PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

University Mohamed Khider of Biskra
Faculty of Exact Sciences and Natural and Life Sciences

A Thesis presented for the degree of

Doctor of Sciences in Mathematics

In the Field of Probability and Statistics

Title

Pointwise Second Order Necessary Conditions for Stochastic Optimal Control with Jump Diffusion

---

Presented by : Abdelhak *Ghoul*

Members of the jury:

Mokhtar HAFAYED	Pr,	University of Biskra	**President**.
Imad Eddine LAKHDARI	Dr,	University of Biskra	**Supervisor.**
Boulakhras GHERBAL	Pr,	University of Biskra	**Examiner.**
Khalil SAADI	Pr,	University of Msila	**Examiner.**
Abdelmoumen TIAIBA	Pr,	University of Msila	**Examiner.**
Youcef DJENAIHI	Dr,	University of Setif	**Examiner.**

College year 2022

Acknowledgement

First, I thank God the Almighty for helping me to succeed in my quest and giving me courage and patience to realize this work.

I would like to express my sincere gratitude to my advisor Dr. *Imad Eddine LAKHDARI* for the continuous support to establish my doctorate thesis and related research, for his patience, motivation and immense knowledge. His advice helped me throughout the research and writing of this thesis. I could not have imagined having a better advisor and mentor for my doctorate thesis.

I would like to thanks, *Mokhtar HAFAYED* Professor at the university of Biskra for his advice and guidance all the time of research.

My sincere thanks to Prof. *Mokhtar HAFAYED*, Prof. *GHERBAL Boulakhras*, Prof. *Youcef DJENAIHI*, Prof. *Abdelmoumen TIAIBA*, and Prof. *Khalil SAADI* , because they agreed to spend their times for reading and evaluating my thesis.

Aspecial thanks to my parents, my wife, my brothers and sisters for trusting and supporting me spiritually throughout writing this thesis.

I thank all my colleagues of the Mathematics Department.

Résumé

Le principe du maximum stochastique est l'une des approches importantes pour discuter les problèmes de contrôle stochastique. Beaucoup de travail a été fait sur ce genre de problème, voir, par exemple, Bensoussan [3], Cadenillas et Karatzas [10], Kushner [31], Peng [41]. Récemment, un autre type de principe du maximum stochastique, les conditions nécessaires ponctuelles du second ordre pour les contrôles optimaux stochastiques a été établi et étudié pour ses applications sur le marché financier par Zhang et Zhang [58] lorsque la région de contrôle est supposée être convexe. Dans Zhang et Zhang [59], les auteurs ont étendu les conditions nécessaires ponctuelles du second ordre pour les contrôles optimaux stochastiques dans le cas général où la région de contrôle est non convexe. Les conditions nécessaires du second ordre pour un contrôle optimal avec des utilitaires récursifs ont été prouvées par Dong et Meng [13].

Dans cette thèse, nous généralisons le travail de Zhang et Zhang [58] pour les systèmes avec saut, nous établissons les conditions nécessaires du second ordre où le système contrôlé est décrit par un système différentiel stochastique gouverné par une mesure aléatoire de Poisson et un mouvement brownien indépendant. Le domaine de contrôle est supposé convexe. La preuve du résultat principal est basée sur une approche variationnelle utilisant le calcul stochastique des diffusions de sauts et quelques estimations sur le processus d'état.

Mots Clés. Contrôle optimal, Systèmes stochastiques avec sauts, Condition nécessaire ponctuelle du second ordre, Principe du maximum, Equation variationnelle.

Abstract

Stochastic maximum principle is one of the important major approaches to discuss stochastic control problems. A lot of work has been done on this kind of problem, see, for example, Bensoussan [3], Cadenillas and Karatzas [10], Kushner [31], Peng [41].

Recently, another kind of stochastic maximum principle, pointwise second order necessary conditions for stochastic optimal controls has been established and studied for its applications in the financial market by Zhang and Zhang [58] when the control region is assumed to be convex. In Zhang and Zhang [59], the authors extended the pointwise second order necessary conditions for stochastic optimal controls in the general cases when the control region is allowed to be non convex. Second order necessary conditions for optimal control with recursive utilities was proved by Dong and Meng [13].

In this thesis, we generalizes the work of Zhang and Zhang [58] for jump diffusions, we establish a second order necessary conditions where the controlled system is described by a stochastic differential systems driven by Poisson random measure and an independent Brownian motion. The control domain is assumed to be convex. Pointwise second order maximum principle for controlled jump diffusion in terms of the martingale with respect to the time variable is proved. The proof of the main result is based on variational approach using the stochastic calculus of jump diffusions and some estimates on the state processes. Our stochastic control problem provides also an interesting models in many applications such as economics and mathematical finance.

Keys words. Optimal control, Stochastic systems with jumps, Pointwise second-order necessary condition, Maximum principle, Variational equation.

Symbols and Acronyms

$(\Omega, \mathcal{F}, \mathbb{F}, \mathcal{P})$	Complete probability space
$\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$	Natural filtration
W	Brownian motion
N	Poisson random measure
$\widetilde{N}$	Compensator jump martingale random measure
$\mu\,(dt, dz)$	Compensator of random measure
$a.e.$	Almost everywhere
$a.s.$	Almost surely
$e.g.$	For example (abbreviation of Latin exempli gratia)
$i.e.$	Abbreviation of Latin (id)
SDE	Stochastic differential equations
$BSDE$	Backward stochastic differential equation
ODE	Ordinary differential equation
$\phi_x\,(t, x, u)$	First parial derivatives of ϕ with respect to x
$\phi_u\,(t, x, u)$	First parial derivatives of ϕ with respect to u
$\phi_{xx}\,(t, x, u)$	The second order derivatives of ϕ with respect to (x, x)
$\phi_{xu}\,(t, x, u)$	First parial derivatives of ϕ with respect to (x, u)
$\phi_{uu}\,(t, x, u)$	Second order derivatives of ϕ with respect to (u, u)
$\phi_{(x,u)^2}\,(t, x, u)$	Second order derivatives of ϕ with respect to (x, u)
$\mathcal{U}_{ad}$	The set of all admissible controls
$L^2_{\mathcal{F}_t}\,(\Omega; \mathbb{R})$	The space of $\mathbb{R}$-valued, $\mathcal{F}_t$-measurable random variables
$\mathbb{L}^2_{\mathbb{F}}\,([0, T]\,; \mathbb{R})$	The space of $\mathbb{R}$-valued, $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable, $\mathbb{F}$-adapted processes
$\mathcal{L}^2\,([0, T]\,; \mathbb{R})$	The space of $\mathbb{R}$-valued, $\mathcal{B}([0, T] \times \Omega) \otimes \mathcal{B}\,(Z)$ measurable processes

Contents

Abstract

Résumé

Symbols and Acronyms

Introduction

1 Introduction to stochastic calculus **5**

1.1 Diffusion process . . . 5

1.1.1 Brownian motion and martingales . . . 5

1.1.2 Quadratic variation . . . 6

1.1.3 Stochastic integrals . . . 7

1.1.4 Stochastic differential equations . . . 10

1.1.5 Itô's lemma . . . 10

1.1.6 Some examples . . . 12

1.2 Jump diffusions . . . 13

1.2.1 Lévy processes . . . 13

1.2.2 Itô Formula with Jumps . . . 17

1.2.3 Stochastic differential equations with jumps . . . 19

2 Stochastic optimal control problems **21**

2.1 Problem formulation . . . 21

2.2 Dynamic programming principle . . . 23

2.3 Stochastic maximum principle . . . 29

2.4 A General stochastic maximum principle for optimal control problems . . . 37

2.4.1 Problem formulation and assumptions 38
2.4.2 Second order expansion 38
2.4.3 Adjoint processes and variational inequality 42
2.4.4 Adjoint equations and the maximum principle 46

3 Pointwise second order necessary conditions for stochastic optimal control 47
3.1 Preliminaries and assumptions 47
3.2 Second order necessary condition in integral form 49
3.3 Pointwise second order maximum principle in terms of the martingale . . . 59

4 Pointwise second order necessary conditions for stochastic optimal control with jump diffusions 66
4.1 Preliminaries and assumptions 66
4.2 Second order necessary condition in integral form with jump Diffusions . . 69
4.3 Pointwise second order maximum principle in terms of the martingale with Jump Diffusions 81

Conclusion 86

Bibliographie 90

Introduction

The main goal of this thesis is to investigate the pointwise second order necessary conditions for stochastic optimal control problem with jump diffusions. The maximum principle is one of the major approaches to discuss this kind of problems. The stochastic optimal control problems for jump processes have been investigated by many authors. Cadenillas [8] proved a stochastic maximum principle for a linear dynamics with jumps and convex state constraint, this result is the first version of stochastic maximum principle that covers the consumption-investment problem in which there are jumps in the price system. The stochastic maximum principle for jump diffusion in general case, where the control domain need not be convex, and the diffusion coefficient depends explicitly on the control variable, was derived via spike variation method by Tang and Li [50], extending the Peng's stochastic maximum principle of optimality developed in Peng [41]. A general linear quadratic optimal stochastic control problem driven by a Brownian motion and a Poisson random martingale measure with random coefficients has been studied in Meng [37]. Optimal control of mean-field jump-diffusion systems with delay was studied by Meng and Shen [38]. Necessary and sufficient conditions for mean-field jump–diffusion stochastic delay differential equations and its application to finance have been obtained in Meng and Shen [46]. Linear quadratic optimal control problems for mean-field stochastic differential equations with jumps have been investigated in Tang and Meng [51]. Necessary and sufficient conditions for stochastic near-optimal singular controls for jump diffusions have been investigated in Hafayed and Abbas [26]. Necessary conditions for partially observed optimal control of general McKean–Vlasov stochastic differential equations with jumps has been studied in Miloudi et al. [39] A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes has been studied by Hafayed [27]. The sufficient conditions for optimality was obtained by Framstad et al. [15]. Maximum principle for forward-backward stochastic control system with random jumps with some application to finance has been investigated by Shi and Wu [48]. Filtering problems for forward-backward stochastic systems with random jumps with applications to partial information stochastic optimal control have been studied in Xiao and Wang [54]. Infinite horizon stochastic optimal control problem of mean-field delay system with semi-Markov modulated jump-diffusion processes has been studied in Deepa and Muthukumar [12]. Discrete time approximation of decoupled forward-backward stochastic systems with jumps was studied in Bouchard and Elie [6]. Stochastic optimal control of evolution equations of jump type with random coefficients has been studied in Tang and Meng [53]. Zhang et al. [57] proved the sufficient maximum principle where the state process is governed by a continuous-time Markov regime-switching jump-diffusion model. A various maximum principles for optimal controls of stochastic with random jumps have been investigated in [45, 47]. An extensive list of references to the stochastic optimal control problem with jumps with some applications in finance and economics can be found in [47, 40].

An integral type second order necessary condition for stochastic optimal control problems under the assumption that the control region is convex have been studied by Bonnans and Silva [7]. Zhang and Zhang [58] established the pointwise second order necessary conditions for stochastic optimal controls when the control region is assumed to be convex. In Zhang and Zhang [59], the authors extended the pointwise second order necessary conditions for stochastic optimal controls in the general cases when the control region is allowed to be non convex. Second order necessary conditions for optimal control with recursive utilities was proved by Dong and Meng [13]. Pointwise second order necessary conditions of optimality for the Mayer-type problem with constraints have been derived by Frankoswka and Tonon [16]. Second order necessary conditions for singular optimal stochastic controls with some examples have been obtained in Tang [52]. First and second order necessary optimality conditions for local minimizers of stochastic optimal control problems with state constraints have been established in Frankowska et al. [17].

Motivated by the works mentioned above, our main goal in this thesis is to prove pointwise second order necessary conditions for stochastic optimal control for jump diffusions. The control variable is allowed to enter into both drift and diffusion terms. Our stochastic control problem provides also an interesting models in many applications such as economics and mathematical finance. Our maximum principle generalizes the work of Zhang and Zhang [58] to jump diffusion, which is a type of stochastic process that has discrete movements called jumps, with random arrival times, rather than continuous movements.

This thesis is organized as follows.

In **Chapter 1**, we give an introduction to stochastic calculus, we presents some concepts and results that allow us to prove our results, such as Diffusion process (Brownian motion and martingales, Stochastic integrals, Stochastic differential equations, Itô formula), Jump diffusions (Lévy processes, Itô formula and related results, Lévy stochastic differential equations).

In **Chapter 2**, we present strong and weak formulations of stochastic optimal control problems. Then, by using the dynamic programming principle (DPP) and the stochastic maximum principle (SMP) in the classical case where the control domain is convex and the system is governed by Brownian motion, we solve our stochastic control problem. Then, we study the maximum principle for nonlinear stochastic optimal control problems in the general case where the control domain need not be convex, and the diffusion coefficient can contain a control variable.

In **Chapter 3,** we discuss pointwise second-order necessary conditions for stochastic singular optimal controls in the classical sense. The controlled system is described by a stochastic differential equation and the control domain is assumed to be convex. This chapter is based on the work of Zhang and Zhang [58].

In **Chapter 4,** we give the main result of this thesis, we establish a second order necessary conditions for stochastic optimal control for jump diffusions. The controlled system is described by a stochastic differential systems driven by Poisson random measure and an independent Brownian motion. The control domain is assumed to be convex. Pointwise second order maximum principle for controlled jump diffusion in terms of the martingale with respect to the time variable is proved. The proof of the main result is based on variational approach using the stochastic calculus of jump diffusions and some estimates on the state processes.

Published Author Papers

The content of this thesis was the subject of the following paper:

A. Ghoul, & I.E. Lakhdari, & M. Hafayed, & S. Meherrem, Pointwise Second Order Necessary Conditions for Stochastic Optimal Control with Jump Diffusion , Communications in Mathematics and Statistics. Springer.

Received: 30 April 2021 / Revised: 3 September 2021 / Accepted: 25 October 2021.

Chapter 1

Introduction to stochastic calculus

1.1 Diffusion process

1.1.1 Brownian motion and martingales

We assume as given a filtered propability space $(\Omega, \mathcal{F}, P)$ where:

1. Ω is the universe of possible outcomes.

2. The set $\mathcal{F}$ represents the set of possible events where an event is a subset of Ω.

3. P is the true probability measure.

i) There is also a filtration, $\{\mathcal{F}_t\}_{t\geq 0}$, that models the evolution of information through time. So for example,if it is known by time t whether or not an event, E, has occurred, then we have $E \in \mathcal{F}_t$. If we are working with a finite horizon, $[0;T]$, then we can take $\mathcal{F} = \mathcal{F}_T$.

ii) We also say that a stochastic process X_t, is $\mathcal{F}_t$-adapted if the value of X_t is known at time t when the information represented by $\mathcal{F}_t$ is known. All the processes we consider will be $\mathcal{F}_t$-adapted so we will not bother to state this in the sequel.

iii) In the continuous-time models that we will study, it will be understood that the filtration $\{\mathcal{F}_t\}_t$ will be the filtration generated by the stochastic processes $\{B_t\}$, that are specified in the model description.

Definition 1.1.1

A stoshastic process $\{B_t : 0 \leq t \leq \infty\}$ is a standard Brownian motion:

1) $B_0 = 0$.

2) With probability 1, the function $t \to B_t$ is continuous in t.

3) The process $\{B_t\}_{t\geq 0}$ has stationary, independent increments.

4) $B_t \sim N(0,t)$.

Definition 1.1.2

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

Definition 1.1.3

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

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

Example 1.1.1

Let B_t *be a Brownian motion.*

Then $B_t^2 - t$ *,*$B_t^3 - 3tB_t$ *and* $\exp(-\lambda^2 \frac{t}{2}) \exp \lambda B_t$ *,are all martingales.*

1.1.2 Quadratic variation

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

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

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

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

Definition 1.1.4

The quadratic variation of a stochastic process, Y_t*, is the process, written as* $[Y]_t$ *is equal to the limit of*

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

Remark 1.1.1

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

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

This follows by observing that

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

If we now let $n \to \infty$ in (1.2) then the continuity of Y_t implies the impossibility of the process having finite total variation and non-zero quadratic variation. Theorem (1.2.1) therefore implies that the total variation of a Brownian motion is infinite. We have the following important result which proves very useful if we need to price options when there are multiple underlying Brownian motions, as is the case with quanto options for example.

1.1.3 Stochastic integrals

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

Definition 1.1.5

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

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

Definition 1.1.6

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

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

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

Definition 1.1.7

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

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

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

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

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

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

Example 1.1.2

We want to compute $\int_0^T B_t dB_t$*.Towards this end, let*

$$0 = t_0^n < t_1^n < t_2^n < ... < t_n^n = T,$$

be a partition of $[0;T]$ *and define*

$$Y_t^n := \sum_{i=0}^{n-1} B_{t_i^n} I_{\left[t_i^n, t_{i+1}^n\right)}(t),$$

where $I_{\left[t_i^n,t_{i+1}^n\right)} = 1$ *if* $t \in \left[t_i^n, t_{i+1}^n\right)$ *and is* 0 *otherwise. Then* Y_t^n *is an adapted elementary process and, by continuity of Brownian motion, satisfies* $\lim_{n \to 0} Y_t^n = B_t$

almost surely as $\max_i \left|t_{i+1}^n - t_i^n\right| \to 0$*.The stochastic integral of* Y_t^n *is given by*

$$\begin{aligned}\int_0^T Y_t^n dB_t &= \sum_{i=0}^{n-1} B_{t_i^n}\left(B_{t_{i+1}^n} - B_{t_i^n}\right) \\ &= \frac{1}{2}\sum_{i=0}^{n-1}\left(B_{t_{i+1}^n}^2 - B_{t_i^n}^2 - \left(B_{t_{i+1}^n} - B_{t_i^n}\right)^2\right) \\ &= \frac{1}{2}B_T^2 - \frac{1}{2}B_0^2 - \frac{1}{2}\sum_{i=0}^{n-1}\left(B_{t_{i+1}^n} - B_{t_i^n}\right)^2.\end{aligned} \tag{1.4}$$

By Theorem (1.2.1) the sum on the right-hand-side of (1.4) converges in probability to T *as* $n \to \infty$*. And since* $B_0 = 0$ *we obtain*

$$\int_0^T B_t dB_t = \lim_{n\to 0}\int_0^T Y_t^n dB_t = \frac{1}{2}B_T^2 - \frac{1}{2}T.$$

Note that we will generally evaluate stochastic integrals using Itô's Lemma (to be discussed later) without having to take limits of elementary processes as we did in Example (1.2.1).

Definition 1.1.8

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

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

Theorem 1.1.1 (*Itô's Isometry*)

For any $Y_t(\omega) \in L^2[0,T]$ *we have*

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

Theorem 1.1.2 (*Martingale Property of Stochastic Integrals*)

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

1.1.4 Stochastic differential equations

Definition 1.1.9

An n-dimensional Itô process, Y_t, is a process of the form

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s, \tag{1.5}$$

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

We often use the notation

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

as shorthand for (1.5) .An n-dimensional stochastic diferential equation (SDE) has the form

$$dX_t = b_t (X_t, t)\, dt + \sigma_t (X_t, t)\, dB_t; \quad X_0 = 0, \tag{1.6}$$

where as before, B_t is an m- dimensional standard Brownian motion, and b and σ are n-dimensional and $n \times m$-dimensional adapted processes, respectively. Once again, (1.6) is shorthand for

$$X_t = x + \int_0^t b_s (X_s, s)\, ds + \int_0^t \sigma_s (X_s, s)\, dB_s. \tag{1.7}$$

While we do not discuss the issue here, various conditions exist to guarantee existence and uniqueness of solutions to (1.7). A useful tool for solving SDE's is Itô's Lemma which we now discuss.

1.1.5 Itô's lemma

Theorem 1.1.3 (*Itô's Lemma for 1-dimensional Brownian Motion*)

Let B_t be a Brownian motion on $[0, T]$ and suppose $f(x)$ is a twice continuously differentiable function on $\mathbb{R}$. Then for any $t \leq T$ we have

$$f(B_t) = f(0) + \frac{1}{2} \int_0^t f''(B_s)\, ds + \int_0^t f'(B_s)\, dB_s. \tag{1.8}$$

Proof: Let $0 = t_0 < t_1 < t_2 < ... < t_n = T$, be a partition of $[0, t]$. Clearly

$$f(B_t) = f(0) + \sum_{i=0}^{n-1} \left(f(B_{t_{i+1}}) - f(B_{t_i})\right). \tag{1.9}$$

Taylor's Theorem implies

$$f(B_{t_{i+1}}) - f(B_{t_i}) = f'(B_{t_i})(B_{t_{i+1}} - B_{t_i}) + \frac{1}{2} f''(\theta_i)(B_{t_{i+1}} - B_{t_i})^2, \tag{1.10}$$

for some $\theta_i \in (B_{t_{i+1}} - B_{t_i})$Substituting (1.10) into (1.9) we obtain

$$f(B_{t_i}) = f(0) + \sum_{i=0}^{n-1} f'(B_{t_i})(B_{t_{i+1}} - B_{t_i}) + \frac{1}{2}\sum_{i=0}^{n-1} f''(\theta_i)(B_{t_{i+1}} - B_{t_i})^2. \tag{1.11}$$

If we let $\delta := \max |t_{i+1} - t_i| \to 0$ then it can be shown that the terms on the right-hand-side of (1.11) converge to the corresponding terms on the right-hand-side of (1.8) as desired. (This should not be surprising as we know the quadratic variation of Brownian motion on $[0, t]$ is equal to t). ■

A more general version of Itô's Lemma can be stated for Itô processes.

Theorem 1.1.4 (*Itô's Lemma for 1-dimensional Itô process*)

Let X_t be 1-dimensional Itô process satisfying the SDE

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

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

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

The "Box" calculus

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

$$dt \times dt = dt \times dB_t = 0,$$

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

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

1.1.6 Some examples

Example 1.1.3

Suppose a stock price,S_t,satisfies the SDE

$$dS_t = \mu_t S_t dt + \int_0^t \sigma_{t_t} S_t dB_t.$$

Then we can use the substitution,$Y_t = \log(S_t)$ and Itô's Lemma applied to the function $f(x) := \log(x)$ to obtain

$$dS_t = S_0 \exp\left(\int_0^t \left(\mu_s - \sigma_s^2/2\right) ds + \int_0^t \sigma_s dB_s\right) \tag{1.12}$$

Note that S_t does not appear on the right-hand-side of (1.12) so that we have indeed solved the SDE. When $\mu_s = \mu$ and $\sigma_s = \sigma$ are constants we obtain

$$S_t = S_0 \exp\left(\left(\mu - \sigma^2/2\right) t + \sigma dB_t\right), \tag{1.13}$$

so that $\log(S_t) \sim N\left(\left(\mu - \sigma^2/2\right) t, \sigma^2 t\right)$.

Example 1.1.4 (*Ornstein-Uhlenbeck Process*)

Let S_t be a security price and suppose $X_t = \log(S_t)$ satisfies the SDE

$$dX_t = \left[-\gamma\left(X_t - \mu t\right) + \mu\right] dt + \sigma dB_t.$$

Then we can apply Itô's Lemma to $Y_t = \exp(\gamma t) X_t$ to obtain

$$\begin{aligned} dY_t &= \exp(\gamma t) dX_t + X_t d\left(\exp(\gamma t)\right) \\ &= \exp(\gamma t)\left(\left[-\gamma\left(X_t - \mu t\right) + \mu\right] dt + \sigma dB_t\right) + X_t \gamma \exp(\gamma t) dt \\ &= \exp(\gamma t)\left(\left[\gamma \mu t + \mu\right] dt + \sigma dB_t\right) \end{aligned}$$

so that

$$Y_t = Y_0 + \mu \int_0^t e^{\gamma s}(\gamma s + 1) ds + \sigma \int_0^t e^{\gamma s} dB_s, \tag{1.14}$$

or alternatively (after simplifying the Riemann integral in (1.14))

$$X_t = X_0 e^{-\gamma t} + \mu t + \sigma e^{-\gamma t} \int_0^t e^{\gamma s} dB_s. \tag{1.15}$$

Once again,note that X_t does note appear on the right-hand-side of (1.15) so that we have indeed solved SDE .We also abtain $\mathbb{E}(X_t) = X_0 e^{-\gamma t} + \mu t$ and

$$\begin{aligned} Var(X_t) = \; & Var\left(\sigma e^{-\gamma t} \int_0^t e^{\gamma s} dB_s\right) = \sigma^2 e^{-2\gamma t} \mathbb{E}\left[\left(\int_0^t e^{\gamma s} dB_s\right)^2\right] \\ & = \sigma^2 e^{-2\gamma t} \int_0^t e^{2\gamma s} ds \quad \text{(by Itô's Isometry)} \\ & = \frac{\sigma^2}{2\gamma}\left(1 - e^{-2\gamma s}\right). \end{aligned}$$

These moments should be compared with the corresponding moments for $log(S_t)$ in the previous example.

For more informations about stochastic calculus, we refer to [25].

1.2 Jump diffusions

In this part, we present the basic concepts needed for the applied calculus of jump diffusions. Since there are several excellent books which give a detailed account of this basic theory, we will just briefly review it here and refer the reader [40] for more information.

1.2.1 Lévy processes

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

Definition 1.2.1

An $\mathcal{F}_t$ adapted process $\{L(t)\}_{t\geq 0} = \{L_t\}_{t\geq 0} \subset \mathbb{R}$ is called a Lévy process if it satisties

1. $L_0 = 0$ *a.s.*
2. L_t *is continuous in probability*
3. L_t *is stationary, independent increments.*

Theorem 1.2.1

Let $\{L_t\}$ be a Lévy process. Then L_t has a càdlàg version (right continuous with left limits) which is also a Lévy process.

The jump of L_t at $t \geq 0$ is defined by

$$\triangle L_t = L_t - L_{t^-} \tag{1.16}$$

Let $\mathbf{B}_0$ be the family of Borel sets $U \subset \mathbb{R}$ whose closure $\overline{U}$ does not contain 0. For $U \in \mathbf{B}_0$ we define

$$N(t, U) = N(t, U, \omega) = \sum_{0<s\leq t} \chi_U(\triangle L_s). \tag{1.17}$$

In other words, $N(t, U)$ is the number of jumps of size $\triangle L_t \in U$ which occur before or at time t. $N(t, U)$ is called the Poisson random measure (or jump measure) of $L(t)$ The differential form of this measure is written $N(dt, dz)$.

Example 1.2.1 (*Brownian motion*)

Brownian motion $\{B(t)\}_{t\geq 0}$ has stationary and independent increments. Thus $B(t)$ is a Lévy process.

Example 1.2.2 (*The Poisson process*)

The Poisson process $\pi(t)$ of intensity $\lambda > 0$ is a Lévy process taking values in $\mathbb{N} \cup \{0\}$ and such that

$$P[\pi(t) = n] = \frac{(\lambda t)^n}{n!} e^{-\lambda t}; \quad n = 0, 1, 2, ...$$

Example 1.2.3 (*The compound Poisson process*)

Let $X(n)$; $n \in \mathbb{N}$ be a sequence of i.i.d. random variables taking values in $\mathbb{R}$ with common distribution $\mu_{X(1)} = \mu_X$ and let $\pi(t)$ be a Poisson process of intensity λ, independent of all the $X(n)$'s.

The compound Poisson process $Y(t)$ is defined by

$$Y(t) = X(1) + ... + X(\pi(t)); \quad t \geq 0. \tag{1.18}$$

An increment of this process is given by

$$Y(s) - Y(t) = \sum_{k=\pi(t+1)}^{\pi(s)} X(k); \quad s > t.$$

This is independent of $X(1), ..., X(\pi(t))$, and depends only on the difference $(s-t)$. Thus $Y(t)$ is a Lévy process.To find the Lévy measure ν of $Y(t)$ note that if $U \in B_0$ then

$$\nu(U) = \mathbb{E}[N(1,U)] = \mathbb{E}\left[\sum_{s;0\leq s\leq 1} \chi_U(\triangle Y(s))\right]$$
$$= \mathbb{E}[(\text{number of jumps}) . \chi_U(\text{jumps})] = \mathbb{E}[\pi(1)\chi_U(X)] = \lambda\mu_X(U),$$

by independence.We conclude that

$$\nu = \lambda\mu_X. \tag{1.19}$$

This shows that a Lévy process can be represented by a compound Poisson process if and only if its Lévy measure is finite. Note, however, that there are many interesting Lévy processes with infinite Lévy measure. See e.g [1]

Theorem 1.2.2 (*Lévy decomposition [30]*)

Let $\{L_t\}$ be a Lévy process. Then L_t has the decomposition

$$L_t = \alpha t + \beta B(t) + \int_{|z|<R} z\tilde{N}(t,dz) + \int_{|z|\geq R} zN(t,dz), \tag{1.20}$$

for some constant $\alpha \in \mathbb{R}, \beta \in \mathbb{R}, R \in [0,\infty]$ .Here

$$\tilde{N}(t,dz) = N(t,dz) - \nu(dz)\,dt, \tag{1.21}$$

is the compensated Poisson random measure of $L(\cdot)$ and $B(t)$ is an independent Brownian motion. For each $A \in B_0$ the process

$$M_t := \tilde{N}(t,A) \ \text{ is a martingale.} \tag{1.22}$$

If $\alpha = 0$ and $R = \infty$,we call L_t a Lévy martingale .

Theorem 1.2.3 ([43], ***Corollary p. 48***)

A Lévy process is a semimartingale.

Definition 1.2.2 ([43])

Let $\mathbb{D}_{ucp}$ denote the space of cadlag adapted processes,equipped with the topology of uniform convergence on compacts in probability (ucp) : $H_n \to H$ ucp if for all $t > 0$ $\sup |H_n(s) - H(s)| \to 0$ in probability ($A_n \to A$) in probability if for all $\theta > 0$ there exists $n_\theta \in \mathbb{N}$ such that $n \geq n_\theta \Rightarrow P(|A_n \to A| > \theta) < \theta$.

Let $\mathbb{L}_{ucp}$ denote the space of adapted caglad processes (left continuous with right limits), equipped with the ucp topology. If $H(t)$ is a step function of the form

$$H(t) = H_0\chi_{\{0\}}(t) + \sum_i H_i\chi_{\{T_i,T_{i+1}\}}(t),$$

where $H_i \in \mathcal{F}_{T_i}$ and $0 = T_0 \leq T_1 \leq ... \leq T_{n+1} < \infty$ are $\mathcal{F}_t$-stopping times and X is cadlag, we define

$$J_X H(t) := \int_0^t H_s dX_s := H_0X_0 + \sum_i H_i \left(X_{T_{i+1}\wedge t} - X_{T_i \wedge t}\right); \qquad t \geq 0.$$

Theorem 1.2.4 ([43] ***,p.51***)

Let X be a semimartingale.Then the mapping J_X can be extended to a continuous linear map

$$J_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}.$$

This construction allows us to define stochastic integrals of the form

$$\int_0^t H(s)\, dL_s,$$

for all $H \in \mathbb{L}_{ucp}$. (See also Remark 1.3.2). In view of the decomposition (1.20) this integral can be split into integrals with respect to ds, $dB(s)$, $N(ds,dz)$ and $\tilde{N}(ds,dz)$.This makes it natural to consider the more general stochastic integrals of the form

$$X(t) = X(0) + \int_0^t \alpha(s,\omega)\, ds + \int_0^t \beta(s,\omega)\, dB(s) + \int_0^t \int_{\mathbb{R}} \gamma(s,z,\omega)\, \bar{N}(ds,dz), \quad (1.23)$$

where the integrands are satisfying the appropriate conditions for the integrals to exist *and we for simplicity have put*

$$\bar{N}(ds,dz)\begin{cases} N(ds,dz)-\nu\,(dz)\,ds & \text{if } |z|<R,\\ N(ds,dz) & \text{if } |z|\geq R,\end{cases}$$

with R as in (Theorem 1.3.3). As is customary we will use the following short hand differential notation for processes $X\,(t)$ satisfying (1.23):

$$dX_t=\alpha\,(t)\,dt+\beta\,(t)\,dB\,(t)+\int_{\mathbb{R}}\gamma\,(t,z)\,\bar{N}(ds,dz). \tag{1.24}$$

We call such processes Itô-Lévy processes .

1.2.2 Itô Formula with Jumps

We now come to the important Itô formula for Itô-Lévy processes:

If $X\,(t)$ is given by (1.24) and $f:\mathbb{R}^2\to\mathbb{R}$ is a C^2 function,is the process $Y\,(t):=f\,(t,X\,(t))$ again an Itô-Lévy process and if so, how do we represent it in the form (1.24)?

If we argue heuristically and use our knowledge of the classical Itô formula it is easy to guess what the answer is:

Let $X^{(c)}\,(t)$ be the continuous part of $X\,(t)$, i.e. $X^{(c)}\,(t)$ is obtained by removing the jumps from $X\,(t)$. Then an increment in $Y(t)$ stems from an increment in $X^{(c)}\,(t)$ plus the jumps (coming from $N(\,\cdot\,,\cdot\,)$). Hence in view of the classical Itô formula we would guess that

$$\begin{aligned} dY\,(t) &= \frac{\partial f}{\partial t}(t,X\,(t))\,dt+\frac{\partial f}{\partial x}(t,X\,(t))\,dX^c\,(t)+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,X_t)\,\beta^2\,(t)\,dt\\ &+\int_{\mathbb{R}}\left\{f\left(t,X\left(t^-\right)+\gamma\,(t,z)\right)-f\left(t,X\left(t^-\right)\right)\right\}N(ds,dz).\end{aligned}$$

It can be proved that our guess is correct. Since

$$dX^c\,(t)=\left(\alpha\,(t)-\int_{|z|<R}\gamma\,(t,z)\,\nu\,(dz)\right)dt+\beta\,(t)\,dB\,(t)\,,$$

this gives the following result:

Theorem 1.2.5 (***The 1-dimensional Itô formula*** *[43]*)

Suppose $X\,(t)\in\mathbb{R}$ is an Itô-Lévy process of the form

$$X\,(t)=\alpha\,(t,\omega)\,dt+\beta\,(t,\omega)\,dB\,(t)+\int_{\mathbb{R}}\int_{|z|\leq R}\gamma\,(t,z,\omega)\,\bar{N}(dt,dz), \tag{1.25}$$

where

$$\bar{N}(ds,dz)\begin{cases} N(ds,dz)-\nu\left(dz\right)ds & \text{if } |z|<R, \\ N(ds,dz) & \text{if } |z|\geq R, \end{cases}, \tag{1.26}$$

for some $R\in[0,\infty]$.

Let $f\in C^2\left(\mathbb{R}^2\right)$ *and define* $Y\left(t\right)=f\left(t,X\left(t\right)\right)$. *Then* $Y\left(t\right)$ *is again an Itô-Lévy process and*

$$\begin{aligned} dY_t &= \frac{\partial f}{\partial t}\left(t,X\left(t\right)\right)dt+\frac{\partial f}{\partial x}\left(t,X\left(t\right)\right)\left(\left[\alpha\left(t,\omega\right)+\beta\left(t,\omega\right)dB_t\right]\right)+\frac{1}{2}\beta^2\left(t,\omega\right)\frac{\partial^2 f}{\partial x^2}\left(t,X_t\right)dt \\ &= \int\limits_{|z|<R}\left\{f\left(t,X\left(t^-\right)+\gamma\left(t,z\right)\right)-f\left(t,X\left(t^-\right)\right)-\frac{\partial f}{\partial x}\left(t,X\left(t^-\right)\right)\gamma\left(t,z\right)\right\}\nu\left(dz\right). \\ &+\int_{\mathbb{R}}\left\{f\left(t,X\left(t^-\right)+\gamma\left(t,z\right)\right)-f\left(t,X\left(t^-\right)\right)\right\}\bar{N}(dt,dz). \end{aligned} \tag{1.27}$$

Note**:** If $R=0$ then $\bar{N}=N$ everywhere.

If $R=0$ then $\bar{N}=\tilde{N}$ everywhere.

Lemma 1.2.1 (*Integration by parts formula for jumps processes*)

Suppose that the processes $x_i(t)$ *are given by: for* $i=1,2,\ t\in[0,T]$:

$$\begin{cases} dx_i(t)=b\left(t,x_i(t),u(t)\right)dt+\sigma\left(t,x_i(t),u(t)\right)dW(t) \\ \qquad\qquad +\int_Z\eta\left(t,x_i(t_-),z\right)\widetilde{N}\left(dz,dt\right), \\ x_i(0)=0. \end{cases}$$

Then we get

$$\begin{aligned} \mathbb{E}\left(x_1(T)x_2(T)\right) &= \mathbb{E}\left[\int_0^T x_1(t)dx_2(t)+\int_0^T x_2(t)dx_1(t)\right] \\ &+\mathbb{E}\int_0^T\sigma^\intercal\left(t,x_1(t),u(t)\right)\sigma\left(t,x_2(t),u(t)\right)dt \\ &+\mathbb{E}\int_0^T\int_Z\eta^\intercal\left(t,x_1(t),z\right)\eta\left(t,x_2(t),z\right)\mu(dz)dt. \end{aligned}$$

Proposition 1.2.1

Let $\mathcal{G}$ be the predictable $\sigma-$field on $\Omega\times[0,T]$, $\mu(Z)<\infty$, and f be a $\mathcal{G}\times\mathcal{B}(Z)-$measurable function such that.

$$\mathbb{E}\int_0^T\int_Z |f(s,z)|^2\,\mu(dz)ds<\infty,$$

then for all $k\geq 2$ there exists a positive constant $C_{(k,\mu(Z))}>0$ such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t\int_Z f(s,z)\,\widetilde{N}(dz,ds)\right|^k\right]\leq C_{(k,\mu(Z))}\mathbb{E}\left[\int_0^T\int_Z |f(s,z)|^k\,\mu(dz)ds\right].$$

Proof: See Bouchard et al., [6, Appendix]. ■

Theorem 1.2.6 (*The Itô-Lévy isometry*)

Let $X(t)\in\mathbb{R}^n$, with $X(t)=0$ and $\alpha=0$. Then

$$\begin{aligned}\mathbb{E}\left[X^2(T)\right] &= \mathbb{E}\left[\int_0^T\left\{\sum_{i=1}^n\sum_{j=1}^m\sigma_{ij}^2(t)+\sum_{i=1}^n\sum_{j=1}^\ell\int_{\mathbb{R}}\gamma_{ij}^2(t,z_j)\,\nu_j(dz_j)\right\}dt\right]\\ &=\sum_{i=1}^n\mathbb{E}\left[\int_0^T\left\{\sum_{j=1}^m\sigma_{ij}^2(t)+\sum_{j=1}^\ell\int_{\mathbb{R}}\gamma_{ij}^2(t,z_j)\,\nu_j(dz_j)\right\}dt\right],\end{aligned}\tag{1.28}$$

provided that the right hand side is finite.

1.2.3 Stochastic differential equations with jumps

The geometric Lévy process is an example of a **Lévy diffusion**, i.e. the solution of a stochastic differential equation (SDE) driven by Lévy processes.

Theorem 1.2.7 (*Existence and uniqueness of solutions of Lévy SDEs*)

Consider the following Lévy SDE in $\mathbb{R}^n$: $X(0)=x_0$ and

$$dX(t)=\alpha(t,X(t))\,dt+\sigma(t,X(t))\,dB(t)+\int_{\mathbb{R}^n}\gamma\left(t,X\left(t^-\right),z\right)\tilde{N}(dt,dz),\tag{1.29}$$

where $\alpha:[0,T]\times\mathbb{R}^n\to\mathbb{R}^n, \sigma:[0,T]\times\mathbb{R}^n\to\mathbb{R}^{n\times m}$ and $\gamma:[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^{n\times\ell}$ satisfy the following conditions

(i) (At most linear growth) There exists a constant $C_1<\infty$ such that

$$\|\sigma(t,x)\|^2+|\alpha(t,x)|^2+\int_{\mathbb{R}}\sum_{j=1}^\ell|\gamma_k(t,x,z)|^2\,\nu_k(dz_k)\leq C_1\left(1+|x|^2\right);\ \textit{for all } x\in\mathbb{R}^n.$$

(ii) (Lipschitz continuity) There exists a constant $C_2 < \infty$ such that

$$\begin{aligned}
&\left\|\sigma(t,x)-\sigma(t,y)\right\|^2+\left|\alpha(t,x)-\alpha(t,y)\right|^2\\
&+\sum_{j=1}^{\ell}\int_{\mathbb{R}}\left|\gamma^{(k)}(t,x,z_k)-\gamma^{(k)}(t,y,z_k)\right|^2\nu_k(dz_k)\le C_2\left(1+|x-y|^2\right);\\
&\text{for all } x,y\in\mathbb{R}^n.
\end{aligned}$$

Then there exists a unique cadlag adapted solution $X(t)$ such that

$$\mathbb{E}\left[X^2(T)\right]<\infty, \text{for all } t.$$

Solutions of Lévy SDEs in the time homogeneous case, i.e. when $\alpha(t,x) = \alpha(x), \sigma(t,x) = \sigma(x)$ and $\gamma(t,x,z) = \gamma(x,z)$, are called jump diffusions (or Lévy diffusions).

Definition 1.2.3

Let $X(t)\in\mathbb{R}^n$ be a jump diffusion.Then the generator A of X is defined on functions $f:\mathbb{R}^n\to\mathbb{R}$ by

$$Af(x)=\lim_{t\to 0^+}\frac{1}{t}\left\{\mathbb{E}^x\left[f(X(t))\right]-f(x)\right\} \qquad \text{(if the limit exists)},$$

where $\mathbb{E}^x\left[f(X(t))\right]=\mathbb{E}\left[f(X^x(t))\right], X^x(0)=x$.

Theorem 1.2.8

Suppose $f\in C_0^2(\mathbb{R}^n)$ .Then $Af(x)$ exists and is given by

$$\begin{aligned}
Af(x) &= \sum_{i=1}^{n}\alpha_i(x)\frac{\partial f}{\partial x_i}(x)+\frac{1}{2}\sum_{i,j=1}^{n}\left(\sigma\sigma^T\right)_{ij}(x)\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\\
&+\int_{\mathbb{R}}\sum_{k=1}^{\ell}\left\{f\left(x+\gamma^{(k)}(x,z)\right)-f(x)-\nabla f(x)\gamma^{(k)}(x,z)\right\}\nu_k(dz_k).
\end{aligned} \tag{1.30}$$

From now on we define $Af(x)$ by the expression (1.30) for all f such that the partial derivatives of f and the integrals in (1.30) exist at x.

Chapter 2

Stochastic optimal control problems

The problem of stochastic optimal control is to control a system in such a way as to do something to it optimally. This theory is part of a larger field called control theory. The applications of this type of problem are very numerous and in very diverse fields, such as finance, mechanics, biology, electricity, chemistry, economics, etc ...

There are two well-known approaches to solving the optimal control problem, which are the stochastic maximum principle and the principle of dynamic programming.

The study of optimal control problems by using Bellman's Dynamic Programming Principle can be linked with the solution of a particular class of nonlinear second order partial differential equations: the Hamilton-Jacobi-Bellman equations.

Stochastic Maximum Principle is to study a set of necessary and sufficient conditions that must be satisfied by any optimal control, the basic idea is by perturbing an optimal control on a small time interval of length θ. Performing a Taylor expansion with respect to θ and then sending θ to zero one obtains a variational inequality. By duality the stochastic maximum principle is obtained. For more informations about the two approaches, we can see, Lakhdari [36].

2.1 Problem formulation

In this section, we present the strong and weak formulations of stochastic control problem.

2.1.1. Strong formulation

Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ be a filtered probability space, on witch we define an m-dimensional standard Brownian motion $B(\cdot)$, denote by U the separable metric space. We denote by $\mathcal{U}_{ad}[0,T]$ the set of all admissible controls.

The state of a controlled diffusion is described by the SDE

$$\begin{cases} dy(t) & = b(t, y(t), u(t))\, dt + \sigma(t, y(t), u(t))\, dB(t) \\ y(0) & = y, \end{cases} \tag{2.1}$$

where $b : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^n, \sigma : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times m}$, are given. $u(\cdot)$ is called the control representing the action of the decision-makers (controllers). At any time instant the controller knowledgeable about some information (as specified by the information filed $\{\mathcal{F}_t\}_{t\geq 0}$) of what has happened up to that moment, but not able to foretell what is going to happen afterwards due to the uncertainty of the system (as a consequence, for any t the controller cannot exercise his/her decision $u(t)$ befor the time t really comes) witche can be expressed in mathematical term as " $u(\cdot)$ is $\{\mathcal{F}_t\}_{t\geq 0}$ adapted", the control u is taken from the set

$$\mathcal{U}[0,T] \triangleq \left\{ u : [0,T] \times \Omega \longrightarrow U \mid u(\cdot) \text{ is } \{\mathcal{F}_t\}_{t\geq 0} \text{ adapted} \right\}.$$

The cost functional has the form:

$$J(u(\cdot)) = \mathbb{E}\left[\int_0^T f(t, y(t), u(t))\, dt + g(y(T)) \right].$$

Definition 2.1.1

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be given filtered probability space satisfying the usual conditions and let $B(t)$ be a given m-dimensional standard $\{\mathcal{F}_t\}_{t\geq 0}$-Brownian motion. A control $u(\cdot)$ called an admissible control, and $(y(\cdot), u(\cdot))$ an admissible pair, if

i) $y(\cdot)$ is the unique solution of equation (2.1).

ii) $f(\cdot, y(\cdot), u(\cdot)) \in L^1_{\mathcal{F}}(0, T, \mathbb{R})$ and $g(y(T)) \in L^1_{\mathcal{F}}(\Omega, \mathbb{R})$.

iii) $u(\cdot) \in \mathcal{U}[0,T]$.

Stochastic control problem is to find an optimal control $\widehat{u}(\cdot) \in \mathcal{U}[0,T]$ (if it ever exists), such that

$$J(\widehat{u}(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}[0,T]} J(u(\cdot)),$$

where $\widehat{u}(\cdot)$ is called an optimal control and the state control pair $(\widehat{y}(\cdot), \widehat{u}(\cdot))$ are called an optimal state process.

2.1.2. Weak formulation

In the strong formulation the filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ on witch we define the Brownian motion B are all fixed. However in the weak formulation, where we consider them as a parts of the control.

Definition 2.1.2

$\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, B(\cdot), u(\cdot)\right)$ is called a w-admissible control, and $y(\cdot), u(\cdot)$ a w-admissible pair if

1. *$\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ is a filtered probability space satisfying the usual conditions.*
2. *$B(\cdot)$ is an m-dimensional standard Brownian motion defined on $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\right)$.*
3. *$u(\cdot)$ is an $\{\mathcal{F}_t\}_{t\geq 0}$-adapted process on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in U.*
4. *$y(\cdot)$ is the unique solution of equation (2.1).*
5. *$f(\cdot, y(\cdot), u(\cdot)) \in L^1_{\mathcal{F}}(0, T, \mathbb{R})$ and $g(y(T)) \in L^1_{\mathcal{F}}(\Omega, \mathbb{R})$.*

The set of all admissible controls is denoted by $\mathcal{U}[0, T]$. Our stochastic optimal control problem under weak formulation is to find an optimal control $\widehat{u}(\cdot) \in \mathcal{U}[0, T]$ (if it ever exists), such that

$$J(\widehat{u}(\cdot)) = {}_{u(\cdot)\in\mathcal{U}}[0, T] \inf J(u(\cdot)).$$

2.2 Dynamic programming principle

2.2.1. The Bellman principle

Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\leq T}, \mathbb{P}\right)$ be a filtered probability space and $B(t)$ a Brownian motion valued in $\mathbb{R}^d$. We denote by A the set of all progressively measurable processes $\{u(t)\}_{t\geq 0}$ valued in $U \subset \mathbb{R}^k$.

The state of the stochastic controlled system has the form:

$$\begin{cases} dy(t) & = b(t, y(t), u(t))\, dt + \sigma(t, y(t), u(t))\, dB(t) \\ y(0) & = y, \end{cases} \tag{2.2}$$

where $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times d}$ be two given functions satisfying, for some constant M

$$|b(t,y(t),u(t)) - b(t,x(t),u(t))| + |\sigma(t,y(t),u(t)) - \sigma(t,x(t),u(t))| \leq M|y-x|, \tag{2.3}$$

$$|b(t,y(t),u(t))| + |\sigma(t,y(t),u(t))| \leq M(1+|y(t)|). \tag{2.4}$$

Under (2.3) and (2.4) the above equation has a unique solution y.

The cost functional $J : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$, defined by

$$J(t,y,u) = \mathbb{E}^{t,y}\left[\int_t^T f(s,y(s),u(s))\,ds + g(y(T))\right], \tag{2.5}$$

where $\mathbb{E}^{t,y}$ is the expectation operator conditional on $y(t) = y$, and $f : [0,T]\times\mathbb{R}^n\times U \longrightarrow \mathbb{R}$, $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, we assume that

$$|f(t,y,u)| + |g(y)| \leq M\left(1+|y|^2\right), \tag{2.6}$$

for some constant M. The quadratic growth condition (2.6), ensure that J is well defined. The purpose of this Section is to study the minimization problem

$$V(t,y) = \inf_{u\in U} J(t,y,u), \quad \text{for } (t,y) \in [0,T]\times\mathbb{R}^n, \tag{2.7}$$

which is called the value function of the problem (2.2) and (2.5).

The dynamic programming is a fundamental principle in the theory of stochastic control, we give a version of the stochastic Bellman's principle of optimality. For mathematical treatments of this problem , we refer the reader to Lions [35], Krylov [34], Yong and Zhou [55], Fleming and Soner [19], Lakhdari [36].

Theorem 2.2.1

Let $(t,y) \in [0,T]\times\mathbb{R}^n$ be given. Then, for every $h \in [0,T-t]$, we have

$$V(t,y) = \inf_{u\in U} \mathbb{E}^{t,y}\left(\int_t^{t+h} f(s,y(s),u(s))\,ds + V(t+h,y(t+h))\right). \tag{2.8}$$

Proof: Suppose that for $h>0$, we given by $\widehat{u}(s) = \widehat{u}(s,y)$ the optimal feedback control for the problem (2.2) and (2.5) over the time interval $[t,T]$ starting at point $y(t+h)$. i.e.

$$J(t+h,y(t+h),\widehat{u}(t+h)) = V(t+h,y(t+h)),\ \mathbb{P}-a.s. \tag{2.9}$$

Now, we consider

$$\widetilde{u} = \begin{cases} u(s,y), & t \le s \le t+h \\ \widehat{u}(s,y), & t+h \le s \le T, \end{cases}$$

for some control u. By definition of $V(t,y)$, and using (2.5), we obtain

$$\begin{aligned} V(t,y) &\le J(t,y,\widetilde{u}) \\ &= \mathbb{E}^{t;y}\left(\int_t^{t+h} f(s,y(s),u(s))\,ds + \int_{t+h}^{T} f(s,y(s),\widehat{u}(s))\,ds + g(y(T))\right). \end{aligned}$$

By the unicity of solution for the SDE (2.2), we have for s$\ge t+h$, $y^{t+h,y^{t,y}(t+h)}(s) = y^{t,y}(s)$, then

$$\begin{aligned} J(t,y,\widetilde{u}) &= \mathbb{E}\left(\int_t^{t+h} f(s,y(s),u(s))\,ds \right. \\ &\quad \left. + \int_{t+h}^{T} f\left(s, y^{t+h,y^{t,y}(t+h)}(s), \widehat{u}(s)\right) ds + g\left(y^{t+h,y^{t,y}(t+h)}(T)\right)\right) \\ &= \mathbb{E}\left(\int_t^{t+h} f(s,y(s),u(s))\,ds \right. \\ &\quad \left. + \mathbb{E}\int_{t+h}^{T} f(s,y(s),\widehat{u}(s))\,ds + g(y(T)) \mid y^{t,y}(t+h)\right) \\ &= \mathbb{E}\left(\int_t^{t+h} f(s,y(s),u(s))\,ds + V\left(t+h, y^{t,y}(t+h)\right)\right). \end{aligned}$$

So we get

$$V(t,y) \le \mathbb{E}\left(\int_t^{t+h} f(s,y(s),u(s))\,ds + V\left(t+h, y^{t,y}(t+h)\right)\right), \tag{2.10}$$

and the equality holds if $\widetilde{u} = \widehat{u}$, which proves (2.8). ■

2.2.2 The Hamilton Jacobi Bellman equation

Now, we introduce the HJB equation by deriving it form the dynamic programming principle under smoothness assumptions on the value function. Let $G : [0,T] \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n\times d}$ into $\mathbb{R}$, be defined by

$$G(t,y,r,p,A) = b(t,y,u)^{\intercal} p + \frac{1}{2} tr\left[\sigma\sigma^{\intercal}(t,y,u)\,A\right] + f(t,y,u), \tag{2.11}$$

we also need to introduce the linear second order operator $\mathcal{L}^u$ associated to the controlled processes $y(t)$, $t \ge 0$, we consider the constant control u

$$\mathcal{L}^u\varphi(t,y) = b(t,y,u)^\intercal D_y\varphi(t,y) + \frac{1}{2}tr\left[\sigma\sigma^\intercal(t,y,u)\,D_{yy}\left(\varphi(t,y)\right)\right], \tag{2.12}$$

where D_y, D_{yy} denote the gradient and the Hessian operator with respect to the y variable. Assume the value function $V \in C\left([0,T],\mathbb{R}^n\right)$, and $f(\cdot,\cdot,u)$ be continuous in (t,y) for all fixed $u \in A$, then we have by Itô's formula

$$V\left(t+h, y\left(t+h\right)\right) = V\left(t,y\right) + \int_t^{t+h}\left(\frac{\partial V}{\partial s} + \mathcal{L}^u V\right)\left(s, y^{t,y}\left(s\right)\right)ds$$
$$+ \int_t^{t+h} D_y V\left(s, y^{t,y}\left(s\right)\right)^\intercal \sigma\left(s, y^{t,y}\left(s\right), u\right)dB\left(s\right),$$

by taking the expectation, we get

$$\mathbb{E}\left(V\left(t+h, y\left(t+h\right)\right)\right) = V\left(t,y\right) + \mathbb{E}\left(\int_t^{t+h}\left(\frac{\partial V}{\partial s} + \mathcal{L}^u V\right)\left(s, y^{t,y}\left(s\right)\right)ds\right),$$

then, we have by (2.10)

$$0 \leq \mathbb{E}\left(\frac{1}{h}\int_t^{t+h}\left(\left(\frac{\partial V}{\partial s} + \mathcal{L}^u V\right)\left(s, y^{t,y}\left(s\right)\right) + f\left(s, y^{t,y}\left(s\right), u\right)\right)ds\right).$$

We now send h to zero, we obtain

$$0 \leq \frac{\partial V}{\partial t}\left(t,y\right) + \mathcal{L}^u V\left(t,y\right) + f\left(t,y,u\right),$$

this provides

$$-\frac{\partial V}{\partial t}\left(t,y\right) - \inf_{u\in U}\left[\mathcal{L}^u V\left(t,y\right) + f\left(t,y,u\right)\right] \leq 0. \tag{2.13}$$

Now we shall assume that $\widehat{u} \in U$, and using the same procedure as above, we conclude that

$$-\frac{\partial V}{\partial t}\left(t,y\right) - \mathcal{L}^{\widehat{u}} V\left(t,y\right) - f\left(t,y,u\right) = 0, \tag{2.14}$$

by (2.13), then the value function solves the HJB equation

$$-\frac{\partial V}{\partial t}\left(t,y\right) - \inf_{u\in U}\left[\mathcal{L}^u V\left(t,y\right) + f\left(t,y,u\right)\right] = 0, \qquad \forall\left(t,y\right) \in [0,T] \times \mathbb{R}^n. \tag{2.15}$$

We give sufficient conditions which allow to conclude that the smooth solution of the HJB equation coincides with the value functionm this is the so-called verification result.

Theorem 2.2.2

Let W be a $C^{1,2}([0,T],\mathbb{R}^n) \cap C([0,T],\mathbb{R}^n)$ function. Assume that f and g are quadratic growth, i.e. there is a constant M such that

$$|f(t,y,u)| + |g(y)| \leq M\left(1+|y|^2\right), \text{ for all } (t,y,u) \in [0,T]\times\mathbb{R}^n\times U.$$

(1) Suppose that $W(T,\cdot) \leq g$, and

$$\frac{\partial W}{\partial t}(t,y) + G(t,y,W(t,y),D_yW(t,y),D_{yy}(W(t,y))) \geq 0, \tag{2.16}$$

on $[0,T]\times\mathbb{R}^n$, then $W \leq V$ on $[0,T]\times\mathbb{R}^n$.

(2) Assume further that $W(T,\cdot) = g$, and there exists a minimizer $\widehat{u}(t,y)$ of

$$\mathcal{L}^u V(t,y) + f(t,y,u),$$

such that

$$0 = \frac{\partial W}{\partial t}(t,y) + G(t,y,W(t,y),D_yW(t,y),D_{yy}(W(t,y))) \tag{2.17}$$

$$= \frac{\partial W}{\partial t}(t,y) + \mathcal{L}^{\widehat{u}(t,y)}W(t,y) + f(t,y,u), \tag{1.22}$$

the stochastic differential equation

$$dy(t) = b(t,y(t),\widehat{u}(t,y))\,dt + \sigma(t,y(t),\widehat{u}(t,y))\,dB(t), \tag{2.18}$$

defines a unique solution $y(t)$ for each given initial data $y(t) = y$, and the process $\widehat{u}(t,y)$ is a well-defined control process in U. Then $W = V$, and $\widehat{u}$ is an optimal Markov control process.

Proof: The function $W \in C^{1,2}([0,T],\mathbb{R}^n) \cap C([0,T],\mathbb{R}^n)$, then for all $0 \leq t \leq s \leq T$, by Itô's Lemma we get

$$W\left(t,y^{t,y}(r)\right) = \int_t^s \left(\frac{\partial W}{\partial t} + \mathcal{L}^{u(r)}W\right)\left(r,y^{t,y}(r)\right)dr + \int_t^s D_yW\left(r,y^{t,y}(r)\right)^{\intercal}\sigma\left(r,y^{t,y}(r),u(r)\right)dB(r),$$

the process $\int_t^s D_yW\left(r,y^{t,y}(r)\right)^{\intercal}\sigma\left(r,y^{t,y}(r),u(r)\right)$, is a martingale, then by taking expectation, it follows that

$$\mathbb{E}\left[W\left(s, y^{t,y}(s)\right)\right] = W(t,y) + \mathbb{E}\left(\int_t^s \left(\frac{\partial W}{\partial t} + \mathcal{L}^{u(r)}W\right)\left(r, y^{t,y}(r)\right) dr\right).$$

By (2.16), we get

$$\frac{\partial W}{\partial t}\left(r, y^{t,y}(r)\right) + \mathcal{L}^{u(r)}W\left(r, y^{t,y}(r)\right) + f\left(r, y^{t,y}(r), u(r)\right) \geq 0, \ \forall u \in A,$$

then

$$\mathbb{E}\left[W\left(s, y^{t,y}(s)\right)\right] \geq W(t,y) - \mathbb{E}\left(\int_t^s f\left(r, y^{t,y}(r), u(r)\right) dr\right), \ \forall u \in A,$$

we now take the limit as $s \longrightarrow T$, then by the fact that $W(T) \leq g$ we obtain

$$\mathbb{E}\left[g\left(y^{t,y}(T)\right)\right] \geq W(t,y) - \mathbb{E}\left(\int_t^s f\left(r, y^{t,y}(r), u(r)\right) dr\right), \ \forall u \in A,$$

then $W(t,y) \leq V(t,y), \forall (t,y) \in [0,T] \times \mathbb{R}^n$. Statement (2) is proved by repeating the above argument and observing that the control $\widehat{u}$ achieves equality at the crucial step (2.16).

We now state without proof an existence result for the HJB equation (2.15), together with the terminal condition $W(T,y) = g(y)$. ■

Theorem 2.2.3

assume that

1. $\exists C > 0 / \xi^\intercal \sigma \sigma^\intercal (t,y,u)\, \xi \geq C |\xi|^2$, *for all* $(t,y,u) \in [0,T] \times \mathbb{R}^n \times U$,
2. U *is compact,*
3. b, σ *and* f *are in* $C_b^{1,2}([0,T], \mathbb{R}^n)$,
4. $g \in C_b^3(\mathbb{R}^n)$,

Then the HJB equation (1.20), with the terminal data $V(T,y) = g(y)$, *has a unique solution* $V \in C_b^{1,2}([0,T], \mathbb{R}^n)$.

Proof: See Fleming and Rischel [18]. ■

2.3 Stochastic maximum principle

The basic idea of the stochastic maximum principle is to derive a set of necessary and sufficient conditions that must be satisfied by any optimal control. The first version of the stochastic maximum principle was established by Bismut [4], Kushner [33], and Haussmann [21], under the condition that there is no control on the diffusion coefficient. Haussman [22], developed a powerful form of stochastic maximum principle for the feedback class of controls by Girsanov's transformation, and applied it to solve some problems in stochastic control.

2.3.1. Formulation of the problem

Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\leq T}, \mathbb{P}\right)$ be a probability space such that $\mathcal{F}_0$ contains the $\mathbb{P}$-null sets, $\mathcal{F}_T = \mathcal{F}$ for an arbitrarily fixed time horizon T. We assume that $\{\mathcal{F}_t\}_{t\leq T}$ is generated by a d-dimensional standard Brownian motion B. We denote by $\mathcal{U}$ the set of all admissible controls. Any element $y \in \mathbb{R}^n$ will be identified to a column vector with n components, and the norm $|y| = \left|x^1\right| + ... + |x^n|$. The scalar product of any two vectors y and x on $\mathbb{R}^n$ is denoted by yx or $\sum_{i=1}^n y^i x^i$. For a function h, we denote by h_y (resp. h_{yy}) the gradient or Jacobian (resp. the Hessian) of h with respect to the variable y.

Definition 2.3.1

An admissible control is a measurable, adapted processes $u : [0, T] \times \Omega \to U$, such that $\mathbb{E}\left[\int_0^T u(s)\, ds\right] < \infty$.

Consider the following stochastic controlled system

$$\begin{cases} dy(t) & = b(t, y(t), u(t))\, dt + \sigma(t, y(t), u(t))\, dB(t) \\ y(0) & = y, \end{cases} \tag{2.19}$$

where $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n, \sigma : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times d}$, are given.

Suppose we are given a performance functional $J(u)$ of the form

$$J(u) = \mathbb{E}\left[\int_0^T f(t, y(t), u(t))\, dt + g(y(T))\right], \tag{2.20}$$

where $f : [0, T] \times \mathbb{R}^n \times U_1 \to \mathbb{R}, \quad g : \mathbb{R}^n \to \mathbb{R}$.

The stochastic control problem is to find an optimal control $\widehat{u} \in \mathcal{U}$ such that

$$J(\widehat{u}) = \inf_{u\in\mathcal{U}} J(u), \tag{2.21}$$

Let us make the following assumptions about the coefficients b, σ, f, and g.

(H1) The maps b, σ, and f are continuously differentiable with respect to (y, u), and g is continuously differentiable in y.

(H2) The derivatives $b_y, b_u, \sigma_y, \sigma_u, f_y, f_u$, and g_y are continuous in (y, u) and uniformly bounded.

(H3) b, σ, f are bounded by $K_1 (1 + |y| + |u|)$, and g is bounded by $K_1 (1 + |y|)$, for some $K_1 > 0$.

2.3.2. *The stochastic maximum principle*

Now, define the Hamiltonian $H : [0, T] \times \mathbb{R}^n \times \overline{U} \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \longrightarrow \mathbb{R}$, by

$$H(t, y, u, p, q) = f(t, y, u) + pb(t, y, u) + \textstyle\sum_{j=1}^{n} q^j \sigma^j (t, y, u), \tag{2.22}$$

where q^j and σ^j for $j = 1, .., n$, denote the jth column of the matrix q and σ, respectively.

Let $\widehat{u}$ be an optimal control and $\widehat{y}$ denote the corresponding optimal trajectory. Then, we consider a pair (p, q) of square integrable adapted processes associated to $\widehat{u}$, with values in $\mathbb{R}^n \times \mathbb{R}^{n \times d}$ such that

$$\begin{cases} dp(t) = & -H_y(t, \widehat{y}(t), \widehat{u}(t), p(t), q(t))dt + q(t)\, dB(t), \\ p(T) = & g_y(\widehat{y}(T)). \end{cases} \tag{2.23}$$

2.3.3. *Necessary conditions of optimality*

The purpose of this part is to find optimality necessary conditions satisfied by an optimal control, assuming that the solution exists. The idea is to use convex perturbation for the optimal control, jointly with some estimations of the state trajectory and performance functional, and by sending the perturbations to zero, one obtains some inequality, then by completing with martingale representation theorem's the maximum principle is expressed in terms of an adjoint process.

We can state the stochastic maximum principle in a stronger form.

Theorem 2.3.1

Let $\widehat{u}$ be an optimal control minimizing the performance functional J over $\mathcal{U}$, and let $\widehat{y}$ be the corresponding optimal trajectory, then there exists an adapted processes

$(p,q) \in \mathbb{L}^2\left(\left([0,T];\mathbb{R}^n\right)\right) \times \mathbb{L}^2\left(\left([0,T];\mathbb{R}^{n\times d}\right)\right)$ *which is the unique solution of the BSDE (2.23), such that for all* $v \in U$

$$H_u\left(t,\widehat{y}(t),\widehat{u}(t),p(t),q(t)\right)\left(v_t-\widehat{u}(t)\right) \leq 0, \quad \mathbb{P}-a.s.$$

In order to give the proof of (theorem 2.2.1), it is convenient to present the following.

2.3.4. *Variational equation*

Let $v \in \mathcal{U}$ be such that $(\widehat{u}+v) \in \mathcal{U}$, the convexity condition of the control domain ensure that, for $\theta \in (0,1)$ the control $(\widehat{u}+\theta v)$ is also in $\mathcal{U}$. We denote by y^θ the solution of the SDE (2.19) correspond to the control $(\widehat{u}+\theta v)$, then by standard arguments from stochastic calculus, it is easy to check the following convergence result.

Lemma 2.3.1

Under assumption $(H1)$ *we have*

$$\lim_{\theta\to 0}\mathbb{E}\left[\sup_{t\in[0,T]}\left|y^\theta(t)-\widehat{y}(t)\right|^2\right]=0. \tag{2.24}$$

Proof: From assumption $(H1)$, we get by using the Burkholder-Davis-Gundy inequality

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|y^\theta(t)-\widehat{y}(t)\right|^2\right] \leq K\int_0^t \mathbb{E}\left[\sup_{\tau\in[0,s]}\left|y^\theta(r)-\widehat{y}(r)\right|^2\right]ds + K\theta^2\left(\int_0^t \mathbb{E}\left[\sup_{r\in[0,s]}|v(r)|^2\right]ds\right). \tag{2.25}$$

From (definition 2.2.1), and Gronwall's lemma, the result follows immediately by letting θ go to zero. ■

We define the process $z(t) = z^{\widehat{u},v}(t)$ by

$$\begin{cases} dz(t) &= \left\{b_y\left(t,\widehat{z}(t),\widehat{u}(t)\right)z(t)+b_u\left(t,\widehat{y}(t),\widehat{u}(t)\right)v(t)\right\}dt \\ & \quad + \sum_{j=1}^d \left\{\sigma_y^j\left(t,\widehat{y}(t),\widehat{u}(t)\right)z(t)+\sigma_u^j\left(t,\widehat{y}(t),\widehat{u}(t)\right)v(t)\right\}dB^j(t), \\ z(0) &= 0. \end{cases} \tag{2.26}$$

From $(H2)$ and (definition 2.2.1), one can find a unique solution z which solves the variational equation (2.26), and the following estimation holds.

Lemma 2.3.2

Under assumption $(H1)$, it holds that

$$\lim_{\theta\to 0}\mathbb{E}\left|\frac{y^{\theta}(t)-\widehat{y}(t)}{\theta}-z(t)\right|^{2}=0. \tag{2.27}$$

Proof: Let

$$\Gamma^{\theta}(t)=\frac{y^{\theta}(t)-\widehat{y}(t)}{\theta}-z(t).$$

Denoting $y^{\mu,\theta}(t)=\widehat{y}(t)+\mu\theta\left(\Gamma^{\theta}(t)+z(t)\right)$, and $u^{\mu,\theta}(t)=\widehat{u}(t)+\mu\theta v(t)$, for notational convenience. Then we have immediately that $\Gamma^{\theta}(0)=0$ and $\Gamma^{\theta}(t)$ fulfills the following SDE

$$\begin{aligned}
d\Gamma^{\theta}(t)=&\left\{\frac{1}{\theta}\left(b\left(t,y^{\mu,\theta}(t),u^{\mu,\theta}(t)\right)-b\left(t,\widehat{y}(t),\widehat{u}(t)\right)\right)\right.\\
&\left.-\left(b_{y}\left(t,\widehat{y}(t),\widehat{u}(t)\right)z(t)+b_{u}\left(t,\widehat{y}(t),\widehat{u}(t)\right)v(t)\right)\right\}dt\\
&+\left\{\frac{1}{\theta}\left(\sigma\left(t,y^{\mu,\theta}(t),u^{\mu,\theta}(t)\right)-\sigma\left(t,\widehat{y}(t),\widehat{u}(t)\right)\right)\right.\\
&\left.-\left(\sigma_{y}\left(t,\widehat{y}(t),\widehat{u}(t)\right)z(t)+\sigma_{u}\left(t,\widehat{y}(t),\widehat{u}(t)\right)v(t)\right)\right\}dB(t)
\end{aligned}$$

Since the derivatives of the coefficients are bounded, and from (definition 2.2.1), it is easy to verify by Gronwall's inequality that

$$\begin{aligned}
\mathbb{E}\left|\Gamma^{\theta}(t)\right|^{2}\leq&\ K\mathbb{E}\int_{0}^{t}\left|\int_{0}^{1}b_{y}\left(s,y^{\mu,\theta}(s),u^{\mu,\theta}(s)\right)\Gamma^{\theta}(s)\,d\mu\right|^{2}ds+K\mathbb{E}\left|\rho^{\theta}(t)\right|^{2}\\
&+K\mathbb{E}\int_{0}^{t}\left|\int_{0}^{1}\sigma_{y}\left(s,y^{\mu,\theta}(s),u^{\mu,\theta}(s)\right)\Gamma^{\theta}(s)\,d\mu\right|^{2}ds,
\end{aligned}$$

where $\rho^{\theta}(t)$ is given by

$$\begin{aligned}
\rho^{\theta}(t) = &-\int_0^t b_y(s,\widehat{y}(s),\widehat{u}(s))\, z(s)\, ds \\
&-\int_0^t \sigma_y(s,\widehat{y}(s),\widehat{u}(s))\, z(s)\, dB(s) \\
&-\int_0^t b_v(s,\widehat{y}(s),\widehat{u}(s))\, v(s)\, ds \\
&-\int_0^t \sigma_v(s,\widehat{y}(s),\widehat{u}(s))\, v(s)\, dB(s) \\
&+\int_0^t\int_0^1 b_y\left(s,y^{\mu,\theta}(s),u^{\mu,\theta}(s)\right) z(s)\, d\mu ds \\
&+\int_0^t\int_0^1 b_v\left(s,y^{\mu,\theta}(s),u^{\mu,\theta}(s)\right) v(s)\, d\mu ds \\
&+\int_0^t\int_0^1 \sigma_y\left(s,y^{\mu,\theta}(s),u^{\mu,\theta}(s)\right) z(s)\, d\mu dB(s) \\
&+\int_0^t\int_0^1 \sigma_v\left(s,y^{\mu,\theta}(s),u^{\mu,\theta}(s)\right) v(s)\, d\mu dB(s).
\end{aligned}$$

Since b_y, σ_y are bounded, then

$$\mathbb{E}\left|\Gamma^{\theta}(t)\right|^2 \leq M\mathbb{E}\int_0^t \left|\Gamma^{\theta}(s)\right|^2 ds + M\mathbb{E}\left|\rho^{\theta}(t)\right|^2,$$

where M is a generic constant depending on the constant K and T. We conclude from lemma 1.4.2 that $\lim_{\theta\to 0}\rho^{\theta}(t) = 0$. Hence (2.27) follows from Gronwall lemma and by letting θ go to 0. ■

Let Φ be the fundamental solution of the linear matrix equation, for $0 \leq s < t \leq T$

$$\begin{cases}
d\Phi_{s,t} &= b_y(t,\widehat{y}(t),\widehat{u}(t))\,\Phi_{s,t}dt + \sum_{j=1}^{d}\sigma_y^j(t,\widehat{y}(t),\widehat{u}(t))\,\Phi_{s,t}dB^j(t), \\
\Phi_{s,s} &= I_d,
\end{cases}$$

where I_d is the $n\times n$ identity matrix, this equation is linear with bounded coefficients, then it admits a unique strong solution.

From Itô's formula we can easily check that $d(\Phi_{s,t}\Psi_{s,t}) = 0$, and $\Phi_{s,s}\Psi_{s,s} = I_d$, where Ψ is the solution of the following equation

$$\begin{cases}
d\Psi_{s,t} &= -\Psi_{s,t}\left\{b_y(t,\widehat{y}(t),\widehat{u}(t)) - \sum_{j=1}^{d}\sigma_y^j(t,\widehat{y}(t),\widehat{u}(t))\,\sigma_y^j(t,\widehat{y}(t),\widehat{u}(t))\right\}dt \\
&\quad -\sum_{j=1}^{d}\Psi_{s,t}\sigma_y^j(t,\widehat{y}(t),\widehat{u}(t))\,dB^j(t), \\
\Psi_{s,s} &= I_d,
\end{cases}$$

so $\Psi=\Phi^{-1}$, if $s=0$ we simply write $\Phi_{0,t}=\Phi_t$, and $\Psi_{0,t}=\Psi_t$. By integrating by part formula we can see that, the solution of (2.26) is given by $z(t)=\Phi_t\eta_t$, where η_t is the solution of the stochastic differential equation

$$\begin{cases} d\eta_t &= \Psi_t\left\{b_u\left(t,\widehat{y}\left(t\right),\widehat{u}\left(t\right)\right)v\left(t\right)-\sum\limits_{j=1}^{d}\sigma_y^j\left(t,\widehat{y}\left(t\right),\widehat{u}\left(t\right)\right)\sigma_u^j\left(t,\widehat{y}\left(t\right),\widehat{u}\left(t\right)\right)v\left(t\right)\right\}dt \\ & \quad +\sum\limits_{j=1}^{d}\Psi_t\sigma_u^j\left(t,x_t^{\star},u_t^{\star}\right)v\left(t\right)dB^j\left(t\right), \\ \eta_0 &= 0. \end{cases}$$

Let us introduce the following convex perturbation of the optimal control $\widehat{u}$ by

$$u^{\theta}=\widehat{u}+\theta v, \tag{2.28}$$

for any $v\in\mathcal{U}$, and $\theta\in(0,1)$. Since $\widehat{u}$ is an optimal control, then $\theta^{-1}\left(J\left(u^{\theta}\right)-J\left(\widehat{u}\right)\right)\geq 0$. Thus a necessary condition for optimality is that

$$\lim_{\theta\to 0}\theta^{-1}\left(J\left(u^{\theta}\right)-J\left(\widehat{u}\right)\right)\geq 0. \tag{2.29}$$

The rest is devoted to the computation of the above limit. We shall see that the expression (2.29) leads to a precise description of the optimal control $\widehat{u}$ in terms of the adjoint process. First, it is easy to prove the following lemma

Lemma 2.3.3

Under assumptions $(H1)$, we have

$$\begin{aligned} I &= \lim_{\theta\to 0}\theta^{-1}\left(J\left(u^{\theta}\right)-J\left(\widehat{u}\right)\right) \\ &= \mathbb{E}\left[\int_0^T\left\{f_y\left(s,\widehat{y}\left(s\right),\widehat{u}\left(s\right)\right)z\left(s\right)+f_u\left(s,\widehat{y}\left(s\right),\widehat{u}\left(s\right)\right)v\left(s\right)\right\}ds+g_y\left(\widehat{y}\left(T\right)\right)z\left(T\right)\right]. \end{aligned} \tag{2.30}$$

Proof: We use the same notations as in the proof of (lemma 2.2.2). First, we have

$$\begin{aligned} &\theta^{-1}\left(J\left(u^{\theta}\right)-J\left(\widehat{u}\right)\right) \\ &= \mathbb{E}\left[\int_0^T\int_0^1\left\{f_y\left(s,y^{\mu,\theta}\left(s\right),u^{\mu,\theta}\left(s\right)\right)z\left(s\right)+f_u\left(s,y^{\mu,\theta}\left(s\right),u^{\mu,\theta}\left(s\right)\right)v\left(s\right)\right\}d\mu ds \right.\\ &\qquad \left.+\int_0^1 g_y\left(y^{\mu,\theta}\left(T\right)\right)z\left(T\right)d\mu\right]+\beta^{\theta}\left(t\right), \end{aligned}$$

where

$$\beta^{\theta}(t)=\mathbb{E}\left[\int_0^T\int_0^1 f_y\left(s,y^{\mu,\theta}(s),u^{\mu,\theta}(s)\right)\Gamma^{\theta}(s)\,d\mu ds+\int_0^1 g_y\left(y^{\mu,\theta}(T)\right)\Gamma^{\theta}(T)\,d\mu\right].$$

By using the (lemma 1.4.2), and since the derivatives f_y, f_u, and g_y are bounded, we have $\lim\limits_{\theta\to 0}\beta^{\theta}(t)=0$. Then, the result follows by letting θ go to 0 in the above equality. ■

Substituting by $z(t)=\Phi_t\eta_t$ in (2.30), this leads to

$$I=\mathbb{E}\left[\int_0^T\left\{f_y(s,\widehat{y}(s),\widehat{u}(s))\,\Phi_s\eta_s+f_u(s,\widehat{y}(s),\widehat{u}(s))\,v(s)\right\}ds+g_y(\widehat{y}(T))\,\Phi_T\eta_T\right].$$

Consider the right continuous version of the square integrable martingale

$$M(t):=\mathbb{E}\left[\int_0^T f_y(s,\widehat{y}(s),\widehat{u}(s))\,\Phi_s ds+g_y(\widehat{y}(T))\,\Phi_T\,|\mathcal{F}_t\right].$$

By the representation theorem, there exist $Q=\left(Q^1,..,Q^d\right)$ where $Q^j\in\mathbb{L}^2$, for $j=1,...,d$,

$$M(t)=\mathbb{E}\left[\int_0^T f_y(s,\widehat{y}(s),\widehat{u}(s))\,\Phi_s ds+g_y(\widehat{y}(T))\,\Phi_T\right]+\sum_{j=1}^d\int_0^t Q^j(s)\,dB^j(s).$$

We introduce some more notation, write $\widehat{y}(t)=M(t)-\int_0^t f_y(s,\widehat{y}(s),\widehat{u}(s))\,\Phi_s ds$. The adjoint variable is the processes defined by

$$\begin{cases} p(t) &=\widehat{y}(t)\,\Psi_t,\\ q^j(t) &=Q^j(t)\,\Psi_t-p(t)\,\sigma_y^j(t,\widehat{y}(t),\widehat{u}(t)),\ \text{for } j=1,...,d.\end{cases}\tag{2.31}$$

Theorem 2.3.2

Under assumptions $(H1)$, *we have*

$$I=\mathbb{E}\left[\int_0^T\left\{f_u(s,\widehat{y}(s),\widehat{u}(s))+p(s)\,b_u(s,\widehat{y}(s),\widehat{u}(s))+\sum_{j=1}^d q^j\sigma_u^j(s,\widehat{y}(s),\widehat{u}(s))\right\}\right].$$

Proof: From the integration by part formula, and by using the definition of $p(t)$, $q^j(t)$ for $j=1,..,d$, we easily check that

$$\begin{aligned} E[y(T)\,\eta(T)] &=\mathbb{E}\left[\int_0^T\left\{p(t)\,b_u(s,\widehat{y}(s),\widehat{u}(s))+\sum_{j=1}^d q^j(s)\,\sigma_u^j(s,\widehat{y}(s),\widehat{u}(s))\right\}v(t)\,dt\right.\\ &\quad\left.-\int_0^T f_y(s,\widehat{y}(s),\widehat{u}(s))\,\eta_t\Phi_t dt.\right.\end{aligned}\tag{2.32}$$

Also we have

$$I=\mathbb{E}\left[y\left(T\right)\eta\left(T\right)+\int_0^T f_y\left(s,\widehat{y}\left(s\right),\widehat{u}\left(s\right)\right)\Phi_t\eta_t dt+\int_0^T f_u\left(s,\widehat{y}\left(s\right),\widehat{u}\left(s\right)\right)v\left(t\right)dt\right], \tag{2.33}$$

substituting (2.32) in (2.33), This completes the proof. ■

2.3.5. Sufficient conditions of optimality

Theorem 2.3.3

Let $\widehat{u}$ be an admissible control, we denote $\widehat{y}$ the associated controlled state process, and let (p,q) be a solution to the corresponding BSDE (2.23). Let us assume that $H\left(t,y,u,p\left(t\right),q\left(t\right)\right)$, and (y) are concave functions. Moreover suppose that for all $t\in[0,T]$,

$$H\left(t,\widehat{y}\left(t\right),\widehat{u}\left(t\right),p\left(t\right),q\left(t\right)\right)=\inf_{u\in U}H\left(t,\widehat{y}\left(t\right),u\left(t\right),p\left(t\right),q\left(t\right)\right). \tag{2.34}$$

Then $\widehat{u}$ is an optimal control.

Proof: We consider the difference

$$J\left(\widehat{u}\right)-J\left(u\right)=\mathbb{E}\left[\int_0^T\left(f\left(t,\widehat{y}\left(t\right),\widehat{u}\left(t\right)\right)-f\left(t,y\left(t\right),u\left(t\right)\right)\right)dt\right]$$
$$+\mathbb{E}\left[g\left(\widehat{y}\left(T\right)\right)-g\left(y\left(T\right)\right)\right].$$

Since g is concave, we get

$$\begin{aligned}\mathbb{E}\left[g\left(\widehat{y}\left(T\right)\right)-g\left(y\left(T\right)\right)\right]&\geq\mathbb{E}\left[\left(\widehat{y}\left(T\right)-y\left(T\right)\right)g_y\left(\widehat{y}\left(T\right)\right)\right]\\&=\mathbb{E}\left[\left(\widehat{y}\left(T\right)-y\left(T\right)\right)p\left(T\right)\right]\\&=\mathbb{E}\left[\int_0^T\left(\widehat{y}\left(t\right)-y\left(t\right)\right)dp\left(t\right)+\int_0^T p\left(t\right)d\left(\widehat{y}\left(t\right)-y\left(t\right)\right)\right]\\&\quad+\mathbb{E}\left[\int_0^T\sum_{j=1}^n\left(\sigma^j\left(t,\widehat{y}\left(t\right),\widehat{u}\left(t\right)\right)-\sigma^j\left(t,y\left(t\right),u\left(t\right)\right)\right)q^j\left(t\right)dt\right],\end{aligned}$$

with

$$\mathbb{E}\left[\int_0^T\left(\widehat{y}\left(t\right)-y\left(t\right)\right)dp\left(t\right)\right]=\mathbb{E}\left[\int_0^T\left(\widehat{y}\left(t\right)-y\left(t\right)\right)\left(-H_y\left(t,\widehat{y}\left(t\right),\widehat{u}\left(t\right),p\left(t\right),q\left(t\right)\right)\right)dt\right]$$
$$+\mathbb{E}\left[\int_0^T\left(\widehat{y}\left(t\right)-y\left(t\right)\right)q\left(t\right)dB\left(t\right)\right],$$

and

$$\mathbb{E}\left[\int_0^T p(t)\, d(\widehat{y}(t) - y(t))\right] = \mathbb{E}\left[\int_0^T p(t)\left(b(t, \widehat{y}(t), \widehat{u}(t)) - b(t, y(t), u(t))\right) dt\right]$$
$$+ \mathbb{E}\left[\int_0^T p(t)\left(\sigma(t, \widehat{y}(t), \widehat{u}(t)) - \sigma(t, y(t), u(t))\right) dB(t)\right].$$

On the other hand, the process

$$\mathbb{E}\left[\int_0^T \left\{p(t)\left(\sigma(t, \widehat{y}(t), \widehat{u}(t)) - \sigma(t, y(t), u(t))\right) + (\widehat{y}(t) - y(t))\, q(t)\right\} dB(t)\right]$$

is a continuous local martingale for all $0 < t \leq T$, by the fact that $(p, q) \in \mathbb{L}^2\left(([0,T]; \mathbb{R}^n)\right) \times \mathbb{L}^2\left(\left([0,T]; \mathbb{R}^{n\times d}\right)\right)$, we deduce that the stochastic integrals with respect to the local martingales have zero expectation. By the concavity of the Hamiltonian H, we get

$$\mathbb{E}\left[g(\widehat{y}(T)) - g(y(T))\right] \geq -\mathbb{E}\left[\int_0^T \left(H(t, \widehat{y}(t), \widehat{u}(t), p(t), q(t)) - H(t, y(t), u(t), p(t), q(t))\right) dt\right]$$
$$+ \mathbb{E}\left[\int_0^T p(t)\left(b(t, \widehat{y}(t), \widehat{u}(t)) - b(t, y(t), u(t))\right) dt\right]$$
$$+ \mathbb{E}\left[\int_0^T \left(\sigma(t, \widehat{y}(t), \widehat{u}(t)) - \sigma(t, y(t), u(t))\right) q(t)\, dt\right].$$

By the definition of the Hamiltonian H, we obtain

$$J(\widehat{u}) - J(u) \geq 0,$$

then $\widehat{u}$ is an optimal control. ■

2.4 A General stochastic maximum principle for optimal control problems

In this part, we will give a detailed proof of maximum principle in optimal control in general case, where the control domain need not be convex, and the diffusion coefficient depends explicitly , This result is the generalization of principle of the maximum was obtained by Peng [41], he introduce a second-order expansion method and strong perturbations .

2.4.1 Problem formulation and assumptions

Let $(\Omega, \mathcal{F}, P)$ be a probability space with filtration $\mathcal{F}_t$,. Let $W(\cdot)$ be an $\mathbb{R}^n$-valued standard Brownien process. We denote the set of all admissible controls by U_{ad}.

We assume that $(\mathcal{F}_t) = \sigma\left(W\left(s\right), 0 \leq s \leq t\right)$, and we consider the following control problem

$$\begin{cases} dx_t = b(x_t, v_t)dt + \sigma(x_t, v_t)dW_t \\ x(0) = x_0, \end{cases} \tag{2.35}$$

where $b\left(x, v\right) : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n, \sigma\left(x, v\right) : \mathbb{R}^n \times \mathbb{R}^k \to L\left(\mathbb{R}^d \times \mathbb{R}^n\right), \sigma = \sigma\left(\sigma^1, \sigma^2, ..., \sigma^d\right).$

Definition 2.4.1 (*Admissible control*)

An admissible control $v(\cdot)$ is an $\mathcal{F}_t$,-adapted process with values in U such that

$$\sup_{0 \leq t \leq T} \mathbb{E}\left|v(t)\right|^m < \infty, \forall m = 1, 2, \ldots$$

where U is a nonempty subset of $\mathbb{R}^k$.

The stochastic optimal control problem is to minimize the following cost functional

$$J(v(\cdot)) = \mathbb{E}(\int_0^T l(x_t, v_t)dt + h(x_T)), \tag{2.36}$$

$$\inf\left\{J(v(\cdot)) : v(\cdot) \in U_{ad}\right\}.$$

Here, $l\left(x, v\right) : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}, \quad h(x) : \mathbb{R}^n \to \mathbb{R}.$

We also assume that:

1. b, σ, l, h are twice continuously differentiable with respect to x. (2.37)
2. All their derivatives $b_x, b_{xx}, \sigma_x, \sigma_{xx}, l_x, l_{xx}, h_x, h_{xx}$ arecontinuous in (x, v).
3. $b_x, b_{xx}, \sigma_x, \sigma_{xx}, , l_{xx}, h_{xx}$ are bounded,and b, σ, l_x, h_x are bounded by $C\left(1 + |x| + |v|\right)$.

2.4.2 Second order expansion

In this part, we derive a kind of variational equation and variational inequality. The control domain U is not necessarily convex, the usual first-order expansion approach does not work. Hence, introducing a second order expansion method.

Let $(y(\cdot), u(\cdot))$ be an optimal solution of the our stochastic control problem. We introduce the following spike variation

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

where $0 \leq \tau < T$ is fixed, $\theta > 0$ is sufficiently small, and v is an arbitrary $\mathcal{F}^{\tau}$-measurable random variable with values in U, such that $\sup_{\omega \in \Omega} |v(\omega)| < \infty$.

Let $y^{\theta}(t)$ be the trajectory of the control system (2.35) corresponding to the control $u^{\theta}(\cdot)$.

Now, we derive the variational inequality from the fact that

$$J(u^{\theta}(\cdot)) - J(u(\cdot)) \geq 0.$$

Lemma 2.4.1

Under assumption (2.37). Then

$$\sup_{0 \leq t \leq T} \theta^{-2} \left| y^{\theta}(t) - y(t) - y_1(t) - y_2(t) \right|^2 \leq C, \tag{2.38}$$

where $y_1(\cdot), y_2(\cdot)$ are solutions of

$$\begin{aligned} y_1(t) &= \int_0^t \left[b_x(y_s, u_s) y_1(s) + \left(b\left(y_s.u_s^{\theta}\right) - b(y_s.u_s) \right) \right] ds \\ &+ \int_0^t \left[\sigma_x(y_s, u_s) y_1(s) + \left(\sigma\left(y_s.u_s^{\theta}\right) - \sigma(y_s.u_s) \right) \right] dW_s, \end{aligned} \tag{2.39}$$

and

$$\begin{aligned} y_2(t) &= \int_0^t \left[b_x(y_s, u_s) y_2(s) + \frac{1}{2} b_{xx}(y_s.u_s) y_1(s) y_1(s) \right] ds \\ &+ \int_0^t \left[\sigma_x(y_s.u_s) y_2(s) - \frac{1}{2} \sigma_{xx}(y_s, u_s) \right] y_1(s) y_1(s) dW_s \\ &+ \int_0^t \left[b_x\left(y_s, u_s^{\theta}\right) + b_x(y_s, u_s) \right] y_1(s) ds \\ &+ \int_0^t \left[\sigma_x\left(y_s, u_s^{\theta}\right) + \frac{1}{2} \sigma_x(y_s, u_s) y_1(s) \right] ds, \end{aligned} \tag{2.40}$$

where

$$f_{xx} y y = \sum_{i,j=1}^{n} f_{x^i x^j} y^i y^j \quad \text{for } f = b, \sigma, l, h.$$

Remark 2.4.1

Equation (2.39) is called the first-order variational equation. We must introduce what *we call "the second order variational equation" (2.40), because with the solution of 2.39, we can only obtain the following estimation:*

$$\theta^{-1}\sup_{0\le t\le T}\left|y^{\theta}(t)-y(t)-y_1(t)\right|^2\le C.$$

It is not enough to derive the variational inequality.

Proof : By Gronwall's inequality and the moment inequality (see Ikeda and Watanabe [29]), it is easy to verify that

$$\sup_{0\le t\le T}\mathbb{E}\left(|y_1(t)|^2\right)\le C\theta, \tag{2.41}$$

$$\sup_{0\le t\le T}\mathbb{E}\left(|y_2(t)|^2\right)\le C\theta^2, \tag{2.42}$$

$$\begin{cases}\sup_{0\le t\le T}\mathbb{E}\left(|y_1(t)|^4\right)\le C\theta^2,\\ \sup_{0\le t\le T}\mathbb{E}\left(|y_2(t)|^4\right)\le C\theta^4,\\ \sup_{0\le t\le T}\mathbb{E}\left(|y_1(t)|^8\right)\le C\theta^4.\end{cases} \tag{2.43}$$

Set $y_3=y_1+y_2$.

We have

$$\begin{aligned}
&\int_0^t b\left(y+y_3,u^{\theta}\right)ds+\int_0^t\sigma\left(y+y_3,u^{\theta}\right)dW_s\\
&=\int_0^t\left[b\left(y,u^{\theta}\right)+b_x\left(y,u^{\theta}\right)y_3+\int_0^1\int_0^1\lambda b_{xx}\left(y+\lambda u y_3,u^{\theta}\right)d\lambda du y_3y_3\right]ds\\
&+\int_0^t\left[\sigma\left(y,u^{\theta}\right)+\sigma_x\left(y,u^{\theta}\right)y_3+\int_0^1\int_0^1\lambda\sigma_{xx}\left(y+\lambda u y_3,u^{\theta}\right)d\lambda du y_3y_3\right]dW_s\\
&=\int_0^t b(y,u)\,ds+\int_0^t\sigma(y,u)\,dW_s+\int_0^t b_x(y,u)\,y_3ds+\int_0^t\sigma_x(y,u)\,y_3dW_s\\
&+\int_0^t\left(b\left(y(s),u^{\theta}(s)\right)-b\left(y(s),u(s)\right)\right)ds\\
&+\int_0^t\left(\sigma\left(y(s),u^{\theta}(s)\right)-\sigma\left(y(s),u(s)\right)\right)dW_s\\
&+\int_0^t\frac{1}{2}b_{xx}(y,u)\,y_3(s)y_3(s)ds+\int_0^t\frac{1}{2}\sigma_{xx}(y,u)\,y_3(s)y_3(s)dW_s\\
&+\int_0^t\left(b_x\left(y,u^{\theta}\right)-b_x(y,u)\right)y_3(s)ds\\
&+\int_0^t\sigma_x\left(y,u^{\theta}\right)-\sigma_x(y,u)\,y_3(s)dW_s\\
&+\int_0^t\int_0^1\int_0^1\lambda\left[b_{xx}\left(y+\lambda\mu y_3,u^{\theta}\right)-b_{xx}(y,u)\right]d\lambda d\mu y_3y_3ds\\
&+\int_0^t\int_0^1\int_0^1\lambda\left[\sigma_{xx}\left(y+\lambda\mu y_3,u^{\theta}\right)-\sigma_{xx}(y,u)\right]d\lambda d\mu y_3y_3dWs\\
&=y(t)+y_3(t)-x_0+\int_0^t b^{\theta}(s)\,ds+\int_0^t\wedge^{\theta}(s)\,dW_s.
\end{aligned}$$

Using (2.39) and (2.40), we have

$$\begin{aligned} G^{\theta}(s) &= \frac{1}{2} b_{xx}(y_s, u_s)(y_2(s) y_2(s) + 2y_1(s) y_2(s)) + \left[b_x\left(y_s, u_s^{\theta}\right) - b_x(y_s, u_s)\right] y_2(s) \\ &+ \int_0^1 \int_0^1 \lambda \left[b_{xx}\left(y + \lambda u y_3, u^{\theta}\right) - b_{xx}(y, v)\right] d\lambda du y_3(s) y_3(s), \end{aligned}$$

and

$$\begin{aligned} \wedge^{\theta}(s) &= \frac{1}{2} \sigma_{xx}(y_s, u_s)(y_2(s) y_2(s) + 2y_1(s) y_2(s)) + \left[\sigma_x\left(y_s, u_s^{\theta}\right) - \sigma_x(y_s, u_s)\right] y_2(s) \\ &+ \int_0^1 \int_0^1 \lambda \left[\sigma_{xx}\left(y + \lambda u y_3, u^{\theta}\right) - \sigma_{xx}(y, v)\right] d\lambda du y_3(s) y_3(s). \end{aligned}$$

From (2.41), (2.42), and (2.43) we can see that

$$\sup_{0 \le t \le T} \mathbb{E}\left(\left|G^{\theta}(s)\, ds\right|^2 + \left|\int_0^1 \wedge^{\theta}(s)\, dW_s\right|^2\right) = o(\theta^2). \tag{2.44}$$

Thus we have

$$\begin{aligned} y(t) + y_3(t) &= x_0 + \int_0^t b\left(y(s) + y_3(s), u^{\theta}(s)\right) ds \\ &+ \int_0^t \sigma\left(y(s) + y_3(s), u^{\theta}(s)\right) dW_s - \int_0^1 G^{\theta}(s)\, ds - \int_0^1 \wedge^{\theta}(s)\, dW_s. \end{aligned}$$

Since

$$y^{\theta}(t) = x_0 + \int_0^t b\left(y^{\theta}(s), u^{\theta}(s)\right) ds \int_0^t \sigma\left(y^{\theta}(s), u^{\theta}(s)\right) dB_s,$$

we can derive

$$\begin{aligned} (y^{\theta} - y - y_3)(t) &= \int_0^t A^{\theta}(s)\left(y^{\theta} - y_3 - y\right)(s) ds \int_0^t D^{\theta}(s)\left(y_3^{\theta} - y - y\right) dW_s \\ &+ \int_0^t b^{\theta}(s)\, ds + \int_0^t \wedge^{\theta}(s)\, dW_s, \end{aligned}$$

with

$$\left|A^{\theta}(s,\omega)\right| + \left|D^{\theta}(s,\omega)\right| \le C \quad \forall s, \forall \omega.$$

By using Itô's formula and Gronwall's inequality, we obtain the estimation (2.38). ■

Lemma 2.4.2

Under the assumption of (Lemma 2.4.1), we have

$$\begin{aligned} &\mathbb{E}\int_0^T \left[l_x(y(s), u(s))(y_1(s) + y_2(s)) + \frac{1}{2} l_{xx}(y(s), u(s)) y_1(s) y_1(s)\right] ds \\ &+ \mathbb{E}\int_0^T \left(l\left(y(s), u^{\theta}(s)\right) - l(y(s), u(s))\right) ds \\ &+ \mathbb{E}(h_x(y(T)))(y(_1(T) + y_2(T))) + \frac{1}{2}\mathbb{E}(h_{xx}(y(T)) y_1(T) y_1(T)) \ge o(\theta). \end{aligned} \tag{2.45}$$

Remark 2.4.2

In the case where σ does not contain the control variable v, the relation (2.45) can be reduced to

$$\begin{aligned}&\mathbb{E}\int_0^T l_x\left(y\left(s\right),u\left(s\right)\right)y_1\left(s\right)ds+\mathbb{E}h_x\left(y\left(T\right)\right)y_1\left(T\right)\\&+\mathbb{E}\int_0^T\left(l\left(y\left(s\right),u^\theta\left(s\right)\right)-l\left(y\left(s\right),u\left(s\right)\right)\right)ds\geq o(\theta).\end{aligned}$$

Thus we need only the first-order variational equation (2.39).

Proof: Since $(y(\cdot),u(\cdot))$ is optimal, we have

$$\mathbb{E}\int_0^T\left[l\left(y^\theta\left(t\right),u^\theta\left(t\right)\right)-l\left(y\left(t\right),u\left(t\right)\right)\right]dt+\mathbb{E}\left(h\left(y^\theta\left(T\right)\right)-h\left(y\left(T\right)\right)\right)\geq 0.$$

Thus from (Lemma 2.4.1)

$$\begin{aligned}0\ &\leq \mathbb{E}\int_0^T\left[l\left(y+y_1+y_2,u^\theta\left(t\right)\right)-l\left(y\left(t\right),u\left(t\right)\right)dt\right]\\&+\mathbb{E}\left(h\left(y+y_1+y_2\right)\left(T\right)-h\left(y\left(T\right)\right)\right)+o(\theta)\\&=\mathbb{E}\int_0^T\left[l\left(y+y_1+y_2,u\right)-l\left(y,u\right)\right]dt\\&+\mathbb{E}\left(h\left(y+y_1+y_2\right)\left(T\right)-h\left(y\left(T\right)\right)\right)\\&+\mathbb{E}\int_0^T\left[l\left(y+y_1+y_2,u^\theta\right)-l\left(y+y_1+y_2,u\right)\right]dt+o(\theta)\end{aligned}$$

$$\begin{aligned}&=\mathbb{E}\int_0^T\left[l_x\left(y,u\right)\left(y_1+y_2\right)+\frac{1}{2}l_{xx}\left(y,u\right)\left(y_1+y_2\right)\left(y_1+y_2\right)\right]ds\\&+\mathbb{E}\int_0^T\left[l\left(y,u^\theta\right)-l\left(y,u\right)\right]ds+\mathbb{E}\int_0^T\left[l_x\left(y,u^\theta\right)-l_x\left(y,u\right)\right]\left(y_1+y_2\right)ds\\&+\frac{1}{2}\mathbb{E}\left(\int_0^T\left[l_{xx}\left(y,u^\theta\right)-l_{xx}\left(y,u\right)\right]y_1\left(s\right)y_1\left(s\right)ds\right)\\&+\mathbb{E}\left(h_xy\left(T\right)\right)\left(y_1\left(T\right)+y_2\left(T\right)\right)+\frac{1}{2}\mathbb{E}h_{xx}\left(y\left(T\right)\right)y_1\left(T\right)y_1\left(T\right)+o(\theta).\end{aligned}$$

Then, (2.45) follows from (2.41) and (2.42). ■

2.4.3 Adjoint processes and variational inequality

For simplicity, we let

$$g_x(t)=g_x(y(t),u(t)),\text{for } g=b,\sigma,l,h,$$

$$g_{xx}(t)=g_{xx}(y(t),u(t)),\ \text{ for } g=b,\sigma,l,h.$$

We consider the linear stochastic system

$$\begin{cases} dz(t) = (b_x(t)z(t) + \phi(t))\, ds + (\sigma_x(t)z(t) + \Psi(t))\, dW_s \\ z(0) = 0, \end{cases} \tag{2.46}$$

$$(\phi(t), \psi(t)) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^n) \times L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)^d, \quad \Psi = (\psi_1, ..., \psi_d),$$

where $L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)$ is the space of all $\mathbb{R}^n$-valued adapted processes such that

$$\mathbb{E}\int_0^T |\phi(t)|^2\, dt \langle \infty$$

We can construct a linear functional on the Hilbert space $L^2_{\mathcal{F}}(0, T; \mathbb{R}^n) \times L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)^d$ as follows:

$$I(\phi(t), \psi(t)) = \mathbb{E}\int_0^T l_x(t)z(t)dt + \mathbb{E}(h_x(T)z(T)),$$

where $(\phi(t), \psi(t))$ and $z(t)$ are related by (2.46). It is easy to verify that $I(\cdot, \cdot)$ is continuous. Then by the Riesz Representation Theorem, there is a unique

$$(p(\cdot), K(\cdot)) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^n) \times L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)^d$$
$$K = (K_1, ..., K_d),$$

such that

$$\begin{aligned} &\mathbb{E}\int_0^T \left[(p(\cdot), \phi(t)) + \sum_{j=1}^d (K_j(t), \psi_j(t))\right] dt = I(\phi(.), \psi(.)) \\ &\forall (\phi(\cdot), \psi(\cdot)) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^n) \times L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)^d. \end{aligned} \tag{2.47}$$

With (2.39) and (2.40), we can apply this result to some of the terms of (2.45):

$$\begin{aligned} &\mathbb{E}\int_0^T l_x(s)y_1(s)ds + \mathbb{E}(h_x y(T))y_1(T) \\ &= \mathbb{E}\int_0^T (p(s), b(y(s), u^\theta(s)) - b(y(s), u(s)))ds \\ &+ \mathbb{E}\int_0^T tr\left[K(s)(\sigma(y(s), u^\theta(s)) - \sigma(y(s), u(s))\right]. \\ &\mathbb{E}\int_0^T l_x(s)y_2(s)ds + \mathbb{E}(h_x y(T))y_2(T) \\ &= \mathbb{E}\int_0^T \frac{1}{2}\left[\left(p(s)b_{xx}(s) + \sum_{j=1}^d K_j(s)\sigma^j_{xx}(s)\right) y_1(s)y_1(s)\right] ds \\ &+ \mathbb{E}\int_0^T p^*(s)(b_x(y(s), u^\theta(s)) - b_x(y(s), u(s)))y_1(s)ds \\ &+ \mathbb{E}\int_0^T \sum_{j=1}^d K^*_j(s)\left(\sigma^j_x(y(s), u^\theta(s)) - \sigma^j_x(y(s), u(s))\right) y_1(s)ds. \end{aligned}$$

Thus we can rewrite (2.45) as

$$\begin{aligned}
&\mathbb{E}\int_0^T (H(y(s),u^\theta(s),p(s),K(s)) - (H(y(s),u(s),p(s),K(s)))ds \\
&+\frac{1}{2}\mathbb{E}\int_0^T y_1^*(t)H_{xx}(y(s),u(s),p(s),K(s))y_1(s) \\
&+\frac{1}{2}\mathbb{E}y_1^*(T)h_{xx}(y(T))y_1(T) \geq o(\theta),
\end{aligned} \tag{2.48}$$

where we denote

$$H(x,v,p,K) = l(x,v) + (p,b(x,v)) + \sum_{j=1}^{d}(K_j,\sigma^j(x,v)).$$

The interesting thing is that the quadratic terms of (2.48) can still be treated by applying the Riesz Representation Theorem. Indeed, applying Itô's formula to the matrix-valued processes

$$Y(s) = y_1(s)y_1^*(s) = \begin{matrix} y_1^1y_1^1 & \dots & y_1^1y_1^n \\ \vdots & & \vdots \\ y_1^1y_1^n & \dots & y_1^ny_1^n \end{matrix},$$

we have

$$\begin{aligned}
dY(t) = &\left[Y(t)b_x^*(t) + b_x(t)Y(t) + \sum_{j=1}^{d}\sigma_x^j(t)Y(t)\sigma_x^{*j}(t) + \Phi^\theta(t)\right]dt \\
&+\left[Y(t)\sigma_x^*(t) + \sigma_x(t)Y(t) + \psi^\theta(t)\right]dW_t,
\end{aligned} \tag{2.49}$$

where

$$\begin{aligned}
\Phi^\theta(t) &= y_1(t)\left(b\left(y(t),u^\theta(t)\right) - b\left(y(t),u(t)\right)\right)^* + \left(b\left(y(t),u^\theta(t)\right) - b\left(y(t),u(t)\right)\right)y_1^*(t) \\
&+\sigma_x(t)\,y_1(s)\left(\sigma\left(y(t),u^\theta(t)\right) - \sigma_x(t)\,y(t),u(t)\right))^* \\
&+\left(\sigma\left(y(t),u^\theta(t)\right) - \sigma\left(y(t),u(t)\right)\right)y_1^*(t)\,\sigma_x^*(t) \\
&+\left(\sigma\left(y(t),u^\theta(t)\right) - \sigma\left(y(t),u(t)\right)\right)\left(\sigma\left(y(t),u^\theta(t)\right) - \sigma\left(y(t),u(t)\right)\right)^* \\
\Psi^\theta(t) &= y_1(t)\left(\sigma\left(y(t),u^\theta(t)\right) - \sigma\left(y(t),u^\theta(t)\right)\right)^* + \left(\sigma\left(y(t),u^\theta(t)\right) - \sigma\left(y(t),u^\theta(t)\right)\right)y_1^*(t).
\end{aligned}$$

Now, we define the following symmetric matrix-valued linear SDE:

$$\begin{cases}
dZ(t) &= \left[Z(t)b_x^*(t) + b_x(t)Z(t) + \sum_{j=1}^{d}\sigma_x^j(t)Z(t)\sigma_x^{j*}(t) + \Phi(t)\right]dt \\
&\quad + \left[Z(t)\sigma_x^*(t) + Z(t)\sigma_x(t) + \Psi(t)\right]dW_t \\
Z(0) &= 0,
\end{cases}$$

$$(\Phi(t),\Psi(t)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n) \times L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)^d, \qquad \Psi = (\psi_1,\dots,\psi_d),$$

where $\mathbb{R}^{n,n}$ is the space of all $n \times n$ real symmetric matrices with the following scalar product:

$$(A_1, A_2)_* = tr(A_1 A_2) \quad \forall A_1, A_2 \in \mathbb{R}^{n,n}.$$

Now, let us construct a linear functional via (2.49)

$$M(\Phi(t), \Psi(t)) = \mathbb{E}\int_0^T (Z(t)H_{xx}(t))_* \, dt + \mathbb{E}(Z(T)h_{xx}(y(T)))_* \tag{2.50}$$

Obviously, $M(\Phi(t), \Psi(t))$ is a linear continuous functional on

$$L^2_{\mathcal{F}}(0, T; \mathbb{R}^{n,n}) \times L^2_{\mathcal{F}}(0, T; \mathbb{R}^{n,n})^d,$$

there is a unique pair $(P(\cdot), Q(\cdot)) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^{n,n}) \times L^2_{\mathcal{F}}(0, T; \mathbb{R}^{n,n})^d$ such that

$$M(\Phi(t), \Psi(t)) = \mathbb{E}\left(\int_0^T \left[(P(t), \Phi(t))_* + \sum_{j=1}^d \left(Q^j(t), \Psi(t)^j\right)_*\right] dt\right) \tag{2.51}$$

Since for all $y \in \mathbb{R}^n$, $A \in \mathbb{R}^{n,n}$

$$(yy^*, A)_* = tr(yy^*A) = y^*Ay,$$

from (2.49), (2.50), (2.51) we can rewrite (2.48)

$$\begin{aligned}&\mathbb{E}\int_0^T \left[H\left(y(s), u^\theta(s), p(s), K(s)\right) - H(y(s), u(s), p(s), K(s))\right] ds \\ &+\mathbb{E}\int_0^T \left[\left(P(s)\phi^\theta(s)\right)_* + \sum_{j=1}^d \left(Q_j(s)\psi(s)^\theta_j\right)_*\right] ds \geq o(\theta).\end{aligned}$$

From the definition of $\Phi^\theta(t)$ et $\Psi^\theta(t)$, we get:

$$\begin{aligned}&\mathbb{E}\left(\int_0^T \left[H\left(y(s), u^\theta(s), p(s), K(s)\right) - H(y(s), u(s), p(s), K(s))\right]\right) ds \\ &+\frac{1}{2}\mathbb{E}\left(\int_0^T tr\left[\left(\sigma\left(y(s), u^\theta(s)\right) - \sigma y(s), u(s))\right)^* P(t)\left(\sigma\left(y(s), u^\theta(s)\right) - \sigma(y(s), u(s))\right)\right] ds\right) \geq o(\theta).\end{aligned}$$

Finally, we have

$$\begin{aligned}&H(y(\tau), v, p(\tau), K(\tau)) - H(y(\tau), u(\tau), p(\tau), K(\tau)) \\ &+\frac{1}{2}tr\left[(\sigma(y(\tau), v) - \sigma(y(\tau), u(\tau)))^* P(\tau)(\sigma(y(\tau), v) - \sigma(y(\tau), u(\tau)))\right] \geq 0 \quad \forall v \in U, \quad a.e, a.s\end{aligned}$$

or, equivalently

$$\begin{aligned}&H(y(\tau), v, p(\tau), K(\tau) - P(\tau)\sigma(y(\tau), u(\tau)))) + \frac{1}{2}tr\left(\sigma\sigma^*(y(\tau), v) P(\tau)\right) \\ &\geq H(y(\tau), u(\tau), p(\tau), K(\tau) - P(\tau)\sigma(y(\tau), u(\tau)))) + \frac{1}{2}tr\left(\sigma\sigma^*(y(\tau), u(\tau)) P(\tau)\right); \forall v \in U, a.e, a.s\end{aligned} \tag{2.52}$$

2.4.4 Adjoint equations and the maximum principle

The first-order adjoint equation is the classical one. In fact, from [2] and [42], the first-order adjoint process $(p(\cdot), K(\cdot))$ described in a unique way by (2.46), (2.47) is the unique solution of

$$\begin{cases} -dp(t) &= \left[b_x^*(y(t),u(t))p(t) + \sum_{j=1}^{d} \sigma_x^{j*}(y(t),u(t))K_j(t) + l_x(y(t),u(t)) \right] dt - K(t)dW(t), \\ p(T) &= h_x(y(T)). \end{cases} \tag{2.53}$$

We can also use this result to obtain an equation for $(P(\cdot), Q(\cdot))$. In fact, $(P(.), Q(.))$ is uniquely determined by (2.50), (2.51). Thus exactly as in [2] and [42], we can obtain

$$\begin{cases} -dP(t) &= \Bigg[b_x^*(y(t),u(t))P(t) + P(t)b_x(y(t),u(t)) + \sum_{j=1}^{d} \sigma_x^{j*}(y(t),u(t))P(t)\sigma_x^j(y(t),u(t)) \\ &\quad + \sum_{j=1}^{d} \sigma_x^{j*}(y(t),u(t))Q_j(t) + \sum_{j=1}^{d} Q_j(t)\sigma_x^j(y(t),u(t)) + H_{xx}(y(t),u(t),p(t),K(t)) \Bigg] dt \\ &\quad - Q(t)dW(t) \\ P(T) &= h_{xx}(T). \end{cases} \tag{2.54}$$

Theorem 2.4.1

Let (2.37) hold. If $(y(\cdot), u(\cdot))$ is a solution of the optimal control problem (2.35), (2.36), then we have

$$(p(\cdot), K(\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n) \times L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)^d,$$
$$(P(\cdot), Q(\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^{n,n}) \times L^2_{\mathcal{F}}(0,T;\mathbb{R}^{n,n})^d,$$

which are, respectively, solutions of (2.53) and (2.54) such that the variational inequality (2.52) holds.

Chapter 3

Pointwise second order necessary conditions for stochastic optimal control

3.1 Preliminaries and assumptions

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathcal{P})$ be a complete probability space with filtration, we assume that $\mathbb{F} = \{\mathcal{F}_t\}_{0\leq t\leq T}$ is the natural filtration generated by one-dimensional standard Brownian motion $W(\cdot)$. For a fuction $\phi : [0,T] \times \mathbb{R} \times U \times \Omega \to \mathbb{R}$, we denote by $\phi_x(t,x,u)$ (resp. $\phi_u(t,x,u)$) the first order derivatives of ϕ with respect to x and u at (t,x,u,ω), and by $\phi_{(x,u)^2}(t,x,u)$ the second order derivatives of ϕ with respect to (x,u) at (t,x,u,ω) and by $\phi_{xx}(t,x,u)$, $\phi_{xu}(t,x,u)$, and $\phi_{uu}(t,x,u)$ the second order derivatives of ϕ at (t,x,u,ω), we denote by $\mathcal{U}_{ad}$ the set of all admissible controls. Note that, we take out the $\omega\,(\in \Omega)$ argument in the defined functions, when the conditions is clear as habitual.

We introduce some spaces of random variable and stochastic processes, for any $t \in [0,T]$, we let

- $L^2_{\mathcal{F}_t}(\Omega;\mathbb{R})$ the space of $\mathbb{R}$-valued, $\mathcal{F}_t$-measurable random variables ζ such that
$$\mathbb{E}\,|\zeta|^2 < \infty.$$
- $\mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})$ the space of $\mathbb{R}$-valued, $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable, $\mathbb{F}$-adapted processes ψ such that
$$\|\psi\|_{\mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})} := \left[\mathbb{E}\left(\int_0^T |\psi(t)|^2\, dt\right)\right]^{\frac{1}{2}} < \infty.$$

Now, we introduce the following definition of singular control in the classical sense for diffusion , which was motivated in [58, 20].

Definition 3.1.1

(Singular control in the classical sense) An admissible control $\widetilde{u}(\cdot)$ is called singular in the classical sense if satisfies

$$\begin{cases} H_u\left(t, \widetilde{x}\left(t\right), \widetilde{u}\left(t\right), \widetilde{p}\left(t\right), \widetilde{q}\left(t\right)\right) = 0 \quad a.s.\ a.e., \\ H_{uu}\left(t, \widetilde{x}\left(t\right), \widetilde{u}\left(t\right), \widetilde{p}\left(t\right), \widetilde{q}\left(t\right)\right) + \widetilde{P}\left(t\right)\left(\sigma_u\left(t, \widetilde{x}\left(t\right), \widetilde{u}\left(t\right)\right)\right)^2 = 0 \;\; a.s.\ a.e., \end{cases} \tag{3.1}$$

where $\left(\widetilde{p}\left(\cdot\right), \widetilde{q}\left(\cdot\right),\right)$ and $\left(\widetilde{P}\left(\cdot\right), \widetilde{Q}\left(\cdot\right)\right)$ are the adjoint processes given respectively by (3.15) and (3.16) with $\left(\overline{x}\left(\cdot\right), \overline{u}\left(\cdot\right)\right)$ replaced by $\left(\widetilde{x}\left(\cdot\right), \widetilde{u}\left(\cdot\right)\right)$. If $\widetilde{u}(\cdot)$ in (3.1) is also optimal, then we call $\widetilde{u}(\cdot)$ a singular optimal control in the classical sense.

We consider the following controlled stochastic differential equations

$$\begin{cases} dx\left(t\right) = b\left(t, x\left(t\right), u\left(t\right)\right) dt + \sigma(t, x\left(t\right), u\left(t\right)) dW\left(t\right), \\ x(0) = x_0, \end{cases} \tag{3.2}$$

where $b : [0, T] \times \mathbb{R} \times U \to \mathbb{R}$ and $\sigma : [0, T] \times \mathbb{R} \times U \to \mathbb{R}$, with a cost functional

$$J(u\left(\cdot\right)) = \mathbb{E}\left[\int_0^T f\left(t, x\left(t\right), u\left(t\right)\right) dt + h(x\left(T\right))\right], \;\; u\left(\cdot\right) \in \mathcal{U}_{ad}, \tag{3.3}$$

and $f : [0, T] \times \mathbb{R} \times U \to \mathbb{R}, h : \mathbb{R} \to \mathbb{R}$ are given functions.

The stochastic optimization problem which we interest is to find a control $\overline{u}\left(\cdot\right) \in \mathcal{U}_{ad}$ such that

$$J(\overline{u}\left(\cdot\right)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}} J\left(u\left(\cdot\right)\right). \tag{3.4}$$

Any admissible control $\overline{u}\left(\cdot\right) \in \mathcal{U}_{ad}$ that achieves the minimum is called an optimal control.

We also assume that

Assumptions (A1)

1. The maps b and σ are $\mathcal{B}([0, T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted.
2. The functions b and σ are continuously differentiable up to the second order with respect to (x, u).
3. All the first order derivatives are continuous in (x, u) and uniformly bounded.
4. There exists a constant $K_1 > 0$ such that for almost all $(t, \omega) \in [0, T] \times \Omega$ and for any

$x, \widetilde{x} \in \mathbb{R}$ and $u, \widetilde{u} \in U$,

$$\begin{cases} |\phi(t,x,u)| \leq K_1, \text{ for } \phi = b, \sigma, \\ |\phi(t,x,u) - \phi(t,\widetilde{x},u)| \leq K_1 |x - \widetilde{x}|, \text{ for } \phi = b, \sigma, \\ \left|\phi_{(x,u)^2}(t,x,u) - \phi_{(x,u)^2}(t,\widetilde{x},\widetilde{u})\right| \leq K_1 (|x - \widetilde{x}| + |u - \widetilde{u}|), \end{cases}$$

Assumptions (A2)

1. The process f is $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted.
2. The random variable h is $\mathbb{F}_T$ -measurable.
3. The process f is bounded by $K_2(1 + |x|^2 + |u|^2)$ and h is bounded by $K_2(1 + |x|^2)$.
4. The maps f and h are continuously differentiable up to the second order.
5. For any $x, \widetilde{x} \in \mathbb{R}$ and $u, \widetilde{u} \in U$,

$$\begin{cases} |f_x(t,x,u)| + |f_u(t,x,u)| \leq K_2(1 + |x| + |u|), \ \ |h_x(x)| \leq K_2(1 + |x|), \\ |f_{xx}(t,x,u)| + |f_{xu}(t,x,u)| + |f_{uu}(t,x,u)| \leq K_2, \\ |h_{xx}(x)| \leq K_2, \ \ |h_{xx}(x) + h_{xx}(\widetilde{x})| \leq K_2 |x - \widetilde{x}|, \\ \left|f_{(x,u)^2}(t,x,u) - f_{(x,u)^2}(t,\widetilde{x},\widetilde{u})\right| \leq K_2 (|x - \widetilde{x}| + |u - \widetilde{u}|). \end{cases}$$

Under assumptions (A1) and (A2), equation (3.2) has a unique strong solution $x(t)$ given by

$$x(t) = x_0 + \int_0^t b(s, x(s), u(s))\, ds + \int_0^t \sigma(s, x(s), u(s))\, dW(s)$$

and by standard arguments it is easy to show that for any $C_k > 0$, it holds that

$$\mathbb{E}(\sup_{t \in [0,T]} |x(t)|^k) < C_k,$$

where C_k is a constant depending only on k. Moreover, the functional (3.3) is well defined from $\mathcal{U}_{ad}$ into $\mathbb{R}$.

3.2 Second order necessary condition in integral form

In this section, we prove an integral type second order necessary condition for stochastic optimal control. We consider a control region U is nonempty and bounded. Moreover, a convex perturbation of the optimal control defined by $u^\theta(t) = \overline{u}(t) + \theta(u(t) - \overline{u}(t))$, for $u(\cdot) \in \mathcal{U}_{ad}$ and $\theta \in (0,1)$. The convexity condition of the control domain ensures that $u^{\theta}(\cdot) \in \mathcal{U}_{ad}$.

For convenience, we will use the following notations, we denote by $x^{\theta}(\cdot)$, $\overline{x}(\cdot)$ the state trajectory of the SDE (3.2) corresponding to $u^{\theta}(\cdot)$ and $\overline{u}(\cdot)$.
To simplify our notation, we write for $\phi = b, \sigma, f$:

$$\begin{cases} \delta\phi(t) = \phi\left(t, x^{\theta}(t), u^{\theta}(t)\right) - \phi(t, \overline{x}(t), \overline{u}(t)), \\ \phi_x(t) = \phi_x(t, \overline{x}(t), \overline{u}(t)), \qquad \phi_u(t) = \phi_u(t, \overline{x}(t), \overline{u}(t)), \\ \phi_{xx}(t) = \phi_{xx}(t, \overline{x}(t), \overline{u}(t)), \quad \phi_{uu}(t) = \phi_{uu}(t, \overline{x}(t), \overline{u}(t)), \\ , \phi_{xu}(t) = \phi_{xu}(t, \overline{x}(t), \overline{u}(t)), \end{cases}$$

We introduce the following variational equations

$$\begin{cases} dy_1(t) = \{b_x(t) y_1(t) + b_u(t) v(t)\} dt + \{\sigma_x(t) y_1(t) + \sigma_u(t) v(t)\} dW(t) \\ y_1(0) = 0, \end{cases} \tag{3.5}$$

and

$$\begin{cases} dy_2(t) = [b_x(t) y_2(t) + b_{xx}(t) y_1(t)^2 + 2b_{xu}(t) y_1(t) v(t) + b_{uu}(t) v(t)^2]dt \\ \qquad +[\sigma_x(t) y_2(t) + \sigma_{xx}(t) y_1(t)^2 + 2\sigma_{xu}(t) y_1(t) v(t) + \sigma_{uu}(t) v(t)^2]dW(t) \\ \qquad y_2(0) = 0. \end{cases} \tag{3.6}$$

Remark 3.2.1

Under assumptions (A1), (A2) the variational equations (3.5) and (3.6) admits a unique strong solutions $y_1(t)$ and $y_2(t)$ respectively.

Next, we prove the proposition which plays a crucial role in obtaining a second order necessary conditions.

We note that unless specified, for each $k \in \mathbb{R}_+$, we will denote by $C_k > 0$ a generic positive constant depending only on k and the constants appearing in Proposition 3.2.1 , which may vary from line to line.

Proposition 3.2.1

Assume that assumptions (A1), (A2) satisfied. Then, for any $k \geq 1$, we have the *following basic estimates*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|x^{\theta}\left(t\right)-\overline{x}\left(t\right)\right|^{2k}\right]\leq C_{k}\theta^{k}, \tag{3.7}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|y_{1}\left(t\right)\right|^{2k}\right]\leq C_{k}, \tag{3.8}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|y_{2}\left(t\right)\right|^{2k}\right]\leq C_{k}, \tag{3.9}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|x^{\theta}\left(t\right)-\overline{x}\left(t\right)-\theta y_{1}\left(t\right)\right|^{2k}\right]\leq C_{k}\theta^{2k}, \tag{3.10}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|x^{\theta}\left(t\right)-\overline{x}\left(t\right)-\theta y_{1}\left(t\right)-\frac{\theta^{2}}{2}y_{2}\left(t\right)\right|^{2k}\right]\leq C_{k}\theta^{2k}. \tag{3.11}$$

Proof: Let $\overline{x}(\cdot)$ and $x^{\theta}(\cdot)$ be the trajectory of (3.2) corresponding to $\overline{u}(\cdot)$ and $u^{\theta}(\cdot)$ resp. Let $y_1(\cdot)$ and $y_2(\cdot)$ be the solution of first and second order adjoint equations (3.5)-(3.6). Noting that estimate (3.7) follows from standard arguments, using *Burkholder-Davis-Gundy inequality* for the martingale part and Propositions 1.3.1 . In what follows we shall refer to equation (3.5) as the first-order variational equation, and the process $y_1(\cdot)$ is called the *first order variational process*. A very important step in Peng [41], and Tang and Li [50] is in light of the Taylor expansion, to find a process $y_2(t)$ so that $x^{\theta}(t)-\overline{x}(t)-\theta y_1(t)-\frac{\theta^2}{2}y_2(t)=o\left(\theta^2\right)$, as $\theta\to 0$, and that the convergence is of an appropriate order. The process $y_2(\cdot)$ is called the *second-order variational process*. So the estimates (3.8), (3.9) and (3.10) are obvious and standard, see also [50, Lemma 2.1].

Now, we start to prove the estimate (3.11). From (3.2), (3.5) and (3.6), we have

$$\begin{aligned}
&\left|x^{\theta}(t)-\overline{x}(t)-\theta y_1(t)-\frac{\theta^2}{2}y_2(t)\right|^{2k}\\
&=\left|\int_0^t[\delta b\left(s\right)-\theta\left[b_x(s)y_1(s)+b_u(s)v(s)\right]\right.\\
&-\frac{\theta^2}{2}\left[b_x(s)y_2(s)+b_{xx}(s)y_1(s)^2+2b_{xu}(s)y_1(s)v(s)+b_{uu}(s)v(s)^2\right]]ds\\
&+\int_0^t[\delta\sigma\left(s\right)-\theta\left[\sigma_x(s)y_1(s)+\sigma_u(s)v(s)\right]\\
&-\frac{\theta^2}{2}\left[\sigma_x(s)y_2(s)+\sigma_{xx}(s)y_1(s)^2+2\sigma_{xu}(s)y_1(s)v(s)\right.\\
&\left.+\sigma_{uu}(s)v(s)^2\right]]dW(s)
\end{aligned}$$

straight forward calculation by applying the Cauchy-Schwarz inequality, we shows that

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|x^{\theta}(t)-\overline{x}(t)-\theta y_1(t)-\frac{\theta^2}{2}y_2(t)\right|^{2k}\right] \tag{3.12}$$
$$\leq I_1,$$

where

$$\begin{aligned} I_1 = \mathbb{E}\Bigg[2\sup_{t\in[0,T]}\Bigg|\int_0^t [\delta b\,(s)-\theta\,[b_x(s)y_1(s)+b_u(s)v(s)] \\ -\frac{\theta^2}{2}\left[b_x(s)y_2(s)+b_{xx}(s)y_1(s)^2+2b_{xu}(s)y_1(s)v(s)+b_{uu}(s)v(s)^2\right]\Big]\,ds \\ +\int_0^t [\delta\sigma\,(s)-\theta\,[\sigma_x(s)y_1(s)+\sigma_u(s)v(s)] \\ -\frac{\theta^2}{2}\left[\sigma_x(s)y_2(s)+\sigma_{xx}(s)y_1(s)^2+2\sigma_{xu}(s)y_1(s)v(s)+\sigma_{uu}(s)v(s)^2\right]\Big]\,dW(s)\Bigg|^{2k}\Bigg], \end{aligned} \tag{3.13}$$

Similar to Bonnans [7], Zhang and Zhang [58], by applying the Cauchy-Schwarz inequality and the Burkhölder–Davis–Gundy inequality, we have

$$I_1 \leq C_k\theta^{2k}. \tag{3.14}$$

By combining (3.12), (3.14) the desired result (3.11) is fulfilled. Thus, the proof of Proposition 3.2.1 is completed. ■

Define the Hamiltonian function $H:[0,T]\times\mathbb{R}\times U\times\mathbb{R}\times\mathbb{R}$ by

$$H\,(t,x,u,p,q) := b\,(t,x,u)\,p+\sigma\,(t,x,u)\,q-f(t,x,u).$$

Now, we introduce the first adjoint equation

$$\begin{cases} dp(t) = -\left\{b_x\,(t)\,p\,(t)+\sigma_x\,(t)\,q\,(t)-f_x\,(t)\right\}dt \\ \qquad\qquad +q\,(t)\,dW\,(t)\,, \\ p\,(T) = -h_x\,(\overline{x}\,(T))\,, \end{cases} \tag{3.15}$$

and the second adjoint equation

$$\begin{cases} dP(t) = -\Big\{2b_x(t)P(t) + \sigma_x(t)^2 P(t) + 2\sigma_x(t)Q(t) \\ \qquad +H_{xx}(t)\,dt \\ \qquad +Q(t)\,dW(t), \\ P(T) = -h_{xx}(\overline{x}(T)), \end{cases} \tag{3.16}$$

where

$$H_{xx}(t) = b_{xx}(t)\,p + \sigma_{xx}(t)\,q - f_{xx}(t).$$

It is easy to prove that under assumptions (A1)-(A2), Eqs-(3.15) and (3.16) are classical linear backward stochastic differential equations (BSDEs in short) admit a unique strong solution such that

$$(p(t), q(t)) \in \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})$$
$$(P(t), Q(t),) \in \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})$$

Also, we define the functional $\mathbb{H} : [0,T] \times \mathbb{R} \times U \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ by

$$\begin{aligned} \mathbb{H}(t,x,u,p,q,P,Q) := {} & H_{xu}(t,x,u,p,q) + b_u(t,x,u)\,P(t) \\ & + \sigma_u(t,x,u)\,Q(t) + \sigma_u(t,x,u)\,P(t)\sigma_x(t,x,u) \end{aligned} \tag{3.17}$$

To simplify our notation, we set

$$\mathbb{H}(t) = \mathbb{H}(t, \overline{x}(t), \overline{u}(t), p(t), q(t), P(t), Q(t)),\ t \in [0,T].$$

Lemma 3.2.1

Let $(p(t), q(t))$ be the solution of the adjoint equation (3.15), $(P(t), Q(t))$ be the solution of the adjoint equation (3.16), and $y_1(t)$, $y_2(t)$ are the solution to the first and second variational equations (3.5) and (3.6) respectively. Then the following duality relations hold:

$$-\mathbb{E}[p(T)y_1(T)] = -\mathbb{E}\left[\int_0^T \{p(t)(b_u(t)v(t)) + q(t)(\sigma_u(t)v(t))\}\,dt\right] - \mathbb{E}\left[\int_0^T f_x(t)y_1(t)\,dt\right], \tag{3.18}$$

$$\mathbb{E}\left[p\left(T\right)y_{2}\left(T\right)\right]=-\mathbb{E}\left[\int_{0}^{T}p\left(t\right)\left\{b_{xx}\left(t\right)y_{1}\left(t\right)^{2}+2b_{xu}\left(t\right)y_{1}\left(t\right)v\left(t\right)+b_{uu}\left(t\right)v\left(t\right)^{2}\right\}dt\right] \tag{3.19}$$

$$\mathbb{E}\left[\int_{0}^{T}q\left(t\right)\left\{\sigma_{xx}\left(t\right)y_{1}\left(t\right)^{2}+2\sigma_{xu}\left(t\right)y_{1}\left(t\right)v\left(t\right)+\sigma_{uu}\left(t\right)v\left(t\right)^{2}\right\}dt\right]$$
$$-\mathbb{E}\left[\int_{0}^{T}f_{x}\left(t\right)y_{2}\left(t\right)dt\right],$$

and

$$-\mathbb{E}\left[P\left(T\right)y_{1}\left(T\right)^{2}\right]=-2\mathbb{E}\left[\int_{0}^{T}\left\{P\left(t\right)y_{1}\left(t\right)\left(b_{u}\left(t\right)v\left(t\right)\right)+P\left(t\right)\sigma_{x}\left(t\right)y_{1}\left(t\right)\left(\sigma_{u}\left(t\right)v\left(t\right)\right)\right\}dt\right] \tag{3.20}$$

$$-2\mathbb{E}\left[\int_{0}^{T}\left\{Q\left(t\right)\sigma_{u}\left(t\right)v\left(t\right)y_{1}\left(t\right)\right\}dt\right]-\mathbb{E}\left[\int_{0}^{T}P\left(t\right)\left(\sigma_{u}\left(t\right)v\left(t\right)\right)^{2}dt\right]$$
$$+\mathbb{E}\left[\int_{0}^{T}H_{xx}\left(t\right)y_{1}\left(t\right)^{2}dt\right].$$

Proof: The proof of this lemma follows directly from Itô's formula to $p\left(t\right)y_{1}\left(t\right)$ and taking expectation where $y_1(0)=0$, we have

$$\mathbb{E}\left[p\left(T\right)y_{1}\left(T\right)\right]=-\mathbb{E}\int_{0}^{T}p\left(t\right)dy_{1}\left(t\right)-\mathbb{E}\int_{0}^{T}y_{1}\left(t\right)dp\left(t\right) \tag{3.21}$$
$$\mathbb{E}\int_{0}^{T}q\left(t\right)\left\{\sigma_{x}\left(t\right)y_{1}\left(t\right)+\sigma_{u}\left(t\right)v\left(t\right)\right\}dt$$

where

$$-\mathbb{E}\int_{0}^{T}p\left(t\right)dy_{1}\left(t\right)=-\mathbb{E}\int_{0}^{T}p\left(t\right)\left[b_{x}\left(t\right)y_{1}\left(t\right)+b_{u}\left(t\right)v\left(t\right)\right]dt. \tag{3.22}$$

Consequently

$$\mathbb{E}\int_{0}^{T}y_{1}\left(t\right)dp\left(t\right) \tag{3.23}$$
$$=\mathbb{E}\int_{0}^{T}y_{1}\left(t\right)\left[b_{x}\left(t\right)p\left(t\right)+\sigma_{x}\left(t\right)q\left(t\right)-f_{x}\left(t\right)\right]dt,$$

substituting (3.22), (3.23), into (3.21), then the desired result (3.18) is fulfilled.

Now, applying Itô's formula to $p\left(t\right)y_{2}\left(t\right)$ and taking expectation where $y_2(0)=0$, we have

$$-\mathbb{E}\left[p\left(T\right)y_2\left(T\right)\right] = -\mathbb{E}\int_0^T p\left(t\right)dy_2\left(t\right) - \mathbb{E}\int_0^T y_2\left(t\right)dp\left(t\right) \tag{3.24}$$
$$-\mathbb{E}\left[\int_0^T q\left(t\right)\left\{\sigma_x\left(t\right)y_2\left(t\right) + \sigma_{xx}\left(t\right)y_1\left(t\right)^2\right.\right.$$
$$\left.\left. + 2\sigma_{xu}\left(t\right)y_1\left(t\right)v\left(t\right) + \sigma_{uu}\left(t\right)v\left(t\right)^2\right\}dt\right]$$

where

$$-\mathbb{E}\int_0^T p\left(t\right)dy_2\left(t\right) = -\mathbb{E}\int_0^T p\left(t\right)\left\{b_x\left(t\right)y_2\left(t\right) + b_{xx}\left(t\right)y_1\left(t\right)^2\right.$$
$$\left. + 2b_{xu}\left(t\right)y_1\left(t\right)v\left(t\right) + b_{uu}\left(t\right)v\left(t\right)^2\right\}dt, \tag{3.25}$$

and

$$-\mathbb{E}\int_0^T y_2\left(t\right)dp\left(t\right) \tag{3.26}$$
$$= \mathbb{E}\int_0^T y_2\left(t\right)\left[b_x\left(t\right)p\left(t\right) + \sigma_x\left(t\right)q\left(t\right) - f_x\left(t\right)\right]dt,$$

substituting (3.25), (3.26), into (3.24), we obtain the desired result (3.19).

Next, applying Itô's formula to $P\left(t\right)y_1\left(t\right)$, where $y_1(0) = 0$, we have

$$\left[P\left(T\right)y_1\left(T\right)\right] = \int_0^T P\left(t\right)dy_1\left(t\right) + \int_0^T y_1\left(t\right)dP\left(t\right)$$
$$+ \int_0^T Q\left(t\right)\left\{\sigma_x\left(t\right)y_1\left(t\right) + \sigma_u\left(t\right)v\left(t\right)\right\}dt$$
$$= I_1 + I_2 + I_3, \tag{3.27}$$

where

$$I_1 = \int_0^T P\left(t\right)dy_1\left(t\right)$$
$$= \int_0^T P\left(t\right)\left\{b_x\left(t\right)y_1\left(t\right) + b_u\left(t\right)v\left(t\right)\right\}dt$$
$$+ \int_0^T P\left(t\right)\left\{\sigma_x\left(t\right)y_1\left(t\right) + \sigma_u\left(t\right)v\left(t\right)\right\}dW(t)$$

by simple computations , we can prove

$$I_2 = \int_0^T y_1\left(t\right)dP\left(t\right)$$
$$= -\int_0^T y_1\left(t\right)\left\{2b_x\left(t\right)P\left(t\right) + \sigma_x\left(t\right)^2 P(t) + 2\sigma_x\left(t\right)Q\left(t\right)\right.$$
$$\left. + H_{xx}\left(t\right)dt + \int_0^T y_1\left(t\right)Q\left(t\right)dW\left(t\right),\right.$$

$$I_3 = \int_0^T \{Q(t)\sigma_x(t)y_1(t) + Q(t)\sigma_u(t)v(t)\}\,dt,$$

Then we can write (3.27) as follows

$$\begin{aligned}[P(T)y_1(T)] = &\int_0^T [P(t)b_u(t)v(t)dt + Q(t)\sigma_u(t)v(t)\\ &- y_1(t)b_x(t)P(t) - y_1(t)\sigma_x(t)^2 P(t) - y_1(t)\sigma_x(t)Q(t)\\ &- y_1(t)H_{xx}(t)]\,dt\\ &+ \int_0^T [P(t)\sigma_x(t)y_1(t) + P(t)\sigma_u(t)v(t) + y_1(t)Q(t)]\,dW(t)\end{aligned}$$

Now, we applying Itô's formula to $(P(t)y_1(t))y_1(t)$ and taking expectation, we obtain

$$\begin{aligned}&-\mathbb{E}\left[P(T)y_1(T)^2\right]\\ &= -\mathbb{E}\int_0^T P(t)y_1(t)dy_1(t) - \mathbb{E}\int_0^T y_1(t)d(P(t)y_1(t))\\ &-\mathbb{E}\left[\int_0^T \{\sigma_x(t)y_1(t) + \sigma_u(t)v(t)\}\{P(t)\sigma_x(t)y_1(t) + P(t)\sigma_u(t)v(t) + y_1(t)Q(t)\}\,dt\right]\\ &= J_1 + J_2 + J_3, \end{aligned} \tag{3.28}$$

where

$$J_1 = -\mathbb{E}\int_0^T P(t)y_1(t)dy_1(t) = -\mathbb{E}\int_0^T P(t)y_1(t)\{b_x(t)y_1(t) + b_u(t)v(t)\}\,dt, \tag{3.29}$$

$$\begin{aligned}J_2 &= -\mathbb{E}\int_0^T y_1(t)d(P(t)y_1(t))\\ &= -\mathbb{E}\int_0^T y_1(t)[P(t)b_u(t)v(t)dt + Q(t)\sigma_u(t)v(t)\\ &- y_1(t)b_x(t)P(t) - y_1(t)\sigma_x(t)^2 P(t) - y_1(t)\sigma_x(t)Q(t)\\ &- -y_1(t)H_{xx}(t)]\,dt, \end{aligned} \tag{3.30}$$

and it is easy to show that

$$\begin{aligned}J_3 &= -\mathbb{E}\left[\int_0^T \{\sigma_x(t)y_1(t) + \sigma_u(t)v(t)\}\{P(t)\sigma_x(t)y_1(t) + P(t)\sigma_u(t)v(t) + y_1(t)Q(t)\}\,dt\right]\\ &= -\mathbb{E}\left[\int_0^T \Big\{P(t)(\sigma_x(t)y_1(t))^2 + 2P(t)\sigma_x(t)y_1(t)\sigma_u(t)v(t) + \sigma_x(t)Q(t)y_1(t)^2\right.\\ &\left. + P(t)(\sigma_u(t)v(t))^2 + Q(t)\sigma_u(t)y_1(t)v(t)\Big\}\,dt\right]. \end{aligned} \tag{3.31}$$

Similarly, we have Finally, substituting (3.29), (3.30), (3.31), into (3.28), then (3.20) is fulfilled.

This completes the proof of Lemma 3.2.1. ■

To prove the main theorem we need the following technical result.

Proposition 3.2.2

Let (A1)-(A2) hold. Then, for any $u(\cdot) \in \mathcal{U}_{ad}$ we have

$$\begin{aligned} & J(u^\theta(\cdot)) - J(\overline{u}(\cdot)) \\ & = -\mathbb{E}\int_0^T \left[\theta \left\{H_u(t)\, v(t)\right\} + \frac{\theta^2}{2}\left\{H_{uu}(t)\, v(t)^2\right\}\right. \\ & \quad \left. + \frac{\theta^2}{2}\left\{P(t)\,(\sigma_u(t)\, v(t))^2\right\} + \theta^2\left\{\mathbb{H}(t)\, y_1(t) v(t)\right\}\right] dt + o\left(\theta^2\right), \quad \left(\theta \to 0^+\right), \end{aligned} \tag{3.32}$$

where

$$H_u(t) = H_u(t, \overline{x}(t), \overline{u}(t), p(t), q(t)),$$
$$H_{uu}(t) = H_{uu}(t, \overline{x}(t), \overline{u}(t), p(t), q(t)).$$

Proof: By applying Taylor's formula, we get

$$\begin{aligned} & J\left(u^\theta(\cdot)\right) - J(\overline{u}(\cdot)) \\ & = \mathbb{E}\left[\int_0^T \{\delta f(t)\}\, dt\right] + \mathbb{E}\left[h(x^\theta(T) - h(\overline{x}(T)\right] \\ & = \mathbb{E}\left[\int_0^T \left\{f_x(t)\left(x^\theta(t) - \overline{x}(t)\right) + f_u(t)\left(u^\theta(t) - \overline{u}(t)\right) + \frac{1}{2} f_{xx}(t)\left(x^\theta(t) - \overline{x}(t)\right)^2\right.\right. \\ & + f_{xu}(t)\left(x^\theta(t) - \overline{x}(t)\right)\left(u^\theta(t) - \overline{u}(t)\right) + \left.\left.\frac{1}{2} f_{uu}(t)\left(u^\theta(t) - \overline{u}(t)\right)^2\right\} dt\right] \\ & + \mathbb{E}\left[h_x(\overline{x}(T))\left(x^\theta(T) - \overline{x}(T)\right) + \frac{1}{2} h_{xx}(\overline{x}(T))\left(x^\theta(T) - \overline{x}(T)\right)^2\right] + o\left(\theta^2\right). \end{aligned}$$

Using Proposition 3.2.1 , we have

$$\begin{aligned}
&J\left(u^{\theta}(\cdot)\right)-J(\overline{u}(\cdot)) \\
&=\mathbb{E}\left[\int_0^T\left\{\theta f_x(t)\, y_1(t)+\frac{\theta^2}{2} f_x(t)\, y_2(t)+\theta f_u(t)\, v(t)\right.\right. \\
&\left.\left.+\ \frac{\theta^2}{2}\left(f_{xx}(t)\, y_1(t)^2+2 f_{xu}(t) y_1(t)\, v(t)+f_{uu}(t)\, v(t)^2\right)\right\} dt\right] \\
&+\mathbb{E}\left[\theta h_x(\overline{x}(T)\, y_1(T)+\frac{\theta^2}{2} h_x(\overline{x}(T))\, y_2(T)+\frac{\theta^2}{2} h_{xx}(\overline{x}(T))\, y_1(T)^2\right]+o\left(\theta^2\right), \quad \left(\theta \longrightarrow 0^{+}\right).
\end{aligned} \tag{3.33}$$

Substituting (3.18), (3.19), and (3.20) into (3.33), we obtain

$$\begin{aligned}
&J(u^{\theta}(\cdot))-J(\overline{u}(\cdot)) \\
&=-\mathbb{E}\left[\int_0^T \theta\left[p(t)\left(b_u(t)\, v(t)\right)+q(t)\left(\sigma_u(t)\, v(t)\right)-f_u(t)\, v(t)\right] dt\right] \\
&-\mathbb{E}\left[\int_0^T \frac{\theta^2}{2}\left[p(t)\, b_{uu}(t)\, v(t)^2+q(t)\, \sigma_{uu}(t)\, v(t)^2-f_{uu}(t) v(t)^2+P(t)\left(\sigma_u(t)\, v(t)\right)^2\right] dt\right] \\
&-\mathbb{E}\left[\int_0^T \theta^2\left\{p(t)\, b_{xu}(t)\, y_1(t)\, v(t)+q(t)\, \sigma_{xu}(t)\, y_1(t)\, v(t)-f_{xu}(t) y_1(t)\, v(t)\right.\right. \\
&+\ P(t)\, y_1(t)\, b_u(t)\, v(t)+P(t)\, \sigma_x(t)\, y_1(t)\, \sigma_u(t)\, v(t)+Q(t)\, \sigma_u(t)\, v(t)\, y_1(t)\} dt]+o\left(\theta^2\right).
\end{aligned}$$

Finally, we get

$$\begin{aligned}
&J(u^{\theta}(\cdot))-J(\overline{u}(\cdot)) \\
&=-\mathbb{E}\int_0^T\left[\theta\left(H_u(t)\, v(t)\right)+\frac{\theta^2}{2}\left[H_{uu}(t)\, v(t)^2\right]\right. \\
&\quad\left.+\ \frac{\theta^2}{2}\left[P(t)\left(\sigma_u(t)\, v(t)\right)^2\right]+\theta^2\left[\mathbb{H}(t)\, y_1(t) v(t)\right]\right] dt+o\left(\theta^2\right), \\
&\quad\left(\theta \rightarrow 0^{+}\right),
\end{aligned}$$

Thus, the proof of Proposition 3.2.2 is completed. ■

Now, by Proposition 3.2.2 , we can establish the following second order necessary condition in integral form for stochastic optimal control (3.2)-(3.3).

Theorem 3.2.1

Let (A1)-(A2) hold. If $\overline{u}(\cdot)$ is a singular optimal control in the classical sense for the control problem (3.2)-(3.3). Then we have

$$\mathbb{E}\int_0^T \mathbb{H}(t)y_1(t)(u(t)-\overline{u}(t))dt \leq 0, \tag{3.34}$$

for any $u(\cdot) \in \mathcal{U}_{ad}$, where the Hamiltonian $\mathbb{H}$ is defined by (3.17) and $y_1(t)$ solution of first order adjoint equation given by

$$y_1(t) = \int_0^t [b_x(s)\, y_1(s) + b_u(s)\, v(s)]\, ds + [\sigma_x(s)\, y_1(s) + \sigma_u(s)\, v(s)]\, dW(s)$$

Proof : The desired result (3.34) and Proposition 3.2.2 follows immediately from 3.1.

This completes the proof of Theorem 3.2.1 ■

3.3 Pointwise second order maximum principle in terms of the martingale

In this section, by using the property of Itô's integrals and the martingale representation theorem, we establish the second order necessary condition for singular optimal controls, which is pointwise maximum principle in terms of the martingale with respect to the time variable t. The following lemma play an important role to establish our result.

Lemma 3.3.1

The first variational equation (3.1) admits a unique strong solution $y_1(\cdot)$, which is represented by the following:

$$y_1(t) = \Phi(t)\left[\int_0^t \Psi(s)\left(b_u(s) - \sigma_x(s)\sigma_u(s)\right)v(s)ds + \int_0^t \Psi(s)\sigma_u(s)v(s)dW(s)\right], \tag{3.35}$$

where $\Phi(t)$ is a defined by the following linear stochastic differential equation:

$$\begin{cases} d\Phi(t) &= b_x(t)\Phi(t)\,dt + \sigma_x(t)\Phi(t)\,dW(t), \\ \Phi(0) &= 1, \end{cases} \tag{3.36}$$

and $\Psi(t)$ its inverse.

Proof: Equation (3.5) is linear with bounded coefficients, then it admits a unique strong solution. Moreover, this solution is inversible and its inverse $\Psi(t) = \Phi^{-1}(t)$ given by:

$$\begin{cases} d\Psi(t) &= \left[\sigma_x^2(t)\Psi(t) - b_x(t)\Psi(t)\right]dt - \left[\sigma_x(t)\Psi(t)\right]dW(t) \\ \Psi(0) &= 1. \end{cases} \tag{3.37}$$

Applying Itô's formula to $\Psi(t)y_1(t)$ we have

$$\begin{aligned} d\left[\Psi(t)y_1(t)\right] &= y_1(t)\,d\Psi(t) + \Psi(t)dy_1(t) \\ &\quad - \left[\sigma_x(t)\Psi(t)\right]\left[\sigma_x(t)y_1(t) + \sigma_u(t)v(t)\right]dt \\ &= I_1 + I_2 + I_3, \end{aligned} \tag{3.38}$$

where

$$\begin{aligned} I_1 &= y_1(t)\,d\Psi(t) \\ &= \left[y_1(t)\,\sigma_x^2(t)\Psi(t) - y_1(t)\,b_x(t)\Psi(t)\right]dt \\ &- y_1(t)\,\sigma_x(t)\Psi(t)dW(t) \\ &- y_1(t)\,. \end{aligned} \tag{3.39}$$

By simple computations, we obtain

$$\begin{aligned} I_2 &= \Psi(t)dy_1(t) \\ &= [\Psi(t)b_x(t)y_1(t) + \Psi(t)b_u(t)v(t)]\,dt \\ &+ [\Psi(t)\sigma_x(t)y_1(t) + \Psi(t)\sigma_u(t)v(t)]\,dW(t) \end{aligned} \tag{3.40}$$

and

$$I_3 = -[\sigma_x(t)\Psi(t)]\,[\sigma_x(t)y_1(t) + \sigma_u(t)v(t)]\,dt. \tag{3.41}$$

Substituting (3.36), (3.37), and (3.38) into (3.36), we get

$$\begin{aligned} &\Psi(t)y_1(t) - \Psi(0)y_1(0) \\ &= \left[\int_0^t \Psi(s)\,[b_u(s) - \sigma_x(s)\sigma_u(s)]\,v(s)ds\right. \\ &+ \int_0^t \Psi(s)\sigma_u(s)v(s)dW(s). \end{aligned} \tag{3.42}$$

Since $y_1(0) = 0$ and $\Psi^{-1}(t) = \Phi(t)$, then from (3.42) the desired result (3.35) is fulfilled. This completes the proof of Lemma 3.3.1 ■

To prove the main theorem we need the following technical Lemma.

Lemma 3.3.2

Let (A1)-(A2) hold. Then we have

(1) $\mathbb{H}(\cdot) \in \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})$

(2) For any $v \in U$, *there exists* $\phi_v(\cdot,t) \in \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})$, *such that*

$$\mathbb{H}(t)(v - \bar{u}(t)) = \mathbb{E}\left[\mathbb{H}(t)(v - \bar{u}(t))\right] + \int_0^t \phi_v(s,t)\,dW(s) \tag{3.43}$$

$$a.e.\ t \in [0,T],\ \ P - a.s.$$

Proof: (1). The proof follows immediately from **Lemma 3.9** in [58].

(2) The proof of (3.43) follows from Tang and Li [50, Appendix] ■

Now, we return to integral type of second order necessary condition and substituting the explicit representation (3.35) of $y_1(\cdot)$ into (3.34), we see that there appears a "bad" term in the form

$$\mathbb{E}\int_0^T \left[\mathbb{H}\left(t\right)\Phi(t)\int_0^t \Psi(s)\sigma_u(s)v(s)dW(s)\right]v(t)dt, \tag{3.44}$$

For more details see [58, p.2278] for this type of integrals.

Now, in order to derive a pointwise second order necessary condition from the integral form in (3.35), we need to choose the following needle variation for the optimal control $\overline{u}(\cdot)$:

$$u(t)=\begin{cases} v, & t\in A_\theta, \\ \overline{u}(t), & t\in[0,T]\mid A_\theta, \end{cases} \tag{3.45}$$

where $\tau\in[0,T)$, $v\in U$, and $A_\theta=[\tau,\tau+\theta)$ so that $\theta>0$ and $\tau+\theta\leq T$. Denote by $\chi_{A_\theta}(\cdot)$ the characteristic function of the set A_θ. Then we have $v(\cdot)=u(\cdot)-\overline{u}(\cdot)=(v-\overline{u}(\cdot))\,\chi_{A_\theta}$.

The following theorem constitutes the main contribution of the result

Theorem 3.3.1

Let (A1)-(A2) hold. If $\overline{u}(\cdot)$ is a singular optimal control in the classical sense for the stochastic control (3.2)-(3.3), then for any $v\in U$, it holds that

$$\mathbb{E}\left(\mathbb{H}(\tau)b_u(\tau)(v-\overline{u}(\tau))^2\right)+\partial_\tau^+\left(\mathbb{H}(\tau)(v-\overline{u}(\tau))^2\sigma_u(\tau)\right)\leq 0 \quad a.e.\ \tau\in[0,T], \tag{3.46}$$

where

$$\begin{aligned}&\partial_\tau^+\left(\mathbb{H}(\tau)(v-\overline{u}(\tau))^2\sigma_u(\tau)\right)\\&:=2\lim_{\theta\to0^+}\sup\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\int_\tau^t\left[\phi_v(s,t)\Phi\left(\tau\right)\Psi\left(s\right)\sigma_u(s)(v-\overline{u}(s))\right]dsdt,\end{aligned} \tag{3.47}$$

$\phi_v(\cdot,t)$ is determined by (3.43), $\Phi(\cdot)$ is given by the following process

$$\Phi\left(t\right)=\Phi\left(0\right)+\int_0^t b_x(s)\Phi\left(s\right)ds+\int_0^t\sigma_x(s)\Phi\left(s\right)dW(s)$$

and $\Psi(\cdot)$ is determined by

$$\begin{aligned}\Psi(t)=\Psi(0)&+\int_0^t\left[\sigma_x^2(s)\Psi(s)-b_x(s)\Psi(s)\right]ds\\&-\int_0^t\left[\sigma_x(s)\Psi(s)\right]dW(s).\end{aligned}$$

Proof: From (3.42), we have $v(\cdot)=u(\cdot)-\overline{u}(\cdot)=(v-\overline{u}(\cdot))\chi_{A_\theta}(\cdot)$ and the corresponding solution $y_1(\cdot)$ to (3.5) is given by

$$\begin{aligned} y_1(t) &= \Phi(t)\int_0^t \Psi(s)\left(b_u(s)-\sigma_x(s)\sigma_u(s)\right)(v-\overline{u}(s))\chi_{A_\theta}(s)ds \\ &\quad + \Phi(t)\int_0^t \Psi(s)\,\sigma_u(s)(v-\overline{u}(s))\chi_{A_\theta}(s)dW(s). \end{aligned} \tag{3.48}$$

Substituting $v(\cdot)=(v-\overline{u}(\cdot))\chi_{A_\theta}(\cdot)$ and (3.48) into (3.34), we have

$$\begin{aligned} 0 &\geq \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left[\mathbb{H}(t)\,y_1(t)\,(v-\overline{u}(t))\right]dt \\ &= \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left[\mathbb{H}(t)\,\Phi(t)\int_\tau^t \Psi(s)\left(b_u(s)-\sigma_x(s)\sigma_u(s)\right)(v-\overline{u}(s))ds\,(v-\overline{u}(t))\right]dt \\ &+ \frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left[\mathbb{H}(t)\,\Phi(t)\int_\tau^t \Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt \\ &= J_1^\theta + J_2^\theta. \end{aligned} \tag{3.49}$$

From [[58, Lemma 4.1], we obtain

$$\begin{aligned} \lim_{\theta\to0^+} J_1^\theta &= \lim_{\theta\to0^+}\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left[\mathbb{H}(t)\,\Phi(t)\int_\tau^t \Psi(s)\left(b_u(s)-\sigma_x(s)\sigma_u(s)\right)(v-\overline{u}(s))ds\,(v-\overline{u}(t))\right]dt \\ &= \frac{1}{2}\mathbb{E}\left(\mathbb{H}(\tau)\left(b_u(\tau)-\sigma_x(\tau)\sigma_u(\tau)\right)(v-\overline{u}(\tau))^2\right). \end{aligned} \tag{3.50}$$

On the other hand, by (3.35), it follows that

$$\begin{aligned} J_2^\theta &= \frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\,\Phi(t)\int_\tau^t \Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt \\ &= \frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\,\Phi(\tau)\int_\tau^t \Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt \\ &+ \frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\int_\tau^t b_x(s)\Phi(s)\,ds\right. \\ &\qquad \left.\times\int_\tau^t \Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt \\ &+ \frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\int_\tau^t \sigma_x(s)\Phi(s)\,dW(s)\right. \\ &\qquad \left.\times\int_\tau^t \Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt \\ &\qquad \left.\times\int_\tau^t \Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right\}dt. \\ &= J_{2,1}^\theta + J_{2,2}^\theta + J_{2,3}^\theta + J_{2,4}^\theta. \end{aligned} \tag{3.51}$$

By Lemma 3.3.2 , we get

$$
\begin{aligned}
&\lim_{\theta\to 0^+}\sup J_{2,1}^{\theta}\\
&=\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_{\tau}^{\tau+\theta}\mathbb{E}\left[\mathbb{H}\left(t\right)\Phi\left(\tau\right)\int_{\tau}^{t}\Psi\left(s\right)\sigma_u(s)(v-\overline{u}(s))dW(s)\left(v-\overline{u}(t)\right)\right]dt\\
&=\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_{\tau}^{\tau+\theta}\mathbb{E}\left[\int_{\tau}^{t}\Phi\left(\tau\right)\Psi\left(s\right)\sigma_u(s)(v-\overline{u}(s))dW(s)\mathbb{E}\left[\mathbb{H}\left(t\right)\left(v-\overline{u}(t)\right)\right]\right]dt\\
&+\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_{\tau}^{\tau+\theta}\mathbb{E}\left\{\int_{\tau}^{t}\Phi\left(\tau\right)\Psi\left(s\right)\sigma_u(s)(v-\overline{u}(s))dW(s)\int_{0}^{t}\phi_v(s,t)dW(s)\right\}dt\\
&=\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_{\tau}^{\tau+\theta}\int_{\tau}^{t}\mathbb{E}\left\{\Phi\left(\tau\right)\Psi\left(s\right)\sigma_u(s)(v-\overline{u}(s))\phi_v(s,t)\right\}dsdt \qquad (3.52)\\
&=\frac{1}{2}\partial_{\tau}^{+}\left(\mathbb{H}\left(\tau\right)\left(v-\overline{u}(\tau)\right)^2\sigma_u(\tau)\right),\qquad\forall\tau\in[0,T).
\end{aligned}
$$

It is crucial that, by the Martingale Representation Theorem in Lemma 3.3.2 , we only know that $\phi_v(\cdot,t)\in\mathbb{L}_{\mathbb{F}}^2\left([0,t];\mathbb{R}\right)$ for any $v\in U$, and hence, for each $\tau\in[0,T]$, the function

$$
\varphi_t\left(s\right)=\mathbb{E}\left[\Phi\left(\tau\right)\Psi\left(s\right)\sigma_u(s)(v-\overline{u}(s))\phi_v(s,t)\right],\quad s\in\left[0,t\right],\ t\in\left[0,T\right],
$$

is in $\mathbb{L}^1\left([0,t];\mathbb{R}\right)$. See [58] for more details for the superior limit

$$
\lim_{\theta\to 0^+}\frac{1}{\theta^2}\int_{\tau}^{\tau+\theta}\int_{\tau}^{t}\varphi_t\left(s\right)dsdt.
$$

By simple computations, the last term in (3.51) is in fact a process with zero expectation. Now, by using similar method in [58], we get

$$
\begin{aligned}
\lim_{\theta\to 0^+}J_{2,2}^{\theta}&=\lim_{\theta\to 0^+}\frac{1}{\theta^2}\int_{\tau}^{\tau+\theta}\mathbb{E}\left\{\mathbb{H}\left(t\right)\int_{\tau}^{t}b_x(s)\Phi\left(s\right)ds\right. \qquad (3.53)\\
&\quad\times\left.\int_{\tau}^{t}\Psi\left(s\right)\sigma_u(s)(v-\overline{u}(s))dW(s)\left(v-\overline{u}(t)\right)\right\}dt\\
&=0,
\end{aligned}
$$

$$
\begin{aligned}
\lim_{\theta\to 0^+}J_{2,3}^{\theta}&=\lim_{\theta\to 0^+}\frac{1}{\theta^2}\int_{\tau}^{\tau+\theta}\mathbb{E}\left\{\mathbb{H}\left(t\right)\int_{\tau}^{t}\sigma_x(s)\Phi\left(s\right)dW(s)\right. \qquad (3.54)\\
&\quad\times\left.\int_{\tau}^{t}\Psi\left(s\right)\sigma_u(s)(v-\overline{u}(s))dW(s)\left(v-\overline{u}(t)\right)\right\}dt\\
&=\frac{1}{2}\mathbb{E}\left(\mathbb{H}\left(\tau\right)\left(\sigma_x(\tau)\sigma_u(\tau)\right)\left(v-\overline{u}(\tau)\right)^2\right),
\end{aligned}
$$

and Finally, substituting (3.49), (3.51), (3.52), (3.53), (3.54) in (3.48), we obtain

$$
\mathbb{E}\left(\mathbb{H}(\tau)b_u(\tau)(v-\overline{u}(\tau))^2\right)+\partial_{\tau}^{+}\left(\mathbb{H}(\tau)(v-\overline{u}(\tau))^2\sigma_u(\tau)\right)\leq 0,\quad a.e.\ \ \tau\in[0,T].
$$

This completes the proof of Theorem 3.3.1 ■

Chapter 4

Pointwise second order necessary conditions for stochastic optimal control with jump diffusions

4.1 Preliminaries and assumptions

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathcal{P})$ be a complete probability space with filtration, we assume that $\mathbb{F} = \{\mathcal{F}_t\}_{0\leq t\leq T}$ is the natural filtration generated by one-dimensional standard Brownian motion $W(\cdot)$ and an independent Poisson random measure N on $\mathbb{R}_+ \times Z$, where Z is a fixed nonempty subset of $\mathbb{R}$ with its Borel σ-field $\mathcal{B}(Z)$ such that $\mu(Z) < \infty$. We denote by $\{\mathcal{F}_t^W\}_{0\leq t\leq T}$ (resp. $\{\mathcal{F}_t^N\}_{0\leq t\leq T}$) the $\mathcal{P}$-augmentation of the natural filtration of W (resp. N), then we have

$$\mathcal{F}_t = \sigma\left\{W(s); s \leq t\right\} \vee \sigma\left\{\int\int_{(0,s]\times A} N(dz, dr); s \leq t, A \in \mathcal{B}(Z)\right\} \vee \mathcal{N},$$

where $\mathcal{N}$ denotes the totality of $\mathcal{P}$-null sets, and $\sigma_1 \vee \sigma_2$ denotes the σ-filed generated by $\sigma_1 \cup \sigma_2$. We assume that the compensator of N has the form $\mu(dt, dz) = \mu(dz)\, dt$ for some positive and σ-finite Lévy measure μ on Z. We suppose that $\int_Z 1 \wedge |z|^2 \mu(dz) < \infty$ and write $\widetilde{N} = N - \mu dt$ for the compensated jump martingale random measure of N.

For a fuction $\phi : [0, T] \times \mathbb{R} \times U \times \Omega \to \mathbb{R}$, we denote by $\phi_x(t, x, u)$ (resp. $\phi_u(t, x, u)$) the first order derivatives of ϕ with respect to x and u at (t, x, u, ω), and by $\phi_{(x,u)^2}(t, x, u)$ the second order derivatives of ϕ with respect to (x, u) at (t, x, u, ω) and by $\phi_{xx}(t, x, u)$, $\phi_{xu}(t, x, u)$, and $\phi_{uu}(t, x, u)$ the second order derivatives of ϕ at (t, x, u, ω), we denote by $\mathcal{U}_{ad}$ the set of all admissible controls. Note that, we take out the $\omega\,(\in \Omega)$ argument in the defined functions, when the conditions is clear as habitual.

We introduce some spaces of random variable and stochastic processes, for any $t \in$

$[0,T]$, we let

- $L^2_{\mathcal{F}_t}(\Omega;\mathbb{R})$ the space of $\mathbb{R}$-valued, $\mathcal{F}_t$-measurable random variables ζ such that
$$\mathbb{E}\left|\zeta\right|^2 < \infty.$$
- $\mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})$ the space of $\mathbb{R}$-valued, $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable, $\mathbb{F}$-adapted processes ψ such that
$$\|\psi\|_{\mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})} := \left[\mathbb{E}\left(\int_0^T \left|\psi(t)\right|^2 dt\right)\right]^{\frac{1}{2}} < \infty.$$
- $\mathcal{L}^2([0,T];\mathbb{R})$ the space of $\mathbb{R}$-valued, $\mathcal{B}([0,T]\times\Omega)\otimes\mathcal{B}(Z)$ measurable processes ϑ such that
$$\|\vartheta\|_{\mathcal{L}^2([0,T];\mathbb{R})} := \mathbb{E}\left[\int_0^T \left|\vartheta_t(z)\right|^2 \mu(dz)\, dt\right]^{\frac{1}{2}} < \infty.$$

We should note that in stochastic control problems, there is an other type of singularity, where the control variable has two components $(u(\cdot),\xi(\cdot))$, the first being absolutely continuous and the second is of bounded variation, non-decreasing left-continuous with right limits and $\xi(0_-)=0$. This singularity come since $d\xi(t)$ may be singular with respect to Lebesgue measure dt. An extensive list of references on the stochastic singular control problem can be found in Haussmann and Suo [24], and Cadenillas and Haussmann [9].

In this chapter, we study pointwise optimal stochastic control problem for systems governed by nonlinear controlled stochastic differential equations (SDEs) with jumps of the form: $t\in[0,T]$

$$\begin{cases} dx(t) = b(t,x(t),u(t))\,dt + \sigma(t,x(t),u(t))dW(t) + \displaystyle\int_Z \eta(t,x(t_-),z)\,\widetilde{N}(dz,dt), \\ x(0) = x_0, \end{cases} \tag{4.1}$$

where the coefficients of the state are given by the functions

$$b : [0,T]\times\mathbb{R}\times U \to \mathbb{R},$$
$$\sigma : [0,T]\times\mathbb{R}\times U \to \mathbb{R},$$
$$\eta : [0,T]\times\mathbb{R}\times Z \to \mathbb{R}.$$

The cost functional to be minimized has the form:

$$J(u(\cdot)) = \mathbb{E}\left[\int_0^T f\left(t, x\left(t\right), u\left(t\right)\right) dt + h(x\left(T\right))\right],\ u\left(\cdot\right) \in \mathcal{U}_{ad}, \tag{4.2}$$

where $f : [0,T] \times \mathbb{R} \times U \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$ are given functions.

The stochastic optimization problem which we interest is to find a control $\overline{u}\left(\cdot\right) \in \mathcal{U}_{ad}$ such that

$$J(\overline{u}\left(\cdot\right)) = \inf_{u(\cdot)\in\mathcal{U}_{ad}} J\left(u\left(\cdot\right)\right). \tag{4.3}$$

Any admissible control $\overline{u}\left(\cdot\right) \in \mathcal{U}_{ad}$ that achieves the minimum is called an optimal control.

Throughout this paper, we also assume that

Assumptions (B1) 1. The maps b and σ are $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable and $\mathbb{F}$-adapted.
2. The functions b and σ are continuously differentiable up to the second order with respect to (x, u).
3. The function η is continuously differentiable up to the second order with respect to x.
4. All the first order derivatives are continuous in (x, u) and uniformly bounded.
5. There exists a constant $K_1 > 0$ such that for almost all $(t, \omega) \in [0,T] \times \Omega$ and for any $x, \widetilde{x} \in \mathbb{R}$ and $u, \widetilde{u} \in U$,

$$\begin{cases} \left|\phi\left(t,x,u\right)\right| \leq K_1,\ \text{for } \phi = b, \sigma, \\ \left|\phi\left(t,x,u\right) - \phi\left(t,\widetilde{x},u\right)\right| \leq K_1 \left|x-\widetilde{x}\right|,\ \text{for } \phi = b, \sigma, \\ \left|\phi_{(x,u)^2}\left(t,x,u\right) - \phi_{(x,u)^2}\left(t,\widetilde{x},\widetilde{u}\right)\right| \leq K_1 \left(\left|x-\widetilde{x}\right| + \left|u-\widetilde{u}\right|\right), \\ \left|\eta\left(t,x\left(t_-\right),z\right) - \eta\left(t,\widetilde{x}\left(t_-\right),z\right)\right| \leq K_1\left|x-\widetilde{x}\right| \ \text{and } \eta\left(t,x\left(t_-\right),z\right) \leq K_1, \\ \left|\eta_{xx}\left(t,x\left(t_-\right),z\right) - \eta_{xx}\left(t,\widetilde{x}\left(t_-\right),z\right)\right| \leq K_1\left|x-\widetilde{x}\right|. \end{cases}$$

Under assumptions (B1) and (A2), equation (4.1) has a unique strong solution $x(t)$ given by

$$\begin{aligned} x(t) = x_0 &+ \int_0^t b\left(s, x(s), u(s)\right) ds + \int_0^t \sigma\left(s, x(s), u(s)\right) dW(s) \\ &+ \int_0^t \int_Z \eta\left(s, x\left(s_-\right), z\right) \widetilde{N}\left(dz, ds\right), \end{aligned}$$

and by standard arguments it is easy to show that for any $C_k > 0$, it holds that

$$\mathbb{E}(\sup_{t\in[0,T]} \left|x(t)\right|^k) < C_k,$$

where C_k is a constant depending only on k. Moreover, the functional (4.2) is well defined from $\mathcal{U}_{ad}$ into $\mathbb{R}$.

4.2 Second order necessary condition in integral form with jump Diffusions

In this section, we prove an integral type second order necessary condition for stochastic optimal control with jump diffusions. We consider a control region U is nonempty and bounded. Moreover, a convex perturbation of the optimal control defined by $u^{\theta}(t) = \overline{u}(t) + \theta(u(t) - \overline{u}(t))$, for $u(\cdot) \in \mathcal{U}_{ad}$ and $\theta \in (0,1)$. The convexity condition of the control domain ensures that $u^{\theta}(\cdot) \in \mathcal{U}_{ad}$.

For convenience, we will use the following notations, we denote by $x^{\theta}(\cdot)$, $\overline{x}(\cdot)$ the state trajectory of the SDE (4.1) corresponding to $u^{\theta}(\cdot)$ and $\overline{u}(\cdot)$.
To simplify our notation, we write for $\phi = b, \sigma, f$ and for $\varphi = \eta$:

$$\begin{cases} \delta\phi(t) = \phi\left(t, x^{\theta}(t), u^{\theta}(t)\right) - \phi(t, \overline{x}(t), \overline{u}(t)), \\ \phi_x(t) = \phi_x(t, \overline{x}(t), \overline{u}(t)), \qquad \phi_u(t) = \phi_u(t, \overline{x}(t), \overline{u}(t)), \\ \phi_{xx}(t) = \phi_{xx}(t, \overline{x}(t), \overline{u}(t)), \quad \phi_{uu}(t) = \phi_{uu}(t, \overline{x}(t), \overline{u}(t)), \\ \varphi_x(t,z) = \varphi_x(t, \overline{x}(t_-), z), \qquad \phi_{xu}(t) = \phi_{xu}(t, \overline{x}(t), \overline{u}(t)), \\ \varphi_{xx}(t,z) = \varphi_{xx}(t, \overline{x}(t_-), z). \end{cases}$$

We introduce the following variational equations

$$\begin{cases} dy_1(t) = \{b_x(t) y_1(t) + b_u(t) v(t)\} dt + \{\sigma_x(t) y_1(t) + \sigma_u(t) v(t)\} dW(t) \\ \qquad\qquad + \int_Z \{\eta_x(t,z) y_1(t_-)\} \widetilde{N}(dz, dt), \\ y_1(0) = 0, \end{cases} \tag{4.4}$$

and

$$\begin{cases} dy_2(t) = [b_x(t)\,y_2(t) + b_{xx}(t)\,y_1(t)^2 + 2b_{xu}(t)\,y_1(t)\,v(t) + b_{uu}(t)\,v(t)^2]dt \\ \qquad +[\sigma_x(t)\,y_2(t) + \sigma_{xx}(t)\,y_1(t)^2 + 2\sigma_{xu}(t)\,y_1(t)\,v(t) + \sigma_{uu}(t)\,v(t)^2]dW(t) \\ \qquad + \int_Z [\eta_x(t,z)\,y_2(t_-) + \eta_{xx}(t,z)\,y_1(t_-)^2]\widetilde{N}(dz,dt)\,, \\ y_2(0) = 0. \end{cases} \tag{4.5}$$

Remark 4.2.1

Under assumptions (B1), (A2) the variational equations (4.4) and (4.5) admits a unique strong solutions $y_1(t)$ and $y_2(t)$ respectively.

Next, we prove the proposition which plays a crucial role in obtaining a second order necessary conditions.

We note that unless specified, for each $k \in \mathbb{R}_+$, we will denote by $C_k > 0$ a generic positive constant depending only on k and the constants appearing in Proposition 4.2.1 , which may vary from line to line.

Proposition 4.2.1

Assume that assumptions (B1), (A2) satisfied. Then, for any $k \geq 1$, we have the following basic estimates

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|x^\theta(t) - \overline{x}(t)\right|^{2k}\right] \leq C_{(k,\mu(Z))}\theta^k, \tag{4.6}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|y_1(t)\right|^{2k}\right] \leq C_{(k,\mu(Z))}, \tag{4.7}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|y_2(t)\right|^{2k}\right] \leq C_{(k,\mu(Z))}, \tag{4.8}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|x^\theta(t) - \overline{x}(t) - \theta y_1(t)\right|^{2k}\right] \leq C_{(k,\mu(Z))}\theta^{2k}, \tag{4.9}$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|x^\theta(t) - \overline{x}(t) - \theta y_1(t) - \frac{\theta^2}{2}y_2(t)\right|^{2k}\right] \leq C_{(k,\mu(Z))}\theta^{2k}. \tag{4.10}$$

Proof: Let $\overline{x}(\cdot)$ and $x^\theta(\cdot)$ be the trajectory of (3.2) corresponding to $\overline{u}(\cdot)$ and $u^\theta(\cdot)$ resp.

Let $y_1(\cdot)$ and $y_2(\cdot)$ be the solution of first and second order adjoint equations (4.4)-(4.5).

Noting that estimate (4.6) follows from standard arguments, using *Burkholder-Davis-Gundy inequality* for the martingale part and Propositions A2 (see *Appendix*). In what follows we shall refer to equation (4.4) as the first-order variational equation, and the process $y_1(\cdot)$ is called the *first order variational process.* A very important step in Peng [41], and Tang and Li [50] is in light of the Taylor expansion, to find a process $y_2(t)$ so that $x^\theta(t) - \overline{x}(t) - \theta y_1(t) - \dfrac{\theta^2}{2} y_2(t) = o\left(\theta^2\right)$, as $\theta \to 0$, and that the convergence is of an appropriate order. The process $y_2(\cdot)$ is called the *second-order variational process.* So the estimates (4.7), (4.8) and (4.9) are obvious and standard, see also [50, Lemma 2.1].

Now, we start to prove the estimate (4.10). From (3.2), (4.4) and (4.5), we have

$$
\begin{aligned}
&\left| x^\theta(t) - \overline{x}(t) - \theta y_1(t) - \frac{\theta^2}{2} y_2(t) \right|^{2k} \\
&= \Big| \int_0^t [\delta b\,(s) - \theta\,[b_x(s) y_1(s) + b_u(s) v(s)] \\
&- \frac{\theta^2}{2} \left[b_x(s) y_2(s) + b_{xx}(s) y_1(s)^2 + 2 b_{xu}(s) y_1(s) v(s) + b_{uu}(s) v(s)^2 \right]] ds \\
&+ \int_0^t [\delta\sigma\,(s) - \theta\,[\sigma_x(s) y_1(s) + \sigma_u(s) v(s)] \\
&- \frac{\theta^2}{2} \left[\sigma_x(s) y_2(s) + \sigma_{xx}(s) y_1(s)^2 + 2\sigma_{xu}(s) y_1(s) v(s) \right. \\
&\left. + \sigma_{uu}(s) v(s)^2 \right]] dW(s) \\
&+ \int_0^t \int_Z \left\{ \eta\left(s, x^\theta(s_-), z\right) - \eta\left(s, \overline{x}\left(s_-\right), z\right) - \theta\left[\eta_x\left(s, z\right) y_1\left(s_-\right)\right] \right. \\
&\left. - \frac{\theta^2}{2} \left[\eta_x\left(s, z\right) y_2\left(s_-\right) + \eta_{xx}\left(s, z\right) y_1(s_-)^2 \right] \right\} \widetilde{N}\left(dz, ds\right) \Big|^{2k}.
\end{aligned}
$$

A straightforward calculation by applying the Cauchy-Schwarz inequality, we shows that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} \left| x^\theta(t) - \overline{x}(t) - \theta y_1(t) - \frac{\theta^2}{2} y_2(t) \right|^{2k} \right] \qquad (4.11) \\
&\leq I_1 + I_2,
\end{aligned}
$$

where

$$\begin{aligned} I_1 = \mathbb{E}\Bigg[2 \sup_{t\in[0,T]} \Bigg| \int_0^t & \left[\delta b\left(s\right) - \theta\left[b_x(s)y_1(s) + b_u(s)v(s)\right]\right. \\ & - \frac{\theta^2}{2}\left[b_x(s)y_2(s) + b_{xx}(s)y_1(s)^2 + 2b_{xu}(s)y_1(s)v(s) + b_{uu}(s)v(s)^2\right]\Big] ds \\ & + \int_0^t \left[\delta\sigma\left(s\right) - \theta\left[\sigma_x(s)y_1(s) + \sigma_u(s)v(s)\right]\right. \\ & - \frac{\theta^2}{2}\left[\sigma_x(s)y_2(s) + \sigma_{xx}(s)y_1(s)^2 + 2\sigma_{xu}(s)y_1(s)v(s) + \sigma_{uu}(s)v(s)^2\right]\Big] dW(s)\Bigg|^{2k}\Bigg], \end{aligned} \tag{4.12}$$

and

$$\begin{aligned} I_2 = \mathbb{E}\Bigg[2 \sup_{t\in[0,T]} \Bigg| \int_0^t \int_Z & \Big\{\eta\left(s, x^\theta\left(s_-\right), z\right) - \eta\left(s, \overline{x}\left(s_-\right), z\right) \\ & - \theta\left[\eta_x\left(s,z\right) y_1\left(s_-\right)\right] - \frac{\theta^2}{2}\left[\eta_x\left(s,z\right) y_2\left(s_-\right) + \eta_{xx}\left(s,z\right) y_1(s_-)^2\right]\Big\} \widetilde{N}\left(dz, ds\right)\Bigg|^{2k}\Bigg]. \end{aligned} \tag{4.13}$$

Similar to Bonnans [7], Zhang and Zhang [58], by applying the Cauchy-Schwarz inequality and the Burkhölder–Davis–Gundy inequality, we have

$$I_1 \leq C_k \theta^{2k}. \tag{4.14}$$

Let us turn to estimate of I_2. By using Proposition 1.3.1, then for all $k \geq 1$ there exists a positive constant $C_{(k,\mu(Z))}$ such that

$$\begin{aligned} I_2 \leq C_{(k,\mu(Z))}\mathbb{E}\Bigg[\Bigg|\int_0^T \int_Z & \Big\{\eta\left(s, x^\theta\left(s\right), z\right) - \eta\left(s, \overline{x}\left(s\right), z\right) \\ & - \theta\left[\eta_x\left(s,z\right) y_1\left(s\right)\right] - \frac{\theta^2}{2}\left[\eta_x\left(s,z\right) y_2\left(s\right) + \eta_{xx}\left(s,z\right) y_1(s)^2\right]\Big\}\Bigg|^{2k} \mu\left(dz\right) ds\Bigg]. \end{aligned}$$

Then we have

$$\begin{aligned} I_2 \leq C_{(k,\mu(Z))}\mathbb{E}\Bigg[\Bigg|\mu(Z) \int_0^T \sup_{z\in Z} & \Big\{\eta\left(s, x^\theta\left(s\right), z\right) - \eta\left(s, \overline{x}\left(s\right), z\right) \\ & - \theta\left[\eta_x\left(s,z\right) y_1\left(s\right)\right] - \frac{\theta^2}{2}\left[\eta_x\left(s,z\right) y_2\left(s\right) + \eta_{xx}\left(s,z\right) y_1(s)^2\right]\Big\}\Bigg|^{2k} ds\Bigg]. \end{aligned}$$

Now, applying similar method developed in I_1 for deterministic integral, we get

$$I_2 \leq C_{(k,\mu(Z))}\theta^{2k}. \tag{4.15}$$

By combining (4.11), (4.14) and (4.15), the desired result (4.10) is fulfilled. Thus, the proof of Proposition 4.2.1 is completed. ■

Define the Hamiltonian function $H : [0,T] \times \mathbb{R} \times U \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ by

$$H(t,x,u,p,q,r) := b(t,x,u)p + \sigma(t,x,u)q - f(t,x,u) + \int_Z \eta(t,x,z)r(t,z)\mu(dz).$$

Now, we introduce the first adjoint equation

$$\begin{cases} dp(t) = -\left\{ b_x(t)p(t) + \sigma_x(t)q(t) - f_x(t) + \int_Z \eta_x(t,z)r(t,z)\mu(dz) \right\} dt \\ \qquad\qquad +q(t)\,dW(t) + \int_Z r(t,z)\widetilde{N}(dz,dt), \\ p(T) = -h_x(\overline{x}(T)), \end{cases} \tag{4.16}$$

and the second adjoint equation

$$\begin{cases} dP(t) = -\Big\{ 2b_x(t)P(t) + \sigma_x(t)^2 P(t) + 2\sigma_x(t)Q(t) \\ \qquad + \int_Z \left[\eta_x(t,z)^2 P(t) + \eta_x(t,z)^2 R(t,z) + 2\eta_x(t,z)R(t,z) \right] \mu(dz) + H_{xx}(t) \Big\} dt \\ \qquad +Q(t)\,dW(t) + \int_Z R(t,z)\widetilde{N}(dz,dt), \\ P(T) = -h_{xx}(\overline{x}(T)), \end{cases} \tag{4.17}$$

where

$$H_{xx}(t) = b_{xx}(t)p + \sigma_{xx}(t)q - f_{xx}(t) + \int_Z \eta_{xx}(t,z)r(t,z)\mu(dz).$$

It is easy to prove that under assumptions (B1)-(A2), Eqs-(4.16) and (4.17) are classical linear backward stochastic differential equations (BSDEs in short) admit a unique strong solution such that

$$(p(t), q(t), r(t,\cdot)) \in \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathcal{L}^2([0,T];\mathbb{R})$$
$$(P(t), Q(t), R(t,\cdot)) \in \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R}) \times \mathcal{L}^2([0,T];\mathbb{R})$$

Also, we define the functional $\mathbb{H} : [0,T] \times \mathbb{R} \times U \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ by

$$\mathbb{H}(t,x,u,p,q,r,P,Q,R) := H_{xu}(t,x,u,p,q,r) + b_u(t,x,u)P(t) + \sigma_u(t,x,u)Q(t) + \sigma_u(t,x,u)P(t)\sigma_x(t,x,u) \tag{4.18}$$

To simplify our notation, we set

$$\mathbb{H}(t) = \mathbb{H}(t, \overline{x}(t), \overline{u}(t), p(t), q(t), r(t,\cdot), P(t), Q(t), R(t,\cdot)),\ t \in [0,T].$$

Lemma 4.2.1

Let $(p(t), q(t), r(t,\cdot))$ be the solution of the adjoint equation (4.16),$(P(t), Q(t), R(t,\cdot))$ be the solution of adjoint equation (4.17), and $y_1(t)$, $y_2(t)$ are the solution to the first and second variational equations (4.4) and (4.5) respectively. Then the following duality relations hold:

$$-\mathbb{E}[p(T) y_1(T)] = -\mathbb{E}\left[\int_0^T \{p(t)(b_u(t) v(t)) + q(t)(\sigma_u(t) v(t))\} dt\right] - \mathbb{E}\left[\int_0^T f_x(t) y_1(t) dt\right], \tag{4.19}$$

$$\begin{aligned} -\mathbb{E}[p(T) y_2(T)] = & -\mathbb{E}\left[\int_0^T p(t)\left\{b_{xx}(t) y_1(t)^2 + 2b_{xu}(t) y_1(t) v(t) + b_{uu}(t) v(t)^2\right\} dt\right] \\ & - \mathbb{E}\left[\int_0^T q(t)\left\{\sigma_{xx}(t) y_1(t)^2 + 2\sigma_{xu}(t) y_1(t) v(t) + \sigma_{uu}(t) v(t)^2\right\} dt\right] \\ & - \mathbb{E}\left[\int_0^T \int_Z r(t,z)\left\{\eta_{xx}(t,z) y_1(t)^2\right\} \mu(dz) dt\right] - \mathbb{E}\left[\int_0^T f_x(t) y_2(t) dt\right], \end{aligned} \tag{4.20}$$

and

$$\begin{aligned} -\mathbb{E}\left[P(T) y_1(T)^2\right] = & -2\mathbb{E}\left[\int_0^T \{P(t) y_1(t)(b_u(t) v(t)) + P(t)\sigma_x(t) y_1(t)(\sigma_u(t) v(t))\} dt\right] \\ & - 2\mathbb{E}\left[\int_0^T \{Q(t)\sigma_u(t) v(t) y_1(t)\} dt\right] - \mathbb{E}\left[\int_0^T P(t)(\sigma_u(t) v(t))^2 dt\right] \\ & + \mathbb{E}\left[\int_0^T H_{xx}(t) y_1(t)^2 dt\right]. \end{aligned} \tag{4.21}$$

Proof: The proof of this lemma follows directly from Itô's formula to $p(t) y_1(t)$ and taking expectation where $y_1(0) = 0$, we have

$$-\mathbb{E}\left[p(T)\, y_1(T)\right] = -\mathbb{E}\int_0^T p(t)\, dy_1(t) - \mathbb{E}\int_0^T y_1(t)\, dp(t) \\ - \mathbb{E}\int_0^T q(t)\left\{\sigma_x(t)\, y_1(t) + \sigma_u(t)\, v(t)\right\} dt \\ - \mathbb{E}\int_0^T \int_Z r(t,z)\left\{\eta_x(t,z)\, y_1(t)\right\} \mu(dz)\, dt, \tag{4.22}$$

where

$$-\mathbb{E}\int_0^T p(t)\, dy_1(t) = -\mathbb{E}\int_0^T p(t)\left[b_x(t)\, y_1(t) + b_u(t)\, v(t)\right] dt. \tag{4.23}$$

Consequently,

$$-\mathbb{E}\int_0^T y_1(t)\, dp(t) \tag{4.24}$$
$$= \mathbb{E}\int_0^T y_1(t)\left[b_x(t)\, p(t) + \sigma_x(t)\, q(t) - f_x(t) + \int_Z \eta_x(t,z)\, r(t,z)\, \mu(dz)\right] dt,$$

substituting (4.23), (4.24), into (4.22), then the desired result (4.19) is fulfilled.

Now, applying Itô's formula to $p(t)\, y_2(t)$ and taking expectation where $y_2(0) = 0$, we have

$$-\mathbb{E}\left[p(T)\, y_2(T)\right] = -\mathbb{E}\int_0^T p(t)\, dy_2(t) - \mathbb{E}\int_0^T y_2(t)\, dp(t) \\ - \mathbb{E}\left[\int_0^T q(t)\left\{\sigma_x(t)\, y_2(t) + \sigma_{xx}(t)\, y_1(t)^2 \right.\right. \\ \left.\left. + \ 2\sigma_{xu}(t)\, y_1(t)\, v(t) + \sigma_{uu}(t)\, v(t)^2\right\} dt\right] \\ - \mathbb{E}\left[\int_0^T \int_Z r(t,z)\left[\eta_x(t,z)\, y_2(t) + \eta_{xx}(t,z)\, y_1(t)^2\right] \mu(dz)\, dt\right], \tag{4.25}$$

where

$$-\mathbb{E}\int_0^T p(t)\, dy_2(t) = -\mathbb{E}\int_0^T p(t)\left\{b_x(t)\, y_2(t) + b_{xx}(t)\, y_1(t)^2 \right. \\ \left. + \ 2b_{xu}(t)\, y_1(t)\, v(t) + b_{uu}(t)\, v(t)^2\right\} dt, \tag{4.26}$$

and

$$-\mathbb{E}\int_0^T y_2(t)\, dp(t) \tag{4.27}$$
$$= \mathbb{E}\int_0^T y_2(t)\left[b_x(t)\, p(t) + \sigma_x(t)\, q(t) - f_x(t) + \int_Z \eta_x(t,z)\, r(t,z)\, \mu(dz)\right] dt,$$

substituting (4.26), (4.27), into (4.25), we obtain the desired result (4.20).

Next, applying Itô's formula to $P(t)\,y_1(t)$, where $y_1(0)=0$, we have

$$\begin{aligned}
\left[P(T)\,y_1(T)\right] &= \int_0^T P(t)\,dy_1(t) + \int_0^T y_1(t)\,dP(t) \\
&+ \int_0^T Q(t)\left\{\sigma_x(t)\,y_1(t)+\sigma_u(t)\,v(t)\right\}dt \\
&+ \int_0^T\int_Z R(t,z)\left\{\eta_x(t,z)\,y_1(t_-)\right\}N(dz,dt) \\
&= I_1+I_2+I_3+I_4, \qquad (4.28)
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \int_0^T P(t)\,dy_1(t) \\
&= \int_0^T P(t)\left\{b_x(t)\,y_1(t)+b_u(t)\,v(t)\right\}dt \\
&+ \int_0^T P(t)\left\{\sigma_x(t)\,y_1(t)+\sigma_u(t)\,v(t)\right\}dW(t) \\
&+ \int_0^T\int_Z P(t)\left\{\eta_x(t,z)\,y_1(t_-)\right\}\widetilde{N}(dz,dt),
\end{aligned}$$

by simple computations , we can prove

$$\begin{aligned}
I_2 &= \int_0^T y_1(t)\,dP(t) \\
&= -\int_0^T y_1(t)\Big\{2b_x(t)\,P(t)+\sigma_x(t)^2\,P(t)+2\sigma_x(t)\,Q(t) \\
&+ \int_Z\Big[\eta_x(t,z)^2\,P(t)+\eta_x(t,z)^2\,R(t,z)+2\eta_x(t,z)\,R(t,z)\Big]\mu(dz)+H_{xx}(t)\Big\}dt \\
&+ \int_0^T y_1(t)\,Q(t)\,dW(t)+\int_0^T\int_Z y_1(t_-)\,R(t,z)\,\widetilde{N}(dz,dt),
\end{aligned}$$

$$I_3 = \int_0^T\left\{Q(t)\,\sigma_x(t)\,y_1(t)+Q(t)\,\sigma_u(t)\,v(t)\right\}dt,$$

and

$$\begin{aligned}
I_4 &= \int_0^T\int_Z R(t,z)\left\{\eta_x(t,z)\,y_1(t_-)\right\}N(dt,dz) \\
&= \int_0^T\int_Z R(t,z)\left\{\eta_x(t,z)\,y_1(t_-)\right\}\widetilde{N}(dz,dt) \\
&+ \int_0^T\int_Z R(t,z)\left\{\eta_x(t,z)\,y_1(t)\right\}\mu(dz).
\end{aligned}$$

Then we can write (4.28) as follows

$$
\begin{aligned}
\left[P(T)\, y_1(T)\right] = & \int_0^T \left[P(t)\, b_u(t)\, v(t)\, dt + Q(t)\, \sigma_u(t)\, v(t)\right. \\
& - y_1(t)\, b_x(t)\, P(t) - y_1(t)\, \sigma_x(t)^2 P(t) - y_1(t)\, \sigma_x(t)\, Q(t) \\
& - \int_Z y_1(t)\, \eta_x(t,z)^2 \left(P(t) + R(t,z)\right) \mu(dz) \\
& \left. - \int_Z y_1(t)\, \eta_x(t,z)\, R(t,z)\, \mu(dz) - y_1(t)\, H_{xx}(t)\right] dt \\
& + \int_0^T \left[P(t)\, \sigma_x(t)\, y_1(t) + P(t)\, \sigma_u(t)\, v(t) + y_1(t)\, Q(t)\right] dW(t) \\
& + \int_0^T \int_Z \left[P(t)\, \eta_x(t,z)\, y_1(t_-) + y_1(t_-)\, R(t,z)\right. \\
& \left. + R(t,z)\, \eta_x(t,z)\, y_1(t_-)\right] \widetilde{N}(dz,dt). \qquad (4.29)
\end{aligned}
$$

Now, we applying Itô's formula to $(P(T)\, y_1(T))\, y_1(T)$ and taking expectation, we obtain

$$
\begin{aligned}
& - \mathbb{E}\left[P(T)\, y_1(T)^2\right] \\
& = -\mathbb{E}\int_0^T P(t)\, y_1(t)\, dy_1(t) - \mathbb{E}\int_0^T y_1(t)\, d\left(P(t)\, y_1(t)\right) \\
& - \mathbb{E}\left[\int_0^T \left\{\sigma_x(t)\, y_1(t) + \sigma_u(t)\, v(t)\right\}\left\{P(t)\, \sigma_x(t)\, y_1(t) + P(t)\, \sigma_u(t)\, v(t) + y_1(t)\, Q(t)\right\} dt\right] \\
& - \mathbb{E}\left[\int_0^T \int_Z \left\{\eta_x(t,z)\, y_1(t)\right\}\left\{P(t)\, \eta_x(t,z)\, y_1(t) + y_1(t)\, R(t,z) + R(t,z)\, \eta_x(t,z)\, y_1(t)\right\} \mu(dz)\, dt\right] \\
& = J_1 + J_2 + J_3 + J_4, \qquad (4.30)
\end{aligned}
$$

where

$$
J_1 = -\mathbb{E}\int_0^T P(t)\, y_1(t)\, dy_1(t) = -\mathbb{E}\int_0^T P(t)\, y_1(t)\left\{b_x(t)\, y_1(t) + b_u(t)\, v(t)\right\} dt, \quad (4.31)
$$

$$
\begin{aligned}
J_2 & = -\mathbb{E}\int_0^T y_1(t)\, d\left(P(t)\, y_1(t)\right) \\
& = -\mathbb{E}\int_0^T y_1(t)\left[P(t)\, b_u(t)\, v(t)\, dt + Q(t)\, \sigma_u(t)\, v(t)\right. \\
& - y_1(t)\, b_x(t)\, P(t) - y_1(t)\, \sigma_x(t)^2 P(t) - y_1(t)\, \sigma_x(t)\, Q(t) \qquad (4.32) \\
& - \int_Z y_1(t)\, \eta_x(t,z)^2 \left(P(t) + R(t,z)\right) \mu(dz) \\
& \left. - \int_Z y_1(t)\, \eta_x(t,z)\, R(t,z)\, \mu(dz) - y_1(t)\, H_{xx}(t)\right] dt, \qquad (4.33)
\end{aligned}
$$

and it is easy to show that

$$J_3 = -\mathbb{E}\left[\int_0^T \{\sigma_x(t)\, y_1(t) + \sigma_u(t)\, v(t)\}\{P(t)\,\sigma_x(t)\, y_1(t) + P(t)\,\sigma_u(t)\, v(t) + y_1(t)\, Q(t)\}\, dt\right]$$
$$= -\mathbb{E}\left[\int_0^T \Big\{P(t)\,(\sigma_x(t)\, y_1(t))^2 + 2P(t)\,\sigma_x(t)\, y_1(t)\,\sigma_u(t)\, v(t) + \sigma_x(t)\, Q(t)\, y_1(t)^2\right.$$
$$\left. +\ P(t)\,(\sigma_u(t)\, v(t))^2 + Q(t)\,\sigma_u(t)\, y_1(t)\, v(t)\Big\}\, dt\right]. \qquad (4.34)$$

Similarly, we have

$$J_4 = -\mathbb{E}\left[\int_0^T \int_Z \{\eta_x(t,z)\, y_1(t)\}\{P(t)\,\eta_x(t,z)\, y_1(t)\right.$$
$$+\ y_1(t)\, R(t,z) + R(t,z)\,\eta_x(t,z)\, y_1(t)\}\,\mu(dz)\, dt]$$
$$= -\mathbb{E}\left[\int_0^T \int_Z \Big\{(P(t) + R(t,z))\,(\eta_x(t,z)\, y_1(t))^2 + \eta_x(t,z)\, R(t,z)\, y_1(t)^2\Big\}\,\mu(dz)\, dt\right]. \qquad (4.35)$$

Finally, substituting (4.31), (4.33), (4.34), (4.35), into (4.30), then (4.21) is fulfilled.

This completes the proof of Lemma 4.2.1. ∎

To prove the main theorem we need the following technical result.

Proposition 4.2.2

Let (B1)-(A2) hold. Then, for any $u(\cdot) \in \mathcal{U}_{ad}$ we have

$$J(u^\theta(\cdot)) - J(\overline{u}(\cdot))$$
$$= -\mathbb{E}\int_0^T \left[\theta\,\{H_u(t)\, v(t)\} + \frac{\theta^2}{2}\Big\{H_{uu}(t)\, v(t)^2\Big\}\right.$$
$$\left. + \frac{\theta^2}{2}\Big\{P(t)\,(\sigma_u(t)\, v(t))^2\Big\} + \theta^2\,\{\mathbb{H}(t)\, y_1(t) v(t)\}\right] dt + o\left(\theta^2\right), \quad \left(\theta \to 0^+\right), \qquad (4.36)$$

where

$$H_u(t) = H_u(t, \overline{x}(t), \overline{u}(t), p(t), q(t), r(t,\cdot)),$$
$$H_{uu}(t) = H_{uu}(t, \overline{x}(t), \overline{u}(t), p(t), q(t), r(t,\cdot)).$$

Proof: By applying Taylor's formula, we get

$$
\begin{aligned}
&J\left(u^{\theta}(\cdot)\right) - J(\overline{u}(\cdot)) \\
&= \mathbb{E}\left[\int_0^T \{\delta f(t)\}\, dt\right] + \mathbb{E}\left[h(x^{\theta}(T) - h(\overline{x}(T)\right] \\
&= \mathbb{E}\Bigg[\int_0^T \Big\{ f_x(t)\left(x^{\theta}(t) - \overline{x}(t)\right) + f_u(t)\left(u^{\theta}(t) - \overline{u}(t)\right) + \frac{1}{2} f_{xx}(t)\left(x^{\theta}(t) - \overline{x}(t)\right)^2 \\
&+ f_{xu}(t)\left(x^{\theta}(t) - \overline{x}(t)\right)\left(u^{\theta}(t) - \overline{u}(t)\right) + \frac{1}{2} f_{uu}(t)\left(u^{\theta}(t) - \overline{u}(t)\right)^2 \Big\}\, dt\Bigg] \\
&+ \mathbb{E}\left[h_x(\overline{x}(T)\left(x^{\theta}(T) - \overline{x}(T)\right) + \frac{1}{2} h_{xx}(\overline{x}(T))\left(x^{\theta}(T) - \overline{x}(T)\right)^2\right] + o\left(\theta^2\right).
\end{aligned}
$$

Using Proposition 4.2.1 , we have

$$
\begin{aligned}
&J\left(u^{\theta}(\cdot)\right) - J(\overline{u}(\cdot)) \\
&= \mathbb{E}\Bigg[\int_0^T \Big\{ \theta f_x(t)\, y_1(t) + \frac{\theta^2}{2} f_x(t)\, y_2(t) + \theta f_u(t)\, v(t) \\
&+ \frac{\theta^2}{2}\left(f_{xx}(t)\, y_1(t)^2 + 2 f_{xu}(t) y_1(t)\, v(t) + f_{uu}(t)\, v(t)^2\right)\Big\}\, dt\Bigg] \\
&+ \mathbb{E}\left[\theta h_x(\overline{x}(T)\, y_1(T) + \frac{\theta^2}{2} h_x(\overline{x}(T))\, y_2(T) + \frac{\theta^2}{2} h_{xx}(\overline{x}(T))\, y_1(T)^2\right] + o\left(\theta^2\right), \quad \left(\theta \longrightarrow 0^+\right).
\end{aligned}
\tag{4.37}
$$

Substituting (4.19), (4.20), and (4.21) into (4.37), we obtain

$$
\begin{aligned}
&J(u^{\theta}(\cdot)) - J(\overline{u}(\cdot)) \\
&= -\mathbb{E}\left[\int_0^T \theta\left[p(t)\left(b_u(t)\, v(t)\right) + q(t)\left(\sigma_u(t)\, v(t)\right) - f_u(t)\, v(t)\right] dt\right] \\
&- \mathbb{E}\left[\int_0^T \frac{\theta^2}{2}\left[p(t)\, b_{uu}(t)\, v(t)^2 + q(t)\, \sigma_{uu}(t)\, v(t)^2 - f_{uu}(t) v(t)^2 + P(t)\left(\sigma_u(t)\, v(t)\right)^2\right] dt\right] \\
&- \mathbb{E}\Bigg[\int_0^T \theta^2 \{p(t)\, b_{xu}(t)\, y_1(t)\, v(t) + q(t)\, \sigma_{xu}(t)\, y_1(t)\, v(t) - f_{xu}(t) y_1(t)\, v(t) \\
&+ P(t)\, y_1(t)\, b_u(t)\, v(t) + P(t)\, \sigma_x(t)\, y_1(t)\, \sigma_u(t)\, v(t) + Q(t)\, \sigma_u(t)\, v(t)\, y_1(t)\}\, dt] + o\left(\theta^2\right).
\end{aligned}
$$

Finally, we get

$$\begin{aligned}
&J(u^\theta(\cdot)) - J(\overline{u}(\cdot)) \\
&= -\mathbb{E}\int_0^T \Big[\theta\left(H_u\left(t\right)v\left(t\right)\right) + \frac{\theta^2}{2}\left[H_{uu}\left(t\right)v\left(t\right)^2\right] \\
&\quad + \frac{\theta^2}{2}\left[P\left(t\right)\left(\sigma_u\left(t\right)v\left(t\right)\right)^2\right] + \theta^2\left[\mathbb{H}\left(t\right)y_1(t)v\left(t\right)\right]\Big] dt + o\left(\theta^2\right), \\
&\quad \left(\theta \to 0^+\right),
\end{aligned}$$

Thus, the proof of Proposition 4.2.2 is completed. ■

Now, by Proposition 4.2.2 , we can establish the following second order necessary condition in integral form for stochastic optimal control with jump diffusions (3.2)-(3.3).

Theorem 4.2.1

Let (B1)-(A2) hold. If $\overline{u}(\cdot)$ is a singular optimal control in the classical sense for the control problem (3.2)-(3.3). Then we have

$$\mathbb{E}\int_0^T \mathbb{H}(t)y_1(t)(u(t) - \overline{u}(t))dt \leq 0, \tag{4.38}$$

for any $u(\cdot) \in \mathcal{U}_{ad}$, where the Hamiltonian $\mathbb{H}$ is defined by (4.18) and $y_1(t)$ solution of first order adjoint equation given by

$$\begin{aligned}
y_1(t) &= \int_0^t \left[b_x(s)y_1(s) + b_u(s)v(s)\right]ds + \left[\sigma_x(s)y_1(s) + \sigma_u(s)v(s)\right]dW(s) \\
&\quad + \int_0^t\int_Z \left[\eta_x(s,z)y_1(s_-)\right]\widetilde{N}(dz,ds).
\end{aligned}$$

Proof: The desired result (4.38) follows immediately from (3.1) and Proposition 4.2. 2.

This completes the proof of Theorem 4.2.1 ■

4.3 Pointwise second order maximum principle in terms of the martingale with Jump Diffusions

In this section, by using the property of Itô's integrals and the martingale representation theorem, we establish the second order necessary condition for singular optimal controls, which is pointwise maximum principle in terms of the martingale with respect to the time variable t. The following lemma play an important role to establish our result.

Lemma 4.3.1

The first variational equation (4.4) admits a unique strong solution $y_1(\cdot)$, which is represented by the following:

$$y_1(t) = \Phi(t)\left[\int_0^t \Psi(s)\left(b_u(s) - \sigma_x(s)\sigma_u(s)\right)v(s)ds + \int_0^t \Psi(s)\sigma_u(s)v(s)dW(s)\right], \tag{4.39}$$

where $\Phi(t)$ is a defined by the following linear stochastic differential equation:

$$\begin{cases} d\Phi(t) &= b_x(t)\Phi(t)\,dt + \sigma_x(t)\Phi(t)\,dW(t) + \int_Z \eta_x(t,z)\,\Phi(t)\,\widetilde{N}(dz,dt), \\ \Phi(0) &= 1, \end{cases} \tag{4.40}$$

and $\Psi(t)$ its inverse .

Proof: Equation (4.4) is linear with bounded coefficients, then it admits a unique strong solution. Moreover, this solution is inversible and its inverse $\Psi(t) = \Phi^{-1}(t)$ given by the following jump diffusion:

$$\begin{cases} d\Psi(t) &= \left[\sigma_x^2(t)\Psi(t) - b_x(t)\Psi(t) + \int_Z \eta_x^2(t,z)\,\Psi(t)\mu(dz)\right]dt - \left[\sigma_x(t)\Psi(t)\right]dW(t) \\ &\quad - \int_Z \eta_x(t,z)\,\Psi(t)\widetilde{N}(dz,dt) \\ \Psi(0) &= 1. \end{cases} \tag{4.41}$$

Applying Itô's formula to $\Psi(t)y_1(t)$ we have

$$\begin{aligned} d\left[\Psi(t)y_1(t)\right] &= y_1(t)\, d\Psi(t) + \Psi(t) dy_1(t) \\ &\quad - \left[\sigma_x(t)\Psi(t)\right]\left[\sigma_x(t)y_1(t) + \sigma_u(t)v(t)\right] dt \\ &\quad - \int_Z \left\{\eta_x(t,z)\, y_1(t)\right\} \eta_x(t,z)\, \Psi(t)\mu(dz)\, dt. \\ &= I_1 + I_2 + I_3, \end{aligned} \tag{4.42}$$

where

$$\begin{aligned} I_1 &= y_1(t)\, d\Psi(t) \\ &= \left[y_1(t)\,\sigma_x^2(t)\Psi(t) - y_1(t)\, b_x(t)\Psi(t) + y_1(t)\int_Z \eta_x^2(t,z)\,\Psi(t)\mu(dz)\right] dt \\ &\quad - y_1(t)\,\sigma_x(t)\Psi(t) dW(t) \\ &\quad - y_1(t)\int_Z \eta_x(t,z)\,\Psi(t)\widetilde{N}(dz,dt). \end{aligned} \tag{4.43}$$

By simple computations, we obtain

$$\begin{aligned} I_2 &= \Psi(t) dy_1(t) \\ &= \left[\Psi(t)b_x(t)y_1(t) + \Psi(t)b_u(t)v(t)\right] dt \\ &\quad + \left[\Psi(t)\sigma_x(t)y_1(t) + \Psi(t)\sigma_u(t)v(t)\right] dW(t) \\ &\quad + \Psi(t)\int_Z \left\{\eta_x(t,z)\, y_1(t)\right\} \widetilde{N}(dz,dt), \end{aligned} \tag{4.44}$$

and

$$\begin{aligned} I_3 &= -\left[\sigma_x(t)\Psi(t)\right]\left[\sigma_x(t)y_1(t) + \sigma_u(t)v(t)\right] dt \\ &\quad - \int_Z \left(\eta_x(t,z)\, y_1(t)\right) \eta_x(t,z)\,\Psi(t)\mu(dz)\, dt. \end{aligned} \tag{4.45}$$

Substituting (4.42), (4.43), and (4.44) into (4.42), we get

$$\begin{aligned} &\Psi(t)y_1(t) - \Psi(0)y_1(0) \\ &= \left[\int_0^t \Psi(s)\left[b_u(s) - \sigma_x(s)\sigma_u(s)\right] v(s) ds \right. \\ &\quad + \int_0^t \Psi(s)\sigma_u(s)v(s)dW(s). \end{aligned} \tag{4.46}$$

Since $y_1(0) = 0$ and $\Psi^{-1}(t) = \Phi(t)$, then from (4.46) the desired result (4.40) is fulfilled. This completes the proof of Lemma 4.3.1 ∎

To prove the main theorem we need the following technical Lemma.

Lemma 4.3.2

Let (B1)-(A2) hold. Then we have

(1) $\mathbb{H}(\cdot) \in \mathbb{L}_{\mathbb{F}}^{2}([0,T];\mathbb{R})$

(2) For any $v \in U$*, there exists* $\phi_v(\cdot,t) \in \mathbb{L}_{\mathbb{F}}^{2}([0,T];\mathbb{R})$*, and* $\gamma_v(\cdot,t,z) \in \mathcal{L}^2([0,T];\mathbb{R})$ *such that*

$$\mathbb{H}(t)(v-\bar{u}(t)) = \mathbb{E}\left[\mathbb{H}(t)(v-\overline{u}(t))\right] + \int_0^t \phi_v(s,t)\, dW(s) + \int_0^t \int_Z \gamma_v(s,t,z)\, \widetilde{N}(dz,ds) \tag{4.47}$$

$$a.e.\ t \in [0,T],\ \ P-a.s.$$

Proof: The proof follows immediately from Lemma 3.9 in [58].

(2) The proof of (4.47) follows from Tang and Li [50, Appendix]. ■

Now, we return to integral type of second order necessary condition and substituting the explicit representation (4.39) of $y_1(\cdot)$ into (4.38), we see that there appears a "bad" term in the form

$$\mathbb{E}\int_0^T \left[\mathbb{H}(t)\,\Phi(t)\int_0^t \Psi(s)\sigma_u(s)v(s)dW(s)\right] v(t)dt, \tag{4.48}$$

For more details see [58, p.2278] for this type of integrals.

Now, in order to derive a pointwise second order necessary condition from the integral form in (4.38), we need to choose the following needle variation for the optimal control $\overline{u}(\cdot)$:

$$u(t) = \begin{cases} v, & t \in A_\theta, \\ \overline{u}(t), & t \in [0,T] \mid A_\theta, \end{cases} \tag{4.49}$$

where $\tau \in [0,T)$, $v \in U$, and $A_\theta = [\tau, \tau+\theta)$ so that $\theta > 0$ and $\tau+\theta \leq T$. Denote by $\chi_{A_\theta}(\cdot)$ the characteristic function of the set A_θ. Then we have $v(\cdot) = u(\cdot) - \overline{u}(\cdot) = (v - \overline{u}(\cdot))\,\chi_{A_\theta}$.

The following theorem constitutes the main contribution of this paper.

Theorem 4.3.1

Let (B1)-(A2) hold. If $\overline{u}(\cdot)$ *is a singular optimal control in the classical sense for the* *stochastic control (3.2)-(3.3), then for any $v \in U$, it holds that*

$$\mathbb{E}\left(\mathbb{H}(\tau)b_u(\tau)(v-\overline{u}(\tau))^2\right)+\partial_\tau^+\left(\mathbb{H}(\tau)(v-\overline{u}(\tau))^2\sigma_u(\tau)\right)\leq 0 \quad a.e.\ \tau\in[0,T], \tag{4.50}$$

where

$$\begin{aligned}&\partial_\tau^+\left(\mathbb{H}(\tau)(v-\overline{u}(\tau))^2\sigma_u(\tau)\right) \\ &:= 2\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\int_\tau^t\left[\phi_v(s,t)\Phi(\tau)\Psi(s)\sigma_u(s)(v-\overline{u}(s))\right]dsdt,\end{aligned} \tag{4.51}$$

$\phi_v(\cdot,t)$ is determined by (4.47), $\Phi(\cdot)$ is given by the following jump process

$$\begin{aligned}\Phi(t)&=\Phi(0)+\int_0^t b_x(s)\Phi(s)\,ds+\int_0^t\sigma_x(s)\Phi(s)\,dW(s)\\&+\int_0^t\int_Z\eta_x(s,z)\,\Phi(s)\,\widetilde{N}(dz,ds).\end{aligned}$$

and $\Psi(\cdot)$ is determined by

$$\begin{aligned}\Psi(t)&=\Psi(0)+\int_0^t\left[\sigma_x^2(s)\Psi(s)-b_x(s)\Psi(s)+\int_Z\eta_x^2(s,z)\,\Psi(s)\mu(dz)\right]ds\\&-\int_0^t\left[\sigma_x(s)\Psi(s)\right]dW(s)-\int_0^t\int_Z\eta_x(s,z)\,\Psi(s)\widetilde{N}(dz,ds).\end{aligned}$$

Proof: From (4.49), we have $v(\cdot)=u(\cdot)-\overline{u}(\cdot)=(v-\overline{u}(\cdot))\chi_{A_\theta}(\cdot)$ and the corresponding solution $y_1(\cdot)$ to (4.4) is given by

$$\begin{aligned}y_1(t)&=\Phi(t)\int_0^t\Psi(s)\left(b_u(s)-\sigma_x(s)\sigma_u(s)\right)(v-\overline{u}(s))\chi_{A_\theta}(s)ds\\&+\Phi(t)\int_0^t\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))\chi_{A_\theta}(s)dW(s).\end{aligned} \tag{4.52}$$

Substituting $v(\cdot)=(v-\overline{u}(\cdot))\chi_{A_\theta}(\cdot)$ and (4.52) into (4.38), we have

$$\begin{aligned}0&\geq\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left[\mathbb{H}(t)\,y_1(t)\,(v-\overline{u}(t))\right]dt\\&=\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left[\mathbb{H}(t)\,\Phi(t)\int_\tau^t\Psi(s)\left(b_u(s)-\sigma_x(s)\sigma_u(s)\right)(v-\overline{u}(s))ds\,(v-\overline{u}(t))\right]dt\\&+\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left[\mathbb{H}(t)\,\Phi(t)\int_\tau^t\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt\\&=J_1^\theta+J_2^\theta.\end{aligned} \tag{4.53}$$

From [58, Lemma 4.1], we obtain

$$\begin{aligned}\lim_{\theta\to 0^+} J_1^\theta &= \lim_{\theta\to 0^+}\frac{1}{\theta^2}\mathbb{E}\int_\tau^{\tau+\theta}\left[\mathbb{H}(t)\,\Phi(t)\int_\tau^t\Psi(s)\,(b_u(s)-\sigma_x(s)\sigma_u(s))\,(v-\overline{u}(s))ds\,(v-\overline{u}(t))\right]dt\\ &=\frac{1}{2}\mathbb{E}\left(\mathbb{H}(\tau)\,(b_u(\tau)-\sigma_x(\tau)\sigma_u(\tau))\,(v-\overline{u}(\tau))^2\right). \qquad (4.54)\end{aligned}$$

On the other hand, by (4.40), it follows that

$$\begin{aligned}J_2^\theta &= \frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\,\Phi(t)\int_\tau^t\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt\\ &=\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\,\Phi(\tau)\int_\tau^t\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt\\ &+\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\int_\tau^t b_x(s)\Phi(s)\,ds\right.\\ &\qquad\times\left.\int_\tau^t\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt \qquad (4.55)\\ &+\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\int_\tau^t\sigma_x(s)\Phi(s)\,dW(s)\right.\\ &\qquad\times\left.\int_\tau^t\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt\\ &+\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left\{\mathbb{H}(t)\int_\tau^t\int_Z\eta_x(s,x(s_-),z)\,\Phi(s)\,\widetilde{N}(dz,ds)\right.\\ &\qquad\times\left.\int_\tau^t\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right\}dt.\\ &=J_{2,1}^\theta+J_{2,2}^\theta+J_{2,3}^\theta+J_{2,4}^\theta.\end{aligned}$$

By Lemma 4.3.2 , we get

$$\begin{aligned}&\lim_{\theta\to 0^+}\sup J_{2,1}^\theta\\ &=\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\mathbb{H}(t)\,\Phi(\tau)\int_\tau^t\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\,(v-\overline{u}(t))\right]dt\\ &=\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left[\int_\tau^t\Phi(\tau)\,\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\mathbb{E}\left[\mathbb{H}(t)\,(v-\overline{u}(t))\right]\right]dt\\ &+\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left\{\int_\tau^t\Phi(\tau)\,\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\int_0^t\phi_v(s,t)dW(s)\right\}dt\\ &+\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left\{\int_\tau^t\Phi(\tau)\,\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))dW(s)\int_0^t\int_Z r(s,z)\,\widetilde{N}(dz,ds)\right\}dt\\ &=\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\int_\tau^t\mathbb{E}\left\{\Phi(\tau)\,\Psi(s)\,\sigma_u(s)(v-\overline{u}(s))\phi_v(s,t)\right\}dsdt \qquad (4.56)\\ &=\frac{1}{2}\partial_\tau^+\left(\mathbb{H}(\tau)\,(v-\overline{u}(\tau))^2\sigma_u(\tau)\right),\qquad\forall\tau\in[0,T).\end{aligned}$$

It is crucial that, by the Martingale Representation Theorem in Lemma 4.3.2 , we only know that $\phi_v(\cdot,t)\in\mathbb{L}_{\mathbb{F}}^2([0,t];\mathbb{R})$ for any $v\in U$, and hence, for each $\tau\in[0,T]$, the function

$$\varphi_t(s) = \mathbb{E}\left[\Phi(\tau)\Psi(s)\sigma_u(s)(v-\overline{u}(s))\phi_v(s,t)\right], \quad s\in[0,t], \ t\in[0,T],$$

is in $\mathbb{L}^1([0,t];\mathbb{R})$. See [58] for more details for the superior limit

$$\lim_{\theta\to 0^+}\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\int_\tau^t \varphi_t(s)\,dsdt.$$

By simple computations, the last term in (4.56) is in fact a process with zero expectation. Now, by using similar method in [58], we get

$$\begin{aligned}\lim_{\theta\to 0^+} J_{2,2}^\theta &= \lim_{\theta\to 0^+}\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left\{\mathbb{H}(t)\int_\tau^t b_x(s)\Phi(s)\,ds\right. \\ &\quad\times\left.\int_\tau^t\Psi(s)\sigma_u(s)(v-\overline{u}(s))dW(s)(v-\overline{u}(t))\right\}dt \\ &= 0,\end{aligned} \tag{4.57}$$

$$\begin{aligned}\lim_{\theta\to 0^+} J_{2,3}^\theta &= \lim_{\theta\to 0^+}\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left\{\mathbb{H}(t)\int_\tau^t \sigma_x(s)\Phi(s)\,dW(s)\right. \\ &\quad\times\left.\int_\tau^t\Psi(s)\sigma_u(s)(v-\overline{u}(s))dW(s)(v-\overline{u}(t))\right\}dt \\ &= \frac{1}{2}\mathbb{E}\left(\mathbb{H}(\tau)(\sigma_x(\tau)\sigma_u(\tau))(v-\overline{u}(\tau))^2\right),\end{aligned} \tag{4.58}$$

and

$$\begin{aligned}\lim_{\theta\to 0^+} J_{2,4}^\theta &= \lim_{\theta\to 0^+}\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\mathbb{E}\left\{\mathbb{H}(t)\int_\tau^t\int_Z \eta_x(s,x(s_-),z)\,\Phi(s)\,\widetilde{N}(dz,ds)\right. \\ &\quad\times\left.\int_\tau^t\Psi(s)\sigma_u(s)(v-\overline{u}(s))dW(s)(v-\overline{u}(t))\right\}dt \\ &= 0.\end{aligned} \tag{4.59}$$

Finally, substituting (4.54), (4.56), (4.57), (4.58), (4.59) in (4.53), we obtain

$$\mathbb{E}\left(\mathbb{H}(\tau)b_u(\tau)(v-\overline{u}(\tau))^2\right)+\partial_\tau^+\left(\mathbb{H}(\tau)(v-\overline{u}(\tau))^2\sigma_u(\tau)\right)\le 0, \quad a.e. \ \ \tau\in[0,T].$$

This completes the proof of Theorem 4.3.1 ■

Conclusion

In this thesis, a second order necessary conditions for stochastic optimal control of jump diffusions have been proved. Pointwise second order maximum principle in terms of the martingale with respect to the time variable has been established. The control variable is allowed to enter both drift and diffusion terms. The control domain is assumed to be convex. When the coefficient $\eta \equiv 0$, our results coincides with pointwise second-order maximum principle developed in Zhang and Zhang [58].

When the control enters into both diffusion and jump terms, and the system has the form:

$$\begin{cases} dx\left(t\right)=b\left(t,x\left(t\right),u\left(t\right)\right)dt+\sigma(t,x\left(t\right),u\left(t\right))dW\left(t\right)+\displaystyle\int_{Z}\eta\left(t,x\left(t_{-}\right),u\left(t\right),z\right)\widetilde{N}\left(dz,dt\right), \\ x(0)=x_{0}, \end{cases} \tag{4.60}$$

the pointwise necessary conditions for optimal stochastic control problem (4.60)-(3.3) becomes very complicated. It leads to many problems that we cannot solve now. But we can only establish a second-order maximum principle in integral form.

Following the ideas considered in [58, 59], and in order to establish the second-order necessary conditions, one needs to assume that the first order condition degenerates in some sense. So we define a new argument of singularity in the classical sense associated to control problem (4.60)-(3.3). An admissible control $\widetilde{u}(\cdot)$ is called singular in the classical sense if satisfies

$$\begin{cases} H_u\left(t,\widetilde{x}\left(t\right),\widetilde{u}\left(t\right),\widetilde{p}\left(t\right),\widetilde{q}\left(t\right),\widetilde{r}\left(t,z\right)\right)=0 \quad a.s.\ a.e.\ \ t\in[0,T]\,, \\ H_{uu}\left(t,\widetilde{x}\left(t\right),\widetilde{u}\left(t\right),\widetilde{p}\left(t\right),\widetilde{q}\left(t\right),\widetilde{r}\left(t,z\right)\right)+\widetilde{P}\left(t\right)\left(\sigma_u\left(t,\widetilde{x}\left(t\right),\widetilde{u}\left(t\right)\right)\right)^2 \\ +\int_Z(\widetilde{P}\left(t\right)\left(\eta_u\left(t,\widetilde{x}\left(t\right),\widetilde{u}\left(t\right),z\right)\right)^2+\widetilde{R}\left(t,z\right)\left(\eta_u\left(t,\widetilde{x}\left(t\right),\widetilde{u}\left(t\right),z\right)\right)^2)\mu\left(dz\right)=0\ \ a.e.\ a.s., \end{cases} \tag{4.61}$$

where $(\widetilde{p}(\cdot),\widetilde{q}(\cdot),\widetilde{r}(\cdot))$ and $(\widetilde{P}(\cdot),\widetilde{Q}(\cdot),\widetilde{R}(\cdot))$ are the adjoint processes given respectively by (4.16) and (4.17) associated to $(\widetilde{x}(\cdot),\widetilde{u}(\cdot))$.

By using similar arguments developed above, and under the conditions (B1), (A2) with some additional assumptions on the jump coefficient η, we can establish second-order necessary conditions in integral form:

Theorem 4.3.2

If $\overline{u}(\cdot)$ is a singular optimal control in the classical sense defined in (4.61), then we have

$$\mathbb{E}\int_0^T \mathbb{H}(t)y_1(t)v(t)dt\leq 0, \tag{4.62}$$

for any $v(\cdot)=u(\cdot)-\overline{u}(\cdot)$ with $u(\cdot)\in\mathcal{U}_{ad}$, where $\mathbb{H}$ has the form

$$\begin{aligned}\mathbb{H}\left(t,x,u,p,q,r,P,Q,R\right)&=H_{xu}\left(t,x,u,p,q,r\right)+b_u\left(t,x,u\right)P(t)+\sigma_u\left(t,x,u\right)Q(t)\\&+\sigma_u\left(t,x,u\right)P(t)\sigma_x\left(t,x,u\right)+\int_Z\eta_u\left(t,x,u,z\right)R\left(t,z\right)\mu\left(dz\right)\\&+\int_Z\eta_u\left(t,x,u,z\right)P(t)\eta_x\left(t,x,u,z\right)\mu\left(dz\right)\\&+\int_Z\eta_u\left(t,x,u,z\right)R\left(t,z\right)\eta_x\left(t,x,u,z\right)\mu\left(dz\right).\end{aligned} \tag{4.63}$$

and $y_1(t)$ is the solution of the first variational equation

$$\begin{cases} dy_1\left(t\right)=\left\{b_x\left(t\right)y_1\left(t\right)+b_u\left(t\right)v\left(t\right)\right\}dt+\left\{\sigma_x\left(t\right)y_1\left(t\right)+\sigma_u\left(t\right)v\left(t\right)\right\}dW(t) \\ \qquad +\int_Z\left\{\eta_x\left(t,z\right)y_1\left(t_-\right)+\eta_u\left(t,z\right)v\left(t\right)\right\}\widetilde{N}\left(dz,dt\right), \\ y_1(0)=0, \end{cases} \tag{4.64}$$

The main difficulties to prove pointwise second order necessary conditions of optimality arise due to the appearance of many new "bad" terms. The presence of control variable in jump coefficient creates some new superior limits, which are difficult to obtain. Moreover, our classical assumptions are not sufficient to ensure the existence of these *superior limits. We hope to solve it in the near future. Another challenging problem left unsolved is to derive a pointwise second order necessary conditions for such control problems in the case where the control domain is not assumed to be convex.*

Bibliography

[1] Barles G. Solutions de Viscosité des Equations de Hamilton-Jacobi. *Math. & Appl* 1994;17.

[2] Bensoussan A. Stochastic maximum principle for distributed parameter system. *Journal of the Franklin Inst* 1983;315:387-406.

[3] Bensoussan A. Lectures on stochastic control. *Nonlinear filtering and stochastic control* 1982: 1-62.

[4] Bismut J.M. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications* 1973;44:384–404.

[5] Bismut J.M. An introductory approach to duality in optimal stochastic control. *SIAM Rev* 1978;20:62-78.

[6] Bouchard B., Elie, R. Discrete time approximation of decoupled forward-Backward SDE with jumps, *Stochastic Process. Appl.* 2008;118:53-75.

[7] Bonnans J.F., Silva F.J. First and second order necessary conditions for stochastic optimal control problems. *Applied Mathematics & Optimization* 2012;65(3):403-439.

[8] Cadenillas A. A stochastic maximum principle for systems with jumps, with applications to finance. *Systems & control letters* 2002;47(5):433-444.

[9] Cadenillas A., Haussmann, U.G. The stochastic maximum principle for a singular control problem. *Stochastics: An International Journal of Probability and Stochastic Processes* 1994;49(3-4):211-237..

[10] Cadenillas A., Karatzas I. The stochastic maximum principle for linear, convex optimal control with random coefficients. *SIAM Journal on control and optimization* 1995;33(2):590-624.

[11] Chan T. Pricing contingent claims on stocks driven by Lévy processes. *Annals of Appl. Prob* 1999;9:504–528.

[12] Deepa R., Muthukumar P. Infinite horizon optimal control of mean-field delay system with semi-Markov modulated jump-diffusion processes, *The Journal of Analysis*, 2018:1–19.

[13] Dong Y., Meng, Q.X. Second-order necessary conditions for optimal control with recursive utilities. *Optim Theory Appl* 2019;182(2), 494-524.

[14] EKELAND I. Sur les problems variationels. *Acad. Sci. Paris* 1972;275:1057-1059.

[15] Framstad N.C., Øksendal, B., Sulem, A. Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance*, J. Optim. Theory. Appl.* 2004;121:77-98.

[16] Frankowska, H., Tonon, D. Pointwise second-order necessary optimality conditions for the Mayer problem with control constraints. *Siam J. Control. Optim.* 2013;51(5):3814-3843.

[17] Frankowska H., Zhang H., Zhang X. Necessary optimality conditions for local minimizers of stochastic optimal control problems with state canstraints. *Transactions of the American mathematical society* 2019;372(2):1289-1331.

[18] Fleming W.H., Rishel R.W. Deterministic and Stochastic Optimal Control 1975.

[19] Fleming W.H., Soner H.M. Controlled Markov processes and viscosity solutions. 1993.

[20] Gabasov, R., Kirillova, F. High order necessary conditions for optimality. *Siam J. Control. Optim* 1972;10(1), 127-168.

[21] Haussmann, U.G. A stochastic maximum principle for optimal control of diffusions. *JohnWiley & Sons, Inc* 1986.

[22] Haussmann U. Some examples of optimal stochastic controls or: The stochastic maximum principle at work. *SIAM review* 1981;23:292-307.

[23] Haussmann U. General necessary conditions for optimal control of stochastic system. *Math. Programming Stud* 1976;6:34-48.

[24] Haussmann U., Suo W. Singular optimal control I, II, *Siam J. Control Optim.* 1995;33(3):916-936, 937-959.

[25] Haugh, Martin. A Brief Introduction to Stochastic Calculus, 2016.

[26] Hafayed M, Abbas S. Stochastic near-optimal singular controls for jump diffusions: necessary and sufficient conditions, *Journal of Dynamical and Control Systems*, 2013;19(4), 503-517.

[27] Hafayed M. A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes, *Int. J. Dynam. Control*, 2013; 1(4):300-315.

[28] Hu Y. Maximum principle of optimal control for Markov processes. *Symposium on Control Theory and Applications* 1987.

[29] Ikeda N , Watanbe S. Stochastic Differential Equations and Diffusion Processes. North Holland, Amsterdam, 1981.

[30] Jacod J, Shiryaev A. Limit Theorems for Stochastic Processes 1987.Springer-Verlag .

[31] Kushner H. Necessary conditions for continuous parameter stochastic optimization problems. *SIAM Journal on Control* 1972;10(3):550-565.

[32] Kallenberg O. Foundations of Modern Probability 2002 Second Edition.Springer-Verlag .

[33] Kushner N.J. Necessary conditions for continuous parameter stochastic optimization problems, *SIAM J Control Optim* 1972;10: 550–565.

[34] Krylov N V. Controlled Diffusion Processes 1980. Springer Verlag, New York.

[35] Lions P. Optimal control of difusion processes and Hamilton-Jacobi-Bellman equations. 1: The dynamic programming principle and application, 2: Viscosity solutions and uniqueness. *Communications in Partial Differential Equations* 1983;8:1101-1174, 1229-1276.

[36] Lakhdari I.E. Optimal control for stochastic differential equations governed by normal martingales. Doctoral thesis 2018; *http://thesis.univ-biskra.dz/3677/.*

[37] Meng Q. General linear quadratic optimal stochastic control problem driven by a Brownian motion and a Poisson random martingale measure with random coefficients. *Stochastic Analysis and Applications* 2014;32(1):88-109.

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

[39] Miloudi H., Meherrem S., Lakhdari I.E., Hafayed M. Necessary conditions for partially observed optimal control of general McKean–Vlasov stochastic differential equations with jumps. *International Journal of Control 2021;*93(5):1-13.

[40] Oksendal B., Sulem A. Applied stochastic control of jump diffusions, *Springer-Verlag Berlin Heidelberg* 2005.

[41] Peng S. A general stochastic maximum principle for optimal control problems. *Siam J. Control. Optim.* 1990;28(4):966-979.

[42] Perturbation Methods in Optimal Control. *Dunod.Gauthier-Villars* 1988.

[43] Protter P. Stochastic Integration and Differential Equations 2003. Second Edition. Springer-Verlag.

[44] Pontrvagin L S., Boltyanskii V G., Gamkrelidze R V., Mischenko E F. The Mathematical Theory of Optimal Control Processes, *John Wiley* 1962.

[45] Rishel R. A minimum principle for controlled jump processes. *Lecture Notes in Economics and Mathematical Systems,* Springer, Berlin, Heidelberg, New York, 1975;107:493–508.

[46] Shen Y., Meng Q., Shi P. Maximum principle for mean-field jump diffusion stochastic delay differential equations and its application to finance. *Automatica* 2014;50(6):1565-1579.

[47] Situ R. A maximum principle for optimal controls of stochastic with random jumps, in *Proceedings of National Conference on Control Theory and its Applications*, Qingdao, China 1991.

[48] Shi J., Wu, Z. Maximum principle for Forward-backward stochastic control system with random jumps and application to finance. *Journal of Systems Science and Complexity* 2010;23:219-231.

[49] Sato K. Lévy Processes and Infinitely Divisible Distributions. Cambridge Univ. Press 1999.

[50] Tang S., Li, X. Necessary conditions for optimal control of stochastic systems with random jumps. *Siam J. Control. Optim.* 1994;32(5):1447-1475.

[51] Tang M., Meng Q. Linear-Quadratic Optimal Control Problems for Mean-Field Stochastic Differential Equations with Jumps. *Asian Journal of Control* 2019; 21(2):809-823.

[52] Tang S. A second order maximum principle for singular optimal stochastic controls, *Dynamical systems, Series B*, 2010;4(4):1581–1599.

[53] Tang M, Meng QX. Stochastic evolution equations of jump type with random coefficients: existence, uniqueness and optimal control, *Science China Information Sciences* 2017;60:118-202.

[54] Xiao H., Wang G. The filtering equations of forward-backward stochastic systems with random jumps and applications to partial information stochastic optimal control, *Stochastic Analysis and Applications,* 2010;28:1003-1019.

[55] J. Yong, X.Y. Zhou, Stochastic Controls, Hamiltonian Systems and HJB Equations, Springer Verlag, Berlin 1999.

[56] Zhou X.Y.,Yin G. Markowitz's mean-variance portfolio selection with regime switching: A continuous-time model. *SIAM Journal on Control and Optimization* 2003;42:1466-1482.

[57] Zhang X., Elliott, R.J., Siu, T.K. A stochastic maximum principle for a Markov regime-switching jump-diffusion model and its application to finance. *Siam J. Control. Optim.* 2012;50(2), 964-990.

[58] Zhang H., Zhang X. Pointwise second-order necessary conditions for stochastic optimal controls, Part I: The case of convex control constraint.*Siam J. Control. Optim.* 2015;53(4):2267-2296.

[59] Zhang H., Zhang, X. Pointwise second-order necessary conditions for stochastic optimal controls, Part II: The general case. *Siam J. Control. Optim.* 2017;55(5), 2841-2875.