Lappeenranta University of Technology
School of Engineering Science
Computational Engineering and Technical Physics
Techno-Mathematics

Arnold Kaynet Muchatibaya

BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS WITH APPLICATIONS

Master's Thesis

Examiners:	Professor Simo Särkkä
	Professor Heikki Haario
Supervisors:	Professor Simo Särkkä
	Professor Heikki Haario

ABSTRACT

Lappeenranta University of Technology
School of Engineering Science
Computational Engineering and Technical Physics
Techno-Mathematics

Arnold Kaynet Muchatibaya

Backward Stochastic Differential Equations with Applications

Master's Thesis

2018

60 pages.

Examiners: Professor Simo Särkkä
Professor Heikki Haario

Keywords: Stochastic differential equations, Backward stochastic differential equations, Doob's h-transform, Feynman-Kac, Lévy process

In this thesis we study backward stochastic differential equations driven by a Brownian motion and by a Lévy process and their applications, focusing on their applications to financial markets. We give results on the existence and uniqueness of solution of backward stochastic differential equations when the drift is Lipschitz continuous and the terminal condition is square integrable and measurable with respect to the terminal filtration. Backward stochastic differential equations associated with a forward stochastic differential equation are investigated. We use the generalisation of the Feynman-Kac formula to show the relationship between a backward stochastic differential equation associated with a forward stochastic differential equation and a partial differential equation in the Brownian motion case and a partial differential integral equation for the Lévy process case. The Doob's h-transform is studied for the Brownian motion and applied to stochastic differential equations. Finally, we conclude with an application to option pricing and hedging of a European calls for both Brownian and Lévy processes.

PREFACE

All glory and praise be given to the most High for giving me the life and sustaining me during my studies until the completion of this thesis. I wish to thank my supervisors Professor Simo Särkkä and Professor Heikki Haario for their help, resources and supervision throughout the preparation of this thesis. I would also like to thank my family for the strength and support during my studies. Finally, thank you to African Institute for Mathematical Sciences (AIMS-Tanzania) and Lappeenranta University of Technology for arranging and awarding of the scholarship to undertake studies at the later.

Lappeenranta, October 19, 2018

Arnold Kaynet Muchatibaya

CONTENTS

ABBREVIATIONS AND SYMBOLS

These are the abbreviations we will use throughout this essay:

a.s.	Almost surely
$\mathbb{C}$	Complex numbers
$A \in B$	Element A in set B
$\forall$	For all
$\Rightarrow$	Implies
$\langle \cdot, \cdot \rangle$	Inner product
$\mapsto$	Maps to
$\|\| \cdot \|\|$	Norm of $(\cdot)$
$\mathbb{R}$	Real numbers
$\exists$	There exists
$\cup$	Union of sets
w.r.t	with respect to
BSDE	Backward stochastic differential equation
EMM	Equivalent martingale measure
FPK	Fokker-Planck-Kolmogorov
FBSDE	Forward-backward stochastic differential equation
FSDE	Forward stochastic differential equation
ODE	Ordinary differential equation
PDE	Partial differential equation
PDIE	Partial differential integral equation
SDE	Stochastic differential equation

1 INTRODUCTION

Allowing randomness in the coefficients of an ordinary differential equation results in realistic mathematical models of physical phenomena. Stochastic differential equations arise when we allow randomness in the coefficients of ordinary differential equations or when the forcing is an irregular stochastic process like Gaussian white noise [1]. There are two types of stochastic differential equations; if the initial condition is specified, we have a forward stochastic differential equation and if the terminal condition is specified, we have a backward stochastic differential equation.

The theory of backward stochastic differential equations has found wide applications in areas such as stochastic optimal control, theoretical economies, and mathematical finance problems such as the theory of hedging and non-linear pricing in incomplete markets [2]. Backward stochastic differential equations can be driven by a Lévy process, Brownian motion, Poisson process, or a combination of these.

Bismut [3] first introduced backward stochastic differential equations in a linear form as the equation for the conjugate variable in the stochastic version of the Pontryagin maximum principle. Pardoux et al. [4] were the first to consider general backward stochastic differential equations. Their main result was the existence and uniqueness of an adapted pair of processes as a solution of a backward stochastic differential equation.

Several authors have extended their results. Peng [5] used backward stochastic differential equations to obtain a probabilistic interpretation for systems of second order quasi-linear parabolic partial differential equations. Pardoux et al. [6] introduced a new class of backward stochastic differential equations, which allowed them to produce a probabilistic representation of a certain quasi-linear stochastic partial differential equation thus extending the Feynman-Kac formula for stochastic partial differential equations. Antonelli [7] showed the existence and uniqueness of a solution of a backward stochastic differential equation inspired from the stochastic differential utility in finance theory. Ma et al. [8] investigated adapted solutions to a class of forward-backward stochastic differential equations in which the forward stochastic differential equation is non-degenerate. They showed that the adapted solution can be sought over an arbitrarily prescribed time duration via a direct four step scheme. Using this scheme, they proved that the backward components of the adapted solution are determined explicitly by the forward component via the solution of a certain quasi-linear parabolic partial differential equation system.

El-Karoui et al. [9] summarized the existence and uniqueness of solutions of backward stochastic differential equations by Pardoux et al. [4] and gave new shorter proofs. They stated the a priori estimates of the difference of two backward stochastic differential equations, and the uniqueness and existence was proved using a fixed point theorem. They also looked at the solution of a backward stochastic differential equation associated with a forward stochastic differential equations. The main property was that the solution is Markovian in the sense that it can be written as a function of time and state process. The generalisation of the Feynman-Kac formula is given, and they also showed that under smoothness assumptions, the solution of the backward stochastic differential equation corresponds to a solution of a system of quasi-linear parabolic partial differential equations. These results could be applied to option pricing of a European call in the constrained Markovian cases.

Buckdahn and Pardoux [10] proved the existence and uniqueness of a solution to a backward stochastic differential equation with respect to both the Brownian motion and Poisson random measure and the associated integro-partial differential equation of parabolic type. They proved that under certain conditions, the solution of a backward stochastic differential equation provides the unique viscosity solution of the associated integro-partial differential equation. Situ [11] studied backward stochastic differential equations driven by Brownian motion and Poisson point process. A new existence and uniqueness result for the solution of the partial differential integral equation with non Lipschitz force is obtained. Oukine [12] considered a backward stochastic differential equation driven by a Poisson random measure. The integral representation of the square integral random variable in terms of a Poisson random measure is the main result.

Nualart and Schoutens [13] proved a martingale representation theorem for a Lévy process satisfying some exponential moment condition. Nuarlart et al. [14] used results from [13] to establish existence and uniqueness of solutions for backward stochastic differential equations driven by a Lévy process. Our work is primarily based on [9] and [14].

1.1 Objectives

This thesis is based on the articles by Nualart et al. [13] and El-Karoui et al.[9]. We expand the proof studied in the articles, especially the existence and uniqueness of solutions of a BSDE under Lipschitz conditions on the drift driven by a Lévy process and Brownian motion. We are also concerned with the application of these BSDEs in finance. To achieve our purpose, we consider some specific objectives:

(i) We study the existence and uniqueness of a general BSDE under the Lipschitz condition on the drift driven by a Lévy process and Brownian motion expanding the proofs in these sections.

(ii) Apply the Feynman-Kac formula to the BSDEs to get the relationship between BSDEs, PDEs and PDIEs.

(iii) Consider the application of the theory above to European call options.

(iv) We study the Doob's h-transform applied to an SDE in the Brownian Motion case.

1.2 Outline

In Chapter 2, we study the existence and uniqueness of a solution of BSDEs with a Lipschitz driver and driven by a Lévy process and Brownian motion in Section 2.3. We also consider the Feynman-Kac formula for the BSDEs to establish the relationship with the partial differential equation in the case of BSDE driven by Brownian motion and the partial differential integro equation in the case of the BSDE driven by a Lévy process in Section 2.4. In Section 2.5 we study the Doob's h-transform and application to stochastic differential equations to come up with BSDE. In Chapter 3, we look at the application of these studied BSDEs to option pricing. In Chapter 4, we discuss the results we have obtained in our work and mention possible future work to be done. Finally we conclude our work in Chapter 5.

2 BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

2.1 Preliminaries

In this section we will give the background of SDEs, theorems and inequalities which will be necessary to refer in the forthcoming chapters. We start by defining the space we will be working on,

Definition 1 (Probability space)**.** If Ω is a given set, then a σ-algebra $\mathcal{F}$ on Ω is a family $\mathcal{F}$ of subsets of Ω with the following properties:

(i) $\emptyset \in \mathcal{F}$,

(ii) $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$ where F^c is the complement of F in Ω and $F^c = \Omega \backslash F$, and

(iii) $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow A := \bigcup_{i=1}^{\infty} \in \mathcal{F}$.

Then the pair $(\Omega, \mathcal{F})$ is called a measurable space. A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \longmapsto [0, 1]$ such that

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

(ii) if $A_1, A_2, \cdots \in \mathcal{F}$ and $\{A_i\}_{i=1}^{\infty}$ is disjoint, then

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty}\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Then the triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space [15].

A martingale is a stochastic process for which at a time in the realised sequence the expectation of the next value is the current observed value given prior observation. Now we give the mathematical definition is as follows.

Definition 2 (Martingale)**.** A filtration on $(\Omega, \mathcal{F})$ is a family $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ such that

$$0 \leq s \leq T \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t.$$

An n-dimensional stochastic process $\{X_t\}_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a martingale [15] with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ (and with respect to $\mathbb{P}$) if

(i) X_t is $\mathcal{F}_t-$measurable for all t,

(ii) $\mathbb{E}[|X_t|] < \infty$ for all t, and

(iii) $\mathbb{E}[X_s|\mathcal{F}_t] = X_t$ a.s. for all $s \geq t$.

Now let us define a Lévy process.

Definition 3 (Lévy process)**.** Given the probability space in Definition 1, a Lévy process $X = \{X_t, t \geq 0\}$ taking values in $\mathbb{R}^d$ is a stochastic process having stationary and independent increments and we always assume $X_0 = 0$ with probability 1. So

- $X_t : \Omega \longrightarrow \mathbb{R}^d$.
- Given any selection of distinct time points $0 \leq t_1 < t_2 < \cdots < t_n$ the random vectors $X_{t_1}, X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \cdots, X_{t_n} - X_{t_{n-1}}$ are all independent.
- Given any two distinct times $0 \leq s < t < \infty$, the probability distribution of $X_t - X_s$ coincides with that of X_{t-s}. [16]

Brownian motion and Poison process are examples of Lévy processes. Thus we have,

Definition 4 (Brownian motion)**.** A standard Brownian motion in $\mathbb{R}^d$ is a Lévy process $W = (W_t, t \geq 0)$ for which

- $W_t \sim N(0, tI)$ for each $t \geq 0$
- W has continuous sample paths

Every Lévy process is characterised by its characteristic function which is defined as follows [7].

Definition 5 (Characteristic function)**.** Let X be a random variable defined on the probability space in Definition 1 taking values in $\mathbb{R}^d$ with the probability law $\mathbb{P}_x$. It's characteristic function $\phi_x : \mathbb{R}^d \longrightarrow \mathbb{C}$ is

$$\begin{aligned}\phi_x(u) &= \mathbb{E}\left(e^{i(u,x)}\right) \\ &= \int\limits_{\Omega} e^{i(u,X(\omega))}\mathbb{P}(d\omega) \\ &= \int\limits_{\mathbb{R}^d} e^{i(u,y)}\mathbb{P}_x(dy)\end{aligned}$$

for each $u \in \mathbb{R}^d$

We define the indicator function χ_A as a function defined to be 1 on A and 0 elsewhere. Now the characteristic function is given by the Lévy -Khintchine formula defined below as.

Definition 6 (Lévy-Khintchine formula)**.** If $X = \{X_t, t \geq 0\}$ is a Lévy process, then it has a specific form for its characteristic function [16]. More precisely $\forall t \geq 0, u \in \mathbb{R}^d$

$$\mathbb{E}(e^{i(u,X_t)}) = e^{t\eta(u)}$$

where

$$\eta(u) = i(b,u) - \frac{1}{2}(u,au) + \int\limits_{\mathbb{R}^d-\{0\}} \left[e^{i(u,y)} - 1 - i(u,y)\chi_{0<|y|<1}(y)\right]\nu(dy).$$

where

- $b \in \mathbb{R}^d$
- a is a positive definite symmetric $d \times d$ matrix
- ν is a Lévy measure on $\mathbb{R}^d - \{0\}$ so that

$$\int\limits_{\mathbb{R}^d-\{0\}} \min\{1, |y|^2\nu(dy)\} < \infty.$$

A Lévy process can be decomposed into a linear drift, Brownian motion, and a pure jump process [16]. This result is called the Lévy-Itô decomposition and defined as follows.

Theorem 2.1.1 (The Lévy-Itô decomposition). *If X is a Lévy process, then there exists $b \in \mathbb{R}^d$, a Brownian motion B with diffusion matrix Q and an independent Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R}^d - \{0\})$ such that for each $t \geq 0$,*

$$X_t = bt + B_t + \int\limits_{|x|<1} x\tilde{N}(t, dx) + \int\limits_{|x|>1} xN(t, dx).$$

Proof. See [16] □

Definition 7. Let us consider the SDE,

$$\begin{aligned} dX_t &= b(t, x)dt + \sigma(t, x)dW_t \\ X(0) &= x, \;\; t \geq 0. \end{aligned}$$

A strong solution of this SDE on the given probability space with respect to the fixed Brownian motion W and initial condition x is a process $X = \{X_s;\ 0 \leq s \leq T\}$ with continuous sample paths and with the following properties:

(i) X is adapted to the filtration $\mathcal{F}_s$,

(ii) $\mathbb{P}[X_0 = x] = 1$,

(iii)
$$\mathbb{P}\left[\int\limits_0^T \{|b(s, X_s)| + \sigma_{ij}^2(s, X_s)\}\, ds < \infty\right] = 1$$
holds for every $1 \leq i \leq d$, and

(iv) The integral version is
$$X_s = X_0 + \int\limits_0^T b(s, X_s)\, ds + \int\limits_0^T \sigma(s, X_s)\, dW_s.$$

[17].

For any martingale adapted with respect to a Brownian motion can be expressed as an Itô integral with respect to the same Brownian motion as follows.

Theorem 2.1.2 (Martingale representation)**.** *Let* $(W_t,\, 0 \leq t \leq T)$ *be a Brownian motion on* $(\Omega, \mathcal{F}, \mathbb{P})$*. Let* $\{\mathcal{F}_t; 0 \leq t \leq T\}$ *be the filtration generated by this Brownian motion. Let* $\{X_t; 0 \leq t \leq T\}$ *be a martingale (under* $\mathbb{P}$*) relative to this filtration (i.e., for every* t*,* X_t *is* $\mathcal{F}_t-$ *measurable, and for* $0 \leq s \leq t \leq T$*,* $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ *a.s.). Then there is an adapted process* $\{A_t; 0 \leq t \leq T\}$*,* A_t *square integrable such that*

$$X_t = X_0 + \int_0^t A_u \, dW_u, \quad 0 \leq t \leq T.$$

Proof. See [15] ☐

The Burkholder-Davis-Gundy inequalities relate the maximum of a local martingale with it's quadratic variation. This result is important in the proofs in the next chapters.

Theorem 2.1.3 (Burkholder-Davis-Gundy inequalities)**.** *Let* $T > 0$ *and* $(M_t)_{0 \leq t \leq T}$ *be a continuous local martingale such that* $M_0 = 0$*. For every* $0 < p < \infty$*, there exists universal constants* c_p*,* C_p *independent of* T *and* $(M_t)_{0 \leq t \leq T}$ *such that,*

$$c_p \mathbb{E}\left(\langle M_t \rangle_T^{\frac{p}{2}}\right) \leq \mathbb{E}\left(\left[\sup_{0 \leq t \leq T} |M_t|\right]^p\right) \leq C_p \mathbb{E}\left(\langle M_t \rangle_T^{\frac{p}{2}}\right).$$

Proof. See [17]. ☐

The Banach fixed point theorem is important in the proof of existence and uniqueness of solution of a BSDE, hence we first define a metric space then give theorem.

Definition 8. Let (X, d) be a metric space. A mapping $T : X \longmapsto X$ is Lipschitz continuous if there exists a constant $\alpha > 0$ such that $d(Tx, Ty) \leq \alpha \, d(x, y)$ for all $x, y \in X$. If $0 \leq \alpha < 1$, then T is called a contraction mapping, and α is called the factor of T [18].

Theorem 2.1.4 (Banach fixed point theorem)**.** *Suppose that* (X, d) *is a generalised complete metric space, and that the function* $T : X \longmapsto X$ *is a contraction.*
Let $x_0 \in X$*, and consider the sequence of successive approximations with initial element* x_0

$$x_0, Tx_0, T^2x_0, \cdots, T^ix_0, \cdots. \tag{1}$$

Then either

1. *For every integer* $i = 0, 1, 2, \cdots$, *one has*

$$d(T^i x_0, T^{i+1} x_0) = \infty, \quad \textit{or}$$

2. *The sequence of approximations, Equation* (1) *is* d*-convergent to a fixed point of* T.

Proof. See [18] □

We need to define an $\mathbb{L}^p$ space before the Hölder inequality as follows.

Definition 9. Consider the measurable space in Definition 1 and $1 \leq p, q \leq \infty$. The space $\mathbb{L}^p(\Omega)$ consists of equivalence classes of measurable functions $f : \Omega \mapsto \mathbb{R}$ such that

$$\int |f|^p \mathbb{P}(d\omega \in \mathcal{F}) < \infty,$$

where $\omega \in \Omega$ and two measurable functions are equivalent if they are equal $\mathbb{P}$ a.e [19]. The $\mathbb{L}^p$ norm of $f \in \mathbb{L}(\Omega)$ is defined by

$$||f||_{\mathbb{L}^p} = \left(\int |f|^p \mathbb{P}(d\omega \in \mathcal{F}) \right)^{\frac{1}{p}}.$$

When $p = 1$ the space $\mathbb{L}^1$ consists of all integrable functions on Ω.

Theorem 2.1.5 (Hölder inequality)**.** *Assume that a function* $f \in \mathbb{L}^p$ *and* $g \in \mathbb{L}^q$, *where* $p, q \in (1, \infty)$ *are conjugate numbers, that is,*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then $fg \in \mathbb{L}^1$, *and the following inequality holds*

$$\left| \int fg \, dx \right| \leq \int |fg| \, dx \leq ||f||_p \, ||g||_q. \tag{2}$$

Proof. See [19] □

Theorem 2.1.6 (Markov inequality)**.** *Suppose* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a measure space,* f *is a measurable extended real valued function, and* $\varepsilon > 0$. *Then*

$$\mathbb{P}(\{\omega \in \Omega : |f(\omega)| \geq \varepsilon\}) \leq \frac{1}{\varepsilon} \int_{\Omega} |f| \, \mathbb{P}(d\omega \in \mathcal{F}).$$

Proof. See [19] □

Definition 10 (Itô process)**.** Let W_t be a one-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. An Itô process (or stochastic integral) is a stochastic process X_t on $(\Omega, \mathcal{F}, \mathbb{P})$ of the form

$$X_t = X_0 + \int_0^t b(s,\omega)\, ds + \int_0^t \sigma(s,\omega)\, dW_s,$$

where

$$\mathbb{P}\left[\int_0^t \sigma(s,\omega)^2\, ds < \infty \text{ for all } t \geq 0\right] = 1,$$

and

$$\mathbb{P}\left[\int_0^t |b(s,\omega)|\, ds < \infty \text{ for all } t \geq 0\right] = 1.$$

[15]

We define the quadratic variation and cross variance as follows.

Definition 11. If $X_t(\cdot) : \Omega \longmapsto \mathbb{R}$ is a continuous stochastic process, then for $p > 0$ the p'th variation process of X_t; $\langle X, X\rangle_t^{(p)}$ is defined by

$$\langle X, X\rangle_t^{(p)}(\omega) = \lim_{\triangle t_k \to 0} \sum_{t_k \leq t} |X_{t_{k+1}}(\omega) - X_{t_k}(\omega)|^p \text{ (limit in probability)}$$

where $0 = t_1 < t_2 < \cdots < t_n = t$ and $\triangle t_k = t_{k+1} - t_k$. If $p = 1$ then it is called total variation and if $p = 2$ is called quadratic variation. We also have

$$\langle X, X\rangle_t^{(p)} = \langle X, X\rangle_t = \langle X\rangle_t.$$

More generally for the cross variation between two processes X_t and Y_t we have

$$\langle X, Y\rangle_t(\omega) = \lim_{\triangle t_k \to 0} \sum_{t_k \leq t} (X_{t_{k+1}}(\omega) - X_{t_k}(\omega))(Y_{t_{k+1}}(\omega) - Y_{t_k}(\omega)).$$

[15].

Let us consider the Itô formula for Brownian motion and Lévy process driven Itô process as follows.

Theorem 2.1.7 (Itô formula). *Let X_t be an Itô process, and let $f(t,x)$ be a function for which the partial derivatives $f_t,\ f_x,\ f_{xx}$ are defined and continuous. Then for every $T \geq 0$,*

$$f(T,X_t) = f(0,X_0) + \int_0^T f_t dt + \int_0^T f_x dX_t + \frac{1}{2}\int_0^T f_{xx} d\langle X\rangle_t. \tag{3}$$

Let $f(t,x,y)$ be a function whose partial derivatives $f_t,\ f_x,\ f_y,\ f_{xx},\ f_{xy},\ f_{yy}$ are defined and are continuous. Let X_t and Y_t be Itô processes. The two dimensional Itô formula in differential form is

$$df(t,x,y) = f_t dt + f_x dX_t + f_y dY_t + \frac{1}{2} f_{xx} d\langle X\rangle_t + f_{xy} d\langle X,Y\rangle_t + \frac{1}{2} f_{yy} d\langle Y\rangle_t. \tag{4}$$

Proof. See [20] □

Theorem 2.1.8 (Itô formula for Lévy process driven SDEs). *Let $X = (X^1, \cdots, X^n)$ be an n-tuple of semi martingales and let $f : \mathbb{R}^n \longmapsto \mathbb{R}$ have continuous second order partial derivatives. Then $f(x)$ is a semi-martingale and the following formula holds*

$$f(X_t) - f(X_0) = \sum_{i=1}^n \int_{0+}^t \frac{\partial f}{\partial x_i}(X_{s-}) dX_s^{(i)} + \frac{1}{2} \sum_{1\leq i,j\leq n} \int_{0+}^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-}) d[X^{(i)}, X^{(j)}]_s^c$$
$$+ \sum_{0<s\leq t} \left\{ f(X_s) - f(X_{s-}) - \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X_{s-}) \triangle X_s^{(i)} \right\}$$

Proof. See [21] □

Theorem 2.1.9 (Doob's maximal inequality). *Let $(X_t;\ \mathcal{F}_t,\ 0 \leq t < \infty)$ be a semi-martingale whose every path is right-continuous. Let $\alpha < \beta$ be real numbers, and $[0,T]$ is a sub interval of $[0,\infty)$. Then*

$$\mathbb{E}\left(\sup_{0\leq t\leq T} X_t\right)^p \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left(X_T^p\right), \quad p > 1,$$

provided $X_t \geq 0$ a.s. $\mathbb{P}$ for every $t \geq 0$ and $\mathbb{E}\left(X_T^p\right) < \infty$.

Proof. See [17] □

Theorem 2.1.10 (Gronwall's inequality)**.** *Let $g(t)$ and $h(t)$ be regular non-negative functions on $[0,T]$. Then for any regular $f(t) \geq 0$ satisfying the inequality for all $0 \leq t \leq T$,*

$$f(t) \leq g(t) + \int_0^t h(s)f(s)\,ds,$$

we have

$$f(t) \leq g(t) + \int_0^t h(s)g(s)\exp\left(\int_s^t h(u)\,du\right)ds. \tag{5}$$

In particular, if g is non-decreasing, Equation (5) simplifies to give

$$f(t) \leq g(t)e^{\int_0^t h(s)\,ds}.$$

In its simplest form when $g = A$ and $h = B$ are constants,

$$f(t) \leq Ae^{Bt}.$$

Proof. See [22] □

Theorem 2.1.11 (Fubini's theorem)**.** *Let $f(x,t)$ be continuous on $[t,T] \times [0,b]$. Then iterated integrals:*

$$\int_t^T \int_0^b f(x,t)\,dx\,dt = \int_0^b \int_t^T f(x,t)\,dt\,dx.$$

Proof. See [23] □

We consider the Fubini theorem for stochastic processes. We first of all define an FV process and some notation we will use for the theorem.

Definition 12. An FV process is a cadlag adapted stochastic process such that all its paths are of finite variation on each compact interval on $\mathbb{R}_+$ [21].

The following Notation is by [21]. Notation: Let A be an FV process and let F be jointly measurable process such that

$$\int_0^t F(s,\omega)dA_s(\omega)$$

exists and is finite $\forall t > 0$, a.s. we let

$$(F \cdot A)_t(\omega) = \int_0^t F(s,\omega) dA_s(\omega).$$

We also write $F \cdot A$ to denote the process $F \cdot A = (F \cdot A_t)_{t \geq 0}$. Then the Fubini theorem is as follows.

Theorem 2.1.12 (Fubini's theorem (stochastic processes)). *Let X be a semi-martingale, $H_t^a = H(a,t,\omega)$ be a bounded $\mathcal{A} \otimes \mathbb{P}$ measurable function, and let μ be a finite measure on $\mathcal{A}$. Let $Z_t^a = \int_0^t H_s^a dX_s$ be $\mathcal{A} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ measurable such that for each a, Z^a is cádlág version of $H^a \cdot X$. Then $Y_t = \int_{\mathcal{A}} Z_t^a \mu(da)$ is cádlág version of $H \cdot X$, where $H_t = \int_{\mathcal{A}} \mu(da)$ [21].*

The Itô isometry is essential for computation of variances of random variables given as an Itô integral. The following theorem is by [20].

Theorem 2.1.13 (Itô Isometry). *Let $W : [0,T] \times \Omega \longrightarrow \mathbb{R}$ be the standard Brownian motion defined to $T > 0$. Let $X : [0,T] \times \Omega \longrightarrow \mathbb{R}$ be a stochastic process that is adapted to the natural filtration of the Brownian motion then,*

$$\mathbb{E}\left[\left(\int_0^T X_s dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t X_s^2 ds\right]$$

Theorem 2.1.14 (Comparison theorem). *Let (f^1, ξ^1) and (f^2, ξ^2) be two standard parameters of BSDEs, and let (X^1, Y^1) and (X^2, Y^2) be the associated square integrable solutions. We suppose that*

1. $\xi^1 \geq \xi^2 \; \mathbb{P} \; a.s.$
2. $\delta_2 f_t = f^1(t, X_t^2, Y_t^2) - f^2(t, X_t^2, Y_t^2) \geq 0, \; d\mathbb{P} \bigotimes dt \; a.s.$

Then we have almost surely for any time t,

$$X_t^1 \geq X_t^2.$$

Moreover, the comparison is strict. That is, if, in addition, $X_0^1 = X_0^2$, then $\xi^1 = \xi^2$, $f^1(t, X_t^2, Y_t^2) = f^2(t, X_t^2, Y_t^2)$, $d\mathbb{P} \bigotimes dt$ a.s., and $X^1 = X^2$ a.s. More generally, *if* $X_t^1 = X_t^2$ *on a set* $A \in \mathcal{F}_t$*, then* $X_s^1 = X_s^2$ *almost surely on* $[0, T] \times A$*,* $\xi^1 = \xi^2$*a.s. on* A*, and* $f^1(t, X_t^2, Y_t^2) = f^2(t, X_t^2, Y_t^2)$ *on* $A \times [t, T] d\mathbb{P} \bigotimes dt$ *a.s.*

Proof. See [9] □

Definition 13 (Generalised generator)**.** The generalised (infinitesimal) generator for a time dependent function $\phi(x, t)$ can be defined as [24]

$$\mathcal{A}_t \phi(x, t) = \lim_{s \downarrow 0} \frac{\mathbb{E}[\phi(x_{t+s}, t + s)] - \phi(x_t, t)}{s} \tag{6}$$

for a time dependent SDE and

$$\mathcal{A}_t(\cdot) = \frac{\partial(\cdot)}{\partial t} + \sum_i \frac{\partial(\cdot)}{\partial x_i} b_i(t, X_t) + \frac{1}{2} \sum_{i,j} \left(\frac{\partial^2(\cdot)}{\partial x_i \partial x_j} \right) [\sigma(t, X_t) Q \sigma^*(t, X_t)]_{ij}$$

where σ^* is the transpose of σ and the SDE is

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

where Q is the diffusion matrix of Brownian motion.

2.2 Background

In this section we will discuss the theory of BSDEs. First, we give some notations used in this chapter and the rest of the work, by [9]. For $x \in \mathbb{R}^d$, $|y|$ denotes the Euclidean norm, and $\langle y, z\rangle$ denotes the inner product. An $n \times d$ matrix will be considered as an element $y \in \mathbb{R}^{n\times d}$, and the Euclidean norm is given by

$$|z| = \sqrt{\text{trace}(zz^*)},$$

where z^* is the transpose of z, and

$$\langle y, z\rangle = \text{trace}(yz^*), \quad \text{for } y, z \in \mathbb{R}^{n\times d}.$$

Given a probability space, and $\mathbb{R}^n$-valued Brownian motion W, we consider the following definitions:

- $\{\mathcal{F}_t;\ 0 \le t \le T\}$, the filtration generated by the Brownian motion W and, $\mathcal{P}$ the σ-field of predictable sets of $\Omega \times [0, T]$.
- $\mathbb{L}^2_T(\mathbb{R}^d)$, the space of all $\mathcal{F}_T$ measurable random variables $X : \Omega \longmapsto \mathbb{R}^d$ satisfying

$$||Y||^2 = \mathbb{E}(|Y|^2) < +\infty.$$

  Usually denoted as $\mathbb{L}^{2,d}_T$.
- $\mathbb{H}^2_T(\mathbb{R}^d)$, the space of all predictable processes $\varphi : \Omega \times [0, T] \longmapsto \mathbb{R}^d$ such that

$$||\varphi||^2 = \mathbb{E}\int_0^T |\varphi_t|^2\, dt < +\infty.$$

  Usually denoted as $\mathbb{H}^{2,d}_T$.
- $\mathbb{H}^1_T(\mathbb{R}^d)$, the space of all predictable processes $\varphi : \Omega \times [0, T] \longmapsto \mathbb{R}^d$ such that

$$\mathbb{E}\sqrt{\int_0^T |\varphi_t|^2\, dt} < +\infty.$$

  Usually denoted $\mathbb{H}^{1,d}_T$.

- For $\beta > 0$ and $\phi \in \mathbb{H}^2_T(\mathbb{R}^d)$,

$$||\phi||^2_\beta = \mathbb{E}\int_0^T e^{\beta t}|\phi_t|^2\,dt,$$

and $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d)$ denotes the space $\mathbb{H}^2_T(\mathbb{R}^d)$ equipped with the norm $||\bullet||_\beta$. Usually denoted $\mathbb{H}^{2,d}_{T,\beta}$.

2.3 Existence and Uniqueness

In this section we are going to prove the existence and uniqueness of the solution of a BSDE driven by Lévy process and another driven by Brownian motion. The drift for both BSDEs is considered to be Lipschitz.

2.3.1 Brownian Motion

Consider a BSDE of the form [9],

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t^* dW_t, \;\; Y_T = \xi. \tag{7}$$

or equivalently,

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s^* dW_s \tag{8}$$

where

- The terminal value is an $\mathcal{F}_T$-measurable random variable, $\xi : \Omega \longmapsto \mathbb{R}^d$.
- The generator $f : \Omega \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^{n \times d} \longmapsto \mathbb{R}^d$ and is $\mathcal{P} \bigotimes \mathcal{B}^d \bigotimes \mathcal{B}^{n \times d}$-measurable.

Here $\mathcal{B}^d$ denotes Borel-measurable sets in $\mathbb{R}^d$, and $\mathcal{B}^{n \times d}$ denotes Borel-measurable sets in $\mathbb{R}^{n \times d}$.

Definition 14. A solution of Equation (7) is a pair (Y, Z) such that $\{Y_t \; ; \; t \in [0, T]\}$ is a continuous $\mathbb{R}^d$-valued adapted process, and $\{Z_t \; ; \; t \in [0, T]\}$ is an $\mathbb{R}^{n \times d}$-valued predictable process satisfying

$$\int_0^T |Z_s|^2 \, ds < +\infty, \;\; \mathbb{P} \; a.s.$$

Definition 15. Suppose that $\xi \in \mathbb{L}_T^{2,d}$, $f(\cdot, 0, 0) \in \mathbb{H}_T^{2,d}$, and f is uniformly Lipschitz that is, there exists $C > 0$ such that $d\mathbb{P} \bigotimes dt$ a.s.

$$|f(\omega, t, x_1, y_1) - f(\omega, t, x_2, y_2)| \leq C(|x_1 - x_2| + |y_1 - y_2|), \;\; \forall (x_1, y_1), \forall (x_2, y_2) \in \mathbb{R}^2.$$

Then (f, ξ) are said to be standard parameters for the BSDE.[9]

Proposition 1. Let $((f^i, \xi^i); i = 1, 2)$ be two standard parameters of the BSDE and $((Y^i, Z^i); i = 1, 2)$ be two square integrable solutions. Let C be a Lipschitz constant for f^1, and put $\delta Y_t = Y_t^1 - Y_t^2$ and $\delta_2 f_t = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)$. For any (λ, μ, β) such that $\mu > 0, \lambda^2 > C,\ \beta \geq C(2 + \lambda^2) + \mu^2$, it follows that

$$||\delta Y||_\beta^2 \leq T \left[e^{\beta T} \mathbb{E}(|\delta Y_T|^2) + \frac{1}{\mu^2} ||\delta_2 f||_\beta^2 \right], \tag{9}$$

$$||\delta Z||_\beta^2 = \frac{\lambda^2}{\lambda^2 - C} \left[e^{\beta T} \mathbb{E}(|\delta Y_T|^2) + \frac{1}{\mu^2} ||\delta_2 f||_\beta^2 \right]. \tag{10}$$

[9].

The Proof of Proposition 1 is done by [9] and is a follows.

Proof. Let $(Y, Z) \in \mathbb{H}_T^{2,d} \times \mathbb{H}_T^2(\mathbb{R}^{n \times d})$ be a solution of Equation (7). From Equation (8) using triangle inequality and Hölder inequality Theorem 2.1.5, we have

$$\begin{aligned} |Y_t| &\leq |\xi| + \left| \int_t^T f(s, Y_s, Z_s)\, ds \right| + \left| \int_t^T Z_s^*\, dW_s \right| \\ &\leq |\xi| + \int_t^T |f(s, Ys, Z_s)|\, ds + \left| \int_t^T Z_s^*\, dW_s \right|. \end{aligned}$$

Now taking the supremum, we have

$$\sup_{0 \leq t \leq T} |Y_t| \leq |\xi| + \int_0^T |f(s, Y_s, Z_s)|\, ds + \sup_{0 \leq t \leq T} \left| \int_t^T Z_s^*\, dW_s \right|. \tag{11}$$

We claim that $Y_t \in \mathbb{L}_T^{2,d}$. In fact, it is enough to show that each component of the right hand side of Equation (11) is in $\mathbb{L}_T^{2,d}$. Thus using the triangular inequality, Itô Isometry

(Theorem 2.1.13)and the Burkholder-Davis Gundy inequality (Theorem 2.1.3), we get

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_t^T Z_s^*\,dW_s\right|^2\right] &= \mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^T Z_s^*\,dW_s - \int_0^t Z_s^*\,dW_s\right|^2\right] \\
&\le \mathbb{E}\left[\sup_{0\le t\le T} 2\left(\left|\int_0^T Z_s^*\,dW_s\right|^2 + \left|\int_0^t Z_s^*\,dW_s\right|^2\right)\right] \\
&= 2\mathbb{E}\left[\left|\int_0^T Z_s^*\,dW_s\right|^2\right] + 2\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t Z_s^*\,dW_s\right|^2\right] \\
&= 2\mathbb{E}\left[\int_0^T |Z_s|^2\,ds\right] + 2\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t Z_s^*\,dW_s\right|^2\right] \\
&\le 2\mathbb{E}\left[\int_0^T |Z_s|^2\,ds\right] + 2C_2\mathbb{E}\left[\int_0^T |Z_s|^2\,ds\right].
\end{aligned}$$

Thus $\sup_{0\le t\le T}|\int_t^T Z_s^*\,dW_s| \in \mathbb{L}_T^{2,1}$. Since $\xi \in \mathbb{L}_T^{2,1}$, f is uniformly Lipschitz, $f(\cdot,0,0) \in \mathbb{H}_T^{2,1}$, $X \in \mathbb{H}_T^{2,d}$, $Y \in \mathbb{H}_T^{2,n\times d}$, then $|\xi| + \int_0^T |f(s,X_s,Y_s)|\,ds \in \mathbb{L}_T^{2,1}$, thus $Y_t \in \mathbb{L}_T^{2,1}$.

Now consider two solutions (Y^1, Z^1) and (Y^2, Z^2) associated with (f^1,ξ^1) and (f^2,ξ^2), respectively. Let $\delta Y_s = Y_s^1 - Y_s^2$ such that

$$\delta Y_s = \xi^1 - \xi^2 + \int_t^T [f^1(s,Y_s^1,Z_s^1) - f^2(s,Y_s^2,Z_s^2)]\,ds - \int_t^T \delta Z_s^*\,dW_s.$$

From the Itô's formula Equation (3) applied from $s=t$ to $s=T$ to the semi-martingale $e^{\beta t}|\delta Y_t|^2$, we let $f(t,x) = e^{\beta t}|x|^2$ then substituting x with δY_t and y with δZ_t, we have

$$\begin{aligned}
e^{\beta T}|\delta Y_T|^2 - e^{\beta t}|\delta Y_t|^2 &= \int_t^T \beta e^{\beta s}|\delta Y_s|^2\,ds + \int_t^T 2e^{\beta s}\langle \delta Y_s, d\delta Y_s\rangle + \frac{1}{2}\int_t^T 2e^{\beta s}d\langle \delta Y_s, \delta Y_s\rangle \\
&= \int_t^T \beta e^{\beta s}|\delta Y_s|^2\,ds + \int_t^T 2e^{\beta s}\langle \delta Y_s, d\delta Y_s\rangle + \int_t^T e^{\beta s}|\delta Z_s|^2\,ds.
\end{aligned}$$

Then

$$e^{\beta T}|\delta Y_T|^2 - 2\int_t^T e^{\beta s}\langle \delta Y_s, d\delta Y_s\rangle = e^{\beta t}|\delta Y_t|^2 + \int_t^T \beta e^{\beta s}|\delta Y_s|^2\,ds + \int_t^T e^{\beta s}|\delta Z_s|^2\,ds.$$

But we have

$$2\int_t^T e^{\beta s}\langle \delta Y_s, d\delta Y_s\rangle = -2\int_t^T e^{\beta s}\langle \delta Y_s, (f^1(s,Y_s^1,Z_s^1) - f^2(s,Y_s^2,Z_s^2))\,ds\rangle + 2\int_t^T e^{\beta s}\langle \delta Y_s, \delta Z_s^*\,dW_s\rangle.$$

Thus we have

$$\begin{aligned}&e^{\beta t}|\delta Y_t|^2 + \int_t^T \beta e^{\beta s}|\delta Y_s|^2\,ds + \int_t^T e^{\beta s}|\delta Z_s|^2\,ds\\ &= e^{\beta T}|\delta Y_T|^2 + 2\int_t^T e^{\beta s}\langle \delta Y_s, (f^1(s,Y_s^1,Z_s^1) - f^2(s,Y_s^2,Z_s^2))\,ds\rangle\\ &- 2\int_t^T e^{\beta s}\langle \delta Y_s, \delta Z_s^*\,dW_s\rangle.\end{aligned}$$

Using the Burkholder-Davis Gundy inequality Theorem 2.1.3, we have

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_t^T e^{\beta s}\delta Z_s \delta Y_s\,dWs\right|\right] &\le C\mathbb{E}\left[\left(\int_0^T |\delta Y_s|^2|\delta Y_s|^2\,ds\right)^{\frac{1}{2}}\right]\\ &\le C\left(\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_s|^2\right]\right)\left(\mathbb{E}\left[\int_0^T |\delta Z_s|^2\,ds\right]\right)^{\frac{1}{2}}\\ &< \infty.\end{aligned}$$

Since $\sup_{s\le T}|\delta Y_s|$ belongs to $\mathbb{L}_T^{2,1}$, and $e^{\beta s}\delta Z_s\delta Y_s$ belongs to $\mathbb{H}_T^{2,n}$, and the stochastic integral
$\int_t^T e^{\beta s}\langle \delta Y_s, \delta Z_s^*\,dW_s\rangle$ is $\mathbb{P}-$integrable, with zero expectation. Using the triangle inequal-

ity and Lipschitz condition on f, we have

$$\begin{aligned}|f^1(s,Y_s^1,Z_s^1)-f^2(s,Y_s^2,Z_s^2)| &= |f^1(s,Y_s^1,Z_s^1)-f^1(s,Y_s^2,Z_s^2)\\ &\quad + f^1(s,Y_s^2,Z_s^2)-f^2(s,Y_s^2,Z_s^2)|,\\ &\le |f^1(s,Y_s^1,Z_s^1)-f^1(s,Y_s^2,Z_s^2)|\\ &\quad + |f^1(s,Y_s^2,Z_s^2)-f^2(s,Y_s^2,Z_s^2)|,\\ &\le C(|\delta Y_s|+|\delta Z_s|)+|\delta_2 f_s|.\end{aligned}$$

Using $2ab \le a^2\epsilon + \frac{b^2}{\epsilon}$ applied to $2yCz$ and $2yt$, we have the inequality,

$$2yCz \le y^2\epsilon + \frac{(Cz)^2}{\epsilon}.$$

Taking $\epsilon = C\lambda^2 > 0$, we have

$$2yCz \le y^2C\lambda^2 + \frac{Cz^2}{\lambda^2}.$$

For $2yt$, we have

$$2yt \le y^2\epsilon + \frac{t^2}{\epsilon}.$$

Taking $\epsilon = \mu^2 > 0$, we have

$$2yt \le y^2\mu^2 + \frac{t^2}{\mu^2}.$$

Now since $2y(Cz+t) = 2Cyz + 2yt$, we have the inequality

$$2y(Cz+t) \le \frac{Cz^2}{\lambda^2} + \frac{t^2}{\mu^2} + y^2(\mu^2 + C\lambda^2). \tag{12}$$

$$\begin{aligned}&\mathbb{E}\left[e^{\beta t}|\delta Y_t|^2\right] + \beta\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Y_s|^2\,ds\right] + \mathbb{E}\left[\int_t^T e^{\beta s}|\delta Z_s|^2\,ds\right]\\ &= \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + 2\mathbb{E}\left[\int_t^T e^{\beta s}\langle \delta Y_s, (f^1(s,Y_s^1,Z_s^1)-f^2(s,Y_s^2,Z_s^2))\,ds\rangle\right]\\ &\le \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \mathbb{E}\left(\int_t^T e^{\beta s}2\langle \delta Y_s, [C(|\delta Y_s|+|\delta Z_s|)+|\delta_2 f_s|]\rangle\,ds\right)\\ &= \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \mathbb{E}\left[\int_t^T e^{\beta s}\left(2C\langle \delta Y_s, |\delta Y_s|\rangle + 2C\langle \delta Y_s, |\delta Z_s|\rangle + 2\langle \delta Y_s, |\delta_2 f_s|\rangle\right)\,ds\right].\end{aligned}$$

Using $|\langle y, z\rangle| = |y||z|$, we have

$$\mathbb{E}\left[e^{\beta t}|\delta Y_t|^2\right] + \beta\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Y_s|^2\, ds\right] + \mathbb{E}\left[\int_t^T e^{\beta s}|\delta Z_s|^2\, ds\right]$$
$$\leq \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \mathbb{E}\left[\int_t^T e^{\beta s}\left(2C|\delta Y_s|^2 + 2|\delta Y_s|(C|\delta Z_s| + |\delta_2 f_s|)\right)\, ds\right].$$

Applying Equation (12) with $y = |\delta Y_s|$, $Z = |\delta Z_s|$, $t = |\delta_2 f_s|$ and $C = C$, we get

$$\mathbb{E}\left[e^{\beta t}|\delta Y_t|^2\right] + \beta\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Y_s|^2\, ds\right] + \mathbb{E}\left[\int_t^T e^{\beta s}|\delta Z_s|^2\, ds\right]$$
$$\leq \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \mathbb{E}\left[\int_t^T e^{\beta s}\left(2C|\delta Y_s|^2 + \frac{C|\delta Z_s|^2}{\lambda^2} + \frac{|\delta_2 f_s|^2}{\mu^2} + |\delta Y_s|^2\left(\mu^2 + C\lambda^2\right)\right)\, ds\right]$$
$$= \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \mathbb{E}\left[\int_t^T e^{\beta s}\left(C|\delta Y_s|^2(2+\lambda^2) + \frac{C|\delta Z_s|^2}{\lambda^2} + \frac{|\delta_2 f_s|^2}{\mu^2} + \mu^2|\delta Y_s|^2\right)\, ds\right]$$
$$= \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + [C(2+\lambda^2) + \mu^2]\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Y_s|^2\, ds\right]$$
$$+ \frac{C}{\lambda^2}\mathbb{E}\int_t^T e^{\beta s}|\delta Z_s|^2\, ds + \frac{1}{\mu^2}\mathbb{E}\int_t^T e^{\beta s}|\delta_2 f_s|^2\, ds,$$

which gives

$$\mathbb{E}[e^{\beta t}|\delta Y_t|^2] + \left(\beta - [C(2+\lambda^2) + \mu^2]\right)\mathbb{E}\int_t^T e^{\beta s}|\delta Y_s|^2\, ds + \left(1 - \frac{C}{\lambda^2}\right)\mathbb{E}\int_t^T e^{\beta s}|\delta Z_s|^2\, ds$$
$$\leq \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \frac{1}{\mu^2}\mathbb{E}\int_t^T e^{\beta s}|\delta_2 f_s|^2\, ds. \tag{13}$$

If we take $\beta \geq [C(2+\lambda^2) + \mu^2]$ and $C \leq \lambda^2$, we have

$$\mathbb{E}[e^{\beta t}|\delta Y_t|^2] \leq \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \mathbb{E}\int_t^T e^{\beta s}|\delta_2 f_s|^2 \frac{1}{\mu^2} ds.$$

Taking the integral from $0 \longrightarrow T$, using Fubini's theorem (Theorem 2.1.11) and Definition 2.2, we have the control of the norm for the process $|\delta Y|$ as

$$\begin{aligned}||\delta Y||_\beta^2 &\le Te^{\beta T}\mathbb{E}[|\delta Y_T|^2] + \int_t^T \frac{1}{\mu^2}||\delta_2 f_s||_\beta^2\, ds \\ &\le T\left[e^{\beta T}\mathbb{E}[|\delta Y_T|^2] + \frac{1}{\mu^2}||\delta_2 f_s||_\beta^2\right].\end{aligned}$$

The control of the process for $|\delta Z|$ from Equations (13) is

$$\frac{\lambda^2 - C}{\lambda^2}\mathbb{E}\int_t^T e^{\beta s}|\delta Z_s|^2\, ds \le e^{\beta T}\mathbb{E}[|\delta Y_T|^2] + \frac{1}{\mu^2}||\delta_2 f_s||_\beta^2.$$

That is,

$$\frac{\lambda^2 - C}{\lambda^2}||\delta Z||_\beta^2 \le e^{\beta T}\mathbb{E}[|\delta Y_T|^2] + \frac{1}{\mu^2}||\delta_2 f_s||_\beta^2.$$

Hence,

$$||\delta Z||_\beta^2 \le \frac{\lambda^2}{\lambda^2 - C}\left[e^{\beta T}\mathbb{E}[|\delta Y_T|^2] + \frac{1}{\mu^2}||\delta_2 f_s||_\beta^2\right].$$

□

We have developed the necessary tools we need to prove uniqueness and existence of a solution. A detailed proof of the following theorem by [4], is given in their article. As in [9] we will prove it using the Banach fixed point theorem and a priori estimates.

Theorem 2.3.2. *Given standard parameters* (f, ξ)*, there exists a unique pair* $(Y, Z) \in \mathbb{H}_T^{2,d} \times \mathbb{H}_T^2(\mathbb{R}^{n\times d})$ *which solves Equation* (7) *[9].*

The solution is often referred to as a square-integrable solution.

Proof. This proof is by [9], we expand on the proof. We use the Banach fixed point theorem (Theorem 2.1.4) for the mapping from $\mathbb{H}_{T,\beta}^{2,d} \times \mathbb{H}_{T,\beta}^2\mathbb{R}^{n\times d}$ onto itself, which maps (y, z) onto the solution (Y, Z) of the BSDE with generator $f(t, xy_t, z_t)$, that is

$$Y_t = \xi + \int_t^T f(s, y_s, z_s)\, ds - \int_t^T Z_s^*\, dW_s.$$

The assumption that (f, ξ) are standard parameters implies f is uniformly Lipschitz, $f(\cdot, 0, 0) \in \mathbb{H}_T^{2,d}$, and $\xi \in \mathbb{L}^{2,d}$. Thus $(f(t, y_t, z_t);\ t \in [0, T])$ belongs to $\mathbb{H}_T^{2,d}$. Now we show why the solution to the BSDE is defined as a pair of adaptable processes. Consider the continuous version M of a square integrable martingale $\mathbb{E}\left[\xi + \int_0^T f(s, Y_s, Z_s)\, ds|\mathcal{F}_t\right]$,

$$M_t = \mathbb{E}\left[\xi + \int\limits_0^T f(s, y_s, z_s)\, ds\middle| \mathcal{F}_t\right].$$

[9]. By the Martingale representation theorem (Theorem 2.1.2) there exists a unique integrable process $Z \in \mathbb{H}_{T,\beta}^{2,n\times d}$ such that

$$M_t = M_0 + \int\limits_0^t Z_s^*\, dW_s.$$

Define the adapted and continuous process [9]

$$Y_t = M_t - \int\limits_0^t f(s, y_s, z_s)\, ds.$$

Substitute for M_t, we have

$$\begin{aligned} Y_t &= M_0 + \int\limits_0^t Z_s^*\, dW_s - \int\limits_0^t f(s, y_s, z_s)\, ds \\ &= \xi + \int\limits_0^T f(s, y_s, z_s)\, ds - \int\limits_0^T Z_s^*\, dW_s + \int\limits_0^t Z_s^*\, dW_s - \int\limits_0^t f(s, y_s, z_s)\, ds \\ &= \xi + \int\limits_t^T f(s, y_s, z_s)\, ds - \int\limits_t^T Z_s^*\, dW_s. \end{aligned}$$

Since Y_t is adapted, we have

$$\begin{aligned} Y_t &= \mathbb{E}\left[Y_t|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[\xi + \int_t^T f(s,y_s,z_s)\,ds - \int_t^T Z_s^*\,dW_s|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[\xi + \int_t^T f(s,y_s,z_s)\,ds|\mathcal{F}_t\right] - \mathbb{E}\left[\int_t^T Z_s^*\,dW_s|\mathcal{F}_t\right] \\ &= \mathbb{E}\left[\xi + \int_t^T f(s,y_s,z_s)\,ds|\mathcal{F}_t\right]. \end{aligned}$$

Y is square integrable since f, ξ are square integrable. Let (y^1, z^1), and (y^2, z^2) be two elements of $\mathbb{H}^{2,d}_{T,\beta} \times \mathbb{H}^2_{T,\beta}\mathbb{R}^{n\times d}$, and let (Y^1, Z^1) and (Y^2, Z^2) be the associated solutions. By Proposition 1 applied with $C = 0$ and $\beta = \mu^2$, we have (f, ξ) standard parameters, $\delta_2 f_s = f(s, y_s^1, z_s^1) - f(s, y_s^2, z_s^2)$ and $\delta Y_T = 0$. Then from Equations (9) and (10), we have

$$\begin{aligned} ||\delta Y||_\beta^2 &\le \frac{T}{\beta}\mathbb{E}\left[\int_0^T e^{\beta s}|f(s,y_s^1,z_s^1) - f(s,y_s^2,z_s^2)|^2\,ds\right], \text{ and} \\ ||\delta Z||_\beta^2 &\le \frac{1}{\beta}\mathbb{E}\left[\int_0^T e^{\beta s}|f(s,y_s^1,z_s^1) - f(s,y_s^2,z_s^2)|^2\,ds\right]. \end{aligned}$$

Since f is Lipschitz with constant C and using $(a+b)^2 \le 2(a^2+b^2)$, we have

$$\begin{aligned} ||\delta Y||_\beta^2 + ||\delta Z||_\beta^2 &\le \left(\frac{T}{\beta} + \frac{1}{\beta}\right) C\mathbb{E}\left[\int_0^T e^{\beta s}(|\delta y| + |\delta z|)^2\,ds\right] \\ &\le \left(\frac{T}{\beta} + \frac{1}{\beta}\right) C\mathbb{E}\left[\int_0^T e^{\beta s}2(|\delta y|^2 + |\delta z|^2)\,ds\right] \\ &= \frac{2(1+T)C}{\beta}\left[||\delta y||_\beta^2 + ||\delta z||_\beta^2\right]. \end{aligned} \tag{14}$$

Choosing $\beta > 2(1+T)C$, we see that this mapping is a contraction from $\mathbb{H}^{2,d}_{T,\beta} \times \mathbb{H}^2_{T,\beta}\mathbb{R}^{n\times d}$ onto itself and that there exists a fixed point, which is a unique continuous solution of the BSDE. □

From Equation (14), we show that the Picard iterative sequence converges almost surely to the solution of the BSDE. The following Corollary is by El-Karoui et al. [9].

Corollary 1. Let β be such that $\beta > 2(1+T)C$. Let (Y^k, Z^k) be the sequence defined recursively by $(Y^0 = 0, Z^0 = 0)$, and

$$-dY_t^{k+1} = f(t, Y_t^k, Z_t^k)dt - \left(Z_t^{k+1}\right)^* dW_t, \;\; Y_T^{k+1} = \xi. \tag{15}$$

Then the sequence (Y^k, Z^k) converges to (Y, Z), $d\mathbb{P} \bigotimes dt$ a.s.(and in $\mathbb{H}^{2,d}_{T,\beta} \times \mathbb{H}^2_{T,\beta}\mathbb{R}^{n\times d}$) as $k \longrightarrow +\infty$.

Proof. The proof is by [9]. Let (Y^k, Z^k) be a sequence defined by Equation (15). Then by Equation (14), we have

$$\begin{aligned}
||Y^{k+1} - Y^k||^2_\beta + ||Z^{k+1} - Z^k||^2_\beta = ||\delta Y^k||^2_\beta + ||\delta Z^k||^2_\beta & \\
\leq \left(\frac{2(1+T)C}{\beta}\right) (||\delta Y^{k-1}||^2_\beta + ||\delta Z^{k-1}||^2_\beta) & \\
\leq \left(\frac{2(1+T)C}{\beta}\right)^2 (||\delta Y^{k-2}||^2_\beta + ||\delta Z^{k-2}||^2_\beta) & \\
\vdots \qquad & \\
\leq \left(\frac{2(1+T)C}{\beta}\right)^k (||\delta Y^{0}||^2_\beta + ||\delta Z^{0}||^2_\beta) & \\
= \left(\frac{2(1+T)C}{\beta}\right)^k (||Y^1 - Y^0||^2_\beta + ||Z^1 - Z^0||^2_\beta & \\
= \epsilon^k K, &
\end{aligned}$$

where $K = ||Y^1 - Y^0||^2_\beta + ||Z^1 - Z^0||^2_\beta$, and $\epsilon = \frac{2(1+T)C}{\beta} < 1$. Thus,

$$\begin{aligned}
\sum_k (||Y^{k+1} - Y^k||^2_\beta + ||Z^{k+1} - Z^k||^2_\beta) \leq K \sum_k \epsilon^k, \text{ and} \\
\sum_k ||Y^{k+1} - Y^k||^2_\beta + \sum_k ||Z^{k+1} - Z^k||^2_\beta \leq \frac{\epsilon(1-\epsilon^k)}{(1-\epsilon)} K.
\end{aligned}$$

As we take limits as $k \longrightarrow +\infty$, we have

$$\begin{aligned}
\sum_k ||Y^{k+1} - Y^k||^2_\beta + \sum_k ||Z^{k+1} - Z^k||^2_\beta \leq \frac{\epsilon}{(1-\epsilon)} K < +\infty \\
\leq \epsilon K,
\end{aligned}$$

and the results follow. $\square$

2.3.3 Levy Process

Let X be a Lévy process and we denote the left limit process as

$$X_{t_-} = \lim_{s \longrightarrow t} X_s,\ t > 0,\ s < t$$

and we denote the jump size at time t [13] as

$$\Delta X_t = X_t - X_{t_-}.$$

According to [13] we define the following transformations of X, set

$$X_t^{(i)} = \sum_{0<s\leq t} (\Delta X_s)^i,\ i \geq 2$$

where

$$X_t^1 = X_t.$$

The processes $X^{(i)} = \{X_t^{(i)}, t \geq 0\},\ i = 1, 2, \cdots$ are Lévy processes and we term them power jump processes [13]. The processes have jumps at the same points as the original Lévy process.

Hypothesis 1 The Lévy measure satisfies for some $\epsilon > 0$ and $\lambda > 0$

$$\int_{(-\epsilon,\epsilon)^c} \exp(\lambda|x|)\nu(dx) < \infty$$

which implies

$$\int_{-\infty}^{\infty} |x|^2\nu(dx) < \infty,\ i \geq 2$$

[13]. We also have

$$\mathbb{E}[X_t] = \mathbb{E}[X_t^1] = tm_1 < \infty$$

and that

$$\mathbb{E}[X_t] = \mathbb{E}\left[\sum_{0<s\leq t} (\Delta X_s)^{(i)}\right] = t\int_{-\infty}^{\infty} x^i\nu(dx) = m_it,\ i \geq 2$$

Then we denote the power jump process of i as

$$Y_t^i := X_t^{(i)} - \mathbb{E}[X_t^{(i)}] = X_t^{(i)} - m_it,\ i = 1, 2, 3, \cdots,$$

this is a normal martingale since for an integrable Lévy process X, the process $X_t - \mathbb{E}[X_t],\ t \geq 0$ is also a martingale. The process $Y^{(i)}$ is also referred to as Teugels martingales of order i [13]. Now lets denote $\mathcal{M}^2$ as the space of square integrable martingales M such that

$$\sup_t \mathbb{E}(M_t^2) < \infty \text{ and } M_0 = 0 \ \ a.s.$$

thus if $M \in \mathcal{M}^2$ then

$$\lim_{t \longrightarrow \infty} \mathbb{E}(M_t^2) = \mathbb{E}(M_\infty^2) < \infty$$

and

$$M_t = \mathbb{E}[M_\infty | \mathcal{F}_t]$$

hence each $M \in \mathcal{M}^2$ can be identified by its terminal value M_∞. Two martingales $M, N \in \mathcal{M}^2$ are strongly orthogonal denoted $M \times N$ if and only if their product is a uniformly integrable martingale [13]. We also have two random variables $X, Y \in L^2(\Omega, \mathcal{F})$ are weakly orthogonal denoted $X \perp Y$ if $\mathbb{E}[XY] = 0$ [13]. In [13] an orthonormalization procedure was applied to the martingales $Y^{(i)}$ to obtain pairwise strongly orthonormal martingales $\{H^{(i)}\}_{i=1}^{\infty}$ such that

$$H_t^{(i)} = C_{i,i}Y^{(i)} + C_{i,i-1}Y^{(i-1)} + \cdots + C_{i,1}Y^{(1)}.$$

They have also shown that

$$H_t^{(i)} = q_{i-1}(0)X_t + \sum_{0<s\leq t} \tilde{p}_i(\Delta X_s) - t\mathbb{E}\left[\sum_{0<s\leq t} \tilde{p}_i(\Delta X_s)\right] - tq_{i-1}(0)\mathbb{E}[X_1]$$

then

$$\Delta H_t^{(i)} = p_i(\Delta X_t), \text{ for each } i \geq 1$$

where

$$\begin{aligned} q_{i-1}(x) &= C_{i,i}X^{i-1} + C_{i,i-1}X^{i-2} + \cdots + C_{i,1} \\ p_i(x) &= xq_{i-1}(x) = C_{i,i}X^i + C_{i,i-1}X^{i-1} + \cdots + C_{i,1}x \\ \tilde{p}_i(x) &= x(q_{i-1}(x) - q_{i-1}(0)) = C_{i,i}X^i + C_{i,i-1}X^{i-1} + \cdots + C_{i,1}x^2. \end{aligned}$$

We do not give any proofs of the above orthogonalization as it is beyond the scope of this thesis. Now consider the following BSDE

$$\begin{aligned} -dY_t &= f(t, Y_{t-}, Z_t)dt - \sum_{i=1}^{\infty} Z_t^{(i)} dH_t^{(i)} \qquad (16) \\ Y_T &= \xi \end{aligned}$$

where $H_t^{(i)}$ is the orthonormalized Teugels martingales of order i associated with the Lévy process X_t [14]. Given the function f is measurable such that $f(\cdot,0,0) \in \mathbb{H}_T^2$ and uniformly Lipschitz in y, z, that is, $\exists C > 0$ such that $dP \otimes dt$ almost surely $\forall(y_1, z_1)$and $(y_2, z_2) \in \mathbb{R} \times l^2$

$$|f(t, y_1, z_1) - f(t, y_2, z_2)| \leq C(|Y_1 - Y_2| + |Z_1 - Z_2|)$$

and $\xi \in L_T^2$. Then we say (f, ξ) is standard data as in Definition 15. The solution of the BSDE (16) is a pair of processes $\{(Y_t, Z_t),\ 0 \leq t \leq T\} \in \mathbb{H}_T^2 \times M_T^2(l^2)$ such that $\forall t \in [0, T]$ we have

$$Y_t = \xi + \int_t^T f(s, Y_{s-}, Z_s)ds - \sum_{i=1}^{\infty} \int_t^T Z_s^{(i)} dH_s^{(i)}.$$

Given standard data (f, ξ) we can apply Theorem 2.3.2 for the existence and uniqueness of the solution. The proof is similar to the previous section. Note the process Z_t is given by the representation theorem by [13] applied to the square integrable variable

$$\xi + \int_0^T f(s, Y_{s-}, Z_s)ds.$$

The following theorem by [14] is similar to Proposition 1 by [9] thus we will give the sketch of the proof here and some of the detail is similar to the proof of Proposition 1 in the preceding Section.

Theorem 2.3.4. *Given standard data* $(f^i, \xi^i), i = 1, 2$ *and* $(Y^i, Z^i), i = 1, 2$ *as the unique solution of BSDE* (16) *and respective data, then*

$$\mathbb{E}\left[\int_0^T \left(|Y_{s-}^1 - Y_{s-}^2|^2 + \sum_{i=1}^{\infty} |Z_s^{1(i)} - Z_s 2(i)|^2\right)\right]$$
$$\leq C\left(\mathbb{E}[|\xi^1 - \xi^2|^2] + \mathbb{E}\left[\int_0^T |f^1(s, Y_{s-}, Z_s) - f^2(s, Y_{s-}, Z_s)|^2 ds\right]\right)$$

The proof was done by [14] and we expand and fill in the detail in the proof as follows.

Proof. Applying the Itô's formula Equation (4) to $|Y_s^1 - Y_s^2|^2$ from $s = t \longrightarrow s = T$ we have

$$|Y_T^1 - Y_T^2|^2 - |Y_t^1 - Y_t^2|^2 = 2\int_t^T |Y_{s-}^1 - Y_{s-}^2| d|Y_s^1 - Y_s^2| + \int_t^T d[Y^1 - Y^2, Y^1 - Y^2]_s \quad (17)$$

since

$$-d(Y_t^1 - Y_t^2) = (f^1(s, Y_{s-}^1, Z_s^1) - f^2(s, Y_{s-}^2, Z_s^2))dt - \sum_{i=1}^{\infty}(Z_t^{1(i)} - Z_t^{2(i)})dH_t^{(i)},$$

$$d[Y^1 - Y^2, Y^1 - Y^2]_t = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(Z_t^{1(i)} - Z_t^{2(i)})(Z_t^{1(j)} - Z_t^{2(j)})d[H^{(i)}, H^{(j)}]$$

and

$$\langle H^{(i)}, H^{(j)}\rangle_t = \delta_{ij} t.$$

Taking expectation to Equation (17) and apply the above relations we have

$$\begin{aligned}
&\mathbb{E}\left[|Y_t^1 - Y_t^2|^2\right] + \sum_{i=1}^{\infty}\mathbb{E}\left[\int_t^T |Z_s^{1(i)} - Z_s^{2(i)}|^2 ds\right] \\
&= \mathbb{E}[|\xi^1 - \xi^2|^2] + 2\mathbb{E}\left[\int_t^T Y_{s-}^1 - Y_{s-}^2||f^1(s, Y_{s-}^1, Z_s^1) - f^2(s, Y_{s-}^2, Z_s^2)|\right]
\end{aligned}$$

Using similar computations as in the proof of Proposition 1 and the fact that f^2 is Lipschitz with Lipschitz constant C we obtain

$$\begin{aligned}
&\mathbb{E}\left[|Y_t^1 - Y_t^2|^2\right] + \frac{1}{2}\mathbb{E}\left[\int_t^T \sum_{i=1}^{\infty} |Z_s^{1(i)} - Z_s^{2(i)}|^2 ds\right] \\
&\leq \mathbb{E}[|\xi^1 - \xi^2|^2] + (1 + 2C + 2C^2)\mathbb{E}\left[|Y_{s-}^1 - Y_{s-}^2|^2 ds\right] \\
&+ \mathbb{E}\left[\int_t^T |f^1(s, Y_{s-}, Z_s) - f^2(s, Y_{s-}, Z_s)|^2 ds\right],
\end{aligned}$$

applying the Gronwall inequality (Theorem 2.1.10) then the result follows. □

2.4 Feynman-Kac Theorem

In this section we discuss the Feynman-Kac theorem for BSDEs driven by a Brownian motion and a Lévy process. The Feynman-Kac Theorems are useful in many applications, for example you have a PDE which you can not solve in closed form or when you have an expectation which you can not solve in closed form [2]. In financial applications the expectation giving the arbitrage price of the contract can not always be evaluated in closed form and numerical approximations are preferred. We either do simulations of the SDEs or solve the resulting PDE (or systems of PDEs) which is provided by the Feynman-Kac theorems [24].

Model 1

Under Model 1, all theorems, definitions and proofs were done by [9] we will expand the proofs in their article. For any $(t, x) \in [0, T] \times \mathbb{R}^p$, consider the following stochastic differential equation on $[0, T]$:

$$\begin{aligned} dX_s = b(s, X_s)\, ds + \sigma(s, X_s)\, dW_s, \quad t \le s \le T, \\ X_s = x, \quad 0 \le s \le t. \end{aligned} \tag{18}$$

We denote the solution of Equation (18) by $(X_s^{t,x},\ 0 \le s \le T)$. We consider the associated BSDE

$$\begin{aligned} -dY_s = f(s, X_s^{t,x}, Y_s, Z_s)\, ds - Z_s^*\, dW_s, \\ Y_T = \Psi(X_T^{t,x}). \end{aligned} \tag{19}$$

We denote the solution of Equation (19) by $\{(Y_s^{t,x}, Z_s^{t,x}), 0 \le s \le T\}$. The coupled system of Equations (18) and (19) is termed a FBSDE or a BSDE associated with a FSDE, and the solution is denoted by $\{(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}), 0 \le s \le T\}$.

The function f is an $\mathbb{R}^d$ valued Borel function defined on $[0, T] \times \mathbb{R}^p \times \mathbb{R}^d \times \mathbb{R}^{n\times d}$, and Ψ is an $\mathbb{R}^d$ valued Borel function defined on $\mathbb{R}^p$. The coefficient b is a $\mathbb{R}^p$ valued function defined on $[0, T] \times \mathbb{R}^p$, and σ is an $\mathbb{R}^{p\times n}$ valued function defined on $[0, T] \times \mathbb{R}^p$. Standard Lipschitz assumptions are required on the coefficients, that is, there exists a Lipschitz constant $C > 0$ such that

$$\begin{aligned} |b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \le C(1 + |x - y|), \text{ and} \\ |f(t, x, y_1, z_1) - f(t, x, y_2, z_2)| \le C(|z_1 - z_2| + |y_1 - y_2|). \end{aligned}$$

Finally, we suppose that there exists a constant C such that for each (s, x, y, z),

$$|b(t,x)| + |\sigma(t,x)| \le C(1+|x|), \text{ and}$$
$$|f(t,x,y,z)| + |\Psi(x)| \le C(1+|x|^p),$$

for real $p \ge \frac{1}{2}$.

Generalization of the Feynman-Kac formula.

Proposition 2. Let ν be a function of class $\mathcal{C}^{1,2}$ (or smooth enough to be able to apply Itô formula to $\nu(s, X_s^{t,x})$) and suppose that there exists a constant C such that, for each (s, x),

$$|\nu(s,x)| + |\sigma(s,x)^* \partial_x \nu(s,x)| \le C(1+|x|).$$

Also, ν is supposed to be the solution of the following quasi-linear parabolic partial differential equation

$$\partial_t \nu(t,x) + \mathcal{L}\nu(t,x) + f(t,x,\nu(t,x),\sigma(t,x)^*\partial_x \nu(t,x)) = 0, \quad (20)$$
$$\nu(T,x) = \Psi(x),$$

where $\partial_x \nu$ is the gradient of ν, and $\mathcal{L}_{(t,x)}$ denotes the second order differential operator

$$\mathcal{L}_{(t,x)} = \sum_{i,j} a_{ij}(t,x)\partial^2_{x_i x_j} + \sum_i b_i(t,x)\partial_{x_i}, \;\; a_{ij} = \frac{1}{2}[\sigma\sigma^*]_{ij}.$$

Then $\nu(t,x) = Y_t^{t,x}$, where $\{(Y_s^{t,x}, Z_s^{t,x}),\ t \le s \le T\}$ is the unique solution of BSDE Equation (19). Also

$$(Y_s^{t,x}, Z_s^{t,x}) = (\nu(s, X_s^{t,x}), \sigma(s, X_s^{t,x})^* \partial_x \nu(s, X_s^{t,x})), \quad t \le s \le T.$$

Proof. $X_s^{t,x}$ is a d-dimensional Itô process. The partial derivatives with respect to time and space are d-dimensional vectors $\partial_t \nu(s, X_s^{t,x})$ and $\partial_x \nu(s, X_s^{t,x})$, respectively. We also have $\partial_x^2 \nu(s, X_s^{t,x})$, a $d \times d$ Hessian matrix of ν with respect to $X_s^{t,x}$. Applying the Itô formula Equation (3) in differential form to $\nu(s, X_s^{t,x})$, we have

$$d\nu(s, X_s^{t,x}) = \partial_t \nu(s, X_s^{t,x})ds + \left[\partial_x \nu(s, X_s^{t,x})\right]^* dX_s^{t,x} + \frac{1}{2}\left[dX_s^{t,x}\right]^* \partial_x^2 \nu(s, X_s^{t,x}) dX_s^{t,x}.$$

Substituting for $dX_s^{t,x} = b(s, X_s^{t,x})ds + \sigma(s, X_s^{t,x})dW_s$, we have

$$\begin{aligned} d\nu(s, X_s^{t,x}) &= \partial_t \nu(s, X_s^{t,x})ds + \left[\partial_x \nu(s, X_s^{t,x})\right]^* \left[b(s, P_s^{t,x})ds + \sigma(s, X_s^{t,x})dW_s\right] \\ &+ \frac{1}{2}\left[dX_s^{t,x}\right]^* \partial_x^2 \nu(s, X_s^{t,x})dX_s^{t,x} \\ &= \partial_t \nu(s, X_s^{t,x})ds + \left[\partial_x \nu(s, X_s^{t,x})\right]^* b(s, X_s^{t,x})ds \\ &+ \left[\partial_x \nu(s, X_s^{t,x})\right]^* \sigma(s, X_s^{t,x})dW_s + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\left[dX_{j,s}^{t,x}\frac{\partial^2 \nu(s, X_s^{t,x})}{\partial x_i \partial x_j}\right] dX_{i,s}^{t,x}. \end{aligned}$$

Since W_i is independent of W_j when $i \neq j$, that is W_i is orthogonal to W_j, then for $i = j$ we have

$$\begin{aligned} d\nu(s, X_s^{t,x}) = \partial_t \nu(s, X_s^{t,x})ds + \left[\partial_x \nu(s, X_s^{t,x})\right]^* b(s, X_s^{t,x})ds \\ + \left[\partial_x \nu(s, X_s^{t,x})\right]^* \sigma(s, X_s^{t,x})dW_s + \sum_{i,j=1}^{d}\frac{1}{2}[(\sigma\sigma^*)_{i,j}(s, X_s^{t,x})]\frac{\partial^2 \nu(s, X_s^{t,x})}{\partial x_i \partial x_j}ds. \end{aligned}$$

However,

$$\mathcal{L}_{s,X_s^{t,x}} = \sum_{i,j}\frac{1}{2}[(\sigma\sigma^*)_{i,j}(s, X_s^{t,x})\partial^2_{x_i x_j} + \sum_i b(s, X_s^{t,x})\partial_{x_i}.$$

Then

$$d\nu(s, X_s^{t,x}) = \left[\partial_t \nu(s, X_s^{t,x}) + \mathcal{L}_{(s,X_s^{t,x})}\nu(s, X_s^{t,x})\right] ds + \left[\partial_x \nu(s, X_s^{t,x})\right]^* \sigma(s, X_s^{t,x})dW_s.$$

Since ν solves Equation (20), it follows that

$$\begin{aligned} -d\nu(s, X_s^{t,x}) &= f(s, , X_s^{t,x}, \nu(s, X_s^{t,x}), \sigma(s, X_s^{t,x})^* \partial_x \nu(s, X_s^{t,x}))ds \\ &\quad - \partial_x \nu(s, X_s^{t,x})^* \sigma(s, X_s^{t,x})dW_s, \\ \nu(T, X_T^{t,x}) &= \Psi(X_T^{t,x}). \end{aligned}$$

Thus by uniqueness of solutions of BSDE, $\{\nu(s, X_s^{t,x}), \sigma(s, X_s^{t,x})^*\partial_x \nu(s, X_s^{t,x}),\ s \in [0, T]\}$ is the unique solution of the BSDE Equation (19), and the result is obtained. □

We now show that conversely, in certain cases the solution of BSDE Equation (19) corresponds to the solution of the PDE Equation (20). If $d = 1$, we can use the comparison theorem to show that if $b,\ \sigma,\ f,$ and Ψ satisfy the assumptions by [9], and if f, and Ψ are supposed to be uniformly continuous with respect to x, then $u(t, x)$ is the viscosity solution of Equation (20) [6].

Definition 16. Suppose $u \in \mathcal{C}([0, T] \times \mathbb{R}^p)$ satisfies $u(T, x) = \Psi(x),\ x \in \mathbb{R}^p$. Then

u is called a viscosity sub-solution or super-solution of PDE Equation (20) if, for each $(t,x) \in [0,T] \times \mathbb{R}^p$ and $\phi \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^p)$ such that $\phi(t,x) = u(t,x)$ and (t,x) is a minimum or maximum of $\phi - u$,

$$\partial_t \phi(t,x) + \mathcal{L}\phi(t,x) + f(t,x,\phi(t,x),\sigma(t,x)^* \partial_x \phi(t,x)) \geq 0, \qquad \text{sub-solution}$$

or

$$\partial_t \phi(t,x) + \mathcal{L}\phi(t,x) + f(t,x,\phi(t,x),\sigma(t,x)^* \partial_x \phi(t,x)) \leq 0. \qquad \text{super-solution}$$

Moreover, u is called a viscosity solution of PDE Equation (20) if it is both a viscosity sub-solution and a viscosity super-solution of Equation (20).

Theorem 2.4.1. *We suppose that $d = 1$ and that f and Ψ are uniformly continuous with respect to x. Then the function u defined by $u(t,x) = Y_t^{t,x}$ is a viscosity solution of PDE* (20)*. Furthermore, if we suppose that for each $R > 0$ there exists a continuous function $m_R : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ such that $m_R(0) = 0$ and*

$$|f(t,x,y,z) - f(t,x',y,z)| \leq m_R(|x - x'|(1 + |z|)), \tag{21}$$

for all $t \in [0,T]$, $|x|$, $|x'| \leq R$, and $|z| \leq R$ for $z \in \mathbb{R}^n$, then u is the unique viscosity solution of Equation (20)*.*

Proof. The countinuity of u with respect to (t,x) is given by [9]. We consider Equation (20) in the viscosity sense to avoid restrictive assumptions on the coefficients of our model. We prove that u is a viscosity sub-solution and the proof for the super-solution is similar. Let $(t,x) \in [0,T] \times \mathbb{R}^p$ and $\phi \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^p)$ be such that

$$\phi(t,x) = u(t,x),$$

and

$$\phi \geq u, \text{ on } [0,T] \times \mathbb{R}^p$$

without loss of generality we suppose that $\phi \in \mathcal{C}^\infty$ with bounded derivatives. For $h > 0$ we have

$$\phi(t+h, X_{t+h}^{t,x}) - \phi(t,x) - \int_t^{t+h} f(s, X_s^{t+h}, Y_s^{t+h}, Z_s^{t,x}) ds + \int_t^{t+h} Z_s^{t,x} dW_s \geq 0,$$

at this moment we are not sure yet if $Z_s^{t,x}$ converges to $\sigma(t,x)^* \partial_x \phi(t,x)$. Now consider

$(\bar{Y}_s, \bar{Z}_s)$ in the interval $t \leq s \leq t+h$ to be the solution of the BSDE

$$\bar{Y}_s = \phi(t+h, X_{t+h}^{t,x}) + \int_s^{t+h} f(r, X_r^{t+h}, \bar{Y}_r, \bar{Z}_r)dr - \int_s^{t+h} \bar{Z}_r dW_r$$

Note $(\bar{Y}, \bar{Z})$ has the same generator as (Y, Z) but different terminal condition which is $\phi(t+h, X_{t+h}^{t,x}) \geq Y_{t+h} = u(t+h, X_{t+h}^{t,x})$. By the Comparison Theorem 2.1.14 and continuity process shown in [9] it follows that $\bar{Y}_t \geq Y_t^{t,x} = u(t,x) = \phi(t,x)$.
We have to show that by letting $h \longrightarrow 0$, $\bar{Y}_s \longrightarrow \phi(t,x)$ and $\bar{Z}_s \longrightarrow \partial_x\phi(t,x)^*\sigma(t,x)$.
Let

$$G(s,x) = \partial_s\phi(s,x) + \mathcal{L}\phi(s,x) + f(s,x,\phi(s,x), \partial_x\phi(s,x)\sigma(s,x))$$

for $t \leq s \leq t+h$. We need to show $G(t,x) \geq 0$. Let

$$\tilde{Y}_s = \bar{Y}_s - \phi(s, X_s^{t,x}) - \int_s^{t+h} G(r,x)dr$$

and

$$\tilde{Z}_s = \bar{Z}_s - \partial_x\phi\sigma(s, X_s^{t,x})$$

We want to show that $\tilde{Y}_t = h\epsilon(h)$ where $\epsilon(h) \longrightarrow 0$ as $h \longrightarrow 0$. Applying the Itô formula, $(\tilde{Y}_s, \tilde{Z}_s)$, $t \leq s \leq t+h$ is the unique solution of the BSDE

$$\tilde{Y}_s = \int_s^{t+h} f\left(r, X_r^{t,x}, \phi(r, X_r^{t,x}) + \tilde{Y}_s + \int_s^{t+h} G(v,x)dv, \partial_x\phi\sigma(r, X_r^{t,x}) + \tilde{Z}_r\right) dr$$
$$\int_s^{t+h} \left[(\partial_r\phi + \mathcal{L}\phi)(r, X_r^{t,x}) - G(r,x)\right] dr - \int_s^{t+h} \tilde{Z}_r dW_r. \tag{22}$$

First we show that $(\tilde{Y}_s, \tilde{Z}_s) \longrightarrow (0,0)$ as $h \longrightarrow 0$. Let $(Y^1, Z^1) = (\tilde{Y}, \tilde{Z})$ and $(Y^2, Z^2) = (0,0)$ applying a priori estimate we have

$$\mathbb{E}\left[\sup_{t \leq s \leq t+h} |\tilde{Y}_s|^2\right] + \mathbb{E}\left[\int_s^{t+h} |\tilde{Z}_s|^2 ds\right] \leq K\mathbb{E}\left[\int_s^{t+h} |\delta(r,h)|^2 dr\right]$$

where

$$\delta(r,h) = -G(r,x) + (\partial_r\phi + \mathcal{L}\phi)(r, X_r^{t,x}) \\ + f(r, X_r^{t,x}, \phi(r, X_r^{t,x}) + \int_r^{t+h} G(v,x)dv, \sigma(r, X_r^{t,x})^* \partial_x \phi(r, X_r^{t,x}))$$

Since

$$\sup_{t\le s\le t+h} \mathbb{E}(|X_s^{t,x} - x|^2) \longrightarrow 0 \text{ as } h \longrightarrow 0$$

since ϕ and all coefficients and their derivatives are uniformly continuous with respect to x thus

$$\lim_{h\longrightarrow 0} \sup_{t\le s\le t+h} \mathbb{E}\left[|\delta(r,h)|^2\right] = 0$$

hence

$$\mathbb{E}\left[\sup_{t\le s\le t+h} |\tilde{Y}_s|^2\right] + \mathbb{E}\left[\int_s^{t+h} |\tilde{Z}_s|^2 ds\right] \le K\mathbb{E}\left[\int_s^{t+h} |\delta(r,h)|^2 dr\right] \le h\epsilon(h). \quad (23)$$

where $\epsilon(h) \longrightarrow 0$ as $h \longrightarrow 0$. □

By taking expectation of Equation (22) we have

$$\tilde{Y}_t = \mathbb{E}(\tilde{Y}_t) = \mathbb{E}\left[\int_t^{t+h} \delta'(r,h)dr\right]$$

where

$$\delta'(r,h) = -G(r,x) + (\partial_r\phi + \mathcal{L}\phi)(r, X_r^{t,x}) \\ + f\left(r, X_r^{t,x}, \phi(r, X_r^{t,x}) + \tilde{Y}_r + \int_r^{t+h} G(v,x)dv, \sigma(r, X_r^{t,x})^* \partial_x \phi(r, X_r^{t,x}) + \tilde{Z}_r\right).$$

Since f is Lipschitz we have

$$|\delta'(r,h) - \delta(r,h)| \le K(|\tilde{Y}_r| + |\tilde{Z}_r|)$$

and by Equation (23) we have $\tilde{Y}_t = h\epsilon(h)$ hence since $\bar{Y}_t \geq \phi(t,x)$ we have

$$\int_t^{t+h} G(r,x)dr \geq -h\epsilon(h)$$

so letting $h \longrightarrow 0$ we obtain

$$G(t,x) = \partial_t\phi(t,x) + \mathcal{L}\phi(t,x) + f(t,x,\phi(t,x),\partial_x\phi(t,x)\sigma(t,x)) \geq 0.$$

Following the same procedure we show that u is a super-solution of Equation (20) therefore we have u is a viscosity solution for Equation (20).

Model 2

All Theorems, definitions and proof under this model were done by [14]. Now consider the BSDE (16) or equivalently

$$Y_t = g(X_T) + \int_t^T f(s, Y_{s-} Z_s)ds - \sum_{i=1}^{\infty}\int_t^T Z_s^{(i)} dH_s^{(i)}$$

we have X_t a Lévy process which does not have a Brownian motion part such that

$$X_t = a + L_t$$

where L_t is a pure Jump process with Lévy measure $\nu(dx)$. We also have $g(X_T)$ is square integrable. Consider the partial differential integral equation (PDIE) satisfied by $\theta = \theta(t,x)$

$$\frac{\partial\theta}{\partial t}(t,x) + \int_{\mathbb{R}} \theta^{(1)}(t,x,y)\nu(dy) + a'\frac{\partial\theta}{\partial x}(t,x) + f(t,\theta(t,x),\{\theta^{(i)}(t,x)\}_{i=1}^{\infty}) = 0. \quad (24)$$

$$\theta(T,x) = g(x)$$

where

$$\theta^{(1)}(t,x,y) = \theta(t,x+y) - \theta(t,x) - \frac{\partial\theta}{\partial x}(t,x)y, \quad (25)$$

and

$$a' = a + \int_{\{|y|\geq 1} y\nu(dy).$$

Thus we have,

$$\theta^{(1)}(t,x)=\int\limits_{\mathbb{R}}\theta^{(1)}(t,x,y)p_1(y)\nu(dy)+\frac{\partial\theta}{\partial x}(t,x)\left(\int\limits_{\mathbb{R}}y^2\nu(dy)\right)^{\frac{1}{2}} \tag{26}$$

for $i\geq 2$ we have

$$\theta^{(i)}(t,x)=\theta^{(1)}(t,x,y)p_i(y)\nu(dy).$$

Proposition 3. Suppose θ is a $\mathcal{C}^{1,2}$ function such that the first and second partial derivatives in x are bounded by a polynomial of x, uniformly in t. Then the unique adapted solution of Equation (16) is given by

$$\begin{aligned}
Y_t &= \theta(t,X_t),\\
Z_t^{(1)} &= \int\limits_{\mathbb{R}}\theta^{(1)}(t,X_{t-},Y)p_1(y)\nu(dy)+\frac{\partial\theta}{\partial x}(t,x)\left(\int\limits_{\mathbb{R}}y^2\nu(dy)\right)^{\frac{1}{2}}\\
Z_t^{(i)} &= \int\limits_{\mathbb{R}}\theta^{(1)}p_i(y)\nu(dy),\ i\geq 2
\end{aligned}$$

where $\theta=\theta(t,x)$ is the solution of the PDIE (24) and $\theta^{(1)}(t,x,y)$ is given by (25) [14].

Proof. We consider the following by [14].

Lemma 1. Let $h:\Omega\times[0,T]\times\mathbb{R}\longmapsto\mathbb{R}$ be a random function measurable with respect to $\mathcal{P}\otimes\mathbb{B}_{\mathbb{R}}$ such that

$$|h(s,y)|\leq a_s(y^2\wedge|y|)\ \ a.s, \tag{27}$$

where $\{a_s;0\leq s\leq T\}$ is the non-negative predictable process such that

$$\mathbb{E}\left[\int\limits_0^T a_s^2ds\right]<\infty$$

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

$$\sum_{t<s\leq T}h(s,\triangle X_s)=\sum_{i=1}^{\infty}\int\limits_t^T\langle h(s,y),p_i\rangle dH_s^{(i)}+\int\limits_t^T\int\limits_{\mathbb{R}}h(s,y)\nu(dy)ds$$

Proof. We prove Lemma 1 then we use it to prove Proposition 3. Since from Equation

(27) it implies that

$$\mathbb{E}\left[\int_0^T\int_{\mathbb{R}} |h(s,y)|^2\nu(dy)ds\right] < \infty$$

thus we have

$$M_t = \sum_{0<s\le t} h(s,\triangle X_s) - \int_0^T\int_{\mathbb{R}} h(s,y)\nu(dy)ds$$

is a square integrable martingale. Then by the predictable representation theorem [13] $\exists\phi$ such that

$$M_t = \sum_{i=1}^{\infty}\int_0^t \phi_s^{(i)} dH_s^{(i)}$$

and recognising that

$$\langle H^{(i)}, H^{(j)}\rangle_t = t\delta_{ji}$$

thus

$$\langle M, H^{(i)}\rangle_t = \int_0^t \phi_s^{(i)} ds \tag{28}$$

We also have

$$\triangle M_s \triangle H_s^{(i)} = h(s,\triangle X_s)p_i(\triangle X_s)$$

thus we have

$$\langle M, H^{(i)}\rangle_t = \int_0^t\int_{\mathbb{R}} h(s,y)p_i(y)\nu(dy)ds \tag{29}$$

from Equations (28) and (29) we have

$$\phi_s^{(i)} = \int_{\mathbb{R}} h(s,y)p_i(y)\nu(dy)$$

then the result follows. □

Now we apply the Itô formula Theorem 2.1.8 to $\theta(s, X_s)$ from $s = t$ to $s = T$ we have

$$\theta(T, X_T) - \theta(t, X_t) = \int_t^T \frac{\partial\theta}{\partial t}(s, X_{s-})ds + \int_t^T \frac{\partial\theta}{\partial x}(s, X_{s-})ds$$
$$+\sum_{t<s\le T}\left[\theta(s,X_s) - \theta(s,X_{s-}) - \frac{\partial\theta}{\partial x}(s,X_{s-})\triangle X_s\right] \tag{30}$$

applying Lemma 1 to $h(s,y) = \theta(s, X_s) - \theta(s, X_{s-}) - \frac{\partial\theta}{\partial x}(s, X_{s-}) \triangle X_s$ we obtain

$$\sum_{t<s\leq T} \left[\theta(s, X_s) - \theta(s, X_{s-}) - \frac{\partial\theta}{\partial x}(s, X_{s-}) \triangle X_s\right]$$
$$= \sum_{i=1}^{\infty} \int_t^T \left(\int_{\mathbb{R}} \theta^{(i)}(s, X_{s-}, y) p_i(y) \nu(dy) \right) dH_s^{(i)}$$
$$+ \int_t^T \int_{\mathbb{R}} \theta^{(i)}(s, X_{s-}, y) p_i(y) \nu(dy) ds \tag{31}$$

Substituting Equation (31) into (30) we get

$$g(X_T) - \theta(t, X_t) = \int_t^T \frac{\partial\theta}{\partial t}(s, X_{s-}) ds + \int_t^T \frac{\partial\theta}{\partial x}(s, X_{s-}) ds$$
$$+ \sum_{i=1}^{\infty} \int_t^T \left(\int_{\mathbb{R}} \theta^{(1)}(s, X_{s-}, y) p_i(y) \nu(dy) \right) dH_s^{(i)}$$
$$+ \int_t^T \int_{\mathbb{R}} \theta^{(i)}(s, X_{s-}, y) p_i(y) \nu(dy) ds$$

Now from Equation (24) we have

$$g(X_T) - \theta(t, X_t)$$
$$= \int_t^T f(s, \theta(s, X_{s-}), \{\theta^k(s, X_{s-})\}_{k=1}^{\infty})$$
$$+ \int_t^T \left[\int_{\mathbb{R}} \theta^{(i)}(s, X_{s-}, y) p_1(y) \nu(dy) + \frac{\partial\theta}{\partial x}(s, X_{s-}) \left(\int_{\mathbb{R}} y^2 \nu(dy) \right)^{\frac{1}{2}} \right] dH_s^{(1)}$$
$$+ \sum_{i=2}^{\infty} \int_t^T \left[\int_{\mathbb{R}} \theta^{(1)}(s, X_{s-}, Y) p_i(y) \nu(dy) \right] dH_s^{(i)}$$

the result follows. □

2.5 Doob's h Transform

Doob's h-transform is a method used for deriving the SDE which is obtained by conditioning another SDE at its end point. It is also used to remove drift from an SDE, analysing hitting times when a process reaches a certain subset of the state space and excursions of SDEs. The method entails multiplying the transition density of original SDE with a suitable term (h-function) such that we can construct an SDE corresponding to the transformed density [24]. This section is based on work done by [24], this is another way of coming up with a BSDE. All definitions, theorems and proofs are by [24] unless otherwise stated.

Theorem 2.5.1 (Fokker-Planck Kolmogorov Equation)**.** *The probability density $p(X_t)$ of the solution of SDE Equation (18) solves the PDE*

$$\frac{\partial p(X_t)}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}[b(t,X_t)p(X_t)] + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}\{[\sigma(t,X_t)Q\sigma^*(t,X_t)]_{i,j}p(X_t)\} \tag{32}$$

The PDE is called the Fokker-Planck-Kolmogorov equation (FPK) usually referred to as forward kolmogorov equation in stochastics. The PDE is an initial value problem with given initial condition $p(X_{t_0})$ at $t = t_0$ [1].

Theorem 2.5.2 (Transition density)**.** *The transition density $p(X_t|X_s)$ of SDE (18) where $t \geq s$ is the solution to the Fokker-Planck-Kolmogorov equation with initial condition $p(X_t|X_s) = \delta(X_t - X_s)$ at $t = s$. More explicitly if we denote the transition density from Y_s to X_t as $p(x_t|y_s)$ then it solves*

$$\frac{\partial p(x_t|y_s)}{\partial t} = \mathcal{A}^* p(x_t|y_s),$$
$$p(x_s|y_s) = \delta(x-y)$$

where

$$\mathcal{A}^*(\cdot) = -\sum_i \frac{\partial}{\partial x_i}[b(t,X_t)(\cdot)] + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}\{[\sigma(t,X_t)Q\sigma^*(t,X_t)]_{i,j}(\cdot)\}$$

and Q is the diffusion matrix for the Brownian motion W. [1]

Let $p(Y_t|X_s)$ denote the transition density of the SDE and let $h(x,t)$ be defined by the state space regularity property

$$h(t,x) = \int p(y_{t+s}|x_t)h(t+s,y)dy. \tag{33}$$

Now lets define anther Markov process with transition kernel $p^h(y_{t'}|x_t) = p(y_{t'}|x_t)$ where $t' \geq t$ as

$$p^h(y_{t+s}|x_t) = p(y_{t+s}|x_t)\frac{h(t+s,y)}{h(t,x)} \tag{34}$$

we now check if this is a genuine probability density by integrating over y and use Equation (33) we have

$$\begin{aligned}
\int p^h(y_{t+s}|x_t)dy &= \int p(y_{t+s}|x_t)\frac{h(t+s,y)}{h(t,x)}dy \\
&= \frac{h(t,x)}{h(t,x)} \\
&= 1.
\end{aligned}$$

thus it is a genuine probability density. $h(t,x)$ obeys $\mathcal{A}_t h = 0$ from Equation (33), where $\mathcal{A}_t$ is as defined in Equation (6). Using Definition 13 the generator of p^h can be computed as

$$\begin{aligned}
\mathcal{A}^h\phi &= \lim_{s\downarrow 0}\frac{\mathbb{E}^h[\phi(x_{t+s})] - \phi(x_t)}{s} \\
&= \lim_{s\downarrow 0}\frac{\mathbb{E}[\phi(x_{t+s})h(t+s,y)] - \phi(x_t)h(t,x)}{sh(t,x)} \\
&= \frac{1}{h(t,x)}\mathcal{A}_t\{h(t,x)\phi(x)\} \\
&= \frac{1}{h(t,x)}\Bigg\{\frac{\partial h(t,x)}{\partial t}\phi + \sum_i\left[\frac{\partial h(t,x)}{\partial x_i}\phi(x) + h(t,x)\frac{\partial\phi(x)}{\partial x_i}\right]b_i(t,x) \\
&+ \frac{1}{2}\sum_{i,j}\frac{\partial^2[h(t,x)\phi(x)]}{\partial x_i\partial x_j}[\sigma(t,x)Q\sigma^*(t,x)]_{ij}\Bigg\} \\
&= \frac{1}{h(t,x)}\Bigg\{\left[\frac{\partial h(t,x)}{\partial t} + \sum_i\frac{\partial h(t,x)}{\partial x_i}b_i(t,x) + \frac{1}{2}\sum_{i,j}\frac{\partial^2 h(t,x)}{\partial x_i\partial x_j}[\sigma(t,x)Q\sigma^*(t,x)]_{ij}\right] \\
&*\phi(x) + \sum_i h(t,x)\frac{\partial\phi(x)}{\partial x_i}b_i(t,x) + \frac{1}{2}\sum_{i,j}\left[\frac{\partial h(t,x)}{\partial x_j}\frac{\partial\phi(x)}{\partial x_i} + \frac{\partial h(t,x)}{\partial x_i}\frac{\partial\phi(x)}{\partial x_j}\right. \\
&\left.+h(t,x)\frac{\partial^2\phi(x)}{\partial x_i\partial x_j}\right][\sigma(t,x)Q\sigma^*(t,x)]_{ij}\Bigg\}
\end{aligned}$$

Since $\mathcal{A}_t h = 0$ thus,

$$\left[\frac{\partial h(t,x)}{\partial t} + \sum_i \frac{\partial h(t,x)}{\partial x_i} b_i(t,x) + \frac{1}{2}\sum_{i,j} \frac{\partial^2 h(t,x)}{\partial x_i \partial x_j}[\sigma(t,x)Q\sigma^*(t,x)]_{ij}\right]\phi(x) = 0$$

therefore

$$\begin{aligned}\mathcal{A}^h\phi = \sum_i \left[b_i(t,x) + \frac{\bigtriangledown h(t,x)}{h(t,x)}[\sigma(t,x)Q\sigma^*(t,x)]_{ij}\right]\frac{\partial\phi(x)}{\partial x_i}\\ + \frac{1}{2}\frac{\partial^2\phi(x)}{\partial x_i\partial x_j}[\sigma(t,x)Q\sigma^*(t,x)]_{ij}\end{aligned}$$

which gives the new SDE as

$$\begin{aligned}dX &= \left[b(t,X) + \sigma(t,x)Q\sigma^*(t,x)\frac{\bigtriangledown h(t,x)}{h(t,x)}\right]dt + \sigma(t,X)dW\\ &= \left[b(t,X) + \sigma(t,x)Q\sigma^*(t,x)\bigtriangledown \log h(t,x)\right]dt + \sigma(t,X)dW \qquad (35)\end{aligned}$$

which we can use to condition at its end point. The following theorem by [24] shows how we can condition an SDE at its end point.

Theorem 2.5.3. *Assume that we have an SDE of the form* (18) *and we wish to condition its solution to hit X_T at time $t = T$. The h-transform gives the following SDE as the end point conditioned process*

$$dX = \left[b(t,X) + \sigma(t,X)Q\sigma^*(t,X)\bigtriangledown \log p(X_T|X_t)\right]dt + \sigma(t,X)dW \qquad (36)$$

The proof was done by [24] and is as follows.

Proof. We use Bayes' rule to obtain a conditioned SDE at its end point as follows:

$$\begin{aligned}p(X_{t+s}|X_t,X_T) &= \frac{p(X_T|X_{t+s},X_t)p(X_{t+s}|X_t)}{p(X_T|X_t)}\\ &= \frac{p(X_T|X_{t+s})p(X_{t+s}|X_t)}{p(X_T|X_t)}.\end{aligned}$$

Now let $h(t,X) = p(X_T|X_t)$, note this function satisfies the regularity property Equation (33) thus

$$p(X_T|X_t) = \int p(X_{t+s}|X_t)p(X_T|X_{t+s})dX_{t+s} \qquad (37)$$

substituting Equation (37) into Equation (35) we have the result Equation (36). □

We have shown the existence and uniqueness and existence of solutions for the BSDEs considered. We have shown the important theorem is Theorem 2.3.2 for this result. We also studied the relationship between a BSDE driven by Lévy process and a PDIE and a BSDE driven by Brownian motion and PDE. We also looked at the h-transform and how we could obtain our BSDE from an appropriate SDE.

3 APPLICATION

In this chapter, we give a simple application to finance which shows that Markovian BSDEs are useful tools in pricing theory since they give a generalisation of the Black-Scholes formula, in the sense the price of a contingent claim which only depends on the prices of the basic securities has the same property. Also, the hedging portfolio depends only on these prices. We consider the pricing of a European call option for both models presented in Section 2.4. Model 1 was done by [9] and Model 2 was done by [14], this will highlight the application of the theory we have studied in Chapter 2.

3.1 Model 1

Consider a financial market with coefficients which only depend on times s and on the vector of stock price processes P_s. Fix $(t,x) \in [0,T] \times \mathbb{R}^{n+1}$. Here, the prices of the basic securities satisfy the following equations on $[0,T]$;

$$dP_s^0 = r(s,P_s)P_s^0\,ds, \text{ and} \tag{38}$$

$$dP_s^i = P_s^i\left[\mu^i(s,P_s)dt + \sum_{j=1}^{n}\sigma_j^i(s,P_s)dW_s^j\right]. \tag{39}$$

Let $(P_s^{t,x},\ t \le s \le T)$ be the vector of stock price processes; $P_s^{t,x} = (P_s^0,\ P_s^1, \cdots, P_s^n)$ with initial condition given by $P_t^{t,x} = x$. In this context, a general setting of the wealth equation is

$$-\,dX_s = b(s,P_s,X_s,\sigma(s,P_s)^*\pi_s)ds - \pi_s^*\sigma(s,P_s)dW_s. \tag{40}$$

Here b is an $\mathbb{R}-$valued continuous function defined on $[0,T]\times\mathbb{R}^{n+1}\times\mathbb{R}\times\mathbb{R}^n$ that is Lipschitz with respect to (x,π) uniformly in t.

Consider a contingent claim $\xi = \phi(P_T^{t,x})$. The function $\phi : \mathbb{R}_+^{n+1} \longrightarrow \mathbb{R}^+$ is Lipschitz. Then there exists a unique square integrable hedging strategy (by Theorem 2.3.2) $(X^{t,x},\pi^{t,x}) \in \mathbb{H}_T^{2,d}\times\mathbb{H}_T^{2,n\times d}$ against ξ such that

$$\begin{aligned} -dX_s^{t,x} &= b(s,P_s^{t,x},X_s^{t,x},\sigma(s,P_s^{t,x})^*\pi_s^{t,x}) - (\pi_s^{t,x})^*\sigma(s,P_s^{t,x})dW_s, \\ X_T^{t,x} &= \phi(P_T^{t,x}), \end{aligned} \tag{41}$$

and $X_s^{t,x}$ is the price of the contingent claim $\phi(P_T^{t,x})$ at time s. Then from the results of

Section 2.4, the value at time s of the contingent claim ξ is

$$X_s^{t,x} = u(s, P_s^{t,x}),$$

where $u(t,x) = X_t^{t,x}$ is the unique viscosity solution of the non-linear parabolic PDE

$$\begin{aligned} \frac{\partial u}{\partial t} + \sum_{i,j=1}^{n} a_{i,j}(t,x) x_i x_j \frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) \\ + \sum_{i=1}^{n} \mu_i(t,x) x_i \frac{\partial u}{\partial x_i}(t,x) + r(t,x) x_0 \frac{\partial u}{\partial x_0}(t,x) \\ = -b\left(t, x, u(t,x), \sigma^*(t,x)\left[x\frac{\partial u}{\partial x}\right]\right), \qquad (42) \\ u(T,x) = \phi(x), \end{aligned}$$

where

$$a_{i,j}(t,x) = \frac{1}{2}[\sigma\sigma^*]_{ij}(t,x),$$

and

$$\left[x\frac{\partial u}{\partial x}\right] = x_i \frac{\partial u}{\partial x_i}(t,x).$$

Also if the function b is $\mathcal{C}^3$ with bounded derivatives, then u belongs to $\mathcal{C}^{1,2}([0,T] \times \mathbb{R}^{n+1}, \mathbb{R})$, and it is a regular solution of the PDE. Notice that the portfolio process of the hedging strategy is then

$$\pi_s^i = P_s^i \frac{\partial u}{\partial x_i}(s, P_s), \quad t \leq s \leq T, \quad 1 \leq i \leq n.$$

3.2 Model 2

Consider a market with a one riskless asset (Bond) and one risky asset (stock). The price process for the assets is $B_t = e^{rt}$ and $S_t = S_0 \exp(X_t)$ respectively, where X_t is a Lévy process, r is the risk free interest rate and S_0 stock price at time 0. The probability measure of X_1 is denoted by $p(dx)$. The pay-off of a derivative at time T(expiry/ maturity) is $G(S_T) = F(X_T)$. Since we considering a European call option the pay-off becomes $G(S_T) = (S_T - K)^+$, where K is the strike price. Substituting for S_T we get

$$F(X_T) = (S_0 \exp(X_T) - K)^+.$$

The arbitrage free price V_t of the derivative at time $t \in [0, T]$ is

$$V_t = \mathbb{E}_Q \left[e^{-r(T-t)} G(S_T) | \mathcal{F}_t \right] \tag{43}$$

where $\{\mathcal{F}_t\}_0^T$ is the natural filtration of $\{X_t\}_0^T$ and $Q(dx)$ is the equivalent martingale measure.

Definition 17. The equivalent martingale measure is a probability measure which is equivalent to the given probability measure and under it the discounted price process $\{e^{-rt} S_t\}$ of the security is a martingale [14].

Our model has more than one equivalent martingale measure (EMM), we say the model is incomplete in this case. To obtain atleast one emm to use for the valuation of the derivative security under our model we use the Esscher Transform. Now let us define a Meixner process as follows.

Definition 18 (Meixner process)**.** A Meixner process $M = \{M_t; t \geq 0\}$ is a bounded variation Lévy process based on the uniformly divisible distribution with density function

$$f(x; m, a) = \frac{(2\cos(\frac{a}{2}))^{2m}}{2\pi\Gamma(2m)} \exp(ax) |\Gamma(m + ix)|^2, \; x \in (-\infty, +\infty)$$

where a is a real constant and $m > 0$ [13].

Now consider the density function

$$f_{\text{Meixner}}(x; \alpha, \beta, \delta, \mu) = \frac{(2\cos(\frac{\beta}{2}))^{2\delta}}{\pi\alpha\Gamma(2\delta)} \exp(\beta(x-\mu)/\alpha) |\Gamma(\delta + i(x-\mu)/\alpha)|^2,$$

the cumulant generating function

$$K_{\text{Meixner}}(\theta;\alpha,\beta,\delta,\mu) = \mu\theta + 2\delta\left\{\log\left[\cos\left(\frac{\beta}{2}\right)\right] - \log\left[\cos\left(\frac{\alpha\theta+\beta}{2}\right)\right]\right\},$$

the drift

$$a_{\text{Meixner}}(\alpha,\beta,\delta,\mu) = \mu + \alpha\delta\tan\left(\frac{\beta}{2}\right) - 2\delta\int_1^\infty \frac{\sinh\left(\beta x/\alpha\right)}{\sinh(\pi x/\alpha)}dx,$$

and the Lévy measure

$$\nu_{\text{Meixner}}(dx;\alpha,\beta,\delta,\mu) = \frac{\delta\exp(\beta x/\alpha)}{x\sinh(\pi x/\alpha)}dx,$$

where $\alpha > 0,\ -\pi < \beta < \pi,\ \mu \in \mathbb{R}$ and $\delta > 0$. Consider K to be the cumulant generating function of X under the measure $p(dx)$. Let θ be the solution of $K(\theta+1) - K(\theta) = r$. The risk neutral measure $Q(dx)$ is the probability measure with the Radon-Nykodym derivative with respect to $p(dx)$ given as

$$\frac{Q(dx)}{p(dx)} = \exp(\theta x - K(\theta)).$$

In the Meixner case if we shift β to $\beta + \alpha\theta$ we obtain the density under the risk neutral measure $Q(dx)$. The the process M_t under the risk neutral measure is a Meixner process. The price at time t of the derivative security, $V_t = V(t, M_t)$ satisfies regularity conditions like having uniformly bounded derivatives first order in time and second order in space for $V(t,x)$, then $V(t,x)$ is the solution for the PDIE,

$$\begin{aligned} rV(t,x) &= a\frac{\partial}{\partial x}V(t,x) + \frac{\partial}{\partial t}V(t,x) \\ &+ \int_{*\infty}^{\infty}\left(V(t,x+y) - V(t,x) - y\frac{\partial}{\partial x}V(t,x)\right)\nu^Q(dy) \\ V(T,x) &= F(x) \end{aligned}$$

where $\nu^Q(dy)$ is the Lévy measure of the risk neutral distribution $Q(dx)$. We have

$$\nu^Q(dx) = d\frac{\exp(a\theta+\beta)x/a}{x\sinh(\pi x/a)},$$

where

$$a = a_{\text{Meixner}}(\alpha, \alpha\theta+\beta, \delta, \mu).$$

The PDIE is obtained by applying the Feynman-Kac formula for Lévy processes and it is the analogue of the Black-Scholes PDE which we got for Model 1.

4 DISCUSSION

We have studied the existence and uniqueness of a solution of a BSDE with a Lipschitz driver and a square integrable terminal condition measurable to the filtration at expiry time. We considered BSDEs of two types, the first is a BSDE driven by a Brownian motion and the second is driven by the a Lévy process, the solution of the BSDEs is a pair of adaptable processes (Y, Z). The process Z is obtained independent of the process Y, for BSDE of Model 1 we have Z obtained through the martingale representation theorem and for BSDE of Model 2 it is obtained through the predictable representation theorem.

In Section 2.3, to prove uniqueness we made use of the Banach fixed point theorem and the Picard iterative sequence for the existence of solution for the Brownian motion driven BSDE. For the Lévy process driven BSDE we had a similar result as to prove existence and uniqueness we showed that we have a contraction mapping onto itself. In both cases the data is said to be standard, that is the terminal condition is square integrable and $\mathcal{F}_T$-measurable and the drift is uniformly Lipschitz and $f(\cdot, 0, 0) \in \mathbb{H}^T_{2,d}$.

In Section 2.4 we consider the relationship between a BSDE and a partial differential equation, with the aid of the Feynman-Kac formula. We did prove that the solution of BSDE in Model 1 is also a solution for a PDE, while the BSDE in Model 2 is also a solution to a PDIE under some smoothness conditions. For Model 1 we prove this using viscosity solution of the PDE to avoid restrictions on the coefficients.

In Section 2.5 the Doob's h-transform is studied. The transform is used to condition an SDE on its end points from another SDE. We considered the definitions of the FPK equation for both Model 1 and Model 2 , but however gave an example with SDE of type Model 1.

Finally in Section 3 we consider an application of Section 2.4 to option pricing of European call options for both the BSDE considered in this thesis. For BSDE of Model 1 we come up with the Black-Scholes PDE and for BSDE of Model 2 we have an analogue of the Black-Scholes PDE as a PDIE. The PDE and PDIE can be solved to come up with the option price in the respective market settings.

4.1 Future Work

In this thesis we considered backward stochastic differential equations driven by a Brownian motion and Lévy processes, with the driver being Lipschitz, the terminal condition square integrable and measurable with respect to the filtration at time T. In future work we could consider BSDEs driven by a non-Lipschitz drift, with a deterministic terminal condition. We could also consider the Lévy process as in Model 2 having a Brownian motion part. Application to financial markets was the main focus, BSDE have wide variety of application even in physics, which is another area we can consider and also in sequential Monte Carlo methods.

5 CONCLUSION

We proved the existence and uniqueness of solution for the BSDE driven by Brownian motion and also driven by a Lévy process. The a prior estimate of the solution is a pair of adapted processes which solve the BSDE. We used the contraction mapping to show uniqueness of the solution as a mapping onto itself. The Picard iterative sequence was employed to show existence of solution.

The relationship between a BSDE driven by a Brownian motion and a PDE and also a BSDE driven by a Lévy process and a PDIE was investigated. The Feynman-Kac formula was used to show and prove that the solution of the BSDE is a solution of PDE or PDIE under some smoothness conditions. Viscosity solutions approach was used to prove this relationship.

In Section 2.5 we considered an SDE which we would want to condition at its end point that is coming up with a BSDE from an initial SDE. We use the Doob's h transform to accomplish it. We apply it to FSDE of Model 1 and come up with an SDE conditioned at it end point.

Finally we apply the Feynman-Kac formulas to option procing of European calls for a Brownian motion market and a Lévy market. The PDE and PDIE is derived for the markets respectively which can be solved to come up with the option price.

REFERENCES

[1] P. E. Kloeden and E. Platen. *Numerical Solutions of Stochastic Differential Equations*. Springer-Verlag Berlin Heidelberg GmbH, 1999.

[2] Etienne Pardoux and Aurel Râscanu. *Stochastic differential equations Backward SDEs Partial differential equations*, volume 69. Springer, 2014.

[3] J. M. Bismut. Conjugate Convex Functions in Optimal Stochastic Control. *Journal of Mathematical Analysis and Applications*, 44.2, 1973.

[4] E. Pardoux and S. Peng. Adapted Solution of a Backward Stochastic Differential Equation. *Systems and Control Letters*, 14(1), 1990.

[5] S. Peng. Probabilistic Interpretation for Systems of Quasilinear Parabolic Partial Differential Equations. *Stochastics Stochastics Rep*, 37(1-2), 1991.

[6] E. Pardoux and S. Peng. Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations. *Stochastic Partial Differential Equations and their Applications*, 200-217, 1992.

[7] F. Antonelli. Backward-Forward Stochastic Differential Equations. *The Annals of Applied Probability*, 1993.

[8] J. Ma, P. Protter, and J. Yong. Solving Forward-Backward Stochastic Differential Equations Explicitly—a Four Step Scheme. *Probability Theory and Related Fields*, 98(3), 1994.

[9] N. El-Karoui, S. Peng, and M.C. Quenez. Backward Stochastic Differential Equations in Finance. *Mathematical Finance*, 7:1–29, 42–53, 1997.

[10] R. Buckdahn and E. Pardoux. BSDE's with Jumps and Associated Integro-Partial Differential Equations. *preprint*, 79, 1994.

[11] Situ Rong. On solutions of backward stochastic differential equations with jumps and applications. *Stochastic Processes and their Applications*, 66(2):209–236, 1997.

[12] Y Ouknine. Reflected backward stochastic differential equations with jumps. *Stochastics: An International Journal of Probability and Stochastic Processes*, 65(1-2):111–125, 1998.

[13] David Nualart and Wim Schoutens. Chaotic and predictable representations for lévy processes. *Stochastic processes and their applications*, 90(1):109–122, 2000.

[14] David Nualart, Wim Schoutens, et al. Backward stochastic differential equations and feynman-kac formula for lévy processes, with applications in finance. *Bernoulli*, 7(5):761–776, 2001.

[15] Brent Øksendal. *Stochastic Differential Equations, An Introduction with Applications*. Springer-Verlag Berlin Heidelberg, 2000.

[16] David Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.

[17] Ioannis Karatzas and Steven E Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag Newyork, 1991.

[18] J. B. Diaz and B. Margolis. A Fixed Point Theorem of the Alternative, for Contractions on a Generalized Complete Metric Space. *Bulletin of the American Mathematical Society*, 74(2):305–309, 1968.

[19] René L Schilling. *Measures, Integrals and Martingales*. Cambridge University press, Cambridge, 2005.

[20] Steven E Shreve. *Stochastic Calculus for Finance II, Continuous Time Models*. Springer Science+Business Media inc, New York, 2004.

[21] Philip Protter. *Stochastic integration and differential equations: a new approach*. Springer, 1990.

[22] F. C. Klebaner. *Introduction To Stochastic Calculus with Applications*. Imperial College Press, London, 2012.

[23] Arthur Mattuck. *Introduction to Analysis*. Prentice Hall, New Jersey, Inc, 1999.

[24] Särkkä Simo and Arno Solin. *Applied Stochastic Differential Equations*. Cambridge University Press, In press.

[25] S. Peng. Stochastic Hamilton–Jacobi–Bellman Equations. *SIAM Journal on Control and Optimization*, 30(2), 1992.