Title of Thesis: A Finance Perspective to Risk-Sensitive Optimal Control with Regime Switching and Applications

Serges Love Teutu Talla, Master of Arts, December 2019

Thesis Chair: Isabelle Kemajou-Brown, PhD

Department of Mathematics

**Abstract**

In this work, we briefly review the literature on stochastic optimization problems using the Pontryagin Maximum Principle. Then, we investigate the procedure for solving a risk-sensitivity stochastic maximum principle problem with regime-switching. Finally, we provide an application in finance.

# A Finance Perspective to Risk-Sensitive Optimal Control with Regime Switching and Applications

by

Serges Love Teutu Talla

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree

Master of Arts

Morgan State University

December 2019

# A Finance Perspective to Risk-Sensitive Optimal Control with Regime Switching and Applications

by

Serges Love Teutu Talla

has been approved

October, 2019

THESIS COMMITTEE APPROVAL:

____________________________, Chair

Isabelle Kemajou-Brown, PhD

____________________________

Gaston N'Guérékata, PhD

____________________________

Xuming Xie, PhD

____________________________

Najat Ziyadi, PhD

# Acknowledgments

I want to express my special gratitude and thanks to my advisor, Dr. Isabelle Kemajou-Brown, for her patience and support. It is also a great pleasure to acknowledge my sincere thanks and gratitude to my committee members: Dr. N'Guérékata, Dr. Xie, and Dr. Ziyadi for their guidance and suggestions. Also, major thank is due to Dr. Cai for the financial support provided during this summer and Dr. Romeo Tedongap, professor at ESSEC Business School France for the documentation provided. Finally, I wish to thank my family members for their love and encouragement, without whom I would never have enjoyed so many opportunities.

# Contents

# 1 Introduction

The financial market is an environment where people trade stocks, financial securities, and other values. This market is characterized by volatility and risk-sensitivity. The volatility is a statistical measure of the degree of uncertainty associated with a particular security or investment portfolio while, the risk sensitivity accounts for the degree to which investors are willing to take risk. Over the last century, mathematicians have developed various models of financial commodities (cf. [25]) as an attempt to predict the prices of these commodities and to optimize trading strategies. Now, since financial objects are subject to random behavior, how can one adjust the model to account this random fluctuations?

Brownian motion marked the first step in the process to bring some solutions to this problem. The Brownian motion was first discovered by the biologist Robert Brown. While examining pollen grains under the microscope, he observed that the grains in the water were moving continuously and randomly on a different way. In fact, the pollen grains were being knocked around by the water molecules. An example of such continuous and random movements in finance would be the increments and decrements of stock prices. The french mathematician Louis Bachelier, was the first to introduce Brownian motion in finance in 1900. In his doctoral thesis, Bachelier introduced a mathematical model of Brownian motion and its use for valuing stock options.

He is historically the first to use advanced mathematics in the study of finance, therefore he is considered as the forefather of mathematical finance and a pioneer in the study of stochastic processes.

For the price of a stock on the market, making analogy to Brown's observation, the traders activities along with other factors have a random impact on the price.

Brownian motion also called the Wiener process, named after Norbert Wiener, is a mathematical tool that formalizes the random behavior observed by Robert Brown in 1827. It can be defined as the limit for scaling random steps such as step size and time interval. The Wiener process or Brownian motion is very important in stochastic processes modeling because it represents the integral of the noise idealized, independently of the frequency, called white noise. The Wiener process is often used to represent random external influences on a system, or more generally, it is an instrument used when the deterministic model fails to capture the presence of uncertainties.

In this thesis, we will provide the principle for a class of diffusion processes that are usually used to describe financial commodities. Probability theory provides tools that can be used to represent and convert our beliefs, about the dynamics of financial instruments, into actions. The purpose of this thesis is to study and understand some works that have been done on the risk-sensitive stochastic optimal control with Regime switching using the Maximum Principle. The work done in this document is a combination of a revision of lectures notes for stochastic differential equations (SDE), and also a survey of some works done in stochastic optimization, risk-sensitivity, and Markov regime switching.

There exist two principal and mostly used approaches in solving stochastic optimal control problems. The first approach is the Bellman's Dynamic Programming, where the system consists of a partial differential equation (PDE), known as the Hamilton-Jacobi-Bellman (HJB) equation (cf. [25]). In a deterministic case, the PDE is of first order, and in a stochastic case, the PDE is of second order. The second approach is the Pontryagin's Maximum Principle, where the system consists of the adjoint equation, the original state equation and the maximum condition referred to an (extended) Hamiltonian system (cf. [25]). In a deterministic case, the adjoint equation is an ordinary differential equation (ODE), and in a stochastic case, the adjoint is a stochastic differential equation (SDE). Since the publication of the deterministic maximum principle by Pontryagin et al., in 1956 (cf. [25]), tremendous researches have been done in the stochastic optimal control theory. One of the big challenge was to determine a general stochastic maximum principle if the diffusion depends on the control and the control domain is not necessarily convex. This problem has been the subject of many researches on the field since 1960 (cf. [25]). Until 1980, most of these researches considered a diffusion term independent of the control variable or a diffusion term that depends on a convex control domain. Under these considerations the results for stochastic and deterministic maximum principle are very similar. In 1990, Peng and other mathematicians from Fudan University (Shanghai) were able to solve a stochastic maximum principle for systems with control dependent diffusion coefficients and possibly non convex control domains, by introducing for the first time, the second order adjoint equation (cf. [17]).

## Literature review

The first to study the necessary condition Stochastic Maximum Principle (SMP) was Krushner in 1965. Krushner's result was extended to more general stochastic control problems with control-free diffusion coefficients by Bismut in 1973, Bensoussam in 1981, and Haussmann in 1986. For stochastic control problems with controlled diffusion coefficients, Peng in 1990 applied the second order-adjoint equation to derive the necessary SMP. Zhou in 1991 simplified Peng's proof. Tang and Li in 1994 extends Peng's SMP to systems with jump diffusions (cf. [12]), Cadenillas and Karatzas in 1995 extends Peng's SMP to systems with random coefficients. Bismut was the first researcher to study the sufficient SMP in 1978. Peng's necessary SMP result was proved to be also sufficient under certain convexity assumption by Zhou in 1996. In 2004, Framstad, Oksendal, and Sulem applied the sufficient stochastic maximum principle (SMP) to jump diffusion systems. In 2011, Donnelly extended the SMP with Markovian regime-switching, and Zhang, Elliot and Siu in 2012 with Markov regime-switching for jump diffusion processes. Peter Whittle was one of the first to work on risk-sensitive stochastic maximum principle in 1990, when he derived a maximum principle based on the theory of large deviations. Then in 2005, Lim and Zhou established a new risk-sensitive maximum principle, based on the general stochastic maximum principle of Peng and the relationship between maximum principle and dynamic programming principle of Yong and Zhou. In 2009, Wang and Wu considered a more general risk-sensitive cost functional, and found applications in finance. Therefore, tremendous works have been done in finance using the risk-sensitive maximum principle. Among them, the derivation of a general stochastic maximum principle for risk-sensitive type optimal control problem of Markov regime-switching jump-diffusion model, developed by Zhongyang Sun, Isabelle Kemajou-Brown, and Olivier Menoukeu-Pamen in 2017 (cf. [21]).

The remainder of this thesis is organised as follows:

Chapter 2 introduces notation and definitions used throughout this document. It presents a brief review of the relevant theoretic background concerning probability theory, and stochastic differential equations; in particular, definitions of Sigma-algebra, measure space, probability measure, probability space, random variable, and differential equations. We also recall in this chapter the axioms of probability as defined by Kolmogorov, and the properties of Wiener process. We finally recall in this chapter the basic concepts on control optimization, referring to the two common methods used for optimization problems: the dynamic programming method and the Pontryagin maximum principle. The two fundamental theorems of maximum principle are stated. In Chapter 3, we state a stochastic optimal problem and we derive the risk-sensitive stochastic maximum principle. In Chapter 4, the concept of regime switching is presented in a stochastic-optimal problem with risk-sensitivity. Chapter 5 presents an application of a risk-sensitive stochastic optimal control problem based on an optimal portfolio choice problem in the financial market, and using the hyperbolic absolute risk aversion (HARA) utility function. We give our conclusion in chapter 6.

# 2 Preliminaries and framework

In this section we recall and define some of the important mathematical concepts used throughout this document. More details on these concepts can be found in [10].

## 2.1 Notation and Acronyms

The following list summarizes notation and acronyms used in this work.

| | |
|---|---|
| $\mathbb{R}$ | the set of reals |
| $\mathbb{R}_+$ | the set of positive reals |
| $\mathbb{R}^n$ | n-dimensional Euclidean space |
| $\mathbb{R}^{n\times m}$ | the set of all $(n \times m)$ real matrices |
| $I_n$ | the $n \times n$ identity matrix |
| $A^\mathsf{T}$ | the transpose of a matrix $A$ |
| $trA$ | the trace of a matrix $A$ |
| $C^k$ | the space of functions with continuous derivatives up to order $k$ |
| $\langle \cdot, \cdot \rangle$ | the inner product in some Hilbert space. |
| $\mathbb{E}[\cdot]$ | mathematical expectation |
| $\mathscr{F}_t$ | filtration at time $t$ |
| $\mathbb{E}\big[\,\cdot\,\lvert\mathscr{F}_t\big]$ | expectation conditioned on $\mathscr{F}_t$ |
| $\triangleq$ | defined as |
| $:=$ | equal by definition |
| $\equiv$ | identically equal to |
| $\approx$ | approximately equal to |
| $\bigvee$ | the join of all elements operated on |
| $\mathcal{N}(\mu, \sigma^2)$ | normal (Gaussian) distribution with mean $\mu$ and variance $\sigma^2$ |
| $W(t)$ | standard Brownian motion process |

| | |
|---|---|
| $L^k_{\mathscr{F}}\big([0,T];\mathbb{R}^n\big)$ | the space of all $\{\mathscr{F}\}_{t\geq 0}$ -adapted $\mathbb{R}^n$-valued processes $f:[0,T]\to\mathbb{R}^n$ such that $\mathbb{E}\Big[\int_0^T |f(t)|^k dt\Big] < \infty$, for $1\leq k<\infty$ |
| $L^k_{\mathscr{F},p}\big([0,T];\mathbb{R}^n\big)$ | the space of all $\{\mathscr{F}\}_{t\geq 0}$ predictable $\mathbb{R}^n$-valued processes $f:[0,T]\to\mathbb{R}^n$ such that $\mathbb{E}\Big[\int_0^T |f(t)|^k dt\Big] < \infty$, for $1\leq k<\infty$ |
| $\mathcal{L}^k\big([0,T];\mathbb{R}^n\big)$ | the space of measurable functions $f:[0,T]\to\mathbb{R}^n$ such that $\mathbb{E}\Big[\int_0^T \|f(t)\|^k dt\Big] < \infty$, for $1\leq k<\infty$ |
| $\mathcal{M}^2\big([0,T];\mathbb{R}^n\big)$ | the set of square-integrable functions |
| $M^2_{\mathscr{F}}\big([0,T];\mathbb{R}^n\big)$ | the space of all $\{\mathscr{F}\}_{t\geq 0}$ predictable $\mathbb{R}^n$-valued processes $f:[0,T]\to\mathbb{R}^n$ such that $\mathbb{E}\Big[\int_0^T \|f(t)\|^2_{\mathcal{M}^2}\, dt\Big] < \infty$, |

## 2.2 Deterministic differential equations

**Definition 2.2.1** *(Differential Equations).*

*A differential equation is a relation of the form, $F(x,y,y',\cdots,y^{(n)})=0$, where $x$ is the independent variable, $y$ is the dependent variable, and $y^{(i)}$ are the successive derivatives of $y$ with respect to $x$.*

In applications, differential equations describe a relationship that involves some physical quantities (the functions) and their rates of change (derivatives).

**Example 1**: The price $X(t)$ per unit time of a riskless asset grows exponentially and satisfies the following differential equation

$$dX_0(t) = rX_0(t)dt,$$

where $r$ is a positive constant.

## 2.3 Probability theory refresher

To adequately model the financial market behavior, we need some tools from the theory of probability and stochastic analysis to capture the uncertainties.

### 2.3.1 Random experiment

A random experiment is an experiment whose outcomes are unpredictable. Example, tossing a coin could lead us with tail or head, but we cannot say exactly which face of the coin will come. Another example, in financial market the price of a stock is a random quantity, since the experiment of looking that value, let say every hour during an open day is a quantity that cannot be predicted in advance. Therefore, the price of stocks in a financial market forms a random experiment. Nevertheless, to work at ease with the notion of random experiment, we make the assumptions that all possible outcomes for a specific experiment are known, those outcomes could be finite or infinite.

### 2.3.2 Probability measure, probability space

Let us consider an experiment for which the set of all the possible outcomes is known, and let us denote by $\Omega$ that set. $\Omega$ is also called the sample space and could be countably finite or infinite, or $\Omega$ could be uncountable.

**Definition 2.3.1** *($\sigma$-algebra, Measurable space)*

*A $\sigma$-algebra denoted by $\mathscr{F}$ is a collection of subsets of $\Omega$ with the following properties.*

*(i)* $\emptyset, \Omega \in \mathscr{F}$.

*(ii)* *If* $A \in \mathscr{F}$, *then* $A^c \in \mathscr{F}$.

*(iii)* *If* $A_1, A_2, A_3, \cdots, A_n, \cdots \in \mathscr{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathscr{F}$.

*The pair* $(\Omega , \mathscr{F})$ *is called a measurable space.*

In 1933 the axiomatic approach to probability was formalized by the Russian mathematician A.N. Kolmogorov (cf. [19]). The probability theory system is constructed from those axioms (cf. [11]).

**Axioms of Probability**

Let $\Omega$ be a sample space, and $\mathscr{F}$ a $\sigma$-algebra associated to $\Omega$. A *probability measure* $\mathbb{P}$ on a measurable space $(\Omega , \mathscr{F})$ is a function $\mathbb{P}$: $\mathscr{F} \to [0, 1]$ satisfying the following axioms.

(i) $\mathbb{P}(\emptyset) = 0$ , $\mathbb{P}(\Omega) = 1$.

(ii) Let $A_1, A_2, \cdots, A_k, \cdots \in \mathscr{F}$, then $\mathbb{P}(\bigcup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} \mathbb{P}(A_k)$.

(iii) If $A_1, A_2, \cdots$ are mutually disjoint sets in $\mathscr{F}$, then $\mathbb{P}(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$.

If those three axioms are met, then the triple $(\Omega, \mathscr{F}, \mathbb{P})$ is called a *probability space.* Furthermore, $(\Omega, \mathscr{F}, \mathbb{P})$ is called a *complete probability space* if $\mathscr{F}$ contains all subsets A of $\Omega$ with $\mathbb{P}$-outer measure zero.

**Definition 2.3.2** *(Measurable Function)*

*Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space where $\Omega$ is a nonempty set, $\mathscr{F}$ a $\sigma$-algebra of subsets of $\Omega$, and $\mathbb{P}$ is a probability measure on $\mathscr{F}$.*

*Then a function $X\colon \Omega \to \mathbb{R}^n$ is called $\mathscr{F}$-measurable if its preimage belongs to $\mathscr{F}$, i.e., $X^{-1}\ (\mathcal{U}) \triangleq \{\omega \in \Omega : X\ (\omega) \in \mathcal{U}\} \in \mathscr{F}$, for any measurable set $\mathcal{U}$ in $\mathbb{R}^n$.*

**Definition 2.3.3** *(Conditional Probability)*

*Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space such that $\mathbb{P}\ (A)>0$, with A an event of $\Omega$. Let B be an event of $\Omega$: If the event A occurs and we want to use that information to find the probability of the event B, we write: $\mathbb{P}(B|A) = \mathbb{P}(A \cap B) \,/\, \mathbb{P}(A)$ called probability of B given A.*

*If A and B are independent then $\mathbb{P}(B|A) = \mathbb{P}\ (B)$.*

**Definition 2.3.4** *(Random Variable)*

*Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, a random variable X is an $\mathscr{F}$-measurable function, or Real valued function $X\colon \Omega \to \mathbb{R}$ such that its domain is the measurable space $(\Omega, \mathscr{F})$ and its range is the measurable space $(\mathbb{R}, \mathscr{B})$ where $\mathscr{B}$ is a $\sigma$-algebra defined on $\mathbb{R}$, satisfying the following: If $B \in \mathscr{B}$, then $X^{-1}\ (B) \in \mathscr{F}$.*

**Definition 2.3.5** *(Filtration)*

*A filtration* $\{\mathscr{F}_t\}_{t\geq 0}$ *on a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *is a collection of sub-sigma-algebras of* $\mathscr{F}$ *satisfying* $\mathscr{F}_s \subseteq \mathscr{F}_t$ *whenever* $s \leq t$. $\mathscr{F}_t$ *represents the set of events observable by the time* $t$. *The probability space taken together with the filtration* $\left(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ *is called a filtered probability space.*

## 2.4 Stochastic Processes

Nowadays, the description of several phenomena in many fields such as engineering, robotics, neuroscience, and finance often requires the use of mathematical tools related to random processes. In this section, we recall the basics of processes and stochastic differential equations.

**Definition 2.4.1** *(Stochastic Process) A stochastic process is a collection of random variables* $\{X(t)\}_{t\in T}$, *where* $T$ *is the parameter space such as* $[0, \infty)$ *or* $[a, b]$, $a, b \in \mathbb{R}$.

*Any (deterministic) function* $f(t)$ *can be considered as an elementary stochastic function process. A typical example that we find very often in physics, engineering, and finance models is the process of Wiener* $W_t$ *also known as Brownian motion.*

**Definition 2.4.2** *(Adapted Process)*

*Let*

- $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space;*
- *Time set* $T$ *be an index set with a total order;*

• $\mathscr{F} = \{\mathscr{F}_t\}_{t \in T}$ *be a filtration of the sigma-algebra* $\mathscr{F}$ *with* $\mathscr{F}_s \subseteq \mathscr{F}_t$ *if* $s < t$. $\mathscr{F}_t$ *is the information available at time* $t$*;*

• $X : T\ x\ \Omega \to \mathbb{R}$ *be a stochastic process.* $X$ *can be also write as* $X_t(\omega)$*, or* $X(t, \omega)$*.*

*The process* $X$ *is said to be adapted to the filtration* $\{\mathscr{F}_t\}_{t \in T}$ *if the random variable* $X_t$ *is* $\mathscr{F}_t$*-measurable for each* $t \in T$*: that is the value of* $X_t(\omega)$ *can be determined by the "information available at time* $t$*."*

**Definition 2.4.3** *(Standard Brownian Motion)*

*A stochastic process* $\{W(t)\}_{t \geq 0}$ *of real-valued random variables defined on some probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *is called a Brownian motion (or Wiener process) if, given any time points* $0 = t_0 < t_1 < t_2 < \cdots < t_n$*, we have the following properties in continuous time:*

**Property 1**. $W(0) = 0$.

**Property 2**. Independent increments:

$W(t_1) - W(t_0)$, $W(t_2) - W(t_1)$, $\cdots$, $W(t_n) - W(t_{n-1})$ are independent random variables.

**Property 3**. Normal distribution: $W(t_{i+1}) - W(t_i) \sim \mathscr{N}(0, t_{i+1} - t_i)$, for every $i$.

**Property 4**. The Wiener process $\{W(t)\}$ can be represented by continuous paths.

Let $\{\mathscr{F}(t)\}_{t \geq 0}$ be a filtration associated with a Brownian motion $W(t)$. $\{\mathscr{F}(t)\}_{t \geq 0}$ is a collection of $\sigma$-algebras such that:

i) For $0 \leq s < t$, $\mathscr{F}(s) \subset \mathscr{F}(t)$.

ii) $\{W(t)\}$ must be adapted to the filtration $\{\mathscr{F}(t)\}_{t\geq 0}$.

iii) For $0 \leq t < u$, $W(u) - W(t)$ is independent of $\mathscr{F}(t)$.

**Definition 2.4.4** *(Predictable Continuous-time Process)*

*Let $\left(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ a filtered probability space, then a continuous-time stochastic process $(X_t)_{t\geq 0}$ is predictable if $X$, considered as a mapping from $\Omega \times \mathbb{R}_+$ is measurable with respect to the sigma-algebra generated by all left-continuous adapted processes.*

### 2.4.1 Stochastic differential equations

The use of stochastic differential equations (SDE) for the modeling of financial quantities such as asset prices, interest rates and their derivatives has become a standard tool in applied mathematics. Unlike deterministic models which have a unique solution for each appropriate initial condition, stochastic differential equations have solutions which are continuous time stochastic processes. However, the resolution methods of the stochastic differential equations are based on similar techniques for Ordinary differential equations, but generalized to support stochastic dynamics.

**Definition 2.4.5** *(Stochastic Differential Equation).*

*In general, we define a stochastic differential equation, as a differential equation in which one or more of the terms is a stochastic process, resulting in a solution which is also a stochastic process.*

**Example 2**: Let $X(t)$ be a stochastic process. The dynamic of this process can be described as the following (SDE):

$$dX(t) = f\big(t, X(t)\big)dt + g\big(t, X(t)\big)dW(t),$$

where $W(t)$ is a Wiener process, f and g are some functions that satisfy some given conditions.

**Definition 2.4.6** *(Mathematical Expectation)*

*Let* $(\Omega, \mathscr{F}, \mathbb{P})$, *be a probability space, and* $X$ *a random variable defined by* $X : \Omega \to \mathbb{R}^n$. *We write* $\int_\Omega X d\mathbb{P} = \int_\Omega X^+ d\mathbb{P} - \int_\Omega X^- d\mathbb{P}$, *provided at least one of the integrals on the right is finite. Where* $X^+ = max(X, 0)$ *and* $X^- = max(-X, 0)$; *so that* $X = X^+ - X^-$. *We call* $\mathbb{E}(X) := \int_\Omega X d\mathbb{P}$ *the expected value of* $X$.

**Definition 2.4.7** *(Conditional Expectation)*

*Given a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$, *a random variable* $X : \Omega \to \mathbb{R}^n$ *such that* $\mathbb{E}[|X|] < \infty$, *and* $\mathcal{H} \subset \mathscr{F}$ *a* $\sigma$*-algebra, then the conditional expectation of* $X$ *given* $\mathcal{H}$, *is denoted by* $\mathbb{E}[X|\mathcal{H}] : \Omega \to \mathbb{R}^n$ *such that:*

- $\mathbb{E}[X|\mathcal{H}]$ *is* $\mathcal{H} -$ *measurable.*
- $\int_H \mathbb{E}[X|\mathcal{H}] d\mathbb{P} = \int_H X d\mathbb{P}$, *for all* $H \in \mathcal{H}$.

**Definition 2.4.8** *(Square-integrability)*

*A stochastic process* $X(t)$ *is called square-integrable if* $\mathbb{E}\big[\int_\tau^T |X(T)|^2 dt\big] < \infty$ *for any* $T > \tau$.

**Definition 2.4.9** *(Martingale)*

*An N-dimensional stochastic process* $\{M_t\}_{t\geq 0}$ *on a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$*, is called a martingale with respect to a filtration* $\{\mathscr{F}_t\}_{t\geq 0}$ *and a probability measure* $\mathbb{P}$ *if:*

- $M_t$ *is* $\mathscr{F}_t$*-measurable for all* $t$*.*
- $\mathbb{E}[|M_t|] < \infty$ *for all* $t$*.*
- $\mathbb{E}[M_s|M_t] = M_t$ *for any* $s \geq t$*.*

**Definition 2.4.10** *(Itô process)*

*An Itô process is a stochastic integral with respect to a Brownian motion defined on* $\left(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P}\right)$ *a filtered probability space. Let* $\left\{X(t)\right\}_{0\leq t\leq T}$ *be a stochastic process, adapted to the natural filtration* $\left\{\mathscr{F}(t)\right\}_{t\geq 0}$ *of a Brownian motion* $\left\{W(t)\right\}_{0\leq t\leq T}$*, that is* $X(t)$ *be* $\mathscr{F}(t)$*-measurable. We have the following Itô process.*

$$I(t) = \int_0^t X(\mathsf{u})dW(\mathsf{u}), \quad 0 \leq t \leq T.$$

## 2.5 Basic concepts on control optimization

An optimal control is a set of differential equations describing the paths of the control variables that optimize the cost functional. The method is largely due to the work of Richard Bellman and Lev Pontryagin in the 1950s. In economic or finance, we have two tools frequently used: the dynamic programming and the maximum principle.

• **Dynamic Programming**: This is a model developed by Richard Bellman. It is mostly used for discrete-time dynamic optimization. The method involves the optimal *value*. If the value depends on state $x \in \mathbb{R}$ and time $t$, and the optimal value is $V_t(x_t)$, then one trades off immediate payoff $f$ (direct utility) against future optimal value (indirect utility) $V_{t+1}(x_{t+1})$. If our control at time t is $u_t$,
$f := f(t, x, u)$ depends on time, state, and control, and so does $x_{t+1} = g(t, x, u)$, then the best we can do with state $x_t = x$ is to maximize $f(t, x, u) + V_{t+1}(g(t, x, u))$ with regard to our control $u$ if $V_{t+1}$ is a known function, that gives us the optimal $\bar{u}_t$ as a function of time and date.

• **Maximum principle**: This is a model developed by Lev Pontryagin. It is used in optimal control theory to find the best possible control for taking a dynamic system (time dependence) from one state to another, especially in the presence of constraints for the state or input controls.

The maximum principle will be the main focus of this thesis. In general terms, an optimal control problem consists of the following elements:

• **State process** $S(\cdot)$. This process must capture the minimal necessary information needed to describe the problem. Typically, $S(t) \in \mathbb{R}^d$ is influenced by the control and given the control process it has a Markovian structure. Usually its time dynamics is prescribed through an equation. In this section, we will consider only the state processes whose dynamics is described through an ordinary or a stochastic differential equation.

- **Control process** $u(\cdot)$. We need to describe the control set, $\mathcal{U}$, in which $u(t)$ takes values for every t. Applications dictate the choice of $\mathcal{U}$. In addition to this simple restriction $u(t) \in \mathcal{U}$, there could be additional constraints placed on control process. For instance, in the stochastic setting, we will require $v$ to be adapted to a certain filtration, to model the flow of information. Also we may require the state process to take values in a certain region (i.e., state constraint). This also places restrictions on the process $u(\cdot)$.

- **Admissible controls** $\mathscr{U}$. A control process satisfying the constraints is called an admissible control. The set of all admissible controls will be denoted by $\mathscr{U}$ and it may depend on the initial value of the state process.

- **Objective functional** $J\big(S(\cdot), u(\cdot)\big)$. This is the functional to be maximized (or minimized). In all of our applications, $J$ has an additive structure, or in other words $J$ is given as an integral over time.

In optimal control, the goal is to minimize (or maximize) the objective functional $J$ over all *admissible controls* by finding the minimizing (maximizing) control process.

### 2.5.1 Maximum principle

The principle states, informally, that the control Hamiltonian must take an extreme value over controls in the set of all permissible controls. Whether the extreme value is maximum or minimum depends on the problem and on the sign convention used for defining the Hamiltonian.

(a) **Necessary conditions.** Let the time frame $[t_0, t_1]$ be given. Consider the problem to maximize with regard to $u(t) \in U$ the functional $\int_{t_0}^{t_1} f(t, x(t), u(t))dt$ where $x$ starts at $x(t_0) = x_0$ (given) and evolves as $\dot{x}(t) = g(t, x(t), u(t))$; we shall consider the following three possible terminal conditions:

(i) $x(t_1) = x_1$ (given), (ii) $x(t_1) \geq x_1$, or (iii) $x(t_1)$ free.

**Step 1:** Form the Hamiltonian $H(t, x, u, p) = f(t, x, u) + pg(t, x, u)$.

**Step 2:** The optimal $\bar{u}$ maximizes $H$. ($p$ is the adjoint satisfying the next step.)

**Step 3:** We have a differential equation for $p$:

$\dot{p} = -\dfrac{\partial H}{\partial x}$ (evaluated at optimum), with the so-called *transversality conditions* on $p(t_1)$:

- In case the terminal value $x(t_1)$ is fixed, there is no condition on $p(t_1)$.
- In case the problem imposes $x(t_1) \geq x_1$, then we get a complementary slackness condition on $p(t_1)$: $p(t_1) \geq 0$.
- If there is no restriction on $x(t_1)$,then $p(t_1)$ must be equal to zero. If we have a function $\bar{u}(t, x, u)$ for the optimal control, then plugging this into $-\dfrac{\partial H}{\partial x}$ will give $\dot{p}$ as a function of $(t, x^*, p)$.

**Step 4:** We have the differential equation for the the state. Inserting $\bar{u}$ there as well, gives a differential equation system: $\begin{cases} dx^* = \phi(t, x^*, p) \\ dp = \psi(t, x^*, p) \end{cases}$, and the conditions on $x(t_0)$, $x(t_1)$, and $p(t_1)$ determine the integration constants.

(b) **Sufficient conditions.** We have two sets of sufficient conditions. Suppose we have found a pair $(\bar{x}, \bar{u})$ which satisfies the necessary conditions. This pair is a candidate for optimality. We can conclude that it is indeed optimal if it satisfies one of the following:

- **Mangasarian sufficiency condition**: For this condition, we use $p = p(t)$ produced by applying the maximum principle to obtain that $H$ is concave with regard to $(x, u)$ for all $t \in (t_0, t_1)$.
- **The Arrow sufficiency condition**: For this condition, we insert the function $\bar{u}(t, x, p)$ for $u$ in the Hamiltonian to get the function $\bar{H}(t, x, p) = H\big(t, x, \bar{u}(t, x, p), p\big)$, with the $p = p(t)$ that the maximum principle produces. Then $\bar{H}$ is concave with regard to $x$ for all $t \in (t_0, t_1)$.

The Arrow sufficiency condition is more powerful and mostly applied in economics problems.

**Example 3: Maximization problem.** Case of a deterministic optimal control problem using the Maximum principle steps.

Let maximize the function J given by: $J = \displaystyle\int_0^2 (2x - 3u - u^2)dt$ subject to: $\dot{x} = x + u$, $x(0) = 5$, and the control constraint $u \in \mathcal{U} = [0, 2]$. The optimal solution $u(t)$ of this problem can be obtained using the Pontryagin maximum principle as follows.

Step 1: We form the Hamiltonian.

$$H = (2x - 3u - u^2) + p(x + u) = (2 + p)x - (u^2 + 3u - pu)$$

Step 2: The optimal control can be calculated by differentiating $H$ with respect to $u$ and equating the result to zero as:

$$\frac{\partial H}{\partial u} = -2u - 3 + p = 0$$

which gives $\bar{u}(t) = \big(p(t) - 3\big)/2$, where $\bar{u}(t)$ must lie in the interval $\mathcal{U} = [0, 2]$.

Step 3: In order to obtain p we derive the adjoint equation as:

$$\dot{p} = -\frac{\partial H}{\partial x} = -2 - p, \quad p(2) = 0$$

or equivalently,

$$\dot{p} + p = -2, \quad p(2) = 0.$$

$p(2) = 0$ since there is no restriction on $x(2)$. The solution of the above equation is $p(t) = 2\big(\mathrm{e}^{2-t} - 1\big)$

Step 4: Considering the fact that the control must always lie in the interval $\mathcal{U} = [0, 2]$, this leads to the following optimal control:

$$u = \begin{cases} 2 & \text{if} \quad \mathrm{e}^{2-t} - 2.5 > 2, \\ \mathrm{e}^{2-t} - 2.5 & \text{if} \quad 0 \leq \mathrm{e}^{2-t} - 2.5 \leq 2, \\ 0 & \text{if} \quad \mathrm{e}^{2-t} - 2.5 < 0. \end{cases}$$

As it is expected, the resultant optimal control is a continuous function in $t$. it can be easily verified that the total cost $J$ using this optimal control is equal to 68.93.

### 2.5.2 Stochastic maximum principle

- **Problem formulation**.

In a stochastic control problem, the state equation is stochastic. Let $X_0$ be the initial state of an Itô controlled process that describes the state of a system in a form:

$$\begin{cases} dX(t) = f\big(t, X(t), u(t)\big)dt + g\big(t, X(t), u(t)\big)dW(t) \\ X(0) = X_0,\ t \in [0, T] \end{cases} \tag{2.5.1}$$

Where $X(t) \in \mathbb{R}^n$, $f : [0,T] \text{ x } \mathbb{R}^n \text{ x } \mathcal{U} \to \mathbb{R}^n$, $g : [0,T] \text{ x } \mathbb{R}^n \text{ x } \mathcal{U} \to \mathbb{R}^{n \times m}$ and $W(t)$ is m-dimensional Brownian motion. Here $u(t) \in \mathcal{U} \subset \mathbb{R}^k$ is a parameter which can be chosen in the given Borel set $\mathcal{U}$ at any time t in order to control the process $X(t)$. Therefore, $u(t) = u(t, w)$ is a stochastic process.

The cost functional $J\big((0, X_0, u(t)\big)$ associated with the initial condition $(0, X_0)$ is given by:

$$J(0, X_0, u) := \mathbb{E}^{X_0}\left[\int_0^T F\big(t, X(t), u(t)\big)dt + G\big(T, X(T)\big)\right] \tag{2.5.2}$$

Where $F : [0,T] \text{ x } \mathbb{R}^n \text{ x } \mathcal{U} \to \mathbb{R}$, $G : \mathbb{R}^n \to \mathbb{R}$ are given. We say that $u$ is an admissible control, if it is predictable such that: $\mathbb{E}\left[\int_0^T |u(t)|^2 dt\right] < \infty$ and the stochastic differential equation has a unique strong solution.

We denote by $\mathscr{U}[0,T]$ the set of all admissible controls. Our stochastic control problem is now defined as follows:

$$\begin{cases} \text{Minimize } \; \mathrm{J}\big(0, X(t), u(\cdot)\big) \\ \text{subject to } \begin{cases} u(\cdot) \in \mathscr{U}[0,T] \\ \big(X(\cdot), u(\cdot)\big) \text{ satisfies (2.5.1).} \end{cases} \end{cases} \tag{2.5.3}$$

We make further assumptions on the above functions.

$(A1)$: f, g are uniformly lipschitz in $(x,u)$ and $f(t,0,0)$ is bounded $\forall t \in [0,T]$;

$(A2)$: f, g, F, G are twice continuously differentiable with respect to $x$, they and their partial derivatives in $x$ are continuous in $(x,u)$;

$(A3)$: $f_x$, $f_{xx}$, $g_x$, $g_{xx}$, $F_x$, $F_{xx}$, $G_x$, $G_{xx}$ are bounded;

$(A4)$: $F$ and $G$ are uniformly bounded;

$(A5)$: $\mathcal{U}$ is a convex subset of $\mathbb{R}^k$

$(A1)$ - $(A3)$ are usual conditions for risk-neutral maximum principles. $(A4)$ ensure that the cost functional is well defined.

Now, we can present a general maximum principle control problem and also sufficient conditions for optimality.

In the following, we take $\varphi = f, g, F$ we define

$$\begin{cases} \bar{\varphi}(t) \triangleq \varphi\big(t, \bar{X}(t), \bar{u}(t)\big), \\ \bar{\varphi}_x(t) \triangleq \varphi_x\big(t, \bar{X}(t), \bar{u}(t)\big), \\ \bar{\varphi}_{xx}(t) \triangleq \varphi_{xx}\big(t, \bar{X}(t), \bar{u}(t)\big). \end{cases}$$

Let $\left(\bar{X}(\cdot), \bar{u}(\cdot)\right)$ be an admissible pair for the system (2.5.1). To construct our Hamiltonian we need to introduce the first order adjoint variable $\left(\bar{p}(\cdot), \bar{q}(\cdot)\right) \in L^2_{\mathscr{F}}\left([0,T]; \mathbb{R}^n\right) \times L^2_{\mathscr{F},p}\left([0,T]; \mathbb{R}^{n\times m}\right)$, and the second order adjoint variable $\left(\bar{P}(\cdot), \bar{Q}(\cdot)\right) \in L^2_{\mathscr{F}}\left([0,T]; \mathbb{R}^{n\,\times\,n}\right) \times \left(L^2_{\mathscr{F},p}\left([0,T]; \mathbb{R}^{n\,\times\,n}\right)\right)^m$ associated with the admissible pair $\left(\bar{X}(\cdot), \bar{u}(\cdot)\right)$, which are the solutions of the following equations respectively:

$$\begin{cases} d\bar{p}(t) = -\Big\{ \bar{f}_x(t)^{\mathsf{T}} \bar{p}(t) - \bar{F}_x(t)^{\mathsf{T}} + \sum_{j=1}^{m} \left[ \bar{g}^j_x(t)^{\mathsf{T}} \bar{q}_j(t) \right] \Big\} dt + \sum_{j=1}^{m} \bar{q}_j(t) dW_j(t) \\ p(T) = -G_x\left(\bar{X}(T)\right), \end{cases} \tag{2.5.4}$$

$$\begin{cases} d\bar{P}(t) = -\Big\{ \left(\bar{f}_x(t)^{\mathsf{T}} \bar{P}(t) + \bar{P}(t) \bar{f}_x(t)\right) + \sum_{j=1}^{m} \left(\bar{g}^j_x(t)^{\mathsf{T}} \bar{P}(t) \bar{g}^j_x(t)\right) \\ \qquad + \sum_{j=1}^{m} \left[ \left(\bar{g}^j_x(t)^{\mathsf{T}} \bar{Q}_j(t)\right) + \left(\bar{Q}_j(t) \bar{g}^j_x(t)\right) \right] + \bar{H}_{xx}\left(t, \bar{X}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t)\right\} dt \\ \qquad + \sum_{j=1}^{m} \bar{Q}_j(t) dW_j(t), \\ \bar{P}(T) = -G_{xx}\left(\bar{X}(T)\right). \end{cases} \tag{2.5.5}$$

The Hamiltonian $\bar{H} : [0,T] \times \mathbb{R}^n \times \mathcal{U} \times \mathbb{R}^n \times \mathbb{R}^{n\times m} \to \mathbb{R}$ is defined as:

$$\bar{H}(t, X_0, u, p, q) := \langle p, f(t, X_0, u) \rangle - F(t, X_0, u) + \sum_{j=1}^{m} g^j(t, X_0, u)^{\mathsf{T}} q_j$$

From (2.5.4) the unknown is the pair of processes $\left(\bar{p}(\cdot), \bar{q}(\cdot)\right)$ which is $\mathscr{F}$-adapted, and from (2.5.5) the unknown is the pair of processes $\left(\bar{P}(\cdot), \bar{Q}(\cdot)\right)$.

We note that equations (2.5.4) and (2.5.5) are backward stochastic differential equations (BSDE). Then, those equations associated with the 6-tuple $\big(\bar{X}, \bar{u}, \bar{p}(\cdot), \bar{q}(\cdot), \bar{P}(\cdot), \bar{Q}(\cdot)\big)$, define

$$\mathcal{H}(t,\bar{X},\bar{u}) := H(t,\bar{X},\bar{u},\bar{p},\bar{q}) + \frac{1}{2}\big\langle \bar{P}(t)\; g(t,X,u), g(t,X,u)\big\rangle - \big\langle \bar{P}(t)\bar{g}(t),\; g(t,X,u)\big\rangle.$$

The main result in [17] asserts that the optimal pair $\big(\bar{X}(\cdot), \bar{u}(\cdot)\big)$ verifies the following stochastic maximum principle.

$$\bar{\mathcal{H}}(t) = \mathcal{H}(t,\bar{X}(t),\bar{u}(t)) = \max_{u\in\mathcal{U}}\; \mathcal{H}(t,\bar{X}(t),u)\quad, a.e.t \in [0,T], \mathbb{P}-a.s.$$

There are two fundamental theorems used in stochastic control.

- **First fundamental theorem in stochastic control:**

This theorem is also called the necessary stochastic maximum principle, the theorem states that: If there exists an optimal control, then it is associated to the easier problem of finding the maximum of a certain real function in a particular control space.

- **Second fundamental theorem in stochastic control:**

The second fundamental theorem in stochastic control is the sufficient stochastic maximum principle. It states that if a certain real function is maximum for a particular control, then that control is optimal.

## 2.6 Markov regime switching

### 2.6.1 Markov chain

Let $S_t$ be a random variable that can assume only an integer $\{1, 2, \cdots, N\}$. Suppose that the probability that $S_t$ equals some particular value $j$ depends on the past only through the most recent value $S_{t-1}$. $P\{S_t = j|S_{t-1} = i, S_{t-2} = k, \cdots\}$ $= P\{S_t = j|S_{t-1} = i\} = p_{ij}$. Such a process is described as an N-state *Markov chain* with transition probabilities $\{p_{ij}\}_{i,j=1,2,\cdots,N}$.

The transition probability $p_{ij}$ gives the probability that: state $i$ will be followed by state $j$. Note that: $p_{i1} + p_{i2} + \cdots + p_{iN} = 1$.

It is often convenient to collect the transition probabilities in a $(N \times N)$ matrix $P$ known as the transition matrix (cf.[8]).

$$\begin{bmatrix} p_{11} & p_{21} & \cdots & p_{N1} \\ p_{12} & p_{22} & \cdots & p_{N2} \\ \vdots & \vdots & \vdots & \vdots \\ p_{1N} & p_{2N} & \cdots & p_{NN} \end{bmatrix}$$

In our approach, we will exploit the Markovian structure of the problem and use the Maximum Principle.

### 2.6.2 Markov-chain regime switching

- **Problem statement**.

A stochastic control problem in a regime switching diffusion model can be formulated as follows.

Let $X = \{X_t\}_{t\in[0,T]}$ be a continuous time, finite Markov chain. We identify the states of this process with the standard unit vector $e_i$ in $\mathbb{R}^N$ where $N$ is the number of states of the chain. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space. Let $T \in (0, \infty)$ be a fixed deterministic time. Let $W(\cdot)$ be an m-dimensional Brownian motion and $\alpha(\cdot)$ a continuous time finite state space Markov chain defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{\mathscr{F}_t\}_{t\in[0,T]}$ be the filtration generated by $W$ and the Markov chain $\alpha$,

$$\mathscr{F}_t := \sigma\big[\big(W(s),\ \alpha(s)\big) : 0 \leq s \leq t\big] \bigvee \mathcal{N}(\mathbb{P}),\ \forall t \in [0, T],$$

where $\mathcal{N}(\mathbb{P})$ is the collection of all $\mathbb{P}$-null sets in the probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Let the Markov chain $\alpha$ takes values in a finite state space $S = \{1, 2, \cdots, N\}$ and starts from initial state $i_0 \in S$ with a $N$ x $N$ generator matrix $G = \big(g_{ij}\big)_{i,j=1}^N$ . For each pair of distinct states $(i, j)$, the Markov chain is a point process, or a counting process,

$$N_{ij}(t) := \sum_{0<s\leq t} \chi[\alpha(s-) = i]\ \chi[\alpha(s) = j], \forall t \in [0, T],$$

where $\chi$ is an indicator function.

The process $N_{ij}(t)$ counts the number of jumps that the Markov chain $\alpha$ has made from state $i$ to state $j$ up to time $t$. Define the intensity process

$$I_{ij}(t) := g_{ij}\ \chi[\alpha(s-) = i].$$

If we compensate $N_{ij}(t)$ by $\displaystyle\int_0^t I_{ij}(s)ds$, then we have the process

$$M_{ij}(t) := N_{ij}(t) - \int_0^t I_{ij}(s)ds,$$

which is a discontinuous square-integrable martingale with initial value zero (cf.[18]).

Consider a stochastic control model where the state of the system is governed by a controlled Markovian regime-switching.

$$\begin{cases} dX(t) = f\big(t, X(t), u(t), \alpha(t-)\big)dt + g\big(t, X(t), u(t), \alpha(t-)\big)dW(t), \\ X(0) = X_0 \in \mathbb{R}^n,\ \alpha(0) = i_0 \in S, \end{cases} \tag{2.6.1}$$

where $u(t) = u(\omega, t)$ is a control process defined by: $u(t) : \Omega \text{ x } [0,T] \to \mathcal{U},\ \ \mathcal{U} \in \mathbb{R}^k$, $T \in (0, \infty)$, $f : [0,T] \times \mathbb{R}^n \times \mathbb{R}^k \times S \to \mathbb{R}^n$ and $g : [0,T] \times \mathbb{R}^n \times \mathbb{R}^k \times S \to \mathbb{R}^{n \times m}$ are giving continuous functions satisfying the following assumptions:

*(A1):* The maps $f$ and $g$ are measurable, and there exists a constant $L > 0$ such that for $\varphi = f, g$ we have:

$$\begin{cases} |\varphi(t, X, u, i) - \varphi(t, Y, v, i)| \leq L\big(|X - Y| + |u - v|\big), \\ \qquad \forall t \in [0,T], X, Y \in \mathbb{R}^n, u, v \in \mathcal{U}, i \in S, \\ |\varphi(t, 0, 0, i)| \leq L;\ \ \forall t \in [0,T], i \in S. \end{cases}$$

*(A2):* The maps $f$ and $g$ are $C^1$ in $x$ and there exists a constant $K > 0$ and a modulus of continuity $\bar{\omega} : [0, \infty) \to [0, \infty)$ such that:

$$\big|\varphi_x(t, X, u, i) - \varphi_x(t, Y, v, i)\big| \leq K|X - Y| + \bar{\omega}\big(d(u - v)\big),$$

$$\forall t \in [0,T], X, Y \in \mathbb{R}^n, u, v \in \mathcal{U}, i \in S,$$

where $\varphi_x(t, X, u, i)$ is the partial derivative of $\varphi$ with respect to $x$ at the point $(t, x, u, i)$.

We consider the cost functional:

$$J(0,X_0,u)=\mathbb{E}\left[\int_0^T F\big((t,X(t),u(t),\alpha(t)\big)dt+G\big((X(T),\alpha(T)\big)\right]$$
$$X(0)=X_0,\ \alpha(0)=i_0, \tag{2.6.2}$$

*(A3):* The maps $F$ and $G$ are measurable, and there exist constants $L1, L2>0$ such that:

$$\begin{cases} |F(t,X,u,i)-F(t,Y,v,i)|\leq \big[L1+L2(|X|+|u|+|v|)\big]|u-v|, \\ \qquad \forall t\in[0,T], X,Y\in\mathbb{R}^n, u,v\in\mathcal{U}, i\in S, \\ |F(t,0,0,i)|+|G(0,i)|<L1;\ \forall t\in[0,T], i\in S. \end{cases}$$

*(A4):* The maps $F$ and $G$ are $C^1$ in $x$ and there exists a constant $K>0$ and a modulus of continuity $\bar{\omega}:[0,\infty)\to[0,\infty)$ such that for $\varphi=F,G$ we have:

$$\begin{cases} \big|\varphi_x(t,X,u,i)-\varphi_x(t,Y,v,i)\big|\leq K|X-Y|+\bar{\omega}\big(d(u,v)\big), \\ \qquad \forall t\in[0,T], X,Y\in\mathbb{R}^n, u,v\in\mathcal{U}, i\in S, \\ \big|\varphi_x(t,0,0,i)\big|\leq K;\ \forall t\in[0,T],\ \ i\in S, \end{cases}$$

where for each $i\in S$ we have that $F(\cdot,\cdot,\cdot,i):[0,T]\text{ x }\mathbb{R}^n\text{ x }\mathcal{U}\to\mathbb{R}$ is continuous and $G(\cdot,i):\mathbb{R}^n\to\mathbb{R}$ is $C^1(\mathbb{R})$ and concave.

*(A5):* suppose that $G(\cdot,\alpha(T))$ is convex and the Hamiltonian $H(t,\cdot,\cdot,\alpha(t-),\bar{p}(t),\bar{q}(t))$ is concave for all $t\in[0,T]$ almost surely.

Our equation (2.6.1) admits a unique solution and the cost functional is well defined. The control process $u$ is said admissible if it is valued in $\mathcal{U}$, a non-empty closed convex subset of $\mathbb{R}^k$.

Let $X(t) = X^{(u)}(t),\ t \in [0, T]$ satisfying both $X(0) = X_0$ almost surely, and

$$\mathbb{E}\left[\int_0^T F\big((t, X(t), u(t), \alpha(t)\big)dt + G\big((X(T), \alpha(T)\big)\right] < \infty.$$

Let us denoted by $\mathcal{U}_{ad}$ the set of admissible controls. If $X$ is a solution of the equation (2.6.1) with the corresponding admissible control $u \in \mathcal{U}_{ad}$, then we call $(X, u)$ an admissible pair and $X$ an admissible state process. Our optimal stochastic control problem is to find an optimal control $\bar{u}$ such that

$$J(0, X_0, \bar{u}) = \inf_{u \in \mathcal{U}_{ad}} J(0, X_0, u) \tag{2.6.3}$$

The corresponding $\bar{X}$ and $(\bar{X}, \bar{u})$ are called an optimal state process and optimal pair, respectively.

## 2.7 Adjoint variable and Hamiltonian

The Hamiltonian $H : [0, T] \times \mathbb{R}^n \times \mathcal{U} \times S \times \mathbb{R}^n \times \mathbb{R}^{n \times n} \to \mathbb{R}$ is defined by:

$$H(t, X_0, u, i, p, q) := -F(t, X_0, u, i) + f^T(t, X_0, u, i)p + tr\big(g^T(t, X_0, u, i)q\big). \tag{2.7.1}$$

We assume that the Hamiltonian $H$ is differentiable with respect to x.

Given an admissible pair $(\bar{X}, \bar{u})$, the adjoint equation corresponding in the adapted process $\bar{p}(t) \in \mathbb{R}^n$, $\bar{q}(t) \in \mathbb{R}^{n \times n}$ and $\eta(t) = \big(\eta^{(1)}(t), \cdots, \eta^{(N)}(t)\big)^T$.

Where $\eta^{(n)} \in \mathbb{R}^{N \times N}$ for $n = 1, \cdots, N$, is the backward stochastic differential equation (BSDE).

$$
\begin{cases} d\bar{p}(t) = -\bar{H}_x\big((t, \bar{X}(t), \bar{u}(t), \alpha(t-), \bar{p}(t), \bar{q}(t)\big)dt + \bar{q}(t)dW(t) + \eta(t) \bullet dM(t) \\ \bar{p}(T) = -G_x\big(\bar{X}(T), \alpha(T)\big), \end{cases} \tag{2.7.2}
$$

$$
\eta(t) \bullet dM(t) := \left( \sum_{j \neq i} \eta_{ij}^{(1)}(t)dM_{ij}(t), \cdots, \sum_{j \neq i} \eta_{ij}^{(N)}(t)dM_{ij}(t) \right)^T, \quad \forall t \in [0, t)
$$

and $\sum_{j \neq i}$ stands for $\sum_{i=1}^{N}\sum_{j=1}^{N}$ with $j \neq i$.

**Remark**: There are jumps in the adjoint equation, even though there are no jumps in the equation (2.6.1) which governs the state variable $X(t)$. This is a consequence of the coefficients $f(t)$ and $g(t)$ being functions of the Markov chain $\alpha(t)$. Moreover, the unknown process $\eta(t)$ in the adjoint equations does not appear in the Hamiltonian (2.7.1).

**Theorem 2.7.1** *(Necessary stochastic maximum principle regime-rwitching, cf.[13]) Let assumptions (A1)-(A4) hold and let* $(\bar{X}, \bar{u})$ *be an optimal pair of* (2.6.1)*. Then, there exists a triplet stochastic process* $\big(\bar{p}(\cdot), \bar{q}(\cdot), \bar{s}(\cdot)\big)$ *which is an adapted solution to* (2.7.2).

**Theorem 2.7.2** *(Sufficient stochastic maximum principle regime-rwitching, cf.[13]) Suppose that (A1)-(A5) hold. Let* $\big(\bar{x}(\cdot), \bar{u}(\cdot)\big)$ *be an admissible pair, and* $\big(\bar{p}(\cdot), \bar{q}(\cdot)\big), \big(\bar{s}(\cdot)\big)$ *being the solution of the corresponding adjoint equation* (2.7.2) *satisfying*

$$\mathbb{E}\int_0^T \|\big(f\big((t,\bar{X}(t),\bar{u}(t)) - g(t,X(t),u(t))\big)^T\bar{p}(t)\|^2 dt < \infty, \tag{2.7.3}$$

$$\mathbb{E}\int_0^T \|\bar{q}(t)^T\big(\bar{X}(t) - X(t)\big)\|^2 dt < \infty, \tag{2.7.4}$$

and

$$\sum_{n=1}^{N}\sum_{j\neq i}^{N}\mathbb{E}\int_0^T \big|\big(\bar{X}_n(t) - \bar{X}_n(t)\big)\bar{s}_{ij}^n(t)\big|^2 d\big\langle M_{ij}\big\rangle < \infty, \tag{2.7.5}$$

for all admissible controls $u(\cdot) \in \mathcal{U}$,

then, $\big(\bar{x}(\cdot),\bar{u}(\cdot)\big)$ is an optimal pair for the problem (2.6.3).

# 3 Risk-sensitive stochastic control optimization

After exploring the basic concepts of control optimization, we add in this chapter a new parameter in stochastic optimal control problems: the risk-sensitivity. Risk-sensitivity can be seen as the behavior of an investor towards risk. Therefore, an optimal risk-sensitive control, from this perspective, will be the balance of an investor's interest in optimizing the cost functional against his aversion to risk due to deviations of the realized cost rate from the expectation.

## 3.1 Statement of the problem

Let $T > 0$ be a fixed time horizon and $\big(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P}\big)$ a filtered probability space on which a n-dimensional standard Brownian motion $W = \{W_t\}_{t\geq 0}$ is given, and the filtration $\{\mathscr{F}_t\}_{t\geq 0}$ is the natural filtration of $W$ augmented by $\mathbb{P}$-null sets of $\mathscr{F}$. We consider the following stochastic control system:

$$\begin{cases} dX(t) = f\big(t, X(t), u(t)\big)dt + g\big(t, X(t), u(t)\big)dW(t), \\ X(0) = X_0, \end{cases} \tag{3.1.1}$$

where

$$f(t, X, u),\ g(t, X, u) : [0, T] \times \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^n,\ t \in [0, T],\ X \in \mathbb{R}^n,\ u \in \mathcal{U}.$$

An admissible control $u$ is an $\mathscr{F}$-adapted and square-integrable process with values in a non-empty subset $\mathcal{U}$ of $\mathbb{R}^m$. Let $\mathscr{U}$ be the set of all admissible controls.

Given $u \in \mathscr{U}$, equation (3.1.1) is a stochastic differential equation with random coefficients.

The risk-sensitive cost functional associated with (3.1.1) is given by:

$$J^\theta\big(0, X_0, u(\cdot)\big) = \mathbb{E}\Big[e^{\theta\left[\int_0^T F\left(t,X(t),u(t)\right)dt+G\left(X(T)\right)\right]}\Big], \tag{3.1.2}$$

where, $\theta$ is the risk-sensitivity parameter,

$$F(t,X,u) : [0,T] \times \mathbb{R}^n \times \mathcal{U} \to \mathbb{R},\ G(X) : \mathbb{R}^n \to \mathbb{R},\ t \in [0,T],\ X \in \mathbb{R},\ u \in \mathcal{U}.$$

The risk-sensitive control problem associated with (3.1.1) – (3.1.2) is defined as follows:

$$\begin{cases} \text{Minimize} \quad J^\theta\big(0, X_0; u(\cdot)\big) \\ \text{subject to} \begin{cases} u(\cdot) \in \mathscr{U}[0,T] \\ \big(X(\cdot), u(\cdot)\big)\ satisfies\ (3.1.1) \end{cases} \end{cases} \tag{3.1.3}$$

Any $\bar{u}(\cdot) \in \mathscr{U}[0,T]$ satisfying

$$J^\theta\big(0, X_0, \bar{u}(\cdot)\big) = \inf_{u(\cdot)\in\mathscr{U}[0,T]} J^\theta\big(0, X_0, u(\cdot)\big) \tag{3.1.4}$$

is called a risk sensitive optimal control. The corresponding state process, solution (3.1.1), is denoted by $\bar{X}(\cdot) := \bar{X}^{\bar{u}}(\cdot)$.

Our next step is to characterize the pair $(\bar{X}, \bar{u})$ solution of problem (3.1.4).

Let $\quad \Gamma_T := \displaystyle\int_0^T F\big(t, X(t), u(t)\big)dt + G\big(X(T)\big).$

Then the risk sensitive functional is given by:

$$\Gamma_\theta := \frac{1}{\theta} \log \mathbb{E}\Big[e^{\theta\left[\int_0^T F\left(t,X(t),u(t)\right)dt+G\left(X(T)\right)\right]}\Big] = \frac{1}{\theta} \log \mathbb{E}[e^{\theta\Gamma_T}].$$

When the risk-sensitive parameter $\theta$ is small, the functional $\Gamma_\theta$ around $\theta = 0$ can be asymptotically expanded as:

$$\begin{aligned}
\Gamma_\theta &= \frac{1}{\theta}\log\,\mathbb{E}\Big[1+\theta\Gamma_T+\frac{\theta^2}{2!}\Gamma_T^2+\cdots\Big]\\
&= \frac{1}{\theta}\log\Big[1+\theta\mathbb{E}[\Gamma_T]+\frac{\theta^2}{2}\mathbb{E}[\Gamma_T^2]+\cdots\Big]\\
&= \frac{1}{\theta}\log\Big[1+\big(\theta\mathbb{E}[\Gamma_T]+\frac{\theta^2}{2}\mathbb{E}[\Gamma_T^2]\big)+\cdots\Big]\\
&= \frac{1}{\theta}\Big[\big(\theta\mathbb{E}[\Gamma_T]+\frac{\theta^2}{2}\mathbb{E}[\Gamma_T^2]\big)-\frac{1}{2}\Big(\theta\mathbb{E}[\Gamma_T]+\frac{\theta^2}{2}\mathbb{E}[\Gamma_T^2]\Big)^2+\cdots\Big]\\
&= \mathbb{E}[\Gamma_T]+\frac{\theta}{2}\mathbb{E}[\Gamma_T^2]-\frac{\theta}{2}\big(\mathbb{E}[\Gamma_T]\big)^2\\
&= \mathbb{E}[\Gamma_T]+\frac{\theta}{2}\mathrm{Var}(\Gamma_T)+O(\theta^2),
\end{aligned}$$

where $\mathrm{var}(\Gamma_T)$ denotes the variance of $\Gamma_T$. If $\theta < 0$, the function $\Gamma_\theta$ increases and the optimizer is called *risk seeker*. If $\theta > 0$, the function $\Gamma_\theta$ decreases and the optimizer is called *risk averse*. If $\theta = 0$, we evaluate the limit of the function $\Gamma_\theta$ when $\theta \to 0$ and the optimizer is called *risk- neutral*.

**Assumptions:**

*(A1):* $\mathcal{U}$ is a separable metric space and $T > 0$.

*(A2):* The maps $f : [0,T]\times\mathbb{R}^n\times\mathcal{U}\to\mathbb{R}^n$, $\;g : [0,T]\times\mathbb{R}^n\times\mathcal{U}\to\mathbb{R}^n$, $F : [0,T]\times\mathbb{R}^n\times\mathcal{U}\to\mathbb{R}$ and $G : \mathbb{R}^n\to\mathbb{R}$ are measurable, and there exists a constant $K > 0$ and a modulus $\bar{\omega} : [0,\infty)\to[0,\infty)$ such that for

$\varphi(t,X,u) = f(t,X,u), g(t,X,u), F(t,X,u), G(X),$

$$
\begin{cases}
|\varphi(t,X,u)-\varphi(t,Y,v)| \le K|X-Y| + \bar{\omega}(d(u,v)), \\
\qquad \forall t\in[0,T], X,Y\in\mathbb{R}^n, u,v\in\mathcal{U}, \\
|\varphi(t,0,u)| \le K;\ \forall t\in[0,T], u\in\mathcal{U}.
\end{cases}
$$

Also, $F$, and $G$ are uniformly bounded.

*(A3):* $f, F, G$ are $C^2$ in $x$, and there exists a modulus of continuity $\bar{\omega}:[0,\infty)\to[0,\infty)$ such that for $\varphi(t,X,u) = f(t,X,u), F(t,X,u), G(X)$,

$$
\begin{cases}
|\varphi_x(t,X,u)-\varphi_x(t,Y,v)| \le K|X-Y| + \bar{\omega}(d(u,v)), \\
|\varphi_{xx}(t,X,u)-\varphi_{xx}(t,Y,v)| \le K|X-Y| + \bar{\omega}(d(u,v)), \\
\qquad \forall t\in[0,T], X,Y\in\mathbb{R}^n, u,v\in\mathcal{U}.
\end{cases}
$$

*(A4):* $J^{\theta}\big(0, X_0, \bar{u}(\cdot)\big) \in C^{1,3}\big([0,T]\times\mathbb{R}^n\big)$.

For the maximum principle to be sufficient, we need the following additional assumption.

*(A5):* $\mathcal{U}$ is a convex subset of $\mathbb{R}^k$. The maps $f, g$, and $F$ are locally Lipschitz in $u$, and their derivatives in $x$ are continuous in $(x,u)$.

Let $\Phi = f, g, F, G$, we define the following.

$$
\begin{cases}
\delta\Phi(t) = \Phi(t, \bar{X}(t), u(t)) - \Phi(t, \bar{X}(t), \bar{u}(t)), \\
\bar{\Phi}(t) = \Phi\big(t, \bar{X}(t), \bar{u}(t)\big), \\
\bar{\Phi}_x(t) = \dfrac{\partial \Phi}{\partial x}(t, \bar{X}(t), \bar{u}(t), \\
\bar{\Phi}_{xx}(t) = \dfrac{\partial^2 \Phi}{\partial x^2}(t, \bar{X}(t), \bar{u}(t),
\end{cases}
$$

where $u$ is an admissible control from $\mathcal{U}$.

## 3.2 Adjoint variables and Hamiltonian

Let $\big(\bar{X}(\cdot), \bar{u}(\cdot)\big)$ be an admissible pair for the system (3.1.1). We introduce the first order adjoint variable $\big(\bar{p}(\cdot), \bar{q}(\cdot)\big) \in L^2_{\mathcal{F}}([0,T];\mathbb{R}^n) \times L^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$ and the second order adjoint variable $\big(\bar{P}(\cdot), \bar{Q}(\cdot)\big) \in L^2_{\mathcal{F}}([0,T];\mathbb{R}^{n\times n}) \times L^2_{\mathcal{F}}([0,T];\mathbb{R}^{n\times n})$ associated with the admissible pair $\big(\bar{X}(\cdot), \bar{u}(\cdot)\big)$, which are the solutions of the following first order and second order adjoint equations respectively:

$$
\begin{cases}
d\bar{p}(t) = -\Big[\bar{f}_x(t)^T \bar{p}(t) - \bar{F}_x(t)^T - \theta\, \bar{p}(t)^T \bar{g}(t) \bar{g}_x(t)^T \bar{p}(t) - \theta\, \bar{p}(t)^T \bar{g}(t) \bar{q}(t) \\
\qquad\qquad + \bar{g}_x(t)^T \bar{q}(t)\Big] dt + \bar{q}(t) dW(t), \\
\bar{p}(T) = -G_x\big(\bar{X}(T)\big),
\end{cases}
\tag{3.2.1}
$$

$$
\begin{cases}
d\bar{P}(t) = -\Big\{ \bar{f}_x(t)^T \bar{P}(t) + \bar{P}(t)\bar{f}_x(t) + \bar{g}_x(t)^T \big[\bar{P}(t) - \theta\, \bar{p}(t)\bar{p}(t)^T\big]\bar{g}_x(t) \\
\qquad\qquad +\bar{g}_x(t)^T \big[\bar{Q}(t) - \theta\, \bar{p}(t)\bar{q}(t)^T - \theta\, \bar{p}(t)^T \bar{g}(t)\bar{P}(t)\big] \\
\qquad\qquad +\big[\bar{Q}(t) - \theta\, \bar{q}(t)\bar{p}(t)^T - \theta\, \bar{p}(t)^T \bar{g}(t)\bar{P}(t)\big]\bar{g}_x(t) \\
\qquad\qquad -\theta\, \bar{p}(t)^T \bar{g}(t)\bar{Q}(t) - \theta\, \bar{q}(t)\bar{q}(t)^T \\
\qquad\qquad +p(t)\, f_{xx}(t) \ + \big(q + \theta\, p^T p\big)\, g_{xx}(t) \ - F_{xx}(t)\big)\Big\} dt \\
\qquad\qquad +\bar{Q}(t) dW(t), \\
\bar{P}(T) = -G_{xx}\big(\bar{X}(T)\big),
\end{cases}
\tag{3.2.2}
$$

Where $\bar{f}_x(t) := f_x\big(t, \bar{X}(t), \bar{u}(t)\big)$ (with similar interpretations for $\bar{g}_x(t), \bar{F}_x(t)$, etc.)

We define the risk-neutral Hamiltonian associated with random variables $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ as follows. For $(p, q) \in \mathbb{R}^n \times \mathbb{R}^n$

$$
H(t, X, u, p, q) := \big\langle p, f(t, X, u)\big\rangle \ + tr\big(q^T\ g(t, X, u)\big) \ - F(t, X, u). \tag{3.2.3}
$$

We also introduce the risk-sensitive Hamiltonian for $\theta \in \mathbb{R}$ and $(p, q) \in \mathbb{R}^n \times \mathbb{R}^n$, by:

$$
H^\theta(t, X, u, p, q) := \big\langle p, f(t, X, u)\big\rangle + g(t, X, u)^T \big[q - \theta p p^T\ \bar{g}(t)\big] - F(t, X, u). \tag{3.2.4}
$$

We denote (following [6]):

$$\begin{cases} H_x(t) := f_x(t)p \ + g_x(t)q \ - F_x(t), \\ \\ H_x^\theta(t) := p(t)\ f_x(t) \ + \ g_x(t)\big(q - \theta\ pp^T \bar{g}_x(t)\big) \ - F_x(t), \\ \\ H_{xx}(t) := f_{xx}(t)p \ + g_{xx}(t)q \ - F_{xx}(t), \\ \\ H_{xx}^\theta(t) := p(t)\ f_{xx}(t) \ + \ g_{xx}(t)\big(q - \theta\ pp^T \bar{g}_{xx}(t)\big) \ - F_{xx}(t), \end{cases}$$

Let us consider that our assumptions are sufficient to guarantee the existence of unique solutions $\big(\bar{p}(\cdot), \bar{q}(\cdot)\big) \in L^2_{\mathscr{F}}(0, T; \mathbb{R}^n) \ \times \ L^2_{\mathscr{F}}(0, T; \mathbb{R}^n)$ and the second order adjoint variable $\big(\bar{P}(\cdot), \bar{Q}(\cdot)\big) \in L^2_{\mathscr{F}}(0, T; \mathbb{R}^{n \times n}) \ \times \ L^2_{\mathscr{F}}(0, T; \mathbb{R}^{n \times n})$ of (3.1.5) and (3.1.6) respectively.

Let the $\mathcal{H}$-function $\bar{\mathcal{H}}^\theta : \mathbb{R} \ \times \ \mathbb{R}^n \ \times \ \mathbb{R}^m \to \mathbb{R}$, associated with the pair $\big(\bar{X}(\cdot), \bar{u}(\cdot)\big)$, be defined as:

$$\begin{cases} \bar{\mathcal{H}}^\theta(t, X, u) := f(t, X, u)\bar{p}(t) - F(t, X, u) + \dfrac{1}{2} g((t, X, u)^T \times \\ \\ \qquad \big[\bar{P}(t) - \theta\bar{p}(t)\bar{p}(t)^T\big] g(t, X, u) \\ \\ + g(t, X, u)^T \big[\bar{q}(t) - \bar{P}(t) g\big(t, \bar{X}(t), \bar{u}(t)\big)\big]. \end{cases} \tag{3.2.5}$$

Under these assumptions, we present a maximum principle for the risk-sensitive control problem (3.1.3) as well as sufficient conditions for optimality.

**Theorem 3.2.1** *(Risk-sensitive maximum principle, cf.[13])*

*Suppose that (A1)-(A4) hold. Let* $\big(\bar{X}(\cdot), \bar{u}(\cdot)\big)$ *be an optimal pair for the risk-sensitive optimal control problem* (3.1.3).

*Then, there are unique solutions* $\big(\bar{p}(\cdot),\bar{q}(\cdot)\big) \in L^2_{\mathscr{F}}(0,T;\mathbb{R}^n) \times L^2_{\mathscr{F}}(0,T;\mathbb{R}^n)$ *and* $\big(\bar{P}(\cdot),\bar{Q}(\cdot)\big) \in L^2_{\mathscr{F}}(0,T;\mathbb{R}^{n\times n}) \times L^2_{\mathscr{F}}(0,T;\mathbb{R}^{n\times n})$ *of the first order and second order adjoint equations* (3.2.1) *and* (3.2.2) *respectively, such that:*

$$\begin{aligned} &\bar{H}^{\theta}\big(t,\bar{X}(t),\bar{u}(t),\bar{p}(t),\bar{q}(t)\big) - \bar{H}^{\theta}\big(t,\bar{X}(t),u(t),\bar{p}(t),\bar{q}(t)\big) \\ &-\frac{1}{2}\big[g(t,\bar{X}(t),\bar{u}(t)) - g(t,\bar{X}(t),u(t))\big]^T \times \ \big[\bar{P}(t) - \theta\bar{p}(t)\bar{p}(t)^T\big]\times \\ &\big[g(t,\bar{X}(t),\bar{u}(t)) - g(t,\bar{X}(t),u(t))\big] \geq 0, a.e. t \in [0,T], \quad P-a.s., \end{aligned} \tag{3.2.6}$$

**Theorem 3.2.2** *(Sufficient conditions for optimality, cf.[13])*

*Suppose that (A1)-(A5) hold. Let* $\big(\bar{X}(\cdot),\bar{u}(\cdot)\big)$ *be an admissible pair, and* $\big(\bar{p}(\cdot),\bar{q}(\cdot)\big), \big(\bar{P}(\cdot),\bar{Q}(\cdot)\big)$ *the associated first and second order adjoint variables respectively. We suppose that* $G(\cdot)$ *is convex,* $\bar{H}^{\theta}(t,\cdot,\cdot,\bar{p}(t),\bar{q}(t))$ *is concave for all* $t \in [0,T]$ *almost surely, and that* (3.2.6) *is satisfied. Then* $\big(\bar{X}(\cdot),\bar{u}(\cdot)\big)$ *is an optimal pair for the problem* (3.1.3).

# 4 Risk-sensitive stochastic optimal control with Markov regime switching

In this chapter we present and formulate a model of maximum principle for stochastic optimization control problem of Markov regime-switching with risk-sensitivity.

## 4.1 Problem Statement

1) Let $T > 0$ be a fixed time horizon.

2) Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space.

3) Let $W(\cdot)$ be a standard Brownian motion defined on $(\Omega, \mathscr{F}, \mathbb{P})$ over $[0, T]$ (with $W(0) = 0$, P-a.s). $W = \{W(t)\}_{t\in[0,T]}$

4) Let $\{\mathscr{F}_t\}_{t\in[0,T]}$ be the natural filtration of $W$ augmented by all the $\mathbb{P}$-null sets in $\mathscr{F}$.

5) Let $u : [0, T] \times \Omega \to \mathcal{U}$ be an $\{\mathscr{F}_t\}_{t\in[0,T]}$ adapted process on $(\Omega, \mathscr{F}, \mathbb{P})$

6) Let $X = \{X(t)\}_{t\in[0,T]}$ be a continuous time, controlled Markov regime-switching.

7) Let $\alpha(\cdot)$ be a continuous time finite state space Markov chain defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$. $\alpha := \{\alpha(t)\}_{t\in[0,T]}$ is an irreducible homogeneous continuous-time Markov chain with finite state space $\mathcal{S} = \{e_1, e_2, \cdots, e_N\} \subset \mathbb{R}^N$, where $N \in \mathbb{N}$, and jth component of $e_i$ is the Kronecker delta $\delta_{ij}$ for each
$i, j = 1, 2, \cdots, N$.

The Markov chain is characterized by an intensity matrix

$\Lambda := \{\lambda_{ij} : \ 1 \leq i, \ j \leq N\}$ under $\mathbb{P}$.

For each $1 \leq i, \ j \leq N\}$, $\lambda_{ij}$ is the constant transition intensity of the chain from state $e_i$ to state $e_j$ at time t. In addition for $i \neq j$, $\lambda_{ij} \geq 0$ and $\sum_{j=1}^{N} \lambda_{ij} = 0$, therefore $\lambda_{ii} \leq 0$. We then have from [7] the dynamics of $\alpha$ given by:

$$\alpha(t) = \alpha(0) + \int_0^t \Lambda^T \alpha(s) ds + M(t), \tag{4.1.1}$$

where $M := \{M(t)\}_{t \in [0,T]}$ is a $\mathbb{R}^N$- valued $(\mathbb{F}, \mathbb{P})$-martingale and $\Lambda^T$ is the transpose of the matrix $\Lambda$. We consider the following stochastic optimization control system. For each $1 \leq i, j \leq N$, with $i \neq j$, and $t \in [0, T]$, let $J^{ij}(t)$ be the number of jumps from state $e_i$ to state $e_j$ up to time t. From [7], it follows that:

$$J^{ij}(t) = \lambda_{ij} \int_0^t \big\langle \alpha(s-), e_i \big\rangle ds + m_{ij}(t), \tag{4.1.2}$$

with $m_{ij} := \{m_{ij}(t)\}_{t \in [0,t]}$, where $m_{ij}(t) := \int_0^t \big\langle \alpha(s-), e_i \big\rangle \big\langle dM(s), e_j \big\rangle$ is a $(\mathscr{F}, \mathbb{P})$-martingale.

Fix $j \in \{1, 2, \cdots, N\}$ and let $\Phi_j(t)$ be the number of jumps into state $e_j$ up to time t. Then

$$\Phi_j(t) := \sum_{i=1, i \neq j}^{N} J^{ij}(t) = \sum_{i=1, i \neq j}^{N} \lambda_{ij} \int_0^t \big\langle \alpha(s-), e_i \big\rangle ds + \tilde{\Phi}_j(t) = \lambda_j(t) + \tilde{\Phi}_j(t), \tag{4.1.3}$$

where $\tilde{\Phi}_j(t) = \sum_{i=1, i \neq j}^{N} mij(t)$ and $\lambda_j(t) = \sum_{i=1, i \neq j}^{N} \lambda_{ij} \int_0^t \big\langle \alpha(s-), e_i \big\rangle ds$.

It is important to mention that for each $j \in \{1, 2, \cdots, N\}, \tilde{\Phi}_j := \{\tilde{\Phi}_j(t)\}_{t\in[0,t]}$ is a $(\mathscr{F}, \mathbb{P})$-martingale.

Suppose that the state process $X(t) = X^{(u)}(t, w);\ 0 \le t \le T,\ w \in \Omega$ is a controlled Markov regime-switching diffusion of the form

$$\begin{cases} dX(t) = f\big(t, X(t), u(t), \alpha(t)\big)dt + g\big(t, X(t), u(t), \alpha(t)\big)dW(t) \\ \qquad\qquad + \gamma\big(t, X(t-), u(t-), \alpha(t-)\big)d\tilde{\Phi}(t), \\ X(0) = X_0, \qquad t \in [0, T]. \end{cases} \tag{4.1.4}$$

Where the functions $f, g$ and $\gamma$ are given such that for all $t,\ f(t, x, e_n, u)$, $g(t, x, e_n, u),$ and $\gamma(t, x, e_n, u),$ for $n = 1, 2, \cdots, N$ are $\mathscr{F}_t$- measurable continuous functions, for all $x \in \mathbb{R}$.

Here, $f : [0, T] \times \mathbb{R}^n \times \mathcal{U} \times \mathcal{S} \to \mathbb{R}^n,\ g : [0, T] \times \mathbb{R}^n \times \mathcal{U} \times \mathcal{S} \to \mathbb{R}^{n \times m}$, $\gamma : [0, T] \times \mathbb{R}^n \times \mathcal{U} \times \mathcal{S} \to \mathbb{R}^{n \times N}$. $W(t) := (W_1(t), W_2(t), \cdots, W_N(t))^T$ is an N-dimensional standard Brownian motion, and $\tilde{\Phi}(t) := (\tilde{\Phi}_1(t), \cdots, \tilde{\Phi}_N(t))^T$ with $\tilde{\Phi}_j(t),\ j = 1, 2, \cdots, N,$ defined by (4.1.3).

The cost functional $J^\theta(0, X, e_i; u(\cdot))$ associated with the initial condition $(0, X_0, e_i) \in [0, T] \times \mathbb{R}^n \times \mathcal{S}$ and control process $u(\cdot) \in \mathcal{U}$ is given by:

$$J^\theta(0, X, e_i; u(\cdot)) := \mathbb{E}\Big[e^{\theta\left[\int_0^T F(t, X(t), u(t), \alpha(t))dt + G(X(T), \alpha(T))\right]}\Big], \tag{4.1.5}$$

where $F : [0, T] \times \mathbb{R}^n \times \mathcal{U} \times \mathcal{S} \to \mathbb{R}, G : \mathbb{R}^n \times \mathcal{S} \to \mathbb{R}$ are given and $\theta > 0$, the risk-sensitive parameter, is a fixed constant.

We say that $u(\cdot)$ is an admissible control, if it belongs to $L^2_{\mathcal{F},p}([0,T];\mathbb{R}^k)$ and the stochastic differential equation (4.1.4) has a unique strong solution. We denote by $\mathscr{U}[0,T]$ the set of all admissible controls. Our risk-sensitive stochastic control problem associated with (4.1.4) − (4.1.5) is defined as follows:

$$\begin{cases} \text{Minimize } J^\theta(0, X, e_i; u(\cdot)) \\ \text{subject to } \begin{cases} u(\cdot) \in \mathscr{U}[0,T] \\ (X(\cdot), u(\cdot)) \text{ satisfies (4.1.4).} \end{cases} \end{cases} \tag{4.1.6}$$

The value $V^\theta : [0,T] \times \mathbb{R}^n \times \mathcal{S} \to \mathbb{R}$ associated with (4.1.6) is defined as follows:

$$V^\theta(0, X_0, e_i) := \inf_{u(\cdot)\in\mathscr{U}[0,T]} J^\theta(0, X_0, e_i, u(\cdot)). \tag{4.1.7}$$

Because of the exponential function,we require $V^\theta(0, X_0, e_i) \geq 0$ for $e_i \in \mathcal{S}$.

## 4.2 Adjoint variables and Hamiltonian

**Assumptions**

*(A1):* $f, g, \gamma$ are uniformly Lipschitz in $(x, u)$

*(A2):* $f, g, \gamma, F, G$ are twice continuously differentiable with respect to $x$, and their partial derivatives in $x$ are continuous in $(x, u)$;

*(A3):* $f_x, f_{xx}, g_x, g_{xx}, F_x, F_{xx}, \|\gamma_x\|_{\mathcal{M}^{2p}}, \|\gamma_{xx}\|_{\mathcal{M}^2}, p = 1, 2$ and $G_x, G_{xx}$ are bounded;

*(A4):* $F$ and $G$ are uniformly bounded;

*(A5):* $V^\theta \in C^{1,3}([0,T] \times \mathbb{R}^n \times \mathcal{S})$;

*(A6):* $\mathcal{U}$ is a convex subset of $\mathbb{R}^k$.

For risk-neutral maximum principle, we need assumptions *(A1)-(A3)*. To ensure the cost functional (4.1.5) to be well defined, *(A4)* is needed for the risk-sensitivity. And for the sufficient maximum principle, we require *(A6)*.

Let $\varphi = f, g, F, G, \gamma$, we define the following.

$$\begin{cases} \bar{\varphi}(t) \triangleq \varphi\big(t, \bar{X}(t), \bar{u}(t), \alpha(t)\big), \quad \bar{\varphi}_x(t) \triangleq \varphi_x\big(t, \bar{X}(t), \bar{u}(t), \alpha(t)\big), \\ \bar{\varphi}_{xx}(t) \triangleq \varphi_{xx}\big(t, \bar{X}(t), \bar{u}(t), \alpha(t)\big), \\ \delta\varphi(t, u) \triangleq \varphi\big(t, \bar{X}(t-), u, \alpha(t-)\big) - \varphi\big(t, \bar{X}(t-), \bar{u}(t), \alpha(t-)\big). \end{cases} \tag{4.2.1}$$

Let $\big(\bar{X}(\cdot), \bar{u}(\cdot)\big)$ be an admissible pair for the system (4.1.4). We introduce the first order adjoint variable:

$\big(\bar{p}(\cdot), \bar{q}(\cdot), \bar{s}(\cdot)\big) \in L^2_{\mathcal{F}}([0,T]; \mathbb{R}^n) \times L^2_{\mathcal{F},\,p}\big([0,T]; \mathbb{R}^{n\times m}\big) \times M^2_p\big([0,T]; \mathbb{R}^{n\times N}\big)$ and the second order adjoint variable:

$\big(\bar{P}(\cdot), \bar{Q}(\cdot), \bar{S}(\cdot)\big) \in L^2_{\mathcal{F}}([0,T]; \mathbb{R}^{n\times n}) \times \Big(L^2_{\mathcal{F},\,p}\big([0,T]; \mathbb{R}^{n\times n}\big)\Big)^m \times \Big(M^2_p\big([0,T]; \mathbb{R}^{n\times n}\big)\Big)^N$

associated with the admissible pair $\big(\bar{X}(\cdot), \bar{u}(\cdot)\big)$, which are the solutions of the following first order and second order adjoint equations respectively:

$$
\begin{cases}
d\bar{p}(t) = -\Big[\bar{f}_x(t)^T\bar{p}(t) + \bar{F}_x(t) + \sum_{j=1}^{m} \bar{g}_x^j(t)^T\bar{q}_j(t) \\
\qquad\qquad + \sum_{j=1}^{m} +\theta\bar{p}(t)^T\bar{g}^j(t)\big(\bar{g}_x^j(t)^T\bar{p}(t) + \bar{q}_j(t)\big) \\
\qquad\qquad + \sum_{j=1}^{N} \big[\bar{\gamma}_x^j(t)^T\bar{s}_j(t) + \Lambda_j(t)\big(\bar{\gamma}_x^j(t)^T\bar{p}(t) + \bar{\gamma}_x^j(t)^T\bar{s}_j(t)\big)\big]\lambda_j(t)\Big]dt \\
\qquad\qquad + \sum_{j=1}^{m} \bar{q}_j(t)dW_j(t) + \sum_{j=1}^{N} \bar{s}_j(t)d\tilde{\Phi}_j(t), \\
\bar{p}(T) = -G_x\big(\bar{X}(T), \alpha(T)\big),
\end{cases}
\tag{4.2.2}
$$

$$
\begin{cases}
d\bar{P}(t) = -\Big\{\bar{f}_x(t)^T\bar{P}(t) + \bar{P}(t)\bar{f}_x(t) + \sum_{j=1}^{m}\ \Big[\bar{g}_x^j(t)^T(\bar{P}(t) + \theta\bar{p}(t)\bar{p}(t)^T)\bar{g}_x^j(t) \\
\qquad +\bar{g}_x^j(t)^T\big(\bar{Q}_j(t) + \theta\bar{p}(t)^T\bar{g}^j(t)\bar{P}(t) + \theta\bar{p}(t)\bar{q}_j(t)^T\big) \\
\qquad +\big(\bar{Q}_j(t) + \theta\bar{p}(t)^T\bar{g}^j(t)\bar{P}(t) + \theta\bar{q}_j(t)\bar{p}(t)^T\big)\bar{g}_x^j(t) \\
\qquad +\theta\bar{p}(t)^T\bar{g}^j(t)\bar{Q}_j(t) + \theta\bar{q}_j(t)\bar{q}_j(t)^T\Big] \\
\qquad +\sum_{j=1}^{N}\ \Big[\gamma_x^j(t)^T\big(1+\Lambda_j(t)\big)\big(\bar{P}(t) + \bar{S}_j(t)\big)\bar{\gamma}_x^j(t) \\
\qquad +\gamma_x^j(t)^T\big(1+\Lambda_j(t)\big)\big(\theta(\bar{p}(t)+\bar{s}_j(t))\ (\bar{p}(t)+\bar{s}_j(t))^T\big)\bar{\gamma}_x^j(t) \\
\qquad +\gamma_x^j(t)^T\big[\big(1+\Lambda_j(t)\big)\big(\bar{P}(t) + \bar{S}_j(t) + \theta(\bar{p}(t)+\bar{s}_j(t))\ \bar{s}_j(t)^T\big) - \bar{P}(t)\big] \\
\qquad +\big[\big(1+\Lambda_j(t)\big)\big(\bar{P}(t) + \bar{S}_j(t) + \theta\bar{s}_j(t)(\bar{p}(t)+\bar{s}_j(t))^T\big) - \bar{P}(t)\big]\bar{\gamma}_x^j(t) \\
\qquad +\Lambda_j(t)\bar{S}_j(t) + \big(1+\Lambda_j(t)\big)\theta\bar{s}_j(t)\bar{s}_j(t)^T\Big]\gamma_j(t) \\
\qquad +\bar{H}_{xx}^{\theta}\big(t,\bar{X}(t),\bar{u}(t),\alpha(t),\bar{p}(t),\bar{q}(t),\bar{s}(t)\big)\Big\}dt \\
\qquad +\sum_{j=1}^{m}\ \bar{Q}_j(t)dW_j(t) + \sum_{j=1}^{N}\ \bar{S}_j(t)d\tilde{\Phi}_j(t), \\
\bar{P}(T) = -G_{xx}\big(\bar{X}(T),\alpha(T)\big).
\end{cases}
\tag{4.2.3}
$$

Where $g^j, \gamma^j$ are the $j^{th}$ columns of the matrices $g, \gamma$, respectively.

For each $t \in [0,T]$,

$$
\Lambda_j(t) = \frac{V^{\theta}(t,\bar{X}(t-)+\bar{\gamma}^j(t),e_j) - V^{\theta}(t,\bar{X}(t-),\alpha(t-))}{V^{\theta}(t,\bar{X}(t-),\alpha(t-))}, \quad j = 1,2,\cdots,N.
\tag{4.2.4}
$$

The Hamiltonian $\bar{H}^\theta : [0,T] \times \mathbb{R}^n \times \mathcal{U} \times \mathcal{S} \times \mathbb{R}^n \times \mathbb{R}^{n\times m} \times \mathcal{M}^2(\mathbb{R}^+; \mathbb{R}^{n\times N}) \to \mathbb{R}$ is defined as:

$$\begin{aligned}\bar{H}^\theta(t,X,u,e_i,p,q,s) := \big\langle p, f(t,X,u,e_i)\big\rangle - F(t,X,u,e_i) \\ + \sum_{j=1}^{m} g^j(t,x,u,e_i)^T(q_j + \theta p p^T \bar{g}^j(t)) \\ + \sum_{j=1}^{N} \big\langle s_j(t) + \Lambda_j(t)(p + s_j(t)), \gamma^j(t,X,u,e_i)\big\rangle \lambda_j(t).\end{aligned} \tag{4.2.5}$$

Note that (4.2.2) is a non linear backward stochastic differential equation (BSDE), which is different from the risk-neutral case. In addition, its generator does not satisfy the classical Lipschitz condition for the existence and uniqueness of solution to nonlinear (BSDE). Our assumptions are sufficient to guaranty the existence of unique solution to (4.2.2) and (4.2.3).

**Remark 4.1:** A key feature of (4.2.2) and (4.2.3) is that it relies on the value function, which involves the function $\Lambda_j(\cdot)$ defined by (4.2.4) the jump proportion process associated with the value function along with the state trajectory $\bar{X}(\cdot)$.

We define $\bar{\mathcal{H}}^\theta : [0,T] \times \mathbb{R}^n \times \mathcal{U} \times \mathcal{S} \to \mathbb{R}$, associated with the pair $(\bar{X}(\cdot), \bar{u}(\cdot))$, as:

$$
\begin{aligned}
\bar{\mathcal{H}}^\theta(t,X,u,e_i) &:= \big\langle \bar{p}(t-), f(t,X,u,e_i)\big\rangle - F(t,X,u,e_i) \\
&+\sum_{j=1}^{m}\Big[g^j(t,X,u,e_i)^T(\bar{q}_j(t)-\bar{P}(t-)\bar{g}^j(t)) \\
&+\frac{1}{2}g^j(t,X,u,e_i)^T(\bar{P}(t-)+\theta\bar{p}(t-)\bar{p}(t-)^T)g^j(t,X,u,e_i)\Big] \\
&+\sum_{j=1}^{N}\gamma^j(t,X,u,e_i)^T\Big[\bar{s}_j(t)+\Lambda_j(t)\big(\bar{p}(t-)+\bar{s}_j(t)\big)+\frac{1}{2}\big(1+\Lambda_j(t)\big)\times \\
&\Big(\bar{P}(t-)+\bar{S}_j(t)+\theta\big(\bar{p}(t-)+\bar{s}(t)\big)\big(\bar{p}(t-)+\bar{s}_j(t)\big)^T\Big)\Big(\gamma^j(t,X,u,e_i)-2\bar{\gamma}^j(t)\Big)\Big]\lambda_j(t).
\end{aligned}
\tag{4.2.6}
$$

**Theorem 4.2.1** *(Risk-sensitive maximum principle, cf. [21])*

*Suppose that the assumptions (A1)-(A5) hold. Let* $(\bar{X}(\cdot),\bar{u}(\cdot))$ *be an optimal pair for the risk-sensitive control problem* (4.1.6). *Then there exist processes* $\big(\bar{p}(\cdot),\bar{q}(\cdot),\bar{s}(\cdot)\big)$ *and* $\big(\bar{P}(\cdot),\bar{Q}(\cdot),\bar{S}(\cdot)\big)$ *satisfying the first and second order adjoint equations* (4.2.2) *and* (4.2.3), *respectively, such that the following inequality holds,*

$$
\begin{aligned}
&\bar{H}^\theta\big(t,\bar{X}(t-),u,\alpha(t-),\bar{p}(t-),\bar{q}(t),\bar{s}(t)\big)-\bar{H}^\theta\big(t,\bar{X}(t-),\bar{u}(t),\alpha(t-),\bar{p}(t-),\bar{q}(t),\bar{s}(t)\big) \\
&+\frac{1}{2}\sum_{j=1}^{m}\delta g^j(t,u)^T\big(\bar{P}(t-)+\theta\bar{p}(t-)\bar{p}(t-)^T\big)\delta g^j(t,u)+\frac{1}{2}\sum_{j=1}^{N}\Big[\delta\gamma^j(t,u)^T\Big(1+\Lambda_j(t)\Big) \\
&\times\Big(\bar{P}(t-)+\bar{S}_j(t)+\theta\big(\bar{p}(t-)+\bar{s}_j(t)\big)\big(\bar{p}(t-)+\bar{s}_j(t)\big)^T\Big)\delta\gamma^j(t,u)\Big]\lambda_j(t)\geq 0. \\
&\qquad\qquad\qquad \forall u\in\mathcal{U}, a.e.\ \ t\in[0,T], \mathbb{P}-a.s.
\end{aligned}
\tag{4.2.7}
$$

*or equivalently*

$$
\bar{\mathcal{H}}^\theta(t,\bar{X}(t-),\bar{u}(t),\alpha(t-)) = \inf_{u\in\mathcal{U}}\bar{\mathcal{H}}^\theta(t,\bar{X}(t-),u,\alpha(t-)), a.e.\ \ t\in[0,T], \mathbb{P}-a.s.
\tag{4.2.8}
$$

**Remark 4.2:** The equation (4.2.8) expressed by the minimum over $\mathcal{U}$ is still called the maximum condition for the control problem just as in the classical optimal control problem.

**Theorem 4.2.2** *(sufficient conditions for optimality, cf.[21])*

*Suppose that assumptions (A1)-(A6) hold. Let* $(\bar{X}(\cdot), \bar{u}(\cdot))$ *be an admissible pair, and* $\big(\bar{p}(\cdot), \bar{q}(\cdot), \bar{s}(\cdot)\big)$ *and* $\big(\bar{P}(\cdot), \bar{Q}(\cdot), \bar{S}(\cdot)\big)$ *be the associated first and second adjoint variables respectively.*

*Suppose for each* $e_i \in \mathcal{S},\ X \to g(X, e_i)$ *is convex,*

$(X, u) \to \bar{H}^{\theta}(t, X, u, e_i, \bar{p}(t-), \bar{q}(t), \bar{s}(t))$ *is convex* $\forall t \in [0, T], \mathbb{P} - a.s.,$ *and* (4.2.8) *holds. Then* $(\bar{X}(\cdot), \bar{u}(\cdot))$ *is an optimal pair for the problem* (4.1.6).

**Remark 4.3:**

(1) Since the jump proportion process $\Lambda(\cdot)$ given by (4.2.4) do not depend on the control variable $u$, we can simply apply the maximum condition (4.2.8) to look for the optimal control.

(2) Observe that if we take $N = 1$ and $\gamma \equiv 0$ in equation (4.1.4), Theorems (4.2.1) and (4.2.2) could be replaced by Theorems (3.1) and (3.2) in [13] and in this case, the first and second adjoint equations (4.2.2), (4.2.3) along with the $\bar{H}^{\theta}$-function and $\bar{\mathcal{H}}^{\theta}$-function (4.2.5), (4.2.6) coincide with equations (5), (6), (7), and (8) in [13], respectively.

## 4.3 Proofs of Theorems

In this section we bring a proof for Theorems (4.2.1) and (4.2.2). The proof is subdivided in four steps. In the first step, we apply the risk-neutral to problem (4.1.6) to be able to use the risk-neutral maximum principle. In steps 2 and 3, we transform the first and second order adjoint equation respectively to a relative simple form. In step 4 we transform the inequality (4.2.7).

**Step 1: Applying Risk-Neutral Maximum Principle.** In order to obtain a risk-sensitive maximum principle of the risk-neutral problem in global form, we define the Hamiltonian and associated second order adjoint equation. Combining the first and second order adjoint equations, a general stochastic maximum principle is obtained in terms of the variational inequality.

Let the following risk-neutral control problem be considered:

$$\begin{cases} \text{Minimize} \quad J^{\theta}(0,X,y,e_i;u(\cdot)) = E\Big[e^{\theta\,[G(X(T),\alpha(T))+\,y(t)]}\Big] \\ \text{subject to} \begin{cases} dX(t) = f\big(t,X(t),u(t),\alpha(t)\big)dt + g\big(t,X(t),u(t),\alpha(t)\big)dW(t) \\ \qquad\qquad + \gamma\,\big(t,X(t-),u(t-),\alpha(t-)\big)d\tilde{\Phi}(t), \\ dY(t) = F(t,X(t),u(t),\alpha(t))dt, \\ X(0)=X_0, \quad Y(0)=Y_0, \quad t\in[0,T], \quad u(\cdot)\in\mathscr{U}[0,T]. \end{cases} \end{cases} \tag{4.3.1}$$

Problem (4.1.6) corresponds to the case when $Y = 0$ in (4.3.1). The value function $V^\theta : [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathcal{S} \to \mathbb{R}$ associated with (4.2.1) is:

$$V^\theta(0, X_0, Y_0, e_i) := \inf_{u(\cdot)\in\mathscr{U}[0,T]} J^\theta(0, X_0, Y_0, e_i; u(\cdot)) \tag{4.3.2}$$

Note that $V^\theta(0, X_0, Y_0, e_i) = e^{\theta y} V^\theta(0, X_0, e_i)$. In particular, assumption *(A5)* implies that $V^\theta(0, X_0, Y_0, e_i) \in C^{1,3,\infty}\big([0,t] \times \mathbb{R}^n \times \mathbb{R} \times \mathcal{S}\big)$. Assume that *(A1)-(A4)* hold, and let $\big(\bar{X}(\cdot), \bar{Y}(\cdot), \bar{u}(\cdot)\big)$ be an optimal triple for the risk-neutral problem (4.3.1).

We introduce the following first and second order adjoint equations:

Let $G_x(\bar{X}(T)), \alpha(T)) := G_x(T)$, and $G_{xx}(\bar{X}(T)), \alpha(T)) := G_{xx}(T)$,

$$\begin{cases} dp(t) = -\Bigg\{ \begin{pmatrix} \bar{f}_x(t) & 0 \\ \bar{F}_x(t) & 0 \end{pmatrix}^T p(t) + \displaystyle\sum_{j=1}^m \begin{pmatrix} \bar{g}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix}^T q_j(t) \\ \qquad + \displaystyle\sum_{j=1}^N \begin{pmatrix} \bar{\gamma}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix}^T s_j(t)\lambda_j(t) \Bigg\} dt \\ \qquad + \displaystyle\sum_{j=1}^m q_j(t) dW_j(t) + \sum_{j=1}^N s_j(t) d\tilde{\Phi}_j(t), \\ p(T) = \theta e^{\theta\, [G(\bar{X}(T), \alpha(T) + \bar{Y}(T)]} \begin{pmatrix} G_x(\bar{X}(T)), \alpha(T)) \\ 1 \end{pmatrix}, \end{cases} \tag{4.3.3}$$

$$
\begin{cases}
dP(t) = -\Bigg\{ \begin{pmatrix} \bar{f}_x(t) & 0 \\ \bar{F}_x(t) & 0 \end{pmatrix}^T P(t) + P(t) \begin{pmatrix} \bar{f}_x(t) & 0 \\ \bar{F}_x(t) & 0 \end{pmatrix} \\
\quad + \sum_{j=1}^{m} \Bigg[ \begin{pmatrix} \bar{g}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix}^T Q_j(t) \\
\quad + Q_j(t) \begin{pmatrix} \bar{g}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \bar{g}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix}^T P(t) \begin{pmatrix} \bar{g}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix} \Bigg] \\
\quad + \sum_{j=1}^{N} \Bigg[ \begin{pmatrix} \bar{\gamma}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix}^T P(t) \begin{pmatrix} \bar{\gamma}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \bar{\gamma}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix}^T S_j(t) \\
\quad + S_j(t) \begin{pmatrix} \bar{\gamma}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \bar{\gamma}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix}^T S_j(t) \begin{pmatrix} \bar{\gamma}_x^j(t) & 0 \\ 0 & 0 \end{pmatrix} \Bigg] \lambda_j(t) \\
\quad + \begin{pmatrix} H_{xx}^{\theta}(t, \bar{X}(t), \bar{u}(t), \alpha(t), p(t), q(t), s(t)) & 0 \\ 0 & 0 \end{pmatrix} \Bigg\} dt \\
\quad + \sum_{j=1}^{m} Q_j(t) dW_j(t) + \sum_{j=1}^{N} S_j(t) d\tilde{\Phi}_j(t), \\
P(T) = \begin{pmatrix} \theta G_x\big(T\big) G_x\big(T\big)^T + G_{xx}\big(T\big) & \theta G_x\big(T\big) \\ \theta G_x\big(T\big)^T & \theta \end{pmatrix} \times \ \theta e^{\theta\, [G(\bar{X}(T), \alpha(T) + \bar{Y}(T)]}
\end{cases}
\tag{4.3.4}
$$

The Hamiltonian function

$\bar{H}^\theta : [0,T] \times \mathbb{R}^n \times \mathcal{U} \times \mathcal{S} \times \mathbb{R}^n \times \mathbb{R}^{n\times m} \times \mathcal{M}^2(\mathbb{R}^+;\mathbb{R}^{n\times N}) \to \mathbb{R}$ is given by:

$$H^\theta(t,X,u,e_i,p,q,s) := \left\langle p, \begin{pmatrix} f(t,X,u,e_i) \\ F(t,X,u,e_i) \end{pmatrix} \right\rangle + \sum_{j=1}^{m} \left\langle q_j, \begin{pmatrix} g^j(t,X,u,e_i) \\ 0 \end{pmatrix} \right\rangle + \sum_{j=1}^{N} \left\langle s_j(t), \begin{pmatrix} \gamma^j(t,X,u,e_i) \\ 0 \end{pmatrix} \right\rangle \lambda_j(t). \tag{4.3.5}$$

The adjoint equations (4.3.3) and (4.3.4) are linear (BSDE). The guaranty of the existence and uniqueness of solution to (4.3.3) and (4.3.4), respectively hold by considering assumptions *(A1)-(A3)*. The proof follows from, (cf. [2]) where jumps only come from the Markov regime. Combining ([22]. Theorem 2.1) with ([26], Theorem 3.1), let us consider the following Maximum Principle for the risk-neutral problem (4.3.1).

**Proposition 4.2.1** Let assumptions *(A1)-(A4)* hold and let $\big(\bar{X}(\cdot), \bar{Y}(\cdot), \bar{u}(\cdot)\big)$ be an optimal triple for risk-neutral problem (4.3.1). Then there exists a unique solution $\big(p(\cdot), q(\cdot), s(\cdot)\big)$ and $\big(P(\cdot), Q(\cdot), S(\cdot)\big)$ of (4.3.3) and (4.3.4), respectively, such that:

$$H^\theta\big(t,\bar X(t-),u,\alpha(t-),p(t-),q(t),s(t)\big)-H^\theta\big(t,\bar X(t-),\bar u(t),\alpha(t-),p(t-),q(t),s(t)\big)$$
$$+\frac{1}{2}\sum_{j=1}^{m}\begin{pmatrix}\delta g^j(t,u)\\0\end{pmatrix}^T P(t-)\begin{pmatrix}\delta g^j(t,u)\\0\end{pmatrix}$$
$$+\frac{1}{2}\sum_{j=1}^{N}\left[\begin{pmatrix}\delta\gamma^j(t,u)\\0\end{pmatrix}^T\big(P(t-)+S_j(t)\big)\begin{pmatrix}\delta\gamma^j(t,u)\\0\end{pmatrix}\right]\lambda_j(t)\ge 0.$$
$$\forall u\in\mathcal{U}, a.e.\ \ t\in[0,T],\mathbb{P}-a.s, \tag{4.3.6}$$

or equivalently,

$$\mathcal{H}^\theta(t,\bar X(t-),\bar u(t),\alpha(t-))=\inf_{u\in\mathcal{U}}\mathcal{H}^\theta(t,\bar X(t-),u,\alpha(t-)), a.e.\ \ t\in[0,T],\mathbb{P}-a.s., \tag{4.3.7}$$

where the $\mathcal{H}^\theta$-function for problem (4.3.1) associated with $\big(\bar X(\cdot),\bar Y(\cdot),\bar u(\cdot)\big)$ is defined by:

$$\mathcal{H}^\theta(t,X,u,e_i)= H^\theta\big(t,X,u,e_i,p(t-),q(t),s(t)\big)$$
$$-\frac{1}{2}\sum_{j=1}^{m}\begin{pmatrix}\bar g^j(t)\\0\end{pmatrix}^T P(t-)\begin{pmatrix}\bar g^j(t)\\0\end{pmatrix}+\frac{1}{2}\sum_{j=1}^{m}\begin{pmatrix}\delta g^j(t,u)\\0\end{pmatrix}^T P(t-)\begin{pmatrix}\delta g^j(t,u)\\0\end{pmatrix}$$
$$-\frac{1}{2}\sum_{j=1}^{N}\left[\begin{pmatrix}\bar\gamma^j(t)\\0\end{pmatrix}^T\big(P(t-)+S_j(t)\big)\begin{pmatrix}\bar\gamma^j(t)\\0\end{pmatrix}\right]\lambda_j(t)$$
$$+\frac{1}{2}\sum_{j=1}^{N}\left[\begin{pmatrix}\delta\gamma^j(t,u)\\0\end{pmatrix}^T\big(P(t-)+S_j(t)\big)\begin{pmatrix}\delta\gamma^j(t,u)\\0\end{pmatrix}\right]\lambda_j(t). \tag{4.3.8}$$

By combining ([22], Theorem 2.1) and ([26], Theorem 3.1), the complete proof can be easily retrieve. In the following, we will derive the sufficient conditions for the optimality of$\big(\bar X(\cdot),\bar Y(\cdot),\bar u(\cdot)\big)$

**Proposition 4.2.2** Suppose that assumptions *(A1)-(A4)* and *(A6)* hold and let $\big(\bar{X}(\cdot), \bar{Y}(\cdot), \bar{u}(\cdot)\big)$ be an admissible triple, $\big(p(\cdot), q(\cdot), s(\cdot)\big)$ and $\big(P(\cdot), Q(\cdot), S(\cdot)\big)$ satisfy (4.3.3) and (4.3.4), $H^{\theta}\big(t, X, u, e_i, p(t), q(t), s(t)\big)$ is convex for all $t \in [0, T]$, $\mathbb{P}$-a.s., and (4.3.7) holds. Then $\big(\bar{X}(\cdot), \bar{Y}(\cdot), \bar{u}(\cdot)\big)$ is an optimal triple for problem (4.3.1)

**Remark 4.4:** Propositions (4.2.1) and (4.2.2) do not depend on assumption *(A5)*.

**Step 2: Transformation of the first order adjoint equation.**

Proposition (4.2.1) can be viewed as a maximum principle for the underlying risk-sensitive control problem (4.1.6). However, it is not what we actually need since the adjoint equations involve additional components. To solve this problem, we need to transform the adjoint variables $\big(p(\cdot), q(\cdot), s(\cdot)\big)$ and $\big(P(\cdot), Q(\cdot), S(\cdot)\big)$.

We use logarithm transformation (cf. [13]) to derive a PDE for the value function and the relationship between the maximum principle and the dynamic programming principle (cf. [26]). We extend the result in [13] to a continuous time Markov-regime switching, but without Poisson processes which is a simplified version of (cf. [21]) where Poisson processes were considered.

**Lemma 4.1: (cf. [21])** Considering *(A1)-(A5)*, the first order adjoint equation (4.3.3) reduces to (4.2.2).

**Proof**. Let $\big(\bar{X}(\cdot), \bar{Y}(\cdot), \bar{u}(\cdot)\big)$ be an optimal triple for problem (4.2.1) and

$$\big(p(\cdot), q(\cdot), s(\cdot), \big) \equiv \left( \begin{bmatrix} p_1(\cdot) \\ p_2(\cdot) \end{bmatrix}, \begin{bmatrix} q_1(\cdot) \\ q_2(\cdot) \end{bmatrix}, \begin{bmatrix} s_1(\cdot) \\ s_2(\cdot) \end{bmatrix} \right)$$

$\big(p(\cdot), q(\cdot), s(\cdot),\big) \in L^2_{\mathcal{F}}\big([0,T];\mathbb{R}^n\big) \;\times\; L^2_{\mathcal{F},\,p}\big([0,T];\mathbb{R}^{n\,\times\,m}\big) \;\times\; M^2_p\big([0,T];\mathbb{R}^{n\,\times\,N}\big)$,

be the first order adjoint variables satisfying equation (4.3.3), where

$\big(p_1(\cdot), q_1(\cdot), s_1(\cdot),\big) \in L^2_{\mathcal{F}}\big([0,T];\mathbb{R}^n\big) \;\times\; L^2_{\mathcal{F},\,p}\big([0,T];\mathbb{R}^{n\,\times\,m}\big) \;\times\; M^2_p\big([0,T];\mathbb{R}^{n\,\times\,N}\big)$,

and $\big(p_2(\cdot), q_2(\cdot), s_2(\cdot),\big) \in L^2_{\mathcal{F}}\big([0,T];\mathbb{R}\big) \times L^2_{\mathcal{F},\,p}\big([0,T];\mathbb{R}^{1\,\times\,m}\big) \times M^2_p\big([0,T];\mathbb{R}^{1\,\times\,N}\big)$.

By the relationship between the maximum principle and the dynamic programming principle (cf. [26], Theorem 4.2), we have:

$$p(t) = V^{\theta}_{(X,Y)}(t, \bar{X}, \bar{Y}, e_i), \tag{4.3.9}$$

where $V^{\theta}_{(X,Y)}$ denotes the gradient of $V^{\theta}$ in $(X,Y)$. Let us introduce the following logarithmic transformation of the value function

$$v^{\theta}(t, X, Y, e_i) = \frac{1}{\theta} \ln V^{\theta}(t, X, Y, e_i) \tag{4.3.10}$$

Taking gradient on the right hand side of (4.2.10) and noting 4.2.9), we have the following transformation of the first order adjoint variable:

$$\tilde{p}(t) = \frac{1}{\theta}\frac{p(t)}{V(t)}, \tag{4.3.11}$$

where $V(t) := V^{\theta}\big(t, \bar{X}(t), \bar{Y}(t), \alpha(t)\big) > 0$.

We then derive the equation for $\tilde{p}(\cdot) \equiv \begin{pmatrix} \bar{p}(\cdot) \\ \tilde{p}^*(\cdot) \end{pmatrix}$, where $\bar{p}(\cdot)$ is $\mathbb{R}^n$-valued. We notice that $V^{\theta}(0, X_0, Y_0, e_i)$ is the value function of the risk-neutral problem (4.3.1). It follows from ([21], equation(4.12)) that $V$ must satisfy:

$$\begin{cases} dV(t) = \sum_{j=1}^{m} p_1(t)^T \bar{g}^j(t) dW_j(t) + \sum_{j=1}^{N} \big[V^\theta(t, \bar{X}(t-) + \bar{\gamma}^j(t), \bar{Y}(t), e_i) \\ \qquad -V^\theta(t, \bar{X}(t-), \bar{Y}(t), \alpha(t-))\big] d\tilde{\Phi}_j(t), \\ V(T) = e^{\theta\big[G(\bar{X}(T),\alpha(T))+\bar{Y}(T)\big]}. \end{cases} \tag{4.3.12}$$

Assume that $\tilde{p}$ satisfies an equation of the following form:

$$d\tilde{p}(t) = \alpha(t)dt + \sum_{j=1}^{m} \tilde{q}_j(t)dW_j(t) + \sum_{j=1}^{N} \tilde{s}_j(t)d\tilde{\Phi}_j(t). \tag{4.3.13}$$

Using It$\hat{o}$'s formula for Markov regime switching process (cf.[26], Theorem 4.1), we obtain:

$$\begin{aligned} dp(t) = {} & d(\theta V(t)\tilde{p}(t)) \\ = {} & \theta V(t-)d\tilde{p}(t) + \sum_{j=1}^{m} \theta p_1(t)^T \bar{g}^j(t)\, \tilde{q}_j(t)dt + \sum_{j=1}^{m} \theta\tilde{p}(t)p_1(t)^T \bar{g}^j(t)dW_j(t) \\ & + \sum_{j=1}^{N} \theta\big[V^\theta(t, \bar{X}(t-) + \bar{\gamma}^j(t), \bar{Y}(t), e_j) - V^\theta(t, \bar{X}(t-), \bar{Y}(t), \alpha(t-))\big] \times \\ & \tilde{s}_j(t)\lambda_j(t)dt \\ & + \sum_{j=1}^{N} \theta\big[V^\theta(t, \bar{X}(t-) + \bar{\gamma}^j(t), \bar{Y}(t), e_j) - V^\theta(t, \bar{X}(t-), \bar{Y}(t), \alpha(t-))\big] \times \\ & \big(\tilde{p}(t-) + \tilde{s}_j(t)\big)d\tilde{\Phi}_j(t). \end{aligned} \tag{4.3.14}$$

Dividing both sides of (4.3.14) by $\theta V(t-)$ and noting that $p_1(t-) = \theta V(t-)\bar{p}(t-)$, we obtain:

$$
\begin{aligned}
d\tilde{p}(t) = \ & \frac{1}{\theta V(t-)} dp(t) - \sum_{j=1}^{m} \theta \bar{p}(t)^T \bar{g}^j(t)\ \tilde{q}_j(t) dt - \sum_{j=1}^{m} \theta \tilde{p}(t)\ \tilde{p}(t)^T \bar{g}^j(t) dW_j(t) \\
& - \sum_{j=1}^{N} \tilde{\Lambda}_j(t) \tilde{s}_j(t) \lambda_j(t) dt - \sum_{j=1}^{N} \tilde{\Lambda}_j(t) \big(\tilde{p}(t-) + \tilde{s}_j(t)\big) d\tilde{\Phi}_j(t),
\end{aligned}
\tag{4.3.15}
$$

where for each $t \in [0, T]$ we have:

$$
\tilde{\Lambda}_j(t) = \frac{V^\theta(t, \bar{X}(t-) + \bar{\gamma}^j(t), \bar{Y}(t), e_j) - V^\theta(t, \bar{X}(t-), \bar{Y}(t), \alpha(t-))}{V^\theta(t, \bar{X}(t-), \bar{Y}(t), \alpha(t-))}, \quad j = 1, \cdots, N
$$

Substituting (4.3.3) into (4.3.15) leads to:

$$
\begin{aligned}
\tilde{q}_j(t) \equiv \begin{pmatrix} \bar{q}_j(t) \\ \tilde{q}_j^*(t) \end{pmatrix} &= \frac{q_j(t)}{\theta V(t-)} - \theta \tilde{p}(t)\ \bar{p}(t)^T \bar{g}^j(t), j = 1, \cdots, m, \\
\tilde{s}_j(t) \equiv \begin{pmatrix} \bar{s}_j(t) \\ \tilde{s}_j^*(t) \end{pmatrix} &= \frac{s_j(t)}{\theta V(t-)} - \tilde{\Lambda}(t) \big(\tilde{p}(t-) + \tilde{s}_j(t)\big), j = 1, \cdots, N,
\end{aligned}
\tag{4.3.16}
$$

where $\bar{q}(\cdot) = \big(\bar{q}_1(\cdot), \cdots, \bar{q}_m(\cdot)\big)$ is $\mathbb{R}^{n \times m}$-valued, and $\bar{s}(\cdot) = \big(\bar{s}_1(\cdot), \cdots, \bar{s}_N(\cdot)\big)$ is $\mathbb{R}^{n \times N}$-valued. Substituting (4.3.16) into (4.3.15) and considering (4.3.3). It follows that the transformed first order adjoint variable $\tilde{p}(\cdot)$ satisfies the following equation:

$$
\begin{cases}
d\tilde{p}(t) = -\Bigg\{ \begin{pmatrix} \bar{f}_x(t) & 0 \\ \bar{F}_x(t) & 0 \end{pmatrix}^T \tilde{p}(t) + \sum_{j=1}^{m} \begin{pmatrix} \bar{g}^j_x(t) & 0 \\ 0 & 0 \end{pmatrix}^T \big(\tilde{q}_j(t) + \theta\tilde{p}(t)\ \tilde{p}(t)^T \bar{g}^j(t)\big) \\
\quad + \sum_{j=1}^{N} \begin{pmatrix} \bar{\gamma}^j_x(t) & 0 \\ 0 & 0 \end{pmatrix}^T \big(\tilde{s}_j(t) +\ \tilde{\Lambda}_j(t)(\tilde{p}(t-) + \tilde{s}_j(t))\big)\lambda_j(t) \\
\quad + \sum_{j=1}^{m} \theta\bar{p}(t)^T \bar{g}^j(t)\tilde{q}_j(t)dt + \sum_{j=1}^{N} \tilde{\Lambda}_j(t)\tilde{s}_j(t)\lambda_j(t) \Bigg\} dt \\
\quad + \sum_{j=1}^{m} \tilde{q}_j(t)dW_j(t) + \sum_{j=1}^{N} \tilde{s}_j(t)d\tilde{\Phi}_j(t), \\
\tilde{p}(T) = \begin{pmatrix} G_x(\bar{X}(T)), \alpha(T)) \\ 1 \end{pmatrix},
\end{cases}
\tag{4.3.17}
$$

where $\tilde{p}(T)$ is determined from (4.3.11) and $p(T)$.

Letting $Y = 0$ in (4.3.1) and expanding (4.3.17), we have:

$$
\tilde{p}^*(t) = 1, \;\; \tilde{q}^*_j(t) = \tilde{s}^*_j(t) = 0, \; \forall t \in [0, T], \tag{4.3.18}
$$

and $\big(\bar{p}(\cdot), \bar{q}(\cdot), \bar{s}(\cdot)$ is a solution of (4.2.2). This explains how equation (4.2.2) is derived. Therefore, from this derivation, it follows from the uniqueness of solution to (4.3.3) that this solution is unique.

**Step 3: Transformation of the second order adjoint equation.**

The transformation of second order adjoint equation is given in the following lemma.

**Lemma 4.2: (cf. [21])** Let assumptions *(A1)-(A5)* hold, then the second order adjoint equation (4.3.4) reduces to (4.2.3).

**Proof**. Let $(P(\cdot), Q(\cdot), S(\cdot)$ be the second order adjoint variable satisfying (4.3.4) and consider the following transformation:

$$\tilde{P}(t) = \frac{1}{\theta}\frac{P(t)}{V(t)} - \theta\tilde{p}(t)\tilde{p}(t)^T \triangleq \Gamma(t) - \theta\tilde{p}(t)\tilde{p}(t)^T, \tag{4.3.19}$$

assume that:

$$d\Gamma(t) = X(t)dt + \sum_{j=1}^{m} Y_j(t)dW_j(t) + \sum_{j=1}^{N} A_j(t)d\tilde{\Phi}_j(t), \tag{4.3.20}$$

using It$\hat{o}$'s formula, (4.3.12) and (4.3.19) we get:

$$\begin{aligned} dP(t) = &\ d(\theta V(t)\Gamma(t)) \\ = &\ \theta V(t-)d\Gamma(t) + \sum_{j=1}^{m} \theta p_1(t)^T\bar{g}^j(t)\ Y_j(t)dt + \sum_{j=1}^{m} \theta\Gamma(t)p_1(t)^T\bar{g}^j(t)dW_j(t) \\ &+ \sum_{j=1}^{N} \theta\big[V^\theta(t,\bar{X}(t-)+\bar{\gamma}^j(t),\bar{Y}(t),e_j) - V^\theta(t,\bar{X}(t-),\bar{Y}(t),\alpha(t-))\big]\times \\ &\quad A_j(t)\lambda_j(t)dt \\ &+ \sum_{j=1}^{N} \theta\big[V^\theta(t,\bar{X}(t-)+\bar{\gamma}^j(t),\bar{Y}(t),e_j) - V^\theta(t,\bar{X}(t-),\bar{Y}(t),\alpha(t-))\big]\times \\ &\quad \big(\Gamma(t-) + A_j(t)\big)d\tilde{\Phi}_j(t). \end{aligned} \tag{4.3.21}$$

Dividing both sides of (4.3.21) by $\theta V(t-)$ and taking $p_1(t-) = \theta V(t-)\bar{p}(t-)$, we obtain:

$$
\begin{aligned}
d\Gamma(t) = & \frac{1}{\theta V(t-)} dP(t) - \sum_{j=1}^{m} \theta \bar{p}(t)^T \bar{g}^j(t)\, Y_j(t) dt - \sum_{j=1}^{m} \theta \Gamma(t)\, \bar{p}(t)^T \bar{g}^j(t) dW_j(t) \\
& - \sum_{j=1}^{N} \tilde{\Lambda}_j(t)\, A_j(t) \lambda_j(t) dt - \sum_{j=1}^{N} \tilde{\Lambda}_j(t) \big( \Gamma(t-) + A_j(t) \big) d\tilde{\Phi}_j(t),
\end{aligned}
\tag{4.3.22}
$$

substituting the expression (4.3.4) with (4.3.22) we obtain:

$$
\begin{aligned}
Y_j(t) &\equiv \frac{Q_j(t)}{\theta V(t-)} - \theta \bar{p}(t)^T \bar{g}^j(t) \big( \tilde{P}(t) + \theta \tilde{p}(t) \tilde{p}(t)^T \big),\ j = 1, \cdots, m, \\
A_j(t) &\equiv \frac{S_j(t)}{\theta V(t-)} - \tilde{\Lambda}(t) \big( \tilde{P}(t-) + \theta \tilde{p}(t-) \tilde{p}(t-)^T + A_j(t) \big),\ j = 1, \cdots, N,
\end{aligned}
\tag{4.3.23}
$$

Combining (4.3.5) with (4.2.5) we have:

$$
\begin{aligned}
& \frac{1}{\theta V(t-)} \begin{pmatrix} H^{\theta}_{xx}\big(t, \bar{X}(t-), \bar{u}(t), \alpha(t-), p(t-), q(t), s(t)\big) & 0 \\ 0 & 0 \end{pmatrix} \\
& \quad = \begin{pmatrix} \bar{H}^{\theta}_{xx}\big(t, \bar{X}(t-), \bar{u}(t), \alpha(t-), \bar{p}(t-), \bar{q}(t), \bar{s}(t)\big) & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}
\tag{4.3.24}
$$

Applying It$\hat{o}$'s formula to (4.3.19) and using (4.3.17) and (4.3.22), we have.

$$
\begin{aligned}
d\tilde{P}(t) = & -\Bigg\{ \begin{pmatrix} \bar{f}_x(t) & 0 \\ \bar{F}_x(t) & 0 \end{pmatrix}^T \tilde{P}(t) + \tilde{P}(t) \begin{pmatrix} \bar{f}_x(t) & 0 \\ \bar{F}_x(t) & 0 \end{pmatrix} \\
& + \sum_{j=1}^{m} \begin{pmatrix} \bar{g}^j_x(t) & 0 \\ 0 & 0 \end{pmatrix}^T \big(\tilde{P}(t) + \theta \tilde{p}(t)\bar{p}(t)^T\big) \begin{pmatrix} \bar{g}^j_x(t) & 0 \\ 0 & 0 \end{pmatrix} \\
& + \sum_{j=1}^{m} \begin{pmatrix} \bar{g}^j_x(t) & 0 \\ 0 & 0 \end{pmatrix}^T \big(\tilde{Q}_j(t) + \theta \bar{p}(t)^T \bar{g}^j(t)\tilde{P}(t) + \theta \tilde{p}(t)\tilde{q}_j(t)^T\big) \\
& + \sum_{j=1}^{m} \big(\tilde{Q}_j(t) + \theta \bar{p}(t)^T \bar{g}^j(t)\tilde{P}(t) + \theta \tilde{q}_j(t)\tilde{p}(t)^T\big) \begin{pmatrix} \bar{g}^j_x(t) & 0 \\ 0 & 0 \end{pmatrix} \\
& + \sum_{j=1}^{m} \big[\theta \bar{p}(t)^T \bar{g}^j(t)\tilde{Q}_j(t) + \theta \tilde{q}_j(t)\bar{q}_j(t)^T\big] \\
& + \sum_{j=1}^{N} \begin{pmatrix} \bar{\gamma}^j_x(t,z) & 0 \\ 0 & 0 \end{pmatrix}^T \big(1 + \tilde{\Lambda}_j(t)\big)\big(\tilde{P}(t) + \tilde{S}_j(t)\big) \begin{pmatrix} \bar{\gamma}^j_x(t,z) & 0 \\ 0 & 0 \end{pmatrix} \lambda_j(t) \\
& \qquad + \begin{pmatrix} \bar{\gamma}^j_x(t,z) & 0 \\ 0 & 0 \end{pmatrix}^T \big(1 + \tilde{\Lambda}_j(t)\big)\big(\theta(\tilde{p}(t) + \tilde{s}_j(t))(\tilde{p}(t) + \tilde{s}_j(t))^T\big) \times \\
& \qquad \begin{pmatrix} \bar{\gamma}^j_x(t,z) & 0 \\ 0 & 0 \end{pmatrix} \lambda_j(t) \\
& + \sum_{j=1}^{N} \begin{pmatrix} \bar{\gamma}^j_x(t,z) & 0 \\ 0 & 0 \end{pmatrix}^T \Big[\big(1 + \tilde{\Lambda}_j(t)\big)\big(\tilde{P}(t) + \tilde{S}_j(t)\big) \\
& \qquad + \big(1 + \tilde{\Lambda}_j(t)\big)\big(\theta\big(\tilde{p}(t) + \tilde{s}_j(t)\tilde{s}_j(t)^T\big) - \tilde{P}(t)\Big]\lambda_j(t) \\
& + \sum_{j=1}^{N} \Big[\big(1 + \tilde{\Lambda}_j(t)\big)\Big(\tilde{P}(t) + \tilde{S}_j(t) + \theta \tilde{s}_j(t)\big(\tilde{p}(t) + \tilde{s}_j(t)\big)^T\Big) - \tilde{P}(t)\Big] \times \\
& \qquad \begin{pmatrix} \bar{\gamma}^j_x(t,z) & 0 \\ 0 & 0 \end{pmatrix} \lambda_j(t) \\
& + \sum_{j=1}^{N} \Big[\tilde{\Lambda}_j(t)\tilde{S}_j(t) + \Big(1 + \tilde{\Lambda}_j(t)\Big)\, \theta \tilde{s}_j(t)\tilde{s}_j(t)^T\Big]\lambda_j(t) \\
& + \begin{pmatrix} \bar{H}^\theta_{xx}\big(t, \bar{X}(t-), \bar{u}(t), \alpha(t), \bar{p}(t), \bar{q}(t), \bar{s}(t)\big) & 0 \\ 0 & 0 \end{pmatrix} \Bigg\} dt \\
& + \sum_{j=1}^{m} \tilde{Q}_j(t) dW_j(t) + \sum_{j=1}^{N} \tilde{S}_j(t) d\tilde{\Phi}_j(t), \\
\tilde{P}(T) = & \begin{pmatrix} G_{xx}(\bar{X}(T)), \alpha(T)) & 0 \\ 0 & 0 \end{pmatrix},
\end{aligned}
\tag{4.3.25}
$$

where,

$$\tilde{Q}_j(t) = \frac{Q_j(t)}{\theta V(t-)} - \theta \bar{p}(t)^T \bar{g}^j(t)\big(\tilde{P}(t) + \theta \tilde{p}(t)\tilde{p}(t)^T\big) - \theta \tilde{p}(t)\tilde{q}_j(t)^T - \theta \tilde{q}_j(t)\tilde{p}(t)^T,$$

$$j = 1, \cdots, m,$$

$$\tilde{S}_j(t) = \frac{S_j(t)}{\theta V(t-)} - \tilde{\Lambda}(t)\big(\tilde{P}(t-) + \theta \tilde{p}(t-)\tilde{p}(t-)^T + A_j(t)\big) - \theta \tilde{p}(t-)\tilde{s}_j(t)^T$$

$$- \theta \tilde{s}_j(t)\tilde{p}(t-)^T - \theta \tilde{s}_j(t)\tilde{s}_j(t)^T, \; j = 1, \cdots, N. \tag{4.3.26}$$

Therefore, it follows that

$$\begin{aligned} \tilde{P}(t) &= \begin{pmatrix} \bar{P}(t) & 0 \\ 0 & 0 \end{pmatrix}, \\ \tilde{Q}(t) &= \begin{pmatrix} \bar{Q}(t) & 0 \\ 0 & 0 \end{pmatrix}, \; j = 1, \cdots, m \\ \tilde{S}(t) &= \begin{pmatrix} \bar{S}(t) & 0 \\ 0 & 0 \end{pmatrix}, \; j = 1, \cdots, N, \end{aligned} \tag{4.3.27}$$

where $\big(\bar{P}(\cdot), \bar{Q}(\cdot), \bar{S}(\cdot)$ is the solution of (4.2.3). This solution is unique.

**Step 4: Maximum Condition:**

The transformation of the variational inequality (4.3.6) and the maximum condition are given in the following lemma.

**Lemma 4.3: (cf. [21])** Under assumptions *(A1)-(A5)*, the variational inequality (4.3.6) and the maximum condition (4.3.7) are reduced to (4.2.7) and (4.2.8), respectively. (cf.[21])

**Proof.** Let consider (4.3.6), in a view of (4.3.11), (4.3.16), and (4.3.18), for each $e_i \in \mathcal{S}$,

$$H^{\theta}\big(t, X, u, e_i, p(t-), q(t), s(t)\big) = \theta V(t-)\bar{H}^{\theta}\big(t, X, u, e_i, \bar{p}(t-), \bar{q}(t), \bar{s}(t)\big), \quad (4.3.28)$$

where $H^{\theta}$ and $\bar{H}^{\theta}$ are defined by (4.3.5) and (4.2.5), respectively.

By (4.3.19), (4.3.26), and (4.3.27) we obtain:

$$\frac{1}{2}\begin{pmatrix}\delta g^j(t,u)\\ 0\end{pmatrix}^T P(t-)\begin{pmatrix}\delta g^j(t,u)\\ 0\end{pmatrix} = \frac{\theta V(t-)}{2}\delta g^j(t,u)^T \times \big(\bar{P}(t-) + \theta\bar{p}(t-)\bar{p}(t-)^T\big)\delta g^j(t,u), \quad (4.3.29)$$

$$\frac{1}{2}\begin{pmatrix}\delta\gamma^j(t,u)\\ 0\end{pmatrix}^T \big(P(t-) + S_j(t)\big)\begin{pmatrix}\delta\gamma^j(t,u)\\ 0\end{pmatrix} = \frac{\theta V(t-)}{2}\,\delta\gamma^j(t,u)^T \times \Big(1 + \Lambda_j(t)\Big)\Big(\bar{P}(t-) + \bar{S}_j(t) + \theta\big(\bar{p}(t-) + \bar{s}_j(t)\big)\big(\bar{p}(t-) + \bar{s}_j(t)\big)^T\Big)\delta\gamma^j(t,u). \quad (4.3.30)$$

Since $V(t-) > 0$, it follows that (4.2.6) is equivalent to:

$$\begin{aligned}
&\bar{H}^{\theta}\big(t, \bar{X}(t-), u(t), \alpha(t-), \bar{p}(t-), \bar{q}(t), \bar{s}(t)\big) - \bar{H}^{\theta}\big(t, \bar{X}(t-), \bar{u}(t), \alpha(t-), \bar{p}(t-), \bar{q}(t), \bar{s}(t)\big)\\
&\quad + \frac{1}{2}\sum_{j=1}^{m} \delta g^j(t,u)^T\Big(\bar{P}(t-) + \theta\bar{p}(t-)\bar{p}(t-)^T\Big)\delta g^j(t,u)\\
&\quad + \frac{1}{2}\sum_{j=1}^{N}\Bigg[\delta\gamma^j(t,u)^T\Big(1 + \Lambda_j(t)\Big)\Big(\bar{P}(t-) + \bar{S}_j(t)\Big)\delta\gamma^j(t,u)\\
&\quad + \delta\gamma^j(t,u)^T\Big(1 + \Lambda_j(t)\Big)\Big(\theta\big(\bar{p}(t-) + \bar{s}_j(t)\big)\big(\bar{p}(t-) + \bar{s}_j(t)\big)^T\Big)\delta\gamma^j(t,u)\Bigg]\lambda_j(t) \geq 0.\\
&\qquad \forall u \in \mathcal{U}, a.e.\ \ t \in [0,T], \mathbb{P} - a.s.
\end{aligned} \quad (4.3.31)$$

This completes the proofs for Theorem (4.2.1) and (4.2.2).

# 5 Application

In this section we present an application of the risk-sensitivity stochastic maximum principle with regime switching.

## 5.1 Application to optimal portfolio choice problem

We will now discuss the application of the results in the previous chapter to the topic of optimal portfolio choice problem in financial market with Markov regime-switching. We will consider the Fleming and Sheu model (cf. [20]) where the price of the risky asset $S_1(t)$ is given by $L(t) = \log S_1(t)$. The Brownian motion will be interpreted as small random shocks that influence the market dynamics. The Markov process will be the transition to one state model to another state, which is simply a shift in the financial market behavior. Finally, we will use the hyperbolic absolute risk aversion (HARA) utility function of wealth with the risk-sensitive $\theta$ in this model to capture the action of the investors regarding their risk level.

• Problem formulation

Let $T > 0$ be a deterministic finite horizon time and let $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ be a complete filtered probability space. We define on $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ a standard one-dimension Brownian motion $W(t) = \{W(t)\}_{t\in[0,T]}$, a continuous time Markov process $\{\alpha(t)\}_{t\in[0,T]}$ with a finite state space $S = \{e_1, e_2\}$. The generator of the Markov chain $\alpha$ is given by the $Q$-matrix $Q = (q_{ij})_{i,j\in S}$. Let $N_{ij}(t)$ denote the counting process given by

$$N_{ij} = \sum_{0<s\leq t} I_{\{\alpha(s-)=i\}} I_{\{\alpha(s)=j\}},$$

where $I_A$ denotes the indicator function of a set $A$. We note that $N_{ij}(t)$ gives the number of jumps of the Markov process $\alpha$ from state $i$ to state $j$ up to time t. We define the intensity process by:

$$\lambda_{ij} = q_{ij} I_{\{\alpha(s-)=i\}}$$

and we introduce the martingale process $M_{ij}(t)$ given by:

$$M_{ij}(t) = N_{ij}(t) - \int_0^t \lambda_{ij}(s)ds.$$

The process $M_{ij}(t)$ is a pure discontinuous, square-integrable martingale which is null at the origin.

We assume that the stochastic processes, $\alpha(\cdot), W(\cdot)$ are independent. Using the stochastic processes described above, we define a financial market consisting of one risk-free asset and one risky-asset. Let us denote by $\{S_0(t)\}_{t\in[0,T]}$ the risk-free asset (bond) and by $\{S_1(t)\}_{t\in[0,T]}$ the traded risky asset (stock market index). We define the following differential equations:

The risk-free asset satisfies

$$dS_0(t) = r\big(t, \alpha(t)\big) S_0(t)dt. \tag{5.1.1}$$

Where $r \geq 0$ is the constant bond rate.

The price of a stock at time $t$ is given by the risky asset $S_1(t)$. Using Fleming and Sheu model (cf. [20]), Let $L(t) = \log S_1(t)$ and assume that

$$dL(t) = c(\bar{L}(t) - L(t))dt + g(t)(d\tilde{W}(t)) + \gamma(t)d\tilde{\Phi}(t), \qquad (5.1.2)$$

where $g$ is the stock price volatility rate, $c > 0$ is some appropriate coefficient, and $\gamma$ represents the jumps related to the Markov regime-switching. $\bar{L}(t)$ is the deterministic log stock price trend and it is linear in $t$. We have $\bar{L}(t) = mt + \bar{L}_0$, where $m$ and $\bar{L}_0$ are constants.

We assume that the financial market coefficients depend both on time and the state of the Markov process $\{\alpha(t)\}_{t\in[0,T]}$; and we also assume that the risk-free interest rate $r(t,e)$ and the risky-asset volatility $g(t,e)$, are deterministic continuous functions on the interval $[0,T]$ for every fixed state $e \in S$.

Let us also assume that $r(t,e) \geq 0$, $\forall (t,e) \in [0,T] \times S$, and

$$\mathbb{E}\Big[\int_0^T |\sigma(t,e)|^2 dt\Big] < \infty, \;\; \forall\, e \in S$$

We now introduce the control variables.

If $X(t)$ is the amount of the investor's wealth and $u(t)$ is the proportion of wealth invested in the stock at time t with $u(t) \in \mathcal{U}$ where $\mathcal{U} = (-\infty, \infty)$, then $u(t)X(t)$ is the amount in the stock and $(1-u(t))X(t)$ is the amount in the bond. The state equation describing the dynamics of wealth process $X(t)$, for $t \in [0,T]$ as in (cf. [20]), on which we add the regime-switching and the control on the diffusion term is given by:

$$
\begin{cases} dX(t) = X(t)\big[r(1-u(t)dt + u(t)\dfrac{dS_t}{S_t}\big] \\ \\ X(0) = X_0, \qquad \alpha(0) = e_1, \quad t \in [0,T] \end{cases} \tag{5.1.3}
$$

where $X_0$, and $e_1$ represent respectively the initial wealth and the initial state of the Markov process $\alpha(\cdot)$.

Since we have $L_t = \log S_t$, we have: $dL_t = d(\log S_t) = \dfrac{dS_t}{S_t}$. Using equation (5.1.2), we obtain:

$$
dS_t = S_t\Big[c(\bar{L}_t - L_t)dt + g(t)d\tilde{W}(t) + \gamma(t)d\tilde{\Phi}(t) + \frac{1}{2}(g(t))^2 dt\Big]
$$

We wish to maximize the long term exponential growth rate of the expected utility (HARA) of $\big(\frac{1}{\theta}X_T^\theta\big)$ as $T \to \infty$, where $T > 0$ is some terminal time.

$$
\tilde{J}(u(\cdot)) = \frac{1}{\theta}\tilde{\mathbb{E}}[X(T)^\theta] \tag{5.1.4}
$$

Applying It$\hat{o}$'s formula to $\log X(t)^\theta = \theta \log X(t)$, we obtain:

$$
\begin{aligned}
d(\theta \log X(t)) &= \theta\Big[\frac{dX(t)}{X(t)} - \frac{1}{2}\,\big(\frac{dX(t)}{X(t)}\big)^2\Big] \\
&= \theta\Big[r\big(1-u(t)\big)dt + u(t)\frac{dS(t)}{S(t)} - \frac{1}{2}u^2(t)\big(\frac{dS(t)}{S(t)}\big)^2\Big] \\
&= \theta\Bigg[\Big(r\big(1-u(t)\big) + u(t)c(\bar{L}_t - L_t) - \frac{1}{2}u^2(t)g^2(t)\Big)dt \\
&\qquad + u(t)g(t)d\tilde{W}(t) + \Big(u(t)\gamma(t) - \frac{1}{2}u^2(t)\gamma^2(t)\Big)d\tilde{\Phi}(t)\Bigg].
\end{aligned}
$$

Therefore, we have the following expected function:

$$\mathbb{E}(X_T^\theta) = (X_0)^\theta \tilde{\mathbb{E}}\Bigg[\exp\Bigg\{\theta\Bigg(\int_0^T \Big(r\big(1-u(t)\big) + u(t)c(\bar{L}_t - L_t) \\ -\frac{1}{2}u^2(t)g^2(t)\Big)\, dt \\ +\theta\int_0^T u(t)g(t)d\tilde{W}(t) + \theta\int_0^T \big(u(t)\gamma(t) - \frac{1}{2}u^2(t)\gamma^2(t)\big)d\tilde{\Phi}(t)\Bigg)\Bigg\}\Bigg]$$

We use the following Girsanov transformation to eliminate the stochastic integral.

$$\frac{dP}{d\tilde{P}} = e^{\left(\theta\int_0^T \sigma u(t)d\tilde{W}(t) - \frac{1}{2}\theta^2\int_0^T \sigma^2 u^2(t)dt - \theta\int_0^T(\tilde{\lambda}-\lambda)dt(\frac{\tilde{\lambda}}{\lambda})^{N_T}\right)}$$

This change of probability measure argument is valid under the following assumption: The Markov regime-switching contains jumps that follows the Poisson distribution. We also consider

$$\tilde{\mathbb{E}}\Big[e^{(\mu u(t)^2)}\Big] \leq C, \tag{5.1.5}$$

where $\mu, C$ are positive constants. The equation (5.1.3) becomes

$$dL(t) = \big[c\alpha(t)\big(\bar{L}(t) - L(t)\big) + \theta g^2(t)u(t)\big]dt + \alpha(t)u(t)g(t)dW(t) + \alpha(t)u(t)d\Phi(t), \tag{5.1.6}$$

where $W(t)$ is a Brownian motion under another probability measure $P$ and

$$\tilde{\mathbb{E}}[X(T)^\theta] = X_0^\theta \mathbb{E}\Big[e^{\theta\left[\int_0^T F\left(X(t),u(t)\right)dt + G\left(X(T),\alpha(T)\right)\right]}\Big],$$

where

$$F(X(t), u(t)) = r(1-u(t)) + cu(t)(\bar{L}_t - L_t) - \frac{1}{2}u^2(t)g^2(t) + \frac{1}{2}u(t)g^2(t) + \theta\frac{1}{2}u^2(t)g^2(t) \\ +\ \theta\frac{1}{2}\big(u(t)\gamma(t) - \frac{1}{2}u^2(t)\gamma^2(t)\big)^2$$

$$G\big(X(T),\alpha(T)\big) = \alpha(T)X(T)$$

$$g\big(u(t),\alpha(t)\big) = \alpha(t)u(t)$$

$$f\big(X(t),u(t),\alpha(t) = -c\alpha(t)X(t) + \theta g^2(t)u(t)$$

We interpret the equation (5.1.6) as the dynamic of a risk-sensitive stochastic control problem with Markov regime-switching, in which $L_t$ is the state and $u(t)$ the control at time $t$. Assuming that $u(\cdot)$ is $\mathscr{F}_t$-progressively measurable for the filtration $\mathscr{F}_t$, we call such $u(\cdot)$ an admissible control and we denote by $\mathscr{U}_{ad}$ the admissible control set. Let us adopt the following transformation for suitable convenience.

$$X(t) = L_t - \bar{L}_t + \frac{m}{c}, \text{ where } \frac{m}{c} \text{ is a constant.}$$

We then obtain stochastic equation defined by:

$$\begin{cases} dX(t) = \big(-c\alpha(t)X(t) + \theta\alpha^2(t)u^2(t)u(t)\big)dt + \big(\alpha(t)u(t)\big)dW(t) + \big(\alpha(t)u(t)\big)d\Phi(t) \\ X(0) = \dfrac{m}{c} \end{cases} \tag{5.1.7}$$

and the cost functional:

$$J\big(u(\cdot)\big) = \mathbb{E}\Big[e^{\theta\big[\int_0^T F\big(X(t),u(t)\big)dt + G\big(X(T),\alpha(T)\big)\big]}\Big]$$

Under the following assumptions

*(A1):* $f, g, \gamma$ are uniformly Lipschitz in $(x, u)$

*(A2):* $f, g, \gamma, F, G$ are twice continuously differentiable with respect to $x$, and their partial derivatives in $x$ are continuous in $(x, u)$;

*(A3):* $f_x, f_{xx}, g_x, g_{xx}, F_x, F_{xx}, \|\gamma_x\|_{\mathcal{M}^{2p}}, \|\gamma_{xx}\|_{\mathcal{M}^2}, p = 1, 2$ and $G_x, G_{xx}$ are bounded;

*(A4):* $F$ and $G$ are uniformly bounded;

*(A5):* $V^\theta \in C^{1,3}([0,T] \times \mathbb{R}^n \times \mathcal{S})$;

*(A6):* $\mathcal{U}$ is a convex subset of $\mathbb{R}^k$.

We define the first and second adjoint variables respectively as follow:

$$\begin{cases} d\bar{p}(t) = \Big[c\alpha(t)\bar{p}(t) + cu(t) - \theta\bar{p}(t)\alpha(t)u(t)\bar{q}(t)\Big]dt + \bar{q}dW(t) + \sum_{j=1}^{2} \bar{s}_j(t)d\Phi_j(t) \\ \bar{p}(T) = \alpha(T) \end{cases} \tag{5.1.8}$$

$$\begin{cases} d\bar{P}(t) = -\Big[ - 2c\alpha(t)\bar{P}(t) + \theta\bar{p}(t)\alpha(t)u(t)\bar{Q}(t) + \theta\bar{q}^2(t) \\ \qquad + \sum_{j=1}^{2} \Big(\Lambda_j(t)\bar{S}_j(t) + \big(1 + \Lambda_j(t)\big)\theta\bar{s}_j^2(t)\Big)\alpha_j(t)u_j(t)\Big]dt \\ \qquad + \bar{Q}_j(t)dW_j(t) + \sum_{j=1}^{2} \bar{S}_j(t)d\Phi_j(t), \\ \bar{P}(T) = 0 \end{cases} \tag{5.1.9}$$

We have the following Hamiltonian $\mathscr{H}$-function:

$$\begin{aligned} \bar{\mathscr{H}}^\theta(t.X, u) := {} & \big( - c\alpha(t)X(t) + \theta\alpha^2(t)u^2(t)\big)\bar{p}(t) + cu(t)X(t) + (r-m)u(t) - r \\ & + \frac{1}{2}\alpha^2(t)u^4(t) - \frac{1}{2}\alpha^2(t)u^3(t) \\ & + \frac{1}{2}\alpha^2(t)u^2(t)\big(\bar{P}(t) - \theta\bar{p}^2(t)\big) + \alpha(t)u(t)\Big(\bar{q}(t) - \bar{P}(t)\alpha(t)u(t)\Big) \end{aligned} \tag{5.1.10}$$

$$\mathscr{H}_u^\theta(t,X,u) = 2\theta\alpha^2(t)p(t)u(t) + cX(t) + (r-m) + 2\alpha^2(t)u^3(t) - \frac{3}{2}\alpha^2(t)u^2(t) + \alpha^2(t)\left(P(t) - \theta p^2(t)\right)u(t) + \alpha(t)q(t) - 2\alpha^2(t)P(t)u(t) \tag{5.1.11}$$

Equating (5.1.11) to zero gives us the following equation to solve for u(t).

$$\mathscr{H}_u^\theta(t,X,u) = 0$$

Unfortunately, we do not have enough resources to solve this equation for $u(t)$ at this time. This will be our next investigation.

# 6 Conclusion

In this thesis, we reviewed the basic concepts of the probability theory, ordinary and stochastic differential equations, and the principles of Markov chain necessary and unavoidable for this work. We conducted a survey of the literature on stochastic optimization problem using the Pontryagin Maximum Principle. Then, we derived a risk-sensitivity stochastic maximum principle problem with regime-switching. And finally, we studied an example of its application on an optimal portfolio choice problem in the financial market using the hyperbolic absolute risk aversion (HARA) utility function and the linear exponential of quadratic Gaussian (LEQG). The maximum principle for this kind of problem is obtained. However, we notice that the first and second adjoint equations depend deliberately on the risk-sensitive parameter, as well as the maximum condition. In future work, we plan to extend our study to a more general case using the Poisson jump diffusion systems and compute our result to explicitly illustrate the optimal solution.

# References

[1] Bielecki, Tomasz R., and Stanley R. Pliska. "Risk-sensitive dynamic asset management." *Applied Mathematics and Optimization* 39, no. 3 (1999): 337-360.

[2] Crépey, Stéphane. "About the pricing equations in finance." *In Paris-Princeton Lectures on Mathematical Finance 2010*, pp. 63-203. Springer, Berlin, Heidelberg, 2011.

[3] Davis, Mark, and Sébastien Lleo. "Risk-sensitive benchmarked asset management." *Quantitative Finance* 8, no. 4 (2008): 415-426.

[4] Davis, Mark, and Sébastien Lleo. "Jump-diffusion risk-sensitive asset management I: diffusion factor model." *SIAM Journal on Financial Mathematics* 2, no. 1 (2011): 22-54.

[5] Davis, Mark, and Sebastien Lleo. "Jump-diffusion risk-sensitive asset management II: jump-diffusion factor model." *SIAM Journal on Control and Optimization* 51, no. 2 (2013): 1441-1480.

[6] Djehiche, Boualem, Hamidou Tembine, and Raul Tempone. "A stochastic maximum principle for risk-sensitive mean-field type control." *IEEE Transactions on Automatic Control* 60, no. 10 (2015): 2640-2649.

[7] Elliott, Robert J., Lakhdar Aggoun, and John B. Moore. *Hidden Markov models: estimation and control*. Vol. 29. Springer Science & Business Media, 2008.

[8] Hamilton, James Douglas. *Time series analysis.* Vol. 2. Princeton, NJ: Princeton university press, 1994.

[9] Jingtao, Shi, and Wu Zhen. "A risk-sensitive stochastic maximum principle for optimal control of jump diffusions and its applications." *Acta Mathematica Scientia* 31, no. 2 (2011): 419-433.

[10] Karatzas, Ioannis, and Steven E. Shreve. "Brownian motion." In *Brownian Motion and Stochastic Calculus*, pp. 47-127. Springer, New York, NY, 1998.

[11] Kolmogorov, Andrei Nikolaevich. "On logical foundations of probability theory." In *Probability theory and mathematical statistics*, pp. 1-5. Springer, Berlin, Heidelberg, 1983.

[12] Li, Yusong, and Harry Zheng. "Weak necessary and sufficient stochastic maximum principle for markovian regime-switching diffusion models." *Applied Mathematics & Optimization* 71, no. 1 (2015): 39-77.

[13] Lim, Andrew EB, and Xun Yu Zhou. "A new risk-sensitive maximum principle." *IEEE transactions on automatic control* 50, no. 7 (2005): 958-966.

[14] Nagai, Hideo, and Shige Peng. "Risk-sensitive dynamic portfolio optimization with partial information on infinite time horizon." *Annals of Applied Probability* (2002): 173-195.

[15] Øksendal, Bernt Karsten, and Agnes Sulem. *Applied stochastic control of jump diffusions*. Vol. 498. Berlin: Springer, 2005.

[16] Oksendal, Bernt. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[17] Peng, Shige. "A general stochastic maximum principle for optimal control problems." *SIAM Journal on control and optimization* 28, no. 4 (1990): 966-979.

[18] Rogers, L. Chris G., and David Williams. *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*. Vol. 2. Cambridge university press, 2000.

[19] Shafer, Glenn, and Vladimir Vovk. "The Sources of Kolmogorov's" Grundbegriffe"." *Statistical Science* (2006): 70-98.

[20] Shi, Jingtao, and Zhen Wu. "Maximum principle for risk-sensitive stochastic optimal control problem and applications to finance." *Stochastic Analysis and Applications* 30, no. 6 (2012): 997-1018.

[21] Sun, Zhongyang, Isabelle Kemajou-Brown, and Olivier Menoukeu-Pamen. "A risk-sensitive maximum principle for a Markov regime-switching jump-

diffusion system and applications." *ESAIM: Control, Optimisation and Calculus of Variations* 24, no. 3 (2018): 985-1013.

[22] Tang, Shanjian, and Xunjing Li. "Necessary conditions for optimal control of stochastic systems with random jumps." *SIAM Journal on Control and Optimization* 32, no. 5 (1994): 1447-1475.

[23] Whittle, Peter. "Risk-sensitive linear/quadratic/Gaussian control." *Advances in Applied Probability* 13, no. 4 (1981): 764-777.

[24] Whittle, P. "A risk-sensitive maximum principle." *Systems & Control Letters* 15, no. 3 (1990): 183-192.

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

[26] Zhang, Xin, Robert J. Elliott, and Tak Kuen Siu. "A stochastic maximum principle for a Markov regime-switching jump-diffusion model and its application to finance." *SIAM Journal on Control and Optimization* 50, no. 2 (2012): 964-990.