq L_N(|y - y'| + |z - z'|); \qquad \mathbb{P}-a.s., a.e.\ t \in [0,T],$$

and $\forall y, y', z, z'$ such that $|y| \leq N$, $|y'| \leq N$, $|z| \leq N$, $|z'| \leq N$.

Examples: To motivate the **Assumption 2.1 and Assumption 2.2**, we exhibit some relevant examples for the coefficient b and the terminal data ξ.

1) Let $g : \mathbb{R}^d \rightarrow \mathbb{R}^n$ be a continuously differentiable function with bounded derivatives. If $\xi = g(W_T)$ then for every $t \leq T$, $D_t\xi = \nabla g(W_T)$, and thus ξ satisfies (**H.7**).

2) Let $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a function defined by

$$h(x) = \begin{cases} x & \text{if } 0 \leq x \leq 1 \\ x^\alpha & \text{if } x > 1 \end{cases} \quad \alpha \in (0,1).$$

Obviously, h is Lipschitz and satisfies the sub-linear growth condition. Define $b(t,y,z,v) := -(1+y)\log|1+y| + h(z) + f(t,v)$ where $f : [0,T] \times U \rightarrow \mathbb{R}$ is a bounded function, then b satisfies the **Assumption 2.1.** Indeed, due to the fact that $|(1+y)\log(1+y)| \leq 1 + \frac{1}{\alpha}|y|^{1+\alpha}$ for all $\alpha > 0$ and $\langle y, -y\log y\rangle \leq 1$, it is not difficult to see that the b is locally Lipschitz on $\mathbb{R}_+$ and satisfies (**H.1), (H.2),** (**H.3**), (**H.5**) and (**H.6**).

3) Define the function $h_1(x) := 2\sqrt{1+\log x} + \log\left|\frac{\sqrt{1+\log x}-1}{\sqrt{1+\log x}+1}\right|$, then $b(t,y,z,v) := h_1(y) + h_1(z) + f(t,v)$ satisfies **Assumptions 2.2**.

2.3 Some Existence and uniqueness Results

In this section, we will state some basic results related to BSDEs theory and prove a new existence and uniqueness results for one kind of linear SDEs with locally Lipschitz coefficients. More precisely, we prove under Assumptions 2.1 or Assumptions 2.2 that the linear adjoint equation Eq.(0.13) has a unique solution. The following two lemmas state some existing results in the literature which are related to BSDEs with locally Lipschitz generators. More precisely, they provide the existence and uniqueness of solutions on top of some estimates satisfied by their solutions.

Lemma 2.3.1

Let ξ be a bounded random variable. Assume that ***Assumptions 2.1*** *or* ***Assumptions 2.2*** *are satisfied. Then, for any $v(\cdot) \in \mathcal{U}_{ad}$ there exists a unique pair $(y(\cdot), z(\cdot)) \in S^2([0,T], \mathbb{R}^n) \times \mathcal{M}^2\left([0,T], \mathbb{R}^{n\times d}\right)$ which solves BSDE Eq.(2.1).*

Proof: The proof under **Assumptions 2.1** can be found in Bahlali [7, Proposition 7], and the proof under **Assumptions 2.2** has been established in [31, Theorem 2.c and Theorem 2.d (ii)].

Before we state and prove an existence and uniqueness result for SDE (0.13). We need the following auxiliary lemmas.

Lemma 2.3.2

Let $(y(\cdot), z(\cdot))$ be the unique solutions of the BSDEs Eq.(2.1). Then, there is a positive constant C, such that

i) Under ***Assumptions 2.1,*** *we have $\mathbb{P}-a.s.$*

$$\sup_{0\leq t\leq T} |y(t)|^2 \leq C \text{ and } \mathbb{E}\int_0^T |z(r)|^2 \mathrm{d}r \leq C; \tag{2.4}$$

ii) Under ***Assumptions 2.2,*** *we have*

$$\sup_{0\leq t\leq T} (|y(t)| + |z(t)|) \leq C; \;\; \mathbb{P}-a.s. \tag{2.5}$$

Proof : The proof of (i) follows by using Ito's formula, the conditional expectation, Jensen's inequality and Gronwall's Lemma. The proof of (ii) is given in [31, Proposition 2.d (i)]. ■

Remark 2.1. Notice that for any $v(\cdot) \in \mathcal{U}_{ad}$ the functions $b_y(t,\cdot,\cdot,v(t))$ and $b_z(t,\cdot,\cdot,v(t))$ are not bounded in general. However, for any $t \in [0,T]$, such that $(y(t), z(t))$ is the unique solution of BSDE Eq.(2.1), $b_y(t,y(t),z(t),v(t))$ and $b_z(t,y(t),z(t),v(t))$ are bounded. Indeed, under **Assumption 2.1**, Lemma 2.3.2 (i), shows that there exists a positive constant such that $\sup_{0\leq t\leq T} |y(t)| \leq \sqrt{C}$. We conclude using the hypothesis **(H.6)**,

$$b_y(t,y(t),z(t),v(t)) \leq M(1+\varphi(|y(t)|)) \leq K(M,\varphi).$$

Besides, $b_z(t,y(t),z(t),v(t))$ is bounded due to the fact that b is Lipschitz in z.

On the other hand, under **Assumption 2.2,** $b_y(t, y(t), z(t), v(t))$ and $b_z(t, y(t), z(t), v(t))$ evaluated at $(y(t), z(t))$ are bounded, it is easy to see that, using **(H.9)** and Lemma 2.3.2 (ii),

$$\begin{aligned} b_z(t, y(t), z(t), v(t)) &\leq M(1 + |y(t)| + |z(t)|) \\ &\leq M(1 + \sup_{0\leq t\leq T}(|y(t)| + |z(t)|)) \\ &\leq M(1 + C). \end{aligned}$$

Finally, the boundedness of b_y goes similarly.

Throughout the following theorem, we claim a new existence and uniqueness results for SDE Eq.(0.13).

Theorem 2.3.1 (*Existence and uniqueness of SDE*)

*Suppose that **Assumptions 2.1** or **Assumptions 2.2** holds. Then, for any $v(\cdot) \in \mathcal{U}_{ad}$, SDE Eq.(0.13) has one and only one solution $x(\cdot) \in \mathcal{S}^2([0, T], \mathbb{R}^n)$. Moreover, there is a positive constant C such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|x(t)|^4\right] \leq C. \tag{2.6}$$

Proof: Under **Assumption 2.1** or **Assumption 2.2,** Remark 2.1 shows that, for any $t \in [0, T]$, $b_y(t, y(t), z(t), v(t))$ and $b_z(t, y(t), z(t), v(t))$ are bounded, where $(y(t), z(t))$ is the unique solution of the BSDE Eq.(2.1). This implies that the coefficients of the Eq.(0.13) satisfy the globally Lipschitz condition on top of the linear growth propriety. Therefore, it has a unique solution such that, for any $p \geq 1$, we have

$$\mathbb{E}\left(\sup_{0\leq t\leq T}|x(t)|^p\right) \leq C_p.$$

In particular, the inequality (2.6) is satisfied. This finishes the proof. ■

2.4 A family of Control Problems

Since the purpose of this chapter is to deal with the control Problem **(A)**. The controller objective is to derive a necessary condition as well as a sufficient condition of optimality under the locally Lipschitz framework. Notice that, because the derivatives of b are not bounded, the standard duality technique can not be directly applicable in our setup. To overcome this difficulty we approximate Problem (A) by a family of perturbed control problems with globally Lipschiz coefficients. Then we apply Ekeland's variational principle in order to derive the necessary and sufficient conditions for near-optimality. Finally, the desired results are obtained by using the limit argument.

Before we state the following lemma which plays an essential role in approximating the initial control Problem (A), we need to define the following family of semi-norms.

For any $p \geq 1$, we define $(\rho_{n,p}(b))_{n\in\mathbb{N}}$ by

$$\rho_{n,p}(b) = \left(\mathbb{E}\int_0^T \sup_{|y|,|z|\leq n} |b(r,y,z)|^p \,\mathrm{d}r\right)^{\frac{1}{p}}. \tag{2.7}$$

Lemma 2.4.1

*Let b be a function which satisfies **Assumptions 2.1** or **Assumptions 2.2**. Then, there exists a sequence of functions b^n such that ,*

(i) For each n, b^n is globally Lipschitz in (y,z) $\mathbb{P}-a.s., a.e.$ $t \in [0,T]$

*(ii) If b satisfies **(H.3)**, then $\sup_n |b^n(t,y,z,v)| \leq M(1+\varphi(|y|)+|z|^\alpha)$ $\mathbb{P}-a.s, a.e.$ $t \in [0,T]$*

*(iii) If b satisfies **(H.8),** then $\sup_n |b^n(t,y,z,v)| \leq M(1+|y|+|z|)$ $\mathbb{P}-a.s, a.e.$ $t \in [0,T]$*

(iv) For every n, $\rho_{n,p}(b^n - b) \to 0$ as $n \to \infty$.

(v) For every n, $\left|b^n_y\right| \leq |b_y| + \beta\,|b|\,\eta_n$ and $|b^n_z| \leq |b_z| + \beta\,|b|\,\eta_n$, where η_n converges to 0 as n tends to $+\infty$.

Proof: Let $(\psi_n)_{n\in\mathbb{N}}$ be a sequence of smooth functions with support in the ball $B(0,n+1)$ and such that $\psi_n = 1$ in the ball $B(0,n)$. Obviously, the sequence of truncated functions, defined by $b_n = b\psi_n$ satisfies the assertions (i), (ii), (iii) and (iv). We give now the proof of (v). By the definition of b^n, we have, for any (t,y,z,v)

$$b^n_y(t,y,z,v) = b_y(t,y,z,v)\psi^n(y,z) + b(t,y,z,v)\psi^n_y(y,z).$$

Set $\beta = \sup_n \sup\left\{\psi^n_y(y,z), |y| \leq n+1, |z| \leq n+1\right\}$ and $\eta_n = 1\!\!1_{]-n-1,-n[\cup]n,n+1[}$. Then,

$$\left|b^n_y(t,y,z,v)\right| \leq |b_y(t,y,z,v)| + \beta\,|b(t,y,z,v)|\,\eta_n.$$

Obviously, $\lim_{n\to\infty}\eta_n = 0$. The proof of Lemma **2.4.1** is complete. ∎

Let us recall Ekeland's variational principle, which plays crucial role in proving the necessary condition of near optimality.

Lemma 2.4.2 (*Ekeland's variational principle*)

Let (V,d) be a complete metric space and $f : V \to \mathbb{R}\cup\{+\infty\}$ be a lower-semicontinuous function, bounded from below. If for each $\varepsilon > 0$ there exists $u(\cdot) \in V$ satisfies $f(u) \leq \inf\{f(v)+\varepsilon; v \in V\}$. Then, there exists u_ε such that

1. $f(u_\varepsilon) \leq f(u)$.

2. $d(u,u_\varepsilon) \leq \varepsilon^{\frac{1}{2}}$

3. $f(v) + \varepsilon^{\frac{1}{2}} \cdot d(v,u_\varepsilon) < f(u_\varepsilon), \forall v \in V$.

Our aim in the next paragraphs is to convert the initial control Problem **(A)** by a family of control problems with globally Lipschitz coefficients. To this end, for any fixed $n \in \mathbb{N}$ and $v(\cdot) \in \mathcal{U}_{ad}$, we denote $(\bar{y}(\cdot), \bar{z}(\cdot))$ the solution of the following controlled BSDE

$$\begin{cases} d\bar{y}(t) = b^n(t,\bar{y}(t),\bar{z}(t),v(t))\mathrm{d}t + \bar{z}(t)\,\mathrm{d}W_t, \\ \bar{y}(T) = \xi, \end{cases} \tag{2.8}$$

and

$$\mathcal{J}^n(v(\cdot)) = \mathbb{E}\left[g(\bar{y}(0))\right]. \tag{2.9}$$

The following lemma gives some estimates that will be used to relate the control problem {Eq.(2.8), (2.9)} with Problem **(A)**.

Lemma 2.4.3

Let $y(\cdot)$ and $y_1(\cdot)$ be respectively the solutions of BSDE Eqs.(2.1) and (2.8) corresponding to the control $v(\cdot) \in \mathcal{U}_{ad}$, then the following estimates hold:

(i) $\mathbb{E}\left[|\bar{y}(t)-y(t)|^2\right] \leq K_{n,N}$, and $\mathbb{E}\left[\int_t^T |\bar{z}(r)-z(r)|^2\,\mathrm{d}r\right] \leq K_{n,N}$;

(ii) $|\mathcal{J}^n(v) - \mathcal{J}(v)| \leq C \cdot \varepsilon_{n,N}$;

where $K_{n,N}$ and $\varepsilon_{n,N}$ converge to 0 as n and N tend successively to $+\infty$, here N stands for the radius of the ball $B(0,N)$.

Proof: Let us first prove, under **Assumption 2.2**, the two inequalities of assertion (i). The proof under **Assumption 2.1** is similar.

Squaring both sides of

$$-\left(\bar{y}(t)-y(t)\right)-\int_t^T\left(\bar{z}(r)-z(r)\right)\mathrm{d}W_r=\int_t^T\left[b^n(r,\bar{y}(r),\bar{z}(r),u(r)-b(r,y(r),z(r),u(r)\right]\mathrm{d}r,$$

Taking the expectation and using the fact that

$$\mathbb{E}\left[\left(\bar{y}(t)-y(t)\right)\int_t^T\left(\bar{z}(r)-z(r)\right)\mathrm{d}W_r\right]=0,$$

we get

$$\mathbb{E}\left|\bar{y}(t)-y(t)\right|^2+\mathbb{E}\int_t^T\left|\bar{z}(r)-z(r)\right|^2\mathrm{d}r=\mathbb{E}\left(\int_t^T b^n(r,\bar{y}(r),\bar{z}(r),u(r)-b(r,y(r),z(r),u(r)\mathrm{d}r\right)^2.$$

Then, the Cauchy-Schwarz inequality leads to

$$\mathbb{E}\left|\bar{y}(t)-y(t)\right|^2+\mathbb{E}\int_t^T\left|\bar{z}(r)-z(r)\right|^2\mathrm{d}r\leq(T-t)\mathbb{E}\int_t^T\left|b^n(r,\bar{y}(r),\bar{z}(r),u(r)-b(r,y(r),z(r),u(r)\right|^2\mathrm{d}r.$$

For a given $N>1$, let L_N be the Lipschitz constant of b in the ball $B(0,N)$, we define $A_N:=\left\{(r,w);\ \ |z(r)|^2+|\bar{z}(r)|^2>N^2\right\}$ and $\bar{A}_N=\Omega\backslash A_N$, it follows by using the Cauchy-Schwarz inequality,

$$\mathbb{E}\left|\bar{y}(t)-y(t)\right|^2+\mathbb{E}\int_t^T\left|\bar{z}(r)-z(r)\right|^2\mathrm{d}r\leq(T-t)\left(J_1^n+J_2^n\right),\tag{2.10}$$

where

$$J_1^n=\mathbb{E}\int_t^T\left|b^n(r,\bar{y}(r),\bar{z}(r),u(r)-b(r,y(r),z(r),u(r)\right|^2\mathbb{1}_{A_N}\mathrm{d}r,$$

and

$$J_2^n=\mathbb{E}\int_t^T\left|b^n(r,\bar{y}(r),\bar{z}(r),u(r)-b(r,y(r),z(r),u(r)\right|^2\mathbb{1}_{\bar{A}_N}\mathrm{d}r.$$

We first estimate J_1^n. Since b satisfies (**H.3**) and $\sup_n|b^n|\leq|b|$, we use Holder's inequality and the fact that $\mathbb{1}_{A^N}<\dfrac{|\bar{z}(r)|^2+|z(r)|^2}{N^{2(1-\alpha)}}$, we obtain

$$J_1^n\leq\frac{K(M,\xi)}{N^{2(1-\alpha)}}.\tag{2.11}$$

Now, we proceed to estimate J_2^n. Taking into consideration that b is Lipschitz in the ball $B(0,N)$, we get

$$J_2^n\leq 2\rho_N^2(b^n-b)+2L_N^2\mathbb{E}\int_t^T\left(|\bar{y}(r)-y(r)|^2+|\bar{z}(r)-z(r)|^2\right)\mathrm{d}r.\tag{2.12}$$

Then, we have, by replacing (2.10) and (2.11) into (2.12)

$$\begin{aligned}\mathbb{E}\left|\bar{y}(t)-y(t)\right|^2+\mathbb{E}\int_t^T\left|\bar{z}(r)-z(r)\right|^2\mathrm{d}r\leq(T-t)[&\frac{K(M,\xi)}{N^{2(1-\alpha)}}+2\rho_N^2(b^n-b)\\&+2L_N^2\mathbb{E}\int_t^T\left(|\bar{y}(r)-y(r)|^2+|\bar{z}(r)-z(r)|^2\right)\mathrm{d}r].\end{aligned}$$

For every δ, such that $T-t=\delta$, we obtain by choosing $\delta=\dfrac{1}{4L_N^2}$,

$$\mathbb{E}\left|\bar{y}(t)-y(t)\right|^2+\mathbb{E}\int_{T-\delta}^{T}\left|\bar{z}(r)-z(r)\right|^2\mathrm{d}r\le\frac{K(M,\xi)}{4L_N^2N^{2(1-\alpha)}}+\frac{\rho_N^2(b^n-b)}{2}$$
$$+\frac{1}{2}\mathbb{E}\int_{T-\delta}^{T}\left(\left|\bar{y}(r)-y(r)\right|^2+\left|\bar{z}(r)-z(r)\right|^2\right)\mathrm{d}r.$$

From the above inequality, we derive two inequalities

$$\mathbb{E}\left|\bar{y}(t)-y(t)\right|^2\le\frac{K(M,\xi)}{4L_N^2N^{2(1-\alpha)}}+\frac{\rho_N^2(b^n-b)}{2}+\frac{1}{2}\mathbb{E}\int_{T-\delta}^{T}\left|\bar{y}(r)-y(r)\right|^2\mathrm{d}r,$$

and

$$\frac{1}{2}\mathbb{E}\int_{T-\delta}^{T}\left|\bar{z}(r)-z(r)\right|^2\mathrm{d}r\le\frac{K(M,\xi)}{4L_N^2N^{2(1-\alpha)}}+\frac{\rho_N^2(b^n-b)}{2}+\mathbb{E}\int_{T-\delta}^{T}\left|\bar{y}(r)-y(r)\right|^2\mathrm{d}r.$$

Set $K_{n,N}:=\frac{K(M,\xi)}{4L_N^2N^{2(1-\alpha)}}+\frac{\rho_N^2(b^n-b)}{2}$. Obviously $K_{n,N}$ tends to 0 as n and N tend successively to $+\infty$. Consequently, by Gronwall's lemma, the first inequality, becomes

$$\mathbb{E}\left|\bar{y}(t)-y(t)\right|^2\le C_1K_{n,N}. \tag{2.13}$$

Then, by (2.13), we obtain

$$\mathbb{E}\int_{T-\delta}^{T}\left|\bar{z}(r)-z(r)\right|^2\mathrm{d}r\le C_2K_{n,N}.$$

Similarly, we get

$$\begin{aligned}\mathbb{E}\left|\bar{y}(t)-y(t)\right|^2&\le CK_{n,N}^1, &\quad t\in[T-2\delta,T-\delta].\\ \mathbb{E}\int_t^T\left|\bar{z}(r)-z(r)\right|^2\mathrm{d}r&\le CK_{n,N}^2, &\quad t\in[T-2\delta,T-\delta].\end{aligned}$$

After a finite number of iterations, we prove the assertion (i). Next, we prove the assertion (ii).

Since g is Lipschitz continuous, then by using the Cauchy-Schwarz inequality and (2.13), one can prove the following

$$\left|\mathcal{J}^n(u(\cdot))-\mathcal{J}(u(\cdot))\right|\le C\cdot\varepsilon_{n,N}.$$

such that $\varepsilon_{n,N}$ tends to 0 as n and N tend successively to $+\infty$. The proof of Lemma 2.4.3 is complete. ■

Let $u(\cdot)$ be an optimal control for the initial control problem, that is $u(\cdot)$ satisfies

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

subject to Eq.(2.1). Note that $u(\cdot)$ is not necessarily optimal for the new perturbed control problem, according to Lemma 2.4.3, there exists a sequence (δ_n) of positive real numbers converging to 0 such that:

$$\mathcal{J}^n(u(\cdot)) \leq \inf_{v(\cdot)\in U_{ad}} \mathcal{J}^n(v(\cdot)) + \delta_{n,N}, \quad \delta_{n,N} = 2C \cdot \varepsilon_{n,N}.$$

To apply Ekeland's lemma 2.4.2, let us define a metric d on the space of admissible controls. For $u(\cdot), v(\cdot) \in \mathcal{U}_{ad}$

$$d(u(\cdot), v(\cdot)) = P \otimes \mathrm{d}t \left\{(w,t) \in \Omega \times [0,T] : u(w,t) \neq v(w,t)\right\}, \tag{2.14}$$

where $P \otimes \mathrm{d}t$ is the product measure of P with the Lebesgue measure on $[0,T]$. According to Ekeland's lemma applied to the continuous cost functional $\mathcal{J}^n(u(\cdot))$, there exists an admissible control $u^n(\cdot)$ such that:

$$d(u^n(\cdot), u(\cdot)) \leq (\delta_{n,N})^{\frac{1}{2}},$$

and

$$\tilde{\mathcal{J}}^n(u^n(\cdot)) \leq \tilde{\mathcal{J}}^n(v(\cdot)) \text{ for any } v(\cdot) \in \mathcal{U}_{ad},$$

where

$$\tilde{\mathcal{J}}^n(v(\cdot)) = \mathcal{J}^n(v(\cdot)) + (\delta_{n,N})^{\frac{1}{2}} \cdot d(v(\cdot), u^n(\cdot)). \tag{2.15}$$

This means that $u^n(\cdot)$ is optimal for control problem {Eq.(2.8), (2.9)} with the new cost function $\tilde{\mathcal{J}}^n$. For each integer n, we denote by $(y^n(\cdot), z^n(\cdot))$ the unique solution of the following BSDE controlled by $u^n(\cdot)$

$$\begin{cases} dy^n(t) = b^n(t, y^n(t), z^n(t), u^n(t))\mathrm{d}t + z^n(t)\,\mathrm{d}W_t, \\ y^n(T) = \xi. \end{cases} \tag{2.16}$$

And its corresponding cost is given by

$$\mathcal{J}^n(v(\cdot)) = \mathbb{E}\left[g(y^n(0))\right]. \tag{2.17}$$

Then, we can formulate the following optimal control problem.

Problem (B): For each integer n, we want to find $u^n(\cdot) \in \mathcal{U}_{ad}$ such that $u^n(\cdot)$ minimizes the cost function Eq.(2.17) subject to Eq.(2.16).

We conclude this subsection by introducing a family of controlled SDEs called adjoint equations. For each integer n, we introduce the following SDE

$$\begin{cases} -dx^n(t) = b^n_y(t, y^n(t), z^n(t), u^n(t))x^n(t)\mathrm{d}t + b^n_z(t, y^n(t), z^n(t), u^n(t))x^n(t)\mathrm{d}W_t, \\ x^n(0) = g_y(y^n(0)). \end{cases} \tag{2.18}$$

Since b^n is globally Lipschitz function, their derivatives b^n_y and b^n_z are bounded. Hence, the coefficients of SDE Eq.(2.18) are globally Lipschitz and of linear growth, which means that, for each integer n, equation Eq.(2.18) admits a unique solution.

We also define the a family of Hamiltonian functions $\mathcal{H}^n : [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times \mathbb{R}^n \times U \to \mathbb{R}$ by

$$\mathcal{H}^n(t, y^n, z^n, x^n, u^n) = x^n b^n(t, y^n, z^n, u^n) \text{ for each } n \in \mathbb{N}.$$

2.5 Maximum Principle for Optimality

The purpose of this section is to derive the necessary conditions of optimality for the aforementioned control Problem (**A**). To this end, we need some auxiliary lemma which will be gathered and proved in the next subsection.

2.5.1 Some Convergence Lemmas

In this subsection, we will summarize and prove some useful lemmas which will be used in the next section to prove the main results.

Lemma 2.5.1

Let (b_n) be the sequence of functions associated to b by Lemma 2.4.1 and $(y^n(\cdot), z^n(\cdot))$ stands for the solution of Eq.(2.16). Then, there exists a constant $C = C(M)$ such that,

*i) Under **Assumption 2.1,** we have $\mathbb{P}-a.s$*

$$\sup_n \left(\sup_{0\le t\le T} |y^n(t)|^2 \right) \le C \text{ and } \sup_n \mathbb{E} \int_0^T |z^n(t)|^2 \le C.$$

*ii) Under **Assumption 2.2,** we have*

$$\sup_n \left(\sup_{0\leq t\leq T} \left(|y^n(t)| + |z^n(t)|\right)\right) \leq C; \qquad \mathbb{P}-a.s.$$

Proof: The proof can be performed as the proof of Lemma 2.4.1. We omit it here. ■

Lemma 2.5.2

*Under **Assumption 2.1** or **Assumption 2.2,** we have*

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\in[0,T]} |y^n(t) - y(t)|^2\right] = 0. \tag{2.19}$$

$$\lim_{n\to\infty} \mathbb{E}\int_0^T |z^n(t) - z(t)|^2 \, \mathrm{d}t = 0. \tag{2.20}$$

Proof: Noting that the coefficient b^n depends explicitly on the control variable $u^n(\cdot)$ and since $d(u^n(\cdot), u(\cdot))$ converges to 0 as n goes to $+\infty$, we may replace $u^n(\cdot)$ by $u(\cdot)$, therefore, the proof goes as in [7, Lemma 1]. ■

Lemma 2.5.3

*Under **Assumption 2.1** or **Assumption 2.2,** the following estimates hold*

$$\lim_{n\to\infty} \mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), u(r)) - b(r, y(r), z(r), u(r))|^2 \, \mathrm{d}r = 0. \tag{2.21}$$

$$\lim_{n\to\infty} \mathbb{E}\int_0^t \left|b_y^n(r, y^n(r), z^n(r), u(r)) - b_y(r, y(r), z(r), u(r))\right|^4 \mathrm{d}r = 0. \tag{2.22}$$

$$\lim_{n\to\infty} \mathbb{E}\int_0^t |b_z^n(r, y^n(r), z^n(r), u(r)) - b_z(r, y(r), z(r), u(r))|^4 \, \mathrm{d}r = 0. \tag{2.23}$$

Proof: We only give the proof under **Assumption 2.1**, the proof under **Assumption 2.2** goes similarly with some suitable changes**.**

First, we shall prove Eq.(2.21). Let $N > 1$, we put $A_n^N := \left\{(r, w); \quad |z^n(r)|^2 + |z(r)|^2 > N^2\right\}$ and $\bar{A}_n^N = \Omega \backslash A_n^N$, then we have

$$\mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), u(r)) - b(r, y(r), z(r), u(r))|^2 \, \mathrm{d}r \leq C(I_1^n + I_2^n), \tag{2.24}$$

where

$$I_1^n = \mathbb{E}\int_0^t |b^n(r,y^n(r),z^n(r),u(r)) - b(r,y(r),z(r),u(r))|^2 \, \mathbb{I}_{A_n^N} \mathrm{d}r,$$
$$I_2^n = \mathbb{E}\int_0^t |b^n(r,y^n(r),z^n(r),u(r)) - b(r,y(r),z(r),u(r))|^2 \, \mathbb{I}_{\bar{A}_n^N} \mathrm{d}r.$$

Since b satisfies (**H.3**), we obtain, using the fact that $|x|^\alpha \leq 1+|x|$,

$$I_1^n \leq K(M,\varphi)\mathbb{E}\int_0^t \left((2+|z^n(r)|^\alpha)^2 + (2+|y(r)|+|z(r)|^\alpha)^2\right) \mathbb{I}_{A_n^N} \mathrm{d}r.$$

In view of the inequality $\mathbb{I}_{A^N} < \dfrac{|z^n(r)|^2+|z(r)|^2}{N^2}$, we get, using Holder's inequality and (2.5),

$$I_1^n \leq \frac{K(M,\varphi)}{N^{2(1-\alpha)}}. \tag{2.25}$$

On the other hand, it is not difficult to see that

$$\begin{aligned} I_2^n &\leq \mathbb{E}\int_0^t |b^n(r,y^n(r),z^n(r),u(r)) - b^n(r,y(r),z(r),u(r))|^2 \, \mathbb{I}_{\bar{A}_n^N} \mathrm{d}r \\ &+ \mathbb{E}\int_0^t |b^n(r,y(r),z(r),u(r)) - b(r,y(r),z(r),u(r))|^2 \, \mathbb{I}_{\bar{A}_n^N} \mathrm{d}r. \end{aligned}$$

Using the fact that b^n is Lipschitz in the ball $B(0,N)$ and the definition of the semi norm (2.7), we get

$$I_2^n \leq L_N^2 \mathbb{E}\int_0^t |y^n(r)-y(r)|^2 + |z^n(r)-z(r)|^2 \mathrm{d}r + \rho_N^2(b^n - b), \tag{2.26}$$

then, plugging (2.25) and (2.26) into (2.24), to obtain

$$\begin{aligned} &\mathbb{E}\int_0^t |b^n(r,y^n(r),z^n(r),u(r)) - b(r,y(r),z(r),u(r))|^2 \mathrm{d}r \\ &\leq C\frac{K(M,\varphi)}{N^{2(1-\alpha)}} + C\rho_N^2(b^n-b) \\ &+ CL_N^2\mathbb{E}\int_0^t \left(|y^n(r)-y(r)|^2 + |z^n(r)-z(r)|^2\right) \mathrm{d}r. \end{aligned}$$

Passing to the limit, successively on n and N, we get (2.21). Next, we only give the proof of Eq.(2.22). The proof of Eq.(2.23) can be performed similarly. A simple computation shows that

$$\mathbb{E}\int_0^t \left|b_y^n(r,y^n(r),z^n(r),u(r)) - b_y(r,y(r),z(r),u(r))\right|^4 \mathrm{d}r \leq C(I_5^n + I_6^n),$$

where

$$I_5^n = \mathbb{E}\int_0^t \left(\left|b_y^n(r,y^n(r),z^n(r),u(r))\right|^4 + |b_y(r,y(r),z(r),u(r))|^4\right) \mathbb{I}_{A_n^N} \mathrm{d}r$$
$$I_6^n = \mathbb{E}\int_0^t \left|b_y^n(r,y^n(r),z^n(r),u(r)) - b_y(r,y(r),z(r),u(r))\right|^4 \mathbb{I}_{\bar{A}_n^N} \mathrm{d}r.$$

Due to the fact that $\mathbb{I}_{A^N} < \dfrac{|z^n(r)|^2 + |z(r)|^2}{N^2}$, using Holder's inequality together with the relation (v) in Lemma 2.4.1 and (2.5), we obtain

$$I_5^n \leq \frac{K(M,\varphi) + \eta_n}{N^2}.$$

On the other hand

$$\begin{aligned} I_6^n &\leq \mathbb{E}\int_0^t \left|b_y^n(r, y^n(r), z^n(r), u(r)) - b_y(r, y^n(r), z^n(r), u(r))\right|^4 \mathbb{I}_{\bar{A}_n^N} \mathrm{d}r. \\ &+ \mathbb{E}\int_0^t |b_y(r, y^n(r), z^n(r), u(r)) - b_y(r, y(r), z(r), u(r))|^4 \mathbb{I}_{\bar{A}_n^N} \mathrm{d}r, \end{aligned}$$

which implies

$$I_6^n \leq \rho_{N,4}^4(b_y^n - b_y) + \mathbb{E}\int_0^t |b_y(r, y^n(r), z^n(r), u(r)) - b_y(r, y(r), z(r), u(r))|^4 \mathbb{I}_{\bar{A}_n^N} \mathrm{d}r.$$

According to Remark 2.1, b_y evaluated at $(y(\cdot), z(\cdot), u(\cdot))$ is bounded. We deduce, using Lebesgue's dominated convergence theorem,

$$\lim_{n\to+\infty} \mathbb{E}\left(\int_0^t |b_y(r, y^n(r), z^n(r), u(r)) - b_y(r, y(r), z(r), u(r))|^4 \mathbb{I}_{\bar{A}_n^N} \mathrm{d}r\right) = 0,$$

passing to the limits successively on n and N one gets (2.22). ■

Lemma 2.5.4

Let $x(\cdot)$ and $x^n(\cdot)$ be respectively the solution of Eqs.(0.13) and (2.18), then under ***Assumption 2.1*** *or* ***Assumption 2.2****, we have*

$$\lim_{n\to\infty} \mathbb{E}\left[|x^n(t) - x(t)|^2\right] = 0. \tag{2.27}$$

Proof: We only give the proof under **Assumption 2.2,** the proof under **Assumption 2.1** is similar. From Eqs.(0.13) and (2.18), we obtain

$$\begin{aligned} \mathbb{E}\left[|x^n(t) - x(t)|^2\right] &\leq C\mathbb{E}\int_0^t \left|b_y^n(r, y^n(r), z^n(r), u^n(r))\,(x^n(r) - x(r))\right|^2 \mathrm{d}r \\ &+ C\mathbb{E}\int_0^t |b_z^n(r, y^n(r), z^n(r), u^n(r))\,(x^n(r) - x(r))|^2 \mathrm{d}r \\ &+ \alpha_1^n(t), \end{aligned} \tag{2.28}$$

where

$$\alpha_1^n(t) = \mathbb{E}\left|g_y(y^n(0)) - g_y(y(0))\right|^2$$
$$+\mathbb{E}\int_0^t \left|\Big(b_y^n(r, y^n(r), z^n(r), u^n(r)) - b_y(r, y(r), z(r), u(r)\Big))x(r)\right|^2 \mathrm{d}r$$
$$+\mathbb{E}\int_0^t \left|(b_z(r, y^n(r), z^n(r), u^n(r)) - b_z(r, y(r), z(r), u(r)))\, x(r)\right|^2 \mathrm{d}r.$$

Since $b_y\,(t, \cdot, \cdot, u^n(t))\,,\ b_z\,(t, \cdot, \cdot, u^n(t))$ and $b\,\,(t, \cdot, \cdot, u^n(t))$ evaluated at $(y^n(t), z^n(t))$ are bounded. Then, taking into account the relation (v) in Lemma 2.4.1, we get

$$\mathbb{E}\left[|x^n(t) - x(t)|^2\right] \leq C\,(2 + \eta_n)\,\mathbb{E}\int_0^t |x^n(r) - x(r)|^2\,\mathrm{d}r + C\alpha_1^n(t). \tag{2.29}$$

Let us prove that $\lim_{n\to\infty} \alpha_1^n(t) = 0$. Since g_y is bounded and continuous, then by Eq.(2.19) and the dominated convergence theorem, we have

$$\lim_{n\to\infty} \mathbb{E}\left|g_y(y^n(0)) - g_y(y(0))\right|^2 = 0. \tag{2.30}$$

Hence, using Cauchy-Schwarz inequality taking account of (2.6), we get

$$\mathbb{E}\int_0^t \left|\Big(b_y^n(r, y^n(r), z^n(r), u^n(r)) - b_y(r, y(r), z(r), u(r))\Big)\, x(r)\right|^2 \mathrm{d}r$$
$$\leq C\left(\mathbb{E}\int_0^t \left|b_y^n(r, y^n(r), z^n(r), u^n(r)) - b_y(r, y(r), z(r), u(r))\right|^4 \mathrm{d}r\right)^{\frac{1}{2}}.$$

By Eq.(2.22), we have

$$\lim_{n\to\infty} \mathbb{E}\int_0^t \left|\Big(b_y^n(r, y^n(r), z^n(r), u^n(r)) - b_y(r, y(r), z(r), u(r)\Big)\, x(r)\right|^2 \mathrm{d}r = 0. \tag{2.31}$$

Similarly

$$\lim_{n\to\infty} \mathbb{E}\int_0^t \left|(b_z^n(r, y^n(r), z^n(r), u^n(r)) - b_z(r, y(r), z(r), u(r)))\, x(r)\right|^2 \mathrm{d}r = 0. \tag{2.32}$$

From Eqs.(2.30), (2.31) and Eq.(2.32), it is easy to see that

$$\lim_{n\to\infty} \alpha_1^n(t) = 0. \tag{2.33}$$

Returning back to (2.29) and using Gronwall's lemma taking account of Eq.(2.33), we obtain Eq.(2.27) by passing to the limit. ■

2.5.2 Necessary Condition for Optimality

To claim and prove the necessary condition of optimality we need the following lemma. Firstly, we set

$$F^n(r) = \left[\mathcal{H}^n(r, y^n(r), z^n(r), x^n(r), u^n(r)) - \mathcal{H}^n(r, y^n(r), z^n(r), x^n(r), v(r))\right],$$

and

$$F(r) = \left[\mathcal{H}(r, y(r), z(r), x(r), u(r)) - \mathcal{H}(r, y(r), z(r), x(r), v(r))\right].$$

Lemma 2.5.5

Assume that ***Assumption 2.1*** *or* ***Assumption 2.2*** *is in force. Then, we have*

$$\lim_{n\to\infty} \mathbb{E}\int_0^t |F^n(r) - F(r)|\,\mathrm{d}r = 0.$$

Proof: A simple computation shows that, using the definition of $\mathcal{H}^n$ and $\mathcal{H}$

$$\begin{aligned}\mathbb{E}\int_0^t |F^n(t) - F(t)|\,\mathrm{d}r &= C\mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), u^n(r))x^n(r) - b(r, y(r), z(r), u(r))x(r)|\,\mathrm{d}r \\ &\quad + C\mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), v(r))x^n(r) - b(r, y(r), z(r), v(r))x(r)|\,\mathrm{d}r \\ &\leq C\left(I_1^n + I_2^n\right),\end{aligned}$$

where

$$I_1^n = \mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), u^n(r))x^n(r) - b(r, y(r), z(r), u(r))x(r)|\,\mathrm{d}r$$

$$I_2^n = \mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), v(r))x^n(r) - b(r, y(r), z(r), v(r))x(r)|\,\mathrm{d}r.$$

Now, let us prove that $\lim\limits_{n\to+\infty} I_1^n = 0$. Applying Schwarz inequality, using the fact that $\mathbb{E}\int_0^t |x(r)|^2\,\mathrm{d}r \leq C$, we get

$$\begin{aligned}I_1^n &\leq \left[\mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), u^n(r))|^2\,\mathrm{d}r\right]^{\frac{1}{2}}\left[\mathbb{E}\int_0^t |x^n(r) - x(r)|^2\,\mathrm{d}r\right]^{\frac{1}{2}} \\ &\quad + C\mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), u^n(r)) - b^n(r, y^n(r), z^n(r), u(r))|^2\,\mathbb{1}_{\{u^n(r)\neq u(r)\}}(r)\,\mathrm{d}r \\ &\quad + C\left[\mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), u(r)) - b(r, y(r), z(r), u(r))|^2\,\mathrm{d}r\right]^{\frac{1}{2}}.\end{aligned} \tag{2.34}$$

Since b^n satisfies **(H.3),** we use Lemma 2.5.1 and the relation (ii) in Lemma 2.4.1, to deduce that $\mathbb{E}\int_0^t |b^n(r, y^n(r), z^n(r), u^n(r))|^2\,\mathrm{d}r$ is bounded, then by Eq.(2.27) the first expression in the right-hand side converges to 0 as $n \to \infty$. On the other hand by using Eq.(2.21), one can confirm the convergence of the third term to 0. We proceed now to estimate the second term, we apply Holder's inequality to get

$$\begin{aligned}
&\mathbb{E}\int_0^t \left|b^n(r,y^n(r),z^n(r),u^n(r)) - b^n(r,y^n(r),z^n(r),u(r))\right|^2 \mathbb{I}_{\{u^n(r)\neq u(r)\}}(r)\,\mathrm{d}r \\
&\leq 2C\left[\mathbb{E}\int_0^t |z^n(r)|^{2\alpha}\,\mathbb{I}_{\{u^n(r)\neq u(r)\}}(r)\,\mathrm{d}r\right]^{\frac{1}{2}} \\
&\leq 2C\left[\mathbb{E}\int_0^t |z^n(r)|^2\,\mathrm{d}r\right]^{\alpha}\left[\mathbb{E}\int_0^t \mathbb{I}_{\{u^n(r)\neq u(r)\}}(r)\,\mathrm{d}r\right]^{1-\alpha} \\
&\leq 2C\left[d\left(u^n(\cdot),u(\cdot)\right)\right]^{1-\alpha}.
\end{aligned}$$

Since $d(u^n(\cdot),u(\cdot))$ converges to 0 as n goes to $+\infty$, the second term in the right-hand side of (2.34) tends to 0. On the other hand, by using similar arguments developed above one can easily show that $\lim\limits_{n\to+\infty} I_2^n = 0$. This completes the proof. ■

Now we are in a position to state and prove the first main result in this paper.

Theorem 2.5.1 (***Necessary optimality conditions for the locally Lipschitz case***)

Let $(u(\cdot),y(\cdot),z(\cdot))$ be an optimal solution of the initial control problem. Then, there exists a unique adapted processes $x(\cdot) \in \mathcal{S}^2([0,T],\mathbb{R}^n)$, solution to the forward stochastic differential equation Eq.(0.13) such that

$$\mathcal{H}(t,y(t),z(t),x(t),u(t)) = \max_{v(\cdot)\in\mathcal{U}_{ad}} \mathcal{H}(t,y(t),z(t),x(t),v(t)); \qquad \mathrm{d}t - a.e,\ P - a.s. \tag{2.35}$$

Proof: To make the main idea of the proof much clear, we start by giving the outlines of the proof:

1) Firstly, since the generator b is differentiable and locally Lipschitz with respect to the state variables, and thus, their derivatives are not bounded, we convert the Problem **(A)** into Problem **(B).**

2) Then, we use a spike variation method to derive the necessary condition of near optimality by handling the Problem **(B).**

3) We get the necessary condition of optimality (2.35), by passing to limits using Lemma 2.5.5.

Now, for each integer n, we suppose that $u^n(\cdot) \in \mathcal{U}_{ad}$ is an optimal control for Problem

(B), in the sense that $\mathcal{J}^n(u^n(\cdot)) \leq \inf_{v(\cdot)\in\mathcal{U}_{ad}} \mathcal{J}^n(v(\cdot))$, and we denote $(y^n(\cdot), z^n(\cdot))$ the solution of BSDE Eq.(2.16) corresponding to $u^n(\cdot)$. Then, we introduce the following spike variation

$$u^{n,\theta}(t) = \begin{cases} v_t \text{ if } t \in [t_0, t_0+\theta], \\ u^n(t) \quad \text{otherwise,} \end{cases}$$

where $0 \leq t_0 \leq T$ is fixed, $\theta > 0$ is sufficient small, and v is an arbitrary $\mathcal{F}_{t_0}-$measurable random variable.

The fact that

$$\tilde{\mathcal{J}}^n(u^n(\cdot)) \leq \tilde{\mathcal{J}}^n(u^{n,\theta}(\cdot)),$$

and

$$d(u^{n,\theta}(\cdot), u^n(\cdot)) \leq \theta,$$

imply that

$$\mathcal{J}^n(u^{n,\theta}(\cdot)) - \mathcal{J}^n(u^n(\cdot)) \geq -(\delta_{n,N})^{\frac{1}{2}}\theta. \tag{2.36}$$

By using standard arguments (see for example [44]), it is easy to show that, the left-hand side of the inequality (2.36) is equal to

$$\mathbb{E}\int_{t_0}^{t_0+\theta} [\mathcal{H}^n(t, y^n(t), z^n(t), x^n(t), u^n(t)) - \mathcal{H}^n(t, y^n(t), z^n(t), x^n(t), v(r))]\,\mathrm{d}r + o(\theta).$$

Dividing the both sides of the inequality (2.36) by θ, we get

$$-(\delta_{n,N})^{\frac{1}{2}} \leq \frac{1}{\theta}\mathbb{E}\int_{t_0}^{t_0+\theta} [\mathcal{H}^n(r, y^n(r), z^n(r), x^n(r), u^n(r)) - \mathcal{H}^n(r, y^n(r), z^n(r), x^n(r), v(r))]\,\mathrm{d}r + \frac{o(\theta)}{\theta}.$$

By using Lemma 2.5.5 and passing to the limits successively on n, N and θ, keeping in mind that t_0 is an arbitrary element of $[0,T]$, we get

$$\mathbb{E}[\mathcal{H}(t, y(t), z(t), x(t), u(t)) - \mathcal{H}(t, y(t), z(t), x(t), v_t)] \geq 0.$$

Now, let $a \in U$ be a deterministic element and B be an arbitrary element of the $\sigma-$algebra $\mathcal{F}_t$, and set

$$\omega(t) = a\mathbb{I}_B + u(t)\mathbb{I}_{\Omega|B}.$$

It is obvious that $\omega(\cdot)$ is an admissible control. Applying the above inequality with $\omega(\cdot)$, we get

$$\mathbb{E}[\mathbb{I}_B(\mathcal{H}(t, y(t), z(t), x(t), u(t)) - \mathcal{H}(t, y(t), z(t), x(t), a))] \geq 0, \qquad \forall B \in \mathcal{F}_t,$$

which implies

$$\mathbb{E}^{\mathcal{F}_t}\left[\mathcal{H}(t,y(t),z(t),x(t),u(t))-\mathcal{H}(t,y(t),z(t),x(t),a)\right]\geq 0.$$

The quantity inside the conditional expectation is $\mathcal{F}_t-$measurable, and thus the result follows immediately. This proves Theorem 4.1. ■

2.5.3 Sufficient Condition of Optimality

In this section, we will prove that under additional hypothesis, a necessary optimality condition Eq.(2.35) becomes sufficient condition of optimality.

Theorem 2.5.2 (*Sufficient optimality conditions for the locally Lipschitz case*)

*Let $(y(\cdot),z(\cdot),u(\cdot))$ be solution of Eq.(2.1), and $x(\cdot)$ is the solutions of the adjoint equation Eq.(0.13), corresponding to $(y(\cdot),z(\cdot),u(\cdot))$. Assume further that $(y,z,u)\rightarrow \mathcal{H}(t,y,z,x,u)$ is convex for a.e. $t\in[0,T]$, $P-a.s.$, $g(.)$ is convex. If the necessary condition of optimality Eq.(2.35) is satisfied, then $(y(\cdot),z(\cdot),u(\cdot))$ is an optimal triplet for Problem **(A)**, in the sense that*

$$\mathcal{J}(u(\cdot))\leq \inf_{v(\cdot)\in\mathcal{U}_{ad}}\mathcal{J}(v(\cdot)).$$

Proof: Let $u(\cdot)\in\mathcal{U}_{ad}$ be candidate to be an optimal control. For any $v(\cdot)\in\mathcal{U}_{ad}$, we have

$$\mathcal{J}(v(\cdot))-\mathcal{J}(u(\cdot))=\mathbb{E}\left[g(y^v(0))-g(y^u(0))\right].$$

Since g is convex, then

$$g(y^v(0))-g(y^u(0))\geq g_y(y^u(0))(y^v(0)-y^u(0)).$$

Thus,

$$\mathcal{J}(v(\cdot))-\mathcal{J}(u(\cdot))\geq\mathbb{E}\left[g_y(y^u(0))(y^v(0)-y^u(0))\right].$$

Using the fact that $x^u(0)=g_y(y^u(0))$, we have

$$\mathcal{J}(v(\cdot))-\mathcal{J}(u(\cdot))\geq\mathbb{E}\left[x^u(0)(y^v(0)-y^u(0))\right].$$

By applying Ito's formula respectively $x^u(\cdot)\,(y^v(\cdot) - y^u(\cdot))$, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[x^u(0)\,(y^v(0) - y^u(0))\right] = \\
&\qquad\qquad \mathbb{E}\int_0^T \mathcal{H}_y\left(t, y^u(t), z^u(t), x^u(t), u(t)\right)\left(y^v(t) - y^u(t)\right) \mathrm{d}t \\
&\qquad\qquad + \mathbb{E}\int_0^T p_t^u \left[b\left(t, y^v(t), z^v(t), v(t)\right) - b\left(t, y^u(t), z^u(t), u(t)\right)\right] \mathrm{d}t \\
&\qquad\qquad + \mathbb{E}\int_0^T \mathcal{H}_z\left(t, y^u(t), z^u(t), x^u(t), u(t)\right)\left(z^v(t) - z^u(t)\right) \mathrm{d}t.
\end{aligned}
$$

Then,

$$
\begin{aligned}
&\mathcal{J}(v(\cdot)) - \mathcal{J}(u(\cdot)) \\
&\geq \mathbb{E}\int_0^T \left[\mathcal{H}\left(t, y^v(t), z^v(t), x^v(t), v(t)\right) - \mathcal{H}\left(t, y^u(t), z^u(t), x^u(t), u(t)\right)\right] \mathrm{d}t \\
&- \mathbb{E}\int_0^T \mathcal{H}_y\left(t, y^u(t), z^u(t), x^u(t), u(t)\right)\left(y^v(t) - y^u(t)\right) \mathrm{d}t \\
&- \mathbb{E}\int_0^T \mathcal{H}_z\left(t, y^u(t), z^u(t), x^u(t), u(t)\right)\left(z^v(t) - z^u(t)\right) \mathrm{d}t.
\end{aligned} \tag{2.37}
$$

Since $\mathcal{H}$ is convex in (y, z, u), then by using the Clarke's generalized gradient of $\mathcal{H}$ evaluated at $(y(\cdot), z(\cdot), u(\cdot))$ and the necessary optimality conditions Eq.(2.35), it follows by [46, Lemmas 2.2 and 2.3], that

$$
\begin{aligned}
0 \leq\ & \mathcal{H}\left(t, y^v(t), z^v(t), p^v(t), v(t)\right) - \mathcal{H}\left(t, y^u(t), z^u(t), p^u(t), u(t)\right) \\
& - \mathcal{H}_y\left(t, y^u(t), z^u(t), p^u(t), u(t)\right)\left(y^v(t) - y^u(t)\right) \\
& - \mathcal{H}_z\left(t, y^u(t), z^u(t), p^u(t), u(t)\right)\left(z^v(t) - z^u(t)\right).
\end{aligned}
$$

We conclude, by replacing the above inequality into (2.37),

$$
\mathcal{J}(v(\cdot)) - \mathcal{J}(u(\cdot)) \geq 0.
$$

Theorem 2.5.2 is proved. ■

Chapter 3

A Stochastic Maximum Principle in Optimal Control of FBSDE with Irregular Coefficients

3.1 Introduction

In this chapter, we outline the necessary conditions for the optimality of a control problem associated with a forward-backward stochastic differential equation featuring irregular coefficients. Our presentation unfolds in several steps. Initially, we articulate the problem statement and introduce our main result. Subsequently, we define a family of smooth control problems designed to approximate the original one. Leveraging distributional derivatives of the coefficients and the Bouleau-Hirsch flow property, we then proceed to define the adjoint process on an extension of the initial probability space. Finally, we establish the stochastic maximum principle.

3.2 Problem Statement and the Main Result

3.2.1 Formulation of Control Problem

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, where $\Omega = \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n)$ be the space, of continous functions $\omega(0) = 0$, endowed with the topology of uniform convergence on compact subsets of $\mathbb{R}_+$. Let $\mathcal{F}$ be the Borel $\sigma-$field over Ω, $\mathbb{P}$ be the Wiener measure on $(\Omega, \mathcal{F})$.and $(\mathcal{F}_t)_{t\geq 0}$ the filtration of coordinate process augmented with $\mathbb{P}-$null sets of $\mathcal{F}$. We define the canonical process $W_t(\omega) = \omega(t)$, $t \geq 0$. Thus, on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, $\left(\tilde{W}_t\right)_{t\geq 0}$ is a Brownian motion.

Let T be a strictly positive real number and U a non-empty set of $\mathbb{R}^k$.

We consider a stochastic control problem, where the control domain need not be convex and the system is governed by the following controlled forward-backward stochastic differential equation (FBSDE for short) of the type

$$\begin{cases} dX(t) = & b\left(t, X(t), v(t)\right) \mathrm{d}t + \sigma\left(t, X(t)\right) \mathrm{d}W_t, \\ X(0) = & X, \\ dY(t) = & -f\left(t, X(t), Y(t), Z(t), v(t)\right) \mathrm{d}t + Z(t) \mathrm{d}W_t, \\ Y(T) = & h\left(X(T)\right), \end{cases} \tag{3.1}$$

where b, σ, f, g and γ are given maps. The control variable $v = (v(t))$ is an $\mathcal{F}_t$-adapted process with values in some set U of $\mathbb{R}^k$. We denote by $\mathcal{U}_{ad}$ the set of all admissible controls.

The expected cost on the time interval $[0, T]$ is

$$J\left(v\left(\cdot\right)\right) = \mathbb{E}\left[\gamma\left(X(T)\right) + g\left(y(0)\right) + \int_0^T l\left(t, X(t), Y(t), Z(t), v(t)\right) \mathrm{d}t\right]. \tag{3.2}$$

In the above statement,

$$\begin{aligned} &b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n, \\ &\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times d}, \\ &f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \times U \to \mathbb{R}^m, \\ &l : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \times U \to \mathbb{R}, \\ &h : \mathbb{R}^n \to \mathbb{R}^m, \gamma : \mathbb{R}^n \to \mathbb{R}, g : \mathbb{R}^m \to \mathbb{R}. \end{aligned}$$

The optimal control problem is to minimize the cost function $J\left(v\left(\cdot\right)\right)$ over $\mathcal{U}_{ad}$. An admissible control $u\left(\cdot\right)$ is called optimal if it satisfies

$$\mathcal{J}\left(u\left(\cdot\right)\right) = \inf_{v \in \mathcal{U}_{ad}} \mathcal{J}\left(v\left(\cdot\right)\right). \tag{3.3}$$

Equation (3.1) is called the state equation, the solution $\left(X\left(\cdot\right), Y\left(\cdot\right), Z\left(\cdot\right)\right)$ corresponding to $u\left(\cdot\right)$ is called an optimal trajectory.

In what follows, we assume that the coefficients satisfy the following assumptions

($\mathbb{A}$.1) b is bounded function and Lipschitz continuous of x, v with a Lipschitz constant L and σ is bounded functions and Lipschitz continuous in x with a Lipschitz constant L.

($\mathbb{A}$.2) h is continuously differentiable of x and its derivative is bounded. l is continuous and continuously differentiable in x, y, z, v and the derivatives of l are bounded.

($\mathbb{A}$.3) γ and g are continuously differentiable and the derivatives in γ and g are bounded.

($\mathbb{A}$.4) f is continuous and continuously differentiable in x, y, z and their partial derivatives are bounded.

Note that since the functions $b(t, x, v)$ and $\sigma^j(t, x)$ are Lipschitz continuous of x, then by Rademarcher's theorem See [23], they are differentiable almost everywhere (in the sense of the Lebesgue measure). Let us denote by b_x, σ_x^j any Borel measurable functions such that

$$\frac{\partial \sigma^j}{\partial x}(t, x) = \sigma_x^j(t, x) \ \ \mathrm{d}x - a.e;$$
$$\frac{\partial b}{\partial x}(t, x, v) = b_x(t, x, v) \quad \mathrm{d}x - a.e.$$

These almost everywhere derivatives are bounded by the Lipschitz constant L. Let us assume that b_x is continuous in v uniformly in (t, x). Under the assumptions ($\mathbb{A}$.1) and ($\mathbb{A}$.2), for every $v(\cdot) \in \mathcal{U}_{ad}$, equation (3.1) admits a unique adapted solution $(X(\cdot), Y(\cdot), Z(\cdot)) \in S^4([0,T]; \mathbb{R}^n) \times S^4([0,T]; \mathbb{R}^m) \times M^4([0,T]; \mathbb{R}^{m\times d})$.

From well-known results on SDE and BSDE, we have the following lemma.

Lemma 3.2.1

for $p \geq 2$, we have the following estimation

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|X(t)|^p + \sup_{0\leq t\leq T}|Y(t)|^p + \left(\int_0^T |Z(t)|^2 \mathrm{d}t\right)^{\frac{p}{2}}\right]$$
$$\leq C\mathbb{E}\left[1 + |x|^p + \mathbb{E}\int_0^T |b(t,0,v(t))|^p \mathrm{d}t + \mathbb{E}\int_0^T |\sigma(t,0)|^p \mathrm{d}t + \int_0^T |f(t,0,,0,v(t))|^p \mathrm{d}t\right]. \tag{3.4}$$

Let h be a continuous positive function on $\mathbb{R}^n$ satisfying

$$\int h(x)\mathrm{d}x = 1 \quad \text{and} \quad \int |x|^2 h(x)\mathrm{d}x < +\infty.$$

Define the space of functions

$$D :\, = \left\{ f \in L^2(h\mathrm{d}x); \quad \text{such that} \quad \frac{\partial f}{\partial x_j} \in L^2(h\mathrm{d}x), j = 1, ..., n \right\},$$

where $\frac{\partial f}{\partial x_j}$ denotes the derivative of f in the sense of distributions. Equipped with the norm

$$\|f\|_D : = \left[\int f^2 h \mathrm{d}x + \sum_{1 \leq j \leq n} \int \left(\frac{\partial f}{\partial x_j} \right)^2 h \mathrm{d}x \right]^{1/2},$$

D is a Hilbert space, which is a classical Dirichlet space. Moreover, D is a subset of the Sobolev space $\mathcal{W}^1_{loc}(\mathbb{R}^n)$.

Let $\tilde{\Omega} : = \mathbb{R}^n \times \Omega$ and $\tilde{\mathcal{F}}$ the Borel σ-field over $\tilde{\Omega}$ and $\tilde{\mathbb{P}} : = h\mathrm{d}x \otimes \mathbb{P}$, $\tilde{W}_t(x, w) : = W_t(w)$ and $\left(\tilde{\mathcal{F}}_t\right)_{t \geq 0}$ be the natural filtration of $\tilde{W}_t$ augmented with $\tilde{\mathbb{P}}$-negligible sets of $\tilde{\mathcal{F}}$. It is clear that on $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left(\tilde{\mathcal{F}}_t\right)_{t \geq 0}, \tilde{\mathbb{P}}\right)$, $\left(\tilde{W}_t\right)_{t \geq 0}$ is a Brownian motion. We introduce the process $\left(\tilde{X}(t), \tilde{Y}(t), \tilde{Z}(t)\right)$ defined on the enlarged space $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left(\tilde{\mathcal{F}}_t\right)_{t \geq 0}, \tilde{\mathbb{P}}, \tilde{W}_t\right)$, solution of the forward-backward stochastic differential equation

$$\begin{cases} d\tilde{X}(t) = & b\left(t, \tilde{X}(t), \tilde{v}(t)\right) \mathrm{d}t + \sigma\left(t, \tilde{X}(t)\right) \mathrm{d}\tilde{W}_t, \\ \tilde{X}(0) = & x, \\ d\tilde{Y}(t) = & -f\left(t, \tilde{X}(t), \tilde{Y}(t), \tilde{Z}t), \tilde{v}(t)\right) \mathrm{d}t + \tilde{Z}(t)\mathrm{d}\tilde{W}_t, \\ \tilde{Y}(T) = & h\left(\tilde{X}(T)\right). \end{cases} \tag{3.5}$$

Since the coefficients are Lipschitz continuous and grow at most linearly, FBSDE (3.5) has a unique $\tilde{\mathcal{F}}_t-$adapted solution with continuous trajectories.

Equations (3.1) and (3.5) are almost the same except that uniqueness for (3.5) is slightly weaker. One can easily prove that the uniqueness implies that for each $t \geq 0$, $\left(\tilde{X}(t), \tilde{Y}(t), \tilde{Z}(t)\right) = (X(t), Y(t), Z(t)), \tilde{\mathbb{P}}-a.s.$

Now we introduce the adjoint equations and the Hamiltonian function for our problem. The adjoint equations are defined by

$$\begin{cases} dP(t) & = \left[f_y^*(t,X(t),Y(t),Z(t),u(t))P(t)-l_y(t,X(t),Y(t),Z(t),u(t))\right]\mathrm{d}t \\ & +\left[f_z^*(t,X(t),Y(t),Z(t),u(t))P(t)-l_z(t,X(t),Y(t),Z(t),u(t))\right]\mathrm{d}\tilde{W}_t, \\ P(0) & = -g_y^*(y(0)), \\ -dQ(t) & = \left[-f_x^*(t,X(t),Y(t),Z(t),u(t))P(t)+b_x^*(t,X(t),u(t))Q(t)\right. \\ & \left.-\sigma_x^*(t,X(t))R(t)+l_x(t,X(t),Y(t),Z(t),u(t))\right]\mathrm{d}t-R(t)\,\mathrm{d}\tilde{W}_t, \\ Q(T)= & \gamma_x^*(X(T))-h_x^*(X(T))P(T), \end{cases} \tag{3.6}$$

and the Hamiltonian function is given by

$$\begin{aligned}\mathcal{H}(t,x,y,z,P,Q,R,u) &:= Qb(t,x,u)-Pf(t,x,y,z,u)\\ &\quad+R\sigma(t,x)+l(t,x,y,z,u),\end{aligned}$$

where $\mathcal{H}:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U\times\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\to\mathbb{R}$.

Let us recall the Bouleau–Hirsch flow property which will be used in the sequel.

Lemma 3.2.2 (*Bouleau–Hirsch flow property*)

Let $\tilde{X}$ be the solution of the forward component of FBSDE (3.5) on $\left(\tilde{\Omega},\tilde{\mathcal{F}},\left(\tilde{\mathcal{F}}_t\right)_{t\geq 0},\tilde{\mathbb{P}},\tilde{W}_t\right)$. Then, for $\tilde{\mathbb{P}}-$almost every w

1) For all $t\geq 0$, $x\to\tilde{X}_t(w)$ is in D^n.

2) For every $t\geq 0$, the image measure of $\tilde{\mathbb{P}}$ through the map $\tilde{X}(t)$ is absolutely continuous with respect to the Lebesgue measure.

Proof: The proof is similar to the deterministic case see [25] ∎

3.3 A Maximum Principle for a Family of Perturbed Control Problems

Let φ be a non-negative smooth function defined on $\mathbb{R}^n$, with support in the unit ball such that,

$$\int_{\mathbb{R}^n}\varphi(x)\,\mathrm{d}x=1.$$

Define the following smooth functions by convolution:

$$b^k(t,x,v) = k^n \int_{\mathbb{R}^n} b\,(t, x-y, v)\,\varphi\,(ky)\,\mathrm{d}y,$$

$$\sigma^k(t,x) = k^n \int_{\mathbb{R}^n} \sigma\,(t, x-y)\,\varphi\,(ky)\,\mathrm{d}y.$$

In the next lemma, we list some properties satisfied by these functions.

Lemma 3.3.1

(a) $b^k(t,x,v), \sigma^{j,k}(t,x)$ are Borel measurable, bounded functions and M-Lipschitz continuous.

(b) There exists a constant $C>0$ such that $\forall t \in [0,T]$:

$$\left|\sigma^{j,k}(t,x) - \sigma^j(t,x)\right| + \left|b^k(t,x,v) - b(t,x,v)\right| \leq C/k = \varepsilon_k. \tag{3.7}$$

(c) $b^k(t,x,v), \sigma^{j,k}(t,x)$ are C^1 functions in x and $\forall t \in [0,T] \times A$:

$$\lim_{k\to+\infty} \sigma_x^{j,k}(t,x) = \sigma_x^j(t,x) \qquad \mathrm{d}x \quad a.e,$$

$$\lim_{k\to+\infty} b_x^k(t,x,v) = b_x(t,x,v) \quad \mathrm{d}x \ a.e,$$

(d) For every $p \geq 1$ and $M>0$

$$\lim_{k\to+\infty} \int\int_{[0,T]\times B(0,M)} \sup_{a\in A} \left|b_x^k(t,x,v) - b_x(t,x,v)\right|^p \mathrm{d}t\mathrm{d}x = 0,$$

where $B(0,M)$ denotes a ball in $\mathbb{R}^n$ of radius M

Proof: Statements $(a), (b)$ and (c) are classical facts (see [29] for the proof). (d) is proved as in [8]. ■

Now, let us consider $\left(X_1^k(\cdot), Y_1^k(\cdot), Z_1^k(\cdot)\right)$ the solutions of FBSDE defined on the enlarged probability space $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \left(\tilde{\mathcal{F}}_t\right)_{t\geq 0}, \tilde{\mathbb{P}}, \tilde{W}_t\right)$ by

$$\begin{cases} dX_1^k(t) = & b^k\left(t, X_1^k(t), v(t)\right)\mathrm{d}t + \sigma^k\left(t, X_1^k(t)\right)\mathrm{d}\tilde{W}_t, \\ X_1^k(0) = & x, \\ dY_1^k(t) = & -f\left(t, X_1^k(t), Y_1^k(t), Z_1^k(t), v(t)\right)\mathrm{d}t + Z_1^k(t)\,\mathrm{d}\tilde{W}_t, \\ Y_1^k(T) = & h\left(X_1^k(T)\right). \end{cases} \tag{3.8}$$

The corresponding cost is given by:

$$J^k(v(\cdot)) = \tilde{\mathbb{E}}\left[\gamma\left(X_1^k(T)\right) + g\left(Y_1^k(0)\right) + \int_0^T l\left(t, X_1^k(t), Y_1^k(t), Z_1^k(t), v(t)\right) \mathrm{d}t\right] \tag{3.9}$$

where b^k and σ^k are the regularized functions of b and σ.

Lemma 3.3.2

Let $u(\cdot) \in \mathcal{U}_{ad}$, $(X(\cdot), Y(\cdot), Z(\cdot))$ and $\left(X_1^k(\cdot), Y_1^k(\cdot), Z_1^k(\cdot)\right)$ the solutions of (3.1) and (3.8) corresponding to the control $v(\cdot)$ then the following estimates hold:

$$\tilde{\mathbb{E}}\left[\sup_{0\le t\le T}\left|X_1^k(t) - X(t)\right|^2\right] + \tilde{\mathbb{E}}\left[\sup_{0\le t\le T}\left|Y_1^k(t) - Y(t)\right|^2\right] + \tilde{\mathbb{E}}\int_0^T \left|Z_1^k(t) - Z(t)\right|^2 dt \le \frac{C}{k}, \tag{3.10}$$

$$\left|\mathcal{J}^k(v) - \mathcal{J}(v)\right| \le \frac{C}{k},$$

where C is a positive constant.

Proof: The proof of (i) is similar to the deterministic case see [25]. Item (ii) is proved by using the approximation (3.7) and Lemma (3.2.1) ■

Let $u(\cdot)$ be an optimal control for the initial control problem, that is

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

Note that $u(\cdot)$ is not necessarily optimal for the new perturbed control problem, according to Lemma 3.3.2, there exists a sequence (δ_k) of positive real numbers converging to 0 such that:

$$\mathcal{J}^k(u(\cdot)) \le \inf_{v(\cdot)\in\mathcal{U}_{ad}} \mathcal{J}^k(v(\cdot)) + \delta_k.$$

Let us consider the metric d defined by (2.14) in Chapter 2. Then, Ekeland's lemma applied to the continuous bounded functional $J^k(u)$ shows that there exists an admissible control u^k such that:

$$d(u^k(\cdot), u(\cdot)) \le (\delta_k)^{\frac{1}{2}},$$

and

$$\tilde{\mathcal{J}}^k(u^k(\cdot)) \le \tilde{\mathcal{J}}^k(v(\cdot)) \text{ for any } v(\cdot) \in \mathcal{U}_{ad},$$

where

$$\tilde{\mathcal{J}}^k(v(\cdot)) = \mathcal{J}^k(v(\cdot)) + (\delta_k)^{\frac{1}{2}} \cdot d(v(\cdot), u^k(\cdot)). \tag{3.11}$$

This means that $u^k(\cdot)$ is optimal for (3.8) with the new cost function $\tilde{\mathcal{J}}^k$. For each integer k, we denote by $(X^k(\cdot), Y^k(\cdot), Z^k(\cdot))$ the unique solutions of the FBSDEs controlled by $u^k(\cdot)$ of the type

$$\begin{cases} dX^k(t) = & b^k\left(t, X^k(t), u^k(t)\right) \mathrm{d}t + \sigma^k\left(t, X^k(t)\right) \mathrm{d}\tilde{W}_t, \\ X^k(0) = & x, \\ dY^k(t) = & -f\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) \mathrm{d}t + Z^k(t)\, \mathrm{d}\tilde{W}_t, \\ Y^k(T) = & h\left(X^k(T)\right), \end{cases} \tag{3.12}$$

and its corresponding cost is given by

$$\mathcal{J}^k(v(\cdot)) = \mathbb{E}\left[\gamma\left(X^k(T)\right) + g\left(Y^k(0)\right) + \int_0^T l\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) \mathrm{d}t\right]. \tag{3.13}$$

Now, we introduce the following adjoint equation

$$\begin{cases} dP^k(t) & = \left[f_y^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) P^k(t) - l_y^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right)\right] \mathrm{d}t \\ & \quad + \left[f_z^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) P^k(t) - l_z^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right)\right] \mathrm{d}\tilde{W}_t \\ dQ^k(t) & = \left[f_x^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) P^k(t) - b_x^{*,k}\left(t, X^k(t), u^k(t)\right) Q^k(t)\right. \\ & \quad \left. -\sigma_x^{*,k}\left(t, X^k(t)\right) R^k(t) + l_x\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right)\right] \mathrm{d}t + R^k(t)\, \mathrm{d}\tilde{W}_t \\ P^k(0) & = -g_y\left(Y^k(0)\right), \; Q^k(T) = \gamma_x\left(X^k(T)\right) - h_x^*(X^k(T)) P^k(T). \end{cases} \tag{3.14}$$

Under the Assumptions $(\mathbb{A}.1)$-$(\mathbb{A}.4)$ It is easy to see that the FBSDE (3.14) admits a unique solution $\left(P^k(\cdot), Q^k(\cdot), R^k(\cdot)\right) \in S^4([0,T];\mathbb{R}^m) \times S^4([0,T];\mathbb{R}^n) \times M^4\left([0,T];\mathbb{R}^{n\times d}\right)$.

We turn our attention to proving the stochastic maximum principle for control Problem {Eq.(3.12), (3.11)}. For this end we define the following family of perturbed controls $u^{k,\theta}(\cdot)$

$$u^{k,\theta}(t) = \begin{cases} v_{t_0} \text{ if } t \in [t_0, t_0 + \theta], \\ u^k(t) \quad \text{otherwise}, \end{cases} \tag{3.15}$$

where $t_0 \in [0, T)$ is an fixed time, $\theta > 0$ is sufficient small, and v_{t_0} is an arbitrary $\mathcal{F}_{t_0}-$measurable random variable.

Since u^k is optimal for $\tilde{\mathcal{J}}^k$ and the functions b^k, σ^k are smooth enough, we can prove the following proposition by utilizing the same steps as in the proof of Theorem 1.4.1 in chapter 1.

Proposition 3.3.1

For each $k \in \mathbb{N}$, there exists $u^k(\cdot) \in \mathcal{U}_{ad}$ with $\left(X^k(\cdot), Y^k(\cdot), Z^k(\cdot)\right)$ the corresponding trajectory and $\left(P^k(\cdot), Q^k(\cdot), R^k(\cdot)\right)$, the solution of (3.14), such that for every $v(\cdot) \in U$

$$\begin{aligned}&\tilde{\mathbb{E}}\int_0^T \left[\mathcal{H}^k\left(t, X^k(t), Y^k(t), Z^k(t), u^{k,\theta}(t), P^k(t), Q^k(t), R^k(t)\right)\right]\\ &-\tilde{\mathbb{E}}\int_0^T \left[\mathcal{H}^k\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t), P^k(t), Q^k(t), R^k(t)\right)\right] \geq o(\theta) - \theta C(\delta_k)^{\frac{1}{2}}.\end{aligned} \tag{3.16}$$

3.3.1 Estimation Between Two Solutions and some Technical Results

Lemma 3.3.3

Let $u(\cdot) \in U_{ad}$, $(X(\cdot), Y(\cdot), Z(\cdot))$ and $(X^k(\cdot), Y^k(\cdot), Z^k(\cdot))$ the solutions of (3.1) and (3.12), then under Assumptions ($\mathbb{A}$.1)-($\mathbb{A}$.4) the following estimates hold

$$\lim_{k\to\infty} \tilde{\mathbb{E}}\left[\sup_{0\leq t\leq T}\left|X^k(t) - X(t)\right|^2\right] = 0, \tag{3.17}$$

$$\lim_{k\to\infty} \tilde{\mathbb{E}}\left[\sup_{0\leq t\leq T}\left|Y^k(t) - Y(t)\right|^2\right] = 0, \tag{3.18}$$

$$\lim_{k\to\infty} \tilde{\mathbb{E}}\int_0^T \left|Z^k(t) - Z(t)\right|^2 \mathrm{d}t = 0. \tag{3.19}$$

Proof: By squaring, taking expectation and Burkholder–Davis–Gundy inequality, we obtain

$$\tilde{\mathbb{E}}\left[\sup_{0\leq t\leq T}\left|X^k(t) - X(t)\right|^2\right] \leq C\left(M_1 + M_2 + M_3\right),$$

where,

$$M_1 = \tilde{\mathbb{E}}\left(\int_0^t \left|b^k\left(r, X^k(r), u^k(r)\right) - b^k\left(r, X^k(r), u(r)\right)\right|^2 \mathbb{I}_{\left\{u^k(r)\neq u(r)\right\}}\mathrm{d}r\right),$$

$$M_2 = \tilde{\mathbb{E}}\left(\int_0^t \left|b^k\left(r, X^k(r), u(r)\right) - b^k(r, X(r), u(r))\right|^2 + \left|\sigma^k\left(r, X^k(r)\right) - \sigma^k(r, X(r))\right|^2 \mathrm{d}r\right),$$

$$M_3 = \tilde{\mathbb{E}}\left(\int_0^t \left|b^k(r, X(r), u(r)) - b(r, X(r), u(r))\right|^2 + \left|\sigma^k(r, X(r)) - \sigma(r, X(r))\right|^2 \mathrm{d}r\right).$$

By the boundedness of the derivative b^k and σ^k and the fact that $d\left(u^k(\cdot), u(\cdot)\right) \to 0$ as $k \to +\infty$, we obtain $\lim_{k\to\infty} M_1^k = 0$. Then, b^k and σ^k are Lipschitz continuous, we have

$$M_2 \leq C\tilde{\mathbb{E}}\int_0^t \sup_{0\leq r\leq t}\left|X^k(r) - X(r)\right|^2 \mathrm{d}r.$$

By (3.7), we have $\lim_{k\to\infty} M_3 = 0 = 0$. Then by Gronwall inequality, we obtain (3.17).

Applying Ito's formula to $\left|Y^k(t) - Y(t)\right|^2$, we obtain by Holder inequality and Burkholder–Davis–Gundy inequality

$$\begin{aligned}
&\tilde{\mathbb{E}}\left(\sup_{0\leq t\leq T}\left|Y^k(t) - Y(t)\right|^2\right) + \tilde{\mathbb{E}}\int_0^T \left|Z^k(t) - Z(t)\right|^2 \mathrm{d}t \\
&\leq \tilde{\mathbb{E}}\left|h\left(X^k(T)\right) - h(X(T))\right|^2 + C\tilde{\mathbb{E}}\int_0^T \sup_{t\leq r\leq T}\left|Y^k(t) - Y(t)\right|^2 \mathrm{d}t.
\end{aligned}$$

By Gronwall inequality, (3.17) and dominated convergence theorem, we obtain (3.18) and (3.19). ∎

The following technical Lemma is needed to prove Lemma.3.3.5.

Lemma 3.3.4

(a) $\tilde{\mathbb{E}}\left[\int_0^T \left|b_x^k\left(t, X^k(t), u^k(t)\right) - b_x(t, X(t), u(t))\right|^4 \mathrm{d}t\right] \to 0$ *as* $k \to +\infty$,

(b) For every $1 \leq j \leq d$, $\tilde{\mathbb{E}}\left[\int_0^T \left|\sigma_x^k\left(t, X^k(t)\right) - \sigma_x(t, X(t))\right|^4 \mathrm{d}t\right] \to 0$ *as* $k \to +\infty$.

Proof: Let us prove the first limit, we have

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\int_0^T \left|b_x^k\left(t, X^k(t), u^k(t)\right) - b_x(t, X(t), u(t))\right|^4 \mathrm{d}t\right] \\
&\leq C\left\{I_1^k + I_2^k + I_3^k\right\},
\end{aligned}$$

where

$$\begin{aligned}
I_1^k &= \tilde{\mathbb{E}}\left[\int_0^T \left|b_x^k\left(t, X^k(t), u^k(t)\right) - b_x\left(t, X^k(t), u(t)\right)\right|^4 \mathbb{1}_{\left\{u^k(t)\neq u(t)\right\}} \mathrm{d}t\right], \\
I_2^k &= \tilde{\mathbb{E}}\left[\int_0^T \left|b_x^k\left(t, X^k(t), u(t)\right) - b_x\left(t, X^k(t), u(t)\right)\right|^4 \mathrm{d}t\right], \\
I_3^k &= \tilde{\mathbb{E}}\left[\int_0^T \left|b_x\left(t, X^k(t), u(t)\right) - b_x(t, X(t), u(t))\right|^4 \mathrm{d}t\right].
\end{aligned}$$

According to the boundedness of the derivative b_x^k by the Lipschitz constant and the fact that $d(u^k(\cdot), u(\cdot))$ converges to 0 as k goes to $+\infty$, we obtain $\lim_{k\to\infty} I_1^k = 0$.

Moreover, we have

$$I_2^k = \int_0^T \int_{\mathbb{R}^d} \sup_{a\in U} \left| b_x^k (t, y(t), a) - b_x (t, y(t), a) \right|^4 \rho_t^k(y) \mathrm{d}y \mathrm{d}t.$$

Let us show that, for all $t \in [0, T]$,

$$\lim_{k\to\infty} \int_{\mathbb{R}^d} \sup_{a\in U} \left| b_x^k (t, y, a) - b_x (t, y, a) \right|^4 \rho_t^k(y) \mathrm{d}y \mathrm{d}t = 0.$$

where $\rho_t^k(y)$ denotes the density of $X^k(t)$ with respect to Lebesgue measure

For each $p > 0$,

$$\tilde{\mathbb{E}} \left[\sup_{0\le t\le T} \left| x^k(t) \right|^p \right] < \infty.$$

Thus,

$$\lim_{R\to+\infty} \tilde{P} \left(\sup_{0\le t\le T} \left| X^k(t) \right| > R \right) = 0,$$

then it is enough to show that for every $R > 0$,

$$\lim_{k\to+\infty} \int_{B(0,R)} \sup_{a\in U} \left| b_x^k (t, y, a) - b_x (t, y, a) \right|^4 \rho_t^k(y) \mathrm{d}y = 0.$$

According to Lemma (3.3.1)

$$\sup_{a\in U} \left| b_x^k (t, y, a) - b_x (t, y, a) \right|^4 \to 0 \quad \mathrm{d}y\text{-}a.e,$$

at least for a subsequence. Then, by Egorov's theorem, for every $\delta > 0$, there exists a measurable set F with $\lambda(F) < \delta$, such that $\sup_{a\in U} \left| b_x^k (t, y, a) - b_x (t, y, a) \right|$ converges uniformly to 0 on the set F^c. Note that, since the Lebesgue measure is regular, F may be chosen closed. This implies that

$$\begin{aligned}
&\lim_{k\to+\infty} \int_{F^c} \sup_{a\in U} \left| b_x^k (t, y, a) - b_x (t, y, a) \right|^4 \rho_t^k(y) \mathrm{d}y \\
&\le \lim_{k\to+\infty} \left(\sup_{y\in F^c} \sup_{a\in U} \left| b_x^k (t, y, a) - b_x (t, y, a) \right|^4 \right) = 0.
\end{aligned}$$

Now, by using the boundedness of the derivatives b_x^k, b_x we have

$$\begin{aligned}
&\int_F \sup_{a\in U} \left| b_x^k (t, y, a) - b_x (t, y, a) \right|^4 \rho_t^k(y) \mathrm{d}y \\
&= \tilde{\mathbb{E}} \left[\sup_{a\in U} \left| b_x^k \left(t, X^k(t), a\right) - b_x \left(t, X^k(t), a\right) \right|^4 \mathbb{I}_{\{X^k(t)\in F\}} \right] \\
&\le 2M^4 \tilde{P} \left(X^k(t) \in F \right).
\end{aligned}$$

Since $X^k(t)$ converges to $X(t)$ in probability, then in distribution. Then, using the Portmanteau-Alexandorv Theorem we get

$$\begin{aligned}
&\lim \int_F \sup_{a\in U} \left| b_x^k(t,y,a) - b_x(t,y,a) \right|^4 \rho_t^k(y)\mathrm{d}y \\
&\leq 2M^4 \limsup \tilde{P}\left(X^k(t) \in F\right). \\
&\leq 2M^4 \tilde{P}(X(t) \in F). = 2M^4 \int_F \rho_t(y)\mathrm{d}y < \varepsilon,
\end{aligned}$$

where $\rho_t(y)$ denotes the density of $X(t)$ with respect to Lebesgue measure.

Now, since

$$\begin{aligned}
&\int_{B(0,R)} \sup_{a\in U} \left| b_x^k(t,y,a) - b_x(t,y,a) \right|^4 \rho_t^k(y)\mathrm{d}y \\
&= \int_F \sup_{a\in U} \left| b_x^k(t,y,a) - b_x(t,y,a) \right|^4 \rho_t^k(y)\mathrm{d}y \\
&+ \int_{F^c} \sup_{a\in U} \left| b_x^k(t,y,a) - b_x(t,y,a) \right|^4 \rho_t^k(y)\mathrm{d}y,
\end{aligned}$$

we get $\lim_{k\to+\infty} I_2^k = 0$.

Let $k_0 \geq 0$ be a fixed integer, then it holds that $I_3^k \leq C\left(J_1^{k_0} + J_2^{k_0} + J_3^{k_0}\right)$, where

$$\begin{aligned}
J_1^{k_0} &= \tilde{\mathbb{E}}\left[\int_0^T \left| b_x\left(t, X^k(t), u(t)\right) - b_x^{k_0}\left(t, X^k(t), u(t)\right)\right|^4 \mathrm{d}t\right]. \\
J_2^{k_0} &= \tilde{\mathbb{E}}\left[\int_0^T \left| b_x^{k_0}\left(t, X^k(t), u(t)\right) - b_x^{k_0}(t, X(t), u(t))\right|^4 \mathrm{d}t\right]. \\
J_3^{k_0} &= \tilde{\mathbb{E}}\left[\int_0^T \left| b_x^{k_0}(t, X(t), u(t)) - b_x(t, X(t), u(t))\right|^4 \mathrm{d}t\right].
\end{aligned}$$

Applying the same arguments used in the first limit (Egorov and Portmanteau-Alexandrov Theorems), we obtain that $\lim_{k\to+\infty} J_1^{k_0} = 0$. We use the continuity of $b_x^{k_0}$ in x and the convergence in probability of $X^k(t)$ to $X(t)$ to deduce that $b_x^{k_0}\left(t, X^k(t), u(t)\right)$ converges to $b_x^{k_0}(t, X(t), u(t))$ in probability as $k_0 \to +\infty$, and to deduce by using the Dominated Convergence Theorem, that $\lim_{k\to+\infty} J_2^{k_0} = 0$. Since $b_x^{k_0}, b_x$ are bounded by the Lipschitz constant and by using the absolute continuity of the law of $X(t)$ with respect to the Lebesgue measure, the convergence of $b_x^{k_0}$ to b_x, and the dominated convergence theorem, we get $\lim_{k\to+\infty} J_3^{k_0} = 0$. The case of the second assertions (b) can be treated by the same technique. ■

Lemma 3.3.5

Let $u(\cdot) \in \mathcal{U}_{ad}$, $(P(\cdot), Q(\cdot), R(\cdot))$ and $(P^k(\cdot), Q^k(\cdot), R^k(\cdot))$ the solutions of (3.1) and (3.12)Assume that Assumptions ($\mathbb{A}$.1)-($\mathbb{A}$.4) hold. Then we have the following estimates

$$\lim_{k\to\infty} \tilde{\mathbb{E}}\left[\sup_{s\leq t\leq T}\left|P^k(t) - P(t)\right|^2\right] = 0, \tag{3.20}$$

$$\lim_{k\to\infty} \tilde{\mathbb{E}}\left[\sup_{s\leq t\leq T}\left|Q^k(t) - Q(t)\right|^2\right] = 0, \tag{3.21}$$

$$\lim_{k\to\infty} \tilde{\mathbb{E}}\left[\int_0^T \left|R^k(t) - R(t)\right|^2 \mathrm{d}t\right] = 0. \tag{3.22}$$

Proof: We first prove (3.20), using standard arguments based on Hölder inequality, we easily get

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[\sup_{0\leq t\leq T}\left|P^k(t) - P(t)\right|^2\right] \\
&\leq C\left(\tilde{\mathbb{E}}\int_0^T \left|f_y^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) - f_y^*(t, X(t), Y(t), Z(t), u(t))\right|^4 \mathrm{d}t\right)^{\frac{1}{2}} \\
&+ C\left(\tilde{\mathbb{E}}\int_0^T \left|f_z^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) - f_z^*(t, X(t), Y(t), Z(t), u(t))\right|^4 \mathrm{d}t\right)^{\frac{1}{2}} \\
&+ C\left(\tilde{\mathbb{E}}\int_0^T \left|l_y^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) - l_y^*(t, X(t), Y(t), Z(t), u(t))\right|^4 \mathrm{d}t\right)^{\frac{1}{2}} \\
&+ C\left(\tilde{\mathbb{E}}\int_0^T \left|l_z^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) - l_z^*(t, X(t), Y(t), Z(t), u(t))\right|^4 \mathrm{d}t\right)^{\frac{1}{2}} \\
&: = C\left(\left(I_4^k\right)^{\frac{1}{2}} + \left(I_5^k\right)^{\frac{1}{2}} + \left(I_6^k\right)^{\frac{1}{2}} + \left(I_7^k\right)^{\frac{1}{2}}\right).
\end{aligned}$$

We have

$$\begin{aligned}
I_4^k \leq\ & \tilde{\mathbb{E}}\int_0^T \left|f_y^*\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t)\right) - f_y^*(t, X(t), Y(t), Z(t), u(t))\right|^4 \mathbb{1}_{\left\{u^k(r)\neq u(r)\right\}}\mathrm{d}t \\
&+ \tilde{\mathbb{E}}\int_0^T \left|f_y^*\left(t, X^k(t), Y^k(t), Z^k(t), u(t)\right) - f_y^*(t, X(t), Y(t), Z(t), u(t))\right|^4 \mathrm{d}t.
\end{aligned}$$

In view of the boundedness of f_y, (3.17), (3.18), (3.19) and $d\left(u^k(\cdot), u(\cdot)\right)$ converges to 0 as k goes to $+\infty$, by dominated convergence theorem, we obtain $\lim_{k\to\infty} I_4^k = 0$. Similarly, we have $\lim_{k\to\infty} I_5^k = \lim_{k\to\infty} I_6^k = \lim_{k\to\infty} I_7^k = 0$. Therefore, (3.20) is proved.

Applying Ito's formula to $\left|Q^k(t)-Q(t)\right|^2$, we obtain

$$\left|Q^k(t)-Q(t)\right|^2+\int_0^T\left|R^k(r)-R(r)\right|^2dr$$
$$=\left|Q^k(T)-Q(T)\right|^2-2\int_t^T\left\langle Q^k(r)-Q(r),\left(R^k(r)-R(r)\right)\right\rangle dW(r)$$
$$+2\int_t^T\left\langle Q^k(r)-Q(r),b_x^{k,*}\left(r,X^k(r),u^k(r)\right)Q^k(r)-b_x^*(r,X(r),u(r))Q(r)\right\rangle dr$$
$$+2\int_t^T\left\langle Q^k(r)-Q(r),\sigma_x^{k,*}\left(r,X^k(r)\right)R^k(r)-\sigma_x^*(r,X(r))R(r)\right\rangle dr$$
$$+2\int_t^T\left\langle Q^k(r)-Q(r),l_x^*\left(r,X^k(r),Y^k(r),Z^k(r),u^k(r)\right)-l_x^*(t,X(r),Y(r),Z(r),u(r))\right\rangle dr$$
$$-2\int_t^T\left\langle Q^k(r)-Q(r),f_x^*\left(r,X^k(r),Y^k(r),Z^k(r),u^k(r)\right)P^k(r)-\right.$$
$$\left.f_x^*(t,X(r),Y(r),Z(r),u(r))P(r)\right\rangle dr.$$

By standard arguments based on Hölder inequality and Burkholder-Davis-Gundy inequality, we easily get that

$$\tilde{\mathbb{E}}\left(\sup_{0\leq t\leq T}\left|Q^k(t)-Q(t)\right|^2\right)+\tilde{\mathbb{E}}\int_0^T\left|R^k(r)-R(r)\right|^2dr$$
$$\leq C\tilde{\mathbb{E}}\int_0^T\sup_{t\leq s\leq T}\left|Q^k(t)-Q(t)\right|^2\mathrm{d}t+C\tilde{\mathbb{E}}\left|g_x\left(X^k(T)\right)-g_x(X(T))\right|^2$$
$$+C\tilde{\mathbb{E}}\left(\sup_{0\leq t\leq T}\left|P^k(t)-P(t)\right|^2\right)$$
$$+C\tilde{\mathbb{E}}\int_0^T\left|f_x\left(t,X^k(t),Y^k(t),Z^k(t),u^k(t)\right)-f_x(t,X(t),Y(t),Z(t),u(t))\right|^2\mathrm{d}t$$
$$+C\tilde{\mathbb{E}}\int_0^T\left|b_x^k\left(t,X^k(t),u^k(t)\right)-b_x(t,X(t),u(t))\right|^2\mathrm{d}t$$
$$+C\left(\tilde{\mathbb{E}}\int_0^T\left|\sigma_x^k\left(t,X^k(t)\right)-\sigma_x(t,X(t))\right|^{\frac{2\alpha}{\alpha-2}}\mathrm{d}t\right)^{\frac{\alpha-2}{2\alpha}}$$
$$+C\tilde{\mathbb{E}}\int_0^T\left|l_x\left(t,X^k(t),Y^k(t),Z^k(t),u^k(t)\right)-l_x(t,X(t),Y(t),Z(t),u(t))\right|^2\mathrm{d}t$$
$$:=C\left(\tilde{\mathbb{E}}\int_0^T\sup_{t\leq s\leq T}\left|Q^k(t)-Q(t)\right|^2\mathrm{d}t+\sum_{i=1}^6 J_i^k\right).$$

By dominated convergence theorem and (3.17), we have $\lim_{k\to\infty}J_1^k=0$.
From (3.20), we have $\lim_{k\to\infty}J_2^k=0$. Similarly as the proof of I_4^k, we have $\lim_{k\to\infty}J_3^k=0=\lim_{k\to\infty}J_6^k=0$. Next, by assertion (a) in Lemma 3.3.4, we get $\lim_{k\to\infty}J_5^k=0$. ■

Define

$$u^{\theta}(t)=\begin{cases} v_{t_0} \text{ if } t\in[t_0,t_0+\theta], \\ u(t) \quad \text{otherwise,}\end{cases}$$

where, $t_0\in[0,T)$ and we are defined v_{t_0} in (3.15). See easily that $d(u^{k,\theta}(\cdot),u^{\theta}(\cdot))\to 0$ as $k\to+\infty$.

3.3.2 Maximum Principle for Optimality

We need the following lemma to prove the necessary condition of optimality

Lemma 3.3.6

We assume ($\mathbb{A}$.1)-($\mathbb{A}$.4). Then, we have for any $\theta>0$,

$$\begin{aligned}\lim_{k\to+\infty}\tilde{\mathbb{E}}\Big|&\int_0^T\mathcal{H}^k\left(t,X^k(t),Y^k(t),Z^k(t),P^k(t),Q^k(t),R^k(t),u^{k,\theta}(t)\right)\mathrm{d}t\\ &-\int_0^T\mathcal{H}\left(t,X(t),Y(t),Z(t),P(t),Q(t),R(t),u^{\theta}(t)\right)\mathrm{d}t\Big|=0,\end{aligned}\tag{3.23}$$

and

$$\begin{aligned}\lim_{k\to+\infty}\tilde{\mathbb{E}}\Big|&\int_0^T\mathcal{H}^k\left(t,X^k(t),Y^k(t),Z^k(t),P^k(t),Q^k(t),R^k(t),u^{k}(t)\right)\mathrm{d}t\\ &-\int_0^T\mathcal{H}(t,X(t),Y(t),Z(t),P(t),Q(t),R(t),u(t))\,\mathrm{d}t\Big|=0.\end{aligned}\tag{3.24}$$

Proof: For simplicity, denote by

$$\mathcal{H}^{k,\theta}(t)=\mathcal{H}^k\left(t,X^k(t),Y^k(t),Z^k(t),P^k(t),Q^k(t),R^k(t),u^{k,\theta}(t)\right)$$
$$\mathcal{H}^{\theta}(t)=\mathcal{H}\left(t,X(t),Y(t),Z(t),P(t),Q(t),R(t),u^{\theta}(t)\right).$$

Therefore,

$$
\begin{aligned}
&\tilde{\mathbb{E}}\left|\int_0^T \mathcal{H}^{k,\theta}(t)\,\mathrm{d}t-\int_0^T \mathcal{H}^{\theta}(t)\,\mathrm{d}t\right| \\
&\leq C\left\{\left(\tilde{\mathbb{E}}\int_0^T \left|P^k(t)-P(t)\right|^2\mathrm{d}t\right)^{\frac{1}{2}}+\left(\tilde{\mathbb{E}}\int_0^T \left|Q^k(t)-Q(t)\right|^2\mathrm{d}t\right)^{\frac{1}{2}}\right. \\
&+\left(\tilde{\mathbb{E}}\int_0^T \left|R^k(t)-R(t)\right|^2\mathrm{d}t\right)^{\frac{1}{2}} \\
&+\left(\tilde{\mathbb{E}}\int_0^T \left|f\left(t,X^k(t),Y^k(t),Z^k(t),u^{k,\theta}(t)\right)-f\left(t,X(t),Y(t),Z(t),u^{\theta}(t)\right)\right|^2\mathrm{d}t\right)^{\frac{1}{2}} \\
&+\left(\tilde{\mathbb{E}}\int_0^T \left|l\left(t,X^k(t),Y^k(t),Z^k(t),u^{k,\theta}(t)\right)-l\left(t,X(t),Y(t),Z(t),u^{\theta}(t)\right)\right|^2\mathrm{d}t\right)^{\frac{1}{2}} \\
&+\left(\tilde{\mathbb{E}}\int_0^T \left|b^k\left(t,X^k(t),u^{k,\theta}(t)\right)-b\left(t,X(t),u^{\theta}(t)\right)\right|^2\mathrm{d}t\right)^{\frac{1}{2}} \\
&\left.+\left(\tilde{\mathbb{E}}\int_0^T \left|\sigma^k\left(t,X^k(t)\right)-\sigma\left(t,X(t)\right)\right|^2\mathrm{d}t\right)^{\frac{1}{2}}\right\} \\
&=C\left(L_1^k+L_2^k+L_3^k+L_4^k+L_5^k+L_6^k+L_7^k\right).
\end{aligned}
$$

First, by Lemma 3.3.5, we obtain that

$$\lim_{k\to+\infty}\left(L_1^k+L_2^k+L_3^k\right)=0.$$

Next, by Lemma 3.3.3, the Lipschitz continuity of f,l and dominated convergence theorem, we have

$$\lim_{k\to+\infty}\left(L_4^k+L_5^k\right)=0.$$

Finally,

$$
\begin{aligned}
(L_6^k)^2 \leq\; & 3\tilde{\mathbb{E}}\int_0^T \left|b^k\left(t,X^k(t),u^{k,\theta}(t)\right)-b\left(t,X(t),u^{\theta}(t)\right)\right|^2 \mathbb{1}_{\left\{u^{k,\varepsilon}\neq u^{\varepsilon}\right\}}\mathrm{d}t \\
&+3\tilde{\mathbb{E}}\int_0^T \left|b^k\left(t,X^k(t),u^{\theta}(t)\right)-b^k\left(t,X(t),u^{\theta}(t)\right)\right|^2\mathrm{d}t \\
&+3\tilde{\mathbb{E}}\int_0^T \left|b^k\left(t,X(t),u^{\theta}(t)\right)-b\left(t,X(t),u^{\theta}(t)\right)\right|^2\mathrm{d}t.
\end{aligned}
$$

According to the boundedness of b^k, b and the fact that $d(u^{k,\theta}(\cdot),u^{\theta}(\cdot))\to 0$ as $k\to+\infty$ guarantee the convergence of the first part on the right-hand side of the above inequality to 0 as $k\to+\infty$. Moreover, by the Lipschitz continuity of b^k and the fact that $X^k(t)\to X(t)$ uniformly in probability, we get the second part tends to 0 as $k\to+\infty$. In view of assertion (b) of Lemma (3.3.1) we have the last part tends to 0 as $k\to+\infty$.Hence

$\lim_{k\to+\infty} L_6^k = 0$. Similarly, we get $\lim_{k\to+\infty} L_7^k = 0$. Therefore, (3.23) is proved. By the same argument, we obtain (3.24). ■

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

Theorem 3.3.1 (*Necessary Condition for Optimality*)

Let $u(\cdot)$ be an optimal control and $(X(\cdot), Y(\cdot), Z(\cdot))$ be the corresponding trajectory. Then, for any $v(\cdot) \in \mathcal{U}_{ad}$, we have

$$\begin{aligned}&\mathcal{H}(t, X(t), Y(t), Z(t), P(t), Q(t), R(t), u(t)) - \mathcal{H}(t, X(t), Y(t), Z(t), P(t), Q(t), R(t), v)\\ &\geq 0, \quad dt - a.e., \tilde{\mathbb{P}}{-}a.s.,\end{aligned} \tag{3.25}$$

where, $(P(\cdot), Q(\cdot), R(\cdot))$ is the solution of the adjoint equation (2.13) with respect to $(X(\cdot), Y(\cdot), Z(\cdot), u(\cdot))$.

Proof: From Proposition 3.3.1, we have

$$\begin{aligned}&\tilde{\mathbb{E}}\int_0^T \left[\mathcal{H}^k\left(t, X^k(t), Y^k(t), Z^k(t), u^{k,\theta}(t), P^k(t), Q^k(t), R^k(t)\right)\right]\\ &-\tilde{\mathbb{E}}\int_0^T \left[\mathcal{H}^k\left(t, X^k(t), Y^k(t), Z^k(t), u^k(t), P^k(t), Q^k(t), R^k(t)\right)\right] \geq o(\theta) - \theta C(\delta_k)^{\frac{1}{2}}.\end{aligned}$$

Letting k goes to 0 and by using Lemma 3.3.6, we obtain that for each $\theta > 0$,

$$\begin{aligned}0 \leq \tilde{\mathbb{E}}\int_0^T &\left[\mathcal{H}\left(t, X(t), Y(t), Z(t), P(t), Q(t), R(t), u^\theta(t)\right)\right.\\ &\left.-\mathcal{H}(t, X(t), Y(t), Z(t), P(t), Q(t), R(t), u(t))\right] \mathrm{d}r + o(\theta).\end{aligned} \tag{3.26}$$

Dividing the both sides of the inequality (3.26) by θ, and passing to the limit on θ, we get

$$\begin{aligned}&\tilde{\mathbb{E}}\left[\mathcal{H}(t_0, X(t_0), Y(t_0), Z(t_0), P(t_0), Q(t_0), R(t_0), v_{t_0})\right.\\ &\left.-\mathcal{H}(t_0, X(t_0), Y(t_0), Z(t_0), P(t_0), Q(t_0), R(t_0), u(t_0))\right] \geq 0.\end{aligned}$$

Keeping in mind that t_0 is an arbitrary element of $[0, T]$, we get

$$\begin{aligned}&\tilde{\mathbb{E}}\left[\mathcal{H}(t, X(t), Y(t), Z(t), P(t), Q(t), R(t), v_t)\right.\\ &\left.-\mathcal{H}(t, X(t), Y(t), Z(t), P(t), Q(t), R(t), u(t))\right] \geq 0.\end{aligned}$$

Then, completing the proof as the proof of Theorem 2.5.1 given in chapter 2, thus the result follows immediately, which achieves the proof. ■

The following theorem is another result of this chapter, where the coefficients of the Forward part b and σ are only Lipschitz (not necessarily differentiable) and the generator f is C^1 function and hence is Locally lipschitz.

Theorem 3.3.2

*Suppose ($\mathbb{A}$.1)-($\mathbb{A}$.3) hold. Assume further that the function f satisfies **Assumption 2.1** or **Assumption 2.2** of chapter 2. Let $u(\cdot)$ be an optimal control and $(X(\cdot), Y(\cdot), Z(\cdot))$ be the corresponding trajectory. Then, for any $v(\cdot) \in \mathcal{U}_{ad}$, we have*

$$\begin{aligned}
&\mathcal{H}(t, X(t), Y(t), Z(t), P(t), Q(t), R(t), u(t)) \\
&-\mathcal{H}(t, X(t), Y(t), Z(t), P(t), Q(t), R(t), v) \geq 0, \quad dt - a.e., \tilde{\mathbb{P}}\text{−}a.s.,
\end{aligned}$$

where, $(P(\cdot), Q(\cdot), R(\cdot))$ is the solution of the adjoint equation (2.13) with respect to $(X(\cdot), Y(\cdot), Z(\cdot), u(\cdot))$.

Proof: The proof of this theorem can be performed as a combination of the proof of Theorem (3.3.1) and the proof of Theorem (2.5.1) in Chapter 2. ■

Conclusion

In this thesis, we establish a set of necessary conditions of stochastic optimal control for different stochastic models. As the first result, we have discussed a stochastic optimal control problem for one type of controlled BSDE with locally Lipschitz coefficient. We strongly believe that it is the first attempt that goes in this direction and it is a new endeavor. Pretty much all of the difficulties come from the fact the BSDE generator and the adjoint equation are only locally Lipschitz which makes it difficult to solve the control problem using the standard duality technique. We have firstly proved an existence and uniqueness result to the related adjoint process which is described by a linear SDE with locally bounded coefficients. Then, by means, of Ekeland's variational principle along with an approximation and limit arguments, both the necessary and the sufficient conditions for optimality are obtained. As a second result, we have developed a stochastic maximum principle for optimal control problems of the degenerate FBSDEs systems, where the coefficients of the forward equation are only Lipschitz continuous with respect to the state variable x. Using Ekeland's variational principle to a sequence of approximated control problems with smooth coefficients of the initial problems and applying the Bouleau-Hirsch flow property to define an adjoint process which is the unique solution of the linear backward-forward SDE defined on an extension of the initial probability space. Several optimal control problems are still open problems. For example, the stochastic maximum principle for locally Lipschitz systems driven by SDEs of Ito's type, coupled or semi-coupled FBSDEs, and so on. We hope that we can extend the results to the classical non-Lipschitz framework of dissipative or one-sided Lipschitz coefficients

Another open remaining problem is the stochastic maximum principle in the case where coefficients of the BSDE are globally Lipschitz and non-differentiable with respect to y and z. We plan to fill the gaps by studying these open problems in our forthcoming research papers.

Bibliography

[1] Khaoula ADELHADI and Nabil KHELFALLAH. Existence and uniqueness of solutions for bsdes associated to jump markov process with locally lipschitz coefficient. In *9 ème édition du colloque Tendances dans les Applications Mathématiques en Tunisie Algérie et Maroc*, page 30, 2019.

[2] Nacira Agram. Stochastic optimal control of Mckean-Vlasov equations with anticipating law. *Afrika Matematika*, 30:879–901, 2019.

[3] Ishak Alia, Farid Chighoub, Nabil Khelfallah, and Josep Vives. Time-consistent investment and consumption strategies under a general discount function. *Journal of Risk and Financial Management*, 14(2):86, 2021.

[4] Vadim Iosifovich Arkin and MT Saksonov. Necessary optimality conditions in control problems for stochastic differential equations. In *Doklady Akademii Nauk*, volume 244, pages 11–15. Russian Academy of Sciences, 1979.

[5] Hanine Azizi and Nabil Khelfallah. The maximum principle for optimal control of BSDEs with locally Lipschitz coefficients. *Journal of Dynamical and Control Systems*, 28(3):565–584, 2022.

[6] Fouzia Baghery, Nabil Khelfallah, Brahim Mezerdi, and Isabelle Turpin. On optimal control of forward-backward stochastic differential equations. *Afrika Matematika*, 28(7-8):1075–1092, 2017.

[7] Khaled Bahlali. Existence and uniqueness of solutions for BSDEs with locally Lipschitz coefficient. *Electronic Communications in Probability [electronic only]*, 7:169–179, 2002.

[8] Khaled Bahlali, Boualem Djehiche, and Brahim Mezerdi. On the stochastic maximum principle in optimal control of degenerate diffusions with Lipschitz coefficients. *Applied mathematics and optimization*, 56:364–378, 2007.

[9] Khaled Bahlali, Nabil Khelfallah, and Brahim Mezerdi. Necessary and sufficient conditions for near-optimality in stochastic control of fbsdes. *Systems & Control Letters*, 58(12):857–864, 2009.

[10] Khaled Bahlali, Nabil Khelfallah, and Brahim Mezerdi. Optimality conditions for partial information stochastic control problems driven by lévy processes. *Systems & control letters*, 61(11):1079–1084, 2012.

[11] Khaled Bahlali, Brahim Mezerdi, and Youssef Ouknine. The maximum principle for optimal control of diffusions with non-smooth coefficients. *Stochastics: An International Journal of Probability and Stochastic Processes*, 57(3-4):303–316, 1996.

[12] Alain Bensoussan. Lectures on stochastic control. In *Nonlinear Filtering and Stochastic Control: Proceedings of the 3rd 1981 Session of the Centro Internazionale Matematico Estivo (CIME), Held at Cortona, July 1–10, 1981*, pages 1–62. Springer, 2006.

[13] Jean-Michel Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.

[14] Jean-Michel Bismut. Linear quadratic optimal stochastic control with random coefficients. *SIAM Journal on Control and Optimization*, 14(3):419–444, 1976.

[15] Jean-Michel Bismut. An introductory approach to duality in optimal stochastic control. *SIAM review*, 20(1):62–78, 1978.

[16] Jean-Michel Bismut. Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 56(4):469–505, 1981.

[17] Saliha Bougherara and Nabil Khelfallah. The maximum principle for partially observed optimal control of FBSDE driven by Teugels Martingales and independent Brownian motion. *Journal of Dynamical and Control Systems*, 24:201–222, 2018.

[18] Nicolas Bouleau and Francis Hirsch. Sur des propriétés du flot d'une équation différentielle stochastique. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, 306(1):421–424, 1988.

[19] Nicolas Bouleau and Francis Hirsch. On the derivability, with respect to the initial data, of the solution of a stochastic differential equation with Lipschitz coefficients. In *Séminaire de Théorie du Potentiel Paris, No. 9*, pages 39–57. Springer, 1989.

[20] Mezerdi Brahim. Necessary conditions for optimality for a diffusion with a non-smooth drift. *Stochastics: An International Journal of Probability and Stochastic Processes*, 24(4):305–326, 1988.

[21] Abel Cadenillas and Ioannis Karatzas. The stochastic maximum principle for linear, convex optimal control with random coefficients. *SIAM journal on control and optimization*, 33(2):590–624, 1995.

[22] Piermarco Cannarsa and Hélène Frankowska. Value function, relaxation, and transversality conditions in infinite horizon optimal control. *Journal of Mathematical Analysis and Applications*, 457(2):1188–1217, 2018.

[23] Frank H Clarke. *Optimization and nonsmooth analysis*. SIAM, 1990.

[24] Boualem Djehiche, Brahim Mezerdi, and Farid Chighoub. The stochastic maximum principle in optimal control of degenerate diffusions with non-smooth coefficients. *Random Operators and Stochastic Equations*, 17:35–53, 2009.

[25] Ivar Ekeland. Nonconvex minimization problems. *Bulletin of the American Mathematical Society*, 1(3):443–474, 1979.

[26] Nicole El Karoui, Shige Peng, and Marie Claire Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

[27] Xinwei Feng. Maximum principle for optimal control problems involving impulse controls with nonsmooth data. *Stochastics*, 88(8):1188–1206, 2016.

[28] M Margarida A Ferreira. On the regularity of optimal controls for a class of problems with state constraints. *International journal of systems science*, 37(8):495–502, 2006.

[29] Halina Frankowska. The first order necessary conditions for nonsmooth variational and control problems. *SIAM journal on control and optimization*, 22(1):1–12, 1984.

[30] Dalila Guerdouh, Nabil Khelfallah, and Josep Vives. Optimal control strategies for the premium policy of an insurance firm with jump diffusion assets and stochastic interest rate. *Journal of Risk and Financial Management*, 15(3):143, 2022.

[31] Said Hamadène. Equations différentielles stochastiques rétrogrades: le cas localement Lipschitzien. *Annales de l'IHP Probabilités et statistiques*, 32(5):645–659, 1996.

[32] Tao Hao and Qingxin Meng. Singular optimal control problems with recursive utilities of mean-field type. *Asian Journal of Control*, 23(3):1524–1535, 2021.

[33] Ulrich G Haussmann. *A stochastic maximum principle for optimal control of diffusions*. John Wiley & Sons, Inc., 1986.

[34] Nabil Khelfallah and Brahim Mezerdi. Near-optimality conditions in stochastic control of linear fully coupled FBSDEs. *Afrika Matematika*, 27(3-4):327–343, 2016.

[35] Michael Kohlmann and Xun Yu Zhou. Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach. *SIAM Journal on Control and Optimization*, 38(5):1392–1407, 2000.

[36] Harold Kushner. Necessary conditions for continuous parameter stochastic optimization problems. *SIAM Journal on Control*, 10(3):550–565, 1972.

[37] Nazim I Mahmudov and Mark A McKibben. On backward stochastic evolution equations in hilbert spaces and optimal control. *Nonlinear Analysis: Theory, Methods & Applications*, 67(4):1260–1274, 2007.

[38] David Nualart. *The Malliavin calculus and related topics*, volume 1995. Springer, 2006.

[39] Carlo Orrieri. A stochastic maximum principle with dissipativity conditions. *arXiv preprint arXiv:1309.7757*, 2013.

[40] Etienne Pardoux and Shige Peng. Adapted solution of a backward stochastic differential equation. *Systems & control letters*, 14(1):55–61, 1990.

[41] Etienne Pardoux and Shige Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic Partial Differential Equations and Their Applications: Proceedings of IFIP WG 7/1 International Conference University of North Carolina at Charlotte, NC June 6–8, 1991*, pages 200–217. Springer, 2005.

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

[43] Ruimin Xu, Tingting Wu, et al. Mean-field backward stochastic evolution equations in hilbert spaces and optimal control for BSPDEs. *Mathematical Problems in Engineering*, 2014, 2014.

[44] Wensheng Xu. Stochastic maximum principle for optimal control problem of forward and backward system. *The ANZIAM Journal*, 37(2):172–185, 1995.

[45] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.

[46] Xun Yu Zhou. Sufficient conditions of optimality for stochastic systems with controllable diffusions. *IEEE Transactions on Automatic Control*, 41(8):1176–1179, 1996.