Portfolio Optimization & Option Pricing in a Component-wise Semi-Markov Modulated Market

A thesis
submitted in partial fulfillment of the requirements
of the degree of

Doctor of Philosophy

by

Milan Kumar Das
ID: 20133275

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

September, 2018

Dedicated to

my mother and my advisor Anindya Goswami

Certificate

Certified that the work incorporated in the thesis entitled submitted by *Milan Kumar Das* was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Date: *Dr. Anindya Goswami*

Thesis Supervisor

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date:

Milan Kumar Das

Roll Number: 20133275

Contents

Acknowledgments

I would like to express my deepest gratitude to my advisor Dr. Anindya Goswami for his excellent guidance, continuous support, enthusiasm, and patience. I feel lucky to be the first graduate student of him. Besides his professional mentorship, I also acknowledge the everyday wisdom which he offered to me. I value his constructive counseling and advice which I received during my nonproductive days. I wish I could live up to his words.

Besides my supervisor, I would like to thank the rest of my research advisory committee members, Prof. Suresh Kumar and Dr. Anup Biswas for their insightful comments and encouragement. I am grateful to Prof. Uttara Naik Nimbalkar for helpful discussion. I would like to thank Dr. Kedar Mukherjee of National Institute of Bank Management for some very needful help.

I am thankful to Dr. Anup Biswas to provide me an opportunity to learn probability theory in an undergraduate course. I am grateful to Dr. Anisa Chorwadwala for teaching me ordinary and partial differential equations in graduate coursework. I acknowledge Dr. Soumen Maity, Dr. Kaneenika Sinha, Dr. Mousomi Bhakta for their graduate courses.

I am grateful to all the faculty members of the mathematics department of IISER Pune. I am thankful to the administrative staff of the institute. I also appreciate the support; I received from the Department of Mathematics throughout my Ph.D. years. I would like to acknowledge UGC for the financial support.

These last five years would not have been more educating and entertaining without an enlightening company of friends. I am indebted to my friends Girish, Jyotirmoy, Tathagata, Debangana, Sushil, Manidipa, Ayesha, Makrand, Rohit, Chaitanya, Amit, Rajesh, Uday, Hitesh, Prahlad, Soumyadip, Neeraj, Supratik. I am indebted to my wonderful flatmate Chitrabhanu Chaudhury.

Finally, I reckon that words are not enough to express my gratitude towards my mother, who dedicated her all life to me. Now big gestures by me for her would ever commensurate her sacrifices for me.

Milan Kumar Das

Abstract

This thesis studies three problems of mathematical finance. We address the appropriateness of the use of semi-Markov regime switching geometric Brownian motion (GBM) to model risky assets using a statistical technique. Component-wise semi-Markov (CSM) process is a further generalization of the semi-Markov process, which becomes relevant when multiple assets are under consideration. In this thesis, we would present the solution to the optimization problem of portfolio-value, consisting of several stocks under risk-sensitive criterion in a component-wise semi-Markov regime-switching jump diffusion market. Finally, the solution to locally risk minimizing pricing of a broad class of European style basket options would be demonstrated under a market assumption where the risky asset prices follow CSM modulated time inhomogeneous geometric Brownian motion.

Introduction

Mathematical finance began its journey with the pioneering work of a French mathematician L. Bachelier in 1900. Bachelier first introduced randomness to model risky asset price. In his work, he used Brownian motion by using the time limit of random walk to model the stock price. After a long time, P. A. Samuelson proposed the geometric Brownian motion (GBM) to model the stock price in 1965 to capture the non-negativity of stock price dynamics. But only after the groundbreaking works of Black, Scholes, and Merton in 1973, GBM became popular in modeling the risky asset price dynamics. Several empirical studies are against the GBM modeling. The main two drawbacks with the GBM hypothesis are:

1. GBM model implies the simple returns are normally distributed,
2. this model assumes that the volatility is constant.

In view of these, the researchers became interested in the regime switching models, introduced in mathematical finance by J. D. Hamilton in 1989 [33]. In regime switching models, it is assumed that there are several unobserved states in the market whose jump is governed by a pure jump process and the market parameters changes their values as the state changes. We call each state of the coefficients as a regime and the dynamics as a regime switching model. Many researchers have implemented the Markov switching models or hidden Markov models in various studies, e.g., see [2], [48], [37], [18], [17], [63], [64] etc. By an empirical study in [55], the authors have claimed that all the stylized facts of daily return series can not be captured by using hidden Markov models. Semi-Markov switching models are the other possible alternatives, relatively new to the theoretical studies. One may find applications of these models in [25], [29], [10], [11]. This list is merely indicative but not exhaustive by any means. In [7], the authors have shown by empirical studies, the hidden semi-Markov models can describe the stylized facts better to the previous model.

As per our knowledge, there is no comprehensive statistical analysis which helps discriminating among the cases of GBM, Markov-modulated GBM(MMGBM) and semi-Markov modulated GBM(SMGBM) for modeling a given asset price time series data. We investigate the appropriateness of the use of SMGBM by developing a statistical technique. We propose a discriminating statistic whose sampling distribution varies drastically, under the regime switching assumption, with varying values of instantaneous rate parameter. We utilize this statistics to test the model hypothesis for Indian sectoral indices. Strictly speaking, modeling of a market consisting of different assets, governed by a single semi-Markov process is rather restrictive. Ideally those could be driven by independent or correlated processes in practice. Although two independent Markov processes jointly becomes a Markov process, the same phenomena is not true for semi-Markov processes. For this reason consideration of independent regimes are important where regimes are not Markov. We call the joint process (with each component as independent semi Markov) as component wise semi-Markov process which is abbreviated as CSM.

In this thesis, we consider a market with several stocks which are governed by a CSM process. Under this market assumptions, we address two theoretical problem (1) a portfolio optimization problem, (2) a European type basket option pricing problem.

A new characterization of general semi-Markov process was explored in [27]. In that, the semi-Markov process $\{X_t\}_{t\geq 0}$ on $\mathcal{X} := \{1,\ldots,k\} \subset \mathbb{R}$ is specified by a collection of measurable function $\lambda : \{(i,j) \in \mathcal{X}^2 | i \neq j\} \times [0,\infty) \to (0,\infty)$ and is defined by the strong solution of the following system of stochastic integral equations

$$X_t = X_0 + \int_{(0,t]} \int_{\mathbb{R}} h_\lambda(X_{u-}, Y_{u-}, z) \wp(du, dz) \tag{1}$$

$$Y_t = t - \int_{(0,t]} \int_{\mathbb{R}} g_\lambda(X_{u-}, Y_{u-}, z) \wp(du, dz), \tag{2}$$

where $\wp(du, dz)$ is a Poisson random measure with intensity $dudz$, independent of X_0 and h_λ, g_λ are appropriately chosen by

$$h_\lambda(i,y,z) := \sum_{j\in\mathcal{X}\setminus\{i\}} (j-i) 1_{\Lambda_{ij}(y)}(z), \quad g_\lambda(i,y,z) := y \sum_{j\in\mathcal{X}\setminus\{i\}} 1_{\Lambda_{ij}(y)}(z),$$

where for each $y \geq 0$, and $i \neq j$, $\Lambda_{ij}(y)$ are the consecutive (with respect to the lexicographic ordering on $\mathcal{X} \times \mathcal{X}$) left closed and right open intervals on the real line, each having length $\lambda_{ij}(y)$ starting from the origin. We clarify that if $\{(X_t, Y_t)\}_{t\geq 0}$ is the solution to (1)-(2), then Y_t is called the age process. It is shown in (Th. 2.1.3, [50]) that λ becomes the instantaneous transition rate of X.

First we describe in brief the original contribution of this thesis in in portfolio optimization problem. Because of the abrupt nature of the stock price return, given an outlay, it is challenging to make an optimal portfolio of an investor. Among a plenty of investment options, a rational investor need to set his/her investment policy according to the risk tolerance. There are few mathematical methods available in order to make decisions under risky investments. Among them

1. the mean-variance analysis,

2. the utility optimization,

are very popular among the market practitioner.

Following the seminal work of Markowitz [45], the problem of optimization of an investor's portfolio based on different criteria and market assumptions are being studied by several authors. In the mean-variance optimization approach, as done by Markowitz, either the expected value of portfolio wealth is optimized by keeping the variance fixed or the variance is minimized by keeping the expectation fixed. Though Markowitz's mean-variance approach to portfolio is immensely useful in practice, its scope is limited by the fact that only Gaussian distributions are completely determined by their first two moments.

The utility optimization technique is easier and robust for decision making than the mean-variance approach. In a pioneering work, Merton [46], [47] has introduced the utility maximization to the optimal portfolio selection. In this approach, instead of optimizing the expected value of wealth R, the expected value of some continuous increasing function $U(R)$ is to be optimized. The function U is known as utility function. Some standard utility functions are $-e^{-ax}$, $\ln(x)$, bx^b, etc, where the parameter a and b are the risk tolerance of the investor.

Merton's approach is based on applying the method of stochastic optimal control via an appropriate Hamilton-Jacobi- Bellman (HJB) equation. The corresponding optimal dynamic portfolio allocation can also be obtained from the same equation. Although this approach has greater mathematical tractability but does not capture the tradeoff between maximizing expectation and minimizing variance of portfolio value.

There is another approach, namely risk sensitive optimization, where a tradeoff between the long run expected growth rate and the asymptotic variance is captured in an implicit way. The aforesaid utility maximization method can be employed to study the risk-sensitive optimization by choosing a parametric family of exponential utility functions. In such optimization, an appropriate value of the parameter is to be chosen by the investor depending on the investors degree of risk tolerance. We refer to [3], [20], [21], [43] for this criterion under the geometric Brownian motion (GBM) market model.

Risk sensitive optimization of portfolio value in a more general type of market is also studied by various authors. Jump diffusion model is one such generalization, which captures the discontinuity of asset dynamics. Empirical results support such models [15]. Terminal utility optimization problem under such a model assumption is studied by [39]. In all these references, it is assumed that the market parameters, i.e., the coefficients in the asset price dynamics, are either constant or deterministic functions of time. We study a class of models where these parameters are allowed to be finite state pure jump processes.

Risk sensitive portfolio optimization in a GBM model with Markov regimes is studied in [24] whereas [26] studies that in a semi-Markov modulated GBM model. In [26] the market parameters, r, μ^l and σ^l are driven by a finite-state semi-Markov process $\{X_t\}_{t\geq 0}$, where μ^l and σ^l denote the drift and volatility parameters of l-th asset in the portfolio. Here we consider a market consisting of several stocks is modeled by a multi-dimensional jump diffusion process with CSM modulated coefficients.

We study the finite horizon portfolio optimization via the risk sensitive criterion under the above market assumption. The optimization problem is solved by studying the corresponding HJB equation, where we employ the technique of separation of variables to reduce the HJB equation to a system of linear first order PDEs containing some non-local terms. In the reduced equation, the nature of non-locality is such that the standard theory of integro-PDE is not applicable to establish the existence and uniqueness of the solution. In this thesis, to show well-posedness of this PDE, a Volterra integral equation(IE) of the second kind is obtained and then the existence of a unique C^1 solution is shown. Then it is proved that the solution to the IE is a classical solution to the PDE under study. The uniqueness of the PDE is proved by showing that any classical solution also solves the IE. In the uniqueness part, we use conditioning with respect to the transition times of the underlying process. Besides, we also obtain the optimal portfolio selection as a continuous function of time and underlying switching process. The expression of this function does not involve the market transition rate parameter λ. Thus the optimal selection is robust. This study, as alluded above is presented in Chapter 4 of this thesis. In the 5-th Chapter we investigate an option pricing problem.

The modern theory of option pricing is fathered by L. Bachelier. Though his work did not get the recognition for a long time. Bachelier derives the theoretical option prices where the stock price is modeled as a Brownian motion with drift. The main flaw of his modeling was the chances of negative stock prices. In 1973 Black, Scholes and Merton considered a different mathematical model of asset price dynamics to find an expression of the price of a European option on the underlying asset. In their model, the stock price process is modeled with a geometric Brownian motion. The drift and the volatility coefficients of the price were taken as constants. Though this model is widely accepted because of simplicity, the variability of market parameters can not be captured by using this model. One serious drawback of their assumptions is the Gaussianity of stock price return.

Since then, numerous different improvements of their theoretical model are being studied. Regime switching models are one such extension of the Black-Scholes-Merton (BSM) model. Extensive research has been done to study markets with Markov-modulated regime switching [2],[6],[12],[13],[16],[31],[32],[36],[44]. However, the consideration of Markov regimes is not confined in generalizing BSM model only. Regime switching GARCH option models has been studied in [14]. There are also some studies, carried out by several authors, involving regime switching extension of other alternative models of asset price.

These include jump diffusion models, stochastic volatility model etc. In all these works the possibility of switching regimes is restricted to the class of finite state Markov Chains.

In comparison with Markov switching, the study of semi-Markov switching is relatively uncommon. In this type of models one has opportunity to incorporate some memory effect of the market. In particular, knowledge of the past stagnancy period can be fed into the option price formula to obtain the price value. Hence this type of models have greater appeal in terms of applicability than the one with Markov switching. It is shown in [7], by studying sectoral daily data, that the sojourn times of certain regimes have heavier tail than exponential. In particular, the standard deviations are consistently larger than mean. Hence, their study suggests that, the semi-Markov switching models have the capabilities to describe the stylized facts better than Markov model.

The pricing problem with semi-Markov regimes was first solved in [25]. It is important to note that the regime switching models lead to incomplete markets. Since there might be multiple no arbitrage prices of a single option, one needs to fix an appropriate notion to obtain an acceptable price. Locally risk minimizing option pricing with a special type of age-independent semi-Markov regime is studied in [25] using Föllmer-Schweizer decomposition [23]. There it is shown that the price function satisfies a non-local system of degenerate parabolic PDE. In a recent paper [30] the same problem for a more general class of age-dependent semi-Markov processes is studied. The option pricing problem under stochastic volatility model with age-dependent semi-Markov parameters is addressed in [4].

In many regime-switching models of asset price dynamics, the volatility coefficients do not posses explicit time dependence (see [2],[6],[12],[13],[16],[25],[30],[31],[32],[36],[44]). In such time homogeneous models the volatility σ can take values from a finite set only. Such models fail to capture many other stylized facts including periodicity feature of σ. In the present model, we allow σ to be time inhomogeneous.

We consider a market with one locally risk free asset with price process S^0, and n risky assets with prices $\{S^l\}_{l=1,\ldots,n}$, and address locally risk-minimizing pricing for a contingent claim $K(S_T)$. Here we consider a class of Lipschitz continuous functions $K : \mathbb{R}_+^n \to \mathbb{R}_+$, which includes vanilla basket options. We show that the locally risk minimizing price of the claim at time t, when (S_t^l, X_t^l, Y_t^l) is (s^l, x^l, y^l), for each l, is a function φ of $(t, s = (s^1, s^2, \ldots, s^n), x = (x^0, x^1, \ldots, x^n), y = (y^0, y^1, \ldots, y^n))$ and that satisfies a Cauchy problem. In order to write the equation we use a notation $R_j^l v$, for a vector $v \in \mathbb{R}^{n+1}$ to denote the vector $v + (j - v^l)e_l$, in which the l-th component of v is replaced with j. The system of PDE is given by

$$\frac{\partial \varphi}{\partial t}(t,s,x,y) + \sum_{l=0}^{n} \frac{\partial \varphi}{\partial y^l}(t,s,x,y) + r(x) \sum_{l=1}^{n} s^l \frac{\partial \varphi}{\partial s^l}(t,s,x,y)$$
$$+ \frac{1}{2} \sum_{l=1}^{n} \sum_{l'=1}^{n} a^{ll'}(t,x) s^l s^{l'} \frac{\partial^2 \varphi}{\partial s^l \partial s^{l'}}(t,s,x,y)$$

$$+\sum_{l=0}^{n}\sum_{j\neq x^l}\lambda^l_{x^l j}(y^l)\left[\varphi(t,s,R^l_j x,R^l_0 y)-\varphi(t,s,x,y)\right]=r(x)\ \varphi(t,s,x,y), \tag{3}$$

defined on

$$\mathcal{D}:=\{(t,s,x,y)\in(0,T)\times(0,\infty)^n\times\mathcal{X}^{n+1}\times(0,T)^{n+1}\mid y\in(0,t)^{n+1}\},$$

and with conditions

$$\varphi(T,s,x,y)=K(s);\qquad s\in[0,\infty)^n;\qquad 0\le y^l\le T;\qquad x^l\in\mathcal{X},\qquad l=0,1,\dots,n, \tag{4}$$

where the diffusion coefficient $a:=(a^{ll'})_{n\times n}$ is continuous in t.

We note that (3) is a linear, parabolic, degenerate and non-local PDE. The non-locality is due to the occurrence of the term $\varphi(t,s,R^l_j x,R^l_0 y)$. Furthermore the terminal data (4) need not be in the domain of the operator in (3). We establish existence and uniqueness of the classical solution of (3)-(4) in this thesis via a Volterra integral equation (VIE) of second kind. Using the Banach fixed point Theorem, we show that the integral equation has a unique solution. We show that the VIE is equivalent to the PDE. Thus we show that one can find the price function by solving the integral equation which is computationally more convenient (see [31] for more details) than solving the PDE. We also obtain an expression of optimal hedging involving integration of price function.

A concise effort is made to prepare this thesis self contained and accessible. This thesis consists of 5 chapters and an appendix. Some background material of probability, stochastic processes and mathematical finance is recalled in Chapter 1. In chapter 2 we describe a continuous time model testing technique and its application in financial market. In this chapter, we also present some results of Indian stock market, which suggests semi-Markov regime switching models are more appealing. In view of the results in Chapter 2, we present some theories of component wise semi-Markov processes (CSM) which is more general than semi-Markov process in Chapter 3. In Chapter 4, we study a risk sensitive portfolio optimization problem in a CSM modulated jump diffusion market. A European type basket option has been studied in a CSM modulated geometric Brownian motion market in Chapter 5. The appendix consists of some algorithms used in Chapter 2.

1

Preliminaries

We introduce the established theories we have used throughout this thesis briefly. This chapter intends to review some basic theories such that this thesis become readable. We also give references to some excellent texts where they can be found.

1.1 Stochastic Process

Let $(\Omega, \mathcal{F}, P)$ be a probability space. The set Ω is said to be the sample space and $\mathcal{F}$ is the σ-algebra containing all the events and P is the probability measure. We say a probability space is complete if $\mathcal{F}$ contains all the P-null sets. All the definitions and theorems can be found in some excellent texts, e.g. Çınlar [9], Shiryaev [58], Protter [52], Karatzas & Shreve [40], Ikeda & Watanbe [35] etc.

1.1.1 Probability

In this subsection, we recall some basic definitions and theorems of probability theory. We also provide brief information of some distribution, which is useful in this thesis. We begin with the definition of a random variable.

Definition 1.1.1. *Let $(E, \mathcal{G})$ be a measurable space. A map $X : \Omega \to E$ is said to be a* ***random variable*** *taking the values in $(E, \mathcal{G})$ such that it is measurable relative to $\mathcal{F}$ and $\mathcal{G}$, i.e., if for any $A \in \mathcal{G}, X^{-1}A \in \mathcal{F}$.*

Now we define one of the most fundamental concepts of the theory of probability, the expectation, and conditional expectation. The concept of conditional expectation is extensively used in applied probability.

Definition 1.1.2. *Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$. Then the integral of X with respect to measure P is said to be **expectation** of X, denoted by $\mathbb{E}X$, and defined by $\mathbb{E}X := \int_{\Omega} X(\omega)P(d\omega)$.*

Definition 1.1.3. *Let $\mathcal{H}$ be a sub σ-algebra of $\mathcal{F}$. The **conditional expectation** of a non-negative random variable X with respect to $\mathcal{H}$ is a non-negative random variable, denoted by $\mathbb{E}_{\mathcal{H}}(X)$ or by $\mathbb{E}[X|\mathcal{H}]$ such that*

i. $\mathbb{E}_{\mathcal{H}}(X)$ *is* $\mathcal{H}$*-measurable.*

ii. for every $A \in \mathcal{H}$,

$$\int_A X dP = \int_A \mathbb{E}_{\mathcal{H}}(X) dP.$$

The conditional expectation of any random variable X with respect $\mathcal{H}$, if $\mathbb{E}X$ exists, is given by $\mathbb{E}_{\mathcal{H}}(X) := \mathbb{E}_{\mathcal{H}}(X^+) - \mathbb{E}_{\mathcal{H}}(X^-)$, where $X^+ := \max\{X, 0\}$ and $X^- := \max\{-X, 0\}$; Otherwise, if $\mathbb{E}X^+ = \mathbb{E}X^- = \infty$, then $\mathbb{E}_{\mathcal{H}}(X)$ is undefined.

Now we state one of the most fundamental properties of conditional expectation, namely the **Tower Property:**.

Theorem 1.1.4. *Let $\mathcal{H}_1, \mathcal{H}_2$ be two sub-σ algebras of $\mathcal{F}$, then the following holds.*

(a) If $\mathcal{H}_1 \subseteq \mathcal{H}_2$, *then* $\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{H}_2\right]|\mathcal{H}_1\right] = \mathbb{E}\left[X|\mathcal{H}_1\right]$ (a.s.).

(b) If $\mathcal{H}_1 \supseteq \mathcal{H}_2$, *then* $\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{H}_2\right]|\mathcal{H}_1\right] = \mathbb{E}\left[X|\mathcal{H}_2\right]$ (a.s.).

Now we shall define two most important distributions which play a crucial role in this thesis namely the exponential distribution and the log-normal distribution.

Definition 1.1.5. *A random variable X taking values in $\mathbb{R}_+$ is said to follow **exponential distribution** with parameter λ (we write $X \sim Exp(\lambda)$), if it has the p.d.f of the following form*

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & otherwise. \end{cases}$$

It is important to note that if $X \sim \text{Exp}(\lambda)$ then $\mathbb{E}X = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$, where $\text{Var}(\cdot)$ denote the variance.

Definition 1.1.6. *A random variable X taking values in $\mathbb{R}$ is said to follow **lognormal distribution** with parameters μ and σ (we write $X \sim LN(\mu, \sigma)$), if it has the p.d.f of the following form*

$$f(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} & x > 0 \\ 0 & otherwise. \end{cases}$$

We note that if $X \sim \mathrm{LN}(\mu, \sigma)$ then its logarithm $\ln X$ is normally distributed,

$$\ln X \sim \mathcal{N}(\mu, \sigma),$$

where $\mathcal{N}(\mu, \sigma)$ to denote normal distribution with mean μ and standard deviation σ. The mean and variance of a random variable X following $\mathrm{LN}(\mu, \sigma)$ is given by

$$\mathbb{E}X = \exp(\mu + \frac{\sigma^2}{2}) \tag{1.1}$$

$$\mathrm{Var}(X) = \Big[\exp(\sigma^2) - 1\Big] \exp(2\mu + \sigma^2). \tag{1.2}$$

1.1.2 Stochastics

In this subsection we present some basics of stochastic process, which would be useful in our future studies.

Definition 1.1.7. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(E, \mathcal{G})$ be a measurable space.*

1. *For each $t \geq 0$, let X_t be a random variable taking values in E. Then the collection $\{X_t\}_{t \geq 0}$ is said to be a* ***stochastic process*** *with state space $(E, \mathcal{G})$.*

2. *For each $t \geq 0$, let $\mathcal{F}_t$ be a sub σ-algebra of $\mathcal{F}$. The family $\{\mathcal{F}_t\}_{t \geq 0}$ is said to be* ***filtration*** *such that $\mathcal{F}_s \subset \mathcal{F}_t$ for $s < t$.*

3. *A random variable $T : \Omega \to [0, \infty]$ is said to be a* ***stopping time****, if the event $\{T \leq t\} \in \mathcal{F}_t$, for all $t \geq 0$.*

4. *A process X is* ***adapted*** *to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if X_t is $\mathcal{F}_t$-measurable for all $t \geq 0$.*

5. *A stochastic process X is said to be* ***rcll*** *or* ***càdlàg*** *or* ***corlol*** [1] *if it has sample paths right continuous and left limit exists almost surely.*

The concept of martingale and local martingale plays a crucial role in this thesis. We first define them and then state the necessary theorems used in this thesis. We refer Protter [52], Karatzas & Shreve [40] for further details.

Definition 1.1.8. *Let $(\Omega, \mathcal{F}, \{F_t\}_{t \geq 0}, P)$ be a filtered probability space on which $X = \{X_t\}_{t \geq 0}$ be an adapted, rcll process.*

1. *Let $\mathbb{E}|X_t| < \infty$, $\forall\, t \geq 0$. Then X is said to be* ***martingale*** *(resp.* ***sub-martingale, super-martingale****) if $\mathbb{E}\left[X_t | \mathcal{F}_s\right] = X_s$ a.s. (resp.$\mathbb{E}\left[X_t | \mathcal{F}_s\right] \geq X_s$, resp. $\mathbb{E}\left[X_t | \mathcal{F}_s\right] \leq X_s$) holds for all $0 \leq s < t < \infty$.*

[1] Throughout this thesis we use the term rcll.

2. *The process X is a* ***local martingale*** *if there exists a sequence of increasing stopping times, $\{T_n\}_{n\geq 1}$, with $\lim\limits_{n\to\infty} T_n = \infty$ a.s. such that $X_{t\wedge T_n}$ is a martingale for each n.*

We present the concept of quadratic variation, which plays a central role in the theory of stochastic integration. For further details, we refer Föllmer [22].

Definition 1.1.9. *Let $X = \{X_t\}_{t\geq 0}$ and $Y = \{Y_t\}_{t\geq 0}$ be two stochastic processes on a probability space $(\Omega, \mathcal{F}, P)$.*

1. *Let $\{\Pi_n\}_{n=1,2,\ldots}$ be a sequence of finite partitions of the form $\Pi_n = \{0 = t_0 \leq t_1 \leq \ldots \leq t_{i_n}\}$ with $|\Pi_n| = \sup\limits_{t_i\in\Pi_n} |t_{i+1} - t_i| \to 0$ and $t_{i_n} \to \infty$. If for all $t \geq 0$, the weak* limit of*

$$\mu_n := \sum_{\substack{t_i\in\Pi_n \\ t_i\leq t}} |X_{t_{i+1}} - X_{t_i}|^2 \delta_{\{t_i\}},$$

exists then the distribution function $t \mapsto [X]_t$ of the limit μ given by $[X]_t := \int\limits_0^t d\mu$, is said to be the ***quadratic variation*** *of X. Furthermore,*

$$[X]_t = [X]_t^c + \sum_{s\leq t} \Delta X_s^2,$$

where, $[X]^c$ denotes the continuous part, $\Delta X_s := X_s - X_{s-}$ its jump and $\Delta X_s^2 := (\Delta X_s)^2$ the quadratic jump of X.

2. *The* ***cross variation*** *of X and Y is denoted by $[X,Y]$ and is defined by $[X,Y]_t := \frac{1}{4}\left[[X+Y]_t - [X-Y]_t\right]$.*

The next theorem demonstrates an excellent property of a martingale which is also well known as Novikov's condition for martingales. We refer to [35] (Theorem 5.2) for more details.

Theorem 1.1.10. *Let $X := \{X_t\}_{t\geq 0}$ be a continuous, square integrable $\{\mathcal{F}_t\}_{t\geq 0}$ martingale in a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$. Let $M_t = \exp\{X_t - \frac{1}{2}[X]_t\}$. If for all $t \geq 0$, $\mathbb{E}[e^{\frac{[X]_t}{2}}] < \infty$, then $\{M_t\}_{t\geq 0}$ is a continuous $\{\mathcal{F}_t\}_{t\geq 0}$- martingale.*

Now we shall state an important theorem involving the conditions for which a local martingale becomes a martingale. We refer to [52] (Theorem I.51 & Corollary 4 of Theorem II.27) for more details.

Theorem 1.1.11. *Let X be a local martingale.*

(1) Then X is a martingale if $\mathbb{E}[\sup\limits_{s\leq t} |X_s|] < \infty$, for all $t \geq 0$.

(2) If $\mathbb{E}[[X]_\infty] < \infty$, then X is a square integrable martingale.

We recall the definition of Brownian motion, the most important stochastic process in our thesis.

Definition 1.1.12. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a filtered probability space. An adapted process $W = \{W_t\}_{t\geq 0}$ with $W_0 = 0$ a.s. is said to be a **Brownian motion** if*

(i) $W_t - W_s$ is independent of $\mathcal{F}_s$ for $0 \leq s < t < \infty$.

(ii) $W_t - W_s$ follows Gaussian distribution with mean 0 and variance $t - s$.

(iii) $t \mapsto W_t$ is continuous with probability 1.

We recall the result of quadratic variation of a Brownian motion. We refer to [53] (Theorem 1.2.4) for details.

Theorem 1.1.13. *Brownian motion is of finite quadratic variation and $[B]_t = t$ a.s.*

Definition 1.1.14. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a filtered probability space. An adapted process $X = \{X_t\}_{t\geq 0}$ with $X_0 = 0$ a.s. is said to be a **Lévy process** if*

(i) $X_t - X_s$ is independent of $\mathcal{F}_s$ for $0 \leq s < t < \infty$.

(ii) $X_t - X_s$ has the same distribution as X_{t-s} for $0 \leq s < t < \infty$.

(iii) X is stochastically continuous, i.e. for all $\delta > 0$ and for all $s \geq 0$

$$\lim_{s \to t} P(|X_t - X_s| > \delta) \to 0.$$

Although in Protter [52], the following class is termed as a decomposable processes, a subclass of semimartingales, but we use the general term following [38].

Definition 1.1.15. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a filtered probability space. An adapted process $X = \{X_t\}_{t\geq 0}$ is said to be a **semimartingale** if it can be decomposed P a.s. as*

$$X_t = X_0 + M_t + A_t \quad t \geq 0,$$

where M_t is a local martingale and A_t is an rcll adapted process with locally bounded variation.

The stochastic integral with respect semimartingale with full generality can be found in Protter [52]. We shall only state Itô's formula below.

Theorem 1.1.16 (Itô's formula). *Let X be a semimartingale and f be real valued, twice continuously differentiable function. Then the following hold*

$$f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_{s-})dX_s + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})d[X]_s^c + \sum_{0<s\leq t} [f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s.]$$

Now we state a very useful theorem, known as **Girsanov's Theorem**. For more details we cite [40] (Theorem 3.5.1).

Theorem 1.1.17. *Let $W = \{W_t = (W_t^{(1)}, \dots, W_t^{(d)})\}$ be a d-dimensional Brownian motion with covariance matrix I in a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$. Let $t \geq 0$, $X = \{X_t = (X_t^{(1)}, \dots, X_t^{(d)})\}_{t\geq 0}$ be a vector valued $\mathcal{F}_t$-adapted process satisfying*

$$P\left[\int_0^T (X_t^{(i)})^2 dt < \infty\right] = 1;\ 1 \leq i \leq d, 0 \leq T < \infty.$$

Let $Z_t(X) := \exp\left[\sum_{i=1}^{d}\int_0^t X_s^{(i)} dW_s^{(i)} - \frac{1}{2}\int_0^t \|X_s\|^2 ds\right]$ be a martingale. Define a $\mathcal{F}_t$-measurable process $\tilde{W} = \{\tilde{W}_t = (\tilde{W}_t^{(1)}, \dots, \tilde{W}_t^{(d)})\}$ by $\tilde{W}_t^{(i)} := W_t^{(i)} - \int_0^t X_s^{(i)} ds$, $1 \leq i \leq d$, $0 \leq t < \infty$. Then for each fixed $T \in [0, \infty)$, the process $\{\tilde{W}_t\}$ is a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}_T, \tilde{P}_T)$, where $\tilde{P}_T(A) := \mathbb{E}\left[\mathbf{1}_A Z_T(X)\right]$, $A \in \mathcal{F}_T$.

1.1.3 Poisson random measure & Integration

In this section, we first prepare ourselves with the definition of Poisson process and Poisson random measure. A nice presentation of Poisson process and Poisson random measure can be found in [9] and [41]. Then we shall concentrate in the construction of pure jump process with Poisson random measure. Throughout this section we assume that $(\Omega, \mathcal{F}, P)$ is the underlying probability space.

Definition 1.1.18. *A random variable X taking values in $\{0, 1, 2, \dots\}$ is said to be follow a **Poisson distribution** with parameter λ if $P(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}$ for $i = 0, 1, 2, \dots$.*

Definition 1.1.19. *The process $X = \{X_t\}_{t\geq 0}$, taking values in non-negative integers, defined by $X_t(\omega) := \sum_{n\geq 1} \mathbf{1}_{\{t\geq T_n(\omega)\}}$, where $\{T_n\}_{n\geq 1}$ is a strictly increasing sequence of stopping times, is called a **counting process** associated to $\{T_n\}_{n\geq 1}$.*

Definition 1.1.20. *A counting process X is said to be a Poisson process if*

(i) $X_t - X_s$ *is independent of* $\mathcal{F}_s$ *for* $0 \leq s < t$.

(ii) X *has a stationary increment.*

From the Definition 1.1.20, it is clear that if $T_1, T_2, \dots$ are the jump times of X, then X_t counts the total number of jumps between $[0, t]$ i.e.

$$X_t := \#\{i \geq 1, T_i \in [0, t]\}.$$

Now we introduce the random measure. For more details we refer [9].

Definition 1.1.21. *Let* $(E, \mathcal{G})$ *be a measurable space. A map* $M : \Omega \times \mathcal{G} \to \mathbb{R}_+$ *is said to be a* ***random measure*** *if* $\omega \to M(\omega, A)$ *is a random variable for all* $A \in \mathcal{G}$ *and* $A \to M(\omega, A)$ *is a measure on* $(E, \mathcal{G})$ *for all* $\omega \in \Omega$.

Definition 1.1.22. *Let* $(E, \mathcal{G})$ *be a measurable space and* μ *be a* σ*-finite measure on it. A random measure* M *on* $(E, \mathcal{G})$ *is said to be a* ***Poisson random measure*** *if*

1. $M(A)$ *is a Poisson random variable for all* A *with mean* $\mu(A)$,
2. *if for disjoint* $A_1, A_2, \dots, A_n \in \mathcal{G}$, *the random variables* $M(A_1), M(A_2), \dots, M(A_n)$ *are independent.*

Integration

Let $(E, \mathcal{G})$ be a measurable space and μ be a σ-finite measure on it. Let M be a Poisson random measure with mean measure (or intensity measure) μ. We now describe the class of functions $L^2(\mu)$ for which the integral with respect to M to be defined is the following

$$L^2(\mu) := \left\{ f : E \to \mathbb{R} : f \text{ is measurable and} \int_E f^2 \, d\mu < \infty. \right\}$$

Then the space $L^2(\mu)$ is a Banach space with under the norm $\|f\| = \int_E f^2 \, d\mu$. It is easy to see that the space of all simple functions on E is dense on $L^2(\mu)$. Now we state the key lemma for integration.

Lemma 1.1.23. *Let* $f = \sum_{j=1}^{n} c_j \mathbb{1}_{A_j}$ *where* $A_1, \dots, A_n$ *are measurable on* $(E, \mathcal{G})$ *be a simple function. Then*

$$M_f(\omega) = \int_E f(x) M(\omega, dx) = \sum_{j=1}^{n} c_j M(\omega, A_j).$$

1.1.4 Stochastic Differential Equation

Let $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ be two Borel-measurable functions. Now consider the following stochastic differential equation

$$\left.\begin{aligned} dX_t &= b(t,X_t)dt + \sigma(t,X_t)dW_t \\ X_0 &= Z, \end{aligned}\right\} \tag{1.3}$$

where $W = \{W_t\}_{t\geq 0}$ is a m-dimensional Brownian motion and $X = \{X\}_{t\geq 0}$ is the **solution** of the equation (1.3), a real valued stochastic process with rcll sample paths. First we recall the definition of strong solution of the SDE (1.3) following [40] (Definition 5.2.1).

Definition 1.1.24 (Strong Solution to SDE)**.** *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ be a filtered probability space, where $\{\mathcal{F}_t\}$ is the augmentation of the filtration generated by Z and W. A rcll process $X = \{X_t\}_{t\geq 0}$ is said to be a strong solution of* (1.3)*, if*

(i) X_t *is* $\{\mathcal{F}_t\}$*-adapted,*

(ii) $P(X_0 = Z) = 1$*,*

(iii) $\int_0^t \left[|b_i(s,X_s)| + \sigma_{ij}^2(s,X_s)\right] ds < \infty$ *a.s. for all* $1 \leq i \leq n$*,* $1 \leq j \leq n$ *and* $t \in [0,T]$,

(iv) the integral version of (1.3)

$$X_t = X_0 + \int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s, \qquad 0 \leq t \leq T,$$

holds almost surely.

The following theorem asserts the existence and uniqueness of strong solution of the SDE (1.3) under certain conditions. For more details we refer [40] (Theorem 5.2.9).

Theorem 1.1.25 (Existence and Uniqueness of Strong Solution to SDE)**.** *Let $b(t,x), \sigma(t,x)$ satisfies Lipschitz and linear growth conditions*

$$\|b(t,x) - b(t,y)\| + \|\sigma(t,x) - \sigma(t,y)\| \leq K\|x-y\|, \tag{1.4}$$

$$\|b(t,x)\|^2 + \|\sigma(t,x)\|^2 \leq K^2\left(1 + \|x\|^2\right), \tag{1.5}$$

for $t \in [0,T]$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, $\|\cdot\|$ denotes the Euclidean norm and K is a positive constant. Let Z be a random vector, independent of the Brownian motion $W = \{W_t, \mathcal{F}_t^W; 0 \leq t < T\}$ and $\mathbb{E}\|Z\|^2 < \infty$. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be as in Definition 1.1.24. Then there exists a continuous, adapted process $X = \{X_t, \mathcal{F}_t; 0 \leq t < T\}$ which is a strong solution of the SDE (1.3).

Let N be a Poisson random measure with intensity measure $\mathcal{L} \times \mu$, where $\mathcal{L}$ denotes the Lebesgue measure and $\tilde{N}$ be its compensated Poisson random measure. Consider the following Lévy SDE

$$\left.\begin{aligned} dX_t &= b(t, X_t)dt + \sigma(t, X_t)dW_t + \int_0^t \int_{\mathbb{R}^n} K(s, X_{s-}, z)\tilde{N}(dt, dz) \\ X_0 &= Z, \end{aligned}\right\} \tag{1.6}$$

where $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ and $K : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times l}$ are Borel measurable functions. We state the following theorem from [49] (Theorem 1.19) to ensure the conditions for which the Lévy SDE has a unique strong solution.

Theorem 1.1.26 (Existence and Uniqueness of Solution to Lévy SDE)**.** *Let $b(t, x), \sigma(t, x)$ satisfies Lipschitz and linear growth conditions*

$$\|b(t, x) - b(t, y)\|^2 + \|\sigma(t, x) - \sigma(t, y)\|^2$$
$$+ \sum_{j=1}^{l} \int_0^t \|K_j(t, x, z_j) - K_j(t, x, z_j)\|^2 \mu_j(dz_j) \leq K\|x - y\|^2, \tag{1.7}$$

$$\|b(t, x)\|^2 + \|\sigma(t, x)\|^2 + \int_0^t \sum_{j=1}^{l} \|K_j(t, x, z_j)\|^2 \mu_j(dz_j) \leq K^2 \left(1 + \|x\|^2\right), \tag{1.8}$$

for $t \in [0, T]$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$ and K is a positive constant. Let Z be a random vector, independent of the Brownian motion $W = \{W_t, \mathcal{F}_t^W; 0 \leq t < T\}$ and $\mathbb{E}\|Z\|^2 < \infty$. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be the augmentation of the filtration generated by Z, W and N. Then there exists a rcll, adapted process $X = \{X_t, \mathcal{F}_t; 0 \leq t < T\}$ which is a strong solution of the SDE (1.6)*.*

Definition 1.1.27. *A process $X = \{X_t\}_{t\geq 0}$ is said to be **Diffusion** if it satisfies the following SDE*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t,$$

*where $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are two Lipschitz functions in the space variable. The $n \times n$ matrix $a(t, x) := \frac{1}{2}\sigma\sigma^T$ is known as **Diffusion matrix**.*

1.2 Semigroup of Operators

Throughout this section we assume that B is a Banach space. We present some important definitions from Ethier & Kurtz [19].

Definition 1.2.1. *A family of bounded linear operators $\{T_t\}_{t\geq 0}$ on B is said to be a **semigroup of operators** if* *(i)* $T_0 = I$, *I is the identity operator,*

(ii) $T_{t+s} = T_t T_s$.

Definition 1.2.2. *A semigroup of operators $\{T_t\}_{t\geq 0}$ on B is said to be a* ***strongly continuous or C_0 semigroup*** *if $\lim_{t\to 0} T_t f = f$ for all $f \in B$.*

Definition 1.2.3. *An operator $A : \mathcal{D}(A) \subset B \to B$ defined by*

$$Af = \lim_{t\to 0} \frac{T_t f - f}{t} \quad \forall f \in \mathcal{D}(A),$$

is said to be a ***generator*** *of the semigroup $\{T_t\}_{t\geq 0}$. The domain of A, $\mathcal{D}(A)$ contains $f \in B$ such that the above limit exists.*

1.3 Analysis and Control Theory

We shall state a convergence theorem namely Vitali convergence theorem, which can be found in chapter 18 of Royden & Fitzpatrick [54].

Theorem 1.3.1 (Vitali Convergence Theorem)**.** *Let $(E, \mathcal{G}, \nu)$ be a measure space and $\{f_n\}$ be a sequence of functions on E which is both uniformly integrable and tight over E. Let $\{f_n\} \to f$ a.e. on E pointwise and f is integrable over E. Then*

$$\lim_{n\to\infty} \int_E f_n \, d\nu = \lim_{n\to\infty} \int_E f \, d\nu.$$

The multivalued function plays a crucial role in the context of control theory. Our main intention is to state the maximum theorem. To do this we need some introductory definitions, which can be found in the lecture note of Srivastava [59]

Definition 1.3.2. *Let Γ and Δ be two topological spaces and $\mathcal{O}$ be the family of all open subsets of Γ.*

1. *A* ***multifunction*** *$\Phi : \Gamma \to \Delta$ is a map from A to non-empty subsets of B.*

2. *A multifunction $\Phi : \Gamma \to \Delta$ is said to be $\mathcal{O}$-**measurable** if for every open subset X in Δ ,*

$$\{x \in \Gamma : \Phi(x) \cap X \neq \emptyset\} \in \mathcal{O}.$$

3. *A multifunction $\Phi : \Gamma \to \Delta$ is said to be **lower semi-continuous**(resp. **upper semi-continuous**) if for every open(resp. closed) subset X in Δ ,*

$$\{x \in \Gamma : \Phi(x) \cap X \neq \emptyset\},$$

is open(resp. closed) in Γ.

4. *A multifunction $\Phi : \Gamma \to \Delta$ is said to be* ***continuous*** *if it is both lower and upper semi-continuous.*

The maximum theorem is one of most useful selection theorem in control theory. The version we would present here can be found in Rangarajan [60] (Theorem 9.14).

Theorem 1.3.3 (The Maximum Theorem)**.** *Let Γ and Δ be two subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$. Let $f : \Gamma \times \Delta \to \mathbb{R}$ be a continuous function, and Φ be a compact valued, continuous multifunction from Δ to Γ. Let a function $f^* : \Delta \to \mathbb{R}$ and a multifunction $\Phi^* : \Delta \to \Gamma$ be defined by*

$$f^*(z) := \max\{f(x,z)|x \in \Phi(z)\}$$
$$\Phi^*(z) := arg\ max\{f(x,z)|x \in \Phi(z)\}.$$

Then f^* is continuous on Δ, and Φ^* compact-valued and upper semi-continuous on Δ.

1.4 Finance

Our aim in this section is to introduce some basics of portfolio optimization and option pricing. Now we shall recall some ideas of option pricing.

Definition 1.4.1. *An* ***option*** *is a contract between two parties which gives the holder the right but not obligation to trade (buy or sell) an underlying asset at a specified price (strike price) on a specified date.*

Classification of options. Options can be classified according to right, styles, underlying assets etc.

1. **Rights.** There are two types of options in this category call and put options.
   (a) A **call option** is an option which gives holder the right but not obligation to buy a stock from the writer at a fixed strike price.
   (b) A **put option** is an option which gives holder the right but not obligation to sell a stock at a fixed price.
2. **Styles.** We can classify options according to the styles as following
   (a) European type options are the options which can only be exercised at expiry.
   (b) American type options are the options which can be exercised on or before expiry.
   (c) There are further classification e.g. Asian options, barrier option, binary options, etc are in this class. In this thesis, we shall not discuss about these type of options. For more details one can check chapter 26 of [34].

Before we end our discussion about the classification of options, we must include vanilla and exotic options. The **vanilla option** is a European/American type call/put options. Whereas the options other than vanilla options is known as **exotic options**. Therefore Asian options, barrier options are exotic options. There are another important exotic option namely the basket option.

Definition 1.4.2 (Basket option)**.** *A **basket option** is an exotic option whose underlying asset is the weighted average of different asset which are grouped together into a basket.*

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a filtered probability space on which the $n+1$ assets $\{S^0, \ldots, S^n\}$ where $S^i = \{S^i_t\}_{t\geq 0}$, be defined. Let S^0 be a locally risk free asset. Now we shall talk about some useful definitions which can be found in [38], [42], [57].

Definition 1.4.3. *(1) An $(n+1)$-dimensional process $\pi = \{\pi_t = (\xi_t, \varepsilon_t), 0 \leq t \leq T\}$ is said to be a **trading strategy** if the n-dimensional process ξ is predictable and ε is adapted.*

(2) Let π be a trading strategy with initial value V_0. Then for $0 \leq t \leq T$ the process

$$V_t(\pi) := \sum_{i=1}^{n} \xi^i_t S^i_t + \varepsilon_t S^0_t,$$

*is said to be the **value process** with initial wealth V_0.*

(3) The process

$$\hat{V}_t(\pi) := \sum_{i=1}^{n} \xi^i_t \hat{S}^i_t + \varepsilon_t,$$

*where $\hat{S}^i = S_i/S^0$ is said to be the **discounted value process**.*

(4) Let $\hat{S} = \{\hat{S}^1, \ldots, \hat{S}^n\}$ be semimartingale, then the process

$$C_t = V_t - \int_0^t \xi_s \, d\hat{S}_s,$$

*is known as **consumption process**.*

(5) A strategy π is said to be self-financing if the consumption process C is constant over time, i.e.

$$V_t = V_0 + \int_0^t \xi_s \, d\hat{S}_s$$

Definition 1.4.4 (Arbitrage)**.** *Let π be trading strategy with initial value $V_0 = 0$. Then π is said to be **arbitrage opportunity** if*

$$\hat{V}_T(\pi) \geq 0 \quad a.s.\ P, \ \ and\ P\left(\hat{V}_T(\pi) > 0\right) > 0.$$

Definition 1.4.5 (Contingent Claim). *An $\mathcal{F}_T$-measurable random variable Z_T is said to be a **contingent claim** if*

$$Z_T \geq 0 \quad a.s.\ P \text{ and } \mathbb{E}Z_T < \infty.$$

Example 1. *One of the well known example of contingent claim is European call option. Let S_t be the stock price at time t and K be the strike price. Then the contingent claim for European call option is given by $Z_T = (S_T - K)^+$.*

Definition 1.4.6. *(1) A trading strategy π is said to be a **hedging strategy** of the contingent claim Z_T if $\hat{V}_T(\pi) = Z_T \quad a.s.\ P$.*

*(2) A market is said to be a **complete market** if for given any contingent claim there exists a hedging strategy.*

Example 2. *The best known example of complete market is BSM market. In this market there is one bond and one stock satisfying geometric Brownian motion.*

Now we recall one of the most important concept of finance, namely risk-neutral measure or equivalent martingale measure (EMM) .

Definition 1.4.7. *(1) Let P and Q be two measures on a measure space $(\Omega, \mathcal{F})$. The measure P is said to be **absolutely continuous** with respect to Q if $P(A) = 0$ for each set $A \in \mathcal{F}$ for which $Q(A) = 0$ and denoted by $P \ll Q$.*

*(2) Two measures P and Q on a measure space $(\Omega, \mathcal{F})$ are said to be **equivalent measure** if $P \ll Q$ and $Q \ll P$ and is denoted by $P \approx Q$.*

*(3) Let $(\Omega, \mathcal{F}, P)$ be a probability space. A probability measure Q is said to be **equivalent (local) martingale measure** if $Q \approx P$ and the discounted stock price is (local) martingale with respect to Q.*

The following theorems are known to be as first and second fundamental theorem. They can be found in [38] (Theorem 7.1 and Theorem 9.7).

Theorem 1.4.8. *Let $(\Omega, \mathcal{F}, P)$ be a probability space on which $\hat{S} = \{\hat{S}_t\}_{0 \leq t \leq T}$ be an $\mathbb{R}^n$ valued discounted stock prices. If there exists an EMM Q, then the market is arbitrage free. Moreover, if there exists a unique EMM Q, then the market is complete.*

2

Testing of binary regime switching models using squeeze durations

2.1 Introduction

This chapter concerns about the testing of appropriateness of the use of binary regime switching models. However in order to avoid computational complexity, we restrict ourselves in a particular parametric class of geometric Brownian motion (GBM) and its subsequent generalizations in terms of binary regime switching. By a binary regime switching, we mean that there exists an unobserved two-state stochastic process whose movement allows to change the market parameters i.e. drift and volatility coefficients. Let S_t denote the asset price at time t, under the binary regime switching GBM. Then

$$dS_t = S_t \left(\mu(X_t)\, dt + \sigma(X_t)\, dB_t \right), \tag{2.1}$$

where $\{X_t\}_{t \geq 0}$ is a $\{1, 2\}$-valued stochastic process and $\mu(X_t), \sigma(X_t)$ are the drift and the volatility coefficients and B_t is the standard Brownian motion. In general $\{X_t\}_{t \geq 0}$ is chosen to be Markov or semi-Markov process.

We construct the discriminating statistics using some descriptive statistics of squeeze duration of Bollinger band, which seems to be the most natural approach. The sampling distribution of the descriptive statistics of occupation measure of Bollinger band under a particular model hypothesis need not have a nice form and thus one may not be able to identify that as one of a few known distributions.

In spite of the lack of mathematical tractability, one may surely obtain empirical distribution using a reliable simulation procedure. This is a standard approach in such circumstances. We refer to [62] for further details. In Theiler et al [61] this is termed as the typical realization surrogate data approach. The surrogate data approach is generally perceived one of the most powerful method for testing of hypothesis. The available algorithms to generate surrogate can be categorized into two classes namely typical realizations and constrained realization (see [56]). In this thesis, we shall be using the typical realization approach.

The rest of this chapter consists of two section. In Section 2, we propose a test statistic using Bollinger band squeeze. Section 3 concerns about the empirical results.

2.2 Discriminating statistics based on Squeeze Duration

In this section, a discriminating statistics is proposed whose sampling distribution varies drastically, under the regime switching assumption, with varying values of instantaneous rate parameter. The discriminating statistics is taken as vector valued where every component is a descriptive statistics of squeeze duration of Bollinger band. This section is dedicated in describing the statistics and the numerical methods for obtaining its sampling distributions. The actual numerical experiments are deferred to the next section. This section is organized in four subsections.

2.2.1 Bollinger Band

Keltner channel and Bollinger bands based on the empirical volatility are the most popular indicators for trading. John Bollinger introduced the concept of Bollinger band for pattern recognition in 1980s. Bollinger bands provide a time varying interval for any financial time series data. The end points of the intervals are computed based on the moving average and the moving sample standard deviation of the past data of fixed window size. Now we present a formal definition of the Bollinger bands of an asset.

Definition 2.2.1. *A Bollinger band of a given time series data consists of three lines on the time series plot, computed based on immediate lag values of fixed length n say. The middle line is the moving average of the time series with window size n. The upper and the lower lines are exactly $k\sigma$ unit away from the middle line where k is a fixed constant and σ is the sample standard deviation obtained from the last n numbers of lag values.*

It is important to note that the main focus of the Bollinger band is to capture the fluctuation, to be more precise, the volatility coefficient of the time series. Hence the closeness of the upper and lower line is termed as squeeze, and is the indication of low volatility of a particular time series. On the contrary, when the boundaries of band are far from each other, that corresponds to a high volatility. For more details about Bollinger bands and squeeze we refer to [5].

2.2.2 p-Squeeze Durations

In this chapter, we consider the Bollinger bands of the simple return of a financial time series. We introduce some important notations and definitions which would be used subsequently. Let $S = \{S_k\}_{k=1}^N$ denote an equispaced financial time series. The simple return of S is defined by

$$r_k := \frac{S_k - S_{k-1}}{S_{k-1}}, \quad k = 2, 3, \dots, N. \tag{2.2}$$

Definition 2.2.2 ($\hat{\mu}, \hat{\sigma}$)**.** *By fixing the window size as n, the moving average $\{m_k\}_{k=n+1}^N$ and the sample standard deviation $\{\sigma_k\}_{k=n+1}^N$ are given by*

$$m_k := \frac{1}{n}\sum_{i=0}^{n-1} r_{k-i}, \tag{2.3}$$

$$\sigma_k := \sqrt{\frac{1}{n-1}\sum_{i=0}^{n-1}(r_{k-i})^2 - \frac{n}{n-1}m_k^2}, \tag{2.4}$$

for $k \geq n+1$. The empirical volatility $\hat{\sigma} = \{\hat{\sigma}_k\}_{k=n+1}^N$ is given by $\hat{\sigma}_k := \dfrac{\sigma_k}{\sqrt{\Delta}}$ for all $k \geq n+1$, where Δ is the length of the time step in year unit. Similarly, the empirical drift $\hat{\mu} = \{\hat{\mu}_k\}_{k=n+1}^N$ is given by $\hat{\mu}_k := \dfrac{m_k}{\Delta}$ for all $k \geq n+1$.

Definition 2.2.3. *Let $y = \{y_k\}_{k=1}^m$ be a random sample of a real valued random variable. Then the empirical cumulative distribution function or ecdf $\hat{F}_y$ is defined as*

$$\hat{F}_y(x) := \frac{1}{m}\sum_{k=1}^{m} \mathbb{1}_{[0,\infty)}(x - y_k),$$

where given a subset A, $\mathbb{1}_A$ denotes the indicator function of A.

Definition 2.2.4 (p-percentile)**.** *Let $\hat{F}_y$ be the ecdf of $y = \{y_k\}_{k=1}^m$. Then for any $p \in (0,1)$, the p-percentile of y, denoted by $\hat{F}_y^{\leftarrow}(p)$, is defined as*

$$\hat{F}_y^{\leftarrow}(p) := \inf\left\{x \middle| \hat{F}_y(x) \geq p\right\}.$$

For mathematical tractability we use a particular percentile of $\hat{\sigma}$ as threshold in defining the squeeze of the Bollinger band.

Definition 2.2.5 (p-squeeze)**.** *Given a $p \in (0,1)$, an asset is said to be in p-squeeze at k-th time step if the empirical volatility $\hat{\sigma}_k$, as defined above, is not more than $\hat{F}_{\hat{\sigma}}^{\leftarrow}(p)$.*

We introduce the sojourn times of the p-squeeze below.

Definition 2.2.6. *For a fixed $p \in (0,1)$ and a given time series $\{S_k\}_{k=1}^N$, let $\{(a_i, b_i)\}_{i=1}^\infty$ be an extended real valued double sequence given by*

$$\begin{cases} a_0 = 0 \\ b_{i-1} := \min\{k \geq a_{i-1} | \hat{\sigma}_k > \hat{F}_{\hat{\sigma}}^{\leftarrow}(p)\} \\ a_i := \min\{k \geq b_{i-1} | \hat{\sigma}_k \leq \hat{F}_{\hat{\sigma}}^{\leftarrow}(p)\}, \end{cases}$$

for $i = 1, 2, \ldots$ and by following the convention of $\min \emptyset = +\infty$, where $\hat{\sigma}$ is as in Definition 2.2.2. Then the sojourn time durations for the p-squeezes are $\{d_i\}_{i=1}^L$, where $d_i := b_i - a_i$ and $L := \max\{i | b_i < \infty\}$, provided $L \geq 1$.

We note that one must multiply each d_i by Δ to obtain the squeeze durations in year unit. We would consider the finite sequence $\{d_i\}_{i=1}^L$ as a single object. In particular, we call d_i as the i-th entry of the p-squeeze duration or p-SqD in short for the given time series $\{S_k\}_{k=1}^N$. We call L to be the length of p-SqD.

Remark 2.2.7. *In a reasonably large and practically relevant time series data, the length of p-SqD is considerably small. Hence a non parametric estimation of the entries of p-SqD using empirical cdf is not practicable as that would have a high standard error. Hence, only a collection of some descriptive statistics such as mean($\bar{d}$), standard deviation(s), skewness(ν), kurtosis(κ) of p-SqD can reliably be obtained and compared.*

2.2.3 A Discriminating Statistics

We first consider a discrete time version of continuous time theoretical asset price model with the time step identical to that of the time series data. We note that, for that theoretical model, the corresponding p-SqD is a random sequence with random length. However, the corresponding descriptive statistics as above would constitute a random vector of fixed length whose sampling distribution would be sought for. A comparison of $(\bar{d}, s, \nu, \kappa)$ with respect to that sampling distribution would be the central idea for statistical inference. However, this should not lead to the only criterion for rejecting a model. Of course there are many other natural criteria for the same. Those criteria are typically considered as constraints on parameterization of the class of models of our interest. Here we illustrate with an example. If we restrict ourselves to all possible MMGBM models with two regime as in [7], then

$$\Theta = \{(\mu(1), \sigma(1), \lambda_1, \mu(2), \sigma(2), \lambda_2) | \mu(i) \in \mathbb{R}, \sigma(i) > 0, \lambda(i) > 0, i = 1, 2\} \tag{2.5}$$

is the class of parameters, where $\mu(i)$ and $\sigma(i)$ are the drift and the volatility coefficients respectively and $\begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}$ is the rate matrix for the Markov chain. Therefore the estimation problem boils down to a constrained minimization on a set $\mathscr{C} \subset \Theta$ of the following functional $f : \Theta \to \mathbb{R}$ given by

$$f(\theta) = E^{\theta}\left[(\bar{d}^{\theta} - \bar{d})^2 + (s^{\theta} - s)^2 + (\nu^{\theta} - \nu)^2 + (\kappa^{\theta} - \kappa)^2\right], \tag{2.6}$$

where $(\bar{d}^{\theta}, s^{\theta}, \nu^{\theta}, \kappa^{\theta})$ is the descriptive statistics vector of a member with parameter $\theta \in \Theta$, provided the minimizer exists.

The main difficulty in taking $\mathscr{C} = \Theta$ is its time complexity due to a large scope of parameter values. We introduce a fixed set of constraints, p-admissible class($\mathscr{C}_p$-class), which is a subclass of all possible regime switching models.

Definition 2.2.8 ($\mathscr{C}_p$-class)**.** *Given a time series data and a fixed $p \in (0,1)$, a regime switching model is said to be in $\mathscr{C}_p$-class of models if the model satisfies the following properties.*

i. The long run average of drift coefficient matches with the time average of empirical drift $\hat{\mu}$.

ii. The long run average of volatility parameters matches with the time average of empirical volatility $\hat{\sigma}$.

iii. The long run proportion of time that the volatility process stays below $\hat{F}_{\hat{\sigma}}^{\leftarrow}(p)$ is p, provided the volatility process is not constant.

In view of Remark 2.2.7, we construct a discriminating statistics $\mathbf{T} = (T_1, T_2, \ldots, T_r)$ using r number of descriptive statistics of the p-SqD. To be more specific we choose $T_1 := \frac{1}{L}\sum_{i=1}^{L} d_i$, $T_2 := \sqrt{\frac{1}{L-1}\sum_{i=1}^{L}(d_i - T_1)^2}$, $T_3 := \dfrac{\frac{1}{L}\sum_{i=1}^{L}(d_i - T_1)^3}{T_2^3}$, $T_4 := \dfrac{\frac{1}{L}\sum_{i=1}^{L}(d_i - T_1)^4}{T_2^4}$ etc. Although our test statistics is based on squeeze durations which are amenable to capture the sojourn times of regime transitions but it is not at all obvious that it would indeed be successful to capture those unobserved switchings. The main difficulty lies in the fact that a larger moving window size (n) in defining $\hat{\sigma}$ ignores more number of intermittent transitions and a smaller window size corresponds to higher standard error. So far window size is concerned there is a popular choice of window size by practitioners, i.e., $n = 20$ for computing the empirical volatility. In view of these, we fix $n = 20$ now onward in the definition of $\mathbf{T}$.

Next we describe the procedure, adopted in this chapter, of obtaining the sampling distribution of $\mathbf{T}$ under binary regime switching model hypothesis.

2.2.4 Sampling distribution of the statistics

In this section we give a detailed description of numerical computation of sampling distribution of $\mathbf{T}$ statistics under the null hypothesis using Monte Carlo method, which is popularly known as typical surrogate approach following [61]. It is important to note that the hypothesis testing, relevant to us, is of composite type (see [62]). The main purpose is to test a meaningful composite null hypothesis. The procedure is as follows

(a) Given a time series S, the p-admissible class $\mathscr{C}_p$ under the null hypothesis is identified. A non-empty subclass $\mathscr{A}$ of $\mathscr{C}_p$ is fixed.

(b) For each $\theta \in \mathscr{A}$, B number of time series $\{X^1, X^2, \dots, X^B\}$ are simulated from the corresponding model θ with the same time step as in S. We call these, the surrogate data of S corresponding to θ.

(c) Let $\mathbf{t}^* := \mathbf{T}(S)$ be the value of $\mathbf{T}$ of the observed data S and $\mathbf{t} := \{\mathbf{t}^1, \mathbf{t}^2, \dots, \mathbf{t}^B\}$ be the values of $\mathbf{T}$ for the surrogate data $\{X^1, X^2, \dots, X^B\}$, where $\mathbf{t}^i = (t_1^i, t_2^i, \dots, t_r^i) = \mathbf{T}(X^i)$.

(d) By keeping a two sided test in mind we define α_r^θ in the following manner

$$\alpha_r^\theta := 2 \min_{j \le r} g_B \left(\sum_{i=1}^{B} \mathbb{1}_{[0,\infty)}(t_j^* - t_j^i) \right),$$

where $g_B(x) := \frac{x \wedge (B-x)}{B}$, and $\mathbf{t}^* = (t_1^*, t_2^*, \dots, t_r^*)$.

(e) Therefore the α-value for the test of the class $\mathscr{A}$ is given by

$$\alpha_r = \max_{\theta \in \mathscr{A}} \alpha_r^\theta.$$

(f) We reject the hypothesis that S is a sample from a model in the class with confidence $100(1-\alpha_r)\%$, provided α_r is reasonably small.

Remark 2.2.9. *It is important to note, the above method has a pathetic limitation due to the "curse of dimensionality." Or in other words for a given model θ, the probability of observing the value of α_r^θ to be smaller than a very small value is not so small when r is large. But it is well known that the curse is not so fatal for the dimension r less than five. Therefore we restrict ourselves in four dimensional testing.*

For the purpose of illustration here we consider a specific time series S and $r = 2$ i.e., the test statistics $\mathbf{T} = (T_1, T_2)$. In the following plot Figure 2.1, the values T_1 and T_2 are plotted against the horizontal and vertical axes respectively. The position of the point $t^* = \mathbf{T}(S)$ is denoted as a circle in the plot. Under the null hypothesis of GBM, the $\mathscr{C}_p$-class turns out to be singleton (see Section 2.3.1 for details). The sampling distribution of $\mathbf{T}$ under the null hypothesis is computed by setting $B = 200$ and that is presented using a two dimensional box plot. Furthermore, under the null hypothesis of MMGBM, followed by fixing $B = 1$, and a subclass $\mathscr{A} := \{\frac{1}{\lambda_1} \in (\frac{1}{20}\mathbb{N}) \cap [5, 15], \sigma(1) = \hat{F}_{\hat{\sigma}}^{\leftarrow}(p)\}$ of $\mathscr{C}_p$, the values of $\mathbf{T}$ are plotted as dots.

In the following section, we implement the ideas developed here for more specific choice of models and discuss the implementation issues in details.

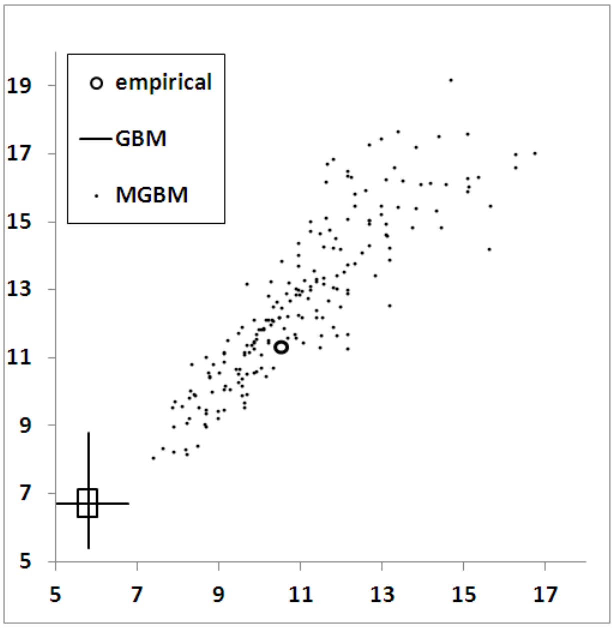

Figure 2.1: Plot of (T_1, T_2) for observation and various different surrogate data.

2.3 Empirical study

For empirical study, we consider, 5-minute data of several Indian stock indices from 1-st December, 2016 to 30-th June, 2017. Assuming there are 250 trading days in a year and 6 hours of trading in each day, we set $\Delta = \frac{5}{250\times360}$. We fix $p = 15\%$ throughout this section. The components of t^* for each index data are given in the Table 2.1 below. Every row of the table corresponds to an index, which is mentioned in the second column with their id's in the first column. The third column gives the value of L, the number of observations of p-squeeze duration of each index data.

2.3.1 Surrogate Data under GBM hypothesis

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which $\{W_t\}_{t\geq0}$ be a Brownian motion. The stock price, modeled by a geometric Brownian motion (or GBM in short) is given by

$$dS_t = S_t\,(\mu\,dt + \sigma\,dW_t)\;\; t \geq 0,\;\; S_0 > 0, \tag{2.7}$$

where μ and σ are the drift and the volatility coefficients respectively. Equation (2.7) has a strong solution of the form

$$S_t = S_0 \exp\left(\mu t - \frac{1}{2}\sigma^2 t + \sigma W_t\right),\;\; t \geq 0. \tag{2.8}$$

Table 2.1: Values of discriminating statistics ($p = 15\%$) of a 5 -minutes data of Indian indices during 1st Dec 2016 to 30th June 2017

Data	Indices	Occurence	t_1^*	t_2^*	t_3^*	t_4^*
I01	NIFTY 100	159	10.52	11.31	1.17	3.41
I02	NIFTY 200	160	10.45	11.18	1.29	3.79
I03	NIFTY 50	155	10.78	11.00	1.08	3.28
I04	NIFTY 500	152	11.01	11.40	1.20	3.63
I05	NIFTY BANK	159	10.52	11.62	1.39	4.03
I06	NIFTY COMMODITY	169	9.89	10.49	1.47	4.59
I07	NIFTY ENERGY	168	9.96	11.44	1.59	4.80
I08	NIFTY FIN. SER.	168	9.95	10.72	1.46	4.39
I09	NIFTY FMCG	178	9.40	10.15	1.58	5.01
I10	NIFTY INFRA	174	9.61	11.70	1.72	5.41
I11	NIFTY IT	159	10.52	11.36	1.19	3.35
I12	NIFTY MEDIA	173	9.66	9.49	1.19	3.77
I13	NIFTY METAL	188	8.89	10.53	1.92	6.52
I14	NIFTY MNC	178	9.40	10.63	1.54	4.67
I15	NIFTY PHARMA	175	9.56	11.13	1.59	4.69
I16	NIFTY PSE	148	11.29	12.73	1.28	3.76
I17	NIFTY REALTY	177	9.45	10.49	1.83	6.02
I18	NIFTY SERVICE SEC.	172	9.72	11.13	1.33	3.77

It is important to note that, the $\mathscr{C}_p$-class is singleton as μ and σ are, by using Definition 2.2.8 (i)-(ii), $\mu = \bar{\hat{\mu}}$ and $\sigma = \bar{\hat{\sigma}}$, where the bar sign represents the time average.

Let S_0 be the initial price of a stock consisting of N number of data points. Let $\{0 = t_1 < t_2 < \ldots < t_N\}$ be a partition of time interval of the observed data series, where $t_{i+1} - t_i = \Delta$ for $i = 1, 2, \ldots, N-1$ and Δ be the length of time step in year unit. Then the $\mathscr{C}_p$-class of surrogate GBM can be generated by using the discretized version of (2.8) which is given by

$$S_{t_{i+1}} = S_{t_i} \exp\left((\bar{\hat{\mu}} - \frac{1}{2}\bar{\hat{\sigma}}^2)\Delta + \bar{\hat{\sigma}}\, Z_i\right), \tag{2.9}$$

where $\{Z_i \mid i = 1, \ldots, N-1\}$ are independent and identically distributed (i.i.d.) normal random variables with mean 0 and variance. We use this notation throughout this chapter.

Testing of hypothesis

We intend to test whether the value of $\mathbf{T}$ of observed index prices are outliers of $\mathbf{T}$ values coming from GBM models. For each index in Table 2.1, we set our null hypothesis,

$$H_0 : \text{the time series is in } \mathscr{C}_p\text{-class of GBM.}$$

We again recall that the $\mathscr{C}_p$-class is indeed singleton. The following figures illustrate results from all 18 indices. Figure 2.2 plots T_1 and Figure 2.3 plots T_2 only. Each box plot is obtained by simulating the GBM model from $\mathscr{C}_p$-class 200 times. The triangle plots are the representative for original data of all the indices. Here we see that the triangles appear far from the box plots.

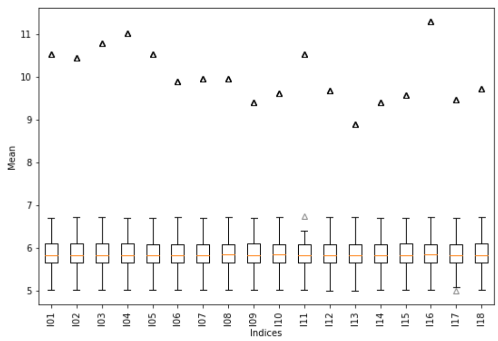

Figure 2.2: Sampling distribution of T_1 under GBM hypothesis

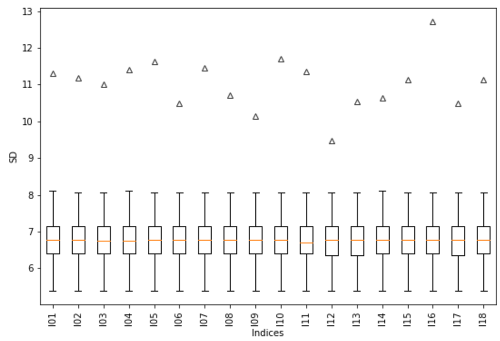

Figure 2.3: Sampling distribution of T_2 under GBM hypothesis

Therefore the above Figures 2.2 and 2.3 indicate a strong rejection for the null hypothesis of GBM model. We continue our investigation with binary regime switching Markov modulated geometric Brownian motion in the following subsection.

2.3.2 Surrogate Data under Markov modulated GBM hypothesis

In this subsection, we study the testing of hypothesis of $\mathscr{C}_p$-class of MMGBM assumption with binary regimes. In this case the equation (2.1) is dependent on a two state Markov process $\{X_t\}_{t\geq 0}$. It is important to note that a continuous time Markov chain can be characterized by its instantaneous transition rate matrix.

In this subsection, we restrict our investigation in a particular subclass $\mathscr{A}$ of $\mathscr{C}_p$ which are embedded in

$$\Theta = \{\theta = (\mu(1), \sigma(1), \lambda_1, \mu(2), \sigma(2), \lambda_2) | \mu(i) \in \mathbb{R}, \sigma(i) > 0, \lambda_i > 0, i = 1, 2\}. \tag{2.10}$$

Here the transition rate matrix for the Markov chain is given by

$$\Lambda := \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}.$$

Therefore the sojourn time distribution of state i is $\text{Exp}(\lambda_i)$ for $i = 1, 2$. Now since $\mathscr{A} \subset \mathscr{C}_p \subset \Theta$, by using Definition 2.2.8(iii), we have

$$\frac{\frac{1}{\lambda_1}}{\frac{1}{\lambda_1} + \frac{1}{\lambda_2}} = p. \tag{2.11}$$

Using Definition 2.2.8(i), (2.11) the drift coefficients $\mu(i)$ satisfy the following relation

$$p\,\mu(1) + (1-p)\,\mu(2) = \bar{\hat{\mu}}. \tag{2.12}$$

Also using Definition 2.2.8(ii) and (2.11) the volatility coefficients $\sigma(i)$ have the relation below

$$p\,\sigma(1) + (1-p)\,\sigma(2) = \bar{\hat{\sigma}}. \tag{2.13}$$

After a simple computation (2.11) becomes

$$\lambda_1 = \left(\frac{1}{p} - 1\right)\lambda_2. \tag{2.14}$$

Thus $\mathscr{C}_p \subset \Theta$ is the set of six parameters satisfying equations (2.12), (2.13) and (2.14). We choose $\mathscr{A}$ by fixing $\mu(1) = \mu(2)$ and $\sigma(1) = \hat{F}_{\hat{\sigma}}^{\leftarrow}(p)$. Thus $\mathscr{A}$ is a subset of the solution space of five equations in six unknowns, or in other words, $\mathscr{A}$ can be viewed as a one-parameter family of models. Next to generate surrogate data, we need to discretize the MMGBM model corresponding to each member of $\mathscr{A}$. To this end we discretize (2.1) and perform Monte Carlo simulation. The discretization of MMGBM surrogate of the form (2.1) is given by

$$\begin{aligned} S_{t_{i+1}} &= S_{t_i} \exp\left(\left(\mu(X_i) - \frac{1}{2}\sigma^2(X_i)\right)\Delta + \sigma(X_i)\,Z_i\right), \\ X_{i+1} &= X_i - (-1)^{X_i}\,P_i, \end{aligned} \tag{2.15}$$

where $\{P_i \mid i = 1, \ldots, N-1\}$ are independent to each other and also to Z_j for all j and each $P_i \sim Bernoulli(\lambda_{X_i}\Delta)$, a Bernoulli random variable with parameter $\lambda_{X_i}\Delta$.

Testing of hypothesis

We intend to test whether the values of **T** of observed index prices are outliers of **T** values coming from MMGBM models. For each index in Table 2.1, we set our null hypothesis,

$$H_0 : \text{the time series is in the class } \mathscr{A} \text{ of MMGBM.}$$

To test H_0, we adopt the typical realization surrogate data approach and consider the discrete model (2.15). For every $\theta \in \mathscr{A}$, we perform Monte Carlo simulation two hundred times ($B = 200$). Then we record the α_1, α_2, α_3 and α_4 values for each index in Table 2.2. From Table 2.2 it is evident that H_0 can be rejected for each index with α value 5% or smaller. We consider the first time series I01 to illustrate the sampling distribution of **T** for some models in $\mathscr{A}$ using the following four plots. In each of the Figures 2.4-2.7, the circle plot represents t^*, the **T** value of I01. Furthermore, there are three two-dimensional box plots corresponding to $\frac{1}{\lambda_1} = 5, 10$ and 15 respectively. It seems from Figures 2.4 that, the least square estimate of $\frac{1}{\lambda_1}$ should be in between 5 and 10, so far (T_1, T_2) is concerned. However, Figure 2.5 implies that the least square estimate of $\frac{1}{\lambda_1} > 15$. On the other hand, Figures 2.6 and 2.7 imply that the data is an outlier. Finally, since all the α_4 are less than 5%, in the MMGBM column of Table 2.2, we reject the null hypothesis with 95% confidence for each index data.

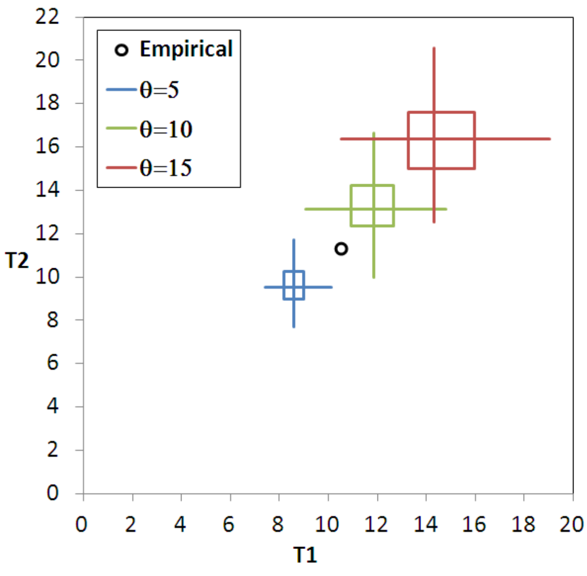

Figure 2.4: T_1 and T_2 under MMGBM hypothesis

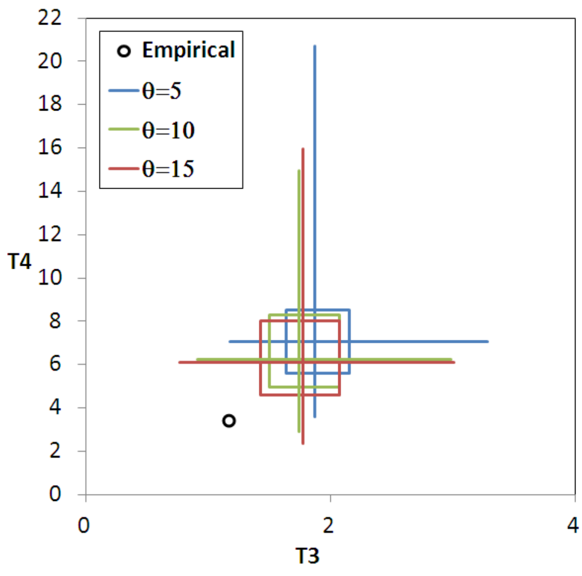

Figure 2.5: T_3 and T_4 under MMGBM hypothesis

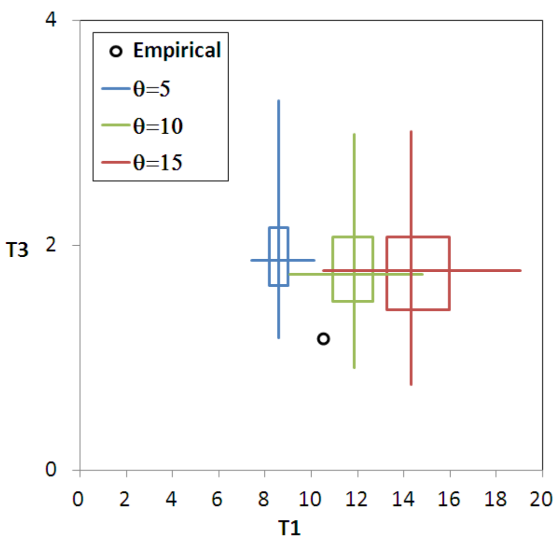

Figure 2.6: T_1 and T_3 under MMGBM hypothesis

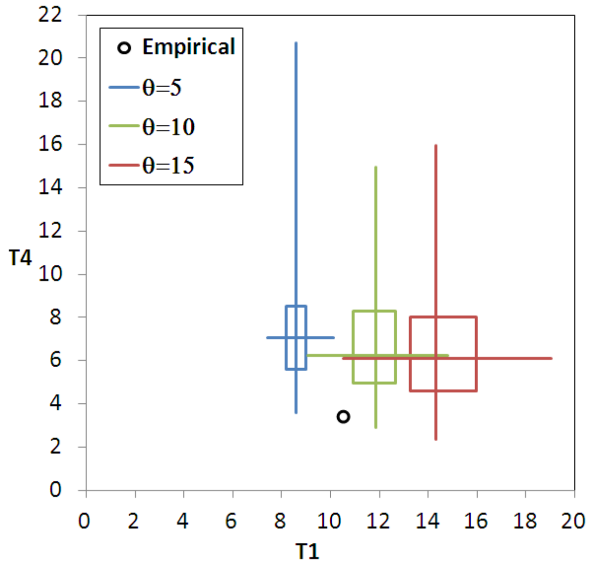

Figure 2.7: T_1 and T_4 under MMGBM hypothesis

2.3.3 Surrogate Data under semi-Markov modulated GBM

In this subsection, we consider a subclass of the following all class of models

$$\Theta = \{\theta = (\mu(1), \sigma(1), \lambda_1(\cdot), \mu(2), \sigma(2), \lambda_2(\cdot) | \mu(i) \in \mathbb{R}, \sigma(i) > 0,\ \lambda_i(\cdot) > 0,\ i = 1, 2\}.$$

For every θ in $\mathscr{C}_p \subset \Theta$, $\mu(i)$ and $\sigma(i)$ satisfy the equation (2.12) and (2.13). The transition rate matrix for the semi-Markov chain is a matrix valued function on $[0, \infty)$, given by

$$\Lambda(y) := \begin{pmatrix} -\lambda_1(y) & \lambda_1(y) \\ \lambda_2(y) & -\lambda_2(y) \end{pmatrix} \quad \forall\ y \in [0, \infty).$$

Now for illustration purpose, $\mathscr{A}$ is chosen in the following manner. The holding time distribution of the state i is $\Gamma(k_i, \lambda_i)$ for $i = 1, 2$, where $\Gamma(k_i, \lambda_i)$ denote the gamma distribution with shape k_i and rate λ_i. Then it follows from [27] that $\lambda_i(y)$ is the hazard rate of $\Gamma(k_i, \lambda_i)$ and is given by $\lambda_i(y) = \frac{\lambda_i^{k_i} y^{k_i - 1} e^{-\lambda_i y}}{\Gamma(k_i) - \gamma(k_i, \lambda_i y)}$, where γ is the lower incomplete gamma function. Since the expectation of $\Gamma(k_i, \lambda_i)$ is $\frac{k_i}{\lambda_i}$, it follows from Definition 2.2.8(iii), that

$$\frac{\frac{k_1}{\lambda_1}}{\frac{k_1}{\lambda_1} + \frac{k_2}{\lambda_2}} = p,$$

i.e.

$$\frac{k_2}{\lambda_2} = \left(\frac{1}{p} - 1\right) \frac{k_1}{\lambda_1}.$$

In addition to these, as before, we further assume that $\mu(1) = \mu(2)$, $\sigma(1) = \hat{F}_{\hat{\sigma}}^{\leftarrow}(p)$ and $k_1 = k_2$. Thus $\mathscr{A}$ is the solution space of six equations in eight unknowns or in other words $\mathscr{A}$ is a two parameter subfamily of Θ. For drawing samples from each member of $\mathscr{A}$ using Monte Carlo simulation, we first discretize (2.1). The discretization scheme for SMGBM surrogate of (2.1) is given by

$$
\begin{aligned}
S_{t_{i+1}} &= S_{t_i} \exp\left(\left(\mu(X_i) - \frac{1}{2}\sigma^2(X_i)\right)\Delta + \sigma(X_i)\, Z_i\right), \\
X_{i+1} &= X_i + (-1)^{X_i}\, P_i, \\
Y_{i+1} &= (Y_i + i\Delta)\,(1 - P_i)\,,
\end{aligned} \tag{2.16}
$$

where P_i and Z_i are as in (2.15). The readers are referred to [27] for more details about similar representation of semi-Markov process.

Testing of hypothesis

We set our null hypothesis for all index

$$H_0 : \text{the time series is in the class } \mathscr{A} \text{ of SMGBM.}$$

From Table 2.2 below, H_0 cannot be rejected for any index with a significant level of confidence. Hence we cannot reject the superset also. Or in other words, we cannot reject the hypothesis that the data, under study, is drawn from a SMGBM population.

2.4 Conclusion

In this chapter, we have developed a statistical technique to test the validity of the use of binary regime switching extension of GBM. By several numerical experiments we have shown that we cannot reject the null hypothesis that the Indian index data is drawn from a SMGBM population.

Table 2.2: The α-vales for all the indices

		MMGBM				SMGBM			
	Index	α_1	α_2	α_3	α_4	α_1	α_2	α_3	α_4
1	I01	0.490	0.395	0.045	0.040	0.495	0.485	0.225	0.095
2	I02	0.490	0.400	0.050	0.045	0.500	0.455	0.195	0.105
3	I03	0.470	0.420	0.050	0.035	0.495	0.470	0.230	0.100
4	I04	0.435	0.395	0.055	0.040	0.500	0.455	0.180	0.105
5	I05	0.415	0.395	0.055	0.040	0.495	0.415	0.195	0.105
6	I06	0.465	0.390	0.055	0.040	0.485	0.415	0.200	0.115
7	I07	0.430	0.430	0.060	0.040	0.500	0.420	0.205	0.100
8	I08	0.475	0.420	0.050	0.035	0.495	0.455	0.235	0.100
9	I09	0.455	0.420	0.085	0.050	0.495	0.425	0.220	0.105
10	I10	0.460	0.390	0.055	0.040	0.485	0.425	0.200	0.115
11	I11	0.455	0.420	0.055	0.040	0.490	0.430	0.190	0.100
12	I12	0.480	0.395	0.050	0.050	0.500	0.470	0.210	0.115
13	I13	0.490	0.405	0.050	0.040	0.500	0.470	0.215	0.105
14	I14	0.430	0.395	0.055	0.040	0.485	0.415	0.195	0.100
15	I15	0.490	0.410	0.050	0.035	0.480	0.480	0.235	0.100
16	I16	0.435	0.395	0.055	0.040	0.495	0.425	0.200	0.105
17	I17	0.470	0.410	0.045	0.030	0.495	0.400	0.195	0.095
18	I18	0.425	0.395	0.055	0.040	0.490	0.430	0.205	0.100

3

The CSM Process

3.1 Introduction

The most common properties of any financial time series are its dramatic change of behavior over the time. This changes in financial markets often occur due to government policies, economic news, etc. The regime switching models allow us to capture these abrupt movements. In this type of models, the market parameters are allowed to be finite state pure jump process. The empirical study based on Indian stock market, presented in the Chapter 2, suggests the appropriateness of use of the semi-Markov process to model the market states. This process is referred to as a hidden semi-Markov process. Age-dependent semi-Markov process is a generalization of the semi-Markov process. In this chapter, we recall the age-dependent process in brief and present some properties of component-wise semi-Markov process (CSM).

3.2 Age-dependent process

The estimation of hidden semi-Markov process from a given time series is generally known to be as a very difficult tusk. However, it is shown in [28] that the instantaneous transition rate, i.e the rate at which the semi-Markov process moves between states, can be estimated. In this section we describe the approach of Ghosh and Saha [27] to construct an age-dependent process on a finite state space $\mathcal{X} := \{1, 2, \ldots, k\}$, specified by instantaneous transition rate λ, which is a collection of measurable functions $\lambda_{ij} : [0, \infty) \to (0, \infty)$, where $(i, j) \in \mathcal{X}_2$ and $\mathcal{X}_2 := \{(i, j) | i \neq j \in \mathcal{X}\}$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. We make following assumption on λ

$$\sup_{y \in (0,\infty)} \sum_{j \neq i} \lambda_{ij}(y) < \infty, \tag{3.1}$$

and we consider $\lambda_{ii}(y) = -\sum_{j\neq i} \lambda_{ij}(y)$.

Now let for $i \neq j$ and for every $y > 0$, $\Lambda_{ij}(y) := \left(\sum_{(i',j')\prec(i,j)} \lambda_{i'j'}(y) \right) + [0, \lambda_{ij}(y))$, using a strict total order $\prec$ on $\mathcal{X}_2$. Now we define $h_\lambda, g_\lambda : \mathcal{X} \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ by

$$h_\lambda(i, y, z) := \sum_{j\in\mathcal{X}\setminus\{i\}} (j - i)1_{\Lambda_{ij}(y)}(z), \quad g_\lambda(i, y, z) := y \sum_{j\in\mathcal{X}\setminus\{i\}} 1_{\Lambda_{ij}(y)}(z).$$

We define the following system of coupled stochastic integral equation

$$\left.\begin{aligned} X_t &= X_0 + \int_{(0,t]} \int_{\mathbb{R}} h_\lambda(X_{u-}, Y_{u-}, z)\wp(du, dz) \\ Y_t &= Y_0 + t - \int_{(0,t]} \int_{\mathbb{R}} g_\lambda(X_{u-}, Y_{u-}, z)\wp(du, dz), \end{aligned}\right\} \tag{3.2}$$

where $\wp$ is a Poisson random measure on $(0, t] \times \mathbb{R}$ with intensity $dt\, dz$. Here $\wp$ is a Poisson random measure in time and space variable. We refer to [50] for the proof that, (3.2) has a unique strong solution under the assumption (3.1). The process $\{X_t\}_{t\geq 0}$ is known to be as an age-dependent semi-Markov process and $\{Y_t\}_{t\geq 0}$ is the age process. We also note that, the joint process $Z_t := \{(X_t, Y_t)\}_{t\geq 0}$ is a time homogeneous Markov process.

The CSM process is further generalization of age dependent semi-Markov process, where each component is an age-dependent semi-Markov process. If we consider a market with multiple asset, then it is better to use a CSM process rather than a single semi-Markov process.

3.3 The CSM Process

In this section we present some properties of CSM process. First we define the CSM process.

Definition 3.3.1. *A pure jump process X on a finite state space $\mathcal{S}$ is called a CSM if there is a bijection $\Gamma : \mathcal{S} \to \mathcal{X}^{n+1}$ for some non-empty finite set $\mathcal{X}$, and some non-negative integer n such that each component of $\Gamma(X)$ is semi-Markov process, independent to each other.*

To model the regimes of the market, we consider a CSM $\{X_t\}_{t\geq 0}$ on $\mathcal{X}^{n+1}$, where X^l, the l-th component of X, is an age-dependent process with instantaneous rate functions λ^l, for every $l = 0, \ldots, n$. We denote the age process of X^l as Y^l and Y defined as $(Y^0, \ldots, Y^n)$ is the age process of X. For each $l = 0, 1, \ldots, n$, let $\{X^l_t\}_{t\geq 0}$ be the solution to (3.2) with

$\wp$ replaced by $\wp^l$, λ by λ^l, X_0 by X_0^l and Y_0 by Y_0^l. In other words

$$\left.\begin{aligned} X_t^l &= X_0^l + \int\limits_{(0,t]}\int\limits_{\mathbb{R}} h^l(X_{u-}^l, Y_{u-}^l, z_0)\wp^l(du, dz_0) \\ Y_t^l &= Y_0^l + t - \int\limits_{(0,t]}\int\limits_{\mathbb{R}} g^l(X_{u-}^l, Y_{u-}^l, z_0)\wp^l(du, dz_0), \end{aligned}\right\} \tag{3.3}$$

where $h^l := h_{\lambda^l}$ and $g^l := g_{\lambda^l}$. We denote the tuple $(X_t^0, X_t^1, \ldots, X_t^n)$ by X_t and $(Y_t^0, Y_t^1, \ldots, Y_t^n)$ by Y_t. Let $\mathcal{X} = \{1, \ldots, k\} \subset \mathbb{R}$. For every $l = 0, 1, \ldots, n$, consider a function $\lambda^l : \{(i,j) \in \mathcal{X}^2 | i \neq j\} \times [0, \infty) \to (0, \infty)$ with the condition (3.1) satisfying the following assumptions:

Assumption 3.1. *(i)* $\lambda_{ii}^l(\bar{y}) = -\sum_{j \neq i} \lambda_{ij}^l(\bar{y})$,

(ii) $\bar{y} \mapsto \lambda_{ij}^l(\bar{y})$ *is continuously differentiable,*

(iii) if $\Lambda_i^l(\bar{y}) := \int\limits_0^{\bar{y}} \sum\limits_{j \neq i} \lambda_{ij}^l(v)dv$, *then* $\lim\limits_{\bar{y} \to \infty} \Lambda_i^l(\bar{y}) = \infty$.

Remark 3.3.2. *It is important to note that Assumptions 3.1 (ii), is not required to construct a CSM process. We need this assumption to establish some regularity of holding time distribution.*

For $l = 0, 1, \ldots, n$, let us consider the following function

$$F^l(\bar{y}|i) := 1 - e^{-\Lambda_i^l(\bar{y})} \text{ for } \bar{y} \geq 0.$$

Let $f^l(\bar{y}|i)$ be the derivative of $F^l(\bar{y}|i)$ with respect to $\bar{y}$. Now for each $l = 0, 1, \ldots, n$, we consider the matrix p^l, where for all i and $\bar{y}$

$$p_{ij}^l(\bar{y}) := \begin{cases} \frac{\lambda_{ij}^l(\bar{y})}{|\lambda_{ii}^l(\bar{y})|} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \tag{3.4}$$

Now by setting $\hat{p}_{ij}^l = \int\limits_0^\infty p_{x^l j}^l(\bar{y})dF^l(\bar{y}|i)$, we make further assumption.

Assumption 3.2. *The matrix* $(\hat{p}_{ij}^l)$ *is irreducible for each* $l = 0, 1, \ldots, n$.

From the definition of F^l and the assumptions on λ we observe $F^l(\bar{y}|i) < 1, \quad \forall \bar{y} > 0$. We also note that $\lambda_{ij}^l(\bar{y}) = p_{ij}^l(\bar{y})\frac{f^l(\bar{y}|i)}{1-F^l(\bar{y}|i)}$ hold for $i \neq j$. It is shown in [27] that $F^l(\bar{y}|i)$ is the conditional c.d.f of the holding time of X^l and $p_{ij}^l(\bar{y})$ is the conditional probability that X^l transits to j given the fact that it is at i for a duration of $\bar{y}$.

Now we will define some important notations for our future work.

Notation 3.1. *T_n^l denote the time of n-th transition of the l-th component of X_t, whereas $T_0^l = 0$ and $\tau_n^l := T_n^l - T_{n-1}^l$.*

Notation 3.2. *For a fixed t, let $n^l(t) := \max\{n : T_n^l \le t\}$. Hence $T_{n^l(t)}^l \le t \le T_{n^l(t)+1}^l$ and $Y_t^l = t - T_{n^l(t)}^l$.*

Notation 3.3. *$\tau^l(t) :=$ time period from time t after which the l-th component of X_t would have a first transition. Note that $\tau^l(t)$ is independent of every component of X other than l-th one.*

Notation 3.4. *We denote the conditional c.d.f of $\tau^l(t)$ given $X_t^l = i$ and $Y_t^l = \bar{y}$ as $F_{\tau^l}(\cdot|i,\bar{y})$, and the conditional p.d.f of $\tau^l(t)$ given $X_t^l = i$ and $Y_t^l = \bar{y}$ as $f_{\tau^l}(\cdot|i,\bar{y})$.*

Notation 3.5. *Let $\ell(t)$ be the component of X_t, where the subsequent transition happens.*

Remark 3.3.3. *It is important to note that this c.d.f does not depend on t mainly because (X_t, Y_t) is time-homogeneous. We also notice that $\tau^l(t) + Y_t^l$ is the duration of stagnancy of X_t^l at present state before it moves to another.*

From now we denote $P(\cdot|X_t = x, Y_t = y)$ by $P_{t,x,y}(\cdot)$ and the corresponding conditional expectation as $\mathbb{E}_{t,x,y}(\cdot)$. Therefore, $P_{t,x,y}(\ell(t) = l)$ represents the conditional probability of observing next transition to occur at the l-th component given that $X_t = x$ and $Y_t = y$. We find the expressions of the c.d.f and the probability defined above and obtain some properties in the following lemmas.

Lemma 3.3.4. *Consider $F^l, f^l, F_{\tau^l}, f_{\tau^l}, P_{t,x,y}$ as given above.*

(i) For each l, $F_{\tau^l}(r|i,\bar{y}) = \frac{F^l(r+\bar{y}|i) - F^l(\bar{y}|i)}{1-F^l(\bar{y}|i)}$, and $f_{\tau^l}(r|i,\bar{y}) = \frac{f^l(r+\bar{y}|i)}{1-F^l(\bar{y}|i)}$, for $r \ge 0$.

(ii) The joint probability distribution of $\tau^l(t)$ and $\ell(t) = l$ is given by

$$P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l) = \int_0^r \prod_{m \ne l} (1 - F_{\tau^m}(s|x^m, y^m)) f_{\tau^l}(s|x^l, y^l) ds.$$

(iii) For each l, $P_{t,x,y}(\ell(t) = l) = \int_0^\infty \prod_{m \ne l} (1 - F_{\tau^m}(s|x^m, y^m)) f_{\tau^l}(s|x^l, y^l) ds.$

(iv) Let $F_{\tau^l|l}(\cdot|x,y)$ be the conditional c.d.f of $\tau^l(t)$ given $X_t = x, Y_t = y$ and $\ell(t) = l$. Then $F_{\tau^l|l}(r|x,y) = \frac{P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l)}{P_{t,x,y}(\ell(t) = l)}$.

Proof. (i) One can compute the conditional c.d.f $F_{\tau^l}(\cdot|i,\bar{y})$ in the following way

$$\begin{aligned} F_{\tau^l}(r|i,\bar{y}) &= P(0 \le \tau^l(t) \le r|X_t^l = i, Y_t^l = \bar{y}) \\ &= P(\tau^l(t) + Y_t^l \le r + \bar{y}|X_t^l = i, Y_t^l = \bar{y}) \\ &= P(Y^l_{T^l_{n^l(t)+1}-} \le r + \bar{y}|Y^l_{T^l_{n^l(t)}-} \ge \bar{y}, X_t^l = i, Y_t^l = \bar{y}) \\ &= \frac{F^l(r+\bar{y}|i) - F^l(\bar{y}|i)}{1 - F^l(\bar{y}|i)} \qquad l = 0, 1, \ldots, n. \end{aligned} \tag{3.5}$$

We denote the derivative of $F_{\tau^l}(r|i,\bar{y})$ by $f_{\tau^l}(r|i,\bar{y})$, given by

$$f_{\tau^l}(\cdot|i,\bar{y}) = \frac{f^l(\cdot+\bar{y}|i)}{1 - F^l(\bar{y}|i)}. \tag{3.6}$$

This proves (i).

(ii) We introduce a new variable $\tau^{-l}(t) := \min_{m \ne l} \tau^m(t)$. We denote the conditional c.d.f of $\tau^{-l}(t)$ given $X_t = x$ and $Y_t = y$ as $F_{\tau^{-l}}(\cdot|x,y)$ which is equal to

$$1 - \prod_{m \ne l} \left(1 - F_{\tau^m}(\cdot|x^m, y^m)\right).$$

It is easy to see that $P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l) = P_{t,x,y}(\tau^{-l}(t) > \tau^l(t), \tau^l(t) \le r)$. To compute this probability we use a conditioning on $\tau^l(t)$. Thus

$$\begin{aligned} P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l) &= \mathbb{E}_{t,x,y}[P_{t,x,y}(\tau^{-l}(t) > \tau^l(t), \tau^l(t) \le r|\tau^l(t))] \\ &= \int_0^r P_{t,x,y}(\tau^{-l}(t) > \tau^l(t)|\tau^l(t) = s) f_{\tau^l}(s|x^l, y^l) ds \\ &= \int_0^r (1 - P_{t,x,y}(\tau^{-l}(t) \le s)) f_{\tau^l}(s|x^l, y^l) ds \\ &= \int_0^r \prod_{m \ne l} (1 - F_{\tau^m}(s|x^m, y^m)) f_{\tau^l}(s|x^l, y^l) ds. \end{aligned} \tag{3.7}$$

This proves (ii).

(iii) It is obvious that, $P_{t,x,y}(\ell(t) = l) = P_{t,x,y}(\tau^l(t) \le \infty, \ell(t) = l)$ and this completes the proof of (iii).

(iv) From the definition of $F_{\tau^l|l}(r|x,y)$ we have,

$$F_{\tau^l|l}(r|x,y) = P_{t,x,y}(\tau^l(t) \le r|\ell(t) = l)$$

$$= \frac{P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l)}{P_{t,x,y}(\ell(t) = l)}. \tag{3.8}$$

Now using (3.7), and (3.8), we completed the proof of (iii).

□

In the next lemma, we shall discuss about the regularity structure of $P_{t,x,y}(\ell(t) = l)$ and $F_{\tau^l|l}(v|x, y)$, which will be used in the forthcoming Chapters.

Lemma 3.3.5. *Let $F_{\tau^l}(r|i, \bar{y}), f_{\tau^l}(r|i, \bar{y}), F_{\tau^l|l}(r|x, \bar{y})$ be as in Lemma 3.3.4.*

(i) Let $f_{\tau^l|l}(r|x, \bar{y})$ is the derivative of $F_{\tau^l|l}(r|x, \bar{y})$ in r variable. Then the following identity holds

$$f_{\tau^l|l}(0|x, \bar{y})P_{t,x,\bar{y}}(\ell(t) = l) = f_{\tau^l}(0|x^l, y^l) = \frac{f^l(y^l|x^l)}{1 - F^l(y^l|x^l)}$$

.

(ii) For each l, $f_{\tau^l}(r|i, \bar{y})$ is differentiable in r. Furthermore, $F_{\tau^l}(r|i, \bar{y})$ and $f_{\tau^l}(r|i, \bar{y})$ are differentiable in $\bar{y}$.

(iii) $P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l)$ is differentiable in y variable and

$$\sum_{i=0}^{n} \frac{\partial}{\partial y^i} P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l) = \sum_{i=0}^{n} f_{\tau^i}(0|x^i, y^i)P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l) + \left[\prod_{m \ne l} (1 - F_{\tau^m}(r|x^m, y^m))\, f_{\tau^l}(r|x^l, y^l) - f_{\tau^l}(0|x^l, y^l)\right].$$

(iv) $P_{t,x,y}(\ell(t) = l)$ is differentiable in y and

$$\sum_{i=0}^{n} \frac{\partial}{\partial y^i} P_{t,x,y}(\ell(t) = l) = \sum_{i=0}^{n} f_{\tau^k}(0|x^i, y^i)P_{t,x,y}(\ell(t) = l) - f_{\tau^l}(0|x^l, y^l).$$

(v) $F_{\tau^l|l}(r|x, y)$ is differentiable with respect to y and

$$\sum_{i=0}^{n} \frac{\partial}{\partial y^i} F_{\tau^l|l}(r|x, y) = f_{\tau^l}(0|x^l, y^l)\frac{F_{\tau^l|l}(r|x, y) - 1}{P_{t,x,y}(\ell(t) = l)} - \prod_{m \ne l} (1 - F_{\tau^m}(r|x^m, y^m))\, f_{\tau^l}(r|x^l, y^l)$$

Proof. (i) Consider the expression of Lemma 3.3.4(iii) and differentiating it with respect to r, we obtain the result.

(ii) Since λ^l is C^1 in r, so is $f_{\tau^l}(r|i,\bar{y})\ \forall\, l$. Again since for all l, F^l and f^l are differentiable in $\bar{y}$, it follows from 3.3.4(i) that $F_{\tau^l}(r|i,\bar{y})$ and $f_{\tau^l}(r|i,\bar{y})$ are differentiable in $\bar{y}$. By a straight forward calculation, we can obtain

$$\frac{\partial}{\partial \bar{y}} F_{\tau^l}(r|i,\bar{y}) = f_{\tau^l}(r|i,\bar{y}) - f_{\tau^l}(0|i,\bar{y})\left(1 - F_{\tau^l}(r|i,\bar{y})\right) \tag{3.9}$$

$$\frac{\partial}{\partial \bar{y}} f_{\tau^l}(r|i,\bar{y}) = f_{\tau^l}(r|i,\bar{y}) f_{\tau^l}(0|i,\bar{y}) + \frac{\bar{f}^l(r+\bar{y}|i)}{1 - F^l(\bar{y}|i)}, \tag{3.10}$$

where $\bar{f}^l$ denotes the derivative of f^l with respect to $\bar{y}$.

(iii) We first recall the Definition of $P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l)$ from Lemma 3.3.4(ii),

$$P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l) = \int_0^r \prod_{m \ne l} (1 - F_{\tau^m}(s|x^m, y^m)) f_{\tau^l}(s|x^l, y^l) ds.$$

Since $\bar{f}_{\tau^l}(s|x^l,\bar{y})$ is bounded for $\bar{y} \in [0,r]$ and it is differentiable with respect to $\bar{y}$, then by using Fatou's lemma we can interchange the limit and integration. Therefore,

$$\frac{\partial}{\partial y^j} P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l)$$
$$= \begin{cases} \displaystyle\int_0^r \prod_{m \ne l} (1 - F_{\tau^m}(s|x^m, y^m)) \frac{\partial}{\partial y^l} f_{\tau^l}(s|x^l, y^l) ds, & j = l \\ \displaystyle -\int_0^r \prod_{m \ne l \ne j} (1 - F_{\tau^m}(s|x^m, y^m)) \frac{\partial}{\partial y^j} F_{\tau^j}(s|x^j, y^j) f_{\tau^l}(s|x^l, y^l) ds, & j \ne l. \end{cases}$$

Now,

$$\sum_j \frac{\partial}{\partial y^j} P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l)$$
$$= \int_0^r \prod_{m \ne l} (1 - F_{\tau^m}(s|x^m, y^m)) \frac{\partial}{\partial y^l} f_{\tau^l}(s|x^l, y^l) ds$$
$$- \sum_{j \ne l} \int_0^r \prod_{m \ne l \ne j} (1 - F_{\tau^m}(s|x^m, y^m)) \frac{\partial}{\partial y^j} F_{\tau^j}(s|x^j, y^j) f_{\tau^l}(s|x^l, y^l) ds$$

Now using (3.9) and (3.10), we have

$$\sum_j \frac{\partial}{\partial y^j} P_{t,x,y}(\tau^l(t) \le r, \ell(t) = l)$$

$$= \int_0^r \prod_{m\neq l} (1 - F_{\tau^m}(s|x^m, y^m)) \left(f_{\tau^l}(s|x^l, y^l) f_{\tau^l}(0|x^l, y^l) + \frac{\bar{f}^l(s + y^l|x^l)}{1 - F^l(y^l|x^l)} \right) ds$$

$$- \int_0^r \sum_{j\neq l} \prod_{m\neq l\neq j} (1 - F_{\tau^m}(s|x^m, y^m)) \left(f_{\tau^j}(s|x^j, y^j) - f_{\tau^j}(0|x^j, y^j) \left(1 - F_{\tau^j}(s|x^j, y^j) \right) \right)$$

$$\times f_{\tau^l}(s|x^l, y^l)\, ds. \tag{3.11}$$

Now by using Lemma 3.3.4(i), (ii) and integration by parts in the first integral of (3.11), we have

$$P_{t,x,y}(\tau^l(t) \leq r, \ell(t) = l) f_{\tau^l}(0|x^l, y^l)$$
$$+ \left[\prod_{m\neq l} (1 - F_{\tau^m}(r|x^m, y^m)) f_{\tau^l}(r|x^l, y^l) - f_{\tau^l}(0|x^l, y^l) \right]$$
$$+ \int_0^r \sum_{i\neq l} \prod_{m\neq l\neq i} (1 - F_{\tau^m}(s|x^m, y^m)) f_{\tau^i}(s|x^i, y^i) f_{\tau^l}(s|x^l, y^l)\, ds \tag{3.12}$$

The second integral of (3.11) can be rewritten as

$$- \int_0^r \sum_{i\neq l} \prod_{m\neq l\neq i} (1 - F_{\tau^m}(s|x^m, y^m)) f_{\tau^i}(s|x^i, y^i) f_{\tau^l}(s|x^l, y^l)\, ds$$
$$+ \sum_{i\neq l} f_{\tau^r}(0|x^i, y^i) \int_0^r \prod_{m\neq l} (1 - F_{\tau^m}(s|x^m, y^m)) f_{\tau^l}(s|x^l, y^l)\, ds \tag{3.13}$$

Therefore by substituting (3.12) and (3.13) in (3.11), we have the desired result.

(iv) This result is obvious.

(v) Using (iii) and (iv), it can be proved. □

Now we state a corollary of the preceding Lemma, which will be used in the succeeding Chapters.

Corollary 3.3.6. *Let $0 \leq t \leq T$, then the following identity holds*

$$(\frac{\partial}{\partial t} + \sum_{i=0}^{n} \frac{\partial}{\partial y^i}) F_{\tau^l|l}(T - t|x, y) = f_{\tau^l|l}(0|x, y)(F_{\tau^l|l}(T - t|x, y) - 1).$$

4

Portfolio Optimization

4.1 Introduction

This chapter concerns with a risk-sensitive portfolio optimization problem. We consider a financial market consisting of several assets, governed by a CSM modulated jump diffusion. Under the above market assumptions, we solve the optimization problem by studying an HJB equation. Using some separation of variables, we reduce the HJB equation to a linear first order system of non-local PDEs. To show the well-posedness of the linear PDE, we study an equivalent Volterra integral equation(VIE) of the second kind.

The rest of the chapter is organized as follows. In the next section we give a rigorous description of the model of a financial market dynamics and then derive the wealth process of an investor's portfolio. In Section 3 we describe the optimization criteria and the equations of corresponding optimal portfolio. The problem of optimizing the portfolio wealth under the risk sensitive criterion on the finite time horizon is presented in Section 4. Section 5 contain some concluding remarks.

4.2 Model Description

Key Assumptions of the chapter

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

4.2.1 Model parameters

Let $\mathcal{X}$ denote a finite subset of $\mathbb{R}$. Without loss of generality, we choose $\mathcal{X} = \{1, 2, \ldots, k\}$ and $\mathcal{X}_2 := \{(i, j) | i \neq j \in \mathcal{X}\}$. Consider for each $l = 0, 1, \ldots, n$, $\lambda^l : \mathcal{X}_2 \times [0, \infty) \to (0, \infty)$ a continuously differentiable function in y with $\lambda^l_{ii}(y) = -\sum_{j\neq i}\lambda^l_{ij}(y)$ and

$$\lim_{y\to\infty}\Lambda^l_i(y) = \infty, \text{where}\, \Lambda^l_i(y) := \int_0^y \sum_{j\neq i}\lambda^l_{ij}(v)dv.$$

Assume that for each $j = 1, 2, \ldots, m_2$, ν_j denotes a finite Borel measure on $\mathbb{R}$. Let $r : [0,T]\times\mathcal{X}^{n+1} \to [0,\infty)$, $\mu^l : [0,T]\times\mathcal{X}^{n+1} \to \mathbb{R}$, and $\sigma^l : [0,T]\times\mathcal{X}^{n+1} \to (0,\infty)^{1\times m_1}$ be continuous functions of the time variable for each $l = 1,\ldots,n$, where m_1 and m_2 are the positive integers. We also consider a collection of measurable functions $\eta_{lj} : \mathbb{R} \to (-1,\infty)$ for each $l = 1,\ldots,n$, $j = 1,\ldots,m_2$.

We further introduce some more notations. Fix $x = (x_0, x_1, \ldots, x_n) \in \mathcal{X}^{n+1}$ and $t \in [0,T]$ and we denote $b(t,x) := [\mu^1(t,x) - r(t,x), \ldots, \mu^n(t,x) - r(t,x)]_{1\times n}$, and $\sigma(t,x) := [\sigma_{lj}(t,x)]_{n\times m_1}$, where σ_{lj} is the j-th component of σ^l function. For each $z = (z_1,\ldots,z_{m_2}) \in \mathbb{R}^{m_2}$, we denote $\eta(z) := [\eta_{lj}(z_j)]_{n\times m_2}$. We use $[\cdot]^*$ to denote transpose of a vector.

4.2.2 Asset price model

Let $\{X^l_0 \mid l = 0,\ldots,n\}$ be a collection of $(n+1)$ $\mathcal{X}$ valued random variables, and $\{Y^l_0 \mid l = 0,\ldots,n\}$ be a collection of $(n+1)$ non negative random variables. Let $W = \{W_t\}_{t\geq 0} = \{[W^1_t,\ldots,W^{m_1}_t]^*\}_{t\geq 0}$ be a standard m_1-dimensional Brownian motion. We further assume that, $\{N_j(dt,dz) | j = 1,\ldots,m_2\}$ on $(0,\infty)\times\mathbb{R}$ and $\{\wp^l(dt,dz_0) \mid l = 0,\ldots,n\}$ on $(0,\infty)\times\mathbb{R}$ are two sets of Poisson random measures with intensity $\nu_j(dz)dt$ and $dtdz_0$ respectively defined on the same probability space. For each j, ν_j denotes a finite Borel measure. It is important to note that the random variables, processes and measures are defined in such a way that they are independent. Let $X := \{X_t\}_{t\geq 0}$ where $X_t = (X^0_t, X^1_t, \ldots, X^n_t)$ be a CSM process whose each component is a solution of the following stochastic differential equation

$$\begin{aligned} X^l_t &= X^l_0 + \int_{(0,t]}\int_{\mathbb{R}} h^l(X^l_{u-}, Y^l_{u-}, z_0)\wp^l(du,dz_0) \\ Y^l_t &= Y^l_0 + t - \int_{(0,t]}\int_{\mathbb{R}} g^l(X^l_{u-}, Y^l_{u-}, z_0)\wp^l(du,dz_0), \end{aligned}$$

where

$$h_{\lambda^l}(i,y,z) := \sum_{j\in\mathcal{X}\setminus\{i\}}(j-i)1_{\Lambda^l_{ij}(y)}(z), \quad g_{\lambda^l}(i,y,z) := \sum_{j\in\mathcal{X}\setminus\{i\}} y1_{\Lambda^l_{ij}(y)}(z).$$

We consider a frictionless market[1] consisting of $(n+1)$ assets whose prices are denoted by $\{S^0_t, S^1_t, \ldots, S^n_t\}$ and are traded continuously. We model the hypothetical state of the

[1] a financial market without any transaction cost.

assets at time t by the pure jump process X_t. The state of the asset indicates its mean growth rate and volatility. We assume

$$dS_t^0 = r(t, X_t)S_t^0 dt, \quad S_0^0 = s_0 \geq 0.$$

Thus the corresponding asset is (locally) risk free, which refers to the money market account with the floating interest rate $r(t, x)$ at time t corresponding to regime x. The other n asset prices are assumed to be given by the following stochastic differential equation

$$\begin{aligned} dS_t^l &= S_{t-}^l \left[\mu^l(t, X_t)dt + \sum_{j=1}^{m_1} \sigma_{lj}(t, X_t)\; dW_t^j + \sum_{j=1}^{m_2} \int_{\mathbb{R}} \eta_{lj}(z_j)\, N_j(dt, dz_j) \right], \qquad (4.1) \\ S_0^l &= s_l, \quad s_l \geq 0,\; l = 1, 2, \ldots, n. \end{aligned}$$

These prices correspond to n different risky assets. Therefore, μ^l represents the growth rate of the l-th asset and σ the volatility matrix of the market. Consider the filtration $\mathscr{F}_t$ be the right continuous augmentation of the filtration generated by $W, X, N_j\; j = 1, \ldots, m_2$ such that $\mathscr{F}_0$ contains all the $\mathbb{P}$-null sets. Here we further assume the following.

Assumption 4.1. *(i) For each $l = 1, \ldots, n$ and $j = 1, \ldots, m_2$, we assume $\eta_{lj} \in L^1(\nu_j) \cap L^2(\nu_j)$.*

(ii) For each $l = 1, \ldots, n$ and $j = 1, \ldots, m_2$, we further assume $\ln(1 + \eta_{lj}) \in L^2(\nu_j)$.

(iii) Let $a(t, x) := \sigma(t, x)\sigma(t, x)^*$ denote the diffusion matrix. Assume that there exist a $\delta_1 > 0$ such that for each t and x, $\xi^* a(t, x)\xi \geq \delta_1 \|\xi\|^2$, where $\|\cdot\|$ denotes the Euclidean norm, and we use $[\cdot]^*$ to denote transpose of a vector.

The next lemma asserts the existence and uniqueness of the solution to the SDE (4.1).

Lemma 4.2.1. *Under the Assumption 4.1(ii) the equation (4.1) has a strong solution, which is adapted, a.s. unique and a rcll process.*

Proof. First we show the uniqueness by assuming that the SDE (4.1) admits a solution, $\{S_t^l\}_{t\geq 0}$, say, the stopping time $\tau = \min\{t \in [0, \infty) \mid S_t^l \leq 0\}$. By applying Itô's lemma (Lemma 1.1.16) on $\ln(S_s^l)$ for $0 \leq s < t \wedge \tau$ we get,

$$\begin{aligned} d\ln(S_s^l) = \frac{S_{s-}^l}{S_{s-}^l} \left[\mu^l(s, X_{s-})ds + \sum_{j=1}^{m_1} \sigma_{lj}(s, X_{s-})\; dW_s^j \right] - \frac{1}{2}(S_{s-}^l)^{-2}(S_{s-}^l)^2 a_{ll}(s, X_{s-})ds \\ + \sum_{j=1}^{m_2} \int_{\mathbb{R}} \Big[\ln(S_{s-}^l + \eta_{lj}(z_j)S_{s-}^l) - \ln(S_{s-}^l) \Big]\; N_j(ds, dz_j) \end{aligned}$$

Integrating both sides from 0 to $t \wedge \tau$ yields,

$$\ln\left(S_{t\wedge\tau}^l\right) - \ln s_l = \int_0^{t\wedge\tau} \Big(\mu^l(s, X_{s-}) - \frac{1}{2} a_{ll}(s, X_{s-})ds \Big) + \sum_{j=1}^{m_1} \int_0^{t\wedge\tau} \sigma_{lj}(s, X_{s-})\, dW_s^j$$

$$+\sum_{j=1}^{m_2}\int_0^{t\wedge\tau}\int_{\mathbb{R}} \ln(1+\eta_{lj}(z_j))\,N_j(ds,dz_j),$$

where all the integrals have finite expectations almost surely by using Assumption 4.1(ii).

$$S^l_{t\wedge\tau} = s_l \exp\Bigg[\int_0^{t\wedge\tau}\Big(\mu^l(s,X_{s-}) - \frac{1}{2}a_{ll}(s,X_{s-})\Big)\,ds + \sum_{j=1}^{m_1}\int_0^{t\wedge\tau}\sigma_{lj}(s,X_{s-})\,dW^j_s \\ + \sum_{j=1}^{m_2}\int_0^{t\wedge\tau}\int_{\mathbb{R}} \ln(1+\eta_{lj}(z_j))\,N_j(ds,dz_j)\Bigg] \tag{4.2}$$

Thus any solution to (4.1) has the above expression. Under Assumption 4.1(ii), $\int_0^{\tau}\int_{\mathbb{R}} \ln(1+\eta_{lj}(z_j))\,N_j(ds,dz_j)$ has finite expectation for any finite stopping time τ.

Let $\Omega_1 := \{\omega\in\Omega : \tau(\omega) < \infty\}$. Now if possible, assume $P(\Omega_1) > 0$. By letting $t\to\infty$ in the above expression, we obtain that $S^l_{\tau(\omega)-}$ is exponential of a random variable which is finite for almost every $\omega\in\Omega_1$. Thus $S^l_{\tau(\omega)-} > 0$. But for almost every $\omega\in\Omega_1$ $S^l_{\tau(\omega)} \le 0$. Hence non-positivity occurred only by jump. In other words $\eta_{lj}(z_j) \le -1$ for some z_j. But that is contrary to the assumption on η. Hence $\tau = \infty$ $\mathbb{P}$ *a.s.* Therefore, $S^l_t > 0$ $\mathbb{P}$ *a.s.* $\forall\, t\in(0,\infty)$ and is given by

$$S^l_t = S^l_0 \exp\Bigg[\sum_{j=1}^{m_1}\int_0^t \sigma_{lj}(s,X_{s-})\,dW^j_s + \sum_{j=1}^{m_2}\int_0^t\int_{\mathbb{R}} \ln(1+\eta_{lj}(z_j))\,\bar{N}_j(ds,dz_j) \\ + \int_0^t\Bigg(\mu^l(s,X_{s-}) - \frac{1}{2}(\sigma_l(s,X_{s-})\sigma_l(s,X_{s-})^*) \\ + \sum_{j=1}^{m_2}\int_{|z_j|<1}(\ln(1+\eta_{lj}(z_j)) - \eta_{lj}(z_j))\;\nu_j(dz_j)\Bigg)\,ds\Bigg]. \tag{4.3}$$

Thus by equation (4.3), $\{S^l_t\}_{t\ge0}$ is an adapted and rcll process and is uniquely determined with the initial condition $S^l_0 = s_0$. Hence the solution is unique.
It is easy to show by a direct calculation that the process S^l_t, given by (4.3) indeed solves the SDE (4.1). □

Remark 4.2.2. *We note that Assumption 4.1(i) and Assumption 4.1(ii) follow for special case where*

$$-1 < \inf_{z\in\mathbb{R}} \eta_{lj}(z) \le \sup_{z\in\mathbb{R}} \eta_{lj}(z) < \infty.$$

By Assumption 4.1(ii) the diffusion matrix $a(t,x)$ is uniformly positive definite, which ensures that $a(t,x)$ is invertible. We will use this condition in Section 3. This condition also implies that $m_1 \geq n$.

4.2.3 Portfolio value process

Consider an investor who is employing a self-financing portfolio of the above $(n+1)$ assets starting with a positive wealth. If the portfolio at time t comprises of π_t^l number of units of l-th asset for every $l = 0, \ldots, n$, then for each $\omega \in \Omega$ the value of the portfolio at time t is given by

$$V_t := \sum_{l=0}^{n} \pi_t^l \; S_t^l.$$

We allow π_t^l be real valued, i.e., borrowing from the money market and short selling of assets are allowed. We further assume that $\{\pi_t^l\}_{t\geq 0}$ is a $\{\mathscr{F}_t\}_{t\geq 0}$ adapted, rcll process for each l. Then the self-financing condition implies that

$$dV_t \;=\; \sum_{l=0}^{n} \pi_{t-}^l \; dS_t^l.$$

If π_t^l are such that V_t remains positive, we can set $u_t^l := \frac{\pi_t^l S^l(t)}{V_t}$, the fraction of investment in the l-th asset. Then we have $\sum_{l=0}^{n} u_t^l = 1$ and hence $u_t^0 = 1 - \sum_{l=1}^{n} u_t^l$. We call $u_t = [u_t^1, u_t^2, \ldots, u_t^n]^*$ as the portfolio strategy of risky assets at time t. Then the wealth process, $\{V_t\}_{t\geq 0}$, now onward denoted by $\{V_t\}_{t\geq 0}$, takes the form

$$\frac{dV_t^u}{V_{t-}^u} = \sum_{l=0}^{n} u_{t-}^l \frac{dS_t^l}{S_{t-}^l}.$$

Thus we would consider the following SDE for the value process,

$$\begin{aligned} dV_t^u = V_t^u \Bigg(& r(t,X_t) + \sum_{l=1}^{n} \Big[\mu^l(t,X_t) - r(t,X_t)\Big] u_t^l \Bigg) dt \\ & + V_t^u \sum_{l=1}^{n} \sum_{j=1}^{m_1} \sigma_{lj}(t,X_t) \; u_t^l dW_t^j \\ & + V_{t-}^u \sum_{l=1}^{n} \sum_{j=1}^{m_2} u_{t-}^l \int_{\mathbb{R}} \eta_{lj}(z_j) N_j(dt, dz_j). \end{aligned} \tag{4.4}$$

Now we denote $b(t,x) := [\mu^1(t,x) - r(t,x), \ldots, \mu^n(t,x) - r(t,x)]_{1\times n}$, and $\sigma(t,x) := [\sigma_{lj}(t,x)]_{n\times m_1}$, where σ_{lj} is the j-th component of σ^l function. For each $z = (z_1, \ldots, z_{m_2}) \in \mathbb{R}^{m_2}$, we denote $\eta(z) := [\eta_{lj}(z_j)]_{n\times m_2}$. Using (4.1), (4.5) can be rewritten as

$$dV_t^u = V_t^u(\; r(t,X_t) + b(t,X_t)u_t)dt + V_t^u u_t^* \sigma(t,X_t) dW_t$$

$$+ V_{t-}^u \sum_{j=1}^{m_2} \int_{\mathbb{R}} \left[u_{t-}^* \eta(z)\right]_j N_j(dt, dz_j), \tag{4.5}$$

where $u_t^* \eta(z) = \left[\sum_{l=1}^n u_t^l \eta_{l1}(z_1), \ldots, \sum_{l=1}^n u_t^l \eta_{lm_2}(z_{m_2})\right]_{1\times m_2}$.

Remark 4.2.3. *It is important to note that for all choices of u, the SDE* (4.5) *need not have a strong solution. Therefore we should restrict ourselves to particular class of portfolio strategy.*

It is clear from the definition and above derivations that V_t^u, the portfolio wealth process, is a controlled process. Let $\mathbb{A} \subseteq \mathbb{R}^n$ be a convex set containing the origin, denoting the range of portfolio. The range is determined based on investment restrictions. For example, $\mathbb{A} = \mathbb{R}^n$ in the case of unrestricted short selling. The restrictions on short selling makes $\mathbb{A} = \{u \in \mathbb{R}^n \mid u^l \geq c_l, \sum_{l\geq 1} u^l \leq 1 - c_0 \forall l\}$, where $c_l \leq 0$ for $l = 0, \ldots, n$. Clearly, $c_l = 0$ for $l = 0, \ldots, n$ correspond to no short selling. Now we shall define the class of admissible portfolio strategy for our problem.

Definition 4.2.4. *An adapted process $u = \{u_t\}_{t\in[0,T]}$ is said to be admissible portfolio strategy if:*

(i) *the process u takes values from the convex set $\mathbb{A}_1 := \mathbb{A} \bigcap \mathfrak{U}_\delta$, where $\mathfrak{U}_\delta := \{u \in \mathbb{R}^n \mid [u^* \eta(z)]_j \geq -1 + \delta, \forall j, z\}$ for some $0 < \delta \leq 1$,*

(ii) (4.5) *has a unique strong solution,*

(iii) $\operatorname{ess\,sup}_{\Omega} \sup_{[0,T]} \|u_t(\omega)\| < \infty$.

Remark 4.2.5. *It is important to note that the set of admissible portfolio strategy is non empty as the constant zero process is in the set of admissible strategies.*

Lemma 4.2.6. *Let $\bar{N}$ be a Poisson random measure on $(\Omega, \mathscr{F}, P)$ with intensity $\bar{\nu}(dz)\, dt$, where $\bar{\nu}$ is a finite measure. If $\bar{\eta} \in L^2(\nu)$, then there exists a positive constant c such that*

$$\mathbb{E}\left[\exp\left(\int_0^t \int_{\mathbb{R}} \ln(1 + \bar{\eta}^2(z))\, \bar{N}(ds, dz)\right)\right] = \exp\left(ct\bar{\nu}(\mathbb{R})\right).$$

Proof. We first note that $|\bar{N}_t| := \bar{N}([0,t] \times \mathbb{R})$ is finite a.s. as $|\bar{\nu}| < \infty$. Therefore the integral $\int_0^t \int_{\mathbb{R}} \ln(1 + \bar{\eta}^2(z))\, \bar{N}(ds, dz)$ can be written as $\sum_{i=1}^{|\bar{N}_t|} \ln(1 + \bar{\eta}^2(z_i))$, where $\{(t_i, z_i)\}$ are the point masses of $\bar{N}$ on $[0,t]\times\mathbb{R}$. To be more precise, $\bar{N}(A) = \sum_{i=1}^{|\bar{N}_t|}\delta_{\{(t_i,z_i)\}}(A)$ for all $A\in\mathscr{B}([0,t]\times\mathbb{R})$. Therefore

$$\mathbb{E}\left[\exp\left(\int_0^t\int_{\mathbb{R}}\ln(1+\bar{\eta}^2(z))\,\bar{N}(ds,dz)\right)\right] = \mathbb{E}\left[\prod_{i=1}^{|\bar{N}_t|}(1+\bar{\eta}^2(z_i))\right]. \tag{4.6}$$

By conditioning on $|\bar{N}_t|$, the right hand side of (4.6) can be rewritten as

$$\mathbb{E}\left[\mathbb{E}\left[\prod_{i=1}^{|\bar{N}_t|}(1+\bar{\eta}^2(z_i))\middle\| \bar{N}_t|\right]\right].$$

Since $(1+\bar{\eta}^2(z_1)),\dots,(1+\bar{\eta}^2(z_{|\bar{N}_t|}))$ are conditionally independent and identically distributed given $|\bar{N}_t| = n$, the above expectation becomes

$$\sum_{n=1}^{\infty}\mathbb{E}[(1+\bar{\eta}^2(z_1))]^n P(|\bar{N}_t|=n).$$

Now using $\mathbb{E}\left[\bar{\eta}^2(z_1)\right] = c$, and $P(|\bar{N}_t| = n) = e^{-t\bar{\nu}(\mathbb{R})}\frac{(t\bar{\nu}(\mathbb{R}))^n}{n!}$ in the above expectation, we have

$$\begin{aligned}&\sum_n (1+c)^n e^{-t\bar{\nu}(\mathbb{R})}\frac{(t\bar{\nu}(\mathbb{R}))^n}{n!}\\ &= e^{-t\bar{\nu}(\mathbb{R})}\exp\left(t\bar{\nu}(\mathbb{R})(1+c)\right)\\ &= \exp\left(ct\bar{\nu}(\mathbb{R})\right).\end{aligned}$$

□

Lemma 4.2.7. *Under Assumption 4.1(i) and with admissible control u, (i) the SDE* (4.5) *has a unique positive strong solution with finite expectation for an admissible control u. (ii) The solution has finite moments of all positive and negative orders, which are also bounded on* $[0,T]$ *uniformly in* u.

Proof. (i) We first note that, since $u_t\in\mathfrak{U}_\delta$ and satisfies Definition 4.2.4(iii),

$$|\ln(1+[u^*_{s-}\eta(z)]_j)| < \max\left(|\ln\delta|, C\|\eta_{\cdot j}(z_j)\|\right),$$

where $C := ess\sup_{\Omega}\sup_{[0,T]}\|u_t(\omega)\|$ and $\eta_{\cdot j}$ is the j-th column of the matrix η. Again using Assumption 4.1(i) and finiteness of measure ν_j, the integration of RHS with respect to N_j has finite expectation. This implies that $\mathbb{E}\int_0^t\int_{\mathbb{R}}\ln(1+[u^*_{s-}\eta(z)]_j)\,N_j(ds,dz_j) < \infty$.

Therefore in the similar line of proof of Lemma 4.2.1, we can show under the Assumption 4.1(i) and admissibility of u, (4.5) has an a.s. unique positive rcll solution, which is an adapted process, and the solution is given by

$$V_t^u = V_0^u \exp\left[\int_0^t \left(r(s,X_s) + b(s,X_s)u_s - \frac{1}{2}u_s^* a(s,X_s)u_s\right) ds + \int_0^t u_s^* \sigma(s,X_s)\, dW_s \right.$$
$$\left. + \sum_{j=1}^{m_2} \int_0^t \int_{\mathbb{R}} \ln(1 + [u_{s-}^* \eta(z)]_j)\, N_j(ds, dz_j)\right]. \tag{4.7}$$

(ii) We first consider the first order moment. To prove for each t, V_t^u has a bounded expectation, we first note that the right hand side can be written as a product of a conditionally log-normal random variable and $\exp\left(\sum_{j=1}^{m_2} \int_0^t \int_{\mathbb{R}} \ln(1 + [u_{s-}^* \eta(z)]_j)\, N_j(ds, dz_j)\right)$, where both are conditionally independent, given the process u. We further note that the log-normal random variable has bounded parameters on $[0,T]$ uniformly in u. Therefore it is sufficient to check if

$$\mathbb{E}\left[\exp\left(\int_0^t \int_{\mathbb{R}} \ln(1 + C\|\eta_{\cdot j}(z_j)\|)\, N_j(ds, dz_j)\right)\right],$$

is bounded on $[0,T]$, for all $j = 1, \ldots, m_2$. By applying Lemma 4.2.6, one can show that the above expectation is bounded on $[0,T]$. Thus V_t^u has bounded expectation on $[0,T]$, uniformly in u. Now for the moments of general order, we note that for any $\alpha \in \mathbb{R}$, $(V_t^u)^\alpha$ can also be written in a similar form of (4.7) where each of the integrals inside the exponential would be multiplied by the constant α. Thus the rest of the proof follows in a similar line of that of first order case, given above. □

It is important to note that for fixed u, $Z = \{t, X_t, Y_t, V_t^u\}_{t\geq 0}$ is a time homogeneous Markov process. Let $\mathscr{A}^u$ be the infinitesimal generator of the process Z. We will derive the generator $\mathscr{A}^u$ in the next proposition.

Proposition 4.2.8. *Let $u \in \mathbb{A}_1$ be fixed and (X_t, Y_t) and V_t^u be as in* (3.3)*and* (4.5)*. Let φ be a C^∞ function with compact support then $\varphi \in dom(\mathscr{A}^u)$ and*

$$\begin{aligned}
&\mathscr{A}^u \varphi(t,x,y,v)\\
&= \frac{\partial}{\partial t}\varphi(t,x,y,v) + \frac{\partial}{\partial y}\varphi(t,x,y,v) + v\,[r(t,x) + b(t,x)\ u]\,\frac{\partial}{\partial v}\varphi(t,x,y,v)\\
&\quad + \frac{1}{2}v^2\,[u^* a(t,x)u]\,\frac{\partial^2}{\partial v^2}\varphi(t,x,y,v)\\
&\quad + \sum_{j=1}^{m_2} \int_{\mathbb{R}} [\varphi(t,x,y,v\,(1 + [u^*\eta(z)]_j)) - \varphi(t,x,y,v)]\ \nu_j(dz_j)
\end{aligned}$$

$$+\sum_{l=0}^{n}\sum_{j\neq x^l}\lambda^l_{x^l j}(y^l)\left[\varphi(t,R^l_j x,R^l_0 y,v)-\varphi(t,x,y,v)\right], \tag{4.8}$$

where the linear operator R^l_z *is given by* $R^l_z x := x+(z-x^l)e_l,\ l=0,\dots,n,\ z\in\mathbb{R}$ *and* $\{e_l : l=0,\dots,n\}$ *is the standard basis of* $\mathbb{R}^{n+1}$.

Proof. Applying Itô's formula (Theorem 1.1.16) on φ, using (3.2), and (4.5), we obtain

$$\begin{aligned}\varphi(r,X_r,Y_r,V^u_r) &= \varphi(t,X_t,Y_t,V^u_t)+\int_t^r\frac{\partial}{\partial s}\varphi(s,X_{s-},Y_{s-},V^u_{s-})ds+\int_t^r\frac{\partial}{\partial y}\varphi(s,X_{s-},Y_{s-},V^u_{s-})ds\\
&+\int_t^r\frac{\partial}{\partial v}\varphi(s,X_{s-},Y_{s-},V^u_{s-})d(V^u_s)^c+\frac{1}{2}\int_t^r\frac{\partial^2}{\partial v^2}\varphi(s,X_{s-},Y_{s-},V^u_{s-})\,d[(V^u)^c]_s\\
&+\sum_{j=1}^{m_2}\int_t^r\int_{\mathbb{R}}\left[\varphi(s,X_{s-},Y_{s-},V^u_{s-}+V^u_{s-}[u^*\eta(z)]_j)-\varphi(s,X_{s-},Y_{s-},V^u_{s-})\right]N_j(ds,dz_j)\\
&+\sum_{l=0}^{n}\int_t^r\int_{\mathbb{R}}\Big[\varphi(s,R^l_{X^l_{s-}+h^l(X^l_{s-},Y^l_{s-},z_0)}(X_{s-}),R^l_{Y^l_{s-}-g^l(X^l_{s-},Y^l_{s-},z_0)}(Y_{s-}),V^u_{s-})\\
&-\varphi(s,X_{s-},Y_{s-},V^u_{s-})\Big]\wp^l(ds,dz_0).\end{aligned} \tag{4.9}$$

It is easy to compute $d[(V^u)^c]_s = u^*a(s,X_s)u\,ds$. Now using (3.2), and (4.5), the right hand side of (4.9), can be rewritten as

$$\begin{aligned}&\varphi(t,X_t,Y_t,V^u_t)+\int_t^r\frac{\partial}{\partial s}\varphi(s,X_{s-},Y_{s-},V^u_{s-})ds+\int_t^r\frac{\partial}{\partial y}\varphi(s,X_{s-},Y_{s-},V^u_{s-})ds\\
&+\int_t^r\frac{\partial}{\partial v}\varphi(s,X_{s-},Y_{s-},V^u_{s-})V^u_t(\ r(t,X_t)+b(t,X_t)u)\,dt\\
&+\int_t^r\frac{\partial}{\partial v}\varphi(s,X_{s-},Y_{s-},V^u_{s-})V^u_t u^*_t\sigma(t,X_t)dW_t\\
&+\frac{1}{2}\int_t^r\frac{\partial^2}{\partial v^2}\varphi(s,X_{s-},Y_{s-},V^u_{s-})u^*a(s,X_s)u\,ds\\
&+\sum_{j=1}^{m_2}\int_t^r\int_{\mathbb{R}}\left[\varphi(s,X_{s-},Y_{s-},V^u_{s-}+V^u_{s-}[u^*\eta(z)]_j)-\varphi(s,X_{s-},Y_{s-},V^u_{s-})\right]\tilde{N}_j(ds,dz_j)\\
&+\sum_{l=0}^{n}\int_t^r\int_{\mathbb{R}}\Big[\varphi(s,R^l_{X^l_{s-}+h^l(X^l_{s-},Y^l_{s-},z_0)}(X_{s-}),R^l_{Y^l_{s-}-g^l(X^l_{s-},Y^l_{s-},z_0)}(Y_{s-}),V^u_{s-})\\
&-\varphi(s,X_{s-},Y_{s-},V^u_{s-})\Big]\tilde{\wp}^l(ds,dz_0)\\
&+\sum_{j=1}^{m_2}\int_t^r\int_{\mathbb{R}}\left[\varphi(s,X_{s-},Y_{s-},V^u_{s-}+V^u_{s-}[u^*\eta(z)]_j)-\varphi(s,X_{s-},Y_{s-},V^u_{s-})\right]ds\,d\nu_j(z_j)\\
&+\sum_{l=0}^{n}\int_t^r\int_{\mathbb{R}}\Big[\varphi(s,R^l_{X^l_{s-}+h^l(X^l_{s-},Y^l_{s-},z_0)}(X_{s-}),R^l_{Y^l_{s-}-g^l(X^l_{s-},Y^l_{s-},z_0)}(Y_{s-}),V^u_{s-})\end{aligned}$$

$$- \varphi(s, X_{s-}, Y_{s-}, V_{s-}^u)\Bigg] dsdz_0 \tag{4.10}$$

We note that the last term of (4.10), can be rewritten as

$$\sum_{l=0}^{n} \int_t^r \sum_{j \neq X_{t-}^l} \lambda_{X_{t-}^l j}(Y_{s-}^l) \Bigg[\varphi(s, R_j^l(X_{s-}), R_0^l(Y_{s-}), V_{s-}^u) - \varphi(s, X_{s-}, Y_{s-}, V_{s-}^u)\Bigg] ds. \tag{4.11}$$

We denote $\int_0^t \frac{\partial}{\partial v}\varphi(s, X_{s-}, Y_{s-}, V_{s-}^u) V_t^u u^* \sigma(s, X_s) dW_s$ by M_t^1. Therefore M_t^1 is a local martingale. Since φ is in C_c^∞, $\sup_t \mathbb{E}[(M_t^1)^2] < \infty$, it follows from Theorem 1.1.11(1) that M_t^1 is a martingale. Let

$$M_t^2 := \int_0^t \int_{\mathbb{R}} \Big[\varphi(s, X_{s-}, Y_{s-}, V_{s-}^u[1 + u^* \eta(z)]_j) - \varphi(s, X_{s-}, Y_{s-}, V_{s-}^u)\Big] \tilde{N}_j(ds, dz_j),$$

for $j = 1, \ldots, m_2$. Again we note that M_t^2 is a local martingale. To show, M_t^2 is a martingale, by using Theorem 1.1.11(1), it is sufficient to check that

$$\mathbb{E}\left[\int_0^t \left(V_{s-}^u \frac{\partial}{\partial v}\varphi(s, X_{s-}, Y_{s-}, V_{s-}^u)\right)^2 ds\right] < \infty,$$

which is readily follows from the previous argument. Consider another term

$$M_t^3 := \int_0^t \int_{\mathbb{R}} \Bigg[\varphi(s, R^l_{X_{s-}^l + h^l(X_{s-}^l, Y_{s-}^l, z_0)}(X_{s-}), R^l_{Y_{s-}^l - g^l(X_{s-}^l, Y_{s-}^l, z_0)}(Y_{s-}), V_{s-}^u)$$
$$- \varphi(s, X_{s-}, Y_{s-}, V_{s-}^u)\Bigg] \tilde{\wp}^l(ds, dz_0).$$

By the similar argument as above, we can show that M_t^3 is martingale. Now taking conditional expectation given $X_t = x, Y_t = y, V_t^u = v$ and denoting it by $\mathbb{E}_{t,x,y,v}$ on both the sides of (4.10) and using (4.11), we have

$$\begin{aligned}
&\mathbb{E}_{r,x,y,v}\left[\varphi(t, X_r, Y_r, V_r^u)\right] \\
&= \varphi(t, X_t, Y_t, V_t^u) + \mathbb{E}_{t,x,y,v}\left[\int_t^r \frac{\partial}{\partial s}\varphi(s, X_{s-}, Y_{s-}, V_{s-}^u) ds\right] \\
&+ \mathbb{E}_{t,x,y,v}\left[\int_t^r \frac{\partial}{\partial y}\varphi(s, X_{s-}, Y_{s-}, V_{s-}^u) ds\right] \\
&+ \mathbb{E}_{t,x,y,v}\left[\int_t^r \frac{\partial}{\partial v}\varphi(s, X_{s-}, Y_{s-}, V_{s-}^u) V_t^u (\, r(t, X_t) + b(t, X_t) u_t)\, dt\right] \\
&+ \mathbb{E}_{t,x,y,v}\left[\frac{1}{2}\int_t^r \frac{\partial^2}{\partial v^2}\varphi(s, X_{s-}, Y_{s-}, V_{s-}^u) u^* a(s, X_s) u\, ds\right]
\end{aligned}$$

$$+\mathbb{E}_{t,x,y,v}\left[\sum_{j=1}^{m_2}\int_t^r\int_{\mathbb{R}}\Big[\varphi(s,X_{s-},Y_{s-},V_{s-}^u+V_{s-}^u[u^*\eta(z)]_j)-\varphi(s,X_{s-},Y_{s-},V_{s-}^u)\Big]\,ds\,d\nu_j(z_j)\right]$$
$$+\mathbb{E}_{t,x,y,v}\left[\sum_{l=0}^{n}\int_t^r\sum_{j\neq X_{t-}^l}\lambda_{X_{t-}^l j}(Y_{s-}^l)\Big[\varphi(s,R_j^l(X_{s-}),R_0^l(Y_{s-}),V_{s-}^u)-\varphi(s,X_{s-},Y_{s-},V_{s-}^u)\Big]\,ds\right] \tag{4.12}$$

Therefore from the Definition of Markov generator,

$$\mathscr{A}^u\varphi(t,x,y,v)=\lim_{r\to t}\frac{\mathbb{E}_{t,x,y,v}\left[\varphi(r,X_r,Y_r,V_r^u)\right]-\varphi(t,X_t,Y_t,V_t^u)}{r}.$$

Using (4.12) and strong Markov property of (X_t,Y_t,V_t^u), we have the desired result. □

In the previous proposition, we use the class of smooth and compactly supported functions. We shall attempt to find the generator in a larger class of functions. In view of (4.12), we introduce a new class of functions $\mathcal{V}$ by

$$\mathcal{V}:=\{\psi\in C\left((0,\infty)\right)\mid\sup_{v\in(0,\infty)}|v^{\frac{\theta}{2}}\psi(v)|<\infty\}.$$

We define a linear operator

$$D_{t,y}g(t,y):=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\{g(t+\varepsilon,y+\varepsilon\mathbb{1})-g(t,y)\}, \tag{4.13}$$

where $\mathrm{dom}(D_{t,y})$, the domain of $D_{t,y}$, contains all measurable functions g on $[0,T]\times[0,T]$ such that above limit exists for every $(t,y)\in[0,T]\times[0,T]$. We shall define a new class of functions $\mathscr{G}$.

Definition 4.2.9. *Let* $\mathscr{G}\subset\{\varphi:\mathscr{D}\times(0,\infty)\to\mathbb{R}\}$ *be such that for every* $\varphi\in\mathscr{G}$ *the following hold*

(i) $\varphi(t,x,y,v)$ *is twice continuously differentiable with respect to* $v\in(0,\infty)$ *for all* $t\in(0,T),x\in\mathcal{X}^{n+1},y\in(0,t)^{n+1}$ *and* φ *is in dom*$(D_{t,y})$ *for each v,x,*

(ii) for fixed $(t,x,y)\in\mathscr{D}$, $\varphi(t,x,y,\cdot)\in\mathcal{V}$,

(iii) for each (t,x,y), $v\mapsto v\frac{\partial\varphi}{\partial v}$ *is in* $\mathcal{V}$.

In the similar line of proof of Proposition 4.2.8, we can prove the following.

Proposition 4.2.10. *Let* $u=\{u_t\}_{t\in[0,T]}$ *be an admissible control and* (X_t,Y_t) *and* V_t^u *be as in* (3.3)*and* (4.5)*. Let* $\varphi\in\mathscr{G}$*, then*

$$\mathscr{A}^u\varphi(t,x,y,v)$$

$$= \frac{\partial}{\partial t}\varphi(t,x,y,v) + \frac{\partial}{\partial y}\varphi(t,x,y,v) + v\left[r(t,x) + b(t,x)\ u\right]\frac{\partial}{\partial v}\varphi(t,x,y,v)$$
$$+ \frac{1}{2}v^2\left[u^* a(t,x) u\right]\frac{\partial^2}{\partial v^2}\varphi(t,x,y,v)$$
$$+ \sum_{j=1}^{m_2}\int_{\mathbb{R}}\left[\varphi\left(t,x,y,v\left(1+[u^*\eta(z)]_j\right)\right) - \varphi(t,x,y,v)\right]\nu_j(dz_j)$$
$$+ \sum_{l=0}^{n}\sum_{j\neq x^l}\lambda^l_{x^l j}(y^l)\left[\varphi(t,R^l_j x, R^l_0 y, v) - \varphi(t,x,y,v)\right],$$

where the linear operator R^l_z *is given by* $R^l_z x := x + (z - x^l)e_l$, $l = 0,\dots,n$, $z \in \mathbb{R}$ *and* $\{e_l : l = 0,\dots,n\}$ *is the standard basis of* $\mathbb{R}^{n+1}$.

4.2.4 Optimal Control Problem

In this chapter we consider a risk sensitive optimization criterion of terminal portfolio wealth corresponding to a portfolio u, that is given by

$$J^{u,T}_\theta(x,y,v) := -\left(\frac{2}{\theta}\right)\ \ln\mathbb{E}\left[\exp\left(-\frac{\theta}{2}\ln\left(V^u_T\right)\right)\ \middle|\ X_0 = x, Y_0 = y, V^u_0 = v\right]$$
$$= -\left(\frac{2}{\theta}\right)\ \ln\mathbb{E}\left[(V^u_T)^{-\frac{\theta}{2}}\ \middle|\ X_0 = x, Y_0 = y, V^u_0 = v\right],$$

which is to be maximized over all admissible portfolio strategies with constant risk aversion parameter $\theta > 0$. Since logarithm is increasing, it suffices to consider the following cost function

$$\mathbb{E}\left[(V^u_T)^{-\frac{\theta}{2}}\ \middle|\ X_0 = x, Y_0 = y, V^u_0 = v\right],$$

which is to be minimized.

For all $(t,x,y,v) \in \mathscr{D}\times(0,\infty)$, let

$$\left.\begin{aligned} \tilde{J}^{u,T}_\theta(t,x,y,v) &:= \mathbb{E}\left[(V^u_T)^{-\frac{\theta}{2}}\ \middle|\ X_t = x, Y_t = y, V^u_t = v\right], \\ \varphi_\theta(t,x,y,v) &:= \inf_u \tilde{J}^{u,T}_\theta(t,x,y,v), \end{aligned}\right\} \tag{4.14}$$

where infimum is taken over all admissible strategies as in Definition 4.2.4. Hence, φ_θ represents the optimal cost.

Let $u = \{u_t\}_{t\in[0,T]}$ be an admissible strategy such that it has the following form $u_t := \tilde{u}(t, X_t, Y_t, V_t)$ for some measurable $\tilde{u} : \mathscr{D}\times(0,\infty) \to \mathbb{A}_1$. We call such controls as Markov feedback control. Then the augmented process $\{(X_t, Y_t, V^u_t)\}_{t\in[0,T]}$ is Markov where, X_t, Y_t, V^u_t are as in (4.1), (4.1), (4.5). We note that for any measurable

$\tilde{u} : \mathscr{D} \times (0, \infty) \to \mathbb{A}_1$, the equation (4.5) may not have a strong solution. However, we will show the existence of a Markov feedback control which is optimal and under which (4.5) has an a.s. unique strong solution. For a given $u \in \mathbb{A}_1$, by abuse of notation, we write $\mathscr{A}^u$, when $\tilde{u}(t, x, y, v) = u$ for all t, x, y, v. We consider the following equation

$$\inf_{u \in \mathbb{A}_1} \mathscr{A}^u \varphi(t, x, y, v) = 0, \tag{4.15}$$

with the terminal condition

$$\varphi(T, x, y, v) = v^{-\frac{\theta}{2}}, \ x \in \mathfrak{X}^{n+1}, \ y \in [0, T]^{n+1}, \quad v > 0. \tag{4.16}$$

We clarify below, what we mean by a classical solution to the problem (4.15)-(4.16).

Definition 4.2.11. *We say $\varphi : \mathscr{D} \times (0, \infty) \to \mathbb{R}$ is a classical solution to* (4.15)-(4.16) *if $\varphi \in \mathscr{G}$ and for all $(t, x, y, v) \in \mathscr{D} \times (0, \infty)$, φ satisfies* (4.15)-(4.16).

4.3 Hamilton-Jacobi-Bellman Equation

We look for a solution to (4.15)-(4.16) of the form

$$\varphi(t, x, y, v) = v^{-\frac{\theta}{2}} \psi(t, x, y), \tag{4.17}$$

where $\psi \in \mathrm{dom}(D_{t,y})$. Clearly, the left hand side of (4.17) is in class $\mathscr{G}$. We will establish the following result in first two subsections.

Theorem 4.3.1. *The Cauchy problem* (4.15)-(4.16) *has a unique classical solution, φ_M, of the form* (4.17).

Substitution of (4.17) into (4.15), yields

$$D_{t,y}\psi(t, x, y) + \sum_{l} \sum_{j \neq x^l} \lambda^l_{x^l j}(y^l) \Big[\psi(t, R^l_j x, R^l_0 y) - \psi(t, x, y)\Big] + h_\theta(t, x)\psi(t, x, y) = 0, \tag{4.18}$$

for each $(t, x, y) \in \mathscr{D}$ with the condition

$$\psi(T, x, y) = 1, \tag{4.19}$$

where the map $h_\theta : [0, T] \times \mathfrak{X}^{n+1} \to \mathbb{R}$ is given by

$$h_\theta(t, x) := \inf_{u \in \mathbb{A}_1} \left[g_\theta(t, x, u)\right], \tag{4.20}$$

the infimum of a family of continuous functions

$$g_\theta(t, x, u) := \left(-\frac{\theta}{2}\right) [r(t, x) + b(t, x)\ u] + \frac{1}{2} \left(-\frac{\theta}{2}\right) \left(-\frac{\theta}{2} - 1\right) [u^*\ a(t, x)\ u]$$

$$+\sum_{j=1}^{m_2}\int_{\mathbb{R}}\left((1+[u^*\eta(z)]_j)^{\left(-\frac{\theta}{2}\right)}-1\right)\nu_j(dz_j).$$

It is important to note that the linear first order equation (4.18) is nonlocal due to the presence of the term $\psi(t,R_j^l x,R_0^l y)$ in the equation. It implies that $D_{t,y}\psi(t,x,y)$ depends on the value of ψ at the point $(t,\cdot,R_0^l y)$, which does not lie in the neighbourhood of $(t,\cdot,y)$. We now define a classical solution to (4.18)-(4.19) below.

Definition 4.3.2. *We say $\varphi:\mathscr{D}\to\mathbb{R}$ is a classical solution to* (4.18)-(4.19) *if $\varphi\in$ dom$(D_{t,y})$ and for all $(t,x,y)\in\mathscr{D}$, φ satisfies* (4.18)-(4.19).

Remark 4.3.3. *It is interesting to note that other than the terminal condition* (4.19)*, no additional boundary conditions are imposed. The remaining parts of the boundary is $\bar{\mathscr{D}}\cap\{(t,x,y)|y^l=0,x\in\mathcal{X}^{n+1},t\in[0,T]\}$. We note from* (4.1) *that, $0\le Y_t^l$, for all $t\in[0,T]$. Hence $\{Y_t\}_{t\ge0}$ does not cross the boundary. Thus the value of solution on the boundary is obtained from the terminal condition* (4.19).

Theorem 4.3.4. *The Cauchy problem* (4.18)-(4.19) *has a unique classical solution in $C_b(\bar{\mathscr{D}})$.*

Remark 4.3.5. *Note that Theorem 4.3.1 may be treated as a corollary of Theorem 4.3.4 in view of the substitution (4.17) and subsequent analysis. Thus it suffices to establish Theorem 4.3.4. We establish Theorem 4.3.4 in the subsection 4.3.2 via a study of an integral equation which is presented in subsection 4.3.1. The following result would be useful to establish well-posedness of* (4.18)-(4.19).

Proposition 4.3.6. *Consider the map $h_\theta:[0,T]\times\mathcal{X}^{n+1}\to\mathbb{R}$, given by,* (4.20)*. Then under Assumption 4.1(iii), we have*

(i) h_θ is continuous, negative valued, bounded below;

(ii) $H_\theta(t_1,t_2,x):=\int_{t_1}^{t_2}h_\theta(s,x)ds$ is C^1 in both t_1 and t_2 for each x;

(iii) For every (t,x), there exists a unique $u^\star(t,x)\in\mathbb{A}_1$ such that $h_\theta(t,x)=g_\theta(t,x,u^\star(t,x))$. and $u^\star:[0,T]\times\mathcal{X}^{n+1}\to\mathbb{A}_1$ is continuous in t;

(iv) $\{u^\star(t,X_t)\}_{t\ge0}$ is admissible.

Proof. (i) We recall that, $\mathbb{A}_1$, the range of portfolio includes origin. Therefore it follows from the Definition of h_θ in (4.20) that

$$h_\theta(t,x)\ \le\ g_\theta(t,x,0)=-\frac{\theta}{2}r(t,x)<0.$$

Thus h_θ is negative valued. By the continuity assumptions on r,b and a, for fixed u and each $x\in\mathcal{X}^{n+1}$, $r(t,x)$, $b(t,x)$, and $a(t,x)$ are bounded on $[0,T]$. Let $M\ge0$ be such that

$$\max_{t\in[0,T]}\{|r(t,x)|,\|b(t,x)\|,\|a(t,x)\|\}\le M.$$

We also observe that for each $u \in \mathbb{A}_1$,

$$\begin{aligned}\sum_j \int_{\mathbb{R}} ((1+[u^*\eta(z)]_j)^{-\frac{\theta}{2}} - 1)\,\nu_j(dz_j) &\geq -\sum_j \int_{\mathbb{R}} \nu_j(dz_j) \\ &= -\sum_j \nu_j(\mathbb{R}) > -\infty,\end{aligned}$$

using the finiteness of the measure ν_j. Also, Assumption 4.1(iii) gives $u^*a(t,x)u \geq \delta_1\|u\|^2$. Hence by using the above mentioned bounds, we can write, $g_\theta(t,x,u) \geq \bar{g}_\theta(u)$, where

$$\bar{g}_\theta(u) = \left(-\frac{\theta}{2}(M + M\|u\|) + \frac{\theta}{4}(1+\frac{\theta}{2})\delta_1\|u\|^2 - \sum_j \nu_j(\mathbb{R})\right).$$

Since $\bar{g}_\theta(u)$ is independent of t and $\uparrow \infty$ as $\|u\| \uparrow \infty$, $h_\theta(t,x)$ is bounded below. Now we will show that for fixed t and x, $g_\theta(t,x,u)$ is a strictly convex function of variable $u \in \mathbb{A}_1$. To see this first we take the derivative of $g_\theta(t,x,u)$ in u_p.

$$\frac{\partial g_\theta}{\partial u^p} = -\frac{\theta}{2} b^p(t,x) + \frac{\theta}{4}\left(\frac{\theta}{2}+1\right)\sum_j a_{pj}u_j - \sum_j \int_{\mathbb{R}} \frac{\theta}{2}\eta_{pj}(z_j)\,(1+[u^*\eta(z)]_j)^{-\frac{\theta}{2}-1}\ \nu_j(dz_j). \tag{4.21}$$

For fixed t and x, let H denote the Hessian matrix for g_θ. Then we can compute (p,q)-th element of H by using (4.21), denote it by H_{pq},

$$\frac{1}{2}\frac{\theta}{2}\left(\frac{\theta}{2}+1\right)a_{pq}(t,x) + \sum_j \int_{\mathbb{R}} \frac{\theta}{2}\left(\frac{\theta}{2}+1\right)\eta_{pj}(z_j)\eta_{qj}(z_j)\,(1+[u^*\eta(z)]_j)^{-\frac{\theta}{2}-2}\ \nu_j(dz_j).$$

Since u is in $\mathbb{A}_1$, $(1+[u^*\eta(z)]_j$ is bounded below by a positive δ. Now, by using Assumption 4.1(iii) we shall show that, there exists $m > 0$ such that $H - mI$ is a positive definite matrix.

$$\begin{aligned}\xi^*H\xi &= \sum_{p,q} H_{pq}\xi_p\xi_q \\ &= \frac{1}{2}\frac{\theta}{2}\left(\frac{\theta}{2}+1\right)\sum_{p,q} a_{pq}(t,x)\xi_p\xi_q \\ &+ \sum_{p,q}\sum_j \int_{\mathbb{R}} \frac{\theta}{2}\left(\frac{\theta}{2}+1\right)\eta_{pj}(z_j)\eta_{qj}(z_j)\,(1+[u^*\eta(z)]_j)^{-\frac{\theta}{2}-2}\,\xi_p\xi_q\,\nu_j(dz_j) \\ &\geq \frac{\theta}{4}\left(\frac{\theta}{2}+1\right)\delta_1\|\xi\|^2 + \frac{\theta}{2}\left(\frac{\theta}{2}+1\right)\delta^{-\frac{\theta}{2}-2}\sum_{p,q}\sum_j \int_{\mathbb{R}} \xi_p\eta_{pj}(z_j)\xi_q\eta_{qj}(z_j)\,\nu_j(dz_j) \\ &= \frac{\theta}{4}\left(\frac{\theta}{2}+1\right)\delta_1\|\xi\|^2 + \frac{\theta}{2}\left(\frac{\theta}{2}+1\right)\delta^{-\frac{\theta}{2}-2}\sum_j \int_{\mathbb{R}} \sum_{p,q} \xi_p\eta_{pj}(z_j)\xi_q\eta_{qj}(z_j)\,\nu_j(dz_j)\end{aligned}$$

$$
\begin{aligned}
&= \frac{\theta}{4}\left(\frac{\theta}{2}+1\right)\delta_1\|\xi\|^2 + \frac{\theta}{2}\left(\frac{\theta}{2}+1\right)\delta^{-\frac{\theta}{2}-2}\sum_j \int_{\mathbb{R}} \left(\xi^* \cdot \eta_{\cdot j}(z_j)\right)^2 \nu_j(dz_j) \\
&= \frac{\theta}{4}\left(\frac{\theta}{2}+1\right)\delta_1\|\xi\|^2.
\end{aligned}
$$

Therefore we can find an $m > 0$ such that $\xi^* H\xi \geq m\|\xi\|^2$, i.e $H - mI$ is positive definite. Therefore $g_\theta(t, x, u)$ is a the strictly convex function on variable u. Therefore $\mathbb{A}_2 := \mathbb{A}_1 \bigcap \bar{g}_\theta^{-1}((-\infty, 1])$ is a non-empty convex compact set. Hence, $(t, x) \twoheadrightarrow \mathbb{A}_2$ is a compact-valued correspondence. Since h_θ is negative, from (4.20), we can write

$$
h_\theta(t, x) \quad = \quad \inf\{g_\theta(t, x, u)|u \in \mathbb{A}_2\}.
$$

We also note that $(t, x, u) \mapsto g_\theta(t, x, u)$ is jointly continuous. Since $(t, x) \twoheadrightarrow \mathbb{A}_2$ is continuous, then it follows from the Maximum Theorem 1.3.3 that $h_\theta(t, x)$ is continuous with respect to (t, x). Hence (i) is proved.

(ii) It follows from continuity of $h_\theta(t, x)$.

(iii) The set of minimizers is defined by

$$
u^\star(t, x) = \operatorname{argmin}\{g_\theta(t, x, u)|u \in \mathbb{A}_2\}.
$$

Again by using Theorem 1.3.3, $(t, x) \twoheadrightarrow u^\star(t, x)$ is upper semi-continuous. Since $g_\theta(t, x, u)$ is strictly convex in u, for each $t \in [0, T]$ and $x \in \mathfrak{X}^{n+1}$ there exist only one element in $u^\star(t, x)$. By abuse of notation, we denote that element by $u^\star(t, x)$ itself. Since a single-valued upper semi-continuous correspondence is continuous, $u^\star(t, x)$ is a continuous function.

(iv) Since $u^\star$ is continuous in t, there exists a positive constant M such that $\|u^\star(t, X_t(\omega))\| < M \quad \forall t \in [0, T], \omega \in \Omega$. Since $u^\star$ does not depend on v, the conditions (1.7) and (1.8) of Theorem 1.1.26 are satisfied. Again since $u^\star$ is bounded all growth conditions are also satisfied. Therefore Definition 4.2.4(ii) satisfied and this completes the proof. □

4.3.1 An Equivalent Volterra Integral equation

In order to study (4.18)-(4.19) we first introduce some notations from Chapter 3.

Notation 4.1.

1. *p^l_{ij} := probability the X^l jumps from state i to state j, defined as in (3.4).*

2. *$\tau^l(t)$:= time period from time t after which the l-th component of X_t would have a first transition.*

3. *We denote the conditional c.d.f of $\tau^l(t)$ given $X^l_t = i$ and $Y^l_t = \bar{y}$ as $F_{\tau^l}(\cdot|i, \bar{y})$, and the conditional p.d.f of $\tau^l(t)$ given $X^l_t = i$ and $Y^l_t = \bar{y}$ as $f_{\tau^l}(\cdot|i, \bar{y})$.*

4. *Let $\ell(t)$ be the component of X_t, where the subsequent transition happens.*

5. *Let $F_{\tau^l|l}(\cdot|x,y)$ and $f_{\tau^l|l}(\cdot|x,y)$ be the conditional c.d.f and conditional p.d.f of $\tau^l(t)$ given $X_t = x, Y_t = y$ and $\ell(t) = l$.*

Using the above notations we introduce the following integral equation on $\mathscr{D}$

$$\psi(t,x,y) = \sum_{l=0}^{n} \mathbb{P}_{t,x,y}(\ell(t)=l)\Big[(1 - F_{\tau^l|l}(T-t|x,y))e^{H_\theta(t,T,x)} + \int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j\neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1}) f_{\tau^l|l}(r|x,y)\Big]dr. \quad (4.22)$$

Remark 4.3.7. *Equation* (4.22) *is a Volterra integral equation of second kind. We note that the boundary of $\bar{\mathscr{D}}$ has many facets. For $t = T$, we directly obtain from* (4.22), *$\psi(T,x,y) = 1$. Hence no additional terminal conditions are required. Although the values of ψ in facets $\bar{\mathscr{D}} \cap \{(t,x,y)|y^l \in \{0,t\}, x \in \mathcal{X}^{n+1}, t \in [0,T]\}$ are not directly followed but can be obtained by solving the integral equation on the facets.*

Now we shall study the regularity properties of (4.22).

Proposition 4.3.8. *The integral equation* (4.22) *(i) has a unique solution in $C(\bar{\mathscr{D}})$, and (ii) the solution is in the* $\mathrm{dom}(D_{t,y})$.

Proof. (i) We first observe that the solution to the integral equation (4.22) is a fixed point of the operator A, where

$$A\psi(t,x,y) := \sum_{l=0}^{n} \mathbb{P}_{t,x,y}(\ell(t)=l)\Big[(1 - F_{\tau^l|l}(T-t|x,y))e^{H_\theta(t,T,x)} + \int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j\neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1})) f_{\tau^l|l}(r|x,y)\Big]dr.$$

It is easy to check that for each $\psi \in C(\bar{\mathscr{D}}), A\psi : \bar{\mathscr{D}} \to (0,\infty)$ is continuous. Now since $h_\theta < 0$ (showed in Proposition 4.3.6(i)),

$$\begin{aligned}
&\|A\psi - A\tilde{\psi}\| \\
&= \sup_{\mathscr{D}} |A\psi - A\tilde{\psi}| \\
&= |\sum_{l=0}^{n} \mathbb{P}_{t,x,y}(\ell(t)=l)[\int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j\neq x^l} p^l_{x^l j}(y^l + r) \\
&\quad \times [\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1})) - \tilde{\psi}(t+v, R^l_j x, R^l_0(y + r\mathbb{1}))] f_{\tau^l|l}(r|x,y)dr]| \\
&\leq \sum_{i=1}^{n} \mathbb{P}_{t,x,y}(\ell(t)=l) \int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j\neq x^l} p^l_{x^l j}(y^l + r) f_{\tau^l|l}(r|x,y)dr \|\psi - \tilde{\psi}\|
\end{aligned}$$

$$< \quad K_1\|\psi - \tilde{\psi}\|,$$

where $K_1 := \sum_{l=0}^{n} P_{t,x,y}(\ell(t) = l) \int_0^{T-t} f_{\tau^l|l}(r|x,y)dr$. Since $F^l(\bar{y})$ is strictly less than 1, $F_{\tau^l|l}(r|x,y) < 1$, $\forall r > 0$. Hence $K_1 < 1$. Therefore, A is a contraction. Thus a direct application of Banach fixed point theorem ensures the existence and uniqueness of the solution to (4.22).
(ii) We denote the unique solution as ψ. Next we show that $\psi \in \text{dom}(D_{t,y})$. To this end, it is sufficient to show that $A : C(\mathscr{D}) \to \text{dom}(D_{t,y})$. The first term of $A\psi$ is in $\text{dom}(D_{t,y})$, which follows from Lemma 3.3.4 (iv) and Proposition 4.3.6 (ii). Now to show that the remaining term

$$\beta_l(t,x,y) := \int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j \neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1}))f_{\tau^l|l}(r|x,y)]dr,$$

is also in the $\text{dom}(D_{t,y})$ for any $\psi \in C(\mathscr{D})$, we need to check if the following limit

$$\begin{aligned}
\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\Big[\int_0^{T-t-\varepsilon} & e^{H_\theta(t+\varepsilon,t+r+\varepsilon,x)} \sum_{j \neq x^l} p^l_{x^l j}(y^l + r + \varepsilon)\psi(t+r+\varepsilon, R^l_j x, R^l_0(y + (r+\varepsilon)\mathbb{1})) \\
&\times f_{\tau^l|l}(r|x,y+\varepsilon))dr - \int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j \neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1})) \\
&\times f_{\tau^l|l}(r|x,y)dr\Big],
\end{aligned}$$

exists and, the limit is continuous in $\mathscr{D}$. If the limit exists the limit value is clearly $D_{t,y}\beta_l(t,x,y)$. By a suitable substitution of variables in the integral, the expression in the above limit can be rewritten, as

$$\begin{aligned}
&\frac{1}{\varepsilon}\Big[\int_\varepsilon^{T-t} e^{H_\theta(t+\varepsilon,t+r,x)} \sum_{j \neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1}))f_{\tau^l|l}(r-\varepsilon|x,y+\varepsilon)dr \\
&\quad - \int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j \neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1}))f_{\tau^l|l}(r|x,y)dr \\
&= \int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j \neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1})) \\
&\quad \times \frac{1}{\varepsilon}\Big(e^{-H_\theta(t,t+\varepsilon,x)} f_{\tau^l|l}(r-\varepsilon|x,y+\varepsilon) - f_{\tau^l|l}(r|x,y)\Big)dr - \frac{1}{\varepsilon}\int_0^{\varepsilon} e^{H_\theta(t+\varepsilon,t+r,x)} \times \\
&\quad \sum_{jx^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1}))f_{\tau^l|l}(r-\varepsilon|x,y+\varepsilon)dr. \qquad (4.23)
\end{aligned}$$

By Lemma 3.3.4 (iv), $f_{\tau^l|l}(T-t|x,y)$ is in $\text{dom}(D_{t,y})$. Thus $D_{t,y}f_{\tau^l|l}(T-t|x,y)$ is bounded on $[0, T-t]$ by a positive constant K_2. Hence by mean value theorem on $f_{\tau^l|l}(T-t|x,y)$, the integrand of the first integral of (4.23) is uniformly bounded. Therefore, using the bounded convergence theorem, the integral converges as $\varepsilon \to 0$. The second integral of (4.23) converges as the integrand is continuous at $r = 0$. Now we compute

$$\begin{aligned}
&D_{t,y}\beta_l(t,x,y) \\
&= \int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j\neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1}))\Big(\frac{d}{dw}e^{-H_\theta(t,t+w,x)}\Big|_{w=0} f_{\tau^l|l}(r|x,y) \\
&\quad + \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[f_{\tau^l|l}(r-\varepsilon|x,y+\varepsilon) - f_{\tau^l|l}(r|x,y+\varepsilon) + f_{\tau^l|l}(r|x,y+\varepsilon) - f_{\tau^l|l}(r|x,y)\right]\Big)dr \\
&\quad - \sum_{j\neq x^l} p^l_{x^l j}(y^l)\psi(t, R^l_j x, R^l_0 y) f_{\tau^l|l}(0|x,y),
\end{aligned}$$

using Lemma 3.3.4 (iii). Therefore $D_{t,y}\beta_l(t,x,y)$ can be rewritten using Lemma 3.3.5(v) as

$$\begin{aligned}
&\int_0^{T-t} e^{H_\theta(t,t+r,x)} \sum_{j\neq x^l} p^l_{x^l j}(y^l + r)\psi(t+r, R^l_j x, R^l_0(y + r\mathbb{1}))(-h_\theta(t,x) + f_{\tau^l|l}(0|x,y))\times \\
&f_{\tau^l|l}(r|x,y)dr - \sum_{j\neq x^l} p^l_{x^l j}(y^l)\psi(t, R^l_j x, R^l_0 y) f_{\tau^l|l}(0|x,y) \\
&= -h_\theta(t,x) + f_{\tau^l|l}(0|x,y)]\beta_l(t,x,y) - \sum_{j\neq x^l} p^l_{x^l j}(y^l)\psi(t, R^l_j x, R^l_0 y) f_{\tau^l|l}(0|x,y). \qquad (4.24)
\end{aligned}$$

Clearly (4.24) is in $C(\mathscr{D})$. Hence $\beta_l(t,x,y)$ is in the dom$(D_{t,y})$. Hence the right hand side of (4.22) is in the dom$(D_{t,y})$ for any $\psi \in C(\bar{\mathscr{D}})$. Thus (ii) holds. □

4.3.2 The linear first order equation

Proposition 4.3.9. *The unique solution to* (4.22) *also solves the initial value problem* (4.18)-(4.19).

Proof. Let ψ be the solutions of the integral equation (4.22). Then by substituting $t = T$ in (4.22), (4.19) follows. Using the results from the proof of Lemma 3.3.4, Proposition 4.3.8, Lemma 3.3.4(iv) and (4.24), we have

$$\begin{aligned}
D_{t,y}\psi(t,x,y) = &\sum_{l=0}^{n}\Big[\sum_r f_{\tau^r}(0|x^r,y^r)P_{t,x,y}(\ell(t)=l) - f_{\tau^l}(0|x^l,y^l)\Big][1 - F_{\tau^l|l}(T-t|x,y)]e^{H_\theta(t,T,x)} \\
&- \sum_{l=0}^{n} P_{t,x,y}(\ell(t)=l)\Big[f_{\tau^l|l}(0|x,y)(F_{\tau^l|l}(v|x,y) - 1)\Big] \\
&\times e^{H_\theta(t,T,x)} - h_\theta(t,x)\sum_{l=0}^{n} P_{t,x,y}(\ell(t)=l)[1 - F_{\tau^l|l}(T-t|x,y) \\
&\times e^{H_\theta(t,T,x)} + \sum_{l=0}^{n}\Big[\sum_r f_{\tau^r}(0|x^r,y^r)P_{t,x,y}(\ell(t)=l) - f_{\tau^l}(0|x^l,y^l)\Big]\beta_l(t,x,y)
\end{aligned}$$

$$+\sum_{l=0}^{n} P_{t,x,y}(\ell(t)=l)\Big(-h_\theta(t,x)+f_{\tau^l|l}(0|x,y)\Big)\beta_l(t,x,y)$$
$$-\sum_{j\neq x^l} p^l_{x^l j}(y^l)\psi(t,R^l_j x,R^l_0 y)f_{\tau^l|l}(0|x,y)\Big).$$

Using the equality in Lemma 3.3.4 (v), the right hand side of above equation can be rewritten as

$$\sum_l \frac{f^l(y^l|x^l)}{1-F^l(y^l|x^l)}\Big[\psi(t,x,y)-\sum_{j\neq x^l} p^l_{x^l j}(y^l)\psi(t,R^l_j x,R^l_0 y)\Big]-h_\theta(t,x)\psi(t,x,y)$$
$$=-\sum_l\sum_{j\neq x^l}\lambda^l_{x^l j}(y^l)\Big[\psi(t,R^l_j x,R^l_0 y)-\psi(t,x,y)\Big]-h_\theta(t,x)\psi(t,x,y).$$

Hence ψ satisfies (4.18). □

Proposition 4.3.10. *Let ψ be a bounded classical solution to* (4.18)-(4.19)*. Then ψ solves the integral equation* (4.22)*.*

Proof. If the PDE has a classical solution ψ, then ψ is also in the domain of $\mathcal{A}$, where $\mathcal{A}$ is the infinitesimal generator of (t, X_t, Y_t). Then we have from the PDE

$$\mathcal{A}\psi + h_\theta(t,x)\psi(t,x,y) = 0. \tag{4.25}$$

Consider

$$N_t := e^{\int_0^t h_\theta(s,X_s)ds}\psi(t,X_t,Y_t).$$

Then by Itô's formula,

$$dN_t = h_\theta(t,X_t)e^{\int_0^t h_\theta(s,X_s)ds}\psi(t,X_t,Y_t)dt + e^{\int_0^t h_\theta(s,X_s)ds}(\mathcal{A}\psi dt + dM^{(1)}_t),$$

where $M^{(1)}_t$ is a local martingale with respect to $\mathcal{F}_t$, the usual filtration generated by (X_t, Y_t). Thus from (4.25) N_t is a local martingale. From definition of N_t,

$$\sup_{[0,T]} N_t < \|\psi\| e^{\|h_\theta\|T} \text{ a.s.}$$

Thus N_t is a martingale. Therefore by using (4.19), we obtain

$$\psi(t,X_t,Y_t) = e^{\int_0^t -h_\theta(s,X_s)ds}N_t = \mathbb{E}[e^{\int_t^T h_\theta(s,X_s)ds}|\mathcal{F}_t].$$

Hence using the Markov property of (X_t, Y_t) irreducibility of probability matrix, we have

$$\psi(t,x,y) = \mathbb{E}_{t,x,y}[e^{\int_t^T h_\theta(s,X_s)ds}],\ \forall (t,x,y)\in\bar{\mathscr{D}}, \tag{4.26}$$

where $\mathbb{E}_{t,x,y}[\cdot] = \mathbb{E}[\cdot|X_t = x, Y_t = y]$. Let $\ell(t)$ be the component of X_t where the transition happens. By conditioning on $\ell(t)$ and using tower property (Theorem 1.1.4)

$$\begin{aligned}\psi(t,x,y) &= \mathbb{E}_{t,x,y}[\mathbb{E}_{t,x,y}[e^{[\int_t^T h_\theta(s,X_s)ds]}|\ell(t)]] \\ &= \sum_{l=0}^{n} \mathbb{P}_{t,x,y}(\ell(t)=l)\mathbb{E}_{t,x,y}[e^{[\int_t^T h_\theta(s,X_s)ds]}|\ell(t)=l],\end{aligned} \tag{4.27}$$

where $\mathbb{P}_{t,x,y}[\cdot] = \mathbb{P}[\cdot|X_t = x, Y_t = y]$. Let $\tau^l(t)$ be the time period from time t after which X^l would have a transition. By conditioning on $\tau^l(t)$ the equation (4.27) can be rewritten as

$$\begin{aligned}&\mathbb{E}_{t,x,y}[e^{\int_t^T h_\theta(s,X_s)ds}|\ell(t)=l] \\ =\ &\mathbb{E}_{t,x,y}[\mathbb{E}_{t,x,y}[e^{\int_t^T h_\theta(s,X_s)ds}|\ell(t)=l,\tau^l(t)]|\ell(t)=l] \\ =\ &P_{t,x,y}(\tau^l(t) > T-t|\ell(t)=l)e^{\int_t^T h_\theta(s,x)ds} \\ &+\int_0^{T-t} \mathbb{E}_{t,x,y}[e^{\int_t^T h_\theta(s,X_s)ds}|\ell(t)=l,\tau^l(t)=r]f_{\tau^l|l}(r|x,y)dr,\end{aligned}$$

where $f_{\tau^l|l}(\cdot|x,y)$ be the conditional p.d.f of $\tau^l(t)$ given $X_t = x, Y_t = y$ and $\ell(t) = l$ and its expression is given in Lemma 3.3.4(iv). Let p_{ij}^l be the probability of transition of X^l from state i to j. Since X_s is constant on $[t, t+r)$ provided $\ell(t) = l, \tau^l(t) = r$, above is equal to

$$\begin{aligned}&[1 - F_{\tau^l|l}(T-t|x,y)]e^{H_\theta(t,T,x)} + \int_0^{T-t} e^{H_\theta(t,t+r,x)} \\ &\times\mathbb{E}_{t,x,y}[\mathbb{E}_{t,x,y}[e^{\int_{t+r}^T h_\theta(s,X_s)ds}|X_{t+r}^l,\ell(t)=l,\tau^l=r]|\ell(t)=l,\tau^l=r]f_{\tau^l|l}(r|x,y)dr \\ =\ &[1 - F_{\tau^l|l}(T-t|x,y)]e^{H_\theta(t,T,x)} \\ &+\int_0^{T-t} e^{H_\theta(t,t+r,x)} \times \sum_{j\neq x^l} p_{x^l j}^l(y^l+r)\psi(t+r, R_j^l x, R_0^l(y+r\mathbb{1}))f_{\tau^l|l}(r|x,y)dr,\end{aligned}$$

where $\mathbb{1} \in \mathbb{R}^{(n+1)}$ with each component 1. Therefore the desired result follows.

□

Proof of Theorem 4.3.4. The result follows from Proposition 4.3.8, Proposition 4.3.9, and Proposition 4.3.10. □

4.3.3 Optimal portfolio and verification theorem

Now we are in a position of deriving the expression of optimal portfolio value under risk sensitive criterion. The optimal value is given by

$$\tilde{\varphi}_\theta(v,x,y) := \sup_u J_\theta^{u,T}(v,x,y)$$

$$
\begin{aligned}
&= -\frac{2}{\theta}\ln(\varphi_\theta(0,x,y,v)) \\
&= \ln(v) - \frac{2}{\theta}\ln(\psi(0,x,y)),
\end{aligned} \tag{4.28}
$$

where the existence and uniqueness of the classical solution to (4.18) - (4.19) follows from Theorem 4.3.4.

Remark 4.3.11. *We note that the study of* (4.18)-(4.19) *becomes much simpler if the coefficients* r, μ, σ *are independent of time* t*. For time homogeneous case, Proposition* 4.3.6 *is immediate. Furthermore, the proof of Theorem* 4.3.4 *does not need the results given in Proposition* 4.3.8*, Proposition* 4.3.9*, and Proposition* 4.3.10*. Indeed Theorem* 4.3.4 *can directly be proved by noting the smoothness of terminal condition.*

We conclude this section with a proof of the verification theorem for optimal control problem (4.14).

Theorem 4.3.12. *Let* φ_M *be the classical solution to* (4.15)-(4.16) *as in* (4.17) *then*

(i) $\varphi_M(t,x,y,v) \leq \tilde{J}_\theta^{\bar{u},T}(t,x,y,v)$ *for every Markov feedback control* $\bar{u}$*.*

(ii) Let $\bar{u}^\star := \{u^\star(t,X_t)\}_{t\geq 0}$ *be as in Proposition* 4.3.6*(iv), then*

$$
\varphi_M(t,x,y,v) = \tilde{J}_\theta^{\bar{u}^\star,T}(t,x,y,v)
$$

.

Proof. (i) Consider an admissible Markov feedback control $\bar{u} := \{\bar{u}_t\}_{t\geq 0}$, where $\bar{u}_t = \tilde{u}(t,X_t,Y_t,V_t)$ and φ_M, the classical solution to (4.15)-(4.16) as in (4.17). Now by Itô's formula

$$
\begin{aligned}
&\varphi_M(s,X_s,Y_s,V_s^{\bar{u}}) - \varphi_M(t,X_t,Y_t,V_t^{\bar{u}}) - \int_t^s [\mathscr{A}^{\bar{u}}\varphi_M(r,X_r,Y_r,V_r^{\bar{u}})]dr \\
&= \sum_{j=1}^{m_1}\int_t^s \frac{\partial}{\partial v}\varphi_M(r,X_r,Y_r,V_r^{\bar{u}})V_r^{\bar{u}}[\tilde{u}(r,X_r,Y_r,V_r)^*\sigma(r,X_r)]_j dW_r^j \\
&+ \sum_{j=1}^{m_2}\int_t^s\int_{\mathbb{R}}\left[\varphi_M(r,X_r,Y_r,V_{r-}^{\bar{u}}(1+[\tilde{u}(r,X_{r-},Y_{r-},V_{r-})^*\eta(z)]_j)) - \varphi_M(r,X_r,Y_r,V_{r-}^{\bar{u}})\right]\tilde{N}_j(dr,dz_j) \\
&+ \sum_{l=0}^{n}\int_t^s\int_{\mathbb{R}}\left[\varphi_M(r,R^l_{X^l_{r-}+h^l(X^l_{r-},Y^l_{r-},z_0)}(X_{r-}),R^l_{Y^l_{r-}-g^l(X^l_{r-},Y^l_{r-},z_0)}(Y_{r-}),V_{r-}^{\bar{u}})\right. \\
&\left. - \varphi_M(r,X_{r-},Y_{r-},V_{r-}^{\bar{u}})\right]\tilde{\wp}^l(dr,dz_0).
\end{aligned} \tag{4.29}
$$

We would first show that the right hand side is an $\mathscr{F}_s$ martingale. Since $\bar{u}$ is admissible, using definition 4.2.4(iii), it is sufficient to show, the following square integrability condition

$$\mathbb{E}\int_0^s \left[V_r^{\bar{u}} \frac{\partial}{\partial v}\varphi_M(r, X_r, Y_r, V_r^{\bar{u}})\right]^2 dr < \infty,$$

to prove that the first term is a martingale. Again since $\varphi_M(t,x,y,v) = v^{-\frac{\theta}{2}}\psi(t,x,y)$. Thus using boundedness of ψ, $v\frac{\partial\varphi_M}{\partial v} = -\frac{\theta}{2}\varphi_M$, the above would follow if $\mathbb{E}\int_0^s [V_r^{\bar{u}}]^{-\theta}\, dr < \infty$, which readily follows from the Lemma 4.2.7(ii) and an application of Tonelli's Theorem. Similarly, using the admissibility of $\bar{u}$ we can show that the last two terms of (4.29) are also martingales. Taking conditional expectation on both sides of (4.29) given $X_t = x, Y_t = y, V_t^{\bar{u}} = v$ and letting $s \uparrow T$, we obtain

$$\begin{aligned} &\mathbb{E}\left[(V_T^{\bar{u}})^{\frac{\theta}{2}} | X_t = x, Y_t = y, V_t^{\bar{u}} = v\right] - \varphi_M(t,x,y,v) \\ &= \mathbb{E}\int_t^T \left[\mathscr{A}^{\tilde{u}}\varphi_M(r, X_r, Y_r, V_r^{\bar{u}})\middle| X_t = x, Y_t = y, V_t^{\bar{u}} = v\right] dr \geq 0. \end{aligned} \tag{4.30}$$

The above non-negativity follows, since φ_M is the classical solution to (4.15)-(4.16) and $\bar{u}_r \in \mathbb{A}_1$ for all r. (4.14) and (4.30) implies result (i).

(ii) The right hand side of (4.30) becomes zero by considering $\bar{u}_t = u^\star(t, X_t)$ and this completes the proof of (ii). □

Theorem 4.3.13. *Let φ, φ_M be classical solutions to* (4.15)-(4.16)*, then $\varphi_M(t,x,y,v) \geq \varphi(t,x,y,v)$, where φ_M as in* (4.17)*.*

Proof. Note that in the Proof of Theorem 4.3.12(i), to show that the RHS of (4.29) is a martingale, we have only effectively used the fact that φ_M satisfies conditions (i),(ii) and (iii) of Definition 4.2.9. Hence for any $\varphi \in \mathscr{G}$ and $\bar{u} := \{\bar{u}_t\}_{t\geq 0}$, where $\bar{u}_t = \tilde{u}(t, X_t, Y_t, V_t)$ a Markov control,

$$\varphi(s, X_s, Y_s, V_s^{\bar{u}}) - \varphi(t, X_t, Y_t, V_t^{\bar{u}}) - \int_t^s [\mathscr{A}^{\tilde{u}}\varphi(r, X_r, Y_r, V_r^{\bar{u}})]dr, \tag{4.31}$$

is an $\mathscr{F}_s$ martingale. Now consider $\bar{u}^\star$ as in Theorem 4.3.12(ii). Taking conditional expectation in (4.31), given $X_t = x, Y_t = y, V_t^{\bar{u}^\star} = v$ and letting $s \uparrow T$, we have

$$\begin{aligned} &\mathbb{E}\left[(V_T^{\bar{u}^\star})^{-\frac{\theta}{2}} | X_t = x, Y_t = y, V_t^u = v\right] - \varphi(t,x,y,v) \\ &= \mathbb{E}\int_t^T \left[\mathscr{A}^{\tilde{u}}\varphi(r, X_r, Y_r, V_r^u)\middle| X_t = x, Y_t = y, V_t^u = v\right] dr, \end{aligned}$$

using $\varphi(T, X_T, Y_T, V_T^{\bar{u}^\star}) = \left(V_T^{\bar{u}^\star}\right)^{-\frac{\theta}{2}}$. Now using RHS is nonnegative and Theorem 4.3.12(ii), we obtain $\varphi_M(t,x,y,v) \geq \varphi((t,x,y,v)$. □

Theorem 4.3.14. *Let* φ_M *be as in Theorem* 4.3.12 *and* $\varphi_A := \inf\{\tilde{J}_\theta^{u,T}(t,x,y,v) : u = u(t,\omega)$ *admissible control*$\}$. *Then* $\varphi_M(t,x,y,v) = \varphi_A(t,x,y,v)$.

Proof. We first note that in the proof of Theorem 4.3.12, we have only used the properties (ii) and (iii) of Definition 4.2.4 of the Markov control $\bar{u}$. Since these two properties are true for a generic admissible control u, we can get as in Theorem 4.3.12(i).

$$\varphi_M(t,x,y,v) \leq \tilde{J}_\theta^{u,T}(t,x,y,v)$$

for every admissible control u. By taking infimum, we get $\varphi_M \leq \varphi_A$. Now using Theorem 4.3.12(ii) and Theorem 4.3.6(iv), $\bar{u}^\star$ is admissible, and $\varphi_M(t,x,y,v) \leq \tilde{J}_\theta^{\bar{u}^\star,T}(t,x,y,v)$. Thus $\varphi_M \geq \varphi_A$. Hence the result is proved. □

4.4 Conclusion

In this chapter a portfolio optimization problem, without any consumption and transaction cost, where stock prices are modeled by multi dimensional geometric jump diffusion market model with semi-Markov modulated coefficients is studied. We find the expression of optimal wealth for expected terminal utility method with risk sensitive criterion on finite time horizon. We have studied the existence of classical solution of HJB equation using a probabilistic approach. We have obtained the implicit expression of optimal portfolio. It is important to note that, the control is robust in the sense that the optimal control does not depend on the transition function of the regime. The corresponding problem in infinite horizon is yet to be investigated. This would require appropriate results on large deviation principle for semi-Markov processes which need to be carried out. The contents of this chapter is from [11].

5

Option Pricing

5.1 Introduction

In this chapter we shall study an option pricing problem. We consider a market where the asset price dynamics governed by a CSM switching geometric Brownian motion. We also allow the volatility coefficient to be time dependent. Under these assumptions the market become incomplete. We have also shown that under admissible strategies the market is arbitrage free . We shall study a locally risk minimizing pricing of European basket options. The option price can be obtained via the classical solution of a non-local partial differential equation. Well posedness of the PDE has been studied. We have also found a Volterra integral equation which is equivalent to the PDE.

The rest of this chapter is arranged in the following manner. We present model description in Section 2. In this section we describe the the asset price dynamics. Section 3 presents the approach of option pricing. In this section we state the main result of the chapter. In Section 4, we establish the existence, uniqueness and regularity of solution of a Volterra integral equation which is shown to be equivalent to the PDE in the next section. Section 5 deals with the well-posedness of the PDE. In this section we also derive certain properties of the solution and its derivative. Using the results of earlier sections, F-S decomposition of contingent claim is obtained in Section 6. In Section 7 we present a sensitivity analysis of the solution to the PDE. We calculate the quadratic residual risk in Section 8. We end this chapter with some concluding remark in Section 9.

5.2 Model description

Key Assumptions of the chapter

Throughout this chapter, we consider $(\Omega, \mathscr{F}, \mathbb{P})$ be a fixed probability space which is complete. We also assume that the market is frictionless.

5.2.1 Model Parameters

Let $\mathcal{X}$ denote a finite subset of $\mathbb{R}$. Without loss of generality, we choose $\mathcal{X} = \{1, 2, \ldots, k\}$ and $\mathcal{X}_2 := \{(i,j) | i \neq j \in \mathcal{X}\}$. Consider for each $l = 0, 1, \ldots, n$, $\lambda^l : \mathcal{X}_2 \times [0, \infty) \to (0, \infty)$ a continuously differentiable function in y with $\lambda^l_{ii}(y) = -\sum_{j \neq i} \lambda^l_{ij}(y)$ and

$$\lim_{y \to \infty} \Lambda^l_i(y) = \infty, \text{where}\, \Lambda^l_i(y) := \int_0^y \sum_{j \neq i} \lambda^l_{ij}(v) dv.$$

Assume that for each $j = 1, 2, \ldots, m_2$, ν_j denotes a finite Borel measure on $\mathbb{R}$. We assume that $r : \mathcal{X}^{n+1} \to [0, \infty)$, $\mu^l : [0, T] \times \mathcal{X}^{n+1} \to \mathbb{R}$, and $\sigma^l : [0, T] \times \mathcal{X}^{n+1} \to \mathbb{R}^n$ are continuous functions for each $l = 1, \ldots, n$.

5.2.2 Regime switching model for asset price dynamics

We also consider the prices of each assets in the market is governed by CSM process $\{X_t\}_{t \geq 0}$ taking values in $\mathcal{X}^{n+1}$, where $\mathcal{X} = \{1, \ldots, k\} \subset \mathbb{R}$. For every $l = 0, 1, \ldots, n$, $X^l := \{X^l_t\}$ is an age dependent process with instantaneous transition rates λ^l and let $Y_l := \{Y^l_t\}$ be the age process and X^l, Y^l satisfies the following SDE

$$\left.\begin{aligned} X^l_t &= X^l_0 + \int_{(0,t]} \int_{\mathbb{R}} h^l(X^l_{u-}, Y^l_{u-}, z_0) \wp^l(du, dz_0) \\ Y^l_t &= Y^l_0 + t - \int_{(0,t]} \int_{\mathbb{R}} g^l(X^l_{u-}, Y^l_{u-}, z_0) \wp^l(du, dz_0). \end{aligned}\right\} \tag{5.1}$$

Suppose that there are a locally risk free asset and n risky assets. Let S^0_t be the price of money market account, with floating interest rate $r(X_t)$ at time t. Therefore its value at time t is given by

$$dS^0_t = r(X_t) S^0_t dt, \quad S^0_0 = 1. \tag{5.2}$$

The prices of the l-th stock governed by X_t is given by the following stochastic differential equation

$$dS^l_t = S^l_t \left[\mu^l(t, X_t) dt + \sum_{j=1}^n \sigma^l_j(t, X_t)\, dW^j_t \right] \tag{5.3}$$

$$S_0^l = s_l, \quad s_l \geq 0,$$

where $\{W_t^j\}_{t\geq 0}$ are n independent standard Wiener processes defined on $(\Omega, \mathfrak{F}, P)$ independent of $\{\wp^l\}_{l=0}^n$. Here μ^l and $\sigma^l = (\sigma_1^l, \ldots, \sigma_n^l)$ represent the growth rate and volatility coefficient of l-th asset respectively. We define the volatility matrix $\sigma(t,x) := (\sigma_{l'}^l(t,x))_{ll'}$ with $\sigma^l(t,x)$ its l-th row vector and we denote $(S_t^1, \ldots, S_t^n)$ by S_t. Let $\{\mathfrak{F}_t\}_{t\geq 0}$ be the completion of filtration generated by S_t, X_t satisfying the usual hypothesis. Let $a(t,x) := \sigma(t,x)\sigma(t,x)^* = \left(\sum_{i=1}^n \sigma_i^l(t,x)\sigma_i^{l'}(t,x)\right)_{ll'}$ denote the diffusion matrix, where $*$ denotes the transpose operation. Then $a(t,x)$ is continuous on $[0,T]$.

Assumption 5.1. *We assume that $\sigma(t,x)$ is invertible for each $(t,x) \in [0,T] \times \mathcal{X}^{n+1}$.*

We first note that, the SDE (5.3) has a unique strong solution with positive continuous paths and is given by

$$S_t^l = s_l \exp\left[\int_0^t \left(\mu^l(u, X_u) - \frac{1}{2}a^{ll}(u, X_u)\right) du + \sum_{j=1}^n \int_0^t \sigma_j^l(u, X_u)\, dW_u^j\right], \text{ for } l \geq 1. \quad (5.4)$$

It follows from (5.4) that

$$\ln \frac{S_{t+v}^l}{S_t^l} = \int_t^{t+v} (\mu^l(u, X_u) - \frac{1}{2}a^{ll}(u, X_u))\, du + \int_t^{t+v} \sum_{j=1}^n \sigma_j^l(t, X_t) dW_t^j.$$

We define $Z := (Z^1, \ldots, Z^n)$, where for each $l = 1, \ldots, n$, $Z^l := \ln \frac{S_{t+v}^l}{S_t^l}$. Now we recall some notations from Chapter 3. Let $\ell(t)$ be the component of X where the subsequent jump happens and $\tau^l(t)$ denotes the life of l-th component of X. Clearly the conditional distribution of Z given $S_t = s, X_t = x, Y_t = y, \ell(t) = m, \tau^m(t) = v$ is conditional normal with mean $\bar{z} := (\bar{z}^1, \ldots, \bar{z}^n)$, where

$$\bar{z}^l := \int_t^{t+v} (\mu^l(u, x) - \frac{1}{2}a^{ll}(u, x))\, du, \quad (5.5)$$

and covariance matrix Σ with $\Sigma^{ll'} := cov\left(Z^l, Z^{l'}\right)$. i.e

$$\begin{aligned}
&\Sigma^{ll'} \\
=&\mathbb{E}\left[\int_t^{t+v} \sigma^l(u, X_u)\, dW_u \times \int_t^{t+v} \sigma^{l'}(u, X_u)\, dW_u \middle| S_t = s, X_t = x, Y_t = y, \ell(t) = m, \tau^m(t) = v\right] \\
=& \int_t^{t+v} a^{ll'}(u, x) du. \quad (5.6)
\end{aligned}$$

In (5.5) and (5.6), we have used the fact that the process X remains constant on $[t, t+v)$ provided $\ell(t) = m, \tau^m(t) \geq v$ hold for some m. We summarize the above derivation in the following lemma where, we use a function $\theta : (0,\infty)^n \times (0,\infty) \times (0,\infty)^n \times \mathcal{X}^{n+1} \times (0,\infty) \to \mathbb{R}$ given by

$$\theta(\varsigma; t, s, x, v) := \frac{1}{\sqrt{(2\pi)^n |\Sigma| \varsigma_1 \varsigma_2 \ldots \varsigma_n}} \exp\left(-\frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll'} (z^l - \bar{z}^l)(z^{l'} - \bar{z}^{l'})\right), \tag{5.7}$$

where $|\Sigma|$ is the determinant of Σ, $z^l = \ln(\frac{\varsigma^l}{s^l})$ and $s \in (0,\infty)^n, t \geq 0, x \in \mathcal{X}^{n+1}, v > 0$ and $\Sigma^{-1}_{ll'}$ is the ll'th element of Σ^{-1} for $l = 1, \ldots, n$.

Lemma 5.2.1. *If S_t satisfies (5.3), then for any $v > 0, t \geq 0$,*

(i)

$$P\left(\left(\frac{S^l_{t+v}}{S^l_t} \leq \varsigma_l\right)_{l=1,\ldots,n} \middle| S_t = s, X_t = x, Y_t = y, \ell(t) = m, \tau^m(t) = v\right)$$
$$= \int_{\prod_{l=1}^n (0,\varsigma_l)} \theta(r; t, s, x, v) dr,$$

(ii) the conditional expectation is given by

$$\mathbb{E}\left[\frac{S^l_{t+v}}{S^l_t} \middle| S_t = s, X_t = x, Y_t = y, \ell(t) = m, \tau^m(t) = v\right] = e^{\int_t^{t+v} \mu^l(u,x) du},$$

(iii) the conditional covariance is given by

$$cov\left(\frac{S^l_{t+v}}{S^l_t}, \frac{S^{l'}_{t+v}}{S^{l'}_t} \middle| S_t = s, X_t = x, Y_t = y, \ell(t) = m, \tau^m(t) = v\right)$$
$$= e^{\int_t^{t+v} \left(\mu^l(u,x) + \mu^{l'}(u,x)\right) du} \left(e^{\int_t^{t+v} a^{ll'}(u,x) du} - 1\right).$$

The following results on conditional moments of first and second order would be useful. In particular the following lemma asserts the square integrability of the asset price process.

Lemma 5.2.2. *Let $\{S^l_t\}_{t\geq 0}$ be as in (5.3) and $\{\mathcal{F}^X_t\}_{t\geq 0}$ be the filtration generated by X.*

(i) Then for each $l = 1, \ldots, n$, and $t \geq 0$,

$$\mathbb{E}\left[S^l_t \middle| \mathcal{F}^X_t\right] \leq s_l e^{\int_0^t \mu^l(u, X_u) du}.$$

(ii) For each l, $\mathbb{E}\left(S_t^{l^2}\middle|\mathcal{F}_t^X\right) < \infty$ for all t.

Proof. (i) Let T_i^l be the time of i-th transition of X^l,

$$\begin{aligned}\mathbb{E}\left[\frac{S_t^l}{S_0^l}\middle|\mathcal{F}_t^X\right] &= \mathbb{E}\left[\prod_{i=1}^{\infty}\frac{S_{T_i^l\wedge t}^l}{S_{T_{i-1}^l\wedge t}^l}\middle|\mathcal{F}_t^X\right]\\ &= \mathbb{E}\left[\lim_{N\to\infty}\prod_{i=1}^{N}\frac{S_{T_i^l\wedge t}^l}{S_{T_{i-1}^l\wedge t}^l}\middle|\mathcal{F}_t^X\right]\\ &\leq \varliminf_{N\to\infty}\mathbb{E}\left[\prod_{i=1}^{N}\frac{S_{T_i^l\wedge t}^l}{S_{T_{i-1}^l\wedge t}^l}\middle|\mathcal{F}_t^X\right],\end{aligned}$$

by Fatou's lemma. Now since for each $i = 1,\ldots,n$, $\frac{S_{T_i^l\wedge t}^l}{S_{T_{i-1}^l\wedge t}^l}$ are conditionally independent to each other given time t, and using Lemma (5.2.1)(ii) the above limit can be rewritten as $\varliminf_{N\to\infty}\prod_{i=1}^N e^{\int_{T_{i-1}^l\wedge t}^{T_i^l\wedge t}\mu^l(u,X_u)du}$, which is same as $e^{\int_0^t \mu(u,X_u)du}$.

(ii) In a similar line of proof (i), using Lemma (5.2.1)(iii), the proof follows. $\square$

We denote the joint process $(\hat{S}_t^1,\ldots,\hat{S}_t^n)$ by $\hat{S}_t$, where $\hat{S}_t^l$ is given by $(S_t^0)^{-1}S_t^l$ and represents the discounted l-th stock price. For each l,

$$d\hat{S}_t^l = \hat{S}_t^l\left[\sum_{j=1}^n \sigma_j^l(t,X_t)\,dW_t^j + \Big(\mu^l(t,X_t) - r(X_t)\Big)\,dt\right], \tag{5.8}$$

with $\hat{S}_0^l = s_l$.

5.2.3 Arbitrage opportunity

In this subsection we show that the market is arbitrage free under admissible strategy. To see this, we seek existence of an EMM (using Theorem 1.4.8) . Consider $\gamma_l(t,x) := \sum_{j=1}^n (\sigma^{-1}(t,x))_j^l\left(\mu^j(t,x) - r(x)\right)$ for each $l = 1,\ldots,n$. Under the Assumption 5.1 and the continuity assumption on parameters, the Novikov's condition (Theorem 1.1.10) holds, i.e., for every $t\in[0,T]$,

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{l=1}^n\int_0^t \gamma_l^2(u,X_u)\,du\right)\right] < \infty.$$

Hence

$$\mathcal{Z}_t := \exp\left(-\sum_{l=1}^n\int_0^t \gamma_l(u,X_u)\,dW_u^j - \frac{1}{2}\sum_{l=1}^n\int_0^t \gamma_l^2(u,X_u)\,du\right),$$

is a square integrable martingale and $\mathbb{E}\mathcal{Z}_T = 1$. Consider an equivalent measure $\mathbb{P}^*$ defined by $d\mathbb{P}^* = \mathcal{Z}_T d\,\mathbb{P}$. It is easy to check that $\mathbb{P}^*$ is a probability measure. Hence by Girsanov's Theorem 1.1.17, $\bar{W}_t$ is a Wiener process under the probability measure $\mathbb{P}^*$, where $\bar{W}_t^l = W_t^l + \int\limits_0^t \gamma_l(u, X_u)du$. Thus (5.3) becomes

$$dS_t^l = S_t^l \left[r(X_t)dt + \sum_{j=1}^n \sigma_j^l(t, X_t)\ d\bar{W}_t^j \right]. \tag{5.9}$$

Therefore under $\mathbb{P}^*$, the discounted stock price $\hat{S}_t^l$ is a martingale and hence $\mathbb{P}^*$ is an equivalent martingale measure. This proves that the market has no arbitrage under admissible strategies. The class of admissible strategy is presented in the next section.

5.3 Pricing Approach and the main result

If ξ_t^l denotes the number of units invested in the l-th stock at time t and ε_t denotes the number of units of the risk free asset, then $\pi = \{\pi_t = (\xi_t, \varepsilon_t)\}_{t\in[0,T]}$ is called a portfolio strategy. For $t \in [0, T]$, $V_t(\pi) := \sum\limits_{l=1}^n \xi_t^l S_t^l + \varepsilon_t S_t^0$ is said to be value process of the portfolio and the discounted value process is given by

$$\hat{V}_t(\pi) = \xi_t \hat{S}_t + \varepsilon_t.$$

Definition 5.3.1. *A portfolio strategy $\pi = \{\pi_t = (\xi_t, \varepsilon_t), 0 \le t \le T\}$ is called* admissible *if it satisfies the following conditions*

(i) $\xi_t = (\xi_t^1, \ldots, \xi_t^n)$ is an n-dimensional predictable process and for each $l = 1, \ldots, n$,

$$\sum_{ll'} \int\limits_0^T \xi_t^l S_t^l a^{ll'}(t, X_t) S_t^{l'} \xi_t^{l'} dt < \infty.$$

(ii) ε is adapted, and $E(\varepsilon_t^2) < \infty\ \forall\ t \in [0, T]$.

(iii) $\mathbb{P}(\hat{V}_t(\pi) \ge -a,\ \forall\ t) = 1$ for some positive a.

An *admissible* strategy ξ is called hedging strategy for an $\mathfrak{F}_T$ measurable claim H if $V_T(\xi) = H$. For example, the claim associated to a European call option on S^1 is $H = (S_T^1 - K)^+$, where K is the strike price and T is the maturity time. To price and hedge an option, an investor prefers an admissible hedging strategy which requires minimal amount of additional cash flow. In [23] the notion of "optimal strtegy" is developed based on this idea. There the initial capital is referred as locally risk-minimizing price of the option. It is shown in [23] that if the market is arbitrage-free, the existence of an optimal strategy for hedging a claim H, is equivalent to the existence of the Föllmer-Schweizer decomposition of the discounted claim $\hat{H} := {S_T^0}^{-1} H$ in the form

$$\hat{H} = H_0 + \sum_{l=1}^{n} \int_0^T \xi_t^l \, d\hat{S}_t^l + L_T^{\hat{H}}, \tag{5.10}$$

where $H_0 \in L^2(\Omega, \mathfrak{F}_0, P)$, $L^{\hat{H}} = \{L_t^{\hat{H}}\}_{0 \le t \le T}$ is a square integrable martingale starting with zero and orthogonal to the martingale part of S_t, and $\{(\xi_t^1, \ldots, \xi_t^n)\}_{t \ge 0}$ satisfies Definition 5.3.1. Further ξ^l, appeared in the decomposition, constitutes the optimal strategy. Indeed the optimal strategy ξ^* is given by

$$\begin{aligned} {\xi_t^*}^l :=& \xi_t^l, \text{ for } l = 1, 2, \ldots, n, \\ \hat{V}_t :=& H_0 + \sum_{l=1}^{n} \int_0^t \xi_u^l \, d\hat{S}_u^l + L_t^{\hat{H}}, \\ {\xi_t^*}^0 :=& \hat{V}_t - \sum_{l=1}^{n} \xi_t^l \hat{S}_t^l, \end{aligned} \tag{5.11}$$

and $S_t^0 \hat{V}_t$ represents the *locally risk minimizing price* at time t of the claim H. Thus the Föllmer-Schweizer decomposition is the key thing in settling down the pricing and hedging problems in a given market. We refer to [57] for more details. In this chapter, we are interested to price a special class of contingent claims, of the form $H = K(S_T)$, where we make the following assumptions on $K : \mathbb{R}_+^n \to \mathbb{R}_+$.

Assumption 5.2. *(i) $K(s)$ is Lipschitz continuous function.*

(ii) There exists $c_1 \in \mathbb{R}^n$, and $c_2 > 0$ such that $|K(s) - c_1^* s| < c_2$ for all $s \in \mathbb{R}_+^n$.

This class includes claims of all types of basket options consisting finitely many vanilla options. As an example, a typical basket call option has a claim $(\sum_{l=1}^{n} c_l S_t^l - \bar{K})^+$, where $\bar{K}$ is the strike price. Our primary goal in this chapter is to obtain expressions for locally risk-minimizing price process and the optimal strategy corresponding to a claim $K(S_T)$.

5.3.1 The pricing equation

In order to study the locally risk minimizing option pricing of the contingent claim $K(S_T)$, we study the following Cauchy problem

$$\frac{\partial \varphi}{\partial t}(t, s, x, y) + \sum_{l=0}^{n} \frac{\partial \varphi}{\partial y^l}(t, s, x, y) + r(x) \sum_{l=1}^{n} s^l \frac{\partial \varphi}{\partial s^l}(t, s, x, y)$$

$$+\frac{1}{2}\sum_{l=1}^{n}\sum_{l'=1}^{n} a^{ll'}(t,x)s^l s^{l'}\frac{\partial^2\varphi}{\partial s^l \partial s^{l'}}(t,s,x,y)$$
$$+\sum_{l=0}^{n}\sum_{j\neq x^l}\lambda^l_{x^l j}(y^l)\left[\varphi(t,s,R^l_j x,R^l_0 y)-\varphi(t,s,x,y)\right]=r(x)\ \varphi(t,s,x,y), \tag{5.12}$$

defined on

$$\mathcal{D}:=\{(t,s,x,y)\in(0,T)\times(0,\infty)^n\times\mathcal{X}^{n+1}\times(0,T)^{n+1}\mid y\in(0,t)^{n+1}\},$$

and with conditions

$$\varphi(T,s,x,y)=K(s);\qquad s\in[0,\infty)^n;\qquad 0\leq y^l\leq T;\qquad x^l\in\mathcal{X},\qquad l=0,1,\ldots,n, \tag{5.13}$$

where $R^l_j v:=v+(j-v^l)e_l$ for $v\in\mathbb{R}^{n+1}$ and e_l is an $n+1$ dimensional vector with only l-th component 1 and rest are zero.

Remark 5.3.2. *It is important to note that, if there is one risky asset and one risk-free asset in the market with the assumption that all the market parameters are constant, then* (5.12) *reduces to the famous B-S-M equation, given by*

$$\varphi_t(t,s)+rs\varphi_x(t,s)+\frac{1}{2}\sigma^2\varphi_{xx}(t,x)=r\varphi(t,x) \tag{5.14}$$

5.3.2 The main result

We study the Cauchy problem (5.12)-(5.13) and obtain expressions of price and hedging using solution of (5.12)-(5.13). We state this result as theorems at the end of this section. But before that we introduce some notation and definition. Again we consider a notation from Chapter 4. We define a linear operator

$$D_{t,y}g(t,y):=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\{g(t+\varepsilon,y+\varepsilon\mathbb{1})-g(t,y)\}, \tag{5.15}$$

where $\mathrm{dom}(D_{t,y})$, the domain of $D_{t,y}$, contains all measurable functions g on $[0,T]\times[0,T]$ such that above limit exists for every $(t,y)\in[0,T]\times[0,T]$. We rewrite (5.12) using the above notation

$$D_{t,y}\varphi(t,s,x,y)+r(x)\sum_{l=1}^{n}s^l\frac{\partial\varphi}{\partial s^l}(t,s,x,y)+\frac{1}{2}\sum_{l=1}^{n}\sum_{l'=1}^{n}a^{ll'}(t,x)s^l s^{l'}\frac{\partial^2\varphi}{\partial s^l\partial s^{l'}}(t,s,x,y)$$
$$+\sum_{l=0}^{n}\sum_{j\neq x^l}\lambda^l_{x^l j}(y^l)\left[\varphi(t,s,R^l_j x,R^l_0 y)-\varphi(t,s,x,y)\right]=r(x)\ \varphi(t,s,x,y). \tag{5.16}$$

Now we define the meaning of classical solution of the PDE.

Definition 5.3.3. *We say, $\varphi : \mathcal{D} \to \mathbb{R}$ is a classical solution of* (5.16)-(5.13) *if $\varphi \in \text{dom}(D_{t,y})$, twice differentiable with respect to s and for all $(t, s, x, y) \in \mathcal{D}$,* (5.16)-(5.13) *are satisfied.*

Theorem 5.3.4. *Under the Assumption 5.1, the initial value problem* (5.16)-(5.13) *has a unique classical solution in the class of functions with at most linear growth.*

We establish this at the end of section 5.5. We present the locally risk-minimizing strategy in terms of the solution to the PDE (5.16)-(5.13). The proof of the following Theorem is deferred to Section 5.6.

Theorem 5.3.5. *Let φ be the unique classical solution of* (5.16)-(5.13) *in the class of functions with at most linear growth and (ξ, ε) be given by*

$$\xi_t^l := \frac{\partial \varphi}{\partial s^l}(t, S_t, X_{t-}, Y_{t-}) \ \forall\ l = 1, \ldots, n, \ \text{and } \varepsilon_t := e^{-\int_0^t r(X_u)du} \left(\varphi(t, S_t, X_t, Y_t) - \sum_{l=1}^{n} \xi_t^l S_t^l \right). \tag{5.17}$$

Then

1. *(ξ, ε) is the optimal admissible strategy,*
2. *$\varphi(t, S_t, X_t, Y_t)$ is the locally risk minimizing price of the claim $K(S_T)$ at time t.*

In order to study the well-posedness of solution of the PDE (5.16)-(5.13), we study a Volterra integral equation of second kind. We prepare ourself by showing the existence and uniqueness of solution of the integral equation in the next section.

5.4 Volterra Integral Equation

For each x, consider the following Cauchy problem which is known as B-S-M PDE as in (5.14),

$$\frac{\partial \rho_x(t,s)}{\partial t} + r(x) \sum_{l=1}^{n} s^l \frac{\partial \rho_x(t,s)}{\partial s^l} + \frac{1}{2} \sum_{l=1}^{n} \sum_{l'=1}^{n} a^{ll'}(t,x) s^l s^{l'} \frac{\partial^2 \rho_x(t,s)}{\partial s^l \partial s^{l'}} = r(x) \rho_x(t,s) \tag{5.18}$$

for $(t, s) \in (0, T) \times (0, \infty)^n$ and $\rho_x(T, s) = K(s)$, for all $s \geq 0$. This has a classical solution with at most linear growth (see [38]), provided K is of at most linear growth. We would like to mention that ρ_x is infinitely many times differentiable with respect to s.

For $\zeta \in \mathbb{R}^n$, let $\|\zeta\|_1$ denote the norm $\sum_{l=1}^{n} |\zeta^l|$. Let

$$\mathcal{B} := \{\varphi : \bar{\mathcal{D}} \to \mathbb{R}, \text{measurable} \mid \|\varphi\|_L := \sup_{\bar{\mathcal{D}}} \frac{|\varphi(t,s,x,y)|}{1 + \|s\|_1} < \infty\}. \tag{5.19}$$

Let $C_s^2(\mathcal{D}) := C^{0,2,0}(\mathcal{D})$ be the set of all measurable functions on $\mathcal{D}$, which are also twice differentiable with respect to s.

Let Σ be an $n \times n$ matrix, whose elements are as in (5.6). We further use the notation Σ, $|\Sigma|$ and Σ^{-1} as in (5.7). By replacing $\mu^l(u, x)$ by $r(x)$ in (5.5), we define a function $\alpha : (0,\infty)^n \times (0,\infty) \times (0,\infty)^n \times \mathcal{X}^{n+1} \times (0,\infty) \to \mathbb{R}$ by

$$\alpha(\varsigma; t, s, x, v) = \frac{1}{\sqrt{(2\pi)^n |\Sigma| \varsigma_1 \varsigma_2 \ldots \varsigma_n}} \exp\left(-\frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll'} (z^l - \bar{z}^l)(z^{l'} - \bar{z}^{l'}) \right), \tag{5.20}$$

where $z^l = \ln(\frac{\varsigma^l}{s^l})$, and $\bar{z}^l := \int_t^{t+v} (r(x) - \frac{1}{2} a^{ll}(u, x))\, du$ for $s \in (0,\infty)^n, t \geq 0, x \in \mathcal{X}^{n+1}, v > 0$ and $\Sigma^{-1}_{ll'}$ is the ll'th element of Σ^{-1} for $l = 1, \ldots, n$. It is clear from (5.20) that $\alpha(\varsigma; t, s, x, v)$ is a log-normal density with respect to ς variable for a fixed (t, s, x, v).

Lemma 5.4.1. *Let $\alpha(\varsigma; t, s, x, v)$ be as in* (5.20). *Then $\alpha(\varsigma; t, s, x, v)$ is C^1 in t, v, and infinite time differentiable in s.*

Proof. From (5.6), we get that Σ^{-1} exists for all $v > 0$ and is differentiable in t and v. Therefore $\alpha(\varsigma; t, s, x, v)$ defined in (5.7) is differentiable in t and v. Taking logarithm on both the sides of (5.7), we have

$$\ln \alpha(\varsigma; t, s, x, v) = -\ln \left(\frac{1}{\sqrt{(2\pi)^n \varsigma_1 \varsigma_2 \ldots \varsigma_n}} \right) - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll'} (z^l - \bar{z}^l)(z^{l'} - \bar{z}^{l'}). \tag{5.21}$$

Now taking derivative on both the sides of (5.21) with respect to t and using Jacobi's formula.

$$\begin{aligned} \alpha_t &= \alpha \left(-\frac{1}{2} \frac{|\Sigma|}{|\Sigma|} tr \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial t} \right) - \frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll't} (z^l - \bar{z}^l)(z^{l'} - \bar{z}^{l'}) + \frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll'} \bar{z}^l_t (z^{l'} - \bar{z}^{l'}) \right. \\ &\quad \left. + \frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll'} (z^l - \bar{z}^l) \bar{z}^{l'}_t \right) \qquad (5.22) \\ &= \alpha \Big(-\frac{1}{2} tr \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial t} \right) - \frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll't} (z^l - \bar{z}^l)(z^{l'} - \bar{z}^{l'}) + \frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll'} \bar{z}^l_t (z^{l'} - \bar{z}^{l'}) \\ &\quad + \frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll'} (z^l - \bar{z}^l) \bar{z}^{l'}_t \Big). \qquad (5.23) \end{aligned}$$

Similarly

$$\alpha_v = \alpha \Big(-\frac{1}{2} tr \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial v} \right) - \frac{1}{2} \sum_{ll'} \Sigma^{-1}_{ll'v} (z^l - \bar{z}^l)(z^{l'} - \bar{z}^{l'})$$

$$+\frac{1}{2}\sum_{ll'}\Sigma_{ll'}^{-1}\bar{z}_v^l(z^{l'}-\bar{z}^{l'})+\frac{1}{2}\sum_{ll'}\Sigma_{ll'}^{-1}(z^l-\bar{z}^l)\bar{z}_v^{l'}\Big), \tag{5.24}$$

where $\Sigma_{ll't}^{-1} := \frac{\partial\Sigma_{ll}^{-1}}{\partial t}$ and $\Sigma_{ll'v}^{-1} := \frac{\partial\Sigma_{ll}^{-1}}{\partial v}$. In similar manner one can show α is infinite times continuously differentiable in s. □

Now we are in a position to introduce the integral equation. To do this, we recall some notations (Notation 3.3, Notation 3.4, Notation 3.5) from Chapter 3.

Notation 5.1.

1. *p_{ij}^l := probability the X^l jumps from state i to state j, defined as in* (3.4).
2. *$\tau^l(t)$:= time period from time t after which the l-th component of X_t would have a first transition.*
3. *We denote the conditional c.d.f of $\tau^l(t)$ given $X_t^l = i$ and $Y_t^l = \bar{y}$ as $F_{\tau^l}(\cdot|i,\bar{y})$, and the conditional p.d.f of $\tau^l(t)$ given $X_t^l = i$ and $Y_t^l = \bar{y}$ as $f_{\tau^l}(\cdot|i,\bar{y})$.*
4. *Let $\ell(t)$ be the component of X_t, where the subsequent transition happens.*
5. *Let $F_{\tau^l|l}(\cdot|x,y)$ and $f_{\tau^l|l}(\cdot|x,y)$ be the conditional c.d.f and conditional p.d.f of $\tau^l(t)$ given $X_t = x, Y_t = y$ and $\ell(t) = l$.*

Using Notation 5.1, consider the following integral equation

$$\varphi(t,s,x,y) = \sum_{l=0}^{n} P_{t,x,y}(\ell(t)=l)\left(\rho_x(t,s)\left(1-F_{\tau^l|l}(T-t\mid x,y)\right)+\int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v\mid x,y)\times\right.$$
$$\left.\sum_{j\neq x^l} p_{x^l j}^l(y^l+v)\int_{\mathbb{R}_+^n}\varphi\left(t+v,\varsigma,R_j^l x,R_0^l(y+v\mathbf{1})\right)\alpha(\varsigma;t,s,x,v)\,d\varsigma\,dv\right). \tag{5.25}$$

Lemma 5.4.2. *The integral equation* (5.25) *has a unique solution in* $\mathcal{B}$ (as in (5.19)).

Proof. We first note that a solution of (5.25) is a fixed point of the operator A and vice versa, where

$$A\varphi(t,s,x,y) := \sum_{l=0}^{n} P_{t,x,y}(\ell(t)=l)\left(\rho_x(t,s)\left(1-F_{\tau^l|l}(T-t\mid x,y)\right)+\int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v\mid x,y)\times\right.$$
$$\left.\sum_{j\neq x^l} p_{x^l j}^l(y^l+v)\int_{\mathbb{R}_+^n}\varphi\left(t+v,\varsigma,R_j^l x,R_0^l(y+v\mathbf{1})\right)\alpha(\varsigma;t,s,x,v)\,d\varsigma\,dv\right).$$

It is simple to verify that for each $\varphi \in \mathcal{B}$, $A\varphi : \bar{\mathcal{D}} \to \mathbb{R}$ is measurable. To prove that A is a contraction in $\mathcal{B}$, we need to show that for $\varphi_1, \varphi_2 \in \mathcal{B}$, $\|A\varphi_1 - A\varphi_2\|_L \leq J\|\varphi_1 - \varphi_2\|_L$

where $J < 1$. In order to show existence and uniqueness in the prescribed class, it is sufficient to show that A is a contraction in $\mathcal{B}$. Then the Banach fixed point Theorem ensures existence and uniqueness of the fixed point in $\mathcal{B}$. To show that for $\varphi_1, \varphi_2 \in \mathcal{B}$, $\|A\varphi_1 - A\varphi_2\|_L \le J\|\varphi_1 - \varphi_2\|_L$ where $J < 1$, we compute

$$\begin{aligned}
\|A\varphi_1 - A\varphi_2\|_L =& \sup_{\mathcal{D}} \left| \frac{A\varphi_1 - A\varphi_2}{1 + \|s\|_1} \right| \\
=& \sup_{\mathcal{D}} \Bigg| \sum_{l=0}^{n} P_{t,x,y}(\ell(t) = l) \int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v \mid x, y) \sum_{j^l \neq x^l} p^l_{x^l j^l}(y^l + v) \times \\
& \int_{\mathbb{R}^n_+} (\varphi_1 - \varphi_2)(t + v, \varsigma, R^l_j x, R^l_0(y + v\mathbf{1})) \frac{\alpha(\varsigma; t, s, x, v)}{1 + \|s\|_1} \, d\varsigma \, dv \Bigg| \\
\le& \sup_{\mathcal{D}} \Bigg| \sum_{l=0}^{n} P_{t,x,y}(\ell(t) = l) \int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v \mid x, y) \sum_{j^l \neq x^l} p^l_{x^l j^l}(y^l + v) \times \\
& \int_{\mathbb{R}^n_+} (1 + \|\varsigma\|_1) \sup_{(t',\varsigma',x',y') \in \mathcal{D}} \left[\frac{(\varphi_1 - \varphi_2)(t', \varsigma', x', y')}{1 + \|\varsigma'\|_1} \right] \frac{\alpha(\varsigma; t, s, x, v)}{1 + \|s\|_1} \, d\varsigma \, dv \Bigg| \\
=& \sup_{\mathcal{D}} \left| \sum_{l=0}^{n} P_{t,x,y}(\ell(t) = l) \int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v \mid x, y) \|\varphi_1 - \varphi_2\|_L \frac{\bar{\alpha}(t, s, x, v)}{1 + \|s\|_1} \, dv \right|,
\end{aligned}$$

where $\bar{\alpha}(t, s, x, v) := \int_{\mathbb{R}^n_+} (1 + \|\varsigma\|_1) \alpha(\varsigma; t, s, x, v) \, d\varsigma$. Replacing $\mu^l(u, x)$ by $r(x)$ in Lemma 5.2.2(i), we get

$$\bar{\alpha}(t, s, x, v) = 1 + \|s\|_1 e^{r(x)v}.$$

Thus, $\|A\varphi_1 - A\varphi_2\|_L \le J\|\varphi_1 - \varphi_2\|_L$, where

$$\begin{aligned}
J =& \sup_{\mathcal{D}} \left| \sum_{l=0}^{n} P_{t,x,y}(\ell(t) = l) \int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v \mid x, y) \frac{1 + \|s\|_1 e^{r(x)v}}{1 + \|s\|_1} \, dv \right| \\
\le& \sup_{\mathcal{D}} \left| \sum_{l=0}^{n} P_{t,x,y}(\ell(t) = l) \int_0^{T-t} f_{\tau^l|l}(v \mid x, y) \, dv \right| \\
=& \sup_{\mathcal{D}} \left| \sum_{l=0}^{n} P_{t,x,y}(\ell(t) = l) F_{\tau^l|l}(v \mid x, y) \right| \\
<& \sup_{\mathcal{D}} \left| \sum_{l=0}^{n} P_{t,x,y}(\ell(t) = l) \right| = 1,
\end{aligned}$$

using $r(x) \geq 0$ and the fact that $F^l(y|i) < 1$ for all l, x, y and i. Thus A is a contraction in $\mathcal{B}$. This completes the proof. □

Remark 5.4.3. *By a direct substitution* $t = T$ *in the* (5.25)*, we obtain* $\varphi(T, s, x, y) = K(s)$*. It is interesting to note that we do not have to impose any other boundary conditions for existence and uniqueness of solution of* (5.25)*. We can directly obtain other boundary values by substituting the boundary in the integral equation.*

Lemma 5.4.4. *Let* φ *be the solution of the integral equation* (5.25)*. Then (i)* $\varphi \in \mathrm{dom}(D_{t,y}) \cap C_s^2(\mathcal{D})$*, and (ii)* $\varphi(t, s, x, y)$ *is non-negative.*

Proof. (i) Using the smoothness of ρ_x for each x, the first term on the right hand side of (5.25) is in $\mathrm{dom}(D_{t,y}) \cap C_s^2(\mathcal{D})$. Thus it is enough to check the desired smoothness of

$$\beta_l(t,s,x,y) = \int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v \mid x, y) \sum_{j \neq x^l} p^l_{x^l j}(y^l + v) \int_{\mathbb{R}^n_+} \varphi\left(t + v, \varsigma, R^l_j x, R^l_0(y + v\mathbf{1})\right) \times \alpha(\varsigma; t, s, x, v)\, d\varsigma\, dv.$$

First we check the applicability of $D_{t,y}$. It is easy to see that $D_{t,y}\beta_l(t,s,x,y)$ is the limit of the following expression

$$\begin{aligned}
&\frac{1}{\varepsilon}\Big[\int_0^{T-t-\varepsilon} e^{-r(x)v} f_{\tau^l|l}(v \mid x, y + \varepsilon\mathbf{1}) \sum_{j \neq x^l} p^l_{x^l j}(y^l + v + \varepsilon) \\
&\times \int_{\mathbb{R}^n_+} \varphi\left(t + v + \varepsilon, \varsigma, R^l_j x, R^l_0(y + (v+\varepsilon)\mathbf{1})\right) \alpha(\varsigma; t+\varepsilon, s, x, v)\, d\varsigma\, dv \\
&- \int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v \mid x, y) \sum_{j \neq x^l} p^l_{x^l j}(y^l + v) \\
&\times \int_{\mathbb{R}^n_+} \varphi\left(t + v, \varsigma, R^l_j x, R^l_0(y + v\mathbf{1})\right) \alpha(\varsigma; t, s, x, v)\, d\varsigma\, dv\Big].
\end{aligned}$$

After a suitable substitution, the above expression becomes

$$\begin{aligned}
&\int_\varepsilon^{T-t} e^{-r(x)v} \sum_{j \neq x^l} p^l_{x^l j}(y^l + v) \int_{\mathbb{R}^n_+} \varphi\left(t + v, \varsigma, R^l_j x, R^l_0(y + v\mathbf{1})\right) \bar{\beta}_\varepsilon(v, \varsigma; t, s, x, y) d\varsigma\, dv \\
&- \frac{1}{\varepsilon}\int_0^{\varepsilon} e^{-r(x)v} f_{\tau^l|l}(v|x,y) \sum_{j \neq x^l} p^l_{x^l j}(y^l + v) \int_{\mathbb{R}^n_+} \varphi\left(t + v, \varsigma, R^l_j x, R^l_0(y + v\mathbf{1})\right) \alpha(\varsigma; t, s, x, v)\, d\varsigma\, dv,
\end{aligned} \tag{5.26}$$

where

$$\bar{\beta}_\varepsilon(v,\varsigma;t,s,x,y) := \frac{1}{\varepsilon}\Big(e^{r(x)\varepsilon} f_{\tau^l|l}(v-\varepsilon \mid x, y+\varepsilon\mathbf{1})\alpha(\varsigma;t+\varepsilon,s,x,v-\varepsilon) - f_{\tau^l|l}(v \mid x,y)\alpha(\varsigma;t,s,x,v)\Big).$$

Now the above defined $\bar{\beta}_\varepsilon$ can be rewritten as

$$\begin{aligned}
&\frac{1}{\varepsilon}\Big[\Big(e^{r(x)\varepsilon}-1+1\Big)\Big(f_{\tau^l|l}(v-\varepsilon \mid x,y+\varepsilon\mathbf{1}) - f_{\tau^l|l}(v \mid x,y+\varepsilon\mathbf{1}) + f_{\tau^l|l}(v \mid x,y+\varepsilon\mathbf{1})\\
&- f_{\tau^l|l}(v \mid x,y) + f_{\tau^l|l}(v \mid x,y)\Big) \times \Big(\alpha(\varsigma;t+\varepsilon,s,x,v-\varepsilon) - \alpha(\varsigma;t,s,x,v-\varepsilon)\\
&+ \alpha(\varsigma;t,s,x,v-\varepsilon) - \alpha(\varsigma;t,s,x,v) + \alpha(\varsigma;t,s,x,v)\Big) - f_{\tau^l|l}(v \mid x,y)\alpha(\varsigma;t,s,x,v)\Big]. \quad (5.27)
\end{aligned}$$

Due to the continuous differentiability results in Lemma 3.3.5 and Lemma 5.4.1, we can use the mean value Theorem to rewrite (5.27) as

$$\begin{aligned}
&\Big[\Big(\varepsilon r(x)e^{r(x)\varepsilon_0}+1\Big)\Big(-f'_{\tau^l|l}(v-\varepsilon_1 \mid x,y+\varepsilon\mathbf{1}) + \sum_{i=1}^n \frac{\partial}{\partial y_i} f_{\tau^l|l}(v \mid x,y+\varepsilon_2\mathbf{1}) + \frac{1}{\varepsilon} f_{\tau^l|l}(v \mid x,y)\Big)\\
&\times\Big(\varepsilon\alpha_t(\varsigma;t+\varepsilon_3,s,x,v-\varepsilon) - \varepsilon\alpha_v(\varsigma;t,s,x,v-\varepsilon_4) + \alpha(\varsigma;t,s,x,v)\Big)\\
&-\frac{1}{\varepsilon} f_{\tau^l|l}(v \mid x,y)\alpha(\varsigma;t,s,x,v)\Big],
\end{aligned}$$

for some $\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_3 < \varepsilon$. After some rearrangement of terms in the above expression, we get

$$\begin{aligned}
\bar{\beta}_\varepsilon(v,\varsigma;t,s,x,y) = \alpha(\varsigma;t,s,x,v)&\Big(r(x)e^{r(x)\varepsilon_0} f_{\tau^l|l}(v|x,y) - f'_{\tau^l|l}(v-\varepsilon_1 \mid x,y+\varepsilon\mathbf{1})\\
&+ \sum_{i=1}^n \frac{\partial}{\partial y_i} f_{\tau^l|l}(v \mid x,y+\varepsilon_2\mathbf{1})\Big)\\
&+ f_{\tau^l|l}(v \mid x,y)\Big(\alpha_t(\varsigma;t+\varepsilon_3,s,x,v-\varepsilon) - \alpha_v(\varsigma;t,s,x,v-\varepsilon_4)\Big)\\
&+ \varepsilon\mathcal{G}_\varepsilon(v,\varsigma;t,s,x,y),
\end{aligned}$$

where

$$\begin{aligned}
&\mathcal{G}_\varepsilon(v,\varsigma;t,s,x,y)\\
&:= r(x)e^{r(x)\varepsilon_0}\Big(-f'_{\tau^l|l}(v-\varepsilon_1 \mid x,y+\varepsilon\mathbf{1}) + \sum_{i=1}^n \frac{\partial}{\partial y_i} f_{\tau^l|l}(v \mid x,y+\varepsilon_2\mathbf{1})\Big) \times\\
&(\varepsilon\alpha_t(\varsigma;t+\varepsilon_3,s,x,v-\varepsilon) - \varepsilon\alpha_v(\varsigma;t,s,x,v-\varepsilon_4) + \alpha(\varsigma;t,s,x,v))\\
&+ \Big(r(x)e^{r(x)\varepsilon_0} f_{\tau^l|l}(v \mid x,y) - f'_{\tau^l|l}(v-\varepsilon_1 \mid x,y+\varepsilon\mathbf{1}) + \sum_{i=1}^n \frac{\partial}{\partial y_i} f_{\tau^l|l}(v \mid x,y+\varepsilon_2\mathbf{1})\Big)
\end{aligned}$$

$$\times \left(\alpha_t(\varsigma; t+\varepsilon_3, s, x, v-\varepsilon) - \alpha_v(\varsigma; t, s, x, v-\varepsilon_4)\right).$$

We also recall from (5.22) and (5.24) that

$$\alpha_t(\varsigma; t, s, x, v) = \alpha(\varsigma; t, s, x, v)O(\log^2|\varsigma|) \text{ and } \alpha_v(\varsigma; t, s, x, v) = \alpha(\varsigma; t, s, x, v))O(\log^2|\varsigma|),$$

where $|\varsigma| := \max_i |\varsigma_i|$. The expression in (5.26) has two additive terms. For showing convergence of the first term, we intend to use above expressions for applying dominated and Vitali convergence theorem 1.3.1 in various cases. For that, as $\varphi \in \mathcal{B}$, it would be sufficient if we have the following three results,

(a) $v \mapsto \int\limits_{\mathbb{R}^n_+} (c_1^*\varsigma + c_2)\log^2|\varsigma|\alpha(\varsigma; t, s, x, v)d\varsigma$ is bounded and left continuous,

(b) $t \mapsto \int\limits_{\mathbb{R}^n_+} (c_1^*\varsigma + c_2)\log^2(|\varsigma|)\alpha(\varsigma; t, s, x, v)d\varsigma$ is continuous uniformly with respect to v,

(c) $\|\varsigma\|^2\alpha(\varsigma; t+\varepsilon_1, s, x, v+\varepsilon_2)$ is uniform integrable and tight w.r.t. ς for $\varepsilon_1, \varepsilon_2 \ll 1$.

To prove the result (a), we introduce a function $B(v) := \int\limits_{\mathbb{R}^n_+} (c_1^*\varsigma + c_2)\log^2(|\varsigma|)\alpha(\varsigma; t, s, x, v)d\varsigma$.
Now for $\varepsilon > 0$ using the mean value theorem, there exist a $0 < \varepsilon' < \varepsilon$ such that

$$\begin{aligned}\frac{1}{\varepsilon}\left(B(v) - B(v-\varepsilon)\right) &= \int\limits_{\mathbb{R}^n_+} (c_1^*\varsigma + c_2)\log^2(|\varsigma|)\alpha_v(\varsigma; t, s, x, v-\varepsilon')d\varsigma \\ &\leq \int\limits_{\mathbb{R}^n_+} \left(c_3\|\varsigma\|_2^2 + c_4\right)\alpha(\varsigma; t, s, x, v-\varepsilon')d\varsigma,\end{aligned}$$

for some positive constants c_3, c_4. Now Lemma 5.2.1(iii) suggests that the right hand side is bounded in v on $[\varepsilon, T]$. This implies that B is left continuous. Using the similar reasoning the boundedness of B also follows from Lemma 5.2.1(iii). Similarly one can prove the result (b). In order to prove (c), we first recall that a family of normal random variables with bounded mean and variance is uniformly integrable and tight. Therefore (c) follows as here a product of a polynomial and a lognormal density function appears.

Now we address the convergence of the second term of (5.26). Clearly the result (a) implies boundedness of $v \mapsto \int\limits_{\mathbb{R}^n_+} \varphi\left(t+v, \varsigma, R_j^l x, R_0^l(y+v\mathbf{1})\right)\alpha(\varsigma; t, s, x, v)\,d\varsigma$, which assures the desired convergence. Thus $\beta_l \in \operatorname{dom}(D_{t,y})$ and hence $\varphi \in \operatorname{dom}(D_{t,y})$.
Now we discuss the smoothness with respect to s. First we observe that $\alpha_{s^{l'}}(\varsigma; t, s, x, v) = \frac{1}{s^{l'}}O(\log(|\varsigma|))\alpha(\varsigma; t, s, x, v)$. Since $\varphi \in \mathcal{B}$, using uniform integrability and tightness of

$\frac{1}{s^{l'}+\varepsilon}\|\varsigma\|^2\alpha(\varsigma;t,s+\varepsilon,x,v)$ and uniform boundedness of $v \mapsto \int\limits_{\mathbb{R}^n_+} \frac{1}{s^{l'}+\varepsilon}\|\varsigma\|^2\alpha(\varsigma;t,s+\varepsilon,x,v)d\varsigma$ for $\varepsilon \ll 1$, we conclude the differentiability of $\beta_l(t,x,y)$ with respect to s^l. Similarly we can establish existence of partial derivatives of any higher order successively. Thus one can obtain twice continuous differentiability of β_l.

(ii) We have already shown that $A : \mathcal{B} \to \mathcal{B}$ is a contraction. From the form of equation (5.18), and non-negativity of K, it is clear that (5.18) admits a non-negative solution. Since all the coefficients in equation (5.25) are non-negative, it follows that $A\varphi \geq 0$ for $\varphi \geq 0$. Furthermore, we have shown that A has a fixed point in $\mathcal{B}$. It can be easily argued that this fixed point is, in fact, non-negative. Hence, φ is non-negative. □

5.5 Study of The Partial Differential Equation

In this section we establish Theorem 5.3.4, i.e uniqueness and existence of (5.16)-(5.13). Before addressing that it is important to clarify few issues regarding boundary conditions. At $s = 0$ facet the partial derivative with respect to s disappear. Since the nature of the domain is triangular, it can be shown by using the method of characteristic that the initial condition would lead to a solution to (5.16)-(5.13). It can also be shown that the PDE would have no solution if we impose a boundary condition which is not obtain from the initial condition. We refer ([50],pp.32) for more details. Let $\tilde{W}$ be a standard n-dimensional Brownian motion on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$. For each $l = 1, 2, \ldots, n$, let $\tilde{S}^l_t$ satisfies

$$d\tilde{S}^l_t = \tilde{S}^l_t \left[r(X_t)dt + \sum_{j=1}^{n} \sigma^l_j(t, X_t)d\tilde{W}^j_t \right], \qquad \tilde{S}_0 > 0, \tag{5.28}$$

where $\{X_t\}_{t\geq 0}$ is the CSM process given by equations (3.3) on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and σ^l is the lth row of σ. We denote $\tilde{S}_t := (\tilde{S}^1_t, \ldots, \tilde{S}^n_t)$.

Proposition 5.5.1. *(i) The Cauchy problem* (5.16)-(5.13) *has a generalized solution,* φ. *(ii) Under Assumption 3.1(i)-(iii) and Assumption 3.2,* φ *solves the integral equation* (5.25).*(iii)* $\varphi \in \mathcal{B}$.

Proof. (i) Let $\tilde{S}_t$ be the strong solution of the SDE (5.28). Let $\tilde{\mathfrak{F}}_t$ be the filtration generated by $\tilde{S}_t$ and X_t, that satisfies the usual hypothesis. Since (t, X_t, Y_t) is Markov, then the process $(t, \tilde{S}_t, X_t, Y_t)$ is Markov process. Let $\mathcal{A}$ be the infinitesimal generator of $(t, \tilde{S}_t, X_t, Y_t)$, where

$$\begin{aligned}
&\mathcal{A}\varphi(t,s,x,y) \\
&= D_{t,y}\varphi(t,s,x,y) + r(x)\sum_{l=1}^{n} s^l \frac{\partial\varphi}{\partial s^l}(t,s,x,y) + \frac{1}{2}\sum_{l=1}^{n}\sum_{l'=1}^{n} a^{ll'}(t,x^l)s^l s^{l'} \frac{\partial^2\varphi}{\partial s^l \partial s^{l'}}(t,s,x,y)
\end{aligned}$$

$$+\sum_{l=0}^{n}\sum_{j\neq x^l}\lambda^l_{x^l j}(y^l)\left[\varphi(t,s,R^l_j x,R^0_l y)-\varphi(t,s,x,y)\right], \tag{5.29}$$

for every function φ which is compactly supported C^2 in s and C^1 in y. Let

$$N_t := \mathbb{E}[e^{-\int_t^T r(X_u)du}K(\tilde{S}_T)\mid \tilde{S}_t=s,X_t=x,Y_t=y]. \tag{5.30}$$

The above expectation is finite due to Assumption 5.2(ii) and Lemma 5.2.2. Thus (5.30) suggests that N_t is a $\tilde{\mathfrak{F}}_t$ martingale. Since $K(s)$ has at-most linear growth, and $\tilde{S}_t$ has finite expectation, (5.30) suggests that $E|N_t|<\infty$ for each t. Hence using the Markov semigroup of $(t,\tilde{S}_t,X_t,Y_t)$ the PDE has a generalized solution $\varphi:\mathcal{D}\to\mathbb{R}$ measurable given by

$$\varphi(t,s,x,y) := \mathbb{E}[e^{-\int_t^T r(X_u)du}K(\tilde{S}_T)\mid \tilde{S}_t=s,X_t=x,Y_t=y]. \tag{5.31}$$

(ii) By conditioning (5.31) on transition times, we get

$$\begin{aligned}
&\varphi(t,\tilde{S}_t,X_t,Y_t)\\
&= \mathbb{E}\left[\mathbb{E}\left[e^{-\int_t^T r(X_u)\,du}K(\tilde{S}_T)\mid \tilde{S}_t,X_t,Y_t,l(t)=l\right]\mid \tilde{S}_t,X_t,Y_t\right]\\
&= \sum_{l=0}^{n}P_{t,x,y}(\ell(t)=l)\;\mathbb{E}\left[e^{-\int_t^T r(X_u)\,du}K(\tilde{S}_T)\mid \tilde{S}_t,X_t,Y_t,l(t)=l\right]\\
&= \sum_{l=0}^{n}P_{t,x,y}(\ell(t)=l)\;\mathbb{E}\left[\mathbb{E}\left[e^{-\int_t^T r(X_u)\,du}K(\tilde{S}_T)\mid \tilde{S}_t,X_t,Y_t,l(t)=l,\tau^l(t)\right]\mid \tilde{S}_t,X_t,Y_t,l(t)=l\right].
\end{aligned}$$

Now,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[e^{-\int_t^T r(X_u)\,du}K(\tilde{S}_T)\mid \tilde{S}_t,X_t,Y_t,l(t)=l,\tau^l(t)\right]\mid \tilde{S}_t,X_t,Y_t,l(t)=l\right]\\
&= P[\tau^l(t)>T-t]\rho_x(t,\tilde{S}_t)\\
&+\int_0^{T-t}\mathbb{E}\left[e^{-\int_t^T r(X_u)\,du}K(\tilde{S}_T)\mid \tilde{S}_t,X_t,Y_t,l(t)=l,\tau^l(t)=v\right]f_{\tau^l|l}(v\mid X_t,Y_t)\,dv.
\end{aligned}$$

We note that

$$\begin{aligned}
&\mathbb{E}\left[e^{-\int_t^T r(X_u)\,du}K(\tilde{S}_T)\mid \tilde{S}_t,X_t,Y_t,l(t)=l,\tau^l(t)=v\right]\\
&= e^{-r(X_t)v}\sum_{j^l\neq X^l}p^l_{X^l j^l}(Y^l+v)\int_{\mathbb{R}^n_+}\mathbb{E}\left[e^{-\int_{t+v}^T r(X_u)\,du}K(\tilde{S}_T)\mid \tilde{S}_{t+v}=\varsigma,X_{t+v}=R^l_j x,\right.\\
&\qquad \left.Y_{t+v}=R^l_0 y,l(t)=l,\tau^l(t)=v\right]\alpha(\varsigma;t,s,x,v)\,d\varsigma.
\end{aligned}$$

Therefore

$$\varphi(t,\tilde{S}_t,X_t,Y_t)$$

$$= \sum_{l=0}^{n} P_{t,x,y}\left(\ell(t)=l\right)\left(\rho_x(t,\tilde{S}_t)\left(1-F_{\tau^l|l}(T-t\mid X_t,Y_t)\right)+\int_0^{T-t} e^{-r(X_t)v} f_{\tau^l|l}(v\mid X_t,Y_t)\times\right.$$

$$\left.\sum_{j\neq x^l} p^l_{x^l j}(y^l+v)\int_{\mathbb{R}^n_+}\varphi\left(t+v,R^l_{\varsigma^l}s,R^l_j x,R^l_0 y\right)\alpha(\varsigma;t,s,x,v)\,d\varsigma\,dv\right).$$

Using Assumption 3.2, and since $\lambda^l_{ij}(y)>0$ for $i\neq j$, we can replace $(\tilde{S}_t,X_t,Y_t)$ by the generic variable (s,x,y) in the above relation. As a conclusion, φ is a solution of (5.25).
(iii) To show φ is of at-most linear growth, it is sufficient to show for all $(t,s,x,y)\in\mathcal{D}$ $|\varphi(t,s,x,y)-c_1^*s|\leq c_2$, where c_1,c_2 is as in Assumption 5.2(ii). We note that, if $\tilde{S}_t$ is the solution of (5.28), $e^{-\int_0^t r(X_u)du}\tilde{S}_t$ is a $\tilde{\mathfrak{F}}_t$ martingale. Therefore by using the Markov property of $\tilde{S}_t,X_t,Y_t$, and the fact $e^{-\int_0^t r(X_u)du}$ is $\tilde{\mathfrak{F}}_t$-measurable, we obtain

$$\begin{aligned}\mathbb{E}\left[e^{-\int_t^T r(X_u)du}\tilde{S}_T\middle|\tilde{S}_t,X_t,Y_t\right]=&\mathbb{E}\left[e^{-\int_t^T r(X_u)du}\tilde{S}_T\middle|\tilde{\mathfrak{F}}_t\right]\\ =&e^{-\int_0^t r(X_u)du}\mathbb{E}\left[e^{-\int_0^T r(X_u)du}\tilde{S}_T\middle|\tilde{\mathfrak{F}}_t\right]\\ =&\tilde{S}_t.\end{aligned}$$

Using this equality, (5.31) and Assumption 5.2(ii), we have

$$\begin{aligned}&|\varphi(t,s,x,y)-c_1^*s|\\ &=\left|\mathbb{E}\left[e^{-\int_t^T r(X_u)du}K(\tilde{S}_T)\mid\tilde{S}_t=s,X_t=x,Y_t=y\right]\right.\\ &\left.-c_1^*\mathbb{E}\left[e^{-\int_t^T r(X_u)du}\tilde{S}_T\mid\tilde{S}_t=s,X_t=x,Y_t=y\right]\right|\\ &\leq\mathbb{E}\left[[e^{-\int_t^T r(X_u)du}|K(\tilde{S}_T)-c_1^*\tilde{S}_T|\mid\tilde{S}_t=s,X_t=x,Y_t=y\right]\\ &\leq c_2.\end{aligned}$$

This completes the proof. □

Proof of Theorem 5.3.4: Proposition 5.5.1 implies that the PDE (5.16)-(5.13) has a generalized solution (see equation (1.18) from [51] for a definition) which is in $\mathcal{B}$, and also solves the integral equation. Lemma 5.4.2 suggests that the integral equation has only one solution in $\mathcal{B}$. Finally Lemma 5.4.4 asserts that this unique solution of the integral equation is in $\mathrm{dom}(D_{t,y})\cap C^2_s$. Therefore using the above results, we conclude that (5.16)-(5.13) has a generalized solution which is in the domain of the operators in (5.16). Hence the generalized solution (5.31) solves (5.16)-(5.13) classically. To prove the uniqueness, first assume that φ_1 and φ_2 are two classical solutions of (5.16)-(5.13) in the prescribed class of functions. Then using Proposition 5.5.1, it follows that both also solve (5.25). By Lemma 5.4.2, there is only one such solution in the prescribed class. Hence $\varphi_1=\varphi_2$. □

Lemma 5.5.2. *Let $\varphi(t, s, x, y)$ be the classical solution of the Cauchy problem* (5.16)-(5.13). *Under Assumption 5.2(i), $\frac{\partial \varphi}{\partial s^m}(t, s, x, y)$ is bounded.*

Proof. Since $\varphi(t, s, x, y)$ is the classical solution of (5.16)-(5.13) it is in $\text{dom}(D_{t,y}) \cap C_s^2$. In fact φ has greater regularity than C_s^2 which is evident in the proof of Lemma 5.4.4. Indeed due to Lemma 5.4.1(iii) and the C^∞ smoothness of ρ, φ is C^∞ in s. Let $\psi^m(t, s, x, y) := \frac{\partial \varphi}{\partial s^m}(t, s, x, y)$, for $m = 1, \dots, n$. Now differentiating equation (5.16) with respect to s^m and using the fact that $a(t, x)$ is symmetric, we obtain

$$\begin{aligned} & D_{t,y}\psi^m(t, s, x, y) + \sum_{l=1}^{n} s^l \left(r(x) + a^{ml}(t, x) \right) \frac{\partial \psi^m}{\partial s^l}(t, s, x, y) \\ & + \frac{1}{2} \sum_{l=1}^{n} \sum_{l'=1}^{n} a^{ll'}(t, x) s^l s^{l'} \frac{\partial^2 \psi^m}{\partial s^l \partial s^{l'}}(t, s, x, y) \\ & + \sum_{l=0}^{n} \sum_{j \neq x^l} \lambda_{x^l j}^l(y^l) \left[\psi^m(t, s, R_j^l x, R_0^l y) - \psi^m(t, s, x, y) \right] = 0. \end{aligned} \tag{5.32}$$

It is easy to check that

$$\begin{aligned} & \hat{A}\psi^m(t, s, x, y) \\ & = D_{t,y}\psi^m(t, s, x, y) + \sum_{l=1}^{n} s^l \left(r(x) + a^{ml}(t, x) \right) \frac{\partial \psi^m}{\partial s^l}(t, s, x, y) \\ & + \frac{1}{2} \sum_{l=1}^{n} \sum_{l'=1}^{n} a^{ll'}(t, x) s^l s^{l'} \frac{\partial^2 \psi^m}{\partial s^l \partial s^{l'}}(t, s, x, y) \\ & + \sum_{l=0}^{n} \sum_{j \neq x^l} \lambda_{x^l j}^l(y^l) \left[\psi^m(t, s, R_j^l x, R_0^l y) - \psi^m(t, s, x, y) \right], \end{aligned}$$

is the infinitesimal generator of the Markov process $(t, \bar{S}_t, X_t, Y_t)$, where $\bar{S}_t = (\bar{S}_t^1, \dots, \bar{S}_t^n)$ and $\bar{S}_t^l$ satisfies the following SDE

$$d\bar{S}_t^l = \bar{S}_t^l \left[\left(r(X_t)I + \text{Diag}(a^l(t, X_t) \right) dt + \sigma(t, X_t) dW_t \right], \tag{5.33}$$

where $\text{Diag}(a^l(t, X_t)$ is the diagonal matrix containing the l-th row of $a(t, x)$ and (X_t, Y_t) is as in (5.1). Therefore the solution of the PDE (5.32) has the stochastic representation of the following form

$$\psi^m(t, s, x, y) = \mathbb{E}\left[K'(\bar{S}_T^m) \middle| \bar{S}_t^m = s, X_t = x, Y_t = y \right], \tag{5.34}$$

where $K' : \mathbb{R}_+^n \to \mathbb{R}$ is defined almost everywhere by $K(s) = \int_0^{s^m} K'(R_r^m s) dr$, for each $s \in \mathbb{R}_+^n$ Since K is of at-most linear growth and it is Lipschitz continuous, K' is in L^∞. Hence (5.34) suggests $\psi^m(t, s, x, y)$ is bounded. □

5.6 Locally risk minimizing pricing and optimal hedging

Proof of Theorem 5.3.5: Using Lemma 5.5.2 we can show that $\pi = (\xi, \varepsilon)$ as given in (5.17) is an admissible portfolio strategy. Indeed ξ_t^l is left continuous and therefore predictable. Hence Assumption 5.2(i) and (ii) holds for this pair $\pi = (\xi, \varepsilon)$. Therefore the discounted value function for this pair of strategy using (5.17) is given by

$$\hat{V}_t(\pi) = \sum_{l=1}^{n} \xi_t^l \hat{S}_t^l + \varepsilon_t = e^{-\int_0^t r(X_u)\,du} \varphi(t, S_t, X_t, Y_t),$$

where φ is the unique classical solution of (5.16)-(5.13). Now we shall find a decomposition for $\hat{V}_t(\pi)$. Under the measure P, we apply Itō's formula to

$$e^{-\int_0^t r(X_u)\,du} \varphi(t, S_t, X_t, Y_t).$$

Using (5.8), (5.16) and (4.1) and after a suitable rearrangement of terms, for all $t < T$, we obtain,

$$\begin{aligned}
&e^{-\int_0^t r(X_u)\,du} \varphi(t, S_t, X_t, Y_t) \\
&= \varphi(0, S_0, X_0, Y_0) + \sum_{l=1}^{n} \int_0^t \frac{\partial \varphi}{\partial s^l}(u, S_u, X_{u-}, Y_{u-}) d\hat{S}_u^l \\
&+ \int_0^t e^{-\int_0^u r(X_v)\,dv} \int_{\mathbb{R}} [\varphi(u, S_u, X_{u-} + h(X_{u-}, Y_{u-}, z), Y_{u-} - g(X_{u-}, Y_{u-}, z)) \\
&- \varphi(u, S_u, X_{u-}, Y_{u-})] \hat{\wp}(du, dz),
\end{aligned} \tag{5.35}$$

where $\hat{\wp}$ is the compensator of $\wp$, i.e. $\hat{\wp}(dt, dz) = \wp(dt, dz) - dtdz$. Therefore from (5.35), we have for each $t \leq T$

$$\frac{1}{S_t^0} \varphi(t, S_t, X_{t-}, Y_{t-}) = H_0 + \sum_{l=1}^{n} \int_0^t \xi_u^l d\hat{S}_u^l + L_t, \tag{5.36}$$

where $H_0 = \varphi(0, S_0, X_0, Y_0)$ and

$$\begin{aligned}
L_t \;:=\; &\int_0^t e^{-\int_0^u r(X_v) dv} \int_{\mathbb{R}} [\varphi(u, S_u, X_{u-} + h(X_{u-}, Y_{u-}, z), Y_{u-} - g(X_{u-}, Y_{u-}, z)) \\
&- \varphi(u, S_u, X_{u-}, Y_{u-})] \hat{\wp}(du, dz).
\end{aligned} \tag{5.37}$$

Clearly the above choice of H_0 is $\mathfrak{F}_0$ measurable and L_T is $\mathfrak{F}_T$ measurable. We know that, the integral with respect to a compensated Poisson random measure is a local martingale.

Hence L_t is a local martingale. The proof of Proposition 5.5.1(iii) suggests that expectation of supremum of L_t is finite. Hence by using Theorem 1.1.11, it is a martingale. Again since W_t and $\wp$ are independent, L_t is orthogonal to $\int_0^t \sigma^l(t, X_t)\hat{S}_t dW_t$. Thus, we obtain the following F-S decomposition by letting $t \uparrow T$ in (5.36),

$$S_T^{0\,-1} K(S_T) = \varphi(0, S_0, X_0, Y_0) + \sum_{l=1}^{n} \int_0^T \xi_t^l d\hat{S}_t^l + L_T. \tag{5.38}$$

This completes the proof. □

Theorem 5.6.1. *Let φ be the unique solution of* (5.25)*. Set*

$$\eta(t,s,x,y) := \sum_{l=0}^{n} P_{t,x,y}(\ell(t)=l) \Bigg(\frac{\partial \rho_x(t,s)}{\partial s^m} \Big(1 - F_{\tau^l|l}(T-t \mid x,y)\Big) + \int_0^{T-t} e^{-r(x)v} f_{\tau^l|l}(v|x,y)$$
$$\times \sum_{j^l \neq x^l} p^l_{x^l j^l}(y^l+v) \int_{\mathbb{R}_+^n} \varphi(t+v, \varsigma, R_j^l x, R_l^0 y) \frac{\partial \alpha(\varsigma; t,s,x,v)}{\partial s^m}\, d\varsigma\, dv \Bigg), \tag{5.39}$$

where $(t,s,x,y) \in \mathcal{D}$*. Then* $\eta(t,s,x,y) = \frac{\partial \varphi}{\partial s^m}(t,s,x,y)$*,*

Proof. We need to show that ψ (as in (5.39)) is equal to $\frac{\partial \varphi}{\partial s^m}$. Indeed, one obtains the RHS of (5.39) by differentiating the right side of (5.25) with respect to s^m. Hence the proof is completed. □

Remark 5.6.2. *We have shown in Theorem 5.3.5 that $\frac{\partial \varphi}{\partial s^m}(t,s,x,y)$ is a necessary quantity to be calculated in order to find the optimal hedging. Attempting to compute $\frac{\partial \varphi}{\partial s^m}(t,s,x,y)$ using numerical differentiation would increase the sensitivity of $\frac{\partial \varphi}{\partial s^m}(t,s,x,y)$ to small errors. Equation* (5.39) *gives a better, more robust approach for computing $\frac{\partial \varphi}{\partial s^m}(t,s,x,y)$, using numerical integration.*

5.7 Sensitivity with respect to the instantaneous rate function

In a recent paper Goswami et al. [28] gave an interesting idea to approximate the solution by approximating the transition rate. In the previous section we have seen that for a class of continuously differentiable transition rate function, there exists a unique classical solution of the PDE (5.16)-(5.13). Let $\lambda := (\lambda^0, \ldots, \lambda^n)$ be a vector where λ^l is as in section 5.2. We state and prove the important result below.

Theorem 5.7.1. *Let φ and $\tilde{\varphi}$ be two solutions of* (5.16)-(5.13) *with parameter λ and $\tilde{\lambda}$ respectively. Then $\|\varphi - \tilde{\varphi}\|_{sup} \leq 2c_2 T \|\lambda - \tilde{\lambda}\|_{sup}$, where c_2 as in Assumption 5.2(ii).*

Proof. We consider

$$\psi(t,s,x,y) := \varphi(t,s,x,y) - \tilde{\varphi}(t,s,x,y). \tag{5.40}$$

Now, it is easy to see that ψ satisfies the following initial value problem,

$$\begin{aligned}
&D_{t,y}\psi(t,s,x,y) + r(x)\sum_{l=1}^{n} s^l \frac{\partial \psi}{\partial s^l}(t,s,x,y) + \frac{1}{2}\sum_{l=1}^{n}\sum_{l'=1}^{n} a^{ll'}(t,x)s^l s^{l'} \frac{\partial^2 \psi}{\partial s^l \partial s^{l'}}(t,s,x,y) \\
&+ \sum_{l=0}^{n}\sum_{j\neq x^l} \lambda^l_{x^l j}(y^l)\Big(\psi(t,s,R^l_j x, R^l_0 y) - \psi(t,s,x,y)\Big) \\
&= r(x)\psi(t,s,x,y) - \sum_{l=0}^{n}\sum_{j\neq x^l}\Big(\lambda^l_{x^l j}(y^l) - \tilde{\lambda}^l_{x^l j}(y^l)\Big)\Big(\tilde{\varphi}(t,s,R^l_j x, R^l_0 y) - \tilde{\varphi}(t,s,x,y)\Big),
\end{aligned} \tag{5.41}$$

defined on

$$\mathcal{D} := \{(t,s,x,y) \in (0,T) \times \mathbb{R}^n_+ \times \mathcal{X}^{n+1} \times (0,T)^{n+1} | y \in (0,t)^{n+1}\},$$

with condition

$$\psi(T,s,x,y) = 0, \quad s \in \mathbb{R}^n_+; \quad 0 \leq y^l \leq T; \quad x = 1,2,\cdots,k.$$

We rewrite (5.41) using (5.40) as

$$\mathcal{A}\psi(t,s,x,y) = r(x)\psi(t,s,x,y) - f(t,s,x,y), \tag{5.42}$$

where

$$f(t,s,x,y) := \sum_{l=0}^{n}\sum_{j\neq x^l}\Big(\lambda^l_{x^l j}(y^l) - \tilde{\lambda}^l_{x^l j}(y^l)\Big)\Big(\tilde{\varphi}(t,s,R^l_j x, R^l_0 y) - \tilde{\varphi}(t,s,x,y)\Big).$$

We recall that $\mathcal{A}$ is the infinitesimal generator of $(t, \tilde{S}_t, X_t, Y_t)$. Using the proof of Proposition 5.5.1(iii), one can show that for all $(t,s,x,y) \in \mathcal{D}$

$$|f(t,s,x,y)| \leq 2c_2 \sum_{l=0}^{n}\sum_{j\neq x^l} \|\lambda^l_{x^l j}(y) - \tilde{\lambda}^l_{x^l j}(y)\|_{\text{sup}}. \tag{5.43}$$

The stochastic representation of the solution of the PDE (5.42) is given by,

$$\psi(t,s,x,y) = \mathbb{E}\int_t^T \exp\left(-\int_t^v r(X_u)du\right) f(v,\tilde{S}_v,X_v,Y_v)dv | \tilde{S}_t = s, X_t = x, Y_t = y]. \tag{5.44}$$

Since $\tilde{\varphi}$ is a solution of (5.16)-(5.13) for parameter $\tilde{\lambda}$, then the proof of Proposition 5.5.1(iii), $|\tilde{\varphi}(t,s,R^l_j x, R^l_0 y) - \tilde{\varphi}(t,s,x,y)| < 2c_2$. Now using (5.43) and $r > 0$ for all $t \leq v \leq T$, we have

$$\begin{aligned}
\|\psi(t,s,x,y)\|_{\sup} &= \sup_{\tilde{\mathcal{D}}} \left| \mathbb{E} \int_t^T \exp\left(-\int_t^v r(X_u)du \right) f(v, \tilde{S}_v, X_v, Y_v) dv | \tilde{S}_t = s, X_t = x, Y_t = y] \right| \\
&\leq 2c_2(T-t) \sum_{l=0}^{n} \sum_{j \neq x^l} \|\lambda^l_{x^l j}(y) - \tilde{\lambda}^l_{x^l j}(y)\|_{\sup} \\
&< 2c_2 T \sum_{l=0}^{n} \sum_{j \neq x^l} \|\lambda^l_{x^l j}(y) - \tilde{\lambda}^l_{x^l j}(y)\|_{\sup}.
\end{aligned}$$

Hence the proof is completed. □

Remark 5.7.2. *It is interesting to note that a weaker variant of Theorem 5.7.1 can also be proved if the Assumption 5.2(ii) is relaxed. Indeed if $K \in \mathcal{B}$ for such case $\|\varphi - \tilde{\varphi}\|_L \leq M\|\lambda - \tilde{\lambda}\|_{sup}$. This readily follows from the fact that $\tilde{\varphi}$ is of at most linear growth and $\tilde{S}_t$ has finite expectation.*

5.8 Calculation of the Quadratic Residual Risk

In this section we find an expression of the quadratic residual risk corresponding to the optimal strategy. Let $\xi_t := (\xi^0_t, \ldots, \xi^n_t)$ be the optimal strategy and V_t be the corresponding value process as defined in Section 5.3. Further we assume that $\{C_t\}_{t\geq 0}$ is the accumulated additional cash flow process associated with the optimal hedging of the contingent claim H, where

$$dC_t := dV_t - \sum_{l=0}^{n} \xi^l_t dS^l_t.$$

Since $\{S^0_t\}_{t\geq 0}$ is continuous finite variation process and $V_t = \sum_{l=0}^n \xi^l_t S^l_t$, from above we get

$$\begin{aligned}
\frac{1}{S^0_t} dC_t &= \left(d\left(\frac{V_t}{S^0_t}\right) - V_t\, d\left(\frac{1}{S^0_t}\right)\right) - \left(\sum_{l=0}^{n} \xi^l_t\, d\left(\frac{S^l_t}{S^0_t}\right) - \sum_{l=0}^{n} \xi^l_t S^l_t\, d\left(\frac{1}{S^0_t}\right)\right) \\
&= d\hat{V}_t - \sum_{l=1}^{n} \xi^l_t d\hat{S}^l_t, \qquad (5.45)
\end{aligned}$$

where $\hat{V}_t$ is the discounted value process as defined in Section 5.3. Now by (5.11), we have

$$d\hat{V}_t = \sum_{l=1}^{n} \xi^l_t d\hat{S}^l_t + dL^{\hat{H}}_t. \qquad (5.46)$$

Now by comparing (5.45) and (5.46), we have

$$\frac{1}{S_t^0} dC_t = dL_t^{\hat{H}}.$$

The discounted value of the accumulated cash flow during $[0, t]$ is

$$\hat{C}_t - \hat{C}_0 := \int_0^t \frac{1}{S_t^0} \, dC_t = L_t^{\hat{H}}.$$

Again, using above and (5.37), we get

$$\int_0^t \frac{1}{S_t^0} \, dC_t = \int_0^t \frac{1}{S_t^0} \sum_{l=0}^{n} \int_{\mathbb{R}} \Big[\varphi\left(t, S_t, R^l_{\{X_{t-}+h(X_{t-},Y_{t-},z)\}}(X_{t-}), R^l_{\{Y_{t-}-g(X_{t-},Y_{t-},z)\}}(Y_{t-})\right) \\ -\varphi(t, S_t, X_{t-}, Y_{t-})\Big] \, \hat{\wp}^l(dt, dz)$$

for all $t \in [0, T]$. Thus

$$dC_t = \sum_{l=0}^{n} \int_{\mathbb{R}} \Big[\varphi\left(t, S_t, R^l_{\{X_{t-}+h(X_{t-},Y_{t-},z)\}}(X_{t-}), R^l_{\{Y_{t-}-g(X_{t-},Y_{t-},z)\}}(Y_{t-})\right) \\ - \varphi(t, S_t, X_{t-}, Y_{t-})\Big] \, \hat{\wp}^l(dt, dz). \tag{5.47}$$

Integrating the above expression, we obtain the external cash flow associated with the optimal hedging. Hence,

$$C_T = C_0 + \sum_{t\in[0,T]} (\varphi(t, S_t, X_t, Y_t) - \varphi(t, S_t, X_{t-}, Y_{t-})) - \int_0^T \sum_{l=0}^{n} \sum_{j \neq X^l_{t-}} \lambda^l_{X^l_{t-}j}(Y^l_{t-}) \times \\ \Big[\varphi(t, S_t, R^l_j(X_{t-}), R^l_0(Y_{t-})) - \varphi(t, S_t, X_{t-}, Y_{t-})\Big] \, dt. \tag{5.48}$$

Lemma 5.8.1. *The quadratic variation process of C_t, the cash flow process corresponding to the optimal hedge, is given by*

$$[C]_t = \sum_{r\in[0,t]} (\varphi(r, S_r, X_r, Y_r) - \varphi(r, S_r, X_{r-}, Y_{r-}))^2 , \tag{5.49}$$

where φ is the unique classical solution of (5.16)-(5.13) *with at most linear growth.*

Proof. It is clear that $\{C_t\}_{t\geq 0}$ as in (5.48) is an *rcll* process. Now, for $r \in (0, T)$ and for sufficiently small Δ, we have

$(C_r - C_{r-\Delta})^2$

$$= (\varphi(r,S_r,X_r,Y_r) - \varphi(r,S_r,X_{r-\Delta},Y_{r-\Delta}))^2 - 2\left(\varphi(r,S_r,X_r,Y_r) - \varphi(r,S_r,X_{r-\Delta},Y_{r-\Delta})\right) \times$$
$$\sum_{l=0}^{n} \sum_{j\neq X^l_{r-\Delta}} \lambda^l_{X^l_{r-\Delta}j}(Y^l_{r-\Delta}) \left[\varphi(r,S_r,R^l_j X_{r-\Delta},R^l_0 Y_{r-\Delta}) - \varphi(r,S_r,X_{r-\Delta},Y_{r-\Delta})\right] \Delta$$
$$+ \left(\sum_{l=0}^{n} \sum_{j\neq X^l_{r-\Delta}} \lambda^l_{X^l_{r-\Delta}j}(Y^l_{r-\Delta}) \left[\varphi(r,S_r,R^l_j X_{r-\Delta},R^l_0 Y_{r-\Delta} - \varphi(r,S_r,X_{r-\Delta},Y_{r-\Delta})\right]\right)^2 \Delta^2$$
$$+ O\left(\Delta^2\right).$$

Since the quadratic variation of C_t is the limit of the sum $\sum_{r\in[0,t]}(C_r - C_{r-\Delta})^2$ over a partition with $\Delta \to 0$, we take the summation both sides. We note that the second term on right, the multiplier of Δ is bounded and is of $O(\Delta)$ except the event of whose probability is $O(\Delta)$. Thus the summation of second, third and fourth terms in the above expression can be ignored. Hence,

$$[C]_t = \sum_{r\in[0,t]} \left[\varphi(r,S_r,X_r,Y_r) - \varphi(r,S_r,X_{r-},Y_{r-})\right]^2. \tag{5.50}$$

□

Given a strategy ξ, the *quadratic residual risk* at $t = 0$, denoted by $\mathcal{R}_0(\xi)$, is defined as $\mathcal{R}_0(\xi) := \mathbb{E}[(\hat{C}_T - \hat{C}_0)^2|\mathcal{F}_0]$. An expression for $\mathcal{R}_0(\pi)$ can be found using Itō's isometry in the following way

$$\mathcal{R}_0(\pi) = \mathbb{E}[(\hat{C}_T - \hat{C}_0)^2|\mathcal{F}_0] = E\left[\left(\int_0^T \frac{1}{S^0_t}\, dC_t\right)^2 \,\middle|\, \mathcal{F}_0\right] = E\left[\int_0^T \frac{1}{S^{0\,2}_t}\, d[C]_t \,\middle|\, \mathcal{F}_0\right].$$

Thus using the above and Lemma 5.8.1, we get

$$\begin{aligned}
\mathcal{R}_0(\pi) =& E\left[\sum_{t\in[0,T]} \frac{1}{S^{0\,2}_t} \left(\varphi(t,S_t,X_t,Y_t) - \varphi(t,S_t,X_{t-},Y_{t-})\right)^2 \,\middle|\, \mathcal{F}_0\right] \\
=& E\left[\sum_{t\in[0,T]} \left(\hat{\varphi}(t,S_t,X_t,Y_t) - \hat{\varphi}(t,S_t,X_{t-},Y_{t-})\right)^2 \,\middle|\, \mathcal{F}_0\right] \\
=& E\left[\sum_{l=1}^{n} \sum_{i=1}^{n^l(T)} \left(\hat{\varphi}(T^l_i,S_{T^l_i},X_{T^l_i},Y_{T^l_i}) - \hat{\varphi}(T^l_i,S_{T^l_i},X_{T^l_i-},Y_{T^l_i-})\right)^2 \,\middle|\, \mathcal{F}_0\right].
\end{aligned} \tag{5.51}$$

5.9 Conclusion

In this chapter, we have studied the locally risk-minimizing pricing of European basket options, under a market where, we allow the drift and volatility coefficients to be time inhomogeneous and CSM modulated. In [25], the above problem was studied assuming all the coefficients are driven by a single semi-Markov process. It is known that, unlike Markov chains, two independent semi-Markov processes is not semi-Markov jointly. Thus the extension to CSM is essential in view of the non-identical regimes of different assets. It is shown in [25], that the option price can be expressed in terms of a price function which depends on the present stock price, present regime, and the sojourn time at the current regime. In this thesis, we have shown that even under the extended CSM setting, option price does have similar representation involving knowledge of each component of the regimes. However, the study of well-posedness of the option price equation turns out to be more involved.

It should be noted that the derivations in this chapter is different from the standard approach. Here we start with a Cauchy problem which we show to possess a classical solution. We then construct a hedging strategy using the first order partial derivatives of the solution so as to obtain Föllmer Schweizer decomposition of contingent claim related to a European option. From the decomposition we conclude that the solution to the Cauchy problem is indeed the locally risk minimizing price of the corresponding European option. This approach avoids an a-priori tacit assumption of desired differentiability of the option price function that is expressed using a conditional expectation with respect to an equivalent minimal martingale measure.

Besides, we have obtained an integral equation of option price. We have also expressed the risk-minimizing hedging strategy as a combination of delta hedge and an integral of the price function. Expressions for the quadratic residual risk and the optimal external cash-flow are obtained at the end. The contents of this chapter is from [10].

6

Conclusion

This thesis concerns the component-wise semi Markov (CSM) modulated market. The main contribution to the thesis consists of three chapters. The first chapter contains an empirical study to validate the appropriateness of the use of CSM process to model the risky asset price dynamics. The rest of the thesis is about the applications of CSM market setup in a portfolio optimization and a pricing problem.

In the first part, we propose a test statistics to examine the usefulness of one component CSM process. We build our test statistics by using the descriptive statistics of the squeeze duration of the famous Bollinger band. In an empirical survey on Indian stock indices, we show using this test statistics that we can not reject our null hypothesis that the asset price dynamics are governed by a semi-Markov process whereas the special subclass namely GBM and the Markov modulated GBM can be rejected. But to limit time complexity of the problem, we restrict our experiment by keeping the drift parameter insensitive to the regime and by equating the volatility coefficient at the first regime to a low percentile of the historical volatility.

In the second part, we solve an optimal investment problem using a utility function based on the risk tolerance of the investor in a CSM modulated multi-dimensional jump diffusion market. Using a stochastic control approach, we identify the value function as a solution of a system of non-local partial differential equation. We use a probabilistic technique to establish the well-posedness of the partial differential equation.

We address a European type basket option pricing problem in a CSM modulated geometric Brownian motion at the end of the thesis. In this part, the drift and volatility coefficients are considered to be time inhomogeneous. Since the market is incomplete, we employ a locally risk minimizing pricing approach. We identify the option price equation as a classical solution of Black-Scholes-Merton (BSM) type partial differential equation. We establish the well-posedness of the PDE equation by using a probabilistic technique. Finally, we calculate the quadratic residual of risk.

It is important to note that the infinite horizon counterpart of the risk sensitive control problem, mentioned above, is open. In a recent work [8], a ground work appears where the controlled problem is studied in one dimension. The question about multidimensional setting is yet unanswered. For modeling purpose, continuous time semi-Markov process is a natural object for many applications including multi class queue with general arrival. In view of this the need for extension of the results in [8] is evident. As per our knowledge, no such study is done yet.

In a recent paper, Biswas et al. [4] have studied the option pricing in a semi-Markov modulated stochastic volatility model. Let $\{X_t\}_{t\geq 0}$ be a semi-Markov process with state space $\{1, 2, \ldots, k\}$. In their work, they have considered one locally risk free asset with spot rate $r(X_t)$ and a risky asset whose price is given by the following stochastic differential equation

$$\begin{aligned} dS_t &= S_t(\mu(X_t)\,dt + \sqrt{V_t}\,dW_t^1),\ \ S_0 > 0 \\ dV_t &= \kappa(X_t)(\theta(X_t) - V_t)dt + \sigma(X_t)\sqrt{V_t}\,dW_t^2,\ \ V_0 > 0, \end{aligned} \tag{6.1}$$

where the market parameters μ, σ, κ and θ are driven by same semi-Markov process $\{X_t\}_{t\geq 0}$. This problem would be more interesting if the market parameters r, μ, σ, κ and θ are governed by a CSM process. The barrier option pricing under CSM modulated market needs to be carried out.

There might be another research direction for enhancing the implementability of the above model. That is including a transaction cost. We have not incorporated the transaction costs in the problem of option pricing. The hedging strategy in our thesis requires continuous trading. Therefore the investor will incur heavy expanses by implementing those strategies in a market with transaction costs. The option pricing with transaction costs via a utility optimization approach in a GBM market is studied by [1]. However this problem is open in semi-Markov modulated GBM market.

A

Appendix

A.1 Algorithms

In this section we present the algorithms used in the empirical study in Chapter 2. We begin this section with the pseudocode of simple return as in (2.2).

Algorithm 1: Simple return of a time series

```
function SimpleReturn (S);
Input  : Time series data of closing prices S
Output: Another series of simple return R
set N ← length(S)
let R[1...N] be a new array
for i = 2 to N do
    R[i] ← (S[i]−S[i−1]) / S[i−1]
end
return R
```

Although the algorithms to calculate the sample standard deviation is well known, we also present the algorithm for the same.

Algorithm 2: Sample standard deviation

```
function Stdev (a);
Input   : An array a
Output: Sample standard deviation
set n ← length(a); m ← 0; v ← 0
for i = 1 to n do
    m ← m + a[i]
end
m ← m/n
for i = 1 to n do
    v ← v + (a[i] − m)^2
end
return sqrt(v/(n−1))
```

Now we present the algorithm to compute the p-squeeze duration as in Definition 2.2.6.

Algorithm 3: Squeeze duration

```
function SqD (p, S);
Input  : A number p ∈ (0, 1) and a time series S
Output: p-Squeeze duration
set N ← length(S)
R ← SimpleReturn(S)
let v[1 ... N] be a new array
for i = 1 to N do
    a ← R[i ... i + n − 1]
    v[i] ← Stdev(a)
end
v ← Sorted(v) \\ create a sorted list
q ← Celing(p ∗ length(v)) \\ the ordinal rank
per ← v[q] \\ number from the list
let T[1 ... N] be a new array
set flag ← 1
set l ← 1
for j = 1 to length(v) do
    if flag ∗ v[j] < flag ∗ per then
        T[l] ← j
        l ← l + 1
        flag ← −flag
    end
end
let diff[1 ... N] be a new array
for k = 1 to length(T) − 1 do
    diff[k] ← T[k + 1] − T[k]
end
let duration[1 ... N] be a new array
for k = 1 to floor((N+1)/2) do
    duration[k] ← diff[2 ∗ k − 1] \\ getting the elements of the odd keys.
end
return duration
```

Now we present the algorithm for geometric Brownian motion. The discretized scheme for the GBM is given in (2.9). We also present the algorithm for MMGBM as in (2.15). In the similar fashion one can write algorithm for SMGBM using (2.16).

Algorithm 4: Geometric Brownian motion

function GBM $(N, S_0, mu, sigma, dt)$;
Input : An integer N, initial price S_0, drift mu, volatility $sigma$, time step dt
Output: An array of length N
let $S[1 \ldots N]$ be a new array
set $S[1] \leftarrow S_0$
for $i = 1$ *to* $N - 1$ **do**
  $w \leftarrow$ generate-normal$(0, dt)$\\ generate a normal random variable with mean zero and variance dt
  $q \leftarrow (mu - \frac{1}{2}sigma^2) * dt + sigma * w$
  $S[i+1] \leftarrow S[i] * \exp(q)$
end
return S

Algorithm 5: Binary Markov modulated Geometric Brownian motion

```
function MMGBM (N, S_0, mu, sigma, la);
Input  : An integer N, initial price S_0, drift vector mu = array[mu_1, mu_2],
         volatility vector sigma = array[sigma_1, sigma_2], transition rate vector
         la = array[la_1, la_2]
Output: An array of length N
let S[1...N] and T[1...N] be two new array
set S[1] ← S_0
set t ← 0; j = 0
while t ≤ N do
    T ← generate-exponential(la[j])\\ generate an exponential random variable
      with rate la[j]
    t ← T + t
    j ← 1 − j
end
set k ← 0; s_0 = S_0; C = 1
let GB[1...N] be a new array
for j = 1 to length(T) do
    n ← floor(T[j])
    if n not equal to 0 then
        GB ← GBM(n + 1, s_0, mu[k], sigma[k])
        for r = 1 to n do
            S[C + r] ← GB[1 + r]
        end
        s_0 ← GB[−1]\\ last element of the array GB
        k ← 1 − k
    end
    C ← C + n
end
return S
```

Bibliography

[1] Guy Barles and Halil Mete Soner. Option pricing with transaction costs and a nonlinear black-scholes equation. *Finance and Stochastics*, 2(4):369–397, 1998.

[2] Gopal K. Basak, Mrinal K. Ghosh, and Anindya Goswami. Risk minimizing option pricing for a class of exotic options in a Markov-modulated market. *Stoch. Anal. Appl.*, 29(2):259–281, 2011.

[3] T. R. Bielecki and S. R. Pliska. Risk-sensitive dynamic asset management. *Appl. Math. Optim.*, 39(3):337–360, 1999.

[4] Arunangshu Biswas, Anindya Goswami, and Ludger Overbeck. Option pricing in a regime switching stochastic volatility model. *Statistics & Probability Letters*, 138:116–126, 2018.

[5] John Bollinger. *Bollinger on Bollinger bands.* McGraw Hill Professional, 2001.

[6] John Buffington and Robert J. Elliott. American options with regime switching. *Int. J. Theor. Appl. Finance*, 5(5):497–514, 2002.

[7] Jan Bulla and Ingo Bulla. Stylized facts of financial time series and hidden semi-markov models. *Computational Statistics & Data Analysis*, 51(4):2192–2209, 2006.

[8] Selene Chávez-Rodríguez, Rolando Cavazos-Cadena, and Hugo Cruz-Suárez. Controlled semi-markov chains with risk-sensitive average cost criterion. *Journal of Optimization Theory and Applications*, 170(2):670–686, 2016.

[9] Erhan Çınlar. *Probability and stochastics*, volume 261. Springer Science & Business Media, 2011.

[10] Milan Kumar Das, Anindya Goswami, and Tanmay S. Patankar. *Pricing Derivatives in a Regime Switching Market with Time Inhomogeneous Volatility*. Stoch. Anal. Appl, 2018.

[11] Milan Kumar Das, Anindya Goswami, and Nimit Rana. *Risk sensitive portfolio optimization in a jump diffusion model with regimes. SIAM J.* Control Optim, 2018.

[12] Amogh Deshpande and Mrinal K. Ghosh. Risk minimizing option pricing in a regime switching market. *Stoch. Anal. Appl.*, 26(2):313–324, 2008.

[13] Giovanni B Di Masi, Yu M Kabanov, and Wolfgang J Runggaldier. Mean-variance hedging of options on stocks with markov volatilities. *Theory of Probability & Its Applications*, 39(1):172–182, 1995.

[14] Jin-Chuan Duan, Ivilina Popova, Peter Ritchken, et al. Option pricing under regime switching. *Quantitative Finance*, 2(116-132):209, 2002.

[15] Mardi Dungey, Michael McKenzie, and L. Vanessa Smith. Empirical evidence on jumps in the term structure of the {US} treasury market. *Journal of Empirical Finance*, 16(3):430 – 445, 2009.

[16] R J Elliott, L Chan, and T K Siu. Option pricing and esscher transform under regime switching. *Annals of Finance 1*, 5:423–432, 2005.

[17] Robert Elliott and Juri Hinz. Portfolio optimization, hidden markov models, and technical analysis of p&f-charts. *International Journal of Theoretical and Applied Finance*, 5(04):385–399, 2002.

[18] Robert J Elliott and John Van der Hoek. An application of hidden markov models to asset allocation problems. *Finance and Stochastics*, 1(3):229–238, 1997.

[19] S.N. Ethier and T.G. Kurtz. *Markov processes: characterization and convergence*. Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Wiley, 1986.

[20] W. H. Fleming and S. J. Sheu. Risk-sensitive control and an optimal investment model. *Math. Finance*, 10(2):197–213, 2000. INFORMS Applied Probability Conference (Ulm, 1999).

[21] W. H. Fleming and S. J. Sheu. Risk-sensitive control and an optimal investment model. II. *Ann. Appl. Probab.*, 12(2):730–767, 2002.

[22] Hans Föllmer. Calcul d'itô sans probabilités. In *Séminaire de Probabilités XV 1979/80*, pages 143–150. Springer, 1981.

[23] Hans Föllmer and Martin Schweizer. Hedging of contingent claims under incomplete information. In *Applied stochastic analysis (London, 1989)*, volume 5 of *Stochastics Monogr.*, pages 389–414. Gordon and Breach, New York, 1991.

[24] M K Ghosh, A Goswami, and S Kumar. Portfolio optimization in a markov modulated market. *Modern Trends in Controlled Stochastic Processes*, pages 181–195, 2010.

[25] Mrinal K. Ghosh and Anindya Goswami. Risk minimizing option pricing in a semi-Markov modulated market. *SIAM J. Control Optim.*, 48(3):1519–1541, 2009.

[26] Mrinal K. Ghosh, Anindya Goswami, and Suresh K. Kumar. Portfolio optimization in a semi-Markov modulated market. *Appl. Math. Optim.*, 60(2):275–296, 2009.

[27] Mrinal K. Ghosh and Subhamay Saha. Stochastic processes with age-dependent transition rates. *Stochastic Analysis and Applications*, 29(3):511–522, 2011.

[28] Anindya Goswami and Sanket Nandan. Convergence of estimated option price in a regime switching market. *Indian J. Pure Appl. Math.*, 47(2):169–182, 2016.

[29] Anindya Goswami, Jeeten Patel, and Poorva Shevgaonkar. A system of non-local parabolic pde and application to option pricing. *Stochastic Analysis and Applications*, 34(5):893–905, 2016.

[30] Anindya Goswami, Jeeten Patel, and Poorva Shevgaonkar. A system of non-local parabolic PDE and application to option pricing. *Stoch. Anal. Appl.*, 34(5):893–905, 2016.

[31] Anindya Goswami and Ravi Kant Saini. Volterra equation for pricing and hedging in a regime switching market. *Cogent Economics & Finance*, 2(1):939769, 2014.

[32] X Guo and Q Zhang. Closed form solutions for perpetual american put options with regime switching. *SIAM J. Appl. Math*, 39:173–181, 2014.

[33] James D Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384, 1989.

[34] John C Hull. *Options, futures, and other derivatives.* Pearson Education India, 2006.

[35] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library.* North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition, 1989.

[36] A. Jobert and L. C. G. Rogers. Option pricing with Markov-modulated dynamics. *SIAM J. Control Optim.*, 44(6):2063–2078, 2006.

[37] Arnaud Jobert and Leonard C. G. Rogers. Option pricing with markov-modulated dynamics. *SIAM Journal on Control and Optimization*, 44(6):2063–2078, 2006.

[38] Gopinath Kallianpur and Rajeeva L. Karandikar. *Introduction to option pricing theory*. Birkhäuser Boston, Inc., Boston, MA, 2000.

[39] Jan Kallsen. Optimal portfolios for exponential Lévy processes. *Math. Methods Oper. Res.*, 51(3):357–374, 2000.

[40] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988.

[41] John Frank Charles Kingman. *Poisson processes*. Wiley Online Library, 1993.

[42] Ralf Korn and Elke Korn. *Option pricing and portfolio optimization: modern methods of financial mathematics*, volume 31. American Mathematical Soc., 2001.

[43] Mario Lefebvre and Pierre Montulet. Risk-sensitive optimal investment policy. *Internat. J. Systems Sci.*, 25(1):183–192, 1994.

[44] Rogemar S. Mamon and Marianito R. Rodrigo. Explicit solutions to European options in a regime-switching economy. *Oper. Res. Lett.*, 33(6):581–586, 2005.

[45] Harry M. Markowitz. *Portfolio selection: Efficient diversification of investments*. Cowles Foundation for Research in Economics at Yale University, Monograph 16. John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London, 1959.

[46] Robert C Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257, 1969.

[47] Robert C Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of economic theory*, 3(4):373–413, 1971.

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

[49] Bernt Øksendal and Agnès Sulem. *Applied stochastic control of jump diffusions*. Universitext. Springer-Verlag, Berlin, 2005.

[50] Tanmay S Patankar. Asset pricing in a semi-markov modulated market with time-dependent volatility. *arXiv preprint arXiv:1609.04907*, 2016.

[51] Amnon Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44. Springer Science & Business Media, 2012.

[52] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[53] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

[54] H.L. Royden and P. Fitzpatrick. *Real Analysis*. Prentice Hall, 2010.

[55] Tobias Rydén, Timo Teräsvirta, and Stefan Åsbrink. Stylized facts of daily return series and the hidden markov model. *Journal of applied econometrics*, pages 217–244, 1998.

[56] Thomas Schreiber and Andreas Schmitz. Surrogate time series. *Physica D: Nonlinear Phenomena*, 142(3):346–382, 2000.

[57] M Schweizer. A guided tour through quadratic hedging approaches. In M. Musiela E. Jouini, J. Cvitanić, editor, *Option Pricing Interest Rates and Risk Management*, pages 538–574. Cambridge University Press, 2001.

[58] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996. Translated from the first (1980) Russian edition by R. P. Boas.

[59] S M Srivastava. A survey of measurable selection theorems. April 2010.

[60] Rangarajan K. Sundaram. *A first course in optimization theory*. Cambridge University Press, Cambridge, 1996.

[61] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, 58(1-4):77–94, 1992.

[62] James Theiler and Dean Prichard. Constrained-realization monte-carlo method for hypothesis testing. *Physica D: Nonlinear Phenomena*, 94(4):221–235, 1996.

[63] W. Wang, Z. Jin, L. Qian, and X. Su. Local risk minimization for vulnerable european contingent claims on nontradable assets under regime switching models. *Stochastic Analysis and Applications*, 34(4):662–678, 2016.

[64] Fei Lung Yuen and Hailiang Yang. Option pricing in a jump-diffusion model with regime switching. *ASTIN Bulletin: The Journal of the IAA*, 39(2):515–539, 2009.

Index