ABSTRACT

Title of dissertation: TWO PRICE ECONOMY IN CONTINUOUS TIME AND ITS APPLICATIONS IN FINANCE

Tong Meng, Doctor of Philosophy, 2013

Dissertation directed by: Professor Dilip B. Madan
Department of Finance

Two price economy provides a new approach to describe incomplete markets. Unlike the classical economy theory, in which the law of one price prevails, a two price economy determines prices by the directions of the trades. Static one period and discrete time two price economies are described and applied in a number of papers.

Following the static and discrete time models, continuous time two price economies are studied in this thesis. Dynamically consistent nonlinear pricing functionals are generated from backward stochastic differential equations (BSDEs) on continuous time Markov chains (CTMCs) and $\mathcal{G}$-expectations. This thesis also includes a convergence theorem of BSDEs on CTMCs, and the existence and uniqueness of solution to the distorted partial integro-differential equation coming from the $\mathcal{G}$-expectation approach.

The continuous time models for two price economies are illustrated through three examples. BSDEs on CTMCs are used to generate bid and ask prices for

option spreads. Then $\mathcal{G}$-expectation theory is applied to produce credit capital commitments for derivatives with bilateral counterparty risk and give bid and ask interest rate swap rates and swaption prices.

TWO PRICE ECONOMY IN CONTINUOUS TIME AND ITS APPLICATIONS IN FINANCE

by

Tong Meng

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor Dilip B. Madan, Chair/Advisor
Professor Leonid Koralov
Professor Mark Loewenstein
Professor Paul J. Smith
Professor Victor M. Yakovenko, Dean's Representative

Dedication

To my family

Acknowledgments

First and foremost I would like to express my utmost appreciation to my advisor, Professor Dilip Madan for his guidance, support, and encouragement on my research. To me, he is not only an extraordinary scholar but also a caring mentor and a good friend. His enthusiasm, wisdom, and depth of knowledge have been the driving forces in my graduate study. He is also generous with his help when it comes to assisting with job opportunities and important decisions in life. I feel extremely lucky having the chance to work with and learn from him.

I would like to thank Professor Leonid Koralov, Professor Mark Loewenstein, Professor Paul J. Smith, and Professor Victor M. Yakovenko for agreeing to serve on my dissertation committee and spending much time reviewing my manuscript. I also benefited from the excellent lecturing of Professors Leonid Koralov, and Professor Mark Loewenstein. I deeply thank Professor Paul J. Smith for providing me the opportunity to study at University of Maryland in the first place, and giving me valuable suggestions throughout these years.

I thank all members in the mathematical finance research interaction team for helpful discussions.

I thank all my good friends for making my graduate life colorful.

I owe my deepest thanks to my parents and Yujie, who have always stood by me and guided me through the hard times. Words cannot express the gratitude I owe them.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

VG	Variance Gamma
FFT	Fast Fourier Transform
BSDE	Backward Stochastic Differential Equation
CTMC	Continuous Time Markov Chain
PIDE	Partial Integro-Differential Equation
CCC	Credit Capital Commitment
CVA	Credit Value Adjustment
with respect to	w.r.t
almost surely	a.s.

Chapter 1

Introduction

1.1 Overview

In classical economic theory, merchandise is traded in both directions at the same price–a phenomenon best known as the law of one price. This price is determined through a market clearing condition (Arrow [4]; Ingersoll [28]) under the framework of economic equilibrium analysis (Arrow and Debreu [5]). In financial economies, the law of one price is explained to be the consequence of the no arbitrage assumption, and derivatives are priced under the risk neutral measure according to the fundamental theorem of asset pricing (Dybvig and Ross [16]).

However, considerable differences between the bid and ask prices (spread) are widely noticed in financial markets, especially during the financial crisis in 2008 (Flannery et al. [20]). Numerous theoretical and statistical studies have been conducted on the bid-ask spread. Some theoretical models emphasize the order processing and inventory holding costs incurred by liquidity providers (Amihud and Mendelson [2]; Demsetz [15]; Stoll [51]), while others concentrate on the adverse selection costs caused by informed traders (Copeland and Galai [13]; Easley and Kiefer [17]; Glosten and Milgrom [23]; Kyle [31]). Statistical methods are applied to measure the components of the bid-ask spread. Roll [48], Choi, Salandro and Shastri [9], Stoll [52], George, Kaul and Nimalendran [21] make inferences about the bid-ask spread from the serial covariance of price changes. A trade indicator regression model is developed and extended in Glosten and Harris [22], Madhavan, Richardson, and Roomans [40], Huang and Stoll [25] [26]. Most of these existing studies focus on finding the source of costs for market makers by modeling the price determination process in liquid markets.

In a recent paper [36] by Madan and Schoutens, a two price economy with an equilibrium model is developed. It is explained in [36] that the bid-ask spread originates from the difference between the event space where contracts are written and the event space where the actual economy lives. This argument is similar to the opaqueness investigated in Flannery *et al.* [20]. Since the actual economy lives in a much larger space, unexpected events may cause endowment loss, and precommitted demands are not cleared. A financial system is thus introduced to approve trades and cover unexpected loss of the markets. In the two price economy, all market participants are modeled to do trades with the same financial system at different prices depending on the directions of the trades. The system would determine the spreads so as to make its loss exposures acceptable. A one period static model for the two price economy is introduced in [36], in which bid and ask prices are defined as the infimum and supremum of test measures' evaluations. This model is generalized to include discrete time case in a subsequent paper [35] by Madan, Pistorius and Schoutens, with bid and ask prices modeled as dynamically consistent nonlinear expectations. Many applications have been conducted under this framework. In [34] and [32] by Cherny and Madan, the static two price model is used to estimate stress levels of distortions from market prices of vanilla options and to define capital requirements and monitor leverage. In [35], Madan, Pistorius and Schoutens employ the discrete time model to price a variety of structured products. These methods are also applied to the pricing of insurance loss liabilities in [38] by Madan, Wang, and Heckman.

Despite the usefulness of static and discrete time two price models in a variety of contexts, they lack the ability of evaluating claims that may be delivered at arbitrary times. As a result, we wants to find ways to conduct dynamically consistent nonlinear pricing in continuous time. Inspired by the discrete time two price model in [35], where the bid and ask prices are generated as nonlinear expectations induced by backward stochastic difference equations, we have developed two approaches to build a two price economy in continuous time. The first approach utilizes properties of backward stochastic differential equation (BSDE). BSDE has been extensively studied during the past two decades since the original paper [44] by Pardoux and Peng appeared in 1990, due to its connection with stochastic optimization problem. The connection between solution to BSDE and nonlinear expectation is first established by Peng in [45] in 1997. The result is obtained in the context of continuous time diffusions, and therefore, is unable to deal with any case when the underlying martingale could jump with positive probability. However, such case may arise in various applications. In a recent paper [11] by Cohen and Elliott, BSDE with randomness generated by continuous time Markov chains (CTMCs) is studied and the connection between BSDE solution and nonlinear expectation is also given. This then allows one to construct dynamically consistent bid and ask prices as solutions to BSDEs on CTMCs, which is also the key idea behind our first approach. The second approach makes use of the $\mathcal{G}$-expectation method proposed by Peng in [46]. In Peng's paper, $\mathcal{G}$-expectation is described as the solution to a nonlinear heat equation with a given infinitesimal generator $\mathcal{G}$. Following this idea, we have extended the $\mathcal{G}$-expectation concept to partial integral differential equations (PIDEs), and the nonlinearity is obtained by distorting the integral term in the PIDE. We can then generate the bid and ask prices as viscosity solutions to the resulting distorted PIDEs.

The two price economy in continuous time adds more flexibility to the two price theory, which now allows construction of dynamically consistent bid and ask prices for a much larger set of derivatives. It also provides approaches to study financial concepts that are related to the bid and ask prices, for example, capital requirements and monitor leverage introduced in [32], and market implied stress levels discussed in [34].

The outline of this thesis is as follows. The rest of Chapter 1 briefly reviews the basics of Lévy processes used in this study, and the Carr-Madan Fast Fourier Transform (FFT) method [7] associated with Lévy based models. Chapter 2 summarizes the static and discrete time models for two price economies. Chapter 3 introduces BSDE on CTMC and its relation with nonlinear expectation. Asymptotic result of solutions to BSDEs on CTMCs under certain conditions is also developed in Chapter 3. Chapter 4 presents the $\mathcal{G}$-expectation methods in the context of Lévy processes. Applications of the continuous time two price economy are conducted in Chapter 5. The model is applied to build bid and ask prices for option spreads, generate bid and ask interest rate swap rates, compute bid and ask swaption prices, and evaluate credit capital commitments for derivatives with bilateral counterparty risk.

1.2 Lévy Processes in Finance

1.2.1 Definition and Lévy-Khintchine Representation

Lévy processes are named after the great French probabilist Paul Lévy, who first studied them back in the 1930s. Although most of their basic structures and properties were understood in the 1930s and 1940s, they have generated great interest during the past two decades, particularly in the field of mathematical finance. Lévy processes, from a probability point of view, are stochastic processes with stationary and independent increments. They could be seen as generalizations of random walks to continuous time. Many well-known processes in the literature fall into this category, such as Brownian motions, Poisson processes, stable processes, and subordinators. A formal definition of Lévy processes can be written as follows. This definition can be found in Section 1.3 in [3].

Definition 1.2.1. *Let $X = (X_t; t \geq 0)$ be a stochastic process taking values in $\mathbb{R}^d$ defined on $(\Omega, \mathcal{F}, P)$. Then X is a Lévy process if:*

1. *$X_0 = 0$ almost surely (a.s.).*

2. *X has independent increments: for any $0 \leq t_0 \leq t_1 \leq ... \leq t_n < \infty$, the random variables $(X_{t_j} - X_{t_{j-1}}, 1 \leq j \leq n)$ are independent.*

3. *X has stationary increments: for any $t \geq 0$ and $h \geq 0$, $X_{t+h} - X_t \overset{d}{=} X_h - X_0$.*

4. X is stochastically continuous: for any $t \geq 0$,$h \geq 0$, and $\epsilon > 0$,

$$\lim_{h \to 0} P(|X_{t+h} - X_t| > \epsilon) = 0.$$

Every Lévy process has a càdlàg (right continuous with left limits) modification that is also a Lévy process. A stochastic process $(Y_t, t \geq 0)$ is said to be a modification of $(X_t, t \geq 0)$ if for each $t \geq 0$, $P(X_t \neq Y_t) = 0$. From now on, for any Lévy process, we will be referring to its cádlág modification. Levy processes are closely connected with infinitely divisible distributions. The following definition is given in Section 1.2.2 in [3].

Definition 1.2.2. *A random variable X taking values in $\mathbb{R}^d$ is called an infinitely divisible if for all $n \in N$, there exists a sequence of independent and identically distributed (i.i.d.) random variables Y_1, Y_2,...,Y_n, such that*

$$X \stackrel{d}{=} Y_1 + \cdots + Y_n.$$

The following proposition is given by Corollary 1.4.6 in [3].

Proposition 1.2.3. *Let $(X_t, t \geq 0)$ be a Lévy process. Then for any $t \geq 0$, the distribution of X_t is infinitely divisible. Conversely, let μ be an infinitely divisible probability measure. Then there exists a Lévy process $(X_t, t \geq 0)$, such that μ is the law of X_1.*

This proposition implies that the characteristic function of a Lévy process $(X_t, t \geq 0)$ at time $t \geq 0$ can be expressed as

$$\phi_{X_t}(u) = E(e^{i(u,X_t)}) = (\phi_{X_1}(u))^t = e^{t\psi_{X_1}(u)},$$

where ψ_{X_1} denotes the characteristic exponent of X_1, that is $\phi_{X_1}(u) = e^{\psi_{X_1}(u)}$. To characterize a Lévy process, it suffices to specify its distribution at unit time. The famous Lévy-Khintchine formula gives a general form of the characteristic component of the distribution of a Lévy process at unit time. The following theorem is given by Theorem 1.2.14 in [3].

Theorem 1.2.4. ***(Lévy-Khintchine Representation)*** *Let $X = (X_t, t \geq 0)$ be a Lévy process that takes values in R^d. Then the characteristic exponent of X_1 can be expressed as*

$$\psi_{X_1}(u) = i < \gamma, u > -\frac{1}{2} < u, Au > + \int_{R^d - \{0\}} (e^{i(u,y)} - 1 - i(u,y)\mathbf{1}_{|y|<1})\nu(dy),$$

where $\gamma \in R^d$, A is a symmetric non-negative definite $d \times d$ matrix, and ν is a measure on $R^d - \{0\}$, named the Lévy measure, with

$$\int_{R^d} (|x|^2 \wedge 1)\nu(dx) < \infty.$$

One calls (γ, A, ν) the Lévy triplet of the process X.

We can see from the Lévy-Khintchine representation that in general, a Lévy process can be decomposed into three independent parts: a linear deterministic drift, a Brownian motion, and a pure jump part.

1.2.2 The Variance Gamma Process

The class of Variance Gamma (VG) processes was first introduced by Madan and Seneta in [37], as an alternative to the Brownian motion in modeling stock market returns. In [37], the symmetric case of VG process is developed, which is later generalized to provide skewness to the model in [33] by Madan, Carr and Chang. Since then, the VG process has become one of the most popular Lévy processes. Before presenting the VG process, we need to introduce the VG distribution. Definitions in this subsection can be found in Section 5.3.7 in [49].

Definition 1.2.5. *Let X be a random variable. We say X is VG distributed with parameters $(\sigma, \nu > 0, \theta)$ if the characteristic function of X satisfies*

$$\phi_X(u) = (1 - iu\theta\nu + \frac{1}{2}u^2\sigma^2\nu)^{-\frac{1}{\nu}}.$$

A VG process $X^{(VG)} = (X_t^{(VG)}, t \geq 0)$ would then be a Lévy process for which the increment $X_{t+s}^{(VG)} - X_s^{(VG)}$ $(t \geq 0, s \geq 0)$ follows a VG $(\sigma\sqrt{t}, \nu/t, \theta t)$ law. An alternative way of defining VG process is as follows.

Definition 1.2.6. *Let $X^{(VG)} = (X_t^{(VG)}, t \geq 0)$ be a VG process with parameters $(\sigma, \nu > 0, \theta)$. Then $X^{(VG)}$ can be written as*

$$X_t^{(VG)} = \theta G_t + \sigma W_{G_t},$$

where $W = (W_t, t \geq 0)$ is a standard Brownian motion and G_t is a Gamma process independent of W, with parameter $(1/\nu, 1/\nu)$.

Definition 1.2.7. *A Lévy process $X = (X_t, t \geq 0)$ is a Gamma process with parameter (γ, λ) if it has Lévy triplet*

$$\gamma(1 - exp(-\lambda))/\lambda, 0, \gamma exp(\lambda x)x^{-1}\mathbf{1}_{\{x>0\}}dx).$$

As a result, a VG process can be seen as a time changed Brownian motion with drift, which is also how the process was invented in the first place. The Lévy measure of $X^{(VG)}$ can be computed to be

$$\nu_{VG}(dx) = \begin{cases} \frac{Ce^{Gx}}{|x|}, & x < 0, \\ \\ \frac{Ce^{-Mx}}{x}, & x > 0, \end{cases}$$

where

$$C = \frac{1}{\nu},$$
$$G = \left(\sqrt{\frac{1}{4}\theta^2\nu^2 + \frac{1}{2}\sigma^2\nu} - \frac{1}{2}\theta\nu \right)^{-1},$$
$$M = \left(\sqrt{\frac{1}{4}\theta^2\nu^2 + \frac{1}{2}\sigma^2\nu} + \frac{1}{2}\theta\nu \right)^{-1}.$$

The Lévy triplet of $X^{(VG)}$ is given by $(\gamma, 0, \nu_{VG}(dx))$, with

$$\gamma = \frac{CG(1-e^{-M}) - CM(1-e^{-G})}{MG}.$$

This leads to another way of explaining the VG process as the difference between two independent Gamma processes:

$$X^{(VG} = X^{(G)}(t; C, M) - X^{(G)}(t; C, G).$$

1.3 The Fast Fourier Transform Method in Option Pricing

Although Lévy based models have shown significant improvements in explaining financial data compared with the renowned Black-Scholes model, their distributions lack closed form expressions in most cases. Numerical methods and Fourier analysis have been applied in the literature to develop calibration methods for non-Gaussian models. Heston in [24] made use of Lévy's inversion formula to evaluate Vanilla options numerically. Carr and Madan showed in [7] a fast Fourier transform

(FFT) method to do option pricing by relating the Fourier tranform of the option price with the characteristic function of the underlying Lévy process. The Carr-Madan FFT method has since been used as a standard engine for calibration due to its fast speed and effectiveness. This method is briefly discussed below.

Consider a European call option with strike K and maturity T. Let k denote $ln(K)$, s_T denote $ln(S_T)$, $q_T(s)$ denote the probability density of s_T under the risk neutral measure, and $C_T(k)$ denote the expected value of the call option. Our goal is to compute $C_T(k)$ in an efficient way. Since $C_T(k)$ is not square integrable, Carr and Madan introduced the modified call price

$$c_T(k) := e^{\alpha k} C_T(k)$$

for some $\alpha > 0$. The α is called the dampening coefficient, and is chosen to make the modified call price square integrable. Now consider the Fourier transform of c_T,

$$\psi_T(v) = \int_{-\infty}^{\infty} e^{ivk} c_T(k) dk.$$

Then $\psi_T(v)$ can be derived analytically as

$$\psi_T(v) = \frac{e^{-rT} \phi_T(v - (\alpha+1)i)}{\alpha^2 + \alpha - v^2 + iv(2\alpha+1)},$$

where $\phi_T(u)$ is the characteristic function of s_T, defined by

$$\phi_T(u) = \int_{-\infty}^{\infty} e^{ius} q_T(s) ds.$$

Therefore, $C_T(k)$ can be recovered from inverse Fourier transform through

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \int_0^{\infty} e^{ivk} \psi_T(v) dv. \tag{1.1}$$

In [7], the integral term in (1.1) is approximated by trapezoidal rule on a well-defined grid for v, and FFT method is applied to obtain values of $C_T(k)$ on a pre-specified grid for k.The grids for v and k are chosen as

$$v_j = (j-1)\eta, \quad \text{for } j = 1, 2, ...N,$$
$$k_u = -\frac{1}{2}N\lambda + (u-1)\lambda, \quad \text{for } u = 1, 2, ...N,$$

where λ, η satisfy $\lambda\eta = \frac{2\pi}{N}$. As a result, we have the following approximation of (1.1),

$$C_T(k_u) \approx \frac{e^{-\alpha k_u}}{\pi} \sum_{j=1}^{N} e^{-\frac{2\pi i}{N}(j-1)(u-1)} e^{i\frac{1}{2}N\lambda v_j} \psi(v_j) \frac{\eta}{3}[3 + (-1)^j - \mathbf{1}_{j=1}]. \tag{1.2}$$

In general, N is chosen to be a power of 2, and is set to 1024 in [7]. The summation in (1.2) is then computed using the FFT method.

Chapter 2

The Two Price Economy

2.1 One Period Two Price Model

In classical economic analysis, the market is modeled as a passive agent, which attains its equilibrium by setting prices so as to ensure market clearing. The two price economy is built upon the belief that market clearing is not always reachable. In the two price economy, the market is defined to achieve its equilibrium by making excess supplies acceptable.

Consider an economy trading bounded cash flows defined on a probability space $(\Omega, \mathcal{F}, P)$. A cash flow X is said to be acceptable if

$$E^Q(X) \geq 0, \text{for all } Q \in \mathcal{N},$$

where $\mathcal{N}$ is a pre-determined convex set of probability measures equivalent to P, and also includes a risk neutral measure. The $\mathcal{N}$ is called the set of test measures. Let $\mathcal{A}$ denote the collection of acceptable cash flows. Then $\mathcal{A}$ would be a cone including all nonnegative $\mathcal{F}$ measurable random variables. The larger the size of $\mathcal{A}$, the bigger the trading opportunities in the market, so the larger the size of the economy.

The two price economy proceeds by offering each market participant the same cone of acceptable cash flows $\mathcal{A}$. It is shown in [34] that for any state contingent claim with $\mathcal{F}$ measurable payoff X, its bid price $b(X)$ and ask price $a(X)$ in the two price system are then given by

$$b(X) = \inf_{Q \in \mathcal{N}} E^Q(X),$$

$$a(X) = \sup_{Q \in \mathcal{N}} E^Q(X).$$

The bid and ask pricing functionals map from the space of bounded $\mathcal{F}$ measurable random variables to $\mathbb{R}$. By construction, bid is concave, ask is convex, and the two functionals satisfy

$$a(X) = -b(-X).$$

As a result, one only needs to develop methods for generating bid prices.

We note that the bid pricing functional is controlled by the set of test measures. In [34], this set of measures is constructed indirectly through acceptability indices. The concept of acceptability index is introduced in [8] by Cherny and Madan as a new measure for performance evaluation. Its definition is given below.

Definition 2.1.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. An acceptability index α is a mapping $\alpha : L^\infty(\Omega, \mathcal{F}, P) \to [0, \infty]$ satisfying*

1. *Quasi-concavity: if $\alpha(X) \geq \gamma$ and $\alpha(Y) \geq \gamma$, then $\alpha(\lambda X + (1-\lambda)Y) \geq \gamma$ for any $\lambda \in [0,1]$,*

2. *Monotonicity: if $X \leq Y$ a.s., then $\alpha(X) \leq \alpha(Y)$,*

3. *Scale invariance: $\alpha(\lambda X) = \alpha(X)$ for any $\lambda > 0$,*

4. *Fatou property: for any sequence of uniformly bounded random variables $(X_i)_{i=1}^\infty$ that converges to X in probability, with $\alpha(X_i) \geq \lambda$ for all i, then $\alpha(X) \geq \lambda$.*

We have the following property regarding acceptability indices. Proofs can be found in [8].

Theorem 2.1.2. *Let α be a mapping from $L^\infty(\Omega, \mathcal{F}, P)$ to $[0, \infty]$, and let $\mathcal{P}$ denote the set of probability measures absolutely continuous with respect to P. The following statements are equivalent.*

(i) α is an acceptability index;

(ii) there exists a family of sets $(\mathcal{D}_\gamma)_{\gamma \in R_+}$, with $\mathcal{D}_\gamma \subseteq \mathcal{P}$ and $\mathcal{D}_{\gamma_1} \subseteq \mathcal{D}_{\gamma_2}$ whenever $\gamma_1 \leq \gamma_2$, such that

$$\alpha(X) = \sup\{\gamma \in R_+ : \inf_{Q \in \mathcal{D}_\gamma} E^Q(X) \geq 0\},$$

where $\inf \emptyset = \infty$ *and* $\sup \emptyset = 0$.

Cherny and Madan [8] proposed a class of acceptability indices termed the Weighted Value at Risk (WVAR) acceptability index. Suppose X is a random variable belonging to $L^\infty(\Omega, \mathcal{F}, P)$ with distribution function $F_X(x)$. Then the WVAR of X is defined as

$$WVAR(X) = -\int_R xd(\Psi(F_X(x))),$$

where Ψ is an increasing, concave continuous function from $[0, 1]$ to $[0, 1]$ called the concave distortion. Let $(\Psi(u)_\gamma)_{\gamma \in R_+}$ be a family of concave distortions with $\Psi_\gamma(u)$ strictly increasing in γ for all $u \in [0, 1]$. The WVAR acceptability index (AIW) is given by

$$AIW(X) = \sup\{\gamma \in R_+ : \int_R xd(\Psi_\gamma(F_X(x))) \geq 0\}.$$

By Theorem 2.1.2, AIW is associated with a sequence of increasing sets of test measures $(\mathcal{D}_\gamma)_{\gamma \in R_+}$. If we fix the set of test measures as $\mathcal{D}_\gamma$ for a given γ, we then get a two price economy in which the bid and ask prices have the following representations,

$$b_\gamma(X) = \int_R x d(\Psi_\gamma(F_X(x))), \tag{2.1}$$

and

$$a_\gamma(X) = -\int_R x d(\Psi_\gamma(F_{-X}(x))). \tag{2.2}$$

Observe that if the concave distortion Ψ_γ is the identity function, equation (2.1) is simply the expectation of X. Therefore, the bid price can be viewed as a distorted expectation under the distribution given by $\Psi_\gamma(F_X(x))$. From (2.1), we can obtain the following two properties about the bid price functional induced by AIW,

1. Law invariant: if $X \overset{d}{=} Y$, then $b(X) = b(Y)$,

2. Linear in comonotone variables: if $(X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \geq 0$ a.s., then $b(X + Y) = b(X) + b(Y)$.

Hence, the bid and ask prices are nonlinear expectations defined on $L^\infty(\Omega, \mathcal{F}, P)$, with linearity preserved only for comonotone random variables. The size of the resulting two price economy is controlled by γ through the function $\Psi_\gamma(u)$. Recall that $(\Psi(u)_\gamma)_{\gamma \in R_+}$ is a family of concave distortions increasing in γ. An example of such family of functions termed minmaxvar is proposed by Cherny and Madan in [8], and is defined by

$$\Psi_\gamma(u) = 1 - (1 - u^{\frac{1}{1+\gamma}})^{1+\gamma}. \tag{2.3}$$

All the computations related with concave distortions in this thesis will employ the minmaxvar function.

2.2 Discrete Time Two Price Model

The discrete time two price model is developed in [35] by Madan, Pistorius and Schoutens. The model could be seen as a multi-period extension of the one period two price model. Consider a discrete time two price economy defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, P)$. Assume the filtration is generated by a discrete time finite state Markov chain $(X_t)_{0\leq t\leq T}$. Without loss of generality, we can use standard basis vectors $\{e_1, ...e_N\} \subset \mathbb{R}^N$ to identify the states of the Markov chain, where $e_i = (0, 0, ...0, 1, 0, ...0)^T$ and N is the total number of states. Suppose the discrete time two price economy trades state contingent terminal cash flows $C \in \mathcal{C} \subseteq L^2(\mathcal{F}_T)$. We note that since there could only be finitely many possible paths for X on $t = 0, 1, ...T$, we have

$$L^2(\mathcal{F}_T) = L^\infty(\mathcal{F}_T).$$

As an analogue to the one period two price economy, in which the bid and ask prices are nonlinear expectations on $(\Omega, \mathcal{F}, P)$, the two prices in discrete time are modeled as dynamically consistent nonlinear expectations defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, P)$. A system of operators

$$\mathcal{E}(\cdot|\mathcal{F}_t) : L^2(\mathcal{F}_T) \to L^2(\mathcal{F}_t), 0 \leq t \leq T$$

is called a dynamically consistent nonlinear expectation if it satisfies the following properties:

1. Monotonicity: if $X \leq Y$ a.s., then $\mathcal{E}(X|\mathcal{F}_t) \leq \mathcal{E}(Y|\mathcal{F}_t)$ a.s.,

2. $\mathcal{F}_t$-triviality: $\mathcal{E}(X|\mathcal{F}_t) = X$ a.s. for any $\mathcal{F}_t$ measurable X,

3. Recursivity: $\mathcal{E}(\mathcal{E}(X|\mathcal{F}_t)|\mathcal{F}_s) = \mathcal{E}(X|\mathcal{F}_s)$ a.s. for any $s \leq t$,

4. Regularity: $\mathbf{1}_A\mathcal{E}(X|\mathcal{F}_t) = \mathcal{E}(\mathbf{1}_A X|\mathcal{F}_t)$ a.s. for any $A \in \mathcal{F}_t$.

Moreover, we want to preserve the concavity and convexity of the bid and ask pricing functionals, respectively. In [35], such nonlinear pricing functionals are constructed as solutions to backward stochastic difference equations. The connection between backward stochastic difference equations and dynamically consistent nonlinear expectations is established by Theorem 7 of Cohen and Elliott [12].

In the discrete two price economy, the uncertainty evolution is described by a N state Markov chain $(X_t)_{0\leq t\leq T}$. The discrete time Markov chain X can be represented as

$$X_t = E(X_t|\mathcal{F}_{t-1}) + M_t,$$

where M_t is a martingale that takes values in $\mathbb{R}^N$. A backward stochastic difference equation based on $(M_t)_{0\leq t\leq T}$ has the following form

$$Y_t - \sum_{t\leq u<T} F(\omega, u, Y_u, Z_u) + \sum_{t\leq u<T} Z_u^T M_{u+1} = C, \tag{2.4}$$

where F is an adapted map $F : \Omega \times \{0, ..., T\} \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ called the driver function, C a bounded $\mathcal{F}_T$ measurable random variable representing the terminal value. The solution to (2.4) is a pair (Y_t, Z_t) that satisfies the equation, with Y taking values in $\mathbb{R}$ and Z taking value in $\mathbb{R}^N$. By Theorem 7 in [12], the solution to

(2.4) is linked to nonlinear expectation through

$$Y_t = \mathcal{E}(C|\mathcal{F}_t).$$

One can rewrite (2.4) as

$$Y_t - F(\omega, t, Y_t, Z_t) + Z_t^T M_{t+1} = Y_{t+1},$$

with $Y_T = C$. Taking $E(\cdot|\mathcal{F}_t)$ on both sides yields

$$Y_t = E(Y_{t+1}|\mathcal{F}_t) + F(\omega, t, Y_t, Z_t). \tag{2.5}$$

As a result, bid and ask prices can be constructed backwardly from (2.5) by carefully selecting the driver functions. In [36], the driver functions for bid and ask prices are chosen to be

$$\begin{aligned} F_b(\omega, t, Y_t^b, Z_t) &= b_\gamma(Z_t^T M_{t+1}) \\ &= b_\gamma(Y_{t+1}^b - E[Y_{t+1}^b|\mathcal{F}_t]), \end{aligned}$$

and

$$\begin{aligned} F_a(\omega, t, Y_t^a, Z_t) &= a_\gamma(Z_t^T M_{t+1}) \\ &= a_\gamma(Y_{t+1}^a - E[Y_{t+1}^a|\mathcal{F}_t]), \end{aligned}$$

where the functions a_γ and b_γ are one step distorted expectations defined in (2.1) and (2.2). We observe from the definition of one step distorted expectations that the bid and ask pricing functionals are locally law invariant and translation invariant, which means

$$\mathcal{E}(C + Q|\mathcal{F}_t) = \mathcal{E}(C|\mathcal{F}_t) + Q \quad \text{for any } Q \in L^2(\mathcal{F}_t).$$

By construction, the bid price process satisfies

$$\begin{aligned}
Y_t^b &= E(Y_{t+1}^b|\mathcal{F}_t) + b_\gamma(Y_{t+1}^b - E[Y_{t+1}^b|\mathcal{F}_t]) \\
&= b_\gamma(Y_{t+1}^b) \\
&\leq E(Y_{t+1}^b|\mathcal{F}_t),
\end{aligned}$$

while the ask price process follows

$$\begin{aligned}
Y_t^a &= E(Y_{t+1}^a|\mathcal{F}_t) + a_\gamma(Y_{t+1}^a - E[Y_{t+1}^a|\mathcal{F}_t]) \\
&= a_\gamma(Y_{t+1}^a) \\
&\geq E(Y_{t+1}^a|\mathcal{F}_t).
\end{aligned}$$

Hence bid prices are submartingales whereas ask prices are supermartingales. We also have

$$Y_t^b \leq E(C|\mathcal{F}_t) \leq Y_t^a,$$

which is consistent with the one period two price economy.

Chapter 3

BSDEs on Markov Chains

3.1 Introduction

The theory of backward stochastic differential equations (BSDEs) has been extensively studied since its initial development in the work of Pardoux and Peng [44]. BSDEs have attracted much attention during the past two decades due to their connections with stochastic optimization problems. Most of the previous results are obtained in the context of continuous time diffusions, and therefore are unable to deal with any case when the underlying martingale could jump with positive probability. However, such cases may arise in various applications. In a series of papers [10], [11] by Cohen and Elliott, BSDEs generated by continuous time Markov chains (CTMCs) are proposed and studied in detail. Their results will be summarized in this section.

Consider a CTMC $X = (X_t)_{0 \le t \le T}$ with N states. Without loss of generality, we can use standard basis vectors $\{e_1, ...e_N\} \subset \mathbb{R}^N$ to identify the states of the Markov chain, where $e_i = (0, 0, ...0, 1, 0, ...0)^T$. This chain generates a filtered probability space $(\Omega,\mathcal{F}_T,\{\mathcal{F}_t\}_{0\le t\le T},P)$. Let $\mathcal{A}_t$ be the rate matrix of X at time t. Then this process has the following representation:

$$X_t = X_0 + \int_{]0,t]} A_u^T X_u du + M_t, \tag{3.1}$$

where M_t is a martingale. The predictable quadratic variation of M_t, $\langle M, M \rangle_t$, satisfies

$$\langle M, M \rangle_t = \int_{]0,t]} \left[diag(A_u^T X_{u-}) - diag(X_{u-}) A_u^T - A_u^T diag(X_{u-}) \right] du. \tag{3.2}$$

Proofs of (3.1) and (3.2) can be found in Appendix B in [19].

A BSDE on the Markov chain X is defined by

$$Y_t - \int_{]t,T]} F(\omega, u, Y_{u-}, Z_u) du + \int_{]t,T]} Z^T{}_u dM_u = Q, \tag{3.3}$$

where Q is a square integrable, $\mathcal{F}_T$ measurable, $\mathbb{R}$ valued random variable, and F is a progressively measurable function which takes values in $\mathbb{R}$, called the driver function. Let ψ_u denote the integrand in (3.2):

$$\psi_u = diag(A_u^T X_{u-}) - diag(X_{u-}) A_u^T - A_u^T diag(X_{u-}).$$

The matrix ψ_u is a nonnegative definite. For any two vectors $\tilde{u}, \tilde{v} \in \mathbb{R}^N$, the inner product induced by ψ_u is defined in [10] as

$$< \tilde{u}, \tilde{v} >_{X_{u-}} = \tilde{u}^T \psi_u(X_{u-}) \tilde{v}. \tag{3.4}$$

With some calculations, we could get for any vector $\tilde{v}$:

$$\|\tilde{v}\|^2_{X_{u-}} = \sum_{j: X_{u-}^T A_u e_j > 0} X_{u-}^T A_u e_j (v_j - v_{X_{u-}})^2.$$

As a result, the norm of $\tilde{v}$ is 0 if and only if $v_j = v_{X_{u-}}$ for all the states j with $X_{u-}^T A_u e_j > 0$.

The solution to the BSDE (3.3) is a pair $(Y_t,\ Z_t)$ that satisfies (3.3) for all $t \in [0, T]$, with Y_t taking values in $\mathbb{R}$ and Z taking values in R^N. Existence and uniqueness of the solution to (3.3) is proved in [10], which is stated below.

Theorem 3.1.1. *Let Q belong to $L^2(\mathcal{F}_T)$, and let F be Lipschitz continuous in the sense that there exists a constant C, such that for any Y^1,Y^2, Z^1, Z^2, square integrable and of appropriate dimension,*

$$|F(\omega, t, Y^1_{t-}, Z^2_t) - F(\omega, t, Y^2_{t-}, Z^2_t)| \leq C(|Y^1_{t-} - Y^2_{t-}| + \|Z^1_t - Z^2_t\|_{X_{t-}}) \quad a.s., dt \times P. \tag{3.5}$$

Then (3.3) has a unique solution (Y, Z), up to indistinguishability for Y and equality $[d\langle M, M\rangle_t \times P]$-a.s. for Z. Moreover, Y_t is an adapted càdlàg process with

$$E\left[\int_{]0,T]} |Y_t|^2 du\right] < +\infty,$$

and Z_t is predictable with

$$E\left[\int_{]0,T]} \|Z_t\|^2_{X_{t-}} du\right] < +\infty.$$

In [11], Cohen and Elliott established the connection between solutions of BSDEs on CTMCs with dynamically consistent nonlinear expectations. Before stating the main result of [11], we'll need the following comparison theorem, which is also proved in [11].

Theorem 3.1.2. *Suppose we have two standard scalar BSDEs with driver functions and terminal values (F^1, Q^1), (F^2, Q^2), and corresponding solutions (Y^1, Z^1) and (Y^2, Z^2). Assume the following conditions hold:*

1. *$Q^1 \geq Q^2$ P-a.s.,*

2. *$F^1(\omega, t, Y^2_{t-}, Z^2_t) \geq F^2(\omega, t, Y^2_{t-}, Z^2_t)$ $dt \times P$-a.s.,*

3. *there exists* $\epsilon > 0$*, such that for all* $t \in [0,T]$*,* P*-a.s., if* Z_t^1, Z_t^2 *satisfies*

$$(e_j^T A_t^T X_{t-})[Z_t^1 - Z_t^2]^T (e_j - X_{t-}) \geq -\epsilon \|Z_t^1 - Z_t^2\|_{X_{t-}} \tag{3.6}$$

for all e_j*,then*

$$F^1(\omega, t, Y_{t-}^2, Z_t^1) - F^1(\omega, t, Y_{t-}^2, Z_t^2) \geq -[Z_t^1 - Z_t^2]^T A_t^T X_{t-} \tag{3.7}$$

holds with equality only if $\|Z_t^1 - Z_t^2\|_{X_{t-}} = 0$.

Then $Y^1 \geq Y^2$ P*-a.s. And this comparison is strict, meaning if* $Y_t^1 = Y_t^2$ *for some* t *and on some set* $U \in \mathcal{F}_t$*, then* $Q^1 = Q^2$ P*-a.s. on* U *and* $F^1(\omega, s, Y_{s-}^2, Z_s^2) = F^2(\omega, s, Y_{s-}^2, Z_s^2)$ $[ds \times P]$*-a.s. on* $[t,T] \times U$*.*

3.2 Nonlinear Expectations Induced by BSDEs on CTMCs

Definition 3.2.1. *A system of operators,* $\mathcal{E}(\cdot|\mathcal{F}_t) : L^2(\mathcal{F}_T) \to L^2(\mathcal{F}_t)$*,* $0 \leq t \leq T$*, is called an* $\mathcal{F}_t$*-consistent nonlinear expectation for* $\mathcal{Q}_t \subseteq L^2(\mathcal{F}_T)$ *defined on* $[0,T]$*, if it satisfies the following properties:*

1. *for* $Q, Q' \in \mathcal{Q}_t$*, with* $Q \geq Q'$ $P-a.s.$*,*

$$\mathcal{E}(Q|\mathcal{F}_t) \geq \mathcal{E}(Q'|\mathcal{F}_t) \quad P-a.s., \tag{3.8}$$

   with equality holding iff $Q = Q'$ $P-a.s.$*,*

2. $\mathcal{E}(Q|\mathcal{F}_t) = Q$*,* $P-a.s.$*, for any* $\mathcal{F}_t$ *measurable* Q*,*

3. $\mathcal{E}(\mathcal{E}(Q|\mathcal{F}_t)|\mathcal{F}_s) = \mathcal{E}(Q|\mathcal{F}_s)$*,* $P-a.s.$*, for any* $s \leq t$*,*

4. *for any* $A \in \mathcal{F}_t$*,* $\mathbf{1}_A \mathcal{E}(Q|\mathcal{F}_t) = \mathcal{E}(\mathbf{1}_A Q|\mathcal{F}_t)$*,* $P-a.s.$*.*

A nonlinear expectation shares all the properties of the traditional expectation except the linearity. In [45], it is shown that under some restrictions, a BSDE on Brownian motion generates a dynamically consistent nonlinear expectation. An analogue to this connection is given in [11] by Cohen and Elliott. Their main result is stated below.

Theorem 3.2.2. *Fix a Lipschitz continuous (3.5) driver F that satisfies $F(\omega, t, Y, 0) = 0$ $dt \times P$-a.s.. Moreover, consider a family of sets $\mathcal{Q}_t \subseteq L^2(\mathcal{F}_T)$, such that for all Q, $Q^{'} \in \mathcal{Q}_t$ with $Q \geq Q^{'}$, the comparison theorem 3.1.2 holds with $F^1 = F^2 = F$. Define a system of operators $\mathcal{E}^F(\cdot|\mathcal{F}_t)$ by*

$$\mathcal{E}^F(Q|\mathcal{F}_t) = Y_t, \tag{3.9}$$

where Y_t is the solution to

$$Y_t - \int_{]t,T]} F(\omega, u, Y_{u-}, Z_u)du + \int_{]t,T]} Z_u^T dM_u = Q.$$

Then $\mathcal{E}^F(\cdot|\mathcal{F}_t)$ is a $\mathcal{F}_t$-consistent nonlinear expectation for $\mathcal{Q}_t$.

3.3 Continuous Time Modeling of Bid and Ask Prices

Recall that in Chapter 2, we have reviewed two models (static and discrete time) for two price economies. In order to extend the theory of two price economy to continuous time, we need to develop continuous time modeling of the bid and ask prices. The idea behind the discrete time two price economy, in which the two prices are constructed as nonlinear expectations, brings us naturally to dynamically consistent nonlinear pricing in continuous time. Theorem 3.2.2 provides a way to generate continuous time nonlinear expectations from BSDEs on Markov chains. Our task reduces to describing the BSDEs corresponding to the bid and ask pricing functionals.

Consider the following BSDE,

$$Y_t - \int_{]t,T]} F(X_u, u, Y_{u-}, Z_u)du + \int_{]t,T]} Z_u^T dM_u = Q(X_T), \tag{3.10}$$

where F is a Markovian driver function, and Q is a square integrable random variable that only depends on the terminal state of the underlying CTMC X. The solution to (3.10) is given by the theorem below.

Theorem 3.3.1. *Let F be a Lipschitz Markovian driver function. Suppose there exists a continuous vector $\tilde{V}$ that satisfies, for $i = 1, 2, ...N$,*

$$\frac{d\tilde{V}_i}{dt} - e_i^T A_{T-t}\tilde{V}(t) - F(e_i, T-t, \tilde{V}_i(t), \tilde{V}(t)) = 0, \tag{3.11}$$

and

$$\tilde{V}_i(0) = Q(e_i).$$

Then $(\tilde{V}^T(T-t)X_t, \tilde{V}(T-t))$ is the solution to (3.10).

Proof. Applying Itô's formula to $\tilde{V}(t)^T X_t$ yields

$$\begin{aligned} & d(\tilde{V}(T-t)^T X_t) && (3.12) \\ = \;& (A_t\tilde{V}(T-t))^T X_t dt + F(X_t, t, \tilde{V}(T-t)^T X_t, \tilde{V}(T-t))dt + \tilde{V}(T-t)^T dX_t. \end{aligned}$$

From (3.1),

$$dX_t = A_t^T X_t dt + dM_t.$$

Substituting into (3.12), we get

$$
\begin{aligned}
d(\tilde{V}^T(T-t)X_t) &= -(\tilde{V}^T(T-t)A_t^T X_t + F(X_t, t, \tilde{V}^T(T-t)X_t, \tilde{V}(T-t)))dt \\
&\quad + \tilde{V}^T(T-t)A_t^T X_t dt + \tilde{V}(T-t)^T dM_t \\
&= -F(X_t, t, \tilde{V}^T(T-t)X_t, \tilde{V}(T-t))dt + \tilde{V}(T-t)^T dM_t.
\end{aligned}
$$

As a result, we have

$$
\begin{aligned}
&\tilde{V}^T(T-t)X_t - \int_{]t,T]} F(X_u, u, \tilde{V}^T(T-u)X_u, \tilde{V}(T-u))du + \int_{]t,T]} \tilde{V}(T-u)^T dM_u \\
&= Q(X_t).
\end{aligned}
$$

Note that $\tilde{V}^T(T-u)X_u = \tilde{V}^T(T-u-)X_{u-}$ except for countably may u s, we get

$$
\begin{aligned}
&\tilde{V}^T(T-t)X_t - \int_{]t,T]} F(X_u, u, \tilde{V}^T(T-u-)X_{u-}, \tilde{V}(T-u))du + \int_{]t,T]} \tilde{V}(T-u)^T dM_u \\
&= Q(X_t),
\end{aligned}
$$

which is in the same form as (3.10) with $Y_t = \tilde{V}^T(T-t)X_t$ and $Z_t = \tilde{V}(T-t)$. $\square$

We will use the BSDE with Markovian driver and terminal value as our tool for modeling continuous time two price economy. For computing bid prices, we pick the driver to be

$$
F_b(X_t, t, Y_t, Z_t) = |X_t^T A_t X_t| \int_R z d(\Psi_\gamma(F_Z(z))) - X_t^T A_t Z_t, \tag{3.13}
$$

where Ψ_γ is the minmaxvar function defined in (2.3), and $F_Z(z)$ is the distribution function of a random variable Z that takes values in $(e_i - X_t)^T Z_t$ for all $e_i \neq X_t$, with

$$
P(Z = (e_i - X_t)^T Z_t) = \frac{X_t^T A_t e_i}{|X_t^T A_t X_t|}. \tag{3.14}
$$

The driver function for generating ask prices is

$$F_a(X_t, t, Y_t, Z_t) = -F_b(X_t, t, Y_t, -Z_t). \tag{3.15}$$

The probability mass function of Z defined in (3.14) is the same as the transition probability at state X_t of the embedded Markov chain associated with the CTMC $(X_t)_{0\le t\le T}$. The driver functions (3.13), (3.15) can be seen as scaled distorted expectations of $Z_t dM_t$. By construction, $F_b \le 0 \le F_a$, we have for any $0 \le s < t \le T$,

$$\begin{aligned} Y_s^b &= E(Y_t^b|\mathcal{F}_s) + E(\int_{]s,t]} F(X_u, u, Y_{u-}^b, Z_u)du|\mathcal{F}_s) \\ &\le E(Y_t^b|\mathcal{F}_s), \end{aligned}$$

and

$$\begin{aligned} Y_s^a &= E(Y_t^a|\mathcal{F}_s) + E(\int_{]s,t]} F(X_u, u, Y_{u-}^a, Z_u)du|\mathcal{F}_s) \\ &\ge E(Y_t^a|\mathcal{F}_s). \end{aligned}$$

Hence the bid prices are submartingales whereas the ask prices are supermartingales, which is consistent with the discrete time two price economy. Moreover, the bid and ask pricing functionals induced by the BSDEs are dynamically consistent nonlinear expectations. This property is given by the proposition below.

Proposition 3.3.2. *Consider a BSDE of the following form*

$$Y_t - \int_{]t,T]} F(X_u, u, Y_{u-}, Z_u)du + \int_{]t,T]} Z_u^T dM_u = Q(X_T),$$

where the driver function F is either F_b or F_a, defined in (3.13) and (3.15). For any $t \in [0, T]$, let $\mathcal{Q}_t$ denote the set of $\mathcal{F}_t$ measurable square integrable random variables *that only depend on X_t. Define a system of operators*

$$\mathcal{E}^F(\cdot|\mathcal{F}_t) : \mathcal{Q}_T \to \mathcal{Q}_t$$

by

$$\mathcal{E}^F(\cdot|\mathcal{F}_t) = Y_t.$$

Then $\mathcal{E}^F$ is a dynamically consistent nonlinear expectation defined for $\{\mathcal{Q}_t\}_{0\le t\le T}$.

We first prove the following lemma.

Lemma 3.3.1. *Let $\tilde{V}^1$ and $\tilde{V}^2$ be solutions to (3.11). If for $i = 1, ..., N$, $\tilde{V}_i^1(0) > \tilde{V}_i^2(0)$, then $\tilde{V}_i^1(t) \ge \tilde{V}_i^2(t)$ for all $t \in [0, T]$, $i = 1, ...N$.*

Proof. For any two vectors $\tilde{U}$ and $\tilde{V}$ in $\mathbb{R}^N$, let $\tilde{U} \succ \tilde{V}$ denote

$$\tilde{U}_i > \tilde{V}_i, \quad i = 1, ..., N,$$

and $\tilde{U} \succeq \tilde{V}$ denote

$$\tilde{U}_i \ge \tilde{V}_i, \quad i = 1, ..., N.$$

Equation (3.11) could be written as

$$\frac{d\tilde{V}}{dt} = \Lambda(t)(\mathcal{E}(\tilde{V}) - \tilde{V}), \tag{3.16}$$

where $\Lambda(t)$ is a diagonal matrix with $\Lambda_{ii}(t) = |e_i^T A_{T-t} e_i|$ for all $t \in [0, T]$, and $\mathcal{E}(\tilde{V})$ an N dimensional vector satisfying

$$\mathcal{E}(\tilde{V})_i = \mathcal{E}(Z_i),$$

for $i = 1, ...N$. The Z_i above denotes a random variable that takes values $e_j^T \tilde{V}$ for $e_j \ne e_i$, with probability mass function

$$P(Z_i = e_j^T \tilde{V}) = \frac{e_i^T A_t e_j}{\Lambda_{ii}},$$

and $\mathcal{E}(Z_i)$ stands for $b_\gamma(Z_i)$ or $a_\gamma(Z_i)$ depending on the form of the driver function F. Therefore, $\tilde{V}^1 - \tilde{V}^2$ satisfies

$$\frac{d(\tilde{V}^1 - \tilde{V}^2)}{dt} = \Lambda(t)[\mathcal{E}(\tilde{V}^1) - \mathcal{E}(\tilde{V}^2) - (\tilde{V}^1 - \tilde{V}^2)]. \tag{3.17}$$

Suppose Lemma 3.3.1 does not hold. Define

$$\tau = \inf\{t \geq 0 | \exists i, V_i^1(t) < V_i^2(t)\}.$$

Since V^1 and V^2 are continuous, we have $\tau > 0$, and $V_i^1(\tau) \geq V_i^2(\tau)$, with equality holding for at least one i. By definition of τ, we infer that when $t \in [0, \tau]$, $V_i^1(t) > V_i^2(t)$ for $i = 1, ..., N$. Consider (3.17) on the interval $[0, \tau]$. Since $\mathcal{E}(\tilde{V}^1) \succeq \mathcal{E}(\tilde{V}^2)$, by the comparison theorem for ordinary differential equations, $(\tilde{V}^1 - \tilde{V}^2)$ is greater than the solution to

$$\frac{d\tilde{U}}{dt} = -\Lambda(t)\tilde{U},$$

with initial condition $\tilde{U}(0) = \tilde{V}^1(0) - \tilde{V}^2(0)$. Thus we have

$$\tilde{V}^1(\tau) - \tilde{V}^2(\tau) \succeq \tilde{U}(\tau) = (\tilde{V}^1(0) - \tilde{V}^2(0))exp\left[\int_0^\tau \Lambda(s)ds\right] \succ 0,$$

which contradicts with the fact that $V_i^1(\tau) = V_i^2(\tau)$ for some i. □

We next prove Proposition 3.3.2.

Proof. Properties 2 through 4 in Definition 3.2.1 follow directly from Theorem 3.2.2. However, our picks of F do not satisfy the requirements for the comparison theorem (Theorem 3.1.2), and so we will prove the monotonicity directly. By Theorem 3.3.1, we need to prove that for any $\tilde{V}^1$ and $\tilde{V}^2$ satisfying (3.11), if $\tilde{V}^1(0) \succeq \tilde{V}^2(0)$, then $\tilde{V}^1(t) \succeq \tilde{V}^2(t)$ for all $t \in [0, T]$.

By definition, for any vector $\tilde{V}$ and any $\epsilon > 0$, we have

$$\mathcal{E}(\tilde{V}) - V = \mathcal{E}(\tilde{V} - \epsilon) - (V - \epsilon).$$

Recall that $\tilde{V}^2$ satisfies (3.16). Hence $\tilde{V}^2 - \epsilon$ also satisfies (3.16), with initial value given by $\tilde{V}^2(0) - \epsilon$. Use Lemma 3.3.1 and $\tilde{V}^1(0) \succ \tilde{V}^2(0) - \epsilon$ to obtain

$$\tilde{V}^1(t) \succeq \tilde{V}^2(t) - \epsilon, \quad \text{for any } t \in [0, T].$$

Since ϵ is arbitrary, we have

$$\tilde{V}^1(t) \succeq \tilde{V}^2(t), \quad \text{for any } t \in [0, T].$$

□

3.4 A Convergence Theorem of BSDEs on CTMCs

3.4.1 Overview

In this section, we are going to prove a convergence theorem of BSDEs on CTMCs. Consider a sequence of BSDEs on CTMCs,

$$Y_t^n = \xi^n + \int_t^T F^n(u, Y_{u-}^n, Z_u^n) - \int_t^T (Z_u^n)^T dM_u^n, \tag{3.18}$$

driven by a sequence of CTMCs X^n with transition rate matrix A_t^n and state value vector V^n.

Also, consider a BSDE generated by a Levy process X,

$$Y_t = \xi + \int_t^T f(s, Y_{s-}, Z_s) ds - \sum_{i=1}^{\infty} \int_t^T Z_s^{(i)} dH_s^{(i)} \tag{3.19}$$

where $H_t^{(i)}$ is the orthonormalized Teugels martingale of order i associated with the Levy process X. The solution to (3.19) is a pair (Y_t, Z_t), in which Y_t has càdlàg trajectories on $[0,T]$ and $Z_t \in \mathbb{R}^\infty$. BSDEs driven by Levy processes are introduced and studied by Nualart and Schoutens in [43]. The existence and uniqueness of the solution of (3.19) is shown in [43].

Let $X^{(n)} = (V^n)^T X^n$, the Markov chain that follows the same dynamics as X^n and takes values in V^n. We are interested in the asymptotic behavior of the solutions to (3.18), as the sequence of Markov chains $X^{(n)}$ converges to a Levy process X.

3.4.2 Preliminaries

Before stating the main result of this section, we'll introduce the preliminary background related to our result. The solutions to both (3.18) and (3.19) are pairs, with their first components having càdlàg trajectories on $[0,T]$. To characterize the convergence of solutions, we'll need a metric to measure the closeness between càdlàg functions. Let $D[0,T]$ denote the set of all càdlàg functions from $[0,T]$ to $\mathbb{R}$. One commonly used metric on $D[0,T]$ is the J_1-Skorokhod metric. It was first introduced by Skorokhod in [50] to characterize convergence of sample paths of stochastic processes . The definition of J_1-Skorokhod metric is given below.

Definition 3.4.1. *Let $f,g \in D[0,T]$, the $J_1 -$ Skorokhod metric d on $D[0,T]$ is defined by*

$$d_{J_1}(f,g) = \inf_{\lambda \in \Lambda} max\{\|\lambda - I\|_\infty, \|f - g \circ \lambda\|_\infty\},$$

where Λ denotes the class of strictly increasing, continuous mappings of $[0,T]$ onto *itself, $\|\cdot\|_\infty$ is the L^∞ norm, and I is the identity function.*

$D[0,T]$ together with the J_1-Skorokhod metric induce a topological space.

In order to characterize the convergence of random processes, we will use the Aldou's extended convergence introduced in [41], which is defined as follows,

Definition 3.4.2. *Given a sequence of random processes X^n defined on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t^n), P)$ and a process X defined on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, we say that $(X^n, (\mathcal{F}_t^n))$ converges to $(X, (\mathcal{F}_t))$ in the Aldous extended sense, if for every bounded Borelian function Φ from $D[0,T]$ to R, the sequence of càdlàg processes $(X_t^n, (E[\Phi(X^n)|\mathcal{F}_t^n))$ converges in probability under J_1-Skorokhod topology to the process $(X_t, (E[\Phi(X)|\mathcal{F}_t]))$.*

There exists many ways to check if a sequence of $(X^n, (F_t^n))$ converges to $(X, (F_t))$ in the Aldous extended sense. We'll use Theorem 1 and Proposition 2(i) in paper [29].

Next we discuss the orthonormalized martingales induced by Teugels martingales of a Levy process. The concept is introduced in [43] by Nualart and Schoutens. For any Levy process X, its power-jump processes are defined as follows,

$$X_t^{(1)} = X_t,$$

$$X_t^{(i)} = \sum_{0<s\le t} (\Delta X_s)^i, \qquad i \ge 2.$$

The Teugels martingales are

$$Y_t^{(i)} = X_t^{(i)} - E(X_t^{(i)}), \qquad \text{for } i \ge 1. \tag{3.20}$$

Similar to the Gram-Schmidt method, orthonormalization could be applied to the Teugels martingales $Y_t^{(i)}$ so as to get a sequence of strongly orthonormal martingales $H_t^{(i)}$, that is $< H_t^{(i)}, H_t^{(j)} >= 0$ for $i \neq j$ and $< H_t^{(i)}, H_t^{(j)} >= t$, where each H^i is a linear combination of $\{Y^{(j)}\}_{j=1}^{i}$. Details of the orthonormalization method can be found in [43].

Now consider an N-state CTMC X, with transition rate matrix $A(t)$ satisfying

$$\forall i, j, \quad a_{ij}(t) \neq 0.$$

Let us also suppose that the state value vector V of X consists of N distinct values. Inspired by the Teugels martingales for a Levy process, we define the generalized Teugels martingales for the CTMC X with N states and state value vector V.

$$X_t^{(i)} = \sum_{0<s\leq t} (\Delta(V^T X_t))^i \quad 1 \leq i \leq N-1.$$

Similar to (3.1), $X_t^{(i)}$ also has a martingale representation form:

$$X_t^{(i)} = \int_0^t ((V - V^T X_{u-})^i)^T A_u^T X_{u-} du + \int_0^t ((V - V^T X_{u-})^i)^T dM_u.$$

Let $Y_t^{(i)}$ be $\int_0^t ((V - V^T X_{u-})^i)^T dM_u$, the martingale part of $X_t^{(i)}$. Similar to the orthonormalization of Teugels martingales induced by a Levy process, we seek to find a set of orthonormal martingales induced by $Y_t^{(i)}$.

For any two martingales of the form $M_1(t) = \int_0^t Z_1(u)^T dM_u$, and $M_2(t) = \int_0^t Z_2(u)^T dM_u$ with predictable vectors Z_1 and Z_2, the predictable cross variation between M_1 and M_2 could be written as

$$< M_1, M_2 >_t = \int_0^t < Z_1(u), Z_2(u) >_{X_{u-}} ds,$$

where $<\cdot,\cdot>_{X_{u-}}$ denotes the inner product induced by M, which is defined in (3.4). Let $V^{(i)}$ denote $(V - V^T X_{u-})^i$. We have the following lemma.

Lemma 3.4.3. *For any fixed time u, the set of vectors $\{V^{(i)}(u)\}_{i=1}^{N-1}$ has rank $N-1$.*

Proof. Without loss of generality, suppose $X_{u-} = j$. Then

$$[V^{(1)}, V^{(2)}, ..., V^{(N)}] = \begin{bmatrix} v_1 - v_j & (v_1 - v_j)^2 & \dots & (v_1 - v_j)^{N-1} \\ v_2 - v_j & (v_2 - v_j)^2 & \dots & (v_2 - v_j)^{N-1} \\ \vdots & \ddots & & \vdots \\ 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ v_N - v_j & (v_N - v_j)^2 & \dots & (v_N - v_j)^{N-1} \end{bmatrix}.$$

If we add a vector of 1s as the first column in front of $\{V^{(i)}\}_{i=1}^{N-1}$, we get a full rank Vandermonde matrix. Therefore the original matrix has rank $N-1$. □

As a result, for each $u \in [0,T]$ we could apply the Gram-Schmidt method to $\{V^{(i)}(u)\}_{i=1}^{N-1}$ to obtain a set of orthonormal vectors $\{C^{(i)}(u)\}_{i=1}^{N-1}$ with respect to (w.r.t) the inner product induced by M. Define $\{H^{(i)}\}_{i=1}^{N-1}$ as

$$H^{(i)} = \int_0^t (C^{(i)}(u))^T dM_u.$$

Then $\{H^{(i)}\}_{i=1}^{N-1}$ forms a set of strongly orthonormal martingales. Let us call $\{H^{(i)}\}_{i=1}^{N-1}$ the generalized orthonormal martingales induced by the CTMC X. We have the following generalized martingale representation theorem for $\{H^{(i)}\}_{i=1}^{N-1}$.

Theorem 3.4.4. *Let K_t be a martingale adapted to $\mathcal{F}_t^n$, the filtration generated by a CTMC $(X_t)_{0\leq t\leq T}$ with N states and nonzero transition rate. Then K_t can be written* *as*

$$K_t = \sum_{i=1}^{N-1} \int_0^t \alpha_i(u) dH_u^{(i)}, \qquad (3.21)$$

where $\{H^{(i)}\}_{i=1}^{N-1}$ *are the generalized orthonormal martingales of* X*.*

Proof. By the martingale representation theorem for CTMC (Lemma 3.1 in [10]), K_t can be written as

$$K_t = \int_0^t D(u)^T dM_u, \qquad (3.22)$$

for some predictable $D(u)$. For each u, $D(u)$ is a vector in $\mathbb{R}^N$. Recall that $\{C^{(i)}\}_{i=1}^{N-1}$ is the orthonormalized $\{V^(i)\}_{i=1}^{N-1}$. For each $1 \leq i \leq N-1$, $(V^i)^T X_{u-} = 0$, which implies $(C^{(i)})^T X_{u-} = 0$. From (3.1), we get

$$dM_u = dX_u - A_u^T X_{u-} dt.$$

Thus

$$\tilde{1}^T dM_u = \tilde{1}^T dX_u - \tilde{1}^T A_u^T X_{u-} dt = \tilde{1}^T dX_u - \tilde{0}^T X_{u-} dt = 0.$$

As a result, if we replace $D(u)$ by $D(u) - D(u)^T X_{u-}$ in (3.22), we will get the same K_t. Use $\tilde{D}(u)$ to denote $D(u) - D(u)' X_{u-}$. Then $\tilde{D}(u)^T X_{u-}$ is 0. Since $\{C^{(i)}\}_{i=1}^{N-1}$ has full rank with $(C^{(i)})^T X_{u-} = 0$, $\tilde{D}(u)$ is in the span of $\{C^{(i)}\}_{i=1}^{N-1}$. Therefore, there exists predictable $\{\alpha_i\}_{1 \leq i \leq N-1}$, such that

$$\tilde{D}(u) = \sum_{i=1}^{N-1} \alpha_i(u) C^{(i)}(u).$$

Thus we obtain

$$\begin{aligned} K_t &= \int_0^t \tilde{D}(u)^T dM_u \\ &= \sum_{i=1}^{N-1} \int_0^t \alpha_i(u) C^{(i)}(u)^T dM_u \\ &= \sum_{i=1}^{N-1} \int_0^t \alpha_i(u) dH_u^{(i)}. \end{aligned}$$

$\square$

3.4.3 A Convergence Theorem

Let $X^{(n)}$ be a sequence of continuous time Markov chain and X a Levy process. Suppose that

1. $(X^{(n)}, \mathcal{F}_t^n)$ converges to $(X, \mathcal{F}_t)$ in the Aldous extended sense,
2. for any N, and sufficiently large n, the first N generalized orthonormal martingales of $X^{(n)}$, $\{H^{n(i)}\}_{i=1}^N$ converge in probability to the first N orthonormal martingales of X $\{H^{(i)}\}_{i=1}^N$ in the J_1-Skorokhod topology. Moreover, $E[|H_T^{(i)}|^3] + \sup_n E[|H_T^{n(i)}|^3] < \infty$ for $1 \le i \le N$,
3. ξ^n converges to ξ in L^2 as $n \to \infty$, where ξ^n is $\mathcal{F}_T^n$ measurable and ξ is $\mathcal{F}_T$ measurable,
4. $E[|\xi|^3] + \sup_n E[|\xi^n|^3] < \infty$,
5. for any $\epsilon > 0$, the jump measure ν of X satisfies for some $\lambda > 0$,

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

6. f is Lipschitz continuous with Lipschitz constant K, and $f(s, 0, 0)$ is square integrable,

7. $F^n(t, Y, Z) = f(t, Y, \tilde{Z})$, where $\tilde{Z}_i =< Z, C^{(i)} >_{X_{t-}}$.

Theorem 3.4.1. *Suppose assumptions* $1-7$ *hold, Then there exists a sequence of solutions* (Y^n, Z^n) *to the BSDEs*

$$Y_t^n = \xi^n + \int_{]t,T]} F^n(s, Y_{s-}^n, Z_s^n)ds - \int_{]t,T]} (Z_s^n)^T dM_t^n$$

that converges to the solution $(Y, (Z^{(i)})_{i=1}^{\infty})$ *to the BSDE (3.19) in the following sense:*

$$d_{J_1}(Y^n, Y) + \int_0^T \|\tilde{Z}^n - Z\|^2 ds \longrightarrow 0 \quad \textit{in probability as } n \to \infty, \tag{3.23}$$

where $\tilde{Z}^n$ *is defined by*

$$\tilde{Z}_i^n =< Z^n, C^{n(i)} >_{X_{t-}^n} .$$

Before proving the theorem, we'll state the following lemmas.

Lemma 3.4.5. *For any sequence of random variables* ζ^n *and* ζ *satisfying condition 3,* $E(\zeta^n|\mathcal{F}_\cdot^n)$ *converges in probability to* $E(\zeta|\mathcal{F}_\cdot)$ *in the* J_1*-Skorokhod topology.*

Proof.

$$\begin{aligned} d(E(\zeta|\mathcal{F}_t), E(\zeta^n|\mathcal{F}_t^n))_{J_1} &\leq d(E(\zeta|\mathcal{F}_t), E(\zeta|\mathcal{F}_t^n))_{J_1} + d(E(\zeta|\mathcal{F}_t^n), E(\zeta^n|\mathcal{F}_t^n))_{J_1} \\ &\leq d(E(\zeta|\mathcal{F}_t), E(\zeta|\mathcal{F}_t^n))_{J_1} + \sup_t |E(\zeta - \zeta^n|\mathcal{F}_t^n)|. \end{aligned}$$

By Doob's martingale inequality, for any ϵ, we have

$$P(\sup_{t \leq T} |E(\zeta - \zeta^n|\mathcal{F}_t^n)| > \epsilon) \leq \frac{E(|\zeta - \zeta^n|^2)}{\epsilon^2}.$$

Since $\mathcal{F}^n$ converges weakly to $\mathcal{F}$, $E(\zeta|\mathcal{F}_t^n)$ converges in probability to $E(\zeta|\mathcal{F}_t)$ in the J_1-Skorokhod topology. In all, we have that $E(\zeta^n|\mathcal{F}_\cdot^n)$ converges in probability to $E(\zeta|\mathcal{F}_\cdot)$ in the J_1-Skorokhod topology. □

Lemma 3.4.6. *Let $\{H_\cdot^n\}_{n=1}^\infty$ be a sequence of $\mathcal{F}_\cdot^n$ martingales that converges in probability to an $\mathcal{F}_\cdot$ martingale $H_\cdot$ in the J_1-Skorokhod topology. Suppose we have extended convergence $(H_\cdot^n, \mathcal{F}_\cdot^n) \to (H_\cdot, \mathcal{F}_\cdot)$. Assume H_T^n and H_T satisfies condition 4, ξ^n and ξ satisfies conditions 3 and 4. Let K_t^n and K_t denote $E(\xi^n|\mathcal{F}_t^n)$ and $E(\xi|\mathcal{F}_t)$ respectively. Then $< K^n, H^n >$ converges in probability to $< K, H >$ uniformly on $[0,T]$ as $n \to \infty$.*

Proof. Since $H_\cdot^n$ converges in probability to $H_\cdot$ in the J_1-Skorokhod topology, by definition of the J_1 metric, H_T^n converges to H_T in probability and then in L^2 by L^3-boundedness. As a result, $\xi^n \pm H_T^n$ converge to $\xi \pm H_T$ in L^2. By Corollary 12 in [41], $< K^n + H^n, K^n + H^n >$ converges in probability to $< K + H, K + H >$ under J_1. Since $< K^n + H^n, K^n + H^n >$ and $< K + H, K + H >$ are continuous, the convergence in J_1-Skorokhod topology actually gives uniform convergence in $[0,T]$. Similarly, $< K^n - H^n, K^n - H^n >$ converges in probability to $< K - H, K - H >$ uniformly in $[0,T]$. By the fact that

$$< K^n, H^n >= \frac{< K^n + H^n, K^n + H^n > - < K^n - H^n, K^n - H^n >}{4}, \tag{3.24}$$

and

$$< K, H >= \frac{< K + H, K + H > - < K - H, K - H >}{4}, \tag{3.25}$$

$< K^n, H^n >$ converges in probability to $< K, H >$ uniformly on $[0,T]$. □

Lemma 3.4.7. *Let $N^{(n)}$ denote the number of states of the CTMC $X^{(n)}$. For any vector $V \in \mathbb{R}^m$, let $\widetilde{V}$ denote its extension to $\mathbb{R}^\infty$, that is $\widetilde{V} = (V^T, 0, 0, \ldots)^T$. Suppose assumptions $1-7$ hold. Then there exists a sequence $(Z_t^n)_{0\le t\le T}$ of $\mathcal{F}^n_{\cdot}$-predictable processes, and an $\mathcal{F}_{\cdot}$-predictable process $(Z_t)_{0\le t\le T}$ such that for any $t \in [0,T]$,*

$$E(\xi^n|\mathcal{F}_t^n) = E(\xi^n) + \sum_{i=1}^{N^{(n)}-1} \int_0^t Z_s^{n(i)} dH_s^{n(i)}, \tag{3.26}$$

$$E(\xi|\mathcal{F}_t) = E(\xi) + \sum_{i=1}^{\infty} \int_0^t Z_s^{(i)} dH_s^{(i)}, \tag{3.27}$$

and

$$\int_0^T \|\widetilde{Z_t^n} - Z_t\|^2 dt \to 0 \quad \textit{in probability.}$$

Moreover, for any $\delta \in (0,1)$, we have

$$E(\sum_{i=1}^{\infty} \int_0^T (\widetilde{Z_t^{n(i)}} - Z_t^{(i)})^{1+\delta} dt) \to 0.$$

Proof. Equation (3.26) and (3.27) come directly from martingale representation theorems for CTMCs and Levy processes. Let K_t denote $E(\xi|\mathcal{F}_t)$. Apply Itô's formula to K_T^2. It follows that

$$\begin{aligned} K_T^2 - K_0^2 &= \int_0^T 2K_{t-} dK_t + \int_0^T d[K,K]_t \\ &= \int_0^T 2K_{t-} \sum_{i=1}^{\infty} Z_s^{(i)} dH_s^{(i)} + \int_0^T \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} Z_t^{(i)} Z_t^{(j)} d[H^{(i)}, H^{(j)}]_t. \end{aligned}$$

Taking expectation on both sides yields

$$E(\xi^2) - E(\xi)^2 = \sum_{i=1}^{\infty} E(\int_0^T |Z_t^{(i)}|^2 dt).$$

By assumption 4, $E(\xi^2)$ and $E(\xi)$ are both bounded, so we have

$$\sum_{i=1}^{\infty} E(\int_0^T |Z_t^{(i)}|^2 dt) < C \tag{3.28}$$

for some constant C. For any $\epsilon > 0$, there exists N_1, such that

$$\sum_{i=N_1+1}^{\infty} E(\int_0^T |Z_t^{(i)}|^2 dt) < \epsilon^2. \tag{3.29}$$

By condition 2, we have $\{H^{n(i)}\}_{i=1}^{N_1}$ converge in probability to $\{H^{(i)}\}_{i=1}^{N_1}$ in the J_1-Skorokhod topology, with $E[|H_T^{(i)}|^3] + \sup_n E[|H_T^{n(i)}|^3] < \infty$ for $1 \leq i \leq N_1$. Let K_t^n denote $E(\xi^n|\mathcal{F}_t^n)$. Since ξ^n and ξ satisfy condition 3, by Lemma 3.4.5, we have K^n converges to K in probability under J_1 Skorokhod topology. Apply Lemma 3.4.6 to K^n, K, $\{H^{n(i)}\}_{i=1}^{N_1}$ and $\{H^{(i)}\}_{i=1}^{N_1}$ and use conditions 2 through 4, it follows that $< K^n, H^{n(i)} >$ converges in probability to $< K, H^{(i)} >$ uniformly in $[0, T]$ for $1 \leq i \leq N_1$, which is equivalent to

$$\sup_{0 \leq t \leq T} |\int_0^t Z_s^{n(i)} ds - \int_0^t Z_s^{(i)} ds| \to 0 \quad \text{in probability for } 1 \leq i \leq N_1.$$

Apply Lemma 3.4.6 to K^n and K to see that $< K^n, K^n >$ converges in probability to $< K, K >$ uniformly in $[0, T]$. That is

$$\sup_{0 \leq t \leq T} |\sum_{i=1}^{N^{(n)}-1} \int_0^t |Z_s^{n(i)}|^2 ds - \sum_{i=1}^{\infty} \int_0^t |Z_s^{(i)}|^2 ds| \to 0 \quad \text{in probability.}$$

Thus, we could extract a subsequence (indexed by n_k), such that

$$\sup_{0 \leq t \leq T} |\int_0^t Z_s^{n_k(i)}(\omega) ds - \int_0^t Z_s^{(i)}(\omega) ds| \to 0 \quad \text{almost surely for } 1 \leq i \leq N_1, \tag{3.30}$$

and

$$\sup_{0 \leq t \leq T} |\sum_{i=1}^{N^{(n_k)}-1} \int_0^t |Z_s^{n_k(i)}(\omega)|^2 ds - \sum_{i=1}^{\infty} \int_0^t |Z_s^{(i)}(\omega)|^2 ds| \to 0 \quad \text{almost surely.}$$

By (3.30), for almost every ω, $Z^{n_k(i)}(\omega)$ converges weakly to $Z^{(i)}(\omega)$ in $L^2([0, T])$ for $1 \leq i \leq N_1$. As a result, we have

$$\liminf_{k \to \infty} \int_0^T |Z_s^{n_k(i)}(\omega)|^2 ds \geq \int_0^T |Z_s^{(i)}(\omega)|^2 ds.$$

Let $M(\omega)$ denote $\sum_{i=1}^{\infty}\int_0^T |Z_s^{(i)}(\omega)|^2 ds$, then there exists some $K_1(\omega)$, such that for all $k > K_1(\omega)$,

$$\sum_{i=1}^{N_1}\int_0^T |Z_s^{n_k(i)}(\omega)|^2 ds \geq \sum_{i=1}^{N_1}\int_0^T |Z_s^{(i)}(\omega)|^2 ds - \epsilon, \tag{3.31}$$

and

$$|\sum_{i=1}^{N^{(n_k)}-1}\int_0^T |Z_s^{n_k(i)}(\omega)|^2 ds - M(\omega)| < \epsilon. \tag{3.32}$$

From (3.29), it follows from Chebyshev's inequality that

$$P(\sum_{i>N_1}\int_0^T |Z^{(i)}|^2 ds \geq \epsilon) < \epsilon.$$

Let O_ϵ denote the collection of all the ωs such that

$$\sum_{i>N_1}\int_0^T |Z^{(i)}|^2 ds < \epsilon. \tag{3.33}$$

Combining (3.31), (3.32), and (3.33), we have for almost every ω in O_ϵ and for all $k > K(\epsilon)$,

$$\begin{aligned} M - 2\epsilon &\leq \sum_{i=1}^{N_1}\int_0^T |Z^{(i)}|^2 ds - \epsilon \\ &\leq \sum_{i=1}^{N_1}\int_0^T |Z^{n_k(i)}(\omega)|^2 ds \\ &\leq \sum_{i=1}^{N^{(n_k)}-1}\int_0^T |Z_s^{n_k(i)}(\omega)|^2 ds \\ &\leq M + \epsilon, \end{aligned} \tag{3.34}$$

which yields

$$\sum_{i>N_1}\int_0^T |Z^{n_k(i)}(\omega)|^2 ds \leq 3\epsilon, \tag{3.35}$$

and

$$|\sum_{i=1}^{N_1}\int_0^T |Z^{n_k(i)}(\omega)|^2 ds - \sum_{i=1}^{N_1}\int_0^T |Z^{(i)}(\omega)|^2 ds| < 2\epsilon. \tag{3.36}$$

For almost every ω in O_ϵ, and sufficiently large k,

$$
\begin{aligned}
&\int_0^T \|\widetilde{Z_t^{n_k}(\omega)} - Z_t(\omega)\|^2 dt \\
&= \sum_{i=1}^{\infty} (Z_t^{n_k(i)}(\omega) - Z_t^{(i)}(\omega))^2 dt \\
&= \sum_{i=1}^{N_1} \int_0^T (Z_t^{n_k(i)}(\omega) - Z_t^{(i)}(\omega))^2 dt + \sum_{i>N_1} \int_0^T (Z_t^{n_k(i)}(\omega) - Z_t^{(i)}(\omega))^2 dt \\
&\leq \sum_{i=1}^{N_1} \int_0^T (Z_t^{n_k(i)}(\omega) - Z_t^{(i)}(\omega))^2 dt + 2\sum_{i>N_1} [\int_0^T (|Z_t^{n_k(i)}(\omega)|^2 + |Z_t^{(i)}(\omega)|^2) dt] \\
&\leq \sum_{i=1}^{N_1} \int_0^T (Z_t^{n_k(i)}(\omega) - Z_t^{(i)}(\omega))^2 dt + 8\epsilon \quad \text{(by (3.33) and (3.35))} \\
&= 8\epsilon + \sum_{i=1}^{N_1} \int_0^T (|Z_t^{n_k(i)}(\omega)|^2 - |Z_t^{(i)}(\omega)|^2) dt + \sum_{i=1}^{N_1} \int_0^T 2Z_t^{(i)}(Z_t^{(i)} - Z_t^{n_k(i)})(\omega) dt \\
&\leq 10\epsilon + \sum_{i=1}^{N_1} \int_0^T 2Z_t^{(i)}(\omega)(Z_t^{(i)}(\omega) - Z_t^{n_k(i)}(\omega)) dt. \quad \text{(by (3.36))} \qquad (3.37)
\end{aligned}
$$

Since $Z^{n_k(i)}(\omega)$ converges weakly to $Z^{(i)}(\omega)$ in $L^2([0,T])$ for $1 \leq i \leq N_1$, by definition of weak convergence, there exists $K_2(\omega)$, such that for all $k > K_2(\omega)$,

$$
\left| \sum_{i=1}^{N_1} \int_0^T 2Z_t^{(i)}(\omega)(Z_t^{(i)}(\omega) - Z_t^{n_k(i)}(\omega)) dt \right| < \epsilon.
$$

Together with (3.37), we get for almost every $\omega \in O_\epsilon$, for all $k > \max(K_1(\omega), K_2(\omega))$,

$$
\int_0^T \|\widetilde{Z_t^{n_k}(\omega)} - Z_t(\omega)\|^2 dt \leq 11\epsilon.
$$

Define for $k \geq 1$

$$
A_k = \bigcup_{m \geq k} \{\omega \in O_\epsilon : \int_0^T \|\widetilde{Z_t^{n_m}(\omega)} - Z_t(\omega)\|^2 dt > 12\epsilon\}.
$$

This is a sequence of decreasing sets satisfying $A_k \supseteq A_{k+1}$, for $k \geq 1$. Let A_∞ denote $\bigcap_{k \leq 1} A_k$. We have

$$
P(A_\infty) = \lim_{k \to \infty} P(A_k).
$$

For all $\omega \in A_\infty$, we have that for any $k > 0$, there exists $m > k$, such that $\int_0^T \|\widetilde{Z_t^{n_m}(\omega)} - Z_t(\omega)\|^2 dt > 12\epsilon$, which implies that A_∞ should have probability 0. As a result, there exists a constant K, such that for all $k > K$, $P(A_k) < \epsilon$. It follows that

$$\begin{aligned} P(\{\omega : \int_0^T \|\widetilde{Z_t^{n_k}(\omega)} - Z_t(\omega)\|^2 dt > 12\epsilon\}) &\leq P(A_k) + P(\Omega \setminus O_\epsilon) \\ &\leq 2\epsilon. \end{aligned}$$

Since ϵ is arbitrary, by the axiom of choice, in all, there exists a sequence of Z_t^n and Z_t that satisfies (3.26) and (3.27), such that

$$\int_0^T \|\widetilde{Z_t^n} - Z_t\|^2 dt \to 0 \quad \text{in probability.}$$

The $L^{1+\delta}(\Omega \times [0,T])$ convergence comes immediately from the L^2 boundedness of $(\int_0^T \|\widetilde{Z_t^n}\| dt)^{\frac{1}{2}}$, which could be derived in the same way as (3.28). □

3.4.4 Proof of Theorem 3.4.1

We follow the same method used in [6]. The main idea is to decompose the differences between solutions into three parts,

$$Y^n - Y = (Y^n - Y^{n,p}) + (Y^{n,p} - Y^{\infty,p}) + (Y^{\infty,p} - Y), \tag{3.38}$$

and

$$\widetilde{Z^n} - Z = (\widetilde{Z^n} - \widetilde{Z^{n,p}}) + (\widetilde{Z^{n,p}} - Z^{\infty,p}) + (Z^{\infty,p} - Z). \tag{3.39}$$

Here ∞ stands for the Levy case and p denotes the approximation of solutions to (3.18) and (3.19) through the Picard iteration, that is

$$Y_t^{n,p+1} = \xi^n + \int_t^T F^n(s, Y_{s-}^{n,p}, Z_s^{n,p}) ds - \int_t^T (Z_s^{n,p+1})^T dM_s, \tag{3.40}$$

and

$$Y_t^{\infty,p+1} = \xi + \int_t^T f(s, Y_{s-}^{\infty,p}, Z_s^{\infty,p})ds - \sum_{i=1}^{\infty} \int_t^T Z_s^{(i)\infty,p+1} dH_s^{(i)}. \tag{3.41}$$

The transformation from $Z^n \in \mathbb{R}^{N^{(n)}}$ to $\widetilde{Z^n} \in \mathbb{R}^\infty$ is defined by

$$\widetilde{Z^n}(i) =< Z^n, C^{n(i)} >_{X_{t-}}, \quad \text{for } 1 \leq i \leq N^{(n)} - 1,$$

and

$$\widetilde{Z^n}(i) = 0 \quad , \text{for } i \geq N^{(n)},$$

such that

$$\int_0^t (Z_s^n)^T dM_s = \sum_{i=1}^{\infty} \int_0^t \widetilde{Z_s^n}(i) dH_s^{n(i)},$$

where $N^{(n)}$ is the number of states in $X^{(n)}$. Moreover, we have

$$\|Z_s^n\|_{X_{t-}}^2 = \sum_{i=1}^{\infty} |\widetilde{Z_s^n}(i)|^2.$$

For the Picard iteration, we set $Y^{\infty,0} = 0$, $Z^{\infty,0} = 0$, $Y^{n,0} = 0$, and $Z^{n,0} = 0$. Let's first consider $Y^{\infty,p} - Y$ and $Z^{\infty,p} - Z$. Define

$$\|(Y^{\infty,p} - Y, Z^{\infty,p} - Z)\|_\beta^2 = E[\int_0^T e^{\beta s}(|Y^{\infty,p} - Y|^2 + \sum_{i=1}^{\infty} |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2)ds].$$

From the proof of Theorem 1 in [43], we have for some $\beta \geq 1$,

$$\|(Y^{\infty,p} - Y, Z^{\infty,p} - Z)\|_\beta^2 \to 0, \quad \text{as } p \to \infty. \tag{3.42}$$

In order to get convergence in the sense of (3.23), we need the following lemma.

Lemma 3.4.8.

$$E\left[\sup_{0 \leq t \leq T} |Y^{\infty,p} - Y| + \sum_{i=1}^{\infty} \int_0^T |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2 ds\right] \to 0, \quad asp \to \infty. \tag{3.43}$$

Proof. It can be easily observed from (3.42) and $\beta \geq 1$ that

$$E\left(\sum_{i=1}^{\infty}\int_0^T |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2 ds\right) \to 0, \quad asp \to \infty. \tag{3.44}$$

What remains is to prove the convergence of $E(\sup_{0\leq t\leq T}|Y^{\infty,p} - Y|)$ to 0. For any $p \geq 0$, we have

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

and

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

Hence

$$\begin{aligned}
&Y_t^{\infty,p+1} - Y_t \\
&= \int_t^T f(s, Y_{s-}^{\infty,p}, Z_s^{\infty,p}) - f(s, Y_{s-}, Z_s)ds - \sum_{i=1}^{\infty}\int_t^T (Z_s^{(i)\infty,p+1} - Z_s^{(i)})dH_s^{(i)} \\
&= \int_t^T f(s, Y_{s-}^{\infty,p}, Z_s^{\infty,p}) - f(s, Y_{s-}, Z_s)ds - \sum_{i=1}^{\infty}\int_0^T (Z_s^{(i)\infty,p+1} - Z_s^{(i)})dH_s^{(i)} \\
&\quad + \sum_{i=1}^{\infty}\int_0^t (Z_s^{(i)\infty,p+1} - Z_s^{(i)})dH_s^{(i)}.
\end{aligned}$$

Take absolute values on both sides of the above equation and use condition 6 to get

$$\begin{aligned}
&|Y_t^{\infty,p+1} - Y_t| \\
&\leq \int_t^T |f(s, Y_{s-}^{\infty,p}, Z_s^{\infty,p}) - f(s, Y_{s-}, Z_s)|ds + \left|\sum_{i=1}^{\infty}\int_0^T (Z_s^{(i)\infty,p+1} - Z_s^{(i)})dH_s^{(i)}\right| \\
&\quad + |\sum_{i=1}^{\infty}\int_0^t (Z_s^{(i)\infty,p+1} - Z_s^{(i)})dH_s^{(i)}| \\
&\leq \int_0^T K\left(|Y_{s-}^{\infty,p} - Y_{s-}| + \left[\sum_{i=1}^{\infty}|Z_s^{(i)\infty,p} - Z_s^{(i)}|^2\right]^{1/2}\right) ds \\
&\quad + 2\sup_{0\leq t\leq T}\left|\int_0^t (Z_s^{(i)\infty,p+1} - Z_s^{(i)})dH_s^{(i)}\right|,
\end{aligned}$$

where K denotes the Lipschitz constant of f. It follows that

$$\sup_{0\le t\le T} |Y_t^{\infty,p+1} - Y_t| \le \int_0^T K\left(|Y_{s-}^{\infty,p} - Y_{s-}| + \left[\sum_{i=1}^{\infty} |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2\right]^{1/2}\right) ds$$
$$+ 2 \sup_{0\le t\le T}\left|\int_0^t (Z_s^{(i)\infty,p+1} - Z_s^{(i)}) dH_s^{(i)}\right|. \tag{3.45}$$

By Hölder's inequality, we have

$$\int_0^T \left(|Y_{s-}^{\infty,p} - Y_{s-}| + \left[\sum_{i=1}^{\infty} |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2\right]^{1/2}\right) ds$$
$$\le \left[T\int_0^T \left(|Y_{s-}^{\infty,p} - Y_{s-}| + \left[\sum_{i=1}^{\infty} |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2\right]^{1/2}\right)^2 ds\right]^{\frac{1}{2}}$$
$$\le \sqrt{2T}\left[\int_0^T |Y_{s-}^{\infty,p} - Y_{s-}|^2 + \sum_{i=1}^{\infty} |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2 ds\right]^{\frac{1}{2}}.$$

Take expectation on both sides and use Jensen's inequality to get

$$E\left[\int_0^T \left(|Y_{s-}^{\infty,p} - Y_{s-}| + \left[\sum_{i=1}^{\infty} |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2\right]^{1/2}\right) ds\right]$$
$$\le \sqrt{2T}\left(E\left[\int_0^T |Y_{s-}^{\infty,p} - Y_{s-}|^2 + \sum_{i=1}^{\infty} |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2 ds\right]\right)^{\frac{1}{2}}. \tag{3.46}$$

By (3.42) and the fact that $\beta \ge 1$, the right hand side of the above inequality should converge to 0 as $p \to \infty$. Applying Burkholder-Davis-Gundy inequality and

Jensen's inequality to $\sup_{0\le t\le T} |\int_0^t (Z_s^{(i)\infty,p+1} - Z_s^{(i)}) dH_s^{(i)}|$, we obtain

$$E\left(2\sup_{0\le t\le T}\left|\int_0^t (Z_s^{(i)\infty,p+1} - Z_s^{(i)})dH_s^{(i)}\right|\right) \tag{3.47}$$

$$\le 2E\left(\sup_{0\le t\le T}\left|\int_0^t (Z_s^{(i)\infty,p+1} - Z_s^{(i)})dH_s^{(i)}\right|^2\right)^{\frac{1}{2}} \tag{3.48}$$

$$\le 4E\left(\int_0^T (Z_s^{(i)\infty,p+1} - Z_s^{(i)})^2 ds\right)^{\frac{1}{2}}. \tag{3.49}$$

By (3.44), the right hand side of 3.47 converges to 0 as $p \to \infty$. Taking expectation to (3.45), substitute in (3.46) and (3.47) to get

$$E(\sup_{0\le t\le T} |Y_t^{\infty,p} - Y_t|) \to 0 \quad \text{as } p \to \infty. \tag{3.50}$$

(3.50) together with (3.44) implies (3.43). $\square$

Next, consider $Y^n - Y^{n,p}$ and $\widetilde{Z^n} - \widetilde{Z^{n,p}}$. By the proof of Theorem 6.2 in [10], we have for some constants C_1 and C_2 that only depend on T and K,

$$\int_0^T E(|Y_s^{n,p+1} - Y_s^{n,p}|^2)ds \le \frac{(C_1)^p}{p!}\int_0^T E(|Y_s^{n,1} - Y_s^{n,0}|^2)ds, \tag{3.51}$$

and

$$\int_0^T E(|\widetilde{Z_s^{n,p+1}} - \widetilde{Z_s^{n,p}}|^2)ds \le C_2 \int_0^T E(|Y_s^{n,p+1} - Y_s^{n,p}|^2) + E(|Y_s^{n,p} - Y_s^{n,p-1}|^2)ds. \tag{3.52}$$

Regarding $Y^{n,p} - Y^n$ and $Z^{n,p} - Z^n$, we have the following lemma.

Lemma 3.4.9.

$$\sup_n E\left[\sup_{0\le t\le T} |Y^{n,p} - Y^n| + \sum_{i=1}^{\infty}\int_0^T |\widetilde{Z_s^{(i)n,p}} - \widetilde{Z_s^{(i)n}}|^2 ds\right] \to 0, \quad asp \to \infty. \tag{3.53}$$

Proof. By Stirling's formula, we could write (3.51) as

$$\int_0^T E(|Y_s^{n,p+1} - Y_s^{n,p}|^2)ds \leq \frac{(C_1)^p}{(p/e)^p} \int_0^T E(|Y_s^{n,1} - Y_s^{n,0}|^2)ds.$$

Thus for $p > 4C_1 e$,

$$\left[\int_0^T E(|Y_s^{n,p+1} - Y_s^{n,p}|^2)ds\right]^{\frac{1}{2}} \leq \left(\frac{1}{2}\right)^p \left[\int_0^T E(|Y_s^{n,1} - Y_s^{n,0}|^2)ds\right]^{\frac{1}{2}}.$$

As a result, we have for $p > 4C_1 e$,

$$\left[\int_0^T E(|Y_s^{n,p} - Y_s^{n}|^2)ds\right]^{\frac{1}{2}} \leq \left(\frac{1}{2}\right)^{p-1} \left[\int_0^T E(|Y_s^{n,1} - Y_s^{n,0}|^2)ds\right]^{\frac{1}{2}}. \tag{3.54}$$

Substitute (3.51) into (3.52) and use Stirling's formula to obtain

$$\int_0^T E(|\widetilde{Z_s^{n,p+1}} - \widetilde{Z_s^{n,p}}|^2)ds \leq 5C_2 \left(\frac{1}{4}\right)^p \int_0^T E(|Y_s^{n,1} - Y_s^{n,0}|^2)ds,$$

for $p > 4C_1 e$. This implies

$$\left[\int_0^T E(|\widetilde{Z_s^{n,p}} - \widetilde{Z_s^{n}}|^2)ds\right]^{\frac{1}{2}} \leq \sqrt{5C_2} \left(\frac{1}{2}\right)^{p-1} \left[\int_0^T E(|Y_s^{n,1} - Y_s^{n,0}|^2)ds\right]^{\frac{1}{2}}. \tag{3.55}$$

For any $p \geq 0$, we have

$$Y_t^{n,p+1} = \xi^n + \int_t^T F^n(s, Y_{s-}^{n,p}, Z_s^{n,p})ds - \int_t^T (Z_s^{n,p+1})^T dM_s,$$

and

$$Y_t^n = \xi^n + \int_t^T F^n(s, Y_{s-}^n, Z_s^n) - \int_t^T (Z_s^n)^T dM_s.$$

Thus

$$\begin{aligned} Y_t^{n,p+1} - Y_t^n = &\int_t^T F^n(s, Y_{s-}^{n,p}, Z_s^{n,p}) - F^n(s, Y_{s-}^n, Z_s^n)ds \\ &- \int_0^T (Z_s^{n,p+1} - Z_s^n)^T dM_s + \int_0^t (Z_s^{n,p+1} - Z_s^n)^T dM_s. \end{aligned}$$

Similar to the proof in Lemma 3.4.8, we can obtain

$$\begin{aligned}
&\sup_{0\le t\le T}|Y_t^{n,p+1}-Y_t^n| \\
&\le K\int_0^T |Y_{s-}^{n,p}-Y_{s-}^n| + \|Z_s^{n,p}-Z_s^n\|_{X_{s-}}ds + 2\sup_{0\le t\le T}\left|\int_0^t (Z_s^{n,p+1}-Z_s^n)^T dM_s\right|.
\end{aligned} \tag{3.56}$$

By Hölder's inequality,

$$\int_0^T |Y_{s-}^{n,p}-Y_{s-}^n| + \|Z_s^{n,p}-Z_s^n\|_{X_{s-}}ds \le \sqrt{2T\int_0^T |Y_{s-}^{n,p}-Y_{s-}^n|^2 + \|Z_s^{n,p}-Z_s^n\|^2_{X_{s-}}ds}.$$

It follows from Jensen's inequality that

$$E[\int_0^T |Y_{s-}^{n,p}-Y_{s-}^n|+\|Z_s^{n,p}-Z_s^n\|_{X_{s-}}ds] \le [2TE(\int_0^T |Y_{s-}^{n,p}-Y_{s-}^n|^2+\|Z_s^{n,p}-Z_s^n\|^2_{X_{s-}}ds)]^{\frac{1}{2}}.$$

Substituting (3.54) and (3.55) into the above inequality yields

$$\begin{aligned}
&E\left[\int_0^T |Y_{s-}^{n,p}-Y_{s-}^n| + \|Z_s^{n,p}-Z_s^n\|_{X_{s-}}ds\right] \\
&\le \sqrt{2T(5C_2+1)}\left(\frac{1}{2}\right)^{p-1}\left[\int_0^T E(|Y_s^{n,1}-Y_s^{n,0}|^2)ds\right]^{\frac{1}{2}}.
\end{aligned} \tag{3.57}$$

Next, apply Burkholder-Davis-Gundy inequality and Jensen's inequality to

$$\sup_{0\le t\le T}\left|\int_0^t (Z_s^{n,p+1}-Z_s^n)^T dM_s\right|$$

to get

$$\begin{aligned}
E\left(\sup_{0\le t\le T}\left|\int_0^t (Z_s^{n,p+1}-Z_s^n)^T dM_s\right|\right) &\le E\left[\left(\sup_{0\le t\le T}\left|\int_0^t (Z_s^{n,p+1}-Z_s^n)^T dM_s\right|\right)^2\right]^{\frac{1}{2}} \\
&\le 2E\left(\int_0^T \|Z_s^{n,p+1}-Z_s^n\|^2_{X_{s-}}ds\right)^{\frac{1}{2}} \\
&= 2\left[\int_0^T E(|\widetilde{Z_s^{n,p+1}}-\widetilde{Z_s^n}|^2)ds\right]^{\frac{1}{2}} \\
&\le \sqrt{5C_2}\left(\frac{1}{2}\right)^{p-1}\left[\int_0^T E(|Y_s^{n,1}-Y_s^{n,0}|^2)ds\right]^{\frac{1}{2}}.
\end{aligned} \tag{3.58}$$

Taking expectation on both sides of (3.56), and substituting (3.57) and (3.58) into the inequality yields

$$E\left(\sup_{0\le t\le T}|Y_t^{n,p+1}-Y_t^n|\right)$$
$$\le (K\sqrt{2T(5C_2+1)}+2\sqrt{5C_2})\left(\frac{1}{2}\right)^{p-1}\left[\int_0^T E(|Y_s^{n,1}-Y_s^{n,0}|^2)ds\right]^{\frac{1}{2}}.$$

In order to get convergence in the sense of (3.53), what remains is to check

$$\sup_n\left[\int_0^T E(|Y_s^{n,1}-Y_s^{n,0}|^2)ds\right]<\infty. \tag{3.59}$$

Apply Itô's formula to $(Y_s^{n,1})^2$ yields

$$|Y_t^{n,1}|^2=|\xi^n|^2-2\int_t^T Y_s^{n,1}(-F^n(s,0,0)ds+Z_s^{n,1}dM_s)-\int_t^T (Z_s^{n,1})^T d[M,M]_s(Z_s^{n,1}).$$

Take expectation and then absolute values on both sides gives

$$\begin{aligned} E(|Y_t^{n,1}|^2) &\le E(|\xi^n|^2)+2E\int_t^T |Y_s^{n,1}F^n(s,0,0)|ds-E\int_t^T \|Z_s^{n,1}\|^2_{X_{s-}}ds\\ &\le E(|\xi^n|^2)+\int_t^T E(|Y_s^{n,1}|^2)ds+\int_t^T |F^n(s,0,0)|^2ds. \end{aligned}$$

Move the second term on the right to the left and multiply both sides by e^t to get

$$\begin{aligned} e^tE(|Y_t^{n,1}|^2)-e^t\int_t^T E(|Y_s^{n,1}|^2)ds &\le e^t\left[E(|\xi^n|^2)+\int_t^T |F^n(s,0,0)|^2ds\right]\\ -(e^t\int_t^T E(|Y_s^{n,1}|^2)ds)' &\le e^t\left[E(|\xi^n|^2)+\int_t^T |F^n(s,0,0)|^2ds\right]. \end{aligned}$$

Integrating the above inequality from 0 to T yields

$$\begin{aligned} \int_0^T E(|Y_s^{n,1}|^2)ds &\le \int_0^T e^t\left[E(|\xi^n|^2)+\int_t^T |F^n(s,0,0)|^2ds\right]dt\\ &\le \int_0^T e^T\left[E(|\xi^n|^2)+\int_t^T |F^n(s,0,0)|^2ds\right]dt\\ &\le Te^T\left[E(|\xi^n|^2)+\int_0^T |F^n(s,0,0)|^2ds\right]\\ &\le Te^T\left[E(|\xi^n|^2)+\int_0^T |f(s,0,0)|^2ds\right]. \end{aligned} \tag{3.60}$$

It follows from conditions 4 and 6 that (3.59) holds. In all, we have convergence in the sense of (3.53). $\square$

Given Lemma 3.4.8 and Lemma 3.4.9, it remains to show the convergence to zero of $Y^{n,p}-Y^{\infty,p}$ and $\widetilde{Z^{n,p}}-Z^{\infty,p}$. We will show the following lemma by induction.

Lemma 3.4.10. *For any $p \geq 0$, we have*

$$\begin{aligned} & d_{J_1}(Y^{n,p}, Y^{\infty,p}) + \int_0^T |Y_s^{n,p} - Y_s^{\infty,p}| ds + \int_0^T \|\widetilde{Z_s^{n,p}} - Z_s^{\infty,p}\|^2 ds \\ & \to 0 \quad \text{in probability as } n \to \infty, \end{aligned} \tag{3.61}$$

as well as

$$\sup_n \{E[\int_0^T (|Y^{n,p}|^3 + |Y^{\infty,p}|^3) dt] + E[(\int_0^T \|\widetilde{Z_s^{n,p}}\|^2 ds)^{\frac{3}{2}}] + E[(\int_0^T \|Z_s^{\infty,p}\|^2 ds)^{\frac{3}{2}}]\} < \infty. \tag{3.62}$$

Proof. The case when $p = 0$ is trivial, since we have both expressions in (3.61) and (3.62) equal to zero ($Y^{n,0} = Y^{\infty,0} = 0$ and $\widetilde{Z_s^{n,0}} = Z_s^{\infty,0} = 0$).

Suppose (3.61) and (3.62) hold for all $p \leq k$. When $p = k+1$, by definition, we have

$$Y_t^{n,k+1} = \xi^n + \int_t^T F^n(s, Y_{s-}^{n,k}, Z_s^{n,k}) ds - \int_t^T (Z_s^{n,k+1})^T dM_s,$$

and

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

Define processes

$$N_t^n = Y_t^{n,k+1} + \int_0^t F^n(s, Y_{s-}^{n,k}, Z_s^{n,k}) ds, \quad \text{and} \quad N_t = Y_t^{\infty,k+1} + \int_0^t f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k}) ds.$$

The process N_t^n satisfies

$$\begin{aligned}
N_t^n &= \xi^n + \int_0^T F^n(s, Y_{s-}^{n,k}, Z_s^{n,k})ds - \int_t^T (Z_s^{n,k+1})^T dM_s \\
&= \xi^n + \int_0^T F^n(s, Y_{s-}^{n,k}, Z_s^{n,k})ds - \int_0^T (Z_s^{n,k+1})^T dM_s + \int_0^t (Z_s^{n,k+1})^T dM_s \\
&= Y_0^{n,k+1} + \int_0^t (Z_s^{n,k+1})^T dM_s \\
&= N_0^n + \int_0^t (Z_s^{n,k+1})^T dM_s \\
&= N_0^n + \sum_{i=1}^{\infty} \int_0^t \widetilde{Z_s^{n,k+1}}(i) dH_s^{n(i)}. \qquad (3.63)
\end{aligned}$$

Hence it is an $\mathcal{F}_\cdot^n$-martingale, with

$$\begin{aligned}
N_t^n &= E(N_T^n | \mathcal{F}_t^n) \\
&= E\left(\xi^n + \int_0^T F^n(s, Y_{s-}^{n,k}, Z_s^{n,k})ds | \mathcal{F}_t^n\right).
\end{aligned}$$

Similarly, N_t is an $\mathcal{F}_\cdot$-martingale, with

$$\begin{aligned}
N_t &= N_0 + \sum_{i=1}^{\infty} \int_0^t (Z_s^{(i)\infty,k+1}) dH_s^{(i)} \qquad (3.64) \\
&= E\left(\xi + \int_0^T f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})ds | \mathcal{F}_t\right).
\end{aligned}$$

At the terminal time, we have

$$\begin{aligned}
|N_T^n - N_T| &= \left|\xi^n - \xi + \int_0^T F^n(s, Y_{s-}^{n,k}, Z_s^{n,k})ds - \int_0^T f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})ds\right| \\
&\leq |\xi^n - \xi| + \int_0^T \left| f(s, Y_{s-}^{n,k}, \widetilde{Z_s^{n,k}}) - f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})\right| ds \\
&\leq |\xi^n - \xi| + K\left(\int_0^T |Y_{s-}^{n,k} - Y_{s-}^{\infty,k}|ds + \int_0^T \|\widetilde{Z_s^{n,k}} - Z_s^{\infty,k}\|ds\right) \\
&\leq |\xi^n - \xi| + K\int_0^T |Y_{s-}^{n,k} - Y_{s-}^{\infty,k}|ds + K\sqrt{T\int_0^T \|\widetilde{Z_s^{n,k}} - Z_s^{\infty,k}\|^2 ds},
\end{aligned}$$

which converges to zero in probability by condition 3 and the induction hypothesis (3.61).

In order to apply Lemma 3.4.7 to N_T^n and N_T, we need to show they satisfy condition 3 and condition 4. Since $|N_T^n - N_T|$ converges to zero in probability, L^2 convergence (condition 3) would immediately follow from L^3 boundedness (condition 4), thus we only need to check condition 4.

$$
\begin{aligned}
&|N_T^n|^3 + |N_T|^3 \\
&= \left|\xi^n + \int_0^T f(s, Y_{s-}^{n,k}, \widetilde{Z_s^{n,k}})ds|^3 + |\xi + \int_0^T f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})ds\right|^3 \\
&\leq 4\left[|\xi^n|^3 + |\xi|^3 + \left(\int_0^T |f(s, Y_{s-}^{n,k}, \widetilde{Z_s^{n,k}})|ds\right)^3 + \left(\int_0^T |f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})|ds\right)^3\right] \\
&\leq C\left[|\xi^n|^3 + |\xi|^3 + \left(\int_0^T |f(s,0,0)|ds\right)^3 + \left(\int_0^T |Y_{s-}^{n,k}|ds\right)^3 + \left(\int_0^T \|\widetilde{Z_s^{n,k}}\|ds\right)^3\right. \\
&\quad \left. + \left(\int_0^T |Y_{s-}^{\infty,k}|ds\right)^3 + \left(\int_0^T \|Z_s^{\infty,k}\|ds\right)^3\right] \\
&\leq \tilde{C}\left[|\xi^n|^3 + |\xi|^3 + \left(\int_0^T |f(s,0,0)|^2 ds\right)^{\frac{3}{2}} + \left(\int_0^T |Y_{s-}^{n,k}|ds\right)^3 + \left(\int_0^T |Y_{s-}^{\infty,k}|ds\right)^3\right. \\
&\quad \left. + \left(\int_0^T \|\widetilde{Z_s^{n,k}}\|^2 ds\right)^{\frac{3}{2}} + \left(\int_0^T \|Z_s^{\infty,k}\|^2 ds\right)^{\frac{3}{2}}\right],
\end{aligned}
$$

for some constants C and $\tilde{C}$. Take expectations and supremum on both sides of the above inequality, use condition 4, condition 6 and (3.62) to get

$$E(|N_T|^3) + \sup_n E(|N_T^n|^3) < \infty.$$

As a result, N_T^n and N_T satisfy the requirements. Apply Lemma 3.4.7 to N_T^n and N_T to obtain

$$\int_0^T \|\widetilde{Z_s^{n,k+1}} - Z_s^{\infty,k+1}\|^2 ds \to 0 \quad \text{as } n \to \infty.$$

Next, by L^2 convergence of N_T^n to N_T and Lemma 3.4.5, it follows that

$$N_t^n \to N_t \quad \text{in probability in the } J_1 \text{ Skorokhod sense.}$$

Since $\int_0^t f(s, Y_{s-}^{n,k}, \widetilde{Z_s^{n,k}})ds$ and $\int_0^t f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})ds$ are continuous in t, we have

$$
\begin{aligned}
& d_{J_1}(Y_t^{n,k+1}, Y_t^{k+1}) \\
& \leq d_{J_1}(N_t^n, N_t) + \sup_t \left| \int_0^t f(s, Y_{s-}^{n,k}, \widetilde{Z_s^{n,k}})ds - \int_0^t f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})ds \right| \\
& \leq d_{J_1}(N_t^n, N_t) + \int_0^T |f(s, Y_{s-}^{n,k}, \widetilde{Z_s^{n,k}} - f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})|ds \\
& \leq d_{J_1}(N_t^n, N_t) + K\int_0^T |Y_{s-}^{n,k} - Y_{s-}^{\infty,k}| + K\int_0^T \|\widetilde{Z_s^{n,k}} - Z_s^{\infty,k}\|ds \\
& \leq d_{J_1}(N_t^n, N_t) + K\int_0^T |Y_{s-}^{n,k} - Y_{s-}^{\infty,k}| + K\sqrt{T\int_0^T \|\widetilde{Z_s^{n,k}} - Z_s^{\infty,k}\|^2)ds},
\end{aligned}
$$

which goes to zero in probability by the induction hypothesis (3.61).

In order to show (3.61) holds for $k+1$, we still need to verify that $\int_0^T |Y_s^{n,k+1} - Y_s^{\infty,k+1}|ds$ converges to zero in probability. By definition,

$$
\int_0^T |Y_s^{n,k+1} - Y_s^{\infty,k+1}|ds \leq \int_0^T |N_s^n - N_s|ds + \int_0^T |f(s, Y_{s-}^{n,k}, \widetilde{Z_s^{n,k}} - f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})|ds.
$$

The convergence of $\int_0^T |f(s, Y_{s-}^{n,k}, \widetilde{Z_s^{n,k}}) - f(s, Y_{s-}^{\infty,k}, Z_s^{\infty,k})|ds$ to zero in probability comes directly from the Lipschitz continuity of f. It suffices to prove the convergence to zero of the process $\int_0^T |N_s^n - N_s|ds$. Since N_T^n converges to N_T in L^2, we can get from Lemma 3.4.5

$$
N_t^n \to N_t \quad \text{in probability under } J_1 \text{ Skorokhod topology.}
$$

As a result,we can extract a subsequence (still indexed by n), so that for almost every ω, $d_{J_1}(N_t^n(\omega), N_t(\omega))$ converges to zero, which means there exists a sequence of strictly increasing homeomorphisms λ^n on $[0, T]$, such that

$$
\sup_t |\lambda^n(t) - t| \to 0, \text{and} \sup_t |N_t^n(\omega) - N_{\lambda^n(t)}(\omega)| \to 0.
$$

For almost every ω,

$$|N_t^n(\omega) - N_t(\omega)| \leq |N_t^n(\omega) - N_{\lambda^n(t)}(\omega)| + |N_{\lambda^n(t)}(\omega) - N_t(\omega)|.$$

By the fact that $\sup_t |N_t^n(\omega) - N_{\lambda^n(t)}(\omega)|$ converges to zero, we have

$$\int_0^T |N_t^n(\omega) - N_{\lambda^n(t)}(\omega)| dt \to 0.$$

As to the part $|N_{\lambda^n(t)}(\omega) - N_t(\omega)|$, we'll make use of one property about càdlàg functions. Since $N_t(\omega)$ is a càdlàg function, by Lemma 6.10 in [30], for any $\alpha > 0$, let $U_\alpha = \{t \in [0,T] : |N_{t_+}(\omega) - N_{t_-}(\omega)| \geq \alpha\}$, which is a finite set. Then

$$\limsup_{\delta \to 0^+} \left(\sup\{|N_v(\omega) - N_u(\omega)| : 0 \leq u < v \leq T, v - u \leq \delta, (u,v] \cap U_\alpha = \emptyset\}\right) \leq \alpha.$$

First, pick $\tilde{\delta}$, such that when $\delta < \tilde{\delta}$,

$$\sup\{|N_v(\omega) - N_u(\omega)| : 0 \leq u < v \leq T, v - u \leq \delta, (u,v] \cap U_\alpha = \emptyset\} \leq 2\alpha,$$

then pick N, st when $n > N$, $\max_t |\lambda^n(t) - t| < \tilde{\delta}$. Therefore, when $n > N$, we have

$$\int_0^T |N_{\lambda^n(t)}(\omega) - N_t(\omega)| dt \leq 2\alpha T + |U_\alpha| \, M_\alpha \max_t |\lambda^n(t) - t|,$$

where M_α denotes the size of the maximum jump in U_α. Since α is arbitrary and $\max_t |\lambda^n(t) - t|$ goes to zero as n goes to infinity,

$$\int_0^T |N_{\lambda^n(t)}(\omega) - N_t(\omega)| dt \to 0.$$

In all, we have for almost all ω,

$$\int_0^T |N_t^n(\omega) - N_t(\omega)| dt \to 0,$$

which then implies

$$\int_0^T |Y_s^{n,k+1} - Y_s^{\infty,k+1}|ds \to 0 \quad \text{in probability.}$$

As a result, we have proved (3.61) holds for $p = k+1$. We next show that when $p = k+1$, the solutions $(Y^{n,k+1}, \widetilde{Z^{n,k+1}})$ and $(Y^{\infty,k+1}, Z^{\infty,k+1})$ satisfy (3.62). By definition of N_t^n,

$$Y_t^{n,k+1} = N_t^n - \int_0^t f(s, Y_s^{n,k}, \widetilde{Z_s^{n,k}})ds,$$

so $Y_t^{n,k+1}$ satisfies

$$\begin{aligned} Y_t^{n,k+1} &\leq |N_t^n| + \int_0^T |f(s, Y_s^{n,k}, \widetilde{Z_s^{n,k}})|ds \\ &\leq \sup_{0\leq t\leq T} |N_t^n| + K\left(\int_0^T f(s,0,0)ds + \int_0^T |Y_s^{n,k}|ds + \int_0^T \|\widetilde{Z_s^{n,k}}\|ds\right). \end{aligned}$$

As a result,

$$\begin{aligned} &\int_0^T |Y_t^{n,k+1}|^3 dt \\ &\leq 3\left[T(\sup_{0\leq t\leq T} |N_t^n|)^3 + KT\left(\int_0^T f(s,0,0)ds + \int_0^T |Y_s^{n,k}|ds + \int_0^T \|\widetilde{Z_s^{n,k}}\|ds\right)\right]^3 \\ &\leq C\left[|\sup_{0\leq t\leq T} |N_t^n||^3 + \left(\int_0^T f(s,0,0)ds\right)^3 + \left(\int_0^T |Y_s^{n,k}|ds\right)^3 + \left(\int_0^T \|\widetilde{Z_s^{n,k}}\|ds\right)^3\right] \\ &\leq \tilde{C}\left[|\sup_{0\leq t\leq T} |N_t^n||^3 + \left(\int_0^T f^2(s,0,0)ds\right)^{\frac{3}{2}} + \left(\int_0^T |Y_s^{n,k}|^3 ds\right) + \left(\int_0^T \|\widetilde{Z_s^{n,k}}\|^2 ds\right)^{\frac{3}{2}}\right], \end{aligned}$$

for some constants C and $\tilde{C}$. Take expectation on both sides yields

$$\begin{aligned} E\left[\left(\int_0^T |Y_t^{n,k+1}|dt\right)^3\right] \leq \tilde{C}\Bigg\{&E\left(|\sup_{0\leq t\leq T} |N_t^n||^3\right) + \left(\int_0^T f^2(s,0,0)ds\right)^{\frac{3}{2}} \\ &+E\left(\int_0^T |Y_s^{n,k}|^3 ds\right) + E\left[\left(\int_0^T \|\widetilde{Z_s^{n,k}}\|^2 ds\right)^{\frac{3}{2}}\right]\Bigg\}. \end{aligned}$$

By induction hypothesis (3.62), it suffices to show that $E(|\sup_{0\leq t\leq T} |N_t^n||^3)$ is bounded. From Doob's martingale property, $\|\sup_{0\leq t\leq T} |N_t^n|\|_3 \leq \frac{3}{2}\||N_T^n|\|_3$, which is equivalent to

$$E(|\sup_{0\leq t\leq T}|N_t^n||^3)\leq\frac{27}{8}E(|N_T^n|^3). \tag{3.65}$$

However, in the previous proof, we have already shown $E(|N_T|^3)+\sup_n E(|N_T^n|^3)<\infty$. Therefore,

$$\sup_n E\left(|\sup_{0\leq t\leq T}|N_t^n||^3\right)<\infty,$$

which then leads to

$$\sup_n E\left(\int_0^T|Y_t^{n,k+1}|^3dt\right)<\infty.$$

Similarly, we can apply the above arguments to $E(\int_0^T|Y_t^{\infty,k+1}|^3dt)$ and then use the fact that $E(|N_T|^3)+\sup_n E(|N_T^n|^3)<\infty$ to derive

$$E\left(\int_0^T|Y_t^{\infty,k+1}|^3dt\right)<\infty.$$

Combined we have

$$\sup_n E\left[\int_0^T(|Y_t^{n,k+1}|^3+|Y_t^{\infty,k+1}|^3)dt\right]<\infty. \tag{3.66}$$

To get the boundedness of

$$\sup_n\left\{E\left[\left(\int_0^T\|\widetilde{Z_s^{n,k+1}}\|^2ds\right)^{\frac{3}{2}}\right]+E\left[\left(\int_0^T\|Z_s^{\infty,k+1}\|^2ds\right)^{\frac{3}{2}}\right]\right\},$$

we use the Theorem 1 in [47], which implies

$$\|<N^n,N^n>_T^{\frac{1}{2}}\|_q\leq C_q\|\sup_t|N_t^n|\|_q\quad for\quad q\geq 2,$$

with C_q a constant depending only on q. Pick $q=3$. By definition of N^n,

$$<N^n,N^n>_T=\int_0^T\|\widetilde{Z_s^{n,k+1}}\|^2ds,$$

substitute back into the previous inequality and use (3.65) to obtain

$$E\left[\left(\int_0^T \|\widetilde{Z_s^{n,k+1}}\|^2\right)^{\frac{2}{3}}\right] \leq \tilde{C}_3 E(|N_T^n|^3).$$

Following the same arguments, we can get for N_t,

$$E\left[\left(\int_0^T \|Z_s^{\infty,k+1}\|^2 ds\right)^{\frac{3}{2}}\right] \leq \tilde{C}_3 E(|N_T|^3).$$

Since $\tilde{C}_3$ is a constant, we finally obtain

$$\begin{aligned}&\sup_n\{E[(\int_0^T \|\widetilde{Z_s^{n,k+1}}\|^2 ds)^{\frac{3}{2}}] + E[(\int_0^T \|Z_s^{\infty,k+1}\|^2 ds)^{\frac{3}{2}}]\}\\ &\leq \tilde{C}_3[E(|N_T|^3) + \sup_n E(|N_T^n|^3)] < \infty.\end{aligned}$$

Thus (3.62) is shown for $p = k+1$. To summarize, we have proved that when $p = k+1$, (3.61) and (3.62) still hold. In all, the induction hypotheses are satisfied for all $p \geq 0$. □

Let's get back to the proof of Theorem 3.4.1. Recall that $Y^n - Y$ and $\widetilde{Z^n} - Z$ are decomposed as

$$Y^n - Y = (Y^n - Y^{n,p}) + (Y^{n,p} - Y^{\infty,p}) + (Y^{\infty,p} - Y),$$

and

$$\widetilde{Z^n} - Z = (\widetilde{Z^n} - \widetilde{Z^{n,p}}) + (\widetilde{Z^{n,p}} - Z^{\infty,p}) + (Z^{\infty,p} - Z).$$

By Lemma 3.4.8, Lemma 3.4.9 and Lemma 3.4.10, we have

$$\begin{aligned}&\sup_n\left[\sup_{0\leq t\leq T}|Y^{\infty,p} - Y| + \sup_{0\leq t\leq T}|Y^{n,p} - Y^n|\right] + \sum_{i=1}^{\infty}\int_0^T |Z_s^{(i)\infty,p} - Z_s^{(i)}|^2 ds\\ &+ \sup_n\left[\sum_{i=1}^{\infty}\int_0^T |\widetilde{Z_s^{(i)n,p}} - \widetilde{Z_s^{(i)n}}|^2 ds\right]\end{aligned}$$

converges to zero in probability as $p \to \infty$, and

$$d_{J_1}(Y^{n,p}, Y^{\infty,p}) + \int_0^T \|\widetilde{Z_s^{n,p}} - Z_s^{\infty,p}\|^2 ds \to 0 \quad \text{in probability as } n \to \infty.$$

For any $\epsilon > 0$, $\delta > 0$, we can find some value $\tilde{p}$, such that

$$P\left(\sup_n \left[\sup_{0\le t\le T} |Y^{\infty,\tilde{p}} - Y| + \sup_{0\le t\le T} |Y^{n,\tilde{p}} - Y^n|\right] + \sum_{i=1}^{\infty} \int_0^T |Z_s^{(i)\infty,\tilde{p}} - Z_s^{(i)}|^2 ds \right.$$
$$\left. + \sup_n \left[\sum_{i=1}^{\infty} \int_0^T |\widetilde{Z_s^{(i)n,\tilde{p}}} - \widetilde{Z_s^{(i)n}}|^2 ds\right] \ge \frac{\epsilon}{6}\right) < \frac{\delta}{2}.$$

Fix $p = \tilde{p}$. There exists some value N, such that when $n > N$,

$$P(d_{J_1}(Y^{n,\tilde{p}}, Y^{\infty,\tilde{p}}) + \int_0^T \|\widetilde{Z_s^{n,\tilde{p}}} - Z_s^{\infty,\tilde{p}}\|^2 ds \ge \frac{\epsilon}{6}) < \frac{\delta}{2}.$$

From decomposition of $Y^n - Y$ and $\widetilde{Z^n} - Z$, we have the following inequalities:

$$\begin{aligned} d_{J_1}(Y^n, Y) &\le d_{J_1}(Y^n, Y^{n,\tilde{p}}) + d_{J_1}(Y^{n,\tilde{p}}, Y^{\infty,\tilde{p}}) + d_{J_1}(Y^{\infty,\tilde{p}} - Y) \\ &\le \sup_{0\le t\le T} |Y^{n,\tilde{p}} - Y^n| + d_{J_1}(Y^{n,\tilde{p}}, Y^{\infty,\tilde{p}}) + \sup_{0\le t\le T} |Y^{\infty,\tilde{p}} - Y|, \end{aligned}$$

and

$$\begin{aligned} \int_0^T \|\widetilde{Z^n} - Z\|^2 ds &\le (\|\widetilde{Z^n} - \widetilde{Z^{n,\tilde{p}}}\| + \|\widetilde{Z^{n,\tilde{p}}} - Z^{\infty,\tilde{p}}\| + \|Z^{\infty,\tilde{p}} - Z\|)^2 ds \\ &\le 3\int_0^T \|\widetilde{Z^n} - \widetilde{Z^{n,\tilde{p}}}\|^2 + \|\widetilde{Z^{n,\tilde{p}}} - Z^{\infty,\tilde{p}}\|^2 + \|Z^{\infty,\tilde{p}} - Z\|^2 ds. \end{aligned}$$

Let $R_{\tilde{p}}$ and $S_{\tilde{p}}$ denote

$$\sup_n \left[\sup_{0\le t\le T} |Y^{\infty,\tilde{p}} - Y| + \sup_{0\le t\le T} |Y^{n,\tilde{p}} - Y^n|\right] + \sum_{i=1}^{\infty} \int_0^T |Z_s^{(i)\infty,\tilde{p}} - Z_s^{(i)}|^2 ds$$
$$+ \sup_n \left[\sum_{i=1}^{\infty} \int_0^T |\widetilde{Z_s^{(i)n,\tilde{p}}} - \widetilde{Z_s^{(i)n}}|^2 ds\right],$$

and

$$d_{J_1}(Y^{n,\tilde{p}}, Y^{\infty,\tilde{p}}) + \int_0^T \|\widetilde{Z_s^{n,\tilde{p}}} - Z_s^{\infty,\tilde{p}}\|^2 ds,$$

respectively. Therefore, for all $n > N$,

$$\begin{aligned}
&P(d_{J_1}(Y^n, Y) + \int_0^T \|\widetilde{Z^n} - Z\|^2 ds \geq \epsilon) \\
&\leq P(3R_{\tilde{p}} + 3S_{\tilde{p}} \geq \epsilon) \\
&\leq P(R_{\tilde{p}} + S_{\tilde{p}} \geq \frac{\epsilon}{3}) \\
&\leq P(R_{\tilde{p}} \geq \frac{\epsilon}{6}) + P(S_{\tilde{p}} \geq \frac{\epsilon}{6}) \\
&< \frac{\delta}{2} + \frac{\delta}{2} \\
&< \delta,
\end{aligned}$$

which finishs the proof of Theorem 3.4.1.

Chapter 4

Nonlinear $\mathcal{G}$-Expectations

4.1 Introduction to $\mathcal{G}$-Expectations

The concept of $\mathcal{G}$-expectation is first introduced in [46] by Peng, where it is used to create a new process termed $\mathcal{G}$-Brownian motion. A $\mathcal{G}$-expectation is defined as a unique viscosity solution to a PDE of the form

$$u_t = \mathcal{G}(u), \tag{4.1}$$

with boundary condition

$$u(0, x) = \phi(x). \tag{4.2}$$

The $\mathcal{G}$ in (4.1) is usually picked to be a nonlinear operator. In [46], $\mathcal{G}$ is given by

$$\mathcal{G}(a) = \frac{1}{2}(a^+ - \sigma_0 a^-),$$

where $a^+ = max(a, 0)$, $a^- = max(-a, 0)$, and $0 \leq \sigma_0 \leq 1$. Consider the $\mathcal{G}$-expectation induced by the following PDE,

$$u_t = \mathcal{G}(u_{xx}), \tag{4.3}$$

with boundary condition (4.2). Peng defines the $\mathcal{G}$-Brownian motion as the process X that satisfies

$$u(t, x) = E^{\mathcal{G}}(\phi(X_t))|X_0 = x),$$

for all $t \geq 0$.

Suppose Y is a Brownian motion, and let us define

$$v(t,x) = E(\phi(x+Y_t)),$$

then v satisfies

$$v_t = \frac{1}{2}v_{xx}. \tag{4.4}$$

Comparing (4.4) and (4.3), it could be observed that the nonlinearity of $\mathcal{G}$-expectation comes from the nonlinearity of the operator $\mathcal{G}$.

4.2 Bid and Ask Prices as $\mathcal{G}$-Expectations

Peng's $\mathcal{G}$-expectation provides us a candidate for the continuous time modeling of two price economy. In this section, we will first describe the basic set up of the two price economy, its underlying driving force and the associated infinitesimal generator. Following Peng's approach in [46], we would then add nonlinearity to the infinitesimal generator using distortion. The bid and ask prices are finally modeled as viscosity solutions to the resulting distorted PIDEs.

4.2.1 The Underlying Uncertainty

Consider a two price economy whose randomness is generated by a pure jump Lévy process $(X_t)_{0\leq t\leq T}$ with jump density k. Moreover, we assume

$$\int_{-\infty}^{\infty} |y|k(y)dy < \infty,$$

and

$$\int_{-\infty}^{\infty} y^2 k(y)dy < \infty.$$

The infinitesimal generator $\mathcal{L}$ of X is given by

$$\mathcal{L}(u) = \int_{-\infty}^{\infty} (u(t, x+y) - u(t, x))k(y)dy. \tag{4.5}$$

Let $u(t, x)$ be the expected value of a claim that pays $\phi(X_t)$ at time t, given $X_0 = x$. Suppose the interest rate r is a constant, a formal definition of $u(t, x)$ is

$$u(t, x) = E(e^{-rt}\phi(X_t)|X_0 = x).$$

The function $u(t, x)$ is the solution to the following partial integro-differential equation (PIDE),

$$u_t = \mathcal{L}(u) - ru, \tag{4.6}$$

with boundary condition

$$u(x, 0) = \phi(x).$$

We shall then construct nonlinear PIDEs from (4.6) that would generates our bid and ask pricing functionals.

4.2.2 $\mathcal{G}$-expectations using Distortions

Inspired by Peng's approach, we will introduce nonlinearity into (4.6) by replacing the linear infinitesimal generator $\mathcal{L}$ by a nonlinear generator $\mathcal{G}$. Let

$$K = \int_{-\infty}^{\infty} y^2 k(y)dy.$$

Then $\mathcal{L}$ could be rewritten as

$$\mathcal{L}(u) = K \int_{-\infty}^{\infty} \frac{u(t, x+y) - u(t, x)}{y^2} \frac{y^2 k(y)}{K} dy. \tag{4.7}$$

We note that $y^2k(y)/K$ is always nonnegative and integrates to one. Thus (4.7) could be seen as a scaled expectation of $(u(t,x+Y)-u(t,x))/Y^2$ for some random variable Y with probability density $y^2k(y)/K$. Define

$$Y_{t,x} = \frac{u(t,x+Y)-u(t,x)}{Y^2},$$

we have

$$\mathcal{L}(u) = KE(Y_{t,x}).$$

As a result, we could define the generators for the bid and ask prices as

$$\mathcal{G}_b^{QV} = Kb_\gamma(Y_{t,x}),$$

and

$$\mathcal{G}_a^{QV} = Ka_\gamma(Y_{t,x}),$$

where b_γ and a_γ are the static one period bid and ask pricing operators defined in (2.1) and (2.2). The superscript QV stands for quadratic variation since this method requires the underlying Lévy process having finite quadratic variation.

Another way of constructing nonlinear operators is to normalize the Lévy density. We start with the following integral

$$\mathcal{L}_\epsilon(u) = \int_{|y|\geq\epsilon} (u(t,x+y)-u(t,x))k(y)dy, \tag{4.8}$$

which could be obtained from (4.5) by truncating a small neighborhood of zero. Although in general, a Lévy density does not integrate to a finite number, it is integrable on $\{|y|\geq\epsilon\}$. Let K_ϵ denote $\int_{|y|\geq\epsilon} k(y)dy$, then (4.8) can be expressed as

$$\mathcal{L}_\epsilon(u) = K_\epsilon \int_{|y|\geq\epsilon} (u(t,x+y)-u(t,x))\frac{k(y)}{K_\epsilon}dy.$$

The above equation can be viewed as

$$K_\epsilon E(u(t, x + Z) - u(t, x)),$$

where Z is a random variable that takes values in $R\backslash(-\epsilon, \epsilon)$ with probability density function given by $k(y)/K_\epsilon$. Let $Z_{t,x} = u(t, x+Z) - u(t, x)$, we could then define our nonlinear generators as

$$\mathcal{G}_b^{NL} = K_\epsilon b_\gamma(Z_{t,x}),$$

and

$$\mathcal{G}_a^{NL} = K_\epsilon a_\gamma(Z_{t,x}),$$

where NL indicates the normalizing Lévy approach.

We finally model continuous time bid and ask prices as $\mathcal{G}$-expectations induced by distorted PIDEs of the following form

$$u_t = \mathcal{G}(u) - ru.$$

We note that the above methods could easily be generalized to construct two price economies driven by non stationary and non homogeneous processes, such as Sato processes and Hunt processes [27]. Existence and uniqueness of solutions to the resulting distorted PIDEs will be discussed in the next section.

4.3 Viscosity Solutions of Distorted PIDEs

4.3.1 Introduction

In this section, existence and uniqueness of solution to the Cauchy problem of a distorted PIDE is studied. Consider a distorted PIDE of form

$$u_t(t,x) = F(t,x,u,u_x(t,x)) + \int_{|y|\geq\epsilon} (u(t,x+y)-u(t,x))\widetilde{k_{t,x}(dy)} \text{ on } (0,T]\times R, \quad (4.9)$$

with initial condition

$$u(0,x) = \varphi(x),$$

where F is a function that is continuous in its every argument, $\varphi(x)$ a continuous function with $|\varphi(x)| \leq M$ for some nonnegative constant M. $k_{t,x}$ denotes a bounded measure on $\{|y| \geq \epsilon\}$, satisfying

$$\lim_{(s,y)\to(t,x)} \int_{|z|\geq\epsilon} |k_{s,y} - k_{t,x}| dz = 0,$$

and there exists a positive constant μ_K, such that

$$\sup_{(t,x)\in[0,T)\times R} \int_{|y|\geq\epsilon} k_{t,x}(dy) \leq \mu_K. \quad (4.10)$$

The $\widetilde{k_{t,x}}$ comes from distortion which depends on the integrand. Pick $C \geq 2M$, we will only consider those solutions that satisfies

$$|u(t,x+y) - u(t,x)| \leq C. \quad (4.11)$$

This section is organized as follows. We first discuss the properties of distorted integrals. Next we introduce several equivalent definitions of viscosity solutions. We would then proceed to prove the existence and uniqueness of solutions to distorted

PIDEs by following Alvarez and Tourin's approach developed in [1]. We will begin with a comparison theorem for distorted PIDEs and then apply Perron's method for the construction of solutions.

4.3.2 Properties of Distorted Expectations

Let X be a random variable defined on a general probability space $(\Omega, \mathcal{F}, P)$, and $F(x)$ its distribution function. The expectation of X is defined as

$$E(X) = \int x dF(x).$$

We know from elementary probability theory that the expectation operator E is linear. Recall from Chapter 2, a distortion is an increasing function from $[0, 1]$ to $[0, 1]$. One example is given by the minmaxvar function,

$$\Psi(u) = 1 - (1 + u^{\frac{1}{1+\gamma}})^{1+\gamma}. \tag{4.12}$$

We define the distorted expectation of X as

$$\mathcal{E}(X) = \int x d\Psi(F(x)), \tag{4.13}$$

where Ψ is the minmaxvar function in (4.12).

We will start with the superadditivity of $\mathcal{E}$.

Theorem 4.3.1. *Suppose X, Y are two random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$ with joint pdf or pmf, then*

$$\mathcal{E}(X + Y) \geq \mathcal{E}(X) + \mathcal{E}(Y)$$

Proof. Suppose X and Y have joint pdf $f(x,y)$. Let Z denotes $X+Y$. The pdf of Z, $f_Z(z)$ satisfies

$$f_Z(z) = \int_{-\infty}^{\infty} f(x, z-x)dx. \tag{4.14}$$

From (4.13), we have

$$\mathcal{E}(X+Y) = \int_{-\infty}^{\infty} zd\Psi(\int_{-\infty}^{z} f_Z(s)ds).$$

Let us define a function $g_Z(z)$ by

$$g_Z(z) = \Psi^{'}\left(\int_{-\infty}^{z} f_Z(s)ds\right) \cdot f_Z(z). \tag{4.15}$$

Thus $g_Z(z)$ is a nonnegative function that integrates to 1, and it satisfies

$$\mathcal{E}(Z) = E_g(Z).$$

Based on (4.14)and (4.15), we could define a function g(x,y) as

$$g(x,y) = \frac{f(x,y)}{f_Z(x+y)} \cdot g_Z(x+y).$$

$g(x,y)$ is a valid pdf since

$$\begin{aligned}
\int\int g(x,y)dxdy &= \int\int \frac{f(x,y)}{f_Z(x+y)} \cdot g_Z(x+y)dxdy \\
&= \int\int \frac{f(x,z-x)}{f_Z(z)} \cdot g_Z(z)dxdz \quad (x=x, z=x+y) \\
&= \int \frac{\int f(x,z-x)dx}{f_Z(z)} g_Z(z)dz \\
&= \int g_Z(z)dz \\
&= 1.
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
E_g(X+Y) &= \int\int (x+y)g(x,y)dxdy \\
&= \int\int (x+y)\frac{f(x,y)}{f_Z(x+y)} \cdot g_Z(x+y)dxdy \\
&= \int\int z\frac{f(x,z-x)}{f_Z(z)} \cdot g_Z(z)dxdz \\
&= \int z\frac{\int f(x,z-x)dx}{f_Z(z)} g_Z(z)dz \\
&= \int z g_Z(z)dz \\
&= \mathcal{E}(Z) \\
&= \mathcal{E}(X+Y).
\end{aligned}$$

Since expectation is linear, we obtain

$$\mathcal{E}(X+Y) = E_g(X+Y) = E_g(X) + E_g(Y). \tag{4.16}$$

Next, we will prove

$$E_g(X) \geq \mathcal{E}(X).$$

Let $F(x)$ and $G(x)$ denote the marginal distribution function of X under $f(x,y)$ and $g(x,y)$ respectively. We have

$$F(u) = \int_{-\infty}^{u}\int f(x,y)dydx,$$

and

$$\begin{aligned}
G(u) &= \int_{-\infty}^{u}\int g(x,y)dydx = \int_{-\infty}^{u}\int \frac{f(x,y)}{f_Z(x+y)} \cdot g_Z(x+y)dydx \\
&= \int_{-\infty}^{u}\int \frac{f(x,z-y)}{f_Z(z)} \cdot g_Z(z)dzdx \\
&= \int_{-\infty}^{\infty} \frac{\int_{-\infty}^{u} f(x,z-x)dx}{f_Z(z)} \Psi'\left(\int_{-\infty}^{z} f_Z(s)ds\right) \cdot f_Z(z)dz \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{u} f(x,z-x)dx\Psi'\left(\int_{-\infty}^{z} f_Z(s)ds\right)dz.
\end{aligned}$$

We note that

$$E_g(X) = E_G(X).$$

Let us define

$$\widetilde{F}(x) = \Psi(F(x)),$$

thus $\widetilde{F}(x)$ is a valid distribution function satisfying

$$E_{\widetilde{F}}(X) = \mathcal{E}(X).$$

$E_g(X) \geq \mathcal{E}(X)$ is equivalent with $E_G(X) \geq E_{\widetilde{F}}(X)$. By the fact that

$$E(X) = \int_0^\infty (1 - F(x))dx - \int_{-\infty}^0 F(x)dx, \tag{4.17}$$

it suffice to prove $G(u) \leq \widetilde{F}(u)$, for any $u \in R$.

Let us write $\widetilde{F}(u)$ in the following form,

$$\begin{aligned} \widetilde{F}(u) &= \Psi\left(\int_{-\infty}^{u}\int_{-\infty}^{\infty} f(x,y)dydx\right) \\ &= \Psi\left(\int_{-\infty}^{\infty}\int_{-\infty}^{u} f(x,z-x)dxdz\right). \end{aligned}$$

Define

$$H(u,v) = \int_{-\infty}^{v}\int_{-\infty}^{u} f(x,z-x)dx\Psi'\left(\int_{-\infty}^{z} f_Z(s)ds\right)dz,$$

and

$$K(u,v) = \Psi\left(\int_{-\infty}^{v}\int_{-\infty}^{u} f(x,z-x)dxdz\right),$$

such that

$$G(u) = \lim_{v\to\infty} H(u,v),$$

and

$$\widetilde{F}(u) = \lim_{Z\to\infty} K(u,v).$$

We have

$$\frac{\partial H(u,v)}{\partial v} = \int_{-\infty}^{u} f(x, v-x)dx \cdot \Psi' \left(\int_{-\infty}^{v} f_Z(z)dz \right), \tag{4.18}$$

and

$$\frac{\partial K(u,v)}{\partial v} = \int_{-\infty}^{u} f(x, v-x) \cdot \Psi' \left(\int_{-\infty}^{v} \int_{-\infty}^{u} f(x, z-x)dxdz \right). \tag{4.19}$$

Comparing (4.18) and (4.19), we notice the only difference lies in Ψ'. It follows from (4.14) that

$$f_Z(z) \geq \int_{-\infty}^{X} f(x, z-x)dx,$$

hence

$$\int_{-\infty}^{Z} f_Z(z)dz \geq \int_{-\infty}^{Z} \int_{-\infty}^{X} f(x, z-x)dxdz.$$

The concavity of $\Psi(u)$ then implies

$$\frac{\partial H(u,v)}{\partial v} \leq \frac{\partial K(u,v)}{\partial v}.$$

Since the limits of $H(u,v)$ and $K(u,v)$ as v goes to negative infinity are both zero, we have

$$H(u,v) = \int_{-\infty}^{v} \frac{\partial H(u,v)}{\partial v}(u,s)ds,$$

and

$$K(u,v) = \int_{-\infty}^{v} \frac{\partial K(u,v)}{\partial v}(u,s)ds.$$

In all, $H(u,v) \leq K(u,v)$ for any value of v. Take limits on both sides as v goes to infinity, we get $G(x) \leq \widetilde{F}(x)$. As a result, $E_g(X) \geq \mathcal{E}(X)$. Similarly, we could also prove $E_g(Y) \geq \mathcal{E}(Y)$, substituting into 4.16 to obtain

$$\mathcal{E}(X+Y) = E_g(X) + E_g(Y) \geq \mathcal{E}(X) + \mathcal{E}(Y).$$

The proof for discrete X and Y would follow the same steps. $\square$

We next state the monotonicity of $\mathcal{E}$, and $\mathcal{E} \leq E$.

Theorem 4.3.2. *Suppose X and Y are two random variables with $X \leq Y$ a.s., we have*

$$\mathcal{E}(X) \leq \mathcal{E}(Y).$$

For any r.v Z,

$$\mathcal{E}(Z) \leq E(Z).$$

Proof. By equation (4.17), it is equivalent to show

$$\Psi(F_X(x)) \geq \Psi(F_Y(x)), \tag{4.20}$$

and

$$\Psi(F_Z(x)) \geq F_Z(x). \tag{4.21}$$

(4.20) is satisfied since $F_X(x) \geq F_Y(x)$, and $\Psi(u)$ is increasing. It follows from $\Psi(u) \geq u$ that (4.21) also holds . $\square$

We finally present an analogue of the dominated convergence theorem.

Theorem 4.3.3. *Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables that converges almost surely to a random variable X defined on the same probability space $(\Omega, \mathcal{F}, P)$. Suppose $|X_n| \leq M$ for some positive constant M. Then,*

$$\mathcal{E}(X_n) \to \mathcal{E}(X)$$

Proof. By Theorem 4.3.1,

$$\mathcal{E}(X - X_n) \leq \mathcal{E}(X) - \mathcal{E}(X_n) \leq -\mathcal{E}(X_n - X).$$

Moreover, from Theorem 4.3.2,

$$\mathcal{E}(-|X - X_n|) \leq (\mathcal{E}(X - X_n), \mathcal{E}(X_n - X)) \leq \mathcal{E}(|X - X_n|).$$

Hence,

$$|\mathcal{E}(X) - \mathcal{E}(X_n)| \leq max(\mathcal{E}(|X - X_n|), |\mathcal{E}(-|X - X_n|)|). \tag{4.22}$$

From Theorem 4.3.2, we have $\mathcal{E}(|X - X_n|) \leq E(|X - X_n|)$. Since $|X - X_n| \leq 2M$ a.s., by dominanted convergence theorem, $\lim_{n\to\infty} E|X - X_n| = 0$, which implies

$$\lim_{n\to\infty} \mathcal{E}(|X - X_n|) = 0. \tag{4.23}$$

By Chebyshev's inequality,

$$P(|X - X_n| \geq \epsilon) \leq \frac{E|X - X_n|}{\epsilon}.$$

As a result, for any $\epsilon > 0$, $\delta > 0$, there exists some N, such that when $n \geq N$,

$$P(|X - X_n| \geq \epsilon) \leq \delta.$$

We have

$$\begin{aligned}
\mathcal{E}(-|X - X_n|) &= \mathcal{E}(-|X - X_n|\mathbf{1}_{|X-X_n|\geq\epsilon} - |X - X_n|\mathbf{1}_{|X-X_n|<\epsilon}) \\
&\geq \mathcal{E}(-|X - X_n|\mathbf{1}_{|X-X_n|\geq\epsilon} + \mathcal{E}(-|X - X_n|\mathbf{1}_{|X-X_n|<\epsilon} \\
&\geq -2M\delta - \epsilon.
\end{aligned}$$

Since ϵ and δ are arbitrary, $\lim_{n\to\infty} |\mathcal{E}(-|X - X_n|)| = 0$. Combined with (4.23) and (4.22), we have $\lim_{n\to\infty} |\mathcal{E}(X) - \mathcal{E}(X_n)| = 0$, thus $\mathcal{E}(X_n) \to \mathcal{E}(X)$. $\square$

4.3.3 Viscosity Solutions

We are going to give two equivalent definitions for semicontinuous viscosity solution of distorted PIDE with boundary condition, that will both be useful in later proofs. Our definitions are analogues to those in [14].

A function f from $\mathbb{R}$ to $\mathbb{R}$ is a upper semicontinuous function (USC) if it satisfies

$$\limsup_{x \to x_0} f(x) \leq f(x_0).$$

The function f is called a lower semicontinuous function (LSC) if it satisfies

$$\liminf_{x \to x_0} f(x) \geq f(x_0).$$

f is continuous if it is both a USC and a LSU. Given any $u \in USC([0,T) \times \mathbb{R})$, and any point $(t,x) \in [0,T) \times \mathbb{R}$, define the first order superjet at point (t,x) as

$$\mathcal{P}^+u(t,x) = \{(p,q) \in \mathbb{R} \times \mathbb{R} | u(s,y) \leq u(t,x) + p(s-t) + q(y-x)$$
$$+ o(|s-t| + |y-x|) \quad \text{as } (s,y) \to (t,x)\},$$

and its closure,

$$\overline{\mathcal{P}}^+u(t,x) = \{(p,q) \in \mathbb{R} \times \mathbb{R} | (p,q) = \lim_{n \to \infty}(p_n, q_n), \quad with \quad (p_n, q_n) \in \mathcal{P}^+u(t_n, x_n),$$
$$and \quad \lim_{n \to \infty}(t_n, x_n, u(t_n, x_n)) = (t, x, u(t,x))\}.$$

For any $u \in LSC([0,T) \times \mathbb{R})$, its first order subjet $\mathcal{P}^-u(t,x)$ and its closure is defined by

$$\mathcal{P}^-u(t,x) = -\mathcal{P}^+(-u)(t,x),$$

and

$$\overline{\mathcal{P}}^-u(t,x) = -\overline{\mathcal{P}}^+(-u)(t,x).$$

We next state the main definition of this subsection.

Definition 4.3.4. *A viscosity subsolution of (4.9) is a locally bounded function $u \in USC([0,T) \times \mathbb{R})$ satisfying (4.11), such that for all $(t,x) \in (0,T) \times \mathbb{R}$, $(p,q) \in \mathcal{P}^{+}u(t,x)$,*

$$p - F(t,x,u,q) - \int_{|y|\geq\epsilon} (u(t,x+y) - u(t,x))\widetilde{k_{t,x}(dy)} \leq 0 \text{ on } (0,T) \times \mathbb{R}, \qquad (4.24)$$

and for all $x \in \mathbb{R}$, $u(0,x) \leq \varphi(x)$.

Similarly, a viscosity supersolution of (4.9) is a locally bounded function// $u \in LSC([0,T) \times \mathbb{R})$ satisfying (4.11), such that for all $(t,x) \in (0,T) \times \mathbb{R}$, $(p,q) \in \mathcal{P}^{-}u(t,x)$,

$$p - F(t,x,u,q) - \int_{|y|\geq\epsilon} (u(t,x+y) - u(t,x))\widetilde{k_{t,x}(dy)} \geq 0 \text{ on } (0,T) \times \mathbb{R}, \qquad (4.25)$$

and for all $x \in \mathbb{R}$, $u(0,x) \geq \varphi(x)$. Finally, $u \in C([0,T) \times R)$ is a viscosity solution of (4.9), if it is a viscosity subsolution and a viscosity supersolution of (4.9).

Remark 4.3.5. *It is equivalent to require the (p,q) in the above definition of subsolution(resp. supersolution) to be in $\overline{\mathcal{P}}^{+}u(t,x)$(resp. $\overline{\mathcal{P}}^{-}u(t,x)$). Since for all $(p,q) \in \overline{\mathcal{P}}^{+}u(t,x)$(resp. $\overline{\mathcal{P}}^{-}u(t,x)$), by definition, there exists a sequence $(t_n, x_n) \in [0,T) \times \mathbb{R}$, as well as $(p_n, q_n) \in \mathcal{P}^{+}u(t_n, x_n)$(resp. $(p_n, q_n) \in \mathcal{P}^{-}u(t_n, x_n)$), such that*

$$\lim_{n\to\infty} (t_n, x_n, u(t_n, x_n), p_n, q_n) = (t, x, u(t,x), p, q).$$

Substitute (t_n, x_n, p_n, q_n) into (4.24) (resp.(4.25)), and take upper limits on both sides(resp. lower limits). By assumption, $u(t_n, x_n + y) - u(t_n, x_n)$ is bounded, thus we could apply Theorem 4.3.3 to get the convergence of the distorted integral term.

Together with continuity of F, we conclude that (4.24) (resp. (4.25)) holds at (t, x) for (p, q).

Definition 4.3.6. *(Equivalence) A viscosity subsolution of (4.9) is a locally bounded function $u \in USC([0,T) \times \mathbb{R})$ satisfying (4.11), such that for all $(t, x) \in (0, T) \times \mathbb{R}$, bounded $\phi \in C^1([0,T) \times \mathbb{R})$ with $u(t, x) = \phi(t, x)$ and $u < \phi$ on $[0,T) \times \mathbb{R}/(t, x)$, we have*

$$\phi_t - F(t, x, \phi, \phi_x) - \int_{|y| \geq \epsilon} (\phi(t, x + y) - u(t, x)) \widetilde{k_{t,x}(dy)} \leq 0 \text{ on } (0, T) \times \mathbb{R}, \quad (4.26)$$

and for all $x \in \mathbb{R}$, $u(0, x) \leq \varphi(x)$.

Similarly, a viscosity supersolution of (4.9) is a locally bounded function $u \in LSC([0,T) \times \mathbb{R})$ satisfying (4.11), such that for all $(t, x) \in (0, T) \times \mathbb{R}$, $\phi \in C^1([0,T) \times \mathbb{R})$ with $u(t, x) = \phi(t, x)$ and $u > \phi$ on $[0,T) \times \mathbb{R}/(t, x)$, we have

$$\phi_t - F(t, x, \phi, \phi_x) - \int_{|y| \geq \epsilon} (\phi(t, x + y) - u(t, x)) \widetilde{k_{t,x}(dy)} \geq 0 \text{ on } (0, T) \times \mathbb{R}, \quad (4.27)$$

and for all $x \in R$, $u(0, x) \geq \varphi(x)$.

Remark 4.3.7. *Definitions 4.3.4 and 4.3.6 are equivalent. Definition 4.3.4 immediately implies Definition 4.3.6 because $(\phi_t, \phi_x) \in \mathcal{P}^+ u(t, x)$(resp. $\mathcal{P}^- u(t, x)$), and the distorted integral is monotone in its integrand. Conversely, given any (t, x) and $(p, q) \in \mathcal{P}^+ u(t, x)$(resp. $\mathcal{P}^- u(t, x)$), we could construct a sequence of functions $\phi_n \in C^1$, such that $u < \phi_n \leq u(t, x) + 2C$ on $[0,T) \times \mathbb{R}/(t, x)$, $\phi \to u$ on $[0,T) \times \mathbb{R}$, with*

$$(\phi_n(t, x), \partial_t \phi_n(t, x), \partial_x \phi_n(t, x)) = (u(t, x), p, q).$$

Since (4.26)(resp. (4.27)) holds for any ϕ_n, as $n \to \infty$, the inequality becomes (4.24)(resp. 4.25), where we have used Theorem 4.3.3 to get the convergence of distorted integral terms.

4.3.4 A Comparison Theorem

We will focus our discussion on equations with F given by

$$F(t, x, u, u_x) = -ru + b(t, x)u_x, \tag{4.28}$$

where r is a nonnegative constant, $b(t, x)$ a function on $\mathbb{R} \times \mathbb{R}$ that is Liptchitz continuous in x.

Before going into details of the comparison theorem, let us first state a fact about the viscosity subsolution and supersolution of (4.9).

Proposition 4.3.8. *Suppose F is of form (4.28), then $-Me^{-rt}$ is a viscosity subsolution of (4.9),and Me^{-rt} is a viscosity supersolution of (4.9).*

Proof. Denote $-Me^{-rt}$ by $\underline{u}$. It is obvious that $-M = \underline{u}(0, x) \leq \varphi(x)$ for all $x \in \mathbb{R}$. By Definition 4.3.6, it suffice to show that for any $(t, x) \in [0, T) \times \mathbb{R}$, and $\phi \in C^1([0, T) \times \mathbb{R})$ with $\underline{u}(t, x) = \phi(t, x)$, the following equation holds,

$$\phi_t - F(t, x, \phi, \phi_x) - \int_{|y| \geq \epsilon} (\phi(t, x + y) - u(t, x))\widetilde{k_{t,x}(dy)} \leq 0. \tag{4.29}$$

Since $\phi_t - \underline{u}$ reaches its local minimum at (t, x), we have $\phi_t(t, x) = \underline{u}_t(t, x) = -r\underline{u}(t, x)$, and $\phi_x(t, x) = \underline{u}_x(t, x) = 0$, which yields $\phi_t - F(t, x, \phi, \phi_x) = 0$. By the requirements of ϕ, we have $\phi(t, x + y) - u(t, x) > 0$ for any y. Thus the distorted integral term is positive, (4.29) holds. Therefore, $\underline{u}$ is a viscosity subsolution. It could be shown using the same arguments that Me^{-rt} is a viscosity supersolution of (4.9). $\square$

From now on, let's denote $-Me^{-rt}$ and Me^{-rt} by $\underline{u}$ and $\overline{v}$ respectively. We have the following comparison theorem.

Theorem 4.3.9. *Suppose F satisfies (4.28). Let $u \in USC([0,T) \times \mathbb{R})$ be a viscosity subsolution, and $v \in LSC([0,T) \times \mathbb{R})$ a viscosity supersolution of (4.9), such that $\underline{u} \leq (u,v) \leq \overline{v}$. Then,*

$$u \leq v \ \ on \ [0,T) \times \mathbb{R} \ . \tag{4.30}$$

Proof. We prove by contradiction. Assume (4.30) doesn't hold, that is

$$\sup_{(t,x)\in[0,T)\times R} u(t,x) - v(t,x) > 0.$$

We could fix for some $\eta > 0$ small enough, such that

$$N := \sup_{(t,x)\in[0,T)\times R} u(t,x) - v(t,x) - \frac{\eta}{T-t} > 0. \tag{4.31}$$

For any $\delta > 0$, set

$$N_\delta = \max_{(t,x)\in[0,T)\times R} u(t,x) - v(t,x) - \frac{\eta}{T-t} - \delta|x|^2. \tag{4.32}$$

Suppose N_δ is achieved at (t_δ, x_δ). Based on N_δ, for any $\tilde{\epsilon} > 0$, define

$$H_{\delta,\tilde{\epsilon}} = u(t,x) - v(t,y) - \frac{\eta}{T-t} - \delta|x|^2 - \frac{|x-y|^2}{\tilde{\epsilon}}.$$

Let's denote a global maximum point of $H_{\delta,\tilde{\epsilon}}$ by $(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}, y_{\tilde{\epsilon}})$. By Proposition 3.7 of [14] we have

$$\lim_{\tilde{\epsilon}\to 0} \frac{|x_{\tilde{\epsilon}} - y_{\tilde{\epsilon}}|^2}{\tilde{\epsilon}} = 0,$$

and there exists a subsequence of $(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}, y_{\tilde{\epsilon}})$, such that

$$\lim_{\tilde{\epsilon}\to 0}(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}) = (t_\delta, x_\delta, x_\delta),$$

and

$$\lim_{\tilde{\epsilon}\to 0} u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}) = u(t_\delta, x_\delta), \quad \lim_{\tilde{\epsilon}\to 0} v(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}) = v(t_\delta, x_\delta).$$

It is easy to check we also have

$$\lim_{\delta\to 0} N_\delta = N, \quad and \quad \lim_{\delta\to 0} \delta|x_\delta|^2 = 0.$$

Following Theorem 8.3 in [14], there exists $(p_{\tilde{\epsilon}}, q_{\tilde{\epsilon}}) \in \mathbb{R}\times\mathbb{R}$, such that

$$(p_{\tilde{\epsilon}} + \frac{\eta}{(T-t_{\tilde{\epsilon}})^2}, q_{\tilde{\epsilon}} + 2\delta x_{\tilde{\epsilon}}) \in \overline{P}^+ u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}), \text{and } (p_{\tilde{\epsilon}}, q_{\tilde{\epsilon}}) \in \overline{P}^- v(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}).$$

Definition 4.3.4 and the fact that u is a subsolution then yield

$$\begin{aligned} &p_{\tilde{\epsilon}} + \frac{\eta}{(T-t_{\tilde{\epsilon}})^2} + ru(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}) - b(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}})(q_{\tilde{\epsilon}} + 2\delta x_{\tilde{\epsilon}}) \\ &- \int_{|z|\geq\epsilon} (u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}} + z) - u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}))\widetilde{k_{t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}}(dz)} \leq 0. \end{aligned} \tag{4.33}$$

Similarly, since v is a supersolution, one can also get

$$p_{\tilde{\epsilon}} + rv(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}) - b(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}})q_{\tilde{\epsilon}} - \int_{|z|\geq\epsilon} (v(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}} + z) - v(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}))\widetilde{k_{t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}}(dz)} \geq 0. \tag{4.34}$$

We could substract (4.34) from (4.33) and take limit superior as $\tilde{\epsilon} \to 0$ on both sides to get

$$\begin{aligned} &ru(t_\delta, x_\delta) - rv(t_\delta, y_\delta) - 2\delta x_\delta b(t_\delta, x_\delta) \\ &\leq \limsup_{\tilde{\epsilon}\to 0} \left[\int_{|z|\geq\epsilon} (u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}} + z) - u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}))\widetilde{k_{t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}}(dz)} \right. \\ &\left. - \int_{|z|\geq\epsilon} (v(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}} + z) - v(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}))\widetilde{k_{t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}}(dz)} \right]. \end{aligned} \tag{4.35}$$

We drop the term $\eta/(T-t_{\tilde{\epsilon}})^2$ since it won't affect the direction of the inequality. Properties of distorted expectation and the fact that $u \in USC$ and $v \in LSC$ then yield

$$\begin{aligned}
&\limsup_{\tilde{\epsilon}\to 0}\left[\int_{|z|\geq\epsilon}(u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}+z)-u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}}))\widetilde{k_{t_{\tilde{\epsilon}},x_{\tilde{\epsilon}}}(dz)}\right.\\
&\qquad\left.-\int_{|z|\geq\epsilon}(v(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}+z)-v(t_{\tilde{\epsilon}}, y_{\tilde{\epsilon}}))\widetilde{k_{t_{\tilde{\epsilon}},y_{\tilde{\epsilon}}}(dz)}\right]\\
&\leq \int_{|z|\geq\epsilon}(u(t_\delta, x_\delta+z)-u(t_\delta, x_\delta))\widetilde{k_{t_\delta,x_\delta}(dz)}\\
&\qquad-\int_{|z|\geq\epsilon}(v(t_\delta, x_\delta+z)-v(t_\delta, x_\delta))\widetilde{k_{t_\delta,x_\delta}(dz)}.
\end{aligned}\tag{4.36}$$

By (4.31) and (4.32), we have for any $z \in \mathbb{R}$,

$$u(t_\delta, x_\delta+z)-v(t_\delta, x_\delta+z)-\frac{\eta}{t_\delta}\leq N, \tag{4.37}$$

and

$$u(t_\delta, x_\delta)-v(t_\delta, x_\delta)-\frac{\eta}{t_\delta}-\delta|x_\delta|^2 = N_\delta. \tag{4.38}$$

Substract (4.38) from (4.37) to get

$$u(t_\delta, x_\delta+z)-u(t_\delta, x_\delta)\leq v(t_\delta, x_\delta+z)-v(t_\delta, x_\delta)+N-N_\delta.$$

By monotonicity and superadditivity of distorted expectation, we could bound from above the right hand side of (4.36) by $(N-N_\delta)\int_{|z|\geq\epsilon}k_{t_\delta,x_\delta}(dz)$, which, by assumption of measure $k_{t,x}$ and the fact that $\lim_{\delta\to 0}N_\delta = N$, goes to zero as $\delta\to 0$. Combined with (4.35), we obtain

$$\limsup_{\delta\to 0}[ru(t_\delta, x_\delta)-rv(t_\delta, y_\delta)-2\delta x_\delta b(t_\delta, x_\delta)]\leq 0.$$

Since $\lim_{\delta\to 0}\delta|x_\delta|^2 = 0$, and $\lim_{\delta\to 0}\delta|x_\delta| = 0$,

$$\lim_{\delta\to 0}2|\delta x_\delta b(t_\delta, x_\delta)|\leq 2\lim_{\delta\to 0}|\delta x_\delta|(|b(0,0)|+C(T+|x_\delta|)) = 0.$$

We conclude that

$$\limsup_{\delta\to 0}[u(t_\delta, x_\delta) - v(t_\delta, y_\delta)] \le 0,$$

which contradicts our assumption that N is positive. □

4.3.5 Perron's Method

In this subsection, we are going to construct the solution to (4.9) using Perron's method.

Theorem 4.3.10. *Suppose all the assumptions of the comparison theorem hold, then there is a unique viscosity solution $u \in C([0,T) \times \mathbb{R})$ satisfying (4.9), with $\underline{u} \le u \le \overline{v}$ on $[0,T) \times \mathbb{R}$.*

Proof. To begin with, let's define a new function $v(t,x)$ in the following way. For all $(t,x) \in [0,T) \times \mathbb{R}$, define

$$v(t,x) = sup\left\{u(t,x) | u \text{ is a subsolution of (4.9) such that } u \le \overline{v} \text{ on } [0,T) \times \mathbb{R}\right\}. \tag{4.39}$$

Let v^* be the upper semicontinuous envelope of v, which is defined as

$$v^*(t,x) = \limsup_{[0,T)\times\mathbb{R}\ni(s,y)\to(t,x)} v(s,y).$$

And define the lower semicontinuous envelope of v, $v_*(t,x) = -(-v)^*(t,x)$.

We first show that v^* and v_* are viscosity subsolution and supersolution, respectively. We prove it in two steps. We begin by showing that v^* (resp. v_*) is a viscosity subsolution (resp. supersolution) with generalized boundary condition, and then prove that v^* is indeed a subsolution in the strict sense(proof for v_* proceeds in the same way). We say that some function $u(t,x)$ is a subsolution of (4.9) with generalized boundary condition, that is the inequality $u(0,x) \leq \varphi(x)$ in Definition 4.3.4 is replaced by the following condition:

$$p - F(t,x,u(0,x),q) - \int_{|y|\geq \epsilon} (u(0,x+y) - u(0,x))\widetilde{k_{0,x}(dy)} \leq 0, \text{ if } u(0,x) > \varphi(x),$$

for all $(p,q) \in \mathcal{P}^{+}(0,x)$.

For any $(t,x) \in [0,T) \times \mathbb{R}$, by definition of v^*, we could find a sequence (s_n, y_n, u_n), where u_n is a subsolution, such that

$$\lim_{n\to\infty} (s_n, y_n, u_n(s_n, y_n)) = (t, x, v^*(t,x)). \tag{4.40}$$

For any bounded $\phi \in C^1([0,T) \times \mathbb{R})$ satisfying $v^*(t,x) = \phi(t,x)$ and $v^* < \phi$ on $[0,T) \times \mathbb{R}/(t,x)$, define

$$N_r = \left\{(s,y) | (s-t)^2 + (y-x)^2 \leq r^2, \quad (s,y) \in [0,T) \times \mathbb{R}\right\}.$$

We could find some r small enough, such that N_r is compact. Let (t_n, x_n) be the maximum point of $u_n - \phi$ in N_r. Since N_r is compact, there exists a converging subsequence (t_n, x_n) (for convenience still indexed by n), and the converging limit is also in N_r. We claim that the limit is (t,x). We prove by contradiction. Suppose the limit is $(s,y) \neq (t,x)$. By (4.40), we could find N, such that when $n \geq N$, $(s_n, y_n) \in N_r$. As a result, for $n \geq N$,

$$u_n(s_n, y_n) \leq u_n(t_n, x_n) - \phi(t_n, x_n) + \phi(s_n, y_n).$$

Take liminf on both sides to get

$$v^*(t,x) \leq \liminf_{n\to\infty} u_n(t_n, x_n) + \phi(t,x) - \phi(s,y),$$

together with

$$\liminf_{n\to\infty} u_n(t_n, x_n) \le \limsup_{n\to\infty} v(t_n, x_n) \le v^*(s, y), \tag{4.41}$$

implies

$$v^*(t, x) \le v^*(s, y) + \phi(t, x) - \phi(s, y).$$

Because $v^*(t, x) = \phi(t, x)$, we have $\phi(s, y) \le v^*(s, y)$, which implies $(s, y) = (t, x)$ and all the inequality signs in (4.41) is actually equality signs. Thus we have found a sequence (t_n, x_n, u_n) with

$$\lim_{n\to\infty} (t_n, x_n, u_n(t_n, x_n)) = (t, x, v^*(t, x)),$$

and for each n, (t_n, x_n) is the local maximum of $u_n - \phi$.

We note that if $t \neq 0$, $t_n > 0$ for large n. When $t = 0$ and $v^*(0, x) > \varphi(x)$, we could also get $t_n > 0$ for large n. If not, then there exists a subsequence such that $u_{n_k}(0, x_{n_k}) \to v^*(0, x)$. However, since each u_{n_k} is a subsolution, $u_{n_k}(0, x_{n_k}) \le \varphi(0, x_{n_k})$. Take limit on both sides to get $v^*(0, x) \le \varphi(0, x)$, contradiction. As a result, $t_n > 0$ for large n, unless $t = 0$ and $v^*(0, x) \le \varphi(x)$. In addition, we have $(\phi_t, \phi_x) \in \mathcal{P}^+ u_n(t_n, x_n)$, which comes from the fact that (t_n, x_n) is a local maximum of $u_n - \phi$.

When $t_n > 0$, by Definition 4.3.4, one has

$$\phi_t - F(t_n, x_n, u_n(t_n, x_n), \phi_x) \le \int_{|y|\ge\epsilon} (\phi(t_n, x_n + y) - u_n(t_n, x_n))\widetilde{k_{t_n,x_n}(dy)},$$

where we have used the monotonicity of distorted expectation together with $u_n \le \phi$. Letting $n \to \infty$, by the continuity of F and Theorem 4.3.3, we conclude that

$$\phi_t - F(t, x, v^*(t, x), \phi_x) \le \int_{|y|\ge\epsilon} (\phi(t, x + y) - v^*(t, x))\widetilde{k_{t,x}(dy)}.$$

Therefore v^* is a viscosity subsolution with generalized boundary condition .

Next let's prove v_* is a viscosity supersolution with generalized boundary condition. We argue by contradiction. Suppose v_* doesn't have the desire property, then exist $(t,x) \in [0,T) \times \mathbb{R}$, a function $\phi \in C_b^1([0,T) \times \mathbb{R}$ with $v_*(t,x) = \phi(t,x)$ and $v_* > \phi$ elsewhere, $v_*(0,x) < \varphi(x)$ if $t = 0$, such that

$$\phi_t - F(t,x,\phi(t,x),\phi_x) < \int_{|y|\geq\epsilon} (\phi(t,x+y) - \phi(t,x))\widetilde{k_{t,x}(dy)}. \tag{4.42}$$

By definition of v_*, $v_*(t,x) \leq \bar{v}(t,x)$. We claim that $v_*(t,x) < \bar{v}(t,x)$. If $v_*(t,x) = \bar{v}(t,x)$, then $(\phi_t, \phi_x) \in \mathcal{P}^-\bar{v}(t,x)$. However, $\bar{v}$ is a supersolution, which contradict (4.42). So we have $v_*(t,x) < \bar{v}(t,x)$, then there exist η_1, δ_1 positive, such that $\phi + \delta_1 \leq \bar{v}$ on $B_{\eta_1}(t,x) \cap [0,T) \times R$, and $\phi(0,y) + \delta_1 \leq \varphi(y)$ for $(0,y) \in B_{\eta_1}(t,x) \cap [0,T) \times \mathbb{R}$ if $t = 0$. By continuity of F and the distorted integral in (4.42), we could find η_2, δ_2 positive, such that for all $(s,y,\delta) \in (B_{\eta_2}(t,x) \cap [0,T) \times \mathbb{R}) \times [0,\delta_2]$, the following inequality holds.

$$\phi_s - F(s,y,\phi(s,y)+\delta,\phi_y) < \int_{|z|\geq\epsilon} (\phi(s,y+z) + \delta - (\phi(s,y)+\delta))\widetilde{k_{s,y}(dz)}. \tag{4.43}$$

Denote $\eta_0 = min(\eta_1, \eta_2)$, there exists $\delta_3 > 0$, such that $v_* > \phi + \delta_3$ on the boundary of $B_{\eta_0}(t,x) \cap [0,T) \times \mathbb{R}$. Now, set $\delta_0 = min(\delta_1, \delta_2, \delta_3)$, and define

$$w = max(v^*, \phi + \delta_0) \text{ on } B_{\eta_0}(t,x) \cap [0,T) \times \mathbb{R}, \text{ and } w = v^* \text{ elsewhere.}$$

Observing that w is upper semicontinuous and $\underline{u} \leq w \leq \bar{v}$, we claim that w is actually a viscosity subsolution of (4.9) with generalized boundary condition. For all $(s,y) \in [0,T) \times \mathbb{R}$, $\psi \in C_b^1([0,T) \times \mathbb{R}$ with $w(s,y) = \psi(s,y)$ and $w < \psi$ elsewhere, we have (ψ_s, ψ_y) in either $\mathcal{P}^+v^*(s,y)$ or $\mathcal{P}^+(\phi+\delta_0)(s,y)$. By our choice of δ_3, $\phi+\delta_0$ is achieved at interior point of $B_{\eta_0}(t,x) \cap [0,T) \times \mathbb{R}$. As a result, if $w(s,y) = \phi(s,y) + \delta_0$, (s,y) is a local maximum of $\phi + \delta_0 - \psi$, we obtain $(\psi_s, \psi_y) = (\phi_s, \phi_y)$. The fact that v^* is a subsolution with generalized boundary condition and (4.43) would then imply (4.26). Therefore w is a subsolution with generalized boundary condition. Let w_* be the lower semicontinuous envelope of w, we have

$$w_*(t,x) \geq max(v_*(t,x), \phi(t,x) + \delta_0) > v_*(t,x).$$

Thus there exists some (s,y), such that $w(s,y) > v(s,y)$.

Recall the definition of v in (4.39), we notice that if w is a viscosity subsolution of (4.9) in the strict sense, then we have obtained a contradiction. Actually one could show that for any u, that is a viscosity subsolution (resp. supersolution) of (4.9) with generalized boundary condition and $\underline{u} \leq u \leq \overline{v}$, it is indeed a viscosity subsolutiton (resp. supersolution) in the strict sense. The rest of the proof is devoted to this claim.

We would only show the case for subsolution, the proof for supersolution proceeds exactly in the same way. Suppose $u \in USC([0,T) \times \mathbb{R})$, $\underline{u} \leq u \leq \overline{v}$, u is a viscosity subsolution of (4.9) with generalized boundary condition, and $u(0,x) > \varphi(x)$ for some $x \in \mathbb{R}$. Since φ is continuous, we could fix some $r > 0$, such that $u(0,x) > \varphi(y)$ for $y \in [x-r, x+r]$ and $\overline{B_r(0,x)} \cap [0,T) \times \mathbb{R}$ compact. For any $\tilde{\epsilon} > 0$, there is some $C_{\tilde{\epsilon}} > 0$, with

$$C_{\tilde{\epsilon}} > \sup_{(s,y) \in \overline{B_r(0,x)} \cap [0,T) \times R} \left\{ F(s,y,u(s,y), \frac{2(y-x)}{\tilde{\epsilon}}) + 2M\mu_K \right\},$$

where μ_K is defined in (4.10). As $\tilde{\epsilon} \to 0$, we could pick $C_{\tilde{\epsilon}}$ such that $C_{\tilde{\epsilon}} \to \infty$. Let

$(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}})$ denote the maximum of

$$u(s,y) - \frac{(y-x)^2}{\tilde{\epsilon}} - sC_{\tilde{\epsilon}}$$

on $B_r(0,x) \cap [0,T) \times \mathbb{R}$. For $\tilde{\epsilon}$ sufficiently small, $(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}})$ is the global maximum, which implies $(C_{\tilde{\epsilon}}, 2(x_{\tilde{\epsilon}} - x)/\tilde{\epsilon}) \in \mathcal{P}^+ u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}})$. When $t_{\tilde{\epsilon}} = 0$, $u(0, x_{\tilde{\epsilon}}) \geq u(0,x) \geq \varphi(x_{\tilde{\epsilon}})$. By definition of the generalized boundary condition, we conclude that for sufficiently small $\tilde{\epsilon}$, (4.24) should hold for $(C_{\tilde{\epsilon}}, 2(x_{\tilde{\epsilon}} - x)/\tilde{\epsilon}) \in \mathcal{P}^+ u(t_{\tilde{\epsilon}}, x_{\tilde{\epsilon}})$. However, (4.24) contradicts our choice of $C_{\tilde{\epsilon}}$. In all, u is a viscosity subsolution in the strict sense.

We have proved that v^* and v_* are viscosity subsolution and supersolution, respectively. It is clear that $\underline{u} \leq (v^*, v_*) \leq \overline{v}$. By Theorem 4.3.9 (comparison theorem), $v^* \leq v_*$, which implies $v^* = v_* = v$. Therefore, $v \in C([0,T) \times \mathbb{R})$ is the unique viscosity solution of (4.9). Uniqueness follows directly from Theorem 4.3.9. $\square$

Chapter 5

Applications of Continuous Time Two Price Economy

5.1 Overview

In this Chapter, three applications of continuous time two price economies will be given. The first example uses BSDEs on CTMCs with Markovian drivers to generate bid and ask prices of option spreads. In the second example, the $\mathcal{G}$-expectation approach is applied to produce credit capital commitments for derivatives with bilateral counterparty risk. The last example computes bid and ask swap rates and swaption prices using the distorted PIDE method.

5.2 Bid and Ask Option Spreads

Consider a two price economy whose randomness is generated by a CTMC $(X_t)_{0\leq t\leq T}$. We'll model bid and ask prices of state contingent claims as solutions to BSDEs on CTMCs with Markovian driver functions. By Theorem 3.3.1, the solution to a BSDE of form

$$Y_t - \int_{]t,T]} F(X_u, u, Y_{u-}, Z_u)du + \int_{]t,T]} Z_u^T dM_u = Q(X_T), \tag{5.1}$$

is given by

$$\frac{d\tilde{V}}{dt} - A_{T-t}\tilde{V}(t) - \tilde{F}(T-t, \tilde{V}(t)) = 0, \tag{5.2}$$

where $\tilde{F}$ satisfies

$$\tilde{F}_i(T-t, \tilde{V}(t)) = F(e_i, T-t, \tilde{V}_i(t), \tilde{V}(t)).$$

We pick the driver functions for bid and ask prices as

$$F_b(X_t, t, Y_t, Z_t) = |X_t^T A_t X_t| \int_R z d(\Psi_\gamma(F_Z(z))) - X_t^T A_t Z_t,$$

and

$$F_a(X_t, t, Y_t, Z_t) = -F_b(X_t, t, Y_t, -Z_t),$$

with Ψ_γ the minmaxvar function defined in (2.3).

We first construct a CTMC $(X_t)_{0\leq t\leq T}$ as an approximation to a Variance Gamma (VG) process. The parameter for the VG process is taken to be ($\sigma = 0.43, \nu = 0.76, \theta = -0.6$). The parameter is obtained by calibrating the VG process to the option surface on JPM on the date of Oct. 18, 2011, with initial stock price equals to 32. We then construct the CTMC X following the method suggested by Mijatović and Pistorious in [42]. The total number of states of X is set to 200, with state value vector V of X given by

$$V = (-2.51 : 0.0178 : 1.04)^T.$$

Suppose the stock price S_t satisfies

$$S_t = 32exp(0.45t + X_t^T V),$$

where V is the state value vector of X.

Consider a strangle with $K_1 = 25$, $K_2 = 39$ and maturity $T = 1$. A strangle is a derivative with terminal payoff

$$Q(S_T) = (K_1 - S_T)^+ + (S_T - K_2)^+, \tag{5.3}$$

which is illustrated in Figure 5.1.

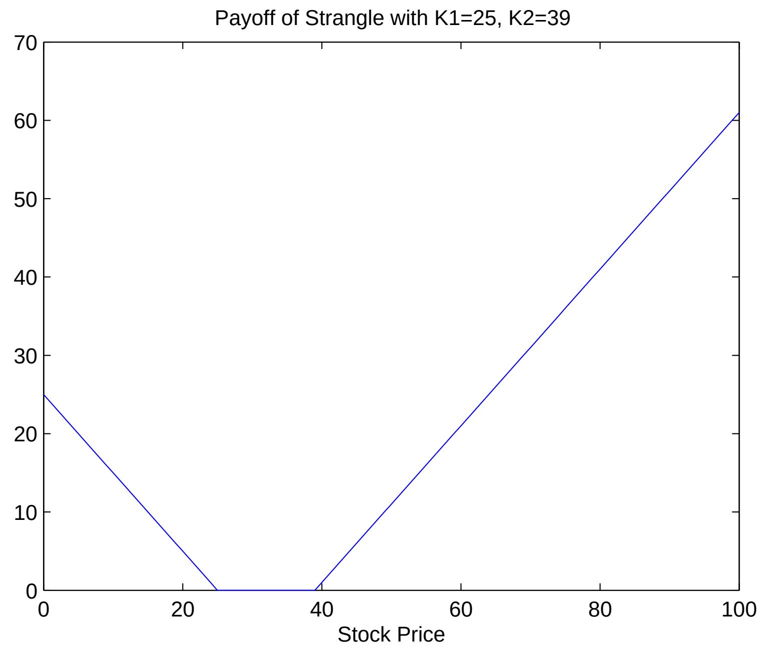


Figure 5.1: Strangle payoff with $K_1 = 25$ and $K_2 = 39$

We then generate the bid and ask prices of the strangle as solutions to (5.2) with initial value (5.3), at three different distortion levels: $\gamma = 0.1$, $\gamma = 0.2$, $\gamma = 0.5$. We have also computed the expected value of the strangle ($\gamma = 0$). Figure 5.2 shows the bid, ask and expected values of the strangle at distortion level $\gamma = 0.1$. In Figure 5.3, the trading advantage of the strangle is illustrated with $\gamma = 0.1$, in which the top curve and the bottom curve correspond to the ask and bid prices of trading the strangle as a put option with strike $K_1 = 25$ and a call option with strike $K_2 = 39$, while the three curves in the middle are the same as in Figure 5.2.

Figure 5.4 compares the bid and ask prices of the strangle under different distortion levels. Table 5.1 presents the bid, ask and expected prices of the strangle

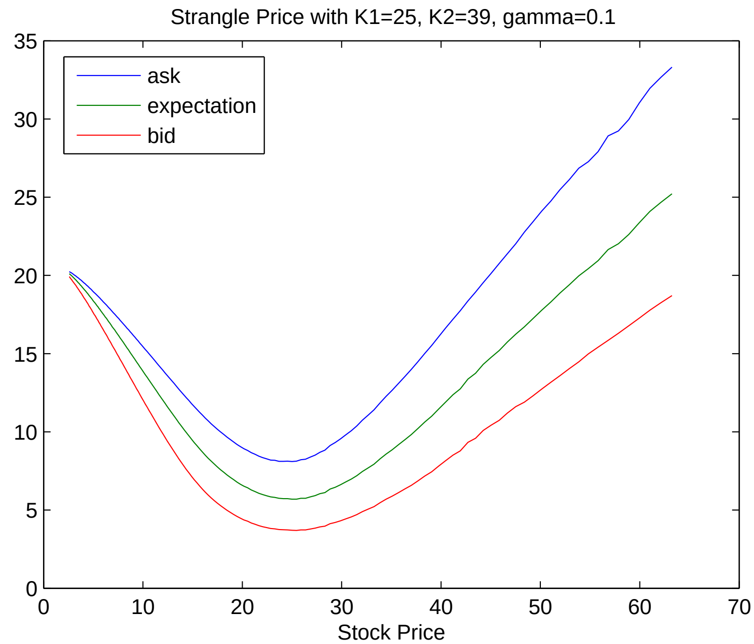


Figure 5.2: Bid, ask and expected prices of a strangle at distortion level 0.1

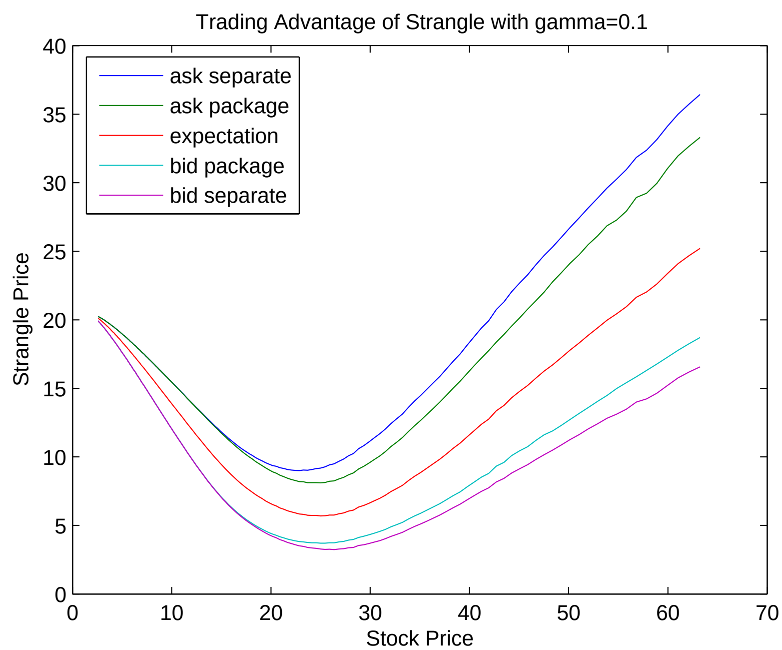
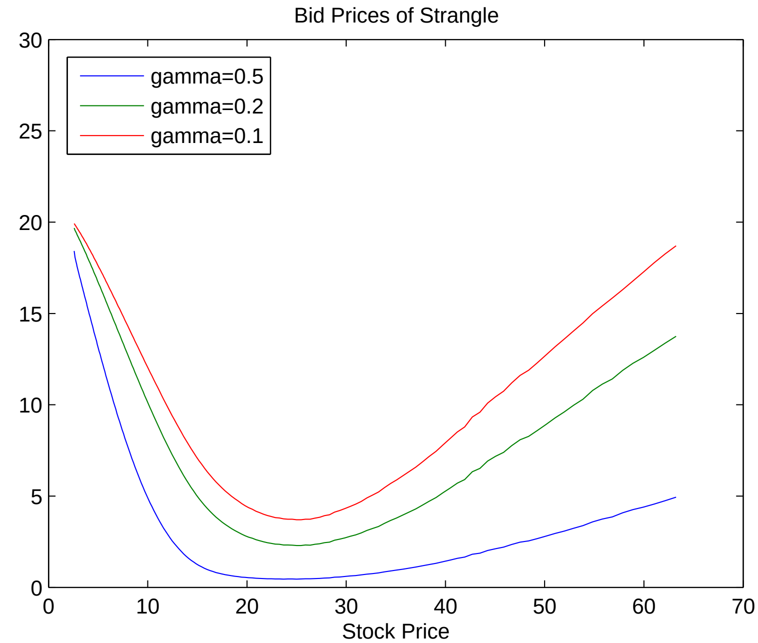
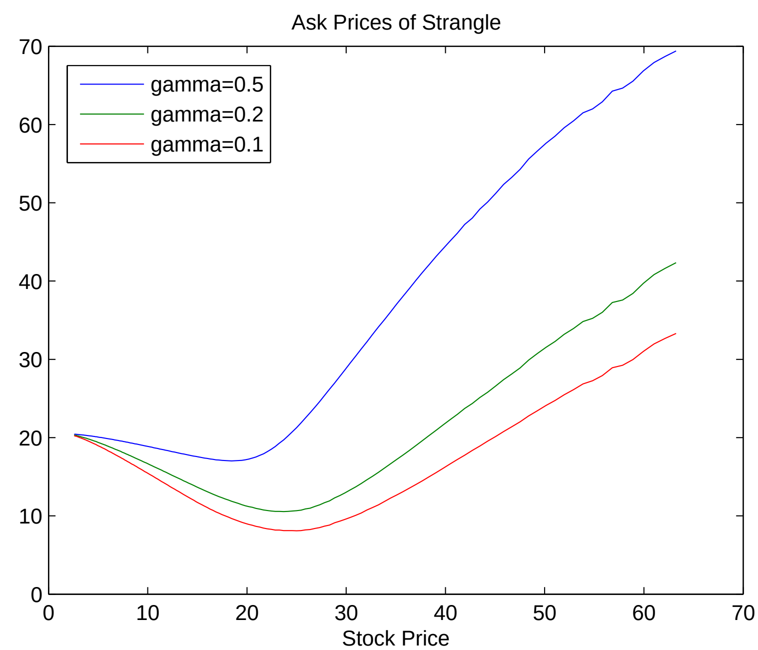


Figure 5.3: Trading advantage of a strangle at distortion level 0.1



(a) Bid prices of a strangle under distortion levels 0.1, 0.2, and 0.5



(b) Ask prices of a strangle under distortion levels 0.1, 0.2, and 0.5

Figure 5.4: Bid and ask prices of a strangle under different distortion levels

at $t = 0$ under different distortion levels and stock prices.

Table 5.1: Prices of a strangle at different stock prices and distortion levels

Distortion Level	Stock Price	Bid	Ask	Expectation
0.1	25	3.7031	8.1031	5.6940
	32	4.9032	10.7494	7.4689
	39	7.4493	15.5189	10.9982
0.2	25	2.2964	10.6567	5.6940
	32	3.1185	14.6207	7.4689
	39	4.9195	20.9310	10.9982
0.5	25	0.4573	19.7900	5.6940
	32	0.7239	30.6385	7.4689
	39	1.3219	41.7017	10.9982

We next consider a butterfly spread with $K_1 = 25$, $K_2 = 39$ and maturity $T = 1$. A butterfly spread is a derivative with terminal payoff given by

$$Q(S_T) = (S_T - K_1)^+ + (S_T - K_2)^+ - 2(S_T - \frac{K_1 + K_2}{2})^+, \tag{5.4}$$

which is illustrated in Figure 5.5.

We computed the bid, ask and expected prices of the butterfly spread at distortion levels $\gamma = 0.1$, $\gamma = 0.2$, $\gamma = 0.5$. Figure 5.6 shows the bid, ask and expected values of the butterfly spread at distortion level $\gamma = 0.1$. Figure 5.7 illustrates the trading advantage of the butterfly spread with $\gamma = 0.01$. In Figure

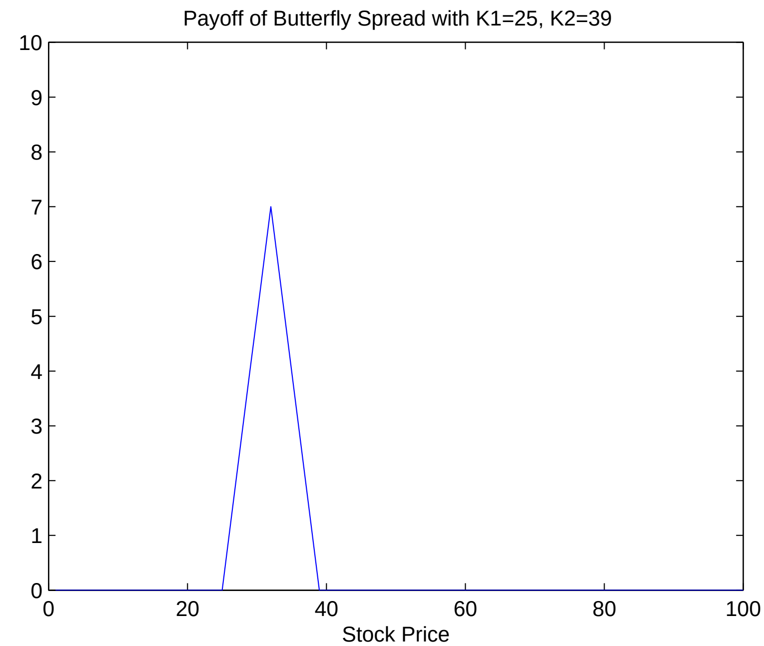


Figure 5.5: Butterfly spread payoff with $K_1 = 25$ and $K_2 = 39$

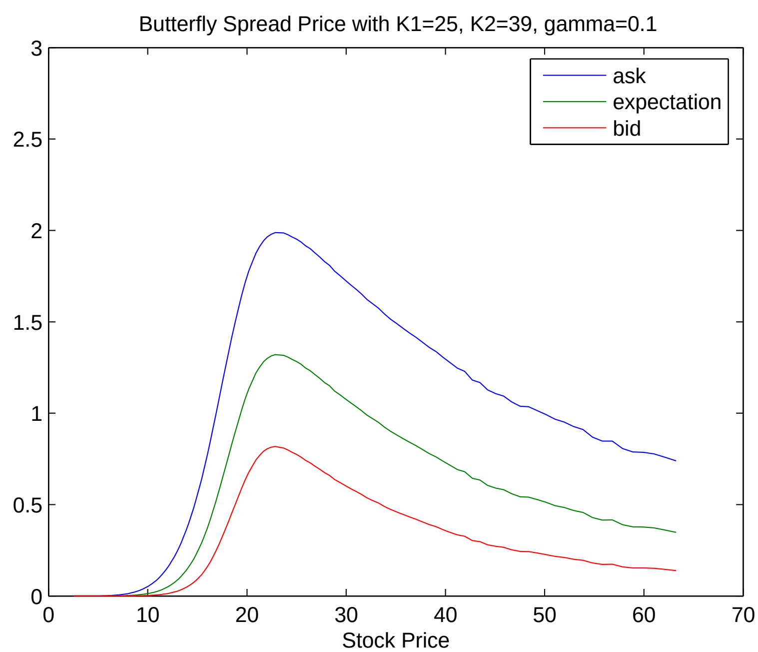


Figure 5.6: Bid, ask and expected prices of a butterfly spread at distortion level 0.1

5.7, the top curve and the bottom curve correspond to the ask and bid prices of trading the butterfly spread as buying two call options with strikes $K_1 = 25$ and $K_2 = 39$, and selling two call options with strike $(K_1 + K_2)/2 = 32$ or *vise versa*, while the three curves in the middle are the bid, ask, and expected values of the butterfly spread computed at distortion level $\gamma = 0.01$. Figure 5.8 compares the

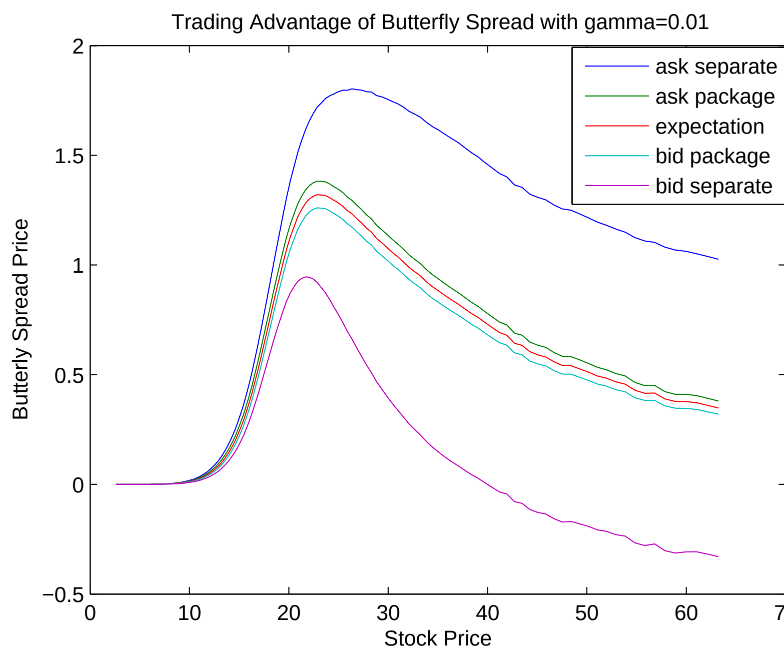
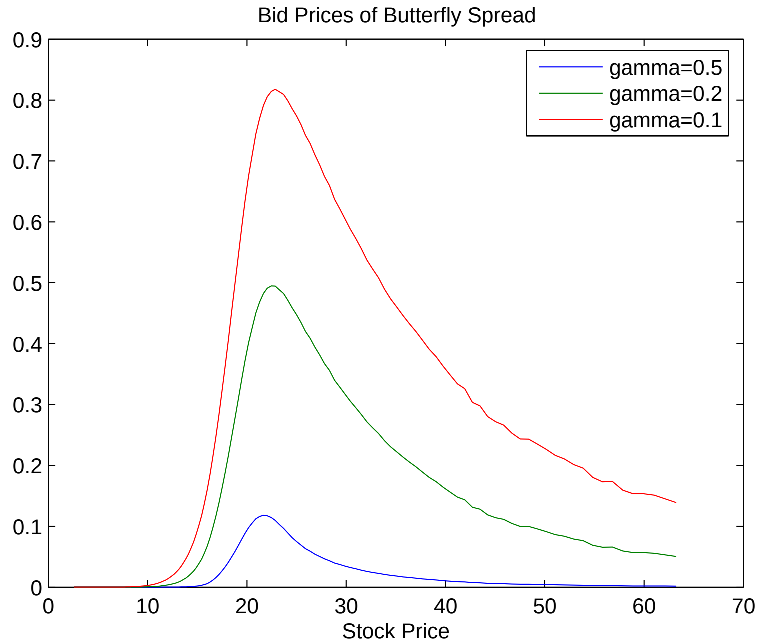
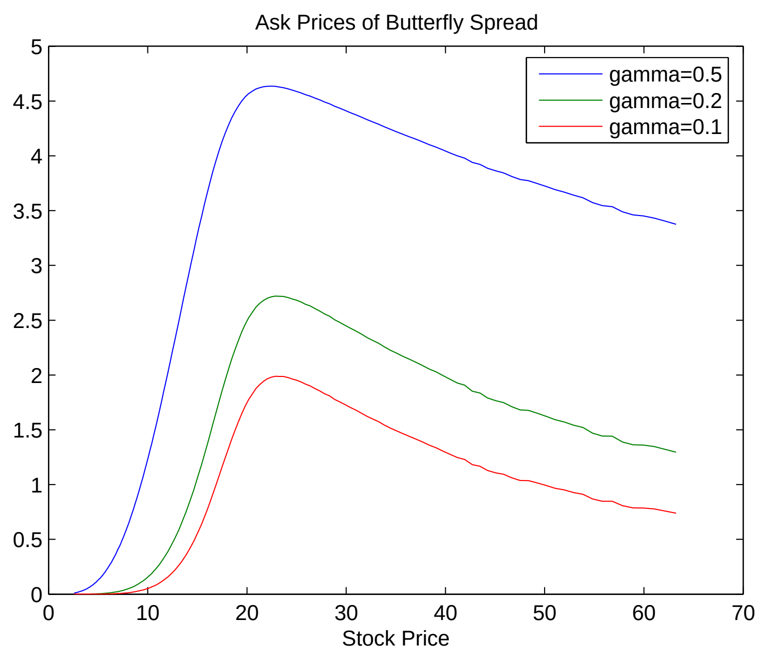


Figure 5.7: Trading advantage of a butterfly spread at distortion level 0.1

bid and ask prices of the butterfly spread under different distortion levels. Table 5.2 presents the bid, ask and expected prices of the butterfly spread at $t = 0$ under different distortion levels and stock prices.



(a) Bid prices of a butterfly spread under distortion levels 0.1, 0.2, and 0.5



(b) Ask prices of a butterfly spread under distortion levels 0.1, 0.2, and 0.5

Figure 5.8: Bid and ask prices of a butterfly spread under different distortion levels

Table 5.2: Prices of a butterfly spread at different stock prices and distortion levels

Distortion Level	Stock Price	Bid	Ask	Expectation
0.1	25	0.7740	1.9526	1.2827
	32	0.5372	1.6226	0.9895
	39	0.3787	1.3365	0.7609
0.2	25	0.4475	2.6826	1.2827
	32	0.2714	2.3420	0.9895
	39	0.1734	2.0281	0.7609
0.5	25	0.0750	4.5879	1.2827
	32	0.0258	4.3306	0.9895
	39	0.0118	4.0794	0.7609

5.3 Credit Capital Commitments in Continuous Time

5.3.1 Introduction

The theory of credit capital commitment (CCC) is introduced in [39] by Madan in place of the credit valuation adjustment (CVA) in incomplete markets. The CCC of a cash flow X is defined as the difference between its bid and ask prices. Some illustrative evaluations are given in [39] in the context of static one period two price economy. We are going to apply the $\mathcal{G}$-expectation method developed in Chapter 4 to compute CCCs in a continuous time two price economy.

5.3.2 PIDE Representation

Consider a derivative contract on an asset S between an issuer I and a counterparty C that may both default. At default, one party would pay the other a predetermined amount. If there is no default, C would pay I $H(S) \in R$ at maturity T. Let J_I and J_C be two independent Poisson processes with intensities λ_I and λ_C that jump from 0 to 1 on default of I and C respectively. Suppose S is driven by a pure jump Levy process $(X_t)_{0 \leq t \leq T}$ independent of the default processes, under the martingale measure. Suppose X has finite variation and quadratic variation. Let V denote the value of the derivative to the issuer I. In fact, V is a function of t, X, J_I and J_C, with boundary conditions

$$
\begin{aligned}
V(t, X, 1, 0) &= M_I(t, X) \\
V(t, X, 0, 1) &= M_C(t, X),
\end{aligned}
$$

and terminal condition

$$V(T, X, 0, 0) = \hat{H}(X),$$

where M_I and M_C are predetermined functions, and $\hat{H}(X)$ satisfying

$$\hat{H}(X) = H(S(X)).$$

We have the following result regarding $V(t, X, J_I, J_C)$.

Proposition 5.3.1. *Let $u(t, X, J_I, J_C)$ be the solution to*

$$\partial_t u + \mathcal{A}_t u - (r + \lambda_I + \lambda_C)u + \lambda_I M_I + \lambda_C M_C = 0$$

$$u(T, X, 0, 0) = \hat{H}_T(X)$$

$$u(t, X, 1, 0) = M_I(t, X), \;\; u(t, X, 0, 1) = M_C(t, X), \tag{5.5}$$

where $\mathcal{A}$ is the infinitesimal generator of X. Let τ denote the time of the first default, we have

$$u(t, X, 0, 0)$$

$$= E(\mathbf{1}_{\tau>T} e^{-rT-t} \hat{H}(X)|\mathcal{F}_t) + E(\mathbf{1}_{\tau\leq T} e^{-r(\tau-t)}((1_{J_I=1})M_I(\tau, X) + 1_{J_C=1}M_C(\tau, X))|\mathcal{F}_t).$$

As a result, $V(t, X, J_I, J_C) = u(t, X, J_I, J_C)$.

Proof. The triplet $\tilde{X} = (X, J_I, J_C)^T$ could be seen as a multivariate jump process with jump intensity $(k(x), \lambda_I \delta(1), \lambda_C \delta(1))^T$, that satisfies the following Levy-Itô decomposition:

$$d\tilde{X}(t) = \int_{R^3} diag(\tilde{z}) \tilde{N}(dt, dz).$$

By Itô's formula, for any $t < \tilde{t}$,

$$e^{-r(\tilde{t}-t)}u(\tilde{t}, X_{\tilde{t}}, J_I(\tilde{t}), J_C(\tilde{t})) - u(t, x, J_I(t), J_C(t))$$

$$= \int_{]t,\tilde{t}]} (\partial_t u - ru) ds$$

$$+ \sum_{t<s\leq\tilde{t}} (u(s, X_s, J_I(s), J_C(s)) - u(s, X_{s-}, J_I(s-), J_C(s-)))$$

$$= \int_{]t,\tilde{t}]} \partial_t u - ru$$

$$+ \int_R u(s, X_{s-} + z, J_I(s-), J_C(s-)) - u(s, X_{s-}, J_I(s-), J_C(s-))k(z)dz$$

$$+ \lambda_I(u(s, X_{s-}, J_I(s-) + 1, J_C(s-)) - u(s, X_{s-}, J_I(s-), J_C(s-)))$$

$$+ \lambda_C(u(s, X_{s-}, J_I(s-), J_C(s-) + 1) - u(s, X_{s-}, J_I(s-), J_C(s-)))ds$$

$$+ \int_{]t,\tilde{t}]} \int_R u(s, X_{s-} + z, J_I(s-), J_C(s-)) - u(s, X_{s-}, J_I(s-), J_C(s-))N_1(ds, dz)$$

$$- k(z)dzds$$

$$+ \int_{]t,\tilde{t}]} (u(s, X_{s-}, J_I(s-) + 1, J_C(s-)) - u(s, X_{s-}, J_I(s-), J_C(s-)))d(N_2(s) - \lambda_I s)$$

$$+ \int_{]t,\tilde{t}]} (u(s, X_{s-}, J_I(s-), J_C(s-) + 1) - u(s, X_{s-}, J_I(s-), J_C(s-)))d(N_3(s) - \lambda_C s)$$

Since the last three expressions in the above equation are martingales, if u satisfies (5.5), then for any $\tilde{t} > t$, with $J_I(t) = J_C(t) = 0$, we have

$$E(e^{-r(\tilde{t}-t)}u(\tilde{t}, X_{\tilde{t}}, J_I(\tilde{t}), J_C(\tilde{t}))) = u(t, x, 0, 0). \tag{5.6}$$

Substitute $\tilde{t} = \tau \wedge T$ into (5.6) to get

$$\begin{aligned} u(t, x, 0, 0) &= E(e^{-r(\tilde{t}-t)}u(\tau \wedge T, X_{\tau\wedge T}, J_I(\tau \wedge T), J_C(\tau \wedge T))) \\ &= E(\mathbf{1}_{\tau>T}e^{-rT-t}\hat{H}(X)|\mathcal{F}_t) \\ &\quad + E(\mathbf{1}_{\tau\leq T}e^{-r(\tau-t)}((\mathbf{1}_{J_I=1})M_I(\tau, X) + \mathbf{1}_{J_C=1}M_C(\tau, X))|\mathcal{F}_t). \end{aligned}$$

$\square$

5.3.3 Distorted PIDEs for Bid and Ask Prices

By Proposition 5.3.1, the risk neutral price of the derivative satisfies (5.5). Following the $\mathcal{G}$-expectation approach discussed in Chapter 4, we could use distorted PIDEs to describe the bid and ask prices. Recall the original PIDE

$$\partial_t u = -\mathcal{A}_t u + (r + \lambda_I + \lambda_C)u - \lambda_I M_I - \lambda_C M_C,$$

its right hand side of the equation could be rewritten as

$$ru - (\mathcal{A}u + \lambda_I(M_I - u) + \lambda_C(M_C - u)).$$

Let's define the operator $\mathcal{L}$ by

$$\mathcal{L} = \mathcal{A}u + \lambda_I(M_I - u) + \lambda_C(M_C - u).$$

$\mathcal{L}$ could be written in the following way,

$$\begin{aligned}\mathcal{L}u = (K + \lambda_I + \lambda_C)(&\int \frac{u(x+y) - u(x)}{y^2} \frac{y^2 k(y)}{K + \lambda_I + \lambda_C} dy \\ &+ \frac{\lambda_I}{K + \lambda_I + \lambda_C}(M_I - u) + \frac{\lambda_C}{K + \lambda_I + \lambda_C}(M_C - u)),\end{aligned} \tag{5.7}$$

where $K = \int y^2 k(y) dy$. Define

$$g(y) = \frac{y^2 k(y)}{K + \lambda_1 + \lambda_2}.$$

Then

$$\int \frac{u(x+y) - u(x)}{y^2} g(y) dy + \frac{\lambda_I}{K + \lambda_I + \lambda_C}(M_I - u) + \frac{\lambda_C}{K + \lambda_B + \lambda_C}(M_C - u)$$

could be seen as the expectation of a random variable Z, that can take values $(u(x+y)-u(x))/y^2$, $M_I - u$, and $M_C - u$, with distribution function

$$F_{QV}(z) = \begin{cases} \int_{A(t,x,z)} g(y)dy & z < M_- - u(x) \\ \int_{A(t,x,z)} g(y)dy + \frac{\lambda_I}{K+\lambda_I+\lambda_C}\mathbf{1}_{M_I<M_C} + \frac{\lambda_C}{K+\lambda_I+\lambda_C}\mathbf{1}_{M_I\geq M_C} & z \text{ elsewhere} \\ \int_{A(t,x,z)} g(y)dy + \frac{\lambda_I+\lambda_C}{K+\lambda_I+\lambda_C} & z \geq M^+ - u(x) \end{cases},$$

where

$$A(t,x,z) = \left\{ \frac{(u(x+y)-u(x)}{y^2} \leq z \right\},$$

$$M_- = min(M_I, M_C),$$

$$M^+ = max(M_I, M_C).$$

We have

$$\mathcal{L}u = (K+\lambda_I+\lambda_C)E(Z).$$

We define $\mathcal{G}_b^{QV}$ and $\mathcal{G}_a^{QV}$ as

$$\mathcal{G}_b^{QV}u = (K+\lambda_I+\lambda_C)\int z d\Psi_\gamma(F_{QV}(z)),$$

and

$$\mathcal{G}_a^{QV}u = -\mathcal{G}_b^{QV}(-u)$$

where Ψ is the minmaxvar function defined in (2.3).

In the above method, the idea is to build a probability measure from an infinite Levy measure by reducing the weights of small jumps and scaling by its quadratic variation. Another way to ignore small jumps is by truncating a small neighborhood of 0 and normalizing the truncated Levy measure. Consider the truncated generator $\mathcal{L}_\epsilon$,

$$\mathcal{L}_\epsilon u = (\hat{K} + \lambda_I + \lambda_C)\Big(\int_{\{|y|>\epsilon\}} (u(x+y) - u(x)) \frac{k(y)}{\hat{K} + \lambda_I + \lambda_C} dy + \frac{\lambda_I(M_I - u)}{\hat{K} + \lambda_I + \lambda_C} + \frac{\lambda_C(M_C - u)}{\hat{K} + \lambda_I + \lambda_C}\Big). \tag{5.8}$$

Similar to the QV method, we can view (5.8) as scaled expectation of a random variable Z, that could take values $u(x+y) - u(x)$, $M_I - u$ and $M_C - u$, with distribution function $F_{NL}(z)$ defined as

$$F_{NL}(z) = \begin{cases} \int_{B(t,x,z)} h(y)dy & z < M_- - u(x) \\ \int_{B(t,x,z)} h(y)dy + \frac{\lambda_I}{\hat{K}+\lambda_I+\lambda_C}\mathbf{1}_{M_I<M_C} + \frac{\lambda_C}{K+\lambda_I+\lambda_C}\mathbf{1}_{M_I\geq M_C} & z \text{ elsewhere} \\ \int_{B(t,x,z)} h(y)dy + \frac{\lambda_I+\lambda_C}{\hat{K}+\lambda_I+\lambda_C} & z \geq M^+ - u(x) \end{cases},$$

where $h(y)$ and $B(t,x,z)$ are defined as

$$h(y) = \frac{k(y)}{\hat{K} + \lambda_I + \lambda_C},$$

$$B(t,x,z) = \{u(x+y) - u(x) \leq z, \quad |y| \geq \epsilon\}.$$

We could then define $\mathcal{G}_b^{NL}$ and $\mathcal{G}_a^{NL}$ by

$$\mathcal{G}_b^{NL} u = (\hat{K} + \lambda_B + \lambda_C) \int z d\Psi_\gamma(F_{NL}(z)),$$

and

$$\mathcal{G}_a^{NL} u = -\mathcal{G}_b^{NL}(-u).$$

Once we have bid and ask generators, the prices are given by

$$\partial_t u - ru + \mathcal{G}(u) = 0.$$

5.3.4 Numerical Results

Consider a one year contract ($T = 1$) between two parties, with terminal payoff $H(S) = (0.9 - S)^+ + (S - 1.1)^+$, and default payment $M_I = M_C = 0$. Suppose $ln(S)$ follows a VG process with drift. We pick the parameter (σ, ν, θ) of the VG process to be $(0.39, 0.51, -0.57)$. The parameter is obtained by calibrating the VG process to the option surface on JPM on the date of Oct. 20, 2008, with initial stock price equals to 40. Assume the interest rate $r = 0.02$. We first choose default intensities as $\lambda_I = \lambda_C = 0.1$, and set the distortion level at 0.1. Figure 5.9 shows the graph of $H(S_T)$, with $ln(S_T)$ ranging from -1.54 to 1.46. Figure 5.10 presents the bid, ask

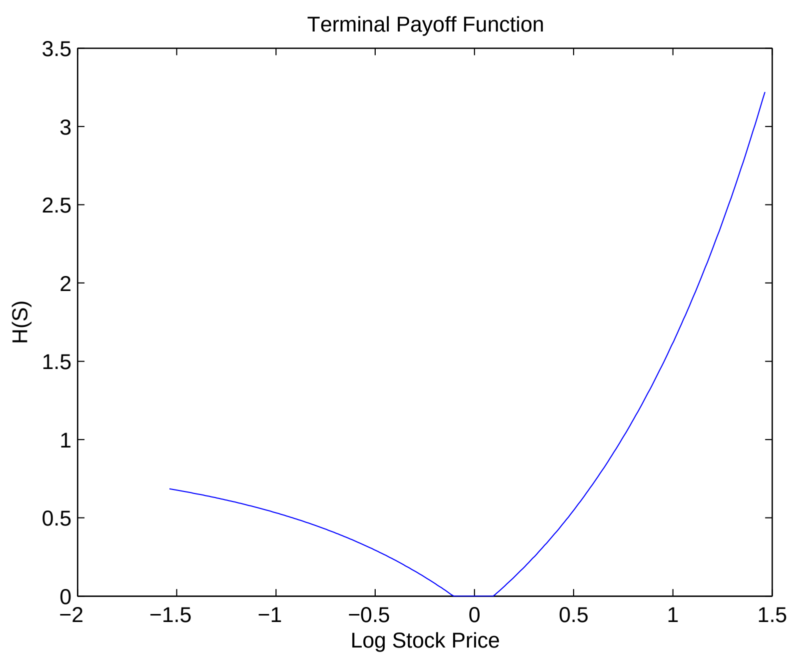
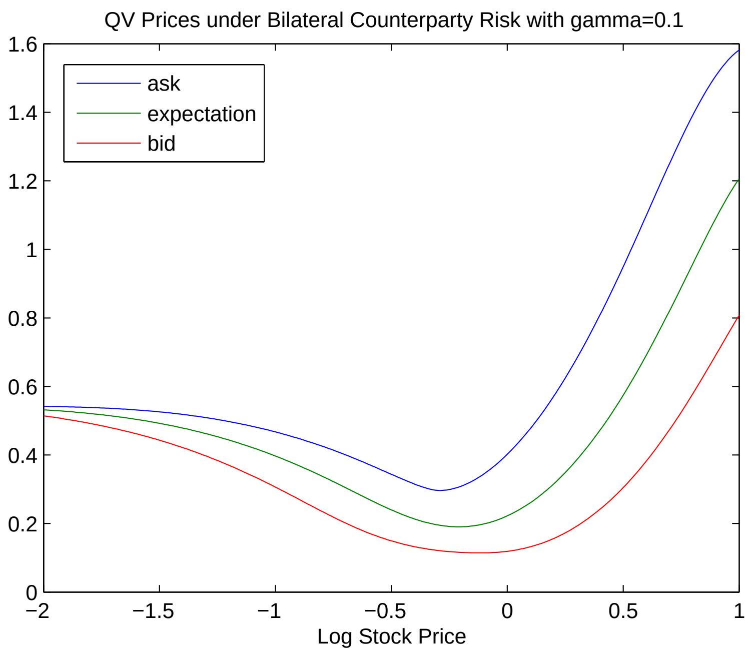
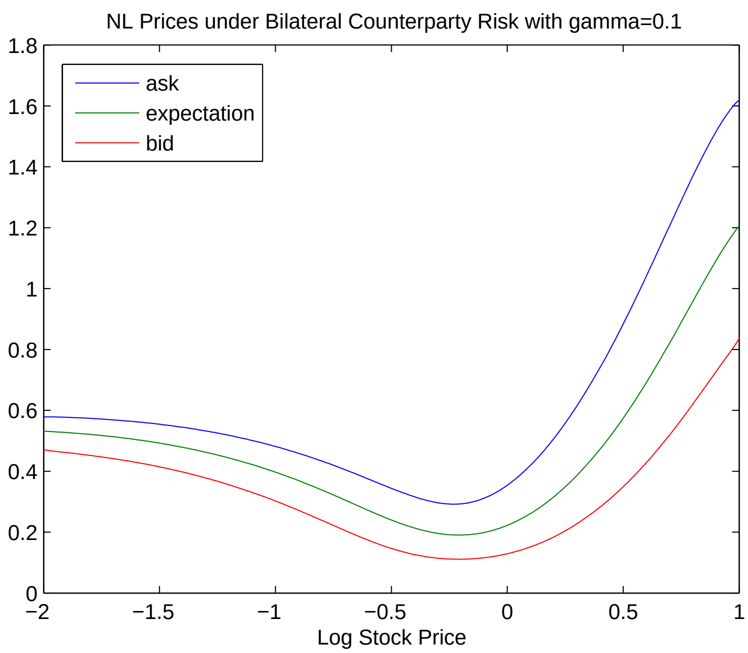


Figure 5.9: Terminal payoff function H(S)

and expected values of the derivative using the QV and NL methods. We compare in Figure 5.11 the CCCs computed using the QV and NL approaches. Figure 5.12



(a) Bid, ask, and expected values computed using QV method at distortion level 0.1



(b) Bid, ask, and expected values computed using NL method at distortion level 0.1

Figure 5.10: Bid, ask and expected values of a derivative with bilateral counterparty risk

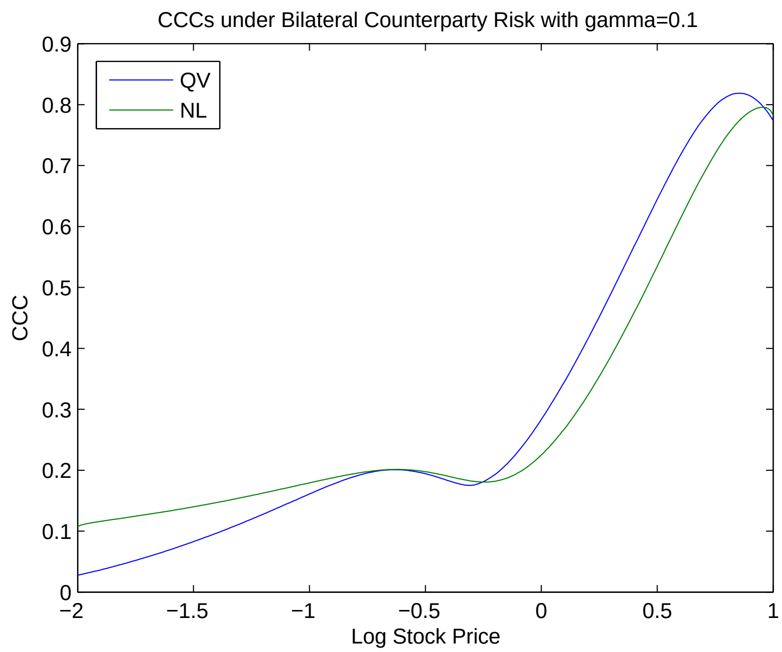
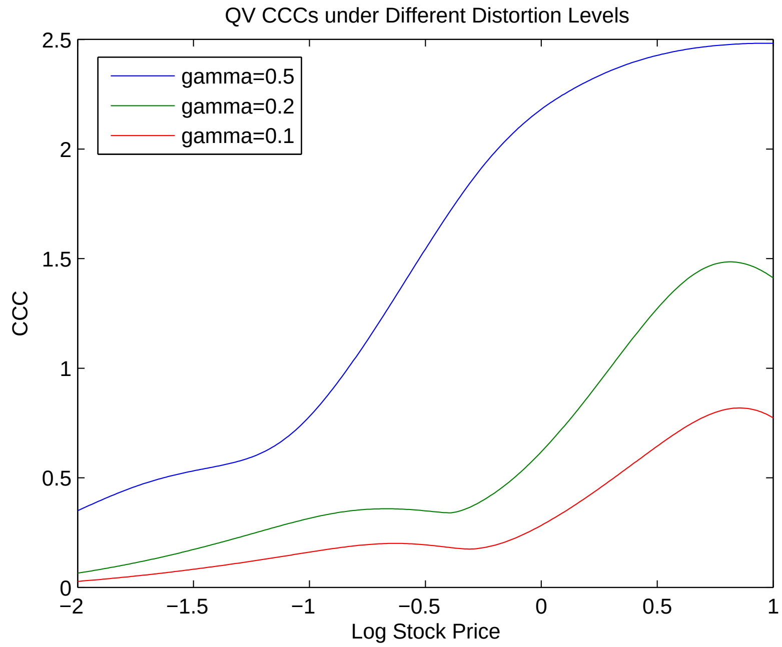


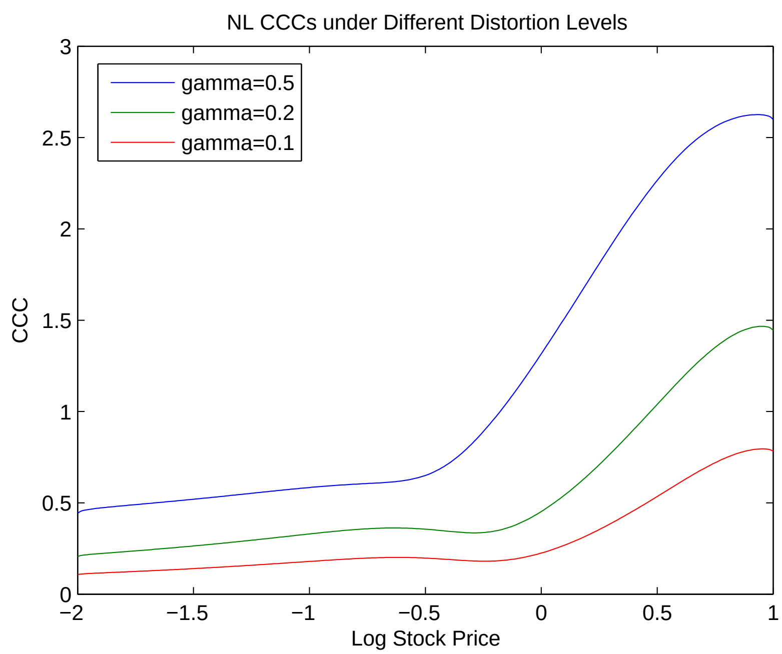
Figure 5.11: CCCs under bilateral counterparty risk

illustrates the effect of distortion levels on the CCCs. Table 5.3 compares the CCCs computed using QV and NL methods under different distortion levels and different initial stock prices.

We next compare the CCCs under bilateral counterparty risk with the CCCs under own default, counterparty default and no default. We set default rates for both I and C to be 0.1, and keep M_B and M_C same as before. The distortion level is set at $\gamma = 0.1$. By symmetry, own default and counterparty default would yield the same result. Figure 5.13 compares the CCCs under different default assumptions. Table 5.4 shows the CCCs at $S_0 = 1$ computed using QV and NL methods under different default assumptions. The distortion level is set to 0.1.



(a) CCCs computed using QV method at distortion levels 0.1, 0.2, and 0.5



(b) CCCs computed using QV method at distortion levels 0.1, 0.2, and 0.5

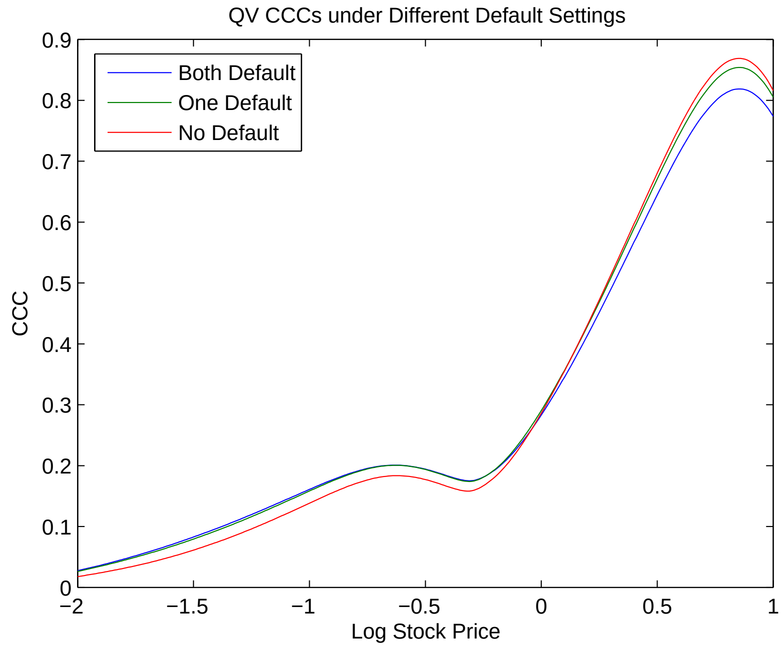
Figure 5.12: CCCs of a derivative with bilateral counterparty risk

Table 5.3: CCCs at different stock prices and distortion levels

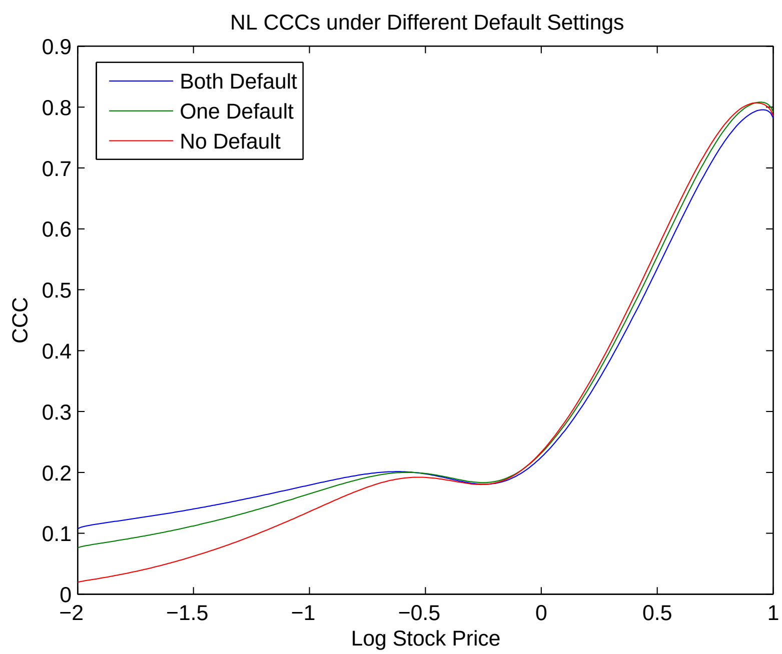
Distortion Level	Stock Price	CCC QV	CCC NL
0.1	0.9	0.2307	0.1957
	1	0.2830	0.2249
	1.1	0.3421	0.2656
0.2	0.9	0.5161	0.3848
	1	0.6185	0.4522
	1.1	0.7309	0.5373
0.5	0.9	2.0945	1.1318
	1	2.1816	1.3158
	1.1	2.2485	1.5006

Table 5.4: CCCs under different default assumptions with $\gamma = 0.1$ and $S_0 = 1$

Assumption	CCC QV	CCC NL
Bilateral Default Risk	0.2830	0.2249
Counterparty Default Risk	0.2904	0.2317
No Default Risk	0.2860	0.2334



(a) CCCs computed using QV under different default assumptions



(b) CCCs computed using NL under different default assumptions

Figure 5.13: CCCs under different default assumptions

5.4 Bid and Ask Swap Rates and Swaption Prices

5.4.1 Introduction

Interest rate swaps are widely used in the market to hedge against changes in interest rates. An interest rate swap is a derivative instrument between two parties, in which they agree to exchange interest rate cash flows based on a specified notional amount from a fixed rate to a floating rate or *vice versa*. In this section, we are going to apply the $\mathcal{G}$-expectation method for the computations of bid and ask interest rate swap rates and swaption prices. We start by deriving PIDE representations of the risk neutral swap rate and swaption prices , and then model the bid and ask swap rates and swaption prices as solutions to the distorted PIDEs build upon the risk neutral representations.

5.4.2 PIDE Representations of Swap Rates

Consider a continuous time interest rate swap between two parties from time 0 to T. Let r denote the interest rate. The swap rate is defined to be the value K, such that the present value of future cash flows,

$$E(\int_0^T (r(s) - K)e^{-\int_0^s r(u)du} ds),$$

equals to zero. Following [18], the instantaneous interest $r(t)$ is modeled as an Ornstein-Uhlenbeck process driven by a gamma process $g(t)$ defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, P)$. Suppose the initial interest rate is r_0, and

$r(t)$ follows

$$dr = -\kappa r dt + dg, \tag{5.9}$$

where g is a gamma process with Levy density

$$k(x) = \gamma \frac{e^{-\lambda x}}{x}, \quad x > 0. \tag{5.10}$$

The solution to (5.9) is

$$r(t) = r_0 e^{-\kappa t} + \int_0^t e^{-\kappa(t-u)} dg(u). \tag{5.11}$$

Integrating both sides of (5.11) yields

$$\int_0^t r(u)du = r_0 \frac{1 - e^{-\kappa t}}{\kappa} + \int_0^t \frac{1 - e^{-\kappa(t-u)}}{k} dg(u).$$

Let $P(0,t)$ denote $E(e^{-\int_0^t r(u)du})$, we can find the analytical expression for $P(0,t)$ as

$$exp[-r_0 \frac{1 - e^{-\kappa t}}{\kappa} + \int_0^t \int_0^\infty (e^{-\frac{1-e^{-\kappa(t-u)}}{k} y} - 1) k(y) dy du], \tag{5.12}$$

where $k(x)$ is defined in (5.10).

Consider the following expression

$$E(\int_t^T (r(s) - K) e^{-\int_t^s r(u)du} ds | \mathcal{F}_t),$$

which is the time t value of the future cash flow of the interest rate swap with swap rate K. Let's call the value of K that makes the above expression 0 the swap rate at time t, denoted by K_t. Then K_t satisfies

$$K_t = \frac{E(\int_t^T r(s) e^{-\int_t^s r(u)du} ds | \mathcal{F}_t)}{E(\int_t^T e^{-\int_t^s r(u)du} | \mathcal{F}_t)} \tag{5.13}$$

We observed that both the numerator and the denominator of (5.13) depend only on t and $r(t)$. As a result, K_t is also a function of t and $r(t)$. Let $V(t,x)$ and $J(t,x)$ denote the functions for K_t and $E(\int_t^T e^{-\int_t^s r(u)du}|\mathcal{F}_t)$ respectively.

Define a random process M_t as follows,

$$M_t = e^{-\int_0^t r(u)du} V(t, r(t)) E(\int_t^T e^{-\int_t^s r(u)du}|\mathcal{F}_t) + \int_0^t r(s) e^{-\int_0^s r(u)du} ds. \tag{5.14}$$

Substitute (5.13) into (5.14) to get

$$M_t = E(\int_0^T r(s) e^{-\int_0^s r(u)du} ds|\mathcal{F}_t).$$

Therefore, M_t is a martingale. By Itô's formula for jump diffusions, we have

$$\begin{aligned} dM_t &= r(t)e^{-\int_0^t r(u)du}dt + e^{-\int_0^t r(u)du}(-r(t))V(t,r(t))J(t,r(t))dt \\ &\quad + e^{-\int_0^t r(u)du}[\frac{\partial(VJ)}{\partial t}dt + \frac{\partial(VJ)}{\partial x}(-\kappa r(t))dt + VJ(t,r(t)+\Delta) - VJ(t,r(t))]. \end{aligned}$$

Since M_t is a martingale, the drift term of dM_t should be zero. We obtain

$$\begin{aligned} & r(t)e^{-\int_0^t r(u)du} + e^{-\int_0^t r(u)du}(-r(t))V(t,r(t))J(t,r(t)) \\ & + e^{-\int_0^t r(u)du}[\frac{\partial(VJ)}{\partial t}dt + \frac{\partial(VJ)}{\partial x}(-\kappa r(t)) \\ & + \int_0^\infty (VJ(t,r(t)+z) - VJ(t,r(t)))k(z)dz] = 0. \end{aligned}$$

Dividing both sides by the common term $e^{-\int_0^t r(u)du}$ and put x in place of $r(t)$ to get

$$\begin{aligned} J(\frac{\partial V}{\partial t} - \kappa x\frac{\partial V}{\partial x}) + \int_0^\infty (VJ(t,r(t)+z) - VJ(t,r(t)))k(z)dz \\ + V(\frac{\partial J}{\partial t} - \kappa x\frac{\partial J}{\partial x} - xJ) + x = 0. \end{aligned} \tag{5.15}$$

The boundary condition is

$$V(T,x) = x. \tag{5.16}$$

We obtain the following expression for $J(t,x)$ from (5.12),

$$J(t,x) = \int_0^{T-t} e^{-x\frac{1-e^{-\kappa u}}{\kappa} + H(u)} du,$$

where

$$H(u) = \int_0^u \int_0^\infty (e^{\frac{e^{-\kappa x}-1}{k} y} - 1) k(y) dy dx.$$

$\frac{\partial J}{\partial t}$ and $\frac{\partial J}{\partial x}$ can be computed as

$$\frac{\partial J}{\partial t} = -e^{-x\frac{1-e^{-\kappa(T-t)}}{\kappa} + H(T-t)},$$

and

$$\frac{\partial J}{\partial x} = \int_0^{T-t} -\frac{1-e^{-\kappa u}}{\kappa} e^{-x\frac{1-e^{-\kappa u}}{\kappa} + H(u)} du.$$

For $t \in [0,T)$, we could divide both sides of (5.15) by J to get

$$\begin{aligned} \frac{\partial V}{\partial t} - \kappa x \frac{\partial V}{\partial x} + \int_0^\infty (V(t,x+z)\frac{J(t,x+z)}{J(t,x)} - V(t,x))k(z)dz \\ + V(\frac{\frac{\partial J}{\partial t} - \kappa x \frac{\partial J}{\partial x}}{J} - x) + \frac{x}{J} = 0 \end{aligned} \tag{5.17}$$

As t goes to T in (5.17), we can compute the limiting values on both sides. We obtain, on the boundary $t = T$, $V(t,x)$ satisfies

$$\frac{\partial V}{\partial t} - \kappa x \frac{\partial V}{\partial x} + \int_0^\infty (V(t,x+z) - V(t,x))k(z)dz = 0. \tag{5.18}$$

In all, (5.17) and (5.18) describes the PIDE for the risk neutral swap rate, with boundary condition given by (5.16).

5.4.3 Bid and Ask Swap Rates

Following the $\mathcal{G}$-expectation method, we model bid and ask swap rates as the solutions to the distorted PIDEs built upon the PIDE representation of the risk neutral swap rate. We construct the distorted PIDEs by distorting the integral terms in (5.17) and (5.18) using the NL method described in Chapter 4. For the dynamics of the interest rate, we use the parameters $\kappa = 0.1868$, $\lambda = 570.3251$, and $\gamma = 4.7936$. These parameters are obtained by calibrating the interest rate model (5.9) to the pure discount curve on Aug. 15, 2011, with details given in Section 3 in [18]. The time interval is set to $[0, 1]$. We apply NL-distortion to the integral term in the PIDE and solve the PIDEs numerically using the Euler method. We build the space grid on the interval $[0, 0.0495]$ with step size 0.0005. The time step equals to 0.1. We present in Figures 5.14 and Figure 5.15 the bid ask and expected swap rates under distortion levels 0.1 and 0.5 respectively. Table 5.5 compares the bid, ask and expected swap rates at $t = 0, 0.5, 1$, $r_t = 0.02$ and $\gamma = 0.1, 0.5$.

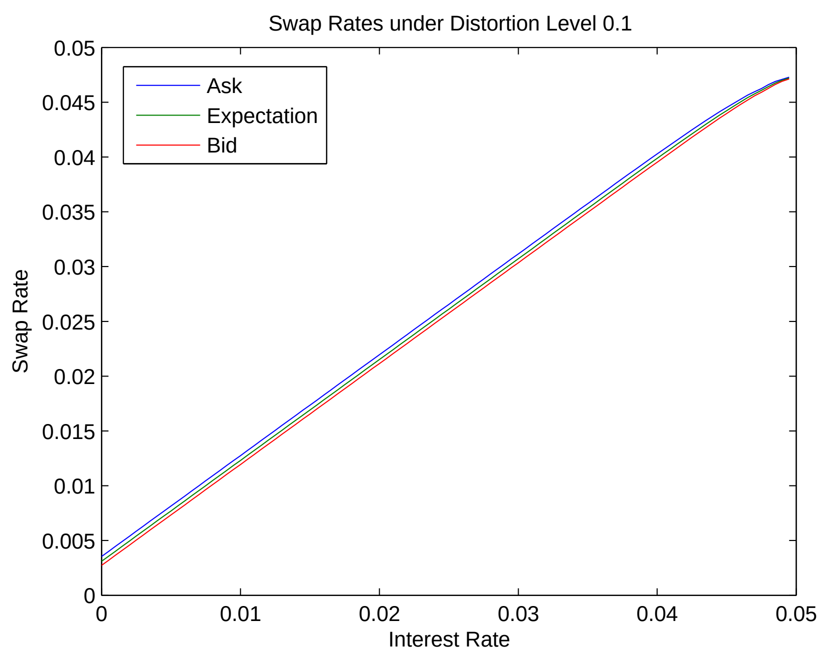


Figure 5.14: Bid, ask, and expected swap rates under distortion level 0.1

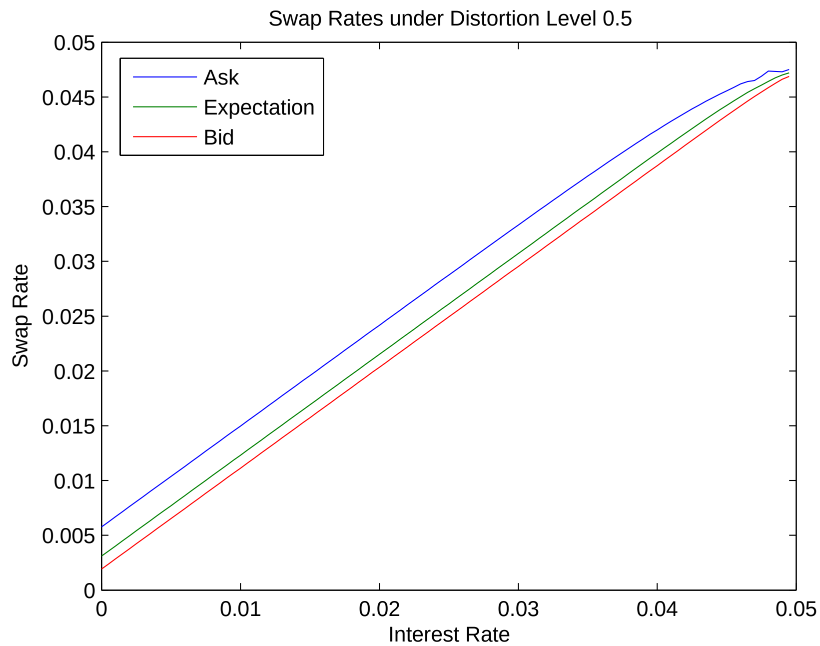


Figure 5.15: Bid, ask, and expected swap rates under distortion level 0.5

Table 5.5: Swap rates at $r = 0.02$, $t = 0, 0.5, 1$ and $\gamma = 0.1, 0.5$

Distortion Level	Time	Bid	Ask	Expectation
0.1	0	0.0211	0.0220	0.0215
	0.5	0.0206	0.0211	0.0209
	1	0.02	0.02	0.02
0.5	0	0.0203	0.0242	0.0215
	0.5	0.0202	0.0224	0.0209
	1	0.02	0.02	0.02

5.4.4 Bid and Ask Prices of Swaptions

Consider a payer swaption on a forward starting swap that starts at time a and ends at time T with strike K^*. The expiration time of this swaption is $t = a$. We are going to derive the bid and ask prices for the swaption. First we need to find a PIDE to describe the risk neutral value of this swaption for $0 \le t \le a$. Let us denote the risk neutral swap rate at time $t = a$ given spot interest rate $r(a)$ by $K(r(a))$. As a result, for $0 \le t \le a$, the risk neutral value of the swaption is

$$E(\int_a^T (r(s) - K^*)e^{-\int_t^s r(u)du} ds \mathbf{1}_{K(r(a))>K*} | \mathcal{F}_t). \tag{5.19}$$

By definition, $K(r(a))$ satisfies

$$E(\int_a^T (r(s) - K(r(a)))e^{-\int_a^s r(u)du} ds | \mathcal{F}_a) = 0,$$

therefore, (5.19) can be rewritten into

$$E(\int_a^T (K(r(a)) - K^*)^+ e^{-\int_t^s r(u)du} ds | \mathcal{F}_t). \tag{5.20}$$

We note that the expression (5.20) should be a function of only t and $r(t)$. Let $V(t, x)$ denote such function, and define a process M_t as follows,

$$M_t = e^{-\int_0^t r(u)du} V(t, r(t)).$$

By Ito's formula, we could get

$$dM_t = e^{-\int_0^t r(u)du}[-r(t)V dt + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}(-\kappa r(t))dt + V(t, r(t-) + \Delta) - V(t-, r(t-))]. \tag{5.21}$$

Since

$$M_t = E(\int_a^T (K(r(a)) - K^*)^+ e^{-\int_0^s r(u)du} ds | \mathcal{F}_t), \tag{5.22}$$

M_t is a martingale, the drift term should equal to zero. We obtain the following PIDE representation of the risk neutral swaption price.

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}(-\kappa x) + \int_0^\infty (V(t, x+z) - V(t, x))k(z)dz - xV = 0, \tag{5.23}$$

with terminal value

$$V(a, x) = (K(x) - K^*)^+ J(a, x). \tag{5.24}$$

We could then generate the bid and ask swaption prices using $\mathcal{G}$-expectation approach. Consider a swaption with maturity $a = 0.5$ and strike $K^* = 0.02$ on a swap that begins at time 0.5 and ends at time 1. We use the same parameters for interest rate as in the last subsection. Figure 5.16 shows the payoff of the swaption at maturity ($t = 0.5$). Figure 5.17 and Figure 5.18 present the bid, ask and expected

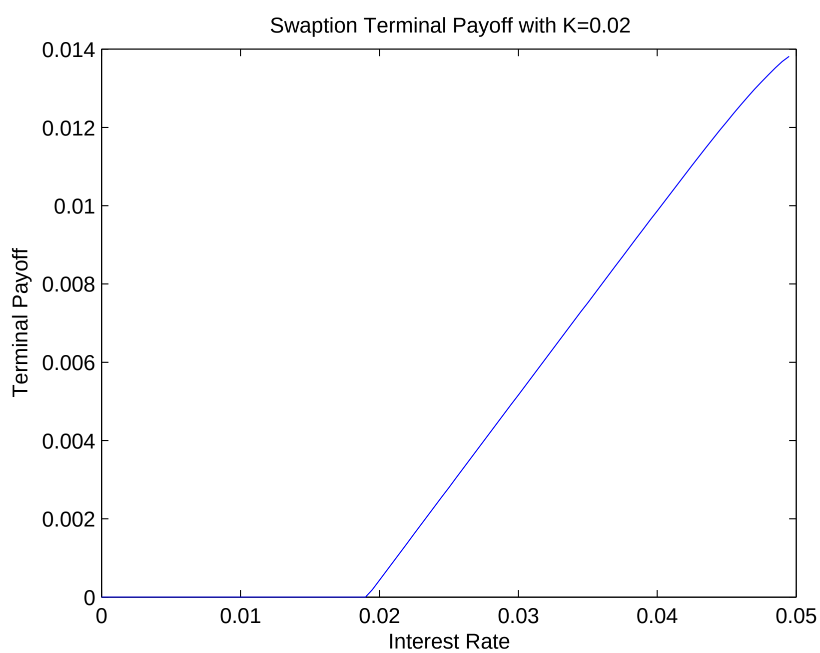


Figure 5.16: Terminal payoff of a swaption with strike 0.02

prices of the swaption at distortion levels 0.1 and 0.5 respectively.

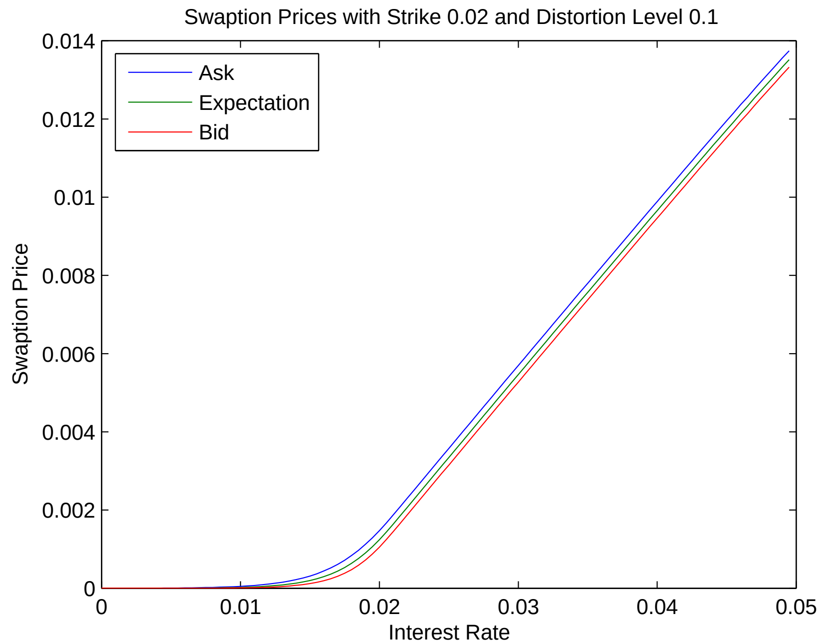


Figure 5.17: Bid, ask, and expected prices of a swaption at distortion level 0.1

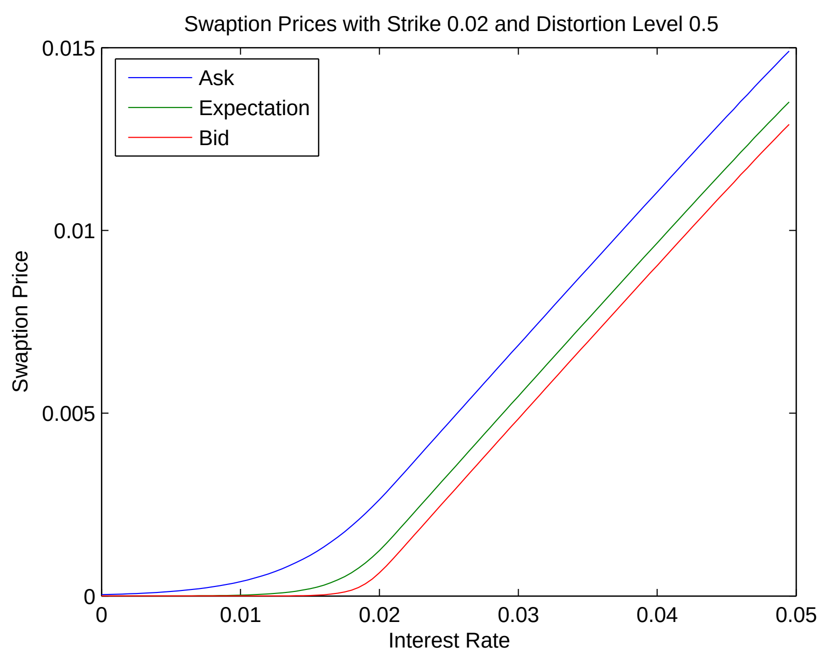


Figure 5.18: Bid, ask, and expected prices of a swaption at distortion level 0.5

Table 5.6 compares the bid, ask and expected prices of the swaption at different spot interest rates under distortion levels 0.1 and 0.5. Finally, in Figure 5.19, we present the implied volatilities of the bid, ask and expected swaption prices computed at distortion level 0.01.

Table 5.6: Swaption prices at different spot interest rates and distortion levels 0.1 and 0.5

Distortion Level	Interest Rate	Bid	Ask	Expectation
0.1	0.01	0	0	0
	0.02	0.0011	0.0015	0.0012
	0.03	0.0053	0.0057	0.0055
	0.04	0.0094	0.0098	0.0096
0.5	0.01	0	0.0004	0
	0.02	0.0006	0.0026	0.0012
	0.03	0.0048	0.0068	0.0055
	0.04	0.0090	0.0107	0.0096

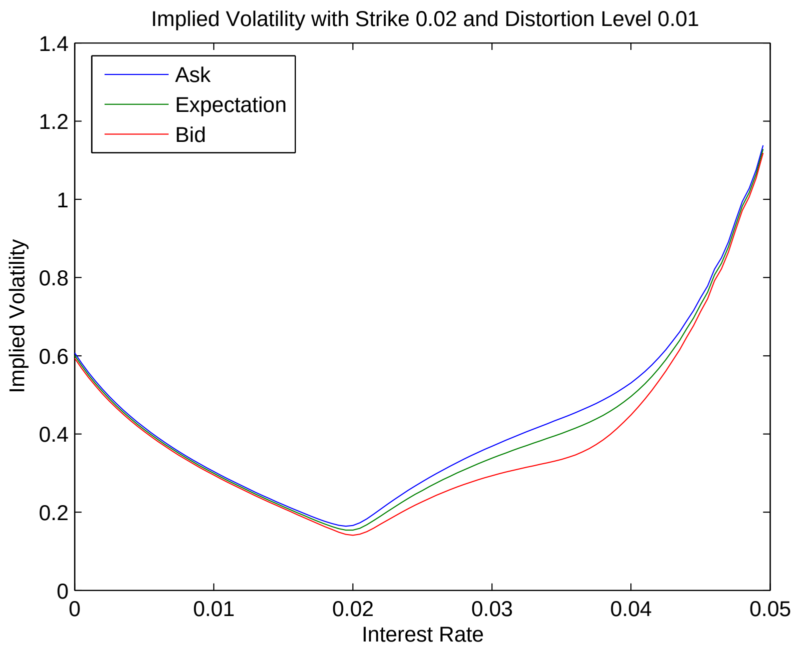


Figure 5.19: Bid, ask, and expected implied volatilities at distortion level 0.01

Bibliography

[1] Alvarez, O., and Tourin, A. Viscosity solutions of nonlinear integro-differential equations. In *Annales de l'Institut Henri Poincaré. Analyse non linéaire* (1996), vol. 13, Elsevier, pp. 293–317.

[2] Amihud, Y., and Mendelson, H. Dealership market: market making with inventory. *Journal of Financial Economics*, 8 (1980), 31–53.

[3] Applebaum, D. *Lévy processes and stochastic calculus.* Cambridge university press, 2009.

[4] Arrow, K. The role of securities in the optimal allocation of risk bearing. *The Review of Economic Studies 31*, 2 (Apr 1964), 91–96.

[5] Arrow, K., and Debreu, G. Existence of an equilibrium for a competitive economy. *Econometrica 22*, 3 (1954), 265–290.

[6] Briand, P., Delyon, B., and Mémin, J. Donsker-type theorem for bsdes. *Electronic Communications in Probability* (2001).

[7] Carr, P., and Madan, D. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2 (1999), 61–73.

[8] Cherny, A., and Madan, D. New measures for performance evaluation. *Review of Financial Studies 22*, 7 (2009), 2571–2606.

[9] Choi, J., Salandro, D., and Shastri, K. On the estimation of bid ask spread: theory and evidence. *Journal of Financial and Quantitative Analysis 23* (1988), 219–230.

[10] Cohen, S., and Elliott, R. Solutions of backward stochastic differential equations on Markov chains. *Communications on Stochastic Analysis 2*, 2 (2008), 251–262.

[11] Cohen, S., and Elliott, R. Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. *The Annals of Applied Probability 20*, 1 (2010), 267–311.

[12] Cohen, S., and Elliott, R. A general theory of finite state backward stochastic difference equations. *Stochastic Processes and their Applications 120*, 4 (2010), 442–466.

[13] Copeland, T., and Galai, D. Information effects on the bid-ask spread. *Journal of Finance 38* (1983), 1457–1469.

[14] Crandall, M. G., Ishii, H., and Lions, P.-L. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society 27*, 1 (1992), 1–67.

[15] Demsetz, H. The cost of transacting. *Quarterly Journal of Economics 82*, 1 (1968), 33–53.

[16] Dybvig, P., and Ross, S. Arbitrage. *The New Palgrave: A Dictionary of Economics 1* (1987), 100–106.

[17] Easley, D., Kiefer, N., O'Hara, M., and Paperman, J. Liquidity, information, and infrequently traded stocks. *Journal of Finance 51* (1996), 1405–1436.

[18] Eberlein, E., Madan, D., Pistorius, M., and Yor, M. A simple stochastic rate model for rate equity hybrid products. *Applied Mathematical Finance*, ahead-of-print (2013), 1–28.

[19] Elliott, R., Aggoun, L., and Moore, J. *Hidden Markov Models: Estimation and Control*, vol. 29. New York: Springer-Verlag, 1995.

[20] Flannery, M., Kwan, S., and Nimalendran, M. The 2007-2009 financial crisis and bank opaqueness. *Journal of Financial Intermediation 22*, 1 (Jan 2013), 55–84.

[21] George, T., Kaul, G., and Nimalendran, M. Estimation of the bid ask spreads and its components: a new approach. *Review of Financial Studies 4* (1991), 623–656.

[22] Glosten, L., and Harris, L. Estimating the components of the bid ask spread. *Journal of Financial Economics 21* (1988), 123–142.

[23] Glosten, L., and Milgrom, P. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14* (1985), 71–100.

[24] Heston, S. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*, 2 (1993), 327–343.

[25] Huang, R., and Stoll, H. Market microstructure and stock return predictions. *The Review of Financial Studies 7* (1994), 179–213.

[26] Huang, R., and Stoll, H. The components of the bid-ask spread: a general approach. *The Review of Financial Studies 10*, 4 (1997), 995–1034.

[27] Hunt, G. A. *Martingales et Processus de Markov*, vol. 1. Dunod, 1963.

[28] Ingersoll, J. J. *Theory of financial decision making*, first ed. Rowman and Littlefiled, New Jersey, 1987.

[29] Jakubowski, A., and Słomiński, L. Extended convergence to continuous in probability processes with independent increments. *Probability Theory and Related Fields 72*, 1 (1986), 55–82.

[30] Kurtz, T. Lectures on stochastic analysis.

[31] Kyle, A. Continuous auctions and insider trading. *Econometrica 53*, 1315-1335 (1985).

[32] Madan, D. Capital requirements, acceptable risks and profits. *Quantitative Finance 7* (2009), 767–773.

[33] Madan, D., Carr, P., and Chang, E. The variance gamma process and option pricing. *European Finance Review 2*, 1 (1998), 79–105.

[34] Madan, D., and Cherny, A. Markets as a counterparty: an introduction to conic finance. *International Journal of Theoretical and Applied Finance* (Aug 2010), 1149–1177.

[35] Madan, D., Pistorius, M., and Schoutens, W. The valuation of structured products using Markov chain models. *Quantitative Finance 13*, 1 (2013), 125–136.

[36] Madan, D., and Schoutens, W. Structured products equilibria in conic two price markets. *Mathematics and Financial Economics 6* (2012), 37–57.

[37] Madan, D., and Seneta, E. The variance gamma (v.g.) model for share market returns. *The Journal of Business 63*, 4 (Oct 1990), 511–524.

[38] Madan, D., Wang, S., and Heckman, P. A theory of risk for two price market equilibria.

[39] Madan, D. B. From credit valuation adjustments to credit capital commitments. *Quantitative Finance 12*, 6 (2012), 839–845.

[40] Madhavan, A., Richardson, M., and Roomans, M. Why do security prices change? a transaction-level analysis of nyse stocks. *The Review of Financial Studies 10*, 4 (1997), 1035–1064.

[41] Mémin, J. Stability of doob-meyer decomposition under extended convergence. *Acta Mathematicae Applicatae Sinica 19*, 2 (2003), 177–190.

[42] Mijatović, A., and Pistorius, M. Continuously monitored barrier options under Markov processes. *Mathematical Finance 23*, 1 (2013), 1–38.

[43] Nualart, D., and Schoutens, W. Backward stochastic differential equations and Feynman-Kac formula for Lévy processes, with applications in finance. *Bernoulli 7*, 5 (2001), 761–776.

[44] Pardoux, E., and Peng, S. Adapted solution of a backward stochastic differential equation. *System and Control Letters*, 14 (1990), 55–61.

[45] Peng, S. Backward sde and related g-expectation. *Pitman Research Notes in Mathematics Series* (1997), 141–160.

[46] Peng, S. G-expectation, g-brownian motion and related stochastic calculus of ito type. *Stochastic Analysis and Applications* (2007).

[47] Ren, Y.-F., and Tian, F.-J. On the Rosenthal's inequality for locally square integrable martingales. *Stochastic Processes and their Applications 104*, 1 (2003), 107–116.

[48] Roll, R. A simple implicit measure of the effective bid ask spread in an efficient market. *Journal of Finance 39* (1984), 1127–1139.

[49] Schoutens, W. *Lévy Processes in Finance: Pricing Financial Derivatives.* John Wiley & Sons Inc., 2003.

[50] Skorokhod, A. Limit theorems for stochastic processes. *Theory of Probability & Its Applications 1*, 3 (1956), 261–290.

[51] Stoll, H. The supply of dealer services in securities markets. *Journal of Finance 33*, 4 (Sep 1978), 1133–1151.

[52] Stoll, H. Inferring the components of the bid ask spread: theory and emipirical tests. *Journal of Finance 44* (1989), 115–134.