THE UNIVERSITY OF CALGARY

European and Swing Option Pricing Under Mean-Reverting Jump Diffusion Models

by

Ouyang, Yuyuan

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

July, 2007

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "European and Swing Option Pricing Under Mean-Reverting Jump Diffusion Models" submitted by Ouyang, Yuyuan in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

Supervisor, Dr. A. Swishchuk
Department of Mathematics and Statistics

Dr. K. C. Cheung
Department of Mathematics and Statistics

Dr. A. Lehar
Haskayne School of Business

Jly 24/07

Date

Abstract

This thesis is devoted to European and Swing option pricing under Pilipović's mean-reverting models, and the associated jump-diffusion model. Currently there is no available analytical option pricing formula under Pilipović models, even the general moments of the models are difficult to calculate. The results of the thesis include an explicit European option pricing formula under the one factor Pilipović's model, and a multinomial tree method, which is developed to price swing options under Pilipović's model with jumps. Here the multinomial tree method is also applicable for other jump-diffusion models. A Monte Carlo simulation method is implemented for the jump diffusion model. The calibration methods of Pilipović's models is also discussed in the thesis.

Acknowledgements

I would like to give my thanks to many individuals, for their generous help in my process of preparing this thesis and in my life.

First and foremost I would like to thank my superviosr Dr. Anatoliy Swishchuk, for his tremendous support, invaluable advice, and great patience, not only on the thesis, but also on my graduate studies and my research life. His guidance was of inestimable worth.

I would also wish to give thanks to my colleagues, Lu Zhao and Matthew Lyle, for their assistance in the project, and Hua Li, Leunglung Chan, Guanghui Quan, for there advices in my graduate studies.

Most importantly, my sincere gratitude goes to my families, for their endless love and boundless support. Thank you for giving me the great confidence in my research and my life.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Energy Markets

Traditional money markets stand for several years. There are many kinds of money markets, such like the interest rate and equity markets, which are all quite mature. Compared to money markets, the energy markets are relatively new, and there are many differences between these two kinds of markets, which make the energy markets harder to model, and therefore the research on energy markets is very important. For example, there can be various regulations in energy markets, while in money markets there are little. The seasonality, which is barely seen in money markets, is an essential part in energy markets. For example, in natural gas markets and electricity markets, there are strong seasonal patterns in market data. The low liquidity and high frequency of events, and the impact of storage also make the energy markets different from money markets. Because of the difference between money markets and energy markets, the framework needs to be changed in the modeling of energy assets.

Most of the characteristics of energy markets is covered in Pilipović's book on energy derivatives (see [32]).

1.1.1 Characteristics in Energy Market Data

Statistical Analysis

Natural gas is bought and sold through a large and integrated continental market with multiple pricing hubs. The main delivery hub in north America is known as the Henry Hub in Louisiana, which is the pricing hub for natural gas traded on the New York

Mercantile Exchange (NYMEX). The Alberta natural gas trading price is closely related to the Henry Hub prices, which is based on a virtual trading hub (AECO) on Natural Gas Exchange system (NGX). In figure 1.1 and figure 1.2 we can see the historical daily natural gas prices in Henry Hub (Jan 1991-Dec 2005) and in AECO (Jan 1994-Dec 2005). We can see that the price paths in both figures are very similar, which is result from the relationship between the two hubs.

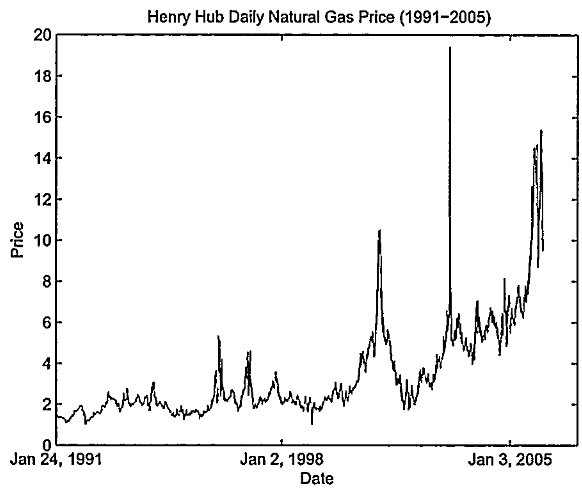

Figure 1.1: Historical Natural Gas Price in Henry Hub (Jan 1991-Dec 2005)

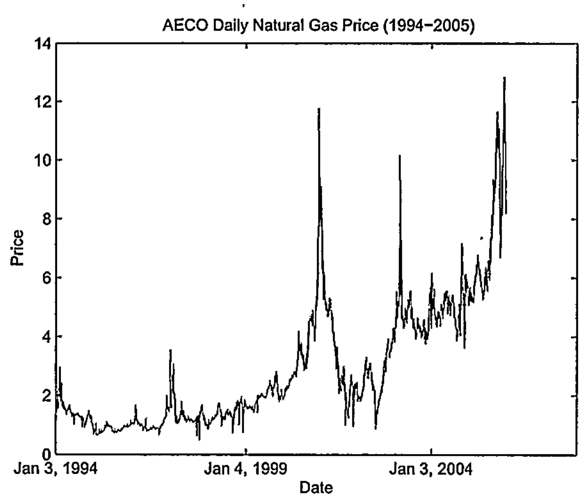

Figure 1.2: Historical Natural Gas Price in AECO (Jan 1994-Dec 2005)

If we look at the histogram of the natural gas prices, we can see high kurtosis and positive skewness in the histogram, which are common in energy markets.

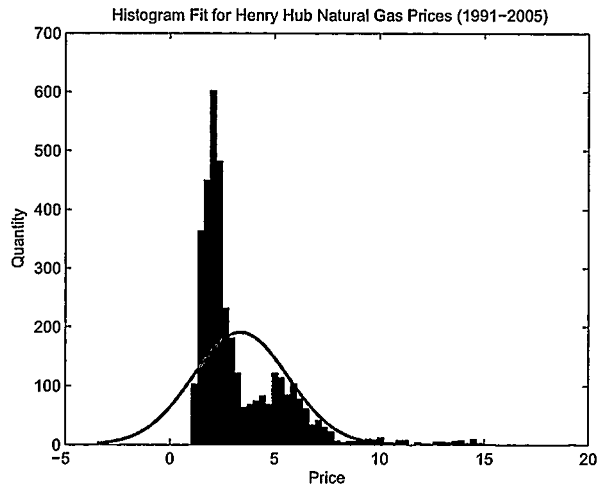

Figure 1.3: Histogram and Normal Distribution Fit for Henry Hub Prices

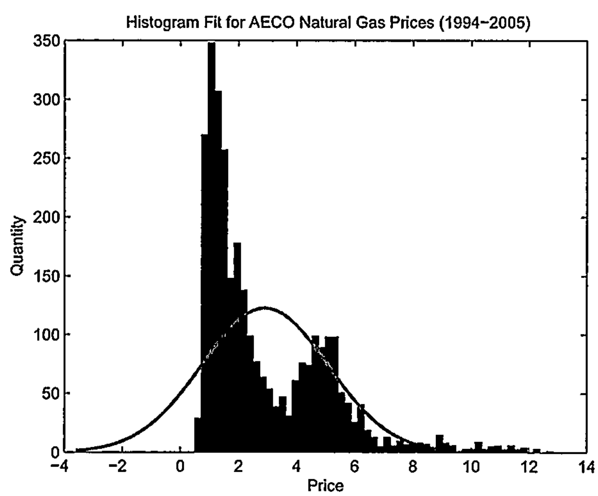

Figure 1.4: Histogram and Normal Distribution Fit for AECO Prices

Now if we take logprice for analysis, we will see that they are not normally distributed, which is an essential assumption in Black-Scholes modeling in money markets. Therefore, Black-Scholes model is not proper for model asset prices in energy markets.

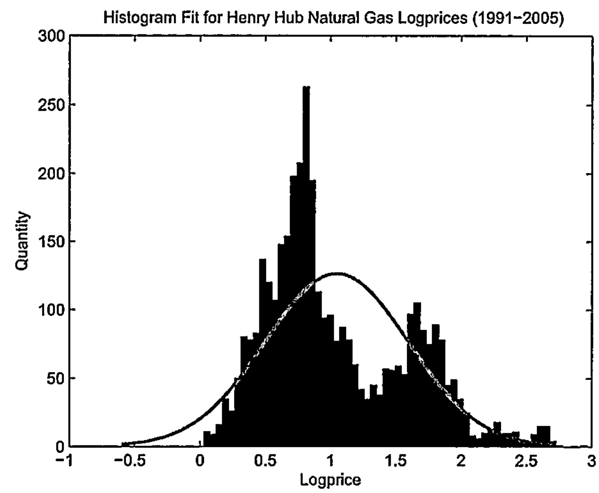

Figure 1.5: Histogram and Normal Distribution Fit for Henry Hub Logrices

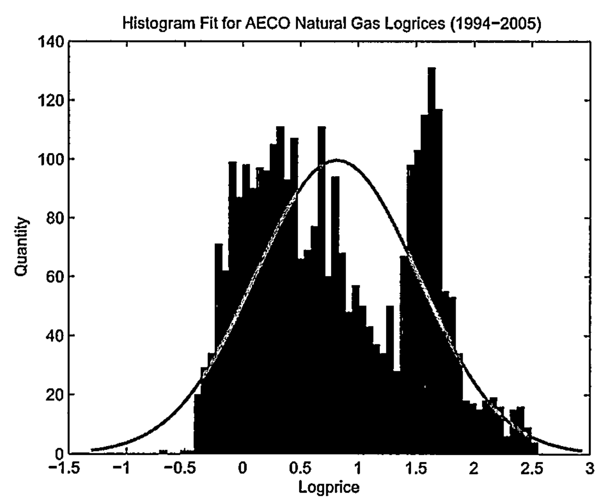

Figure 1.6: Histogram and Normal Distribution Fit for AECO Logprices

Mean-Reversion

The Black-Scholes framework for spot price model is the most famous model in mathematical finance research, and there have already been several mature results derived under this framework. However, the Black-Scholes model is not applicable in energy markets, since it requires the spot prices to be a lognormal process. As we analyzed above, in energy markets, the spot price data has high kurtosis and non-zero skewness, which are different from price data that can be derived under Black-Scholes framework. In fact, the mean-reverting price model turns out to be the best model in capturing the distribution of energy prices.

In a mean-reverting model, there is an equilibrium level around which the underlying mean-reverting variables oscillate. When the underlying variables of the model get away from the equilibrium level, we will notice them in the later time that they go back to the level again. The equilibrium level varies in different models. The level may be a constant, or a deterministic curve, or even itself a stochastic process. There is a drift term in the model which controls the speed of the variables to go back. There will also be noises and discrete jumps in the model. However, the expectation of variables will still converge to the equilibrium level.

We can use the term half-life to calibrate the mean-reverting speed of a mean-reverting process. Similar to the meaning in physics (the expected time for a given amount of substance to decay to half of its mass), the half-life of a mean-reverting model is the expected time that the variable in the model to go half way back to the equilibrium level (see [10]).

The underlying mean-reverting variable in energy markets can be the price, the log-price, or the volatility, etc. Here in the thesis we consider the price to be mean-reverting. The mean-reverting models will be introduced in chapter 2.

1.1.2 Impact of Demand Drivers in Energy Markets

In energy markets, the impact from the demand side can not be ignored. Different from money markets, there are convenience yields and seasonality in energy markets, and both of which result from the impact of demand drivers.

Convenience Yield

For industrial users, they always prefer to keep there plants running, since the cost of shutting down and restarting is relatively higher than constantly running. To maintain constant production and avoid certain cases such like the lack of available supply, they are willing to pay a premium to have the necessary energy for running their plants. The premium they pay is called the convenience yield. The convenience yield can be seen as the value of production by purchasing energy before the users run out of energy supply. There are several factors which have effect on the value of convenience yield, including the benefit of smooth production, storage cost, etc.

Seasonality

The demand of energy can be largely changed at different times. This is mainly from the aggregate consumption of residential users. For example, in electricity markets, the consumption of residential users at midnight are generally lower, since most utilities in houses are shut down. The consumption of natural gas in winter will be higher than the case in summer, due to the demand of heat in winter. Because of the impact of seasonality, we can see significant patterns in price data.

1.1.3 Exotic Contracts in Energy Markets

Due to the various needs of end users in energy markets, there are lots of sophisticated energy contracts in energy markets. Average of prices, barriers, customized delivery requirements, etc., are commonly used in energy contracts. These kinds of typical energy contracts are of great challenge to mathematical finance analysis in energy markets. Most of these kinds of contracts are considered as "exotic contracts", compared to the relatively simpler contracts, like forwards, vanilla options, swaps, swaptions, etc., in money markets.

It is always hard to achieve explicit analytical solutions to the price of exotic options, and they are often priced under numerical methods. Pricing of exotic options is important in energy market trades. The swing options, which we will discuss in the thesis, is a kind of exotic options used mainly in naturally gas markets.

1.2 Financial Derivatives

In finance, a derivative, also called contingent claim, is a financial instrument that is based on an underlying asset's value; rather than trade or exchange the asset at present, market participants enter into an agreement to exchange money, assets or some other value at some future date based on the terms of the derivatives. Examples of assets could be anything from bar of a gold, to a stock, or even an interest rate. Common derivatives include put and call options, forward contracts, future contracts, swaptions, etc. A simple example is a futures contract: an agreement to exchange the underlying asset (or equivalent cash flows) at a future date. The exact terms of the derivative (the payments between the counterparties) depend on, but may or may not exactly correspond to, the behaviour or performance of the underlying asset. The creation of financial derivatives is mainly on the purpose of hedging, but the financial derivatives can be used for both speculation and hedging. Some derivatives can be traded on exchanges, while others are developed as private contracts between the financial agents and clients.

The widely use of financial derivatives leads to a mathematical problem: pricing of derivatives. Although the price of exchangeable financial derivatives can be just driven by the relationship of supply and demand, there are also certain contracts that works privately between two financial institutions. In the latter case, the two sides will be hoping to achieve a reasonably "fair" price for the contract. Even in the former case, when exchangeable derivative are priced due to the supplies and demands, there still exists certain arbitragers that are willing to estimate a relatively "fair" price for the public contracts, to help them to gain benefit from the difference between the real price and estimated price.

Under certain conditions, there are theories which provide ways to price financial derivatives, for example by using the expectation of the discounted payoff. We will discuss the condition and theory in section 4.2.

1.2.1 Options, European Options, Swing Options, Option Pricing

There are many types of financial instruments that are grouped under the term derivatives, and we concentrate on options in the thesis. Options are contracts where one party agrees to pay a fee to another for the right (but not the obligation) to buy something from or sell something to the other. Generally, there are several basic components for an option contract.

The asset being traded in the option contract is called the underlying asset.

The fixed price, which is decided in the contract as the price for the holder to trade the underlying asset, is the strike price, or the exercise price.

The given dates for the trade to be finished is termed as the maturity dates, or expiry dates.

The European option is a kind of options which is widely used. The European options provides an opportunity, but not an obligation, for the holder to buy (European call option) or sell (European put option) a given asset at the strike price. The holder can only exercise his right on the expiry date.

Options like European options are very simple, well understood and traded, and therefore named as "vanilla options".

Swing options are different from vanilla options, which have various versions in practice. Generally, swing options provide option holders multiple exercise rights, and the holders can choose to exercise the rights on several given exercise dates before the end of the contract.

There are several ways in option pricing. For example, analytical solution through expected payoff or partial differential equations. In the cases when an analytical solution can not be found, or the analytical form is to complicated to be practical, numerical methods can be applied, such like the finite difference method, finite element method, Monte Carlo simulation, binomial tree method, etc.

There is an important component in option pricing, which is the risk-neutral measure. The risk-neutral measure is a probability measure, under which the non-arbitrage price of a derivative at present is equal to the risk-neutral measured expected value of the future payoff of the derivative discounted at the risk-free rate. All the option pricing methods are based on the prerequisite that the model is risk-neutral.

By using Girsanov theorem, it is possible to transform the measure from physical measure to risk-neutral measure. Girsanov theorem is a theorem on change of measure in probability theory, which provides the way of the conversion (see [30]). The Radon-Nikodym derivative can be calculated through Girsanov theorem, which is the change of density of one measure to the other (see [30], [33]). However, the Girsanov theorem can not be applied to the models in this thesis, and an alternative method is used (see section 2.2).

Chapter 2

Mean-Reverting Stochastic Models

As introduced in chapter 1, There is mean-reverting patterns in the behavior of the prices of energy assets, and we can assume that there exists equilibrium prices in the asset price curves. In this chapter we will discuss three kinds of mean-reverting models, which will be used in this thesis: one factor, two factor, and one factor with jumps. D. Pilipovic first provided the one factor and two factor model in [32], with their solutions and some related analysis. There is also a solution to the one factor model via change-of-time method in [38] by A. Swishchuk. Some other models used in energy market research will also be introduced in of this chapter.

2.1 Mean-Reverting Models without Jumps

In this section we give a brief introduction of some widely used mean-reverting models without jumps.

2.1.1 One Factor Models without Jumps

Vasicek's Model

Vasicek's model is used to describe the evolution of interest rates in financial markets. It was introduced in 1977 by O. Vasicek ([39]). The underlying process is the Ornstein-Uhlenbeck process, which is an one-factor mean-reverting process.

The stochastic differential equation has the form

$$dr_t = a(b - r_t)dt + \sigma dW_t.$$

The solution of equation (2.1.1) is

$$r_t = r_0 e^{at} + b(1 - e^{-at}) + \int_0^t \sigma e^{a(t-s)} dW_s.$$

The expectation $\mathbb{E}[r_t]$ is

$$\mathbb{E}[r_t] = r_0 e^{-at} + b(1 - e^{-at}),$$

and the variance satisfies

$$\mathrm{Var}[r_t] = \frac{\sigma^2}{2a}(1 - e^{-2at}).$$

We can see that

$$\lim_{t\to\infty} \mathbb{E}[r_t] = b,$$

which implies that r_t is following a mean-reverting process.

Schwartz's One Factor Model

E. S. Schwartz provide three models in his paper on commodity prices ([34]): one factor, two factor, and three factor models. Here we give an introduction of his one factor model.

In one factor model he assumed that the logprice follows the Ornstein-Uhlenbeck process. The underlying asset price S_t has the form

$$dS_t = \alpha(L - \ln S_t)S_t dt + \sigma S_t dW_t,$$

Where $\alpha > 0$ is the drift speed, L is the equilibrium level. α, L, σ are constant. If we let $X_t = \ln S_t$, by Itô's lemma (see appendix A.1.4) we will have

$$dX_t = \alpha(\mu - X_t)dt + \sigma dW_t,$$

where

$$\mu = L - \frac{\sigma^2}{2\alpha}.$$

For risk-neutral measure, he assumes that X_t follows

$$dX_t = \alpha(\mu^* - X_t)dt + \sigma dW_t$$

under risk-neutral measure, where $\mu^* = \mu - \lambda$, λ is the market price of risk.

This one factor model is widely used, since it can be analytically solved, and an explicit European option pricing formula can be calculated (see [34]).

2.1.2 Multi-Factor Models without Jumps

Schwartz's two factor model and three factor model are also widely used. Here we give an introduction of these models.

Schwartz's Two Factor Model

In Schwartz's two factor model ([34]), the asset price follows

$$\begin{cases} dS_t &= (\mu - \sigma_t)S_t dt + \sigma_1 S_t dW_t^1, \\ d\sigma_t &= \kappa(\alpha - \sigma_t)dt + \sigma_2 dW_t^2, \end{cases} \tag{2.1}$$

where σ_t follows the Ornstein-Uhlenbeck process, and the two Wiener process W_t^1 and W_t^2 are correlated with

$$dW_t^1 dW_t^2 = \rho dt.$$

Schwartz provided the solution and European option pricing formula of this model (see [34]).

Schwartz's Three Factor Model

In this model Schwartz assumes that the asset price follows (see [34])

$$\begin{cases} dS_t &= (r_t - \sigma_t)S_t dt + \sigma_1 dW_t^1, \\ d\sigma_t &= \kappa(\alpha - \sigma_t)dt + \sigma_2 dW_t^2, \\ dr_t &= a(m - r_t)dt + \sigma_3 dW_t^3, \end{cases} \tag{2.2}$$

where

$$dW_t^1 dW_t^2 = \rho_1 dt, \; dW_t^2 dW_t^3 = \rho_2 dt, \; dW_t^1 dW_t^3 = \rho_3 dt.$$

Schwartz provided the solution and European option pricing formula of this model (see [34]).

2.2 Pilipović's One Factor Model (Model 1F)

Pilipović introduced the mean-reverting one factor model and two factor model in [32]. In one factor model, it is assumed that the equilibrium price is a constant L, and the spot price S_t follows the stochastic differential equation

$$dS_t = \alpha(L - S_t)dt + \sigma S_t dZ_t \tag{2.3}$$

In equation (2.3), Z_t is a standard Wiener process, L is the equilibrium price, α is the mean reverting speed, σ is the volatility. Here $\alpha > 0$ and $\sigma > 0$.

Note that the difference between model 1F and Vasicek's model is that there is an additional term in the diffusion part in model 1F.

We can easily see that the spot price process defined by equation (2.3) is a mean-reverting process: for any fixed S_t such that $S_t > L$, $\mathbb{E}[dS_t] < 0$, therefore S_t tends to decrease, and when $S_t < L$, $\mathbb{E}[dS_t] > 0$, thus S_t tends to be increasing. Therefore, S_t will be oscillating around the constant line L.

For the solution of this stochastic differential equation, we have (see [32], [38])

Theorem 2.2-1. *In equation(2.3), the explicit solution of S_t with initial value S_0, is the time integral*

$$S_t = S_0 e^{-\alpha t} e^{\sigma Z_t - \frac{\sigma^2 t}{2}} + \alpha L e^{-\alpha t} e^{\sigma Z_t - \frac{\sigma^2 t}{2}} \int_0^t e^{\alpha s} e^{-\sigma Z_s + \frac{\sigma^2 s}{2}} ds, \tag{2.4}$$

and

$$\mathbb{E}[S_t] = e^{-\alpha t}(S_0 - L) + L.$$

Model 1F is not risk-neutral, and there can be infinitely many Radon-Nikodym derivatives for the process S_t, and it can not be simply transformed to a risk-neutral process by applying Girsanov theorem. Therefore, there could be arbitrage in the market if the spot price is under model 1F. In practical use, to make it risk-neutral, the market price of risk is introduced. For a given market under model 1F, with the market price of risk λ, the risk-neutral stochastic differential equation will be (see [32])

$$dS_t = (\alpha(L - S_t) - \lambda\sigma S_t)dt + \sigma S_t dZ_t, \tag{2.5}$$

where $\lambda\sigma$ is the risk premium. If we let $\alpha^* = \alpha + \lambda\sigma$, $L^* = \frac{\alpha L}{\alpha+\lambda\sigma}$, we will have a risk-neutral spot price model

$$dS_t = \alpha^*(L^* - S_t)dt + \sigma S_t dZ_t. \tag{2.6}$$

We can see that equation (2.6) has the same form as equation (2.3).

The motivation of using one factor models is to model the mean-reverting spot price behavior in energy markets. However, there are some problems when implementing one factor models to energy markets. One problem is that one factor models are not quite able to catch the spikes in the price behavior. Therefore, two factor models and jump models are introduced to fix this problem.

2.3 Pilipović's Two Factor Model (Model 2F)

In the two factor case by Pilipović, the equilibrium price L_t follows a stochastic process, and the spot price S_t follows

$$\begin{cases} dS_t = \alpha(L_t - S_t)dt + \sigma S_t dZ_t, \\ dL_t = \mu L_t dt + \xi L_t dW_t. \end{cases} \tag{2.7}$$

In equation (2.7), L_t is the stochastic process of the equilibrium price, which is a geometric Brownian motion with drift μ and volatility ξ. W_t is another Wiener process, which is independent of Z_t. When $\mu = \xi = 0$, the model reduces to model 1F.

If we replace L_t by some other processes other than a geometric Brownian motion, we can achieve several other kinds of models, for example we can use Ornstein-Uhlenbeck process. It is still not quite clear what kind of stochastic process can be the best model for the equilibrium price in energy markets, and there are several kinds of models in modeling L_t.

Below we provide the solution of this model ([32]).

Theorem 2.3-1. *With initial value S_0 for S_t, initial value L_0 for L_t, The explicit solution of equation (2.7) is the time integral*

$$\begin{aligned} S_t &= S_0 e^{-(\alpha+\frac{\sigma^2}{2})t} e^{\sigma Z_t} \\ &+ \alpha L_0 e^{-(\alpha+\frac{\sigma^2}{2})t} e^{\sigma Z_t} \int_0^t e^{(\mu-\frac{\xi^2}{2})s+\xi W_s} e^{(\alpha+\frac{\sigma^2}{2})s-\sigma Z_s} ds \end{aligned} \tag{2.8}$$

and

$$\mathbb{E}\left[S_t\right] = S_0 e^{-\alpha t} S_0 + \frac{\alpha}{\alpha+\mu} L_0 (e^{\mu t} - e^{-\alpha t}) \tag{2.9}$$

To estimate the statistical property of Model 1F and Model 2F, we can calculate the moments of the spot prices. For the moments of this model, first we have the following theorem

Theorem 2.3-2. *If S_t follows the stochastic differential equation (2.7) with equilibrium price L_t, for any $t > 0$, for any positive integer n and positive integer l, $\mathbb{E}\left[S_t^n L_t^l\right]$ follows*

$$\begin{aligned} &\mathbb{E}\left[S_t^n L_t^l\right] \\ = &\sum_{k=0}^{n} \sum_{i=0}^{k} \frac{(-1)^n n! \alpha^{n-k+i} S_0^{k-i} L_0^{n+l-k+i} e^{f(k,n+l-k)t}}{(k-i)! \prod_{j=k-i, j\neq k}^{n} (f(j,n+l-j) - f(k,n+l-k))} \end{aligned} \tag{2.10}$$

where

$$f(n,l) = \frac{1}{2}n(n-1)\sigma^2 - n\alpha + \mu l + \frac{1}{2}l(l-1)\xi^2.$$

For the proof of theorem 2.3-2, see appendix B.1.

Now when $l = 0$ in theorem 2.3-2, we have

Corollary 2.3-3. *The nth moment of* S_t *in equation (2.7) is*

$$\begin{aligned} & \mathbb{E}\left[S_t^n\right] \\ = & \sum_{k=0}^{n}\sum_{i=0}^{k} \frac{(-1)^n n! \alpha^{n-k+i} S_0^{k-i} L_0^{n-k+i} e^{f(k,n-k)t}}{(k-i)! \prod_{j=k-i, j\neq k}^{n} (f(j,n-j) - f(k,n-k))}, \end{aligned} \tag{2.11}$$

where

$$f(n,l) = \frac{1}{2}n(n-1)\sigma^2 - n\alpha + \mu l + \frac{1}{2}l(l-1)\xi^2$$

In corollary (2.3-3), consider the case when $n = 1$, we can get

$$\mathbb{E}\left[S_t\right] = S_0 e^{-\alpha t} S_0 + \frac{\alpha}{\alpha+\mu} L_0 (e^{\mu t} - e^{-\alpha t}),$$

which is exactly the same as equation (2.9).

Similar to model 1F, we can also introduce the market price of risk for model 2F, to make the model of S_t in equation (2.7) be risk-neutral. Let the market price of risk be λ, we can get the risk neutral model

$$\begin{cases} dS_t &= (\alpha(L_t - S_t) - \lambda\sigma S_t)dt + \sigma S_t dZ_t \\ dL_t &= (\mu - \lambda\xi)L_t dt + \xi L_t dW_t \end{cases} \tag{2.12}$$

or

$$\begin{cases} dS_t &= \alpha^*(L_t^* - S_t)dt + \sigma S_t dZ_t \\ dL_t &= \mu^* L_t dt + \xi L_t dW_t \end{cases} \tag{2.13}$$

where $\alpha^* = \alpha + \lambda\sigma$, $L_t^* = \frac{\alpha L_t}{\alpha+\lambda\sigma}$, $\mu^* = \mu - \lambda\xi$. This has the same form as equation (2.7).

2.4 Pilipović's One Factor Model with Jumps (Model 1FJ)

If we introduce a jump process in equation (2.3), we can get a new model

$$dS_t = \alpha(L - S_t)dt + \sigma S_t dZ_t + U_t S_t dP_t \tag{2.14}$$

Here P_t is a Poisson process with intensity η, U_t is the distribution of the proportional jump size. Depends on different markets, U_t may have different distributions.

Note that here we assume the jump part to be of the form $U_tS_tdP_t$, but there can still be other kinds of assumptions for the jump part, for example U_tdP_t. This assumption in equation (2.14) is to make the stochastic differential equation of $\log S_t$ simpler.

Also, note that the only difference between model 1FJ and model 1F is the jump part introduced in model 1FJ, and therefore they will produce similar stochastic path under same initial value, same parameters α, L, σ and same path of Z_t. The difference between the paths is because that when a jump happens, the value of S_t is changed to $(1+U_t)S_t$ at the same time.

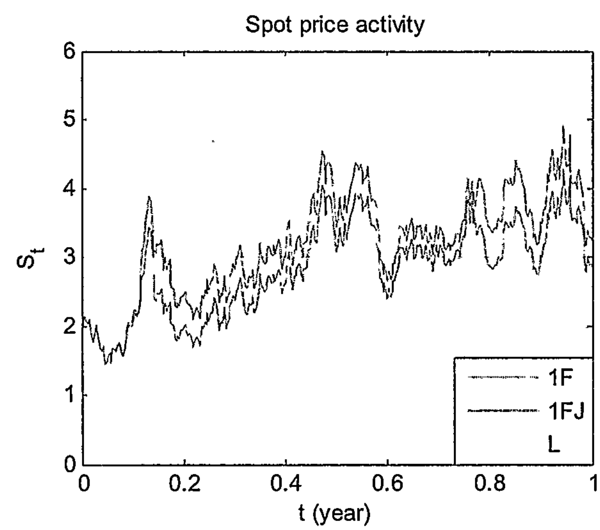

Figure 2.1: The Path of Model 1F and 1FJ with Same Parameters

For the solution of this model, we have

Theorem 2.4-1. *The explicit solution of equation (2.14) with initial value S_0 is*

$$\begin{aligned} S_t &= \prod_{i=1}^{P_t}(1+U_{\tau_i})(S_0 e^{-\alpha t} e^{\sigma Z_t - \frac{\sigma^2 t}{2}} \\ &+ \alpha L e^{-\alpha t} e^{\sigma Z_t - \frac{\sigma^2 t}{2}} \int_0^t g(s) e^{\alpha s} e^{-\sigma Z_s + \frac{\sigma^2 s}{2}} ds), \end{aligned} \tag{2.15}$$

where τ_i is the ith arrival time of the Poisson process P_t, and

$$g(s) = \prod_{i=1}^{P_s} \frac{1}{1+U_i}.$$

The expectation of S_t follows

$$\mathbb{E}[S_t] = e^{-\alpha t + \eta t \mathbb{E}[U_t]}(S_0 - L) + Le^{-\alpha t}\mathbb{E}[U_t]\, J_t + L, \tag{2.16}$$

where

$$J_t = \sum_{n=0}^{\infty} \frac{e^{-\eta t}(\eta t)^n}{n!} \mathbb{E}\left[\sum_{i=1}^{n} (1+\mathbb{E}[U_t])^{-i} e^{\alpha \tau_i} | P_t = n\right].$$

For the proof of theorem 2.4-1, see appendix B.2.

Now If $\mathbb{E}[U_t] = 0$, we will have the following corollary.

Corollary 2.4-2. *When $\mathbb{E}[U_t] = 0$, for equation (2.14),*

$$\mathbb{E}[S_t] = e^{-\alpha t}(S_0 - L) + L \tag{2.17}$$

This is the same as $\mathbb{E}[S_t]$ under equation (2.3).

Chapter 3

Calibration of Stochastic Models

3.1 Maximum Likelihood Estimate (MLE)

Suppose that the spot price follows the stochastic process S_t. The concept of maximum likelihood method is to make the conditional probability (for conditional probability, see appendix A.1.1) of $p(S_{t+\Delta t}|S_t)$ be the most likely to the historical data.

Now if we assume that the probability density function be defined as

$$f(\boldsymbol{\theta}, S_{t+\Delta t}|S_t)$$

where $\boldsymbol{\theta}$ is the set of parameters to be calibrated, then the most likelihood estimator would be

$$\mathcal{L}(\boldsymbol{\theta}) = \log \prod_{i=1}^{n} f(\boldsymbol{\theta}, S_{t+i\Delta t}|S_{t+(i-1)\Delta t}).$$

Now the goal is to make $\mathcal{L}(\boldsymbol{\theta})$ have the maximum value, i.e., to make the probability be most likely to the historical data.

Since the stochastic differential equation in one factor case is simple, it is possible to use the most likelihood method on the historical data to get the parameters.

To get the maximum of the likelihood function, we can use the Downhill Simplex Optimization. The method do not require evaluation of partial derivatives.

For application of maximum likelihood estimation in calibration Pilipović models, see [40].

3.2 Extended Kalman Filter (EKF)

In model 2F, there are more parameters, and therefore the performance of simple most likelihood method is not accurate enough to capture the path character, and it is necessary to use the extend Kalman filter to make the result more accurate (see [4], [19]).

If we consider a discrete model which satisfies the state process

$$\begin{cases} \hat{X}_k &= g(\hat{X}_{k-1}, w_k) \\ X_k &= h(\hat{X}_k, v_k) \end{cases} \tag{3.1}$$

where X_k is the observation value, $\hat{X}_k$ is the estimation from the model, and w_k and v_k are random distributions, $w_k \sim N(0, Q_k)$, $v_k \sim N(0, R_k)$. We can use the Kalman filter to get the true value of $\hat{X}_k$ with least error.

The extended Kalman filter includes two different phases, the predict phase, and the update phase.

In the predict phase,

$$\begin{cases} \hat{X}_{k|k-1} &= g(\hat{X}_{k-1}, 0) \\ P_{k|k-1} &= F_k P_{k-1|k-1} F_k^T + Q_k, \text{ where } F_k = \frac{\partial g}{\partial X}|_{X=\hat{X}_{k-1}} \end{cases} \tag{3.2}$$

In the update phase,

$$\begin{cases} \tilde{y}_k &= X_k - h(\hat{X}_{k|k-1}, 0) \\ V_k &= H_k P_{k|k-1} H_k^T + R_k, \text{ where } H_k = \frac{\partial h}{\partial X}|_{X=\hat{X}_{k-1}} \\ K_k &= P_{k|k-1} H_k^T V_k^{-1} \\ \hat{X}_k &= \hat{X}_{k|k-1} + K_y \tilde{y}_k \\ P_{k|k} &= (I - K_k H_k) P_{k|k-1} \end{cases} \tag{3.3}$$

Here $\tilde{y}_k \sim N(0, V_k)$ is the error of the prediction. When applying EKF in calibration of stochastic models, $\hat{X}_k$ is the vector of stochastic process, for example in model 1F, $\hat{X}_k$ is the values of S_t in time k, and in model 2F, $\hat{X}_k = (S_k, L_k)^T$. The random component

w_k is the random vectors of randomness in stochastic models, i.e, dZ_t and dW_t. X_k is the historical prices of S_t.

The algorithm of extended Kalman filter is as below.

- Start with an initial observation value X_0. We use this value as the initial value of the model, i.e., $\hat{X}_0 = X_0$. $P_{0|0}$ is a k by k matrix, where k is the dimension of $\hat{X}_0$. Start from time $t = 1$.
- Prediction phase: For the given $\hat{X}_{t-1}$ and $P_{t-1|t-1}$, follow the equation (3.2), calculate the predicted values $\hat{X}_{t|t-1}$ and $P_{t|t-1}$.
- Update phase: Calculate the error of prediction $\tilde{y}_t$ by equation (3.3). Now by equation (3.3), calculate $\hat{X}_{t|t}$ and $P_{t|t}$, which are the values of the best estimate.
- Repeat the prediction phase and update phase for $t = 2, 3, \ldots$, until all data points in the historical data are used.

For each set of parameters in function g and h, a series of $\tilde{y}_t$ is calculated. Since $\tilde{y}_t$ is normally distributed for any t, we can apply the maximum likelihood method on $\tilde{y}_t$, and then get the estimator of the parameters.

3.3 Euler Discretization

In calibration, we need to calculate the values of the parameters given the historical data. The stochastic models we are using are continuous, while the historical data is in discrete form. Therefore discretization of models are required for calibration.

The Euler discretization method of Pilipovic models is provided in several papers. Here we use the description in [4].

For model 1F, let Δt be a small enough number, and we can rewrite equation (2.3) as

$$S_{t+\Delta t} = S_t + \alpha(L - S_t)\Delta t + \sigma\sqrt{\Delta t} S_t \tilde{B}, \tag{3.4}$$

where $\tilde{B} \sim N(0,1)$.

However, for model 2F, we may get negative value of S_t using Euler discretization (see [4]). To avoid this, let $X_t = \log S_t, Y_t = \log L_t$, and by Itô's lemma (see A.1.4), equation (2.7) is transformed to

$$\begin{cases} dX_t &= (\alpha e^{Y_t - X_t} - \alpha - \frac{1}{2}\sigma^2)dt + \sigma dZ_t, \\ dY_t &= (\mu - \frac{1}{2}\xi^2)dt + \sigma dW_t. \end{cases} \tag{3.5}$$

By using logprices, we can guarantee the positive definite values in Euler discretization.

$$\begin{cases} X_{t+\Delta t} &= X_t + (\alpha e^{Y_t - X_t} - \alpha - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}\tilde{B}_1, \\ Y_{t+\Delta t} &= Y_t + (\mu - \frac{1}{2}\xi^2)\Delta t + \sigma\sqrt{\Delta t}\tilde{B}_2, \end{cases} \tag{3.6}$$

where $\tilde{B}_1, \tilde{B}_2 \sim N(0,1)$.

3.4 Parameter Estimation Through Spot Prices

3.4.1 One Factor Model (Model 1F)

It is possible to estimate the parameters when there is no jump part in Pilipovic model using maximum likelihood method.

For one factor case, we use the approximation from Euler discretization

$$S_{t+\Delta t} = S_t + \alpha(L - S_t)\Delta t + \sigma\sqrt{\Delta t} S_t \tilde{B}.$$

For this equation, $\tilde{B} \sim N(0,1)$, and therefore we can see

$$S_{t+\Delta t} \sim N(S_t + \alpha(L - S_t)\Delta t, \sigma^2 S_t^2 \Delta t)$$

Therefore for Δt small enough, it is possible to get the conditional probability density function

$$f(\boldsymbol{\theta}, S_{t+\Delta t}|S_t) = \frac{1}{\sigma\sqrt{2\pi\Delta t}} \exp(-\frac{(S_{t+\Delta t} - S_t - \alpha(L - S_t)\Delta t)^2}{2\sigma^2 S_t^2 \sqrt{\Delta t}}), \tag{3.7}$$

where $\boldsymbol{\theta}$ denotes all the parameters to be estimated. Now we choose the likelihood function to be

$$\mathcal{L}(\boldsymbol{\theta}) = \log \prod_{i=1}^{n} f(\boldsymbol{\theta}, S_{t+i\Delta t}|S_{t+(i-1)\Delta t}). \tag{3.8}$$

Here $S_{t+i\Delta t}$ is the spot price observation, or the discrete ith day spot price, from the first day to nth day.

For the simulated spot price path in this section, we assume that there are 360 days a year, therefore $\Delta t = \frac{1}{360}$. Note that Δt can be changed due to different business days in different markets.

Here after the calibration, we simulation the spot price path again, and compare the four moments of the simulated path to the historical path, to see the calibration error through the deviation.

Below is the simulation result using the natural gas spot price data in Henry Hub, Louisiana, USA.

We can calculate the simulated moments, and then compare the moments with the historical moments. Here we use the term "deviation" to define the difference between the generated moments and the historical moments.

The simulated moments (with deviation compared to historical data) are:

Mean: 4.4996(1.1275),

Variance: 48.478(43.3001),

Skewness: 1.3789(−0.74572),

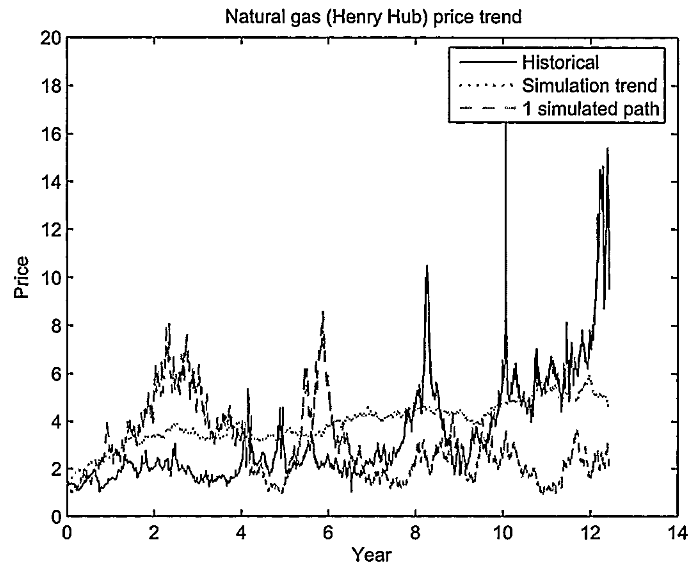

Figure 3.1: Calibration of Henry Hub Natural Gas Prices (Model 1F, MLE)

Kurtosis: 5.3117(−3.5875).

We can see that the simulated variance is not good comparing to the historical data. This is because the one factor model is to simple to capture the realistic movement of spot prices, and therefore it is important to use multi-factor models or models with jumps to improve the accuracy.

3.4.2 Two Factor Model (Model 2F)

In model 2F, we use the extended Kalman filter to do the calibration. By equation (3.6), the model can be expressed by logprices $X_t = \log S_t$, $Y_t = \log L_t$, and

$$\begin{cases} X_k &= X_{k-1} + (\alpha e^{Y_{k-1}-X_{k-1}} - \alpha - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}\tilde{B}_1, \\ Y_k &= Y_{k-1} + (\mu - \frac{1}{2}\xi^2)\Delta t + \sigma\sqrt{\Delta t}\tilde{B}_2. \end{cases} \tag{3.9}$$

Here $\tilde{B}_1, \tilde{B}_2 \sim N(0,1)$. By fitting equation (3.6) into the extended Kalman filter scheme in equation (3.1), it is possible to use the Kalman filter to do the calibration.

Figure 3.2 shows the simulation result using the natural gas spot price data in Henry

Hub, Louisiana, USA.

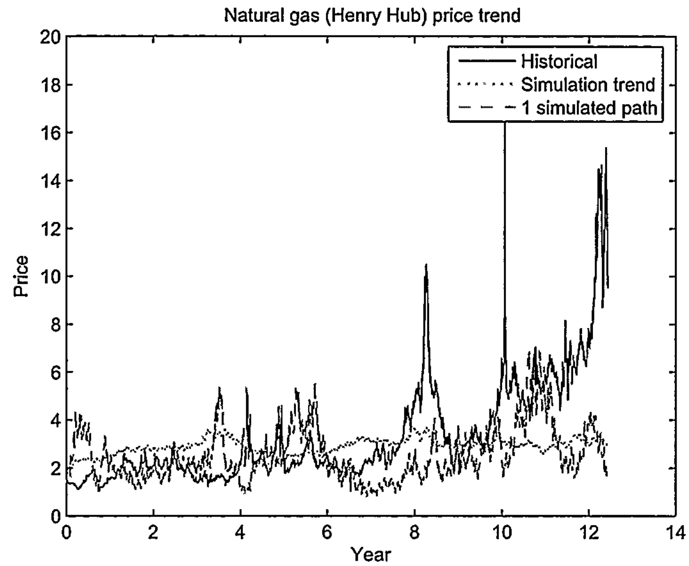

Figure 3.2: Calibration of Henry Hub Natural Gas Prices (Model 2F, EKF)

Simulation results (deviation):

Mean: 3.0023(−0.36982),

Variance: 5.8369(0.65897),

Skewness: 1.3184(−0.80616),

Kurtosis: 5.0346(−3.8646)

Extend Kalman filter method could also be used in other two factor models. For example, the calibration of AECO natural gas spot price data on model 2F has large deviation, however, if we assume that the second factor L_t follows the exponential Ornstein-Uhlenbeck process, we will have better calibration error.

Simulation results (deviation):

Mean: 3.1335(0.2306),

Variance: 4.0177(−0.66612),

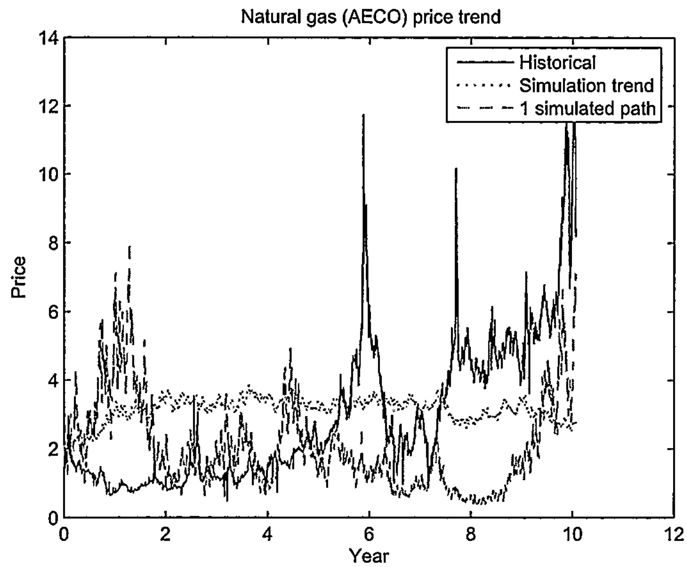

Figure 3.3: Calibration of AECO Natural Gas Prices (Model 2F, EKF)

Skewness: 1.4157(−0.00082024),

Kurtosis: 5.5767(0.47749)

Note that the calibration results here are still not applicable for practical use, since the seasonality components are not considered in the calibration. Furthermore, there is one more component needs to be calibrated before option pricing, which is the market price of risk. Only when all these components are well calibrated, we can price the options under risk-neutral measure and get the right prices.

Chapter 4

European Option Pricing for Model 1F

4.1 Payoffs of European Options

In European contracts, the holder can only exercise the right on the expiry date. Since the European option provides an opportunity to make profit from the difference between the strike price and the asset price on the expiry date, the payoff value for the holder of a European option on the expiry date T follows

$$V(T) = \max(S_T - K, 0) = (S_T - K)_+, \text{ for European call option,}$$

and

$$V(T) = \max(K - S_T, 0) = (K - S_T)_+, \text{ for European put option.}$$

European options are very simple, well understood and widely traded options, they are also called vanilla options.

4.2 Closed Form Option Pricing Method

There are two popular methods in closed-form pricing of financial derivatives. Suppose that the stochastic differential equation of the underlying spot price S_t is

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t, \tag{4.1}$$

we can use the martingale method or the partial differential equation method to price the European option.

Here we assume that the equation (4.1) is risk-neutral, and that the risk-free interest rate is a constant r.

For European call option and European put option, there is a put-call parity (see [15]).

Theorem 4.2-1. *(Put-Call Parity) Given the constant risk-free interest rate r, for the underlying asset that has spot price S_t, consider the European call option price $C_t(S_t, K, T)$ at time t with maturity date T and strike price K, and the European put option price $P_t(S_t, K, T)$, and they have the following parity*

$$C_t(S_t, K, T) + Ke^{-r(T-t)} = S_t + P_t(S_t, K, T).$$

Therefore, we only need to calculate the closed form solution to one of the two kinds of options, and the other one can be achieved by put-call parity. We pick up European call option for pricing in this chapter.

The first method is the martingale approach. Under the assumption that the equation (4.1) is risk-neutral, the discounted spot price is a martingale (see section A.1.6). As we know that the payoff function $V(S_T, K, T)$ of European call options at maturity T with strike price K is

$$V(S_T, K, T) = (S_T - K)_+,$$

the price of the European call option can be achieved by calculating the discounted expect payoff of the option based on all the information $\mathcal{F}_t$ at time t, i.e.,

$$C_t(S_t, K, T) = \mathbb{E}\left[e^{-r(T-t)} V(S_T, K, T) | \mathcal{F}_t\right],$$

or by substituting the payoff function,

$$C_t(S_t, K, T) = \mathbb{E}\left[e^{-r(T-t)} (S_T - K)_+ | \mathcal{F}_t\right].$$

If the price process S_t has Markov property, and we use the induced filtration of S_t for $\mathcal{F}_t$, by the Markov property we can have

$$C_t(S_t, K, T) = \mathbb{E}\left[e^{-r(T-t)} (S_T - K)_+ | S_t\right].$$

The other method is by using partial differential equations. The partial differential equation can be derived from portfolio hedging (see [6], [29]).

If the model of the underlying asset is simple, it is easy to derive the analytical form solution (e.g., model 1F) by either method. However, when the model is complicated (e.g., model 2F and model 1FJ), the closed form solution can not easily be derived, and we need to apply numerical methods. The first method is based on the expectation of the discounted payoff function, and Monte Carlo method can be used. For the second method, there are also several numerical methods in solving the partial differential equation, such like finite difference method, finite element method, etc.

The martingale approach is based on the calculation of the expectation,

4.3 Explicit Formula

In equation (2.4), the solution of the stochastic differential equation (2.3) is given by a time integral of a exponential geometric Brownian motion. M. Yor started the research on this kind of time integrals in 1992 (see [41]), and in his summary paper in 2005 ([42]) he pointed out that currently there is still no closed form probability density function, while the best result is a function with a double integral in it.

With the probability density function provided in M. Yor's paper (see [41], [42]), we can write the European option pricing formula in the form of a triple integral. If we transform equation (2.4) using the scaling property of Wiener process and a change of variables, we can have

$$\begin{aligned} S_T &= S_0 e^{-\alpha T} e^{\sigma Z_T - \frac{\sigma^2 T}{2}} + \alpha L e^{-\alpha T} e^{\sigma Z_T - \frac{\sigma^2 T}{2}} \int_0^T e^{\alpha s} e^{-\sigma Z_s + \frac{\sigma^2 s}{2}} ds \\ &= S_0 e^{-2(B_{T_0}^{(v)})} + \frac{4}{\sigma^2} \alpha L e^{-2(B_{T_0}^{(v)})} A_{T_0}^{(v)}, \end{aligned} \tag{4.2}$$

where

$$\begin{aligned} T_0 &= \frac{\sigma^2}{4}T, \\ v &= \frac{2}{\sigma^2}\alpha + 1, \\ B_s &= -\frac{\sigma}{2}Z_{\frac{4}{\sigma^2}s}, \\ B_{T_0}^{(v)} &= vT_0 + B_{T_0}, \\ A_{T_0}^{(v)} &= \int_0^{T_0} e^{2B_s^{(v)}}ds. \end{aligned}$$

Note that B_s here is still a Brownian motion (see A.1.2). Therefore $A_{T_0}^{(v)}$ is an integral of exponential Brownian motion. Now by [41], there is a theorem on the joint probability density function of $A_{T_0}^{(v)}$ and $B_{T_0}^{(v)}$.

Theorem 4.3-1. *(M. Yor, see [41], [42]) For any one-dimensional Brownian motion B_t, if we consider the drifted Brownian motion $B_t^{(\mu)}$ and the integral*

$$A_t^{(\mu)} = \int_0^t e^{2B_s^{(\mu)}}ds, \ t \geq 0, \tag{4.3}$$

and assume that $t > 0$, $u > 0$ and $x \in \mathbb{R}$, then the joint probability density function of $A_t^{(\mu)}$ and $B_t^{(\mu)}$ satisfies

$$\mathbb{P}(A_t^{(\mu)} \in du, B_t^{(\mu)} \in dx) = e^{\mu x - \mu^2\frac{t}{2}}\exp(-\frac{1+e^{2x}}{2u})\theta(\frac{e^x}{u}, t)\frac{dxdu}{u}, \tag{4.4}$$

where θ is defined by

$$\theta(r,t) = \frac{r}{(2\pi^3 t)^{1/2}}e^{\frac{\pi^2}{2t}}\int_0^\infty e^{-\frac{\xi^2}{2t} - r\cosh(\xi)}\sinh(\xi)\sin(\frac{\pi\xi}{t})d\xi. \tag{4.5}$$

By applying the result in theorem (4.3-1), we can get

Corollary 4.3-2. *If the price process S_t follows equation (2.3), the explicit European option pricing formula under model 1F is*

$$C_T = e^{-\frac{4r}{\sigma^2}T_0}c_{v,T_0}\int_0^{+\infty}\int_0^{+\infty}(S_0x^{-2} + \frac{4}{\sigma^2}\alpha L x^{-2}u^{-1} - K)_+ x^v e^{-\frac{1}{2}u(1+x^2)}\psi_{xu}(T_0)dudx$$

where

$$
\begin{aligned}
c_{v,T_0} &= \frac{1}{\sqrt{2\pi^3 T_0}} e^{\frac{\pi^2}{2T_0} - \frac{v^2 T_0}{2}} \\
\psi_{xu}(T_0) &= \int_0^{+\infty} e^{-\frac{s^2}{2T_0} - xu\cosh s} \sinh s \sin \frac{\pi s}{T_0} ds
\end{aligned}
$$

Note that the European option price in the corollary is of an integral form, and not quite convenient for application in European option prices. For numerical method, Monte Carlo simulation and binomial tree model (see chapter 6) are better choice, since the concepts of the methods are easier to understand.

Chapter 5

Swing Options

5.1 Swing Option Contracts

Swing options provide option holders multiple exercise rights, and the holders can choose to exercise the rights on several given exercise dates before the end of the contract. However, on each exercise date, the holder can only exercise at most once. Furthermore, there could be penalties if the exercise volume exceeds some given constraint. There are several different versions of swing options, depending on the payoff features, the penalty constraints, the flexibility of exercise volumes, etc. There several papers on modelling and pricing swing options, see [12], [17], [23]. Although there are several versions of swing option contracts, however, it is still possible to give a general model for swing options (see [12], [17]).

Definition 5.1-1. We define a general swing contract as a contract with all conditions and assumptions below:

1. Current time when the contract is written and priced is assumed to be the current time 0. This is because the price of swing option does not change when shifting time scale.

2. The contract starts at time T_1, and ends at time T_2.

3. Possible exercise dates: $\{\tau_1, \ldots, \tau_n\} \in [T_1, T_2]$. We assume that

$$T_1 = \tau_1 < \tau_2 < \ldots < \tau_n = T_2.$$

4. Refraction period is Δ. That is, if the holder exercises the right at time τ, the next exercise date should be after time $\tau + \Delta$.

5. Number of rights: $N \leq n$.

6. Exercise volume at τ_i: ξ_{τ_i}, with constraint $\xi_{\tau_i} \in [p_i, q_i]$. If ξ_{τ_i} exceeds the constraint, there would be a penalty charge.

7. Total volume constraints: $\sum_{i=1}^{N} \xi_{\tau_i} \in [P, Q]$. There are penalty charges if the total volume exceeds the constraint.

8. Penalty function when the volume exceeds the constraints: $\phi_{\tau_i}(\xi_i)$ at time τ_i (for constraints in assumption 6), total penalty $\Phi(\sum_{i=1}^{N} \xi_i)$ (for constraint in assumption 7).

9. Strike price: K or variant K_t, $t \in \{\tau_1, \ldots, \tau_n\}$.

Note that the constraints and penalties are important in a swing option, or it will reduce to an American option. R. Elliott and A. Cadenillas pointed out in [12] that if $N = n$, and there is no volume constraints (in definition 6,8 above), and the strike price is constant K, the holder only need to exercise once with volume Q to maximize the portfolio value, and therefore the swing option value is just the same as the American option.

5.1.1 Payoff Function of Swing Options

Let S_t follows a risk neutral model, the holder uses a maximization strategy, then the value of swing option follows (see [12])

$$\max \mathbb{E}[\sum_{\imath=1}^{N} e^{-r\tau_{d_i}}(\xi_{\tau_{d_i}} G(S_{\tau_{d_i}}) - \phi_{\tau_{d_i}}(\xi_{\tau_{d_i}}))] - \Phi(\sum_{i=1}^{N} \xi_i), \tag{5.1}$$

where $G(S_{\tau_{d_i}})$ is the payoff function, and

$$[d_1, \ldots, d_N] \subseteq [1, \ldots, n], \text{ where } d_i \neq d_j \text{ for any } i \neq j$$

In a simpler case, if we let $N = n$, $K_t = K, p_i = p_0, q_i = q_0$ be constant, and assume that the holder is not allowed to exercise volumes out of constraints, and consider a call swing option, we will have the option value

$$\max \mathbb{E}[\sum_{i=1}^{N} e^{-r\tau_i}\xi_{\tau_i}(S_{\tau_i} - K)_+] \tag{5.2}$$

such that

$$\xi_{\tau_i} \in [p_0, q_0],\ i = 1, \ldots, N,$$
$$P \leq \sum_{i=1}^{N} \xi_{\tau_i} \leq Q.$$

R. Elliott and A. Cadenillas pointed out in [12] that the solution of this option value is a solution of variational inequalities.

5.2 Fixed Exercise Volume Swing Options

There are many conditions defined in definition (5.1-1). It is very complicated if we model the swing option using all the assumptions. By altering the restrictions in the definition, there can be several different versions of swing options.

Here in the thesis we assume that the volume exercised on each exercise date τ_i be a fixed amount, i.e., $\xi_i = \xi$ for all $i = 1, 2, \ldots, n$. In this case, on each exercise date, the volume exercised is either 0 or ξ, and the single exercise volume constraint $\phi_{\tau_i}(\xi_i)$ is not used. For the overall penalty function Φ, we can see that it depends only on the number of exercises. By the fixed exercise volume assumption, the model is greatly simplified.

A. Lari-lavassani et al. provided the binomial tree method in pricing this version of swing option under model 1F and model 2F ([23]). Here in the thesis we will use this swing option model, and provide the multinomial method for pricing under model 1FJ.

Below are the assumptions for pricing fixed volume swing options.

- Discrete time scale is used for the pricing. The length between to time scale is Δt. This is to discretize the continuous time model, which is essential for binomial and multinomial tree method, since this Δt decides the time scale between each two consequent steps in the tree. The smaller Δt is, the more accuracy we can get.

- The contract starts at time 0 (current time). This is also the time when the option is priced. At time 0, the underlying asset price is S_0.
  The time 0 assumption is just by shifting of the time scale. Since the option pricing remains the same if we just shift the time scales, this will not change the result of the option price.

- The time scale between two consequent exercise dates is $k\Delta t$. In discrete trees, it takes k steps in the tree.

- There are N exercise dates, starting from time $k\Delta t$, and ending at time $Nk\Delta t$. i.e, $\tau_1 = k\Delta t$, $\tau_2 = 2k\Delta t$, ..., $\tau_N = Nk\Delta t$. Clearly τ_N is the expiry date of the contract.

- The strike price at all exercise dates are the same and constant, K.

- The option holder has at most n exercise rights. Here $n \leq N$, so the holder needs to decide on which dates to exercise the rights.

- When the holder exercises the right at an exercise date, the holder can sell ξ volume of underlying asset at price K. Here ξ is constant at each exercise date.

- Throughout the whole time of the contract, the total exercised volume needs to be in interval $[P, Q]$. If the total volume exceeds Q or is less than P, the penalty price will be charged. The penalty is charged at the expiry date.
  We can see that the penalty only depends on the total rights the holder exercised. Define the penalty function as $\Phi(x)$, where x is the rights the holder exercised. Actually, we can see that the penalty function only affects the payoff function of

the option at the maturity date, therefore when we use backward derivation, it is easy to estimate the effect of the existence of penalties.

- The risk-free interest rate is r.

5.3 The Snell Envelope

Swing options are in the family of multi-exercise options, and the optimal stopping strategy is essential for solving swing option prices. The Snell envelope, which is a strong tool in solving optional stopping problems, can be applied in pricing swing options. By calculating the Snell envelope, one can get an optimal exercise strategy, as well as the option value under the optimal strategy, which is the swing option price.

For definition and discussion of Snell envelope, see [31]. Generally, the Snell envelope can be defined as below.

Definition 5.3-1. (The one-dimensional Snell envelope) Given an filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t=0}^T$, an $\mathcal{F}$-adapted process $G = \{G_t\}_{t=0}^T$, the Snell envelope of G is defined as the process $X = \{X_t\}_{t=0}^T$, such that

$$X_T = G_T, \tag{5.3}$$

$$X_t = \max(G_t, \mathbb{E}[X_{t+1}|\mathcal{F}_t]) \text{ for all } 0 \le t < T. \tag{5.4}$$

When applying Snell envelope in swing option pricing, G_t is the payoff function of the underlying asset at time t, and T the expiry date of the contract. Since the price of the asset at time t is S_t, we can define $\mathcal{F}_t = \sigma\{S_0, S_1, \dots, S_t\}$, i.e., the filtration $\mathcal{F}_t$ is all the information of the asset price we have until time t. Actually, since the price process S_t has the Markov property, we can have

$$\mathbb{E}[X_{t+1}|\mathcal{F}_t] = \mathbb{E}[X_{t+1}|S_t].$$

Take American call option as an example. In American options, the payoff function G_t at time t is

$$G_t = (S_t - K)_+,$$

and therefore the Snell envelope X is

$$X_T = (S_T - K)_+, \tag{5.5}$$

$$X_t = \max((S_t - K)_+, \mathbb{E}[X_{t+1}|S_t]). \tag{5.6}$$

Intuitively, the definition of the Snell envelope contains two parts. First, equation (5.3) shows that at the expiry date T, the holder will always exercise his remaining rights, and X_T is the payoff at time T. Second, at each time $t < T$, there are two strategies, and the holder needs to decide whether to exercise or to hold the right. Therefore the holder needs to compare the benefits he can achieve by using each strategy. Depending on the current price, if the payoff he can get at time t is higher, then he will exercise the right, otherwise it is better to hold the remaining rights. Therefore, the option value at time t is the maximum between the immediate payoff and the conditional future option value, which is described in equation (5.4). Finally, note that X_t is actually the option value at time t under optimal strategy, X_0 is therefore the option price we are looking for.

The Snell envelope can be extended to multi-exercise cases by expanding the dimension of X_t. If the holder is allowed to exercise at most m times during the life of the contract, $m > 1$, and that he could only exercise once at each time, then the Snell envelope can be defined as

Definition 5.3-2. (The multi-dimensional Snell envelope) Given an filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t=0}^{T}$, an $\mathcal{F}$-adapted process $G = \{G_t\}_{t=0}^{T}$, the m-dimensional Snell envelope ($m > 0$) of G is defined as the process $X = \{X_t\}_{t=0}^{T}$, such that

$$X_k^0 = 0, \text{ for all } k = 0, 1, \ldots, T, \tag{5.7}$$

$$
\begin{aligned}
X_T^i &= G_T, \text{ for all } i = 1, 2, \ldots, m, && (5.8)\\
&\text{for all } 0 \le t < T,\\
X_t^i &= \max(G_t + \mathbb{E}[X_{t+1}^{i-1}|\mathcal{F}_t], \mathbb{E}[X_{t+1}^i|\mathcal{F}_t]), \text{ for all } i = 1, 2, ..., m. && (5.9)
\end{aligned}
$$

Note that in option pricing, X_t^i is the option value at time t with i rights left, under optimal exercise strategy. When $m = 1$, this definition reduces to definition (5.3-1).

By the definition of Snell envelope, before the expiry date, since the holder will only exercise when

$$G_t + \mathbb{E}[X_{t+1}^{i-1}|\mathcal{F}_t] \ge \mathbb{E}[X_{t+1}^i|\mathcal{F}_t], \tag{5.10}$$

we can actually get the optimal stopping strategy if we can calculate the conditional expectations $\mathbb{E}[X_{t+1}^{i-1}|\mathcal{F}_t]$ and $\mathbb{E}[X_{t+1}^i|\mathcal{F}_t]$ in equation (5.10). Actually, if all these conditional expectations are known at step $t+1$, by equation (5.9), all the Snell envelopes X_t^i at time t can be evaluated, which will lead to the value of all conditional expectations at step t. At the expiry date T, we can see from equation (5.8) that the Snell envelopes are known, and therefore the conditional expectations at step T can be calculated. In all, starting from the expiry date, by backward derivation, all the Snell envelopes can be finally calculated, as well as the swing option price.

The two kinds of numerical methods, Monte Carlo simulation and binomial and multinomial tree method, which will be discussed chapter 6, are based on the calculation of the conditional expectation in equation (5.10).

The conditional expectations in equation (5.10) are sometimes named as "the continuation value". We will use this term in the following chapters.

Chapter 6

Numerical Option Pricing Method

Here in this chapter we use two methods that can be used in swing option pricing, the Monte Carlo simulation and the binomial/multinomial tree method. Here in the thesis the Monte Carlo simulation is applied in both no jump and jump cases, and the binomial tree method is applied in one-factor model and the one-factor model with jumps. In two-factor case, the binomial tree extends to a forest (see [23]), and in jump case, it is extended to a multinomial tree.

6.1 Monte Carlo Method

The Monte Carlo method is widely used in several disciplines, including physics, finance, and other computational fields. In Monte Carlo method, the stochastic behavior of the underlying systems is simulated by using random numbers (more often pseudo-random numbers), and the results are groups of generally nondeterministic sample data, which can be used for statistical inference. The method can be implemented in mathematical problems that can not easily be solved, and it is more efficient when the dimension of the system increases, compared to other numerical methods.

In our swing option pricing problem, the system of the asset price is itself a stochastic system, and we use Monte Carlo simulation to generate the path of the prices. In 2001, F. A. Longstaff and E. S. Schwartz used a least-squares regression ([25]) to calibration the conditional expectations for evaluating the Snell envelope in one-dimensional case, and American option price can be estimated in this way. In 2004, N. Meinshausen and B. M. Hambly ([27]) used the least-squares scheme to calibrate the continuation values in multi-dimensional case. We can apply this method on swing option pricing in jump cases.

The other implementation of Monte Carlo method in this thesis, is the numerical estimation of European option prices. We can apply the method to estimate the expectation of the payoff function, and the European option price can be evaluated.

6.1.1 Monte Carlo and Randomized Quasi-Monte Carlo Simulation

There are several articles on the Monte Carlo simulations. Here we use the description in [7]. For a given expectation

$$\mu = \mathbb{E}f(\mathbf{u})$$

where $\mathbf{u}$ is a $(0,1)$ uniformly distributed random vector, the concept of Monte Carlo method is to generate n identically independently distributed (i.i.d) vectors $\mathbf{u}_i$, $i = 1, 2, \ldots, n$, which are uniformly $(0,1)$ distributed, and the approximation of the expectation is

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} f(\mathbf{u}_i).$$

For quasi-Monte Carlo method, the difference is that $\mathbf{u}_i$ are not i.i.d $(0,1)$ uniformly distributed vectors, but become highly uniformly quasi-Monte Carlo sequences, which can be generated in a given procedure (see [11]). Examples of this kind of sequences are Halton sequence, Sobol sequence, etc.

For randomized quasi-Monte Carlo simulation, the quasi-Monte Carlo sequences are randomly shifted, i.e.,

$$\mathbf{v}_i = \mathrm{mod}(\mathbf{u}_i + \mathbf{r}_i, 1),$$

where $\mathbf{r}_i$ are $(0,1)$ uniformly distributed vectors (see [24]).

The approximation of expectation now becomes

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} f(\mathbf{v}_i).$$

Monte Carlo method is a probabilistic approximation, and its accuracy can be given by calculating the confidence interval of the simulation results. A $(1-\alpha)$ confidence interval means that the probability that the true value is not in the interval is α. If the number of samples in the simulation is n, and the results of the n samples have mean $\hat{\mu}$ and variance $\hat{\sigma}^2$, the $(1-\alpha)$ confidence interval is given by

$$CI = (\hat{\mu} - \frac{\hat{\sigma}}{\sqrt{n}} z_{\alpha/2},\ \hat{\mu} + \frac{\hat{\sigma}}{\sqrt{n}} z_{\alpha/2}), \tag{6.1}$$

where $z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution. For example, when $\alpha = 0.05$, we can have the 95% confidence interval with half length $1.96\hat{\sigma}/\sqrt{n}$.

From equation (6.1) we can see that the algorithm has probabilistic error $O(\hat{\sigma}/\sqrt{n})$. The error of the estimator depends only on the variance of sample results and sample number n, which has advantage in high-dimensional problems.

6.1.2 Path Simulation

In equation (2.4) and equation (2.8), there are integrals of stochastic terms. These integrals can not be solved analytically, which would be a problem for us to generation spot prices. Although we only need the spot price S_T in European option pricing, it is hard to apply the probability density function to generate values of S_T. One direct way to get S_T is to generate the whole spot price path through time $[0, T]$, and we can simply use the approximate Euler discretization.

For Euler discretization, see section 3.3.

To generate the path for the cases with jumps in equation (2.14), we need to generate the Poisson process P_t, which is discrete. Here we use Bernoulli distribution as the approximation, i.e., from time t to time $t + \Delta t$, the probability of a jump is $\eta\Delta t$. If the jump occurs, we multiply S_t by $(1 + U_t)$.

In figure 6.1 we can see a jump with large size near step 500.

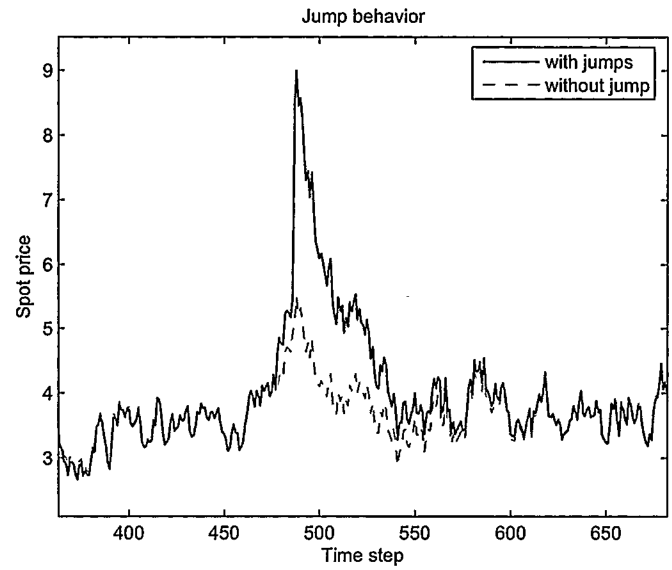

Figure 6.1: Large Jump Generated by Monte Carlo Simulation

Figure 6.2 shows the spot price behavior under model 1F and model 1FJ, generated with the same Wiener process Z_t. Here in the simulation we assume that $\log(1+U_t) \sim N(\mu_J, \sigma_j^2)$. Figure 6.3 shows the spot price behavior of model 2F.

6.1.3 Variance Reduction Techniques

From equation (6.1) we can see that there are two ways to improve the accuracy of Monte Carlo simulations: increment of sample numbers and reduction of variances. While the sample number is often fixed and depends only on the strength of computing machines, the variance of Monte Carlo results can be reduced by better design of simulation algorithms. Variance reduction techniques are widely used to increase the accuracy of Monte Carlo method. We implement antithetic and control variate variance reduction techniques in the thesis. The description of these methods are provided in [7].

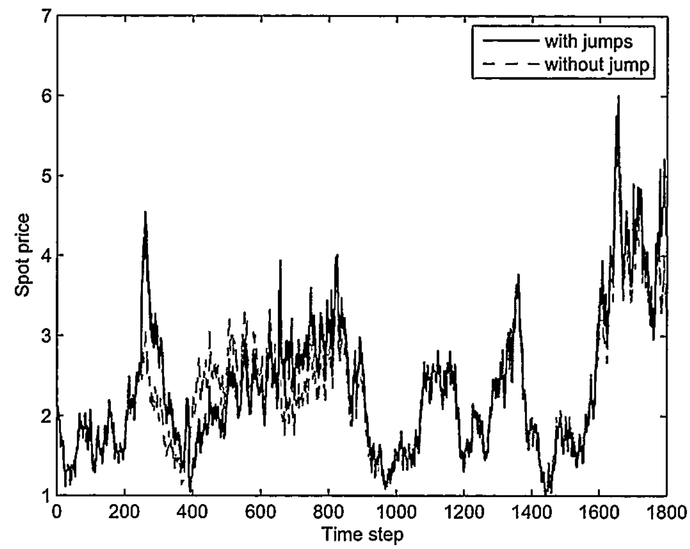

Figure 6.2: Sample Path of Monte Carlo Simulation (Model 1F and Model 1FJ)
5-year simulation(1800 steps),
$S_0 = 2, L = 3, \alpha = 2, \sigma = 1, \lambda = 1.5, \mu_J = 0.2, \sigma_J = 0.2$

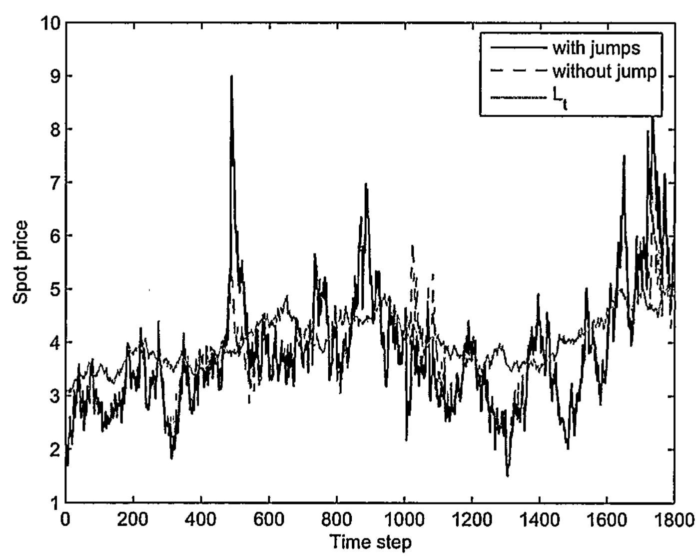

Figure 6.3: Sample Path of Monte Carlo Simulation (Model 2F)
5-year simulation(1800 steps),
$S_0 = 2, L_0 = 3, \alpha = 10, \sigma = 1, \lambda = 1.5, \mu = 0.2, \xi = 0.2, \mu_J = 0, \sigma_J = 0.2$

For antithetic variates, based on the fact that $1-\mathbf{u}$ is also (0,1) uniformly distributed vectors, we generate $\frac{n}{2}$ vectors $\mathbf{u}_i$, and get the estimation of expectation by

$$\hat{\mu} = \frac{2}{n}\sum_{i=1}^{\frac{n}{2}}(f(\mathbf{u}_i) + f(1-\mathbf{u}_i)).$$

For control variates, the estimation of expectation becomes

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}(f(\mathbf{u}_i) + \beta(\mathbb{E}g(\mathbf{u}) - g(\mathbf{u}_i))),$$

where β satisfies

$$\beta = \frac{Cov(f(\mathbf{u}_i), g(\mathbf{u}_i))}{Var(g(\mathbf{u}_i))}.$$

It was proved ([7]) that these two methods could reduce the variance of $f(\mathbf{u}_i)$ and therefore increase the accuracy (see also [2]). And it is possible to combine these two methods together to get better results. Furthermore, they are also available if we use randomized quasi-Monte Carlo method.

The control variates in the thesis are:

- Model 1F:

$$\mathbb{E}S_T = e^{-\alpha T}(S_0 - L) + L$$

- Model 2F:

$$\mathbb{E}S_T = S_0 e^{-\alpha T} S_0 + \frac{\alpha}{\alpha+\mu} L_0 (e^{\mu T} - e^{-\alpha t})$$

- Model 1FJ:

  As there is no closed form of expectation of S_T in this case, we use the same random vector to generate a path of Q_T, where Q_T is the spot price under one factor model without jumps with initial value S_0, and use

$$\mathbb{E}Q_T = e^{-\alpha T}(S_0 - L) + L$$

  as the control variate.

The variance are based on the results of $f(\mathbf{u}_i)$ in Monte Carlo simulation. However, when we apply randomized quasi-Monte Carlo method, it is not proper to use $Var(f(\mathbf{u}_i))$ directly, since $\mathbf{u}_i$ are not independent in this case. Alternatively, we could do m times of simulations, and get m estimates $\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_m$. Now the variance can be represented as $Var(\hat{\mu}_i)$.

6.2 Monte Carlo Method on Swing Option Pricing

In chapter 5 we introduce the Snell envelope for swing option pricing:

$$\begin{aligned}
X_k^0 &= 0, \text{ for all } k = 0, 1, \ldots, T, && (6.2)\\
X_T^i &= G_T, \text{ for all } i = 1, 2, \ldots, n, && (6.3)\\
&\text{for all } 0 \le t \le T,\\
X_t^i &= \max(G_t + \mathbb{E}[X_{t+1}^{i-1}|\mathcal{F}_t], \mathbb{E}[X_{t+1}^i|\mathcal{F}_t]), \text{ for all } i = 1, 2, ..., n. && (6.4)
\end{aligned}$$

The assumptions for the swing option model is given in section 5.2. With unit time interval Δt, the time scales we need to simulate is from time 0 to time $Nk\Delta t$, and the exercise dates are dates $k, 2k, \ldots, Nk$. Actually, in Monte Carlo method, we only need to simulate the prices at exercise dates.

First we modify the Snell envelope description, to fit the notations in the assumptions in section 5.2.

Given the initial price S_0, assume that the prices at exercise dates are $S_k, S_{2k}, \ldots, S_{Nk}$. By substituting the penalty function in the payoff on expiry date, we can write the Snell envelope for the swing put option price as

$$\begin{aligned}
X_{tk}^0 &= 0, \text{ for all } t = 0, 1, \ldots, N, && (6.5)\\
X_{Nk}^i &= \max\left(\Phi(n-i), \Phi(n-i+1) + (K - S_{Nk})_+\right), \text{ for all } i = 1, 2, \ldots, n, && (6.6)\\
&\text{for all } 0 < t < T, \text{ for all } i = 1, 2, ..., n,
\end{aligned}$$

$$X^i_{tk} \quad = \max\left((K - S_{tk})_+ + e^{-rk\Delta t}\mathbb{E}[X^{i-1}_{(t+1)k}|S_{tk}], e^{-rk\Delta t}\mathbb{E}[X^i_{(t+1)k}|S_{tk}]\right). \qquad (6.7)$$

Here the penalty charges are calculated through equation (6.6). Also note that we do not calculate X^i_0 here, since it can not be regressed. The option price is achieved by applying an early exercise strategy.

We apply the Monte Carlo method of swing option pricing in [27]. The method consists of two procedures:

- Regression of the continuation value $\mathbb{E}[X^i_{(t+1)k}|S_{tk}]$ for all $i = 1, 2, \ldots, n$, and for all $t = 1, 2, \ldots, N-1$. This is by extending the method provided in [25].
- Calculation of the negative-biased estimate. This is the option value by early exercising, which is the method introduced in [25].

The convergence of the Monte Carlo method is proved in [27].

6.2.1 Regression of Continuation Value

The first step is to do the regression and get the relationship between S_{tk} and the continuation value $\mathbb{E}[X^i_{(t+1)k}|S_{tk}]$. If we can calculate the value of $X_{(t+1)k}$ by simulation, let

$$\mathbb{E}[X^i_{(t+1)k}|S_{tk} = x] = g^i_t(x),$$

and pick up proper basis function for the function $g^i_t(x)$, for example $g^i_t(x) = a + bx$. Now by regression we can get the parameters a, b and then the form of $g^i_t(x)$, which is an approximation of the continuation value function.

The procedure of the progression is as below ([27]).

- Generate p samples of price paths, and we get p paths of prices on exercise dates. Let the paths be

$$S^{(i)} = \{S^{(i)}_{tk}\}^N_{t=0}, \ i = 1, 2, \ldots, p.$$

- For each path, let the corresponding Snell envelope be $X_{tk}^{i,j}$ at each exercise time tk. Here $i = 1, 2, \dots, n$ denotes the remaining rights, and $j = 1, 2, \dots, p$ is the index of the sample.

- The value $X_{Nk}^{i,j}$ is the Snell envelope value on maturity date, and can be calculated:

$$X_{Nk}^{i,j} = \max\left(\Phi(n-i), \Phi(n-i+1) + (K - S_{Nk}^{(j)})_+\right)$$

- Now we have the sample values $X_{Nk}^{i,j}$ and $S_{(N-1)k}^{(j)}$, for the continuation value function $y = g_t^i(x)$, use the samples values $X_{Nk}^{i,j}$ for y, and $S_{(N-1)k}^{(j)}$ for x, pick up a proper basis function (here we use $a+bx$), and we can use linear regression to get $g_{N-1}^i(x)$.

- For each $S_{(N-1)k}^{(j)}$, calculate the approximate value of $\mathbb{E}[X_{Nk}^i | S_{(N-1)k}]$ by

$$\mathbb{E}[X_{Nk}^i | S_{(N-1)k}] = g_{N-1}^i(S_{(N-1)k}).$$

  Now from the equation

$$X_{tk}^i = \max\left((K - S_{tk})_+ + e^{-rk\Delta t}\mathbb{E}[X_{(t+1)k}^{i-1} | S_{tk}], e^{-rk\Delta t}\mathbb{E}[X_{(t+1)k}^i | S_{tk}]\right),$$

  given $\mathbb{E}[X_{Nk}^i | S_{(N-1)k}]$, we can calculate $X_{(N-1)k}^{i,j}$ for each sample path.

- Repeat the regression at dates $(N-2), (N-3), \dots, 2$, and get all approximate values of the continuation value $\mathbb{E}[X_{tk}^i | S_{(t-1)k}] = g_{t-1}^i(S_{(t-1)k})$.

By using this algorithm, we can get $g_t^i(x)$ for all $t = 1, 2, \dots, N-1$, $i = 1, 2, \dots, n$. We do not need to do the regression for time 0 values, since we do not need to do make exercise decision at time 0.

6.2.2 Negative-Biased Estimation

The negative-biased estimation is based on early exercise strategy. As we already have the regression functions of continuation value, at each exercise date, we can make exercise decision based on the regression function.

Below is the algorithm of the negative-biased estimation ([27]).

- Generate p samples of price paths. These path are different from the paths generated in regression procedure. Still, let the paths be

$$S^{(i)} = \{S_{tk}^{(i)}\}_{t=0}^{N}, \ i = 1, 2, \ldots, p.$$

- For each sample path $S^{(j)}$, starting from date k, with n remaining rights, do the simulation as follows:

  For each exercise opportunity $t = 1, 2, \ldots, N$, compare $(K - S_{tk}^{(j)})_+$ and $g_t^i(S_{tk}^{(j)})$, where i is the remaining rights on the sample path j.

  If $(K - S_{tk}^{(j)})_+ > g_t^i(S_{tk}^{(j)})$, exercise the right, and move on to time $t+1$ with $i-1$ remaining rights.

  If $(K - S_{tk}^{(j)})_+ \leq g_t^i(S_{tk}^{(j)})$, hold the right, and move on to time $t+1$ with i remaining rights.

- For each sample path j, the discounted payoff of the swing option can be calculated. The option price is the mean value of all the payoffs.

Note that by using this method, we are actually applying the stopping strategy $\tau^{(j)} = \{\tau_i^{(j)}\}_{i=1}^{n}$, where $\tau_i^{(j)}$ gives the stopping time with i rights remaining in sample path j. Here

$$\tau_n^{(j)} = \min\{t | (K - S_{tk}^{(j)})_+ > g_t^n(S_{tk}^{(j)}), \ t = 1, 2, \ldots, N\},$$

and for $\tau_{n-1}^{(j)}, \tau_{n-2}^{(j)}, \ldots, \tau_1^{(j)}$, we have

$$\tau_i^{(j)} = \min\{t | (K - S_{tk}^{(j)})_+ > g_t^i(S_{tk}^{(j)}), \tau_t < t \leq N\}$$

for all $i = n-1, n-2, \ldots, 1$.

Since the swing option value is the maximum payoff of all possible stopping strategies, this strategy $\tau^{(j)}$ provides a negative-biased estimate of the swing option price. Actually, if the fitted continuation value function is the real continuation value function, the strategy $\tau^{(j)}$ will be the optimal stopping strategy for the swing option (see [27]).

6.3 Binomial Tree and Multinomial Tree Method

In binomial tree method, it is assumed that the price of the asset has a binomial behavior, i.e., the price will either be increasing or decreasing for certain amount in each unit time. Based on this assumption, if we simulate the price behavior in the length of the contract by a discrete binomial tree, and calculate the probabilities for all the increasing and decreasing behaviors, by using the fact that the option value is equal to the payoff at the final date, deriving all the option value backwards, the option price could be calculated.

When there is jump in the price behavior, we can extend the binomial tree to a grid, and we assume that at each time step, there are several nodes that the price can achieve, other than the "up" or "down" nodes in binomial tree case. For all the exceptional nodes, we assume that the behavior is caused by the jump of the price. Under this assumption, it is still possible to take backward derivation and calculate the final option price.

The binomial pricing model arises from discrete random walk models of the underlying asset. Jarrow and Turnbull (1996, see [18]) pointed out that this method provides reasonable estimation when the number of trading intervals is large and the time between trades is small. The model was first introduced by Sharpe (1978), and later developed in detail in the most famous binomial tree model, the CRR (Cox-Ross-Rubinstein) model, in 1979. Binomial tree models are easy to program, and we can easily understand the result from the structure of the tree. At the start, the practical use of binomial tree models is the pricing of American-style options, and it is still of great importance now.

K. I. Amin developed the multinomial tree model in 1993 ([3]), to fit the binomial tree scheme to models with jumps.

A. Lari-Lavassani et al. ([23]) developed the method for constructing a binomial tree under model 1F and the forest version under model 2F, and they applied these models to price swing option prices. The convergence of the method was also proved via finite-difference method. For the jump case, K. I. Amin developed a discrete multinomial tree scheme for pricing derivatives with asset prices under jump-diffusion process. This method can be an extension for binomial tree methods. Combining these two methods together, the multinomial tree under model 1FJ can be developed.

6.3.1 One-Step Binomial Model

An one step binomial model is shown in figure 6.4. Consider the asset price to be S at the start, and after one step, the price can have only two states: rising up to S_u or dropping down to S_d. Assume that there is a risk-neutral measure, such that the probability of rising up is p_u, and the probability of dropping down is p_d. Here $p_u + p_d = 1$ since there could only be two possible states in the next step. Suppose that there exists a derivative, which is exercisable at the future state, and that the payoff at the future state is given by $g(X)$ for asset price X. Finally, we assume that the risk-free interest rate is a constant r, exercised continuously.

Under the risk-neutral measure, given the payoff function g, the price of the derivative can be give by the discounted value of expectation of the future payoffs at the initial state. i.e., If the time between the two states is Δt, the price of the derivative can be given as

$$V = p_u g(S_u) + p_d g(S_d).$$

Note that the Markov property is satisfied in our models, i.e., in the previous model, after one step, if we use S_u and S_d as new initial values, we can derive a two-period

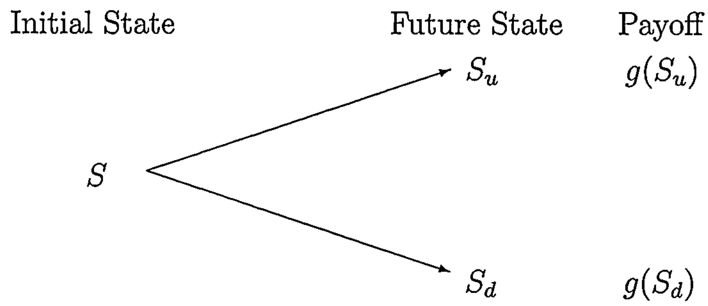

Figure 6.4: An One Step Binomial Model
(see [15])

binomial model. If we continue developing new states, we can construct a multi-period binomial model, i.e., the binomial tree model.

6.3.2 Construction of Binomial Trees

There can be two different kinds of binomial tree models: the nonrecombining tree and the recombining tree (see [23]). In nonrecombining tree, the nodes indicating the prices can never be coincided, while this is possible in recombining tree (see figure 6.5 and figure 6.6). In either cases, the tree can be expressed as $\{S_j^{(i)}\}$, where i is the state (or step), j is the index of possible values.

In this thesis we choose the recombing tree model, since this reduces the time and space complexity, while the accuracy could still be guaranteed.

In the construction of a recombining binomial tree, the prices at state i can be given as

$$S^{(i)} = \{S_j^{(i)}\}_{j=1}^{i},$$

i.e., there are i nodes in state i. The starting price is $S_1^{(1)} = S_0$. We assume that the transition probabilities are $p_u^{i,j}$ for the price $S_j^{(i)}$ to be rising, and $p_d^{i,j}$ for descending.

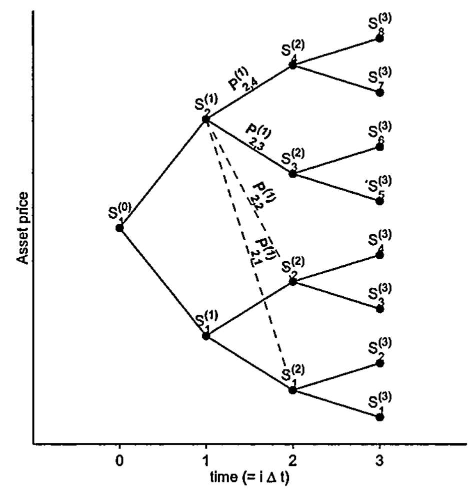

Figure 6.5: A Nonrecombining Binomial Tree
Note: Dashed lines indicate zero values for the possibilities. Solid lines indicate possible transactions. (from [23])

Binomial tree models can be a good approximation of stochastic models, since sometimes it might be hard to give an analytical solution of the price of a derivative. Examples are American options and swing options. We assume that the time length between two states be Δt. Note that the accuracy depends only on Δt, and the smaller Δt is, the more accurate results we could get.

6.3.3 Binomial Tree under Model 1F

For model 1F, A. Lari-Lavassani et al. proved the following equations in [23], which provides the essential theorem for binomial tree models under model 1F.

Theorem 6.3-1. *(Binomial tree component for Model 1F) For the stochastic differential in equation (2.3),*

$$dS_t = \alpha(L - S_t)dt + \sigma S_t dZ_t,$$

with constant coefficients α, L, σ over the time interval $[0, T]$, Let $S^{(i)} = (S_j^{(i)})$ be a

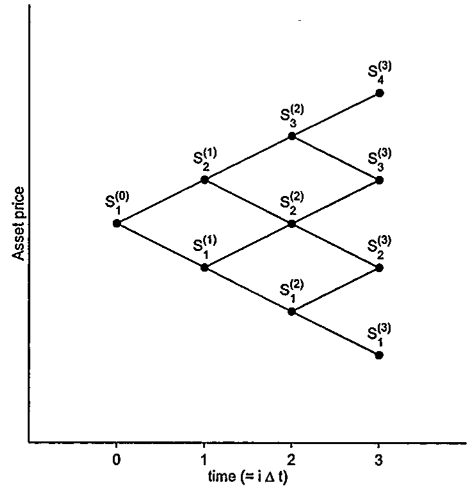

Figure 6.6: A Recombining Binomial Tree
(from [23])

recombining binomial tree, which describes this jump-diffusion equation discretely, with time step Δt. Then the transition probabilities and the prices satisfies

$$\begin{aligned} p_u^j &= \tfrac{1}{2} + (\alpha \tfrac{L - S_j^{(i)}}{2\sigma S_j^{(i)}} - \tfrac{\sigma}{4})\sqrt{\Delta t}, \\ S_{j+1}^{(i+1)} &= S_j^{(i)} e^{\sigma\sqrt{\Delta t}}, \\ S_j^{(i+1)} &= S_j^{(i)} e^{-\sigma\sqrt{\Delta t}}. \end{aligned} \tag{6.8}$$

By applying theorem (6.3-1), a binomial tree can be constructed. In the equation, Δt is independent of the nodes of the tree, and therefore the up and down sizes u_j^i and d_j^i are constant throughout the whole tree. This forces the tree to be recombining. Furthermore, $p_u^j \in [0, 1]$ when Δt is small enough. Thus, the constructed tree is numerically efficient. (see [23])

Now consider a swing option contract following the assumptions in section 5.2. For the recombining tree, we can construct the Snell envelope X as

$$X = \{X_{i,j}^{(k)}\},$$

where $X_{i,j}^{(k)}$ is the corresponding Snell envelope of node j in state i, when the holder still has k rights remaining. Now if we assume the swing option to be a swing put option, the Snell envelope will satisfy the equations below.

$$X_{i,j}^{(0)} = 0, \text{ for any state } i \text{ and any node } j, \tag{6.9}$$

$$X_{Nk,j}^{(i)} = (K - S_{Nk}^{(j)})_+ - \Phi(i), \text{ for any node } j, \text{ for all } i = 1, 2, \ldots, n, \tag{6.10}$$

$$\begin{aligned} &\text{for all } 1 \leq t < Nk, \text{ for any } j = 1, 2, \ldots, t, \text{ for all } i = 1, 2, ..., n, \\ X_{t,j}^{(i)} &= \max((K - S_t^{(j)})_+ + (p_u^{t,j} X_{n+1,j+1}^{(i-1)} + p_d^{t,j} X_{t+1,j-1}^{(i-1)}), \\ &\quad p_u^{t,j} X_{t+1,j+1}^{(i)} + p_d^{t,j} X_{t+1,j-1}^{(i)}). \end{aligned} \tag{6.11}$$

Note that in binomial tree models, there can only be two possible transactions for each node, and therefore the conditional expectation in equation (5.9) can be simply expressed as equation (6.11).

The algorithm of the binomial tree method on swing option pricing is as below.

- Decide the number of steps between two exercise dates. Let it be k. Given the length of the contract T, and the number of exercisable dates N, the time length of one step is $\Delta t = \frac{T}{Nk}$.
- Construct a recombining binomial tree under the given parameters in equation (2.3).
- Calculate all the values of the Snell envelope $X = \{X_{i,j}^{(k)}\}$ associated with the binomial tree.
- The swing option price is given by $X_{1,1}^{(n)}$.

6.3.4 Multinomial Tree under Model 1FJ

The binomial tree constructed using theorem (6.3-1) is for model 1F. Now if we consider model 1FJ, the only difference between model 1F and model 1FJ is the jump part, and the stochastic differential equation is

$$dS_t \;=\; \alpha(L - S_t)dt + \sigma S_t dZ_t + U_t S_t dP_t.$$

As showed in chapter 2, if we assume the price at t is S_t, after a small interval Δt, if the price is $S_{t+\Delta t}$ by model 1F, the price by model 1FJ would be $(1 + U_{t+\Delta t})\Delta P_t S_{t+\Delta t}$.

Now since the jumps are independent of asset prices, and by theorem 6.3-1, the possible prices in each state i will always be of the form $S_0 e^{j\sigma\sqrt{\Delta t}}$, where j is an integer, and that $-i \leq j \leq i$. If there are Nk states, j will be between $-Nk$ and Nk, and there are $2Nk + 1$ accessible prices. In binomial tree models, however, there will be only i prices that can be accessed with non-zero probability. Now if we assume that all the accessible prices have non-zero transaction probability, it is possible to introduce jumps in the pricing process, and the model will be a multinomial tree.

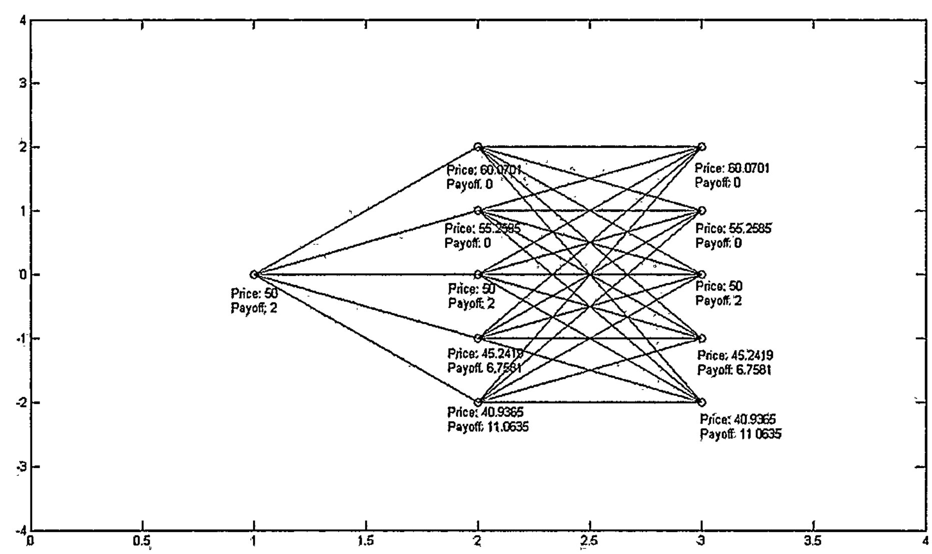

Figure 6.7: A Multinomial Tree

In figure 6.7 we can see a multinomial tree with initial price 50, and we consider a two-step swing put option with strike price 52, with 5 accessible nodes in each states after the initial state. The payoff in each node is calculated, and by calculating the expected payoff, it is possible to find out the swing option price. Now the only problem is the transaction probabilities.

First, we present the tree as $\{S_j^{(i)}\}_{j=-D}^{D}$, where $S_j^{(i)}$ is the price of node j in state i, and

$$S_j^{(i)} = S_0 e^{j\sigma\sqrt{\Delta t}}.$$

There are $2D+1$ possible nodes in each state, except that in state 1 there can be only 1 initial node. Clearly, if j is positive, $S_j^{(i)} > S_0$. If it is negative, $S_j^{(i)} < S_0$, and $S_0^{(i)} = S_0$. We write the transaction probabilities as $p_k^{i,j}$, which is the probability of the transition from node j in state i to node k in state $i+1$.

Second, since P_t is a Poisson process, dP_t could be approximated by a binomial distribution, with probability $\eta\Delta t$ with time step Δt. i.e., the probability of jump in each step is $\eta\Delta t$.

For the transition, we assume that the transition from node j in state i to node $j+1$ or node $j-1$ in state $i+1$ are the "no jump" transitions, i.e., they are up and down transitions which are the same as the moves in binomial tree models. Therefore by theorem 6.3-1, for any state $i = 1, 2, \ldots, Nk$, for any $j = -D, -D+1, \ldots, 0, 1, \ldots, D$,

$$p_{j+1}^{i,j} = \begin{cases} (1-\eta\Delta t)\left(\frac{1}{2} + \left(\alpha\frac{L-S_{(i)}^{(j)}}{2\sigma S_j^{(i)}} - \frac{\sigma}{4}\right)\sqrt{\Delta t}\right) & \text{when } -D < j < D \\ 1-\eta\Delta t & \text{when } j = D \end{cases} \tag{6.12}$$

$$p_{j-1}^{i,j} = \begin{cases} (1-\eta\Delta t)\left(\frac{1}{2} - \left(\alpha\frac{L-S_{(i)}^{(j)}}{2\sigma S_j^{(i)}} - \frac{\sigma}{4}\right)\sqrt{\Delta t}\right) & \text{when } -D < j < D \\ 1-\eta\Delta t & \text{when } j = -D \end{cases} \tag{6.13}$$

Note that in no jump cases, the probability of transition is equal to the probability of both no jump and the transition probabilities in theorem 6.3-1. In node N and node $-N$ (top and bottom nodes), there can be only one possible transition.

Now by [3], the transition probabilities in jump case can be developed in as below.

First, by the equation of model 1FJ, in equation (2.14), for a given price S_t at time

t, after a jump at that time, the price will be $(1+U_t)S_t$. i.e., $(1+U_t)$ is the distribution which produces the jump amplitude. Let the cumulative density function of $\log(1+U_t)$ be $\zeta(x)$, where

$$\zeta(x) = \mathbb{P}(\log(1+U_t) \leq x).$$

Second, the jump probability in any state will always be Δt. The transition probability in jump case is the probability of jump time the probability of jump amplitude, and by [3], in state i, node j,

$$p^{i,j}_{j+k} \tag{6.14}$$
$$= \begin{cases} \eta\Delta t\left(\zeta\left(\left(1+\frac{1}{2}\right)\sigma\sqrt{\Delta t}\right) - \zeta\left(-\left(1+\frac{1}{2}\right)\sigma\sqrt{\Delta t}\right)\right), & \text{when } k=0, \\ \eta\Delta t\left(\zeta\left(\left(k+\frac{1}{2}\right)\sigma\sqrt{\Delta t}\right) - \zeta\left(\left(k-\frac{1}{2}\right)\sigma\sqrt{\Delta t}\right)\right), & \text{when } k \notin \{-1,0,1\}. \end{cases}$$

Finally, by equation (6.12) and equation 6.14, for each state i, any node j, the transition probabilities can be given as

$$p^{i,j}_{j+k}$$
$$= \begin{cases} \eta\Delta t\left(\zeta\left(\left(1+\frac{1}{2}\right)\sigma\sqrt{\Delta t}\right) - \zeta\left(-\left(1+\frac{1}{2}\right)\sigma\sqrt{\Delta t}\right)\right), & \text{when } k=0, \\ (1-\eta\Delta t)\left(\frac{1}{2} + \left(\alpha\frac{L-S^{(j)}_{(i)}}{2\sigma S^{(i)}_{j}} - \frac{\sigma}{4}\right)\sqrt{\Delta t}\right), & \text{when } k=1,\ -D<j<D \\ 1-\eta\Delta t, & \text{when } k=1,\ j=D \\ (1-\eta\Delta t)\left(\frac{1}{2} - \left(\alpha\frac{L-S^{(j)}_{(i)}}{2\sigma S^{(i)}_{j}} - \frac{\sigma}{4}\right)\sqrt{\Delta t}\right), & \text{when } k=-1,\ -D<j<D \\ 1-\eta\Delta t, & \text{when } k=-1,\ j=-D \\ \eta\Delta t\left(\zeta\left(\left(k+\frac{1}{2}\right)\sigma\sqrt{\Delta t}\right) - \zeta\left(\left(k-\frac{1}{2}\right)\sigma\sqrt{\Delta t}\right)\right), & \text{otherwise.} \end{cases}$$

For the swing put option under this model, the Snell envelope will be

$$X^{(0)}_{i,j} = 0, \text{ for any state } i \text{ and any node } j, \tag{6.15}$$

for any node j, for all $i = 1,2,\ldots,n$,

$$X^{(i)}_{Nk,j} = \max(\Phi(i), (K - S^{(j)}_{Nk})_+ - \Phi(i-1)), \tag{6.16}$$

for all $0 \leq t < Nk$, for any $j = 1,2,\ldots,t$, for all $i = 1,2,\ldots,n$,

$$X_{t,j}^{(i)} = \max((K - S_t^{(j)})_+ + \sum_{k=-D}^{D} p_k^{t,j} X_{t+1,k}^{(i-1)}, \sum_{k=-D}^{D} p_k^{t,j} X_{t+1,k}^{(i)}). \tag{6.17}$$

The algorithm is as below.

- Decide the number of steps between two exercise dates. Let it be k. Given the length of the contract T, and the number of exercisable dates N, the time step is $\Delta t = \frac{T}{Nk}$.

- Decide D. Note that D should be chosen carefully, and should be associated with Δt, otherwise when applying equation (6.12), there might be values of probabilities that are not in $[0, 1]$.

- Construct a multinomial tree under the given parameters in equation (2.3).

- Calculate all the values of the Snell envelope $X = \{X_{i,j}^{(k)}\}$ associated with the binomial tree.

- The swing option price is given by $X_{1,1}^{(n)}$.

6.4 Numerical Example: European Call Option Pricing

In this section, we generate an "empirical" spot price path under model 1FJ, and calibrate the parameters for model 1F and model 2F, then use Monte Carlo method to calculate European call option prices.

The parameters and the term of European call option contracts are as below.

- Parameters

  $S_0 = 2, L = 3, \alpha = 2, \sigma = 1, \eta = 1.5, \mu_J = 0.2, \sigma_J = 0.2, \Delta t = \frac{1}{360}$

  Constant risk free interest rate: $r = 5\%$

  Expiry: 6 months($T = 0.5$)

Strike price: K=2.1

Simulation time: n=16384, m=10

- "Empirical data" simulated by one factor model with jumps
- Maximum likelihood estimation for one factor model without jumps $\alpha = 2.3218$, $\sigma = 1.0224$, $L = 3.1698$
- Maximum likelihood estimation for two factor model without jumps $\alpha = 15.147$, $\sigma = 1.0857$, $\mu = 0.19223$, $\xi = 0.30846$, $L_0 = 2.9257$

Figure 6.8 shows the "Empirical" price path used in this section.

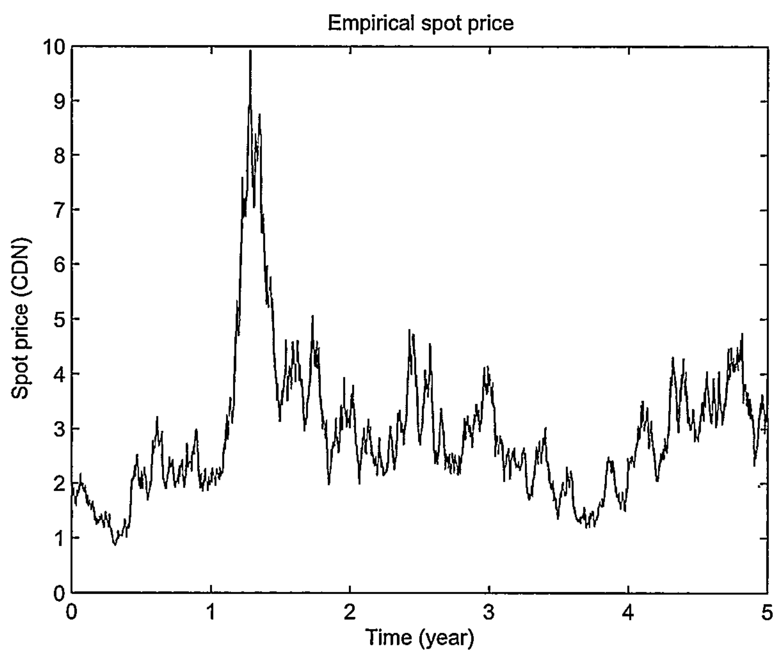

Figure 6.8: The "Empirical" Price Path for European Call Option Pricing

The simulation results[1] of European call option prices are given in table 6.1. The results are achieved by Monte Carlo and quasi-Monte Carlo simulation with variance reduction techniques, and the estimate of European call option price $\hat{C}_T$, the variance of

[1]The results are calculated under Matlab 7.0. Program codes are available if requested.

the simulation results, the length of the 95% confidence interval and the 95% confidence interval are provided in the table. Here for randomized quasi-Monte Carlo method, we use the Sobol sequence (see section 6.1.1 and section 6.1.3).

	Model	$\hat{C}_T$	Variance	95% half CI	95% CI
MC C+A	1F	0.81725	$1.2584e-06$	0.0006953	$(0.81655, 0.81795)$
RQ C+A	1F	0.81732	$7.788e-07$	0.00054698	$(0.81677, 0.81787)$
MC C+A	1FJ	0.96055	$2.9206e-05$	0.0033496	$(0.9572, 0.9639)$
RQ C+A	1FJ	0.96084	$2.008e-05$	0.0027774	$(0.95806, 0.96361)$
MC C+A	2F	1.0725	$7.5093e-08$	0.00016985	$(1.0723, 1.0726)$
RQ C+A	2F	1.0725	$8.3094e-08$	0.00017867	$(1.0724, 1.0727)$

Table 6.1: The Monte Carlo Simulation Results of European Call Option Pricing (C: control variate, A: antithetic variates, RQ: randomized quasi-MC)

We can see from table 6.1 that actually randomized Sobol sequence could improve the simulation results, but not much, and sometimes even a little worse (in model 2F). Actually this often happens when we combine randomized quasi-monte carlo method with control variates.

The accuracy of results under model 1FJ is the worst among the three cases, this could result from the highest dimension of random variables in model 1FJ. It may also be because that we do not have good control variate for S_t in model 1FJ.

We can also see that the simulation results are different for different models. This is because of the loss of accuracy in the calibration of parameters. It is actually more important to improve the calibration methods to get better estimation of option prices.

6.5 Numerical Example: Swing Put Option Pricing

In this section we provide some sample results of swing option pricing under Monte Carlo method and binomial tree method.

6.5.1 Swing Put Option Pricing under Model 1F

Here we implement model 1F for asset price, and we assume that the model is risk-neutral, with form

$$dS_t = \alpha(L - S_t)dt + \sigma S_t dZ_t,$$

where the parameters of the model and the parameters in the contract are given as

Parameter	Value
α	10
L	45
σ	1.5
S_0	40
K	50

Table 6.2: Parameters for Swing Put Option Pricing Under Model 1F

We assume that the contract lasts for 24 months, with 30 days for each month, and the holder has the right to exercise at the first day of each month. The risk free interest rate $r = 0$.

The results from Monte Carlo simulation and binomial tree are given in table 6.3.

Rights	Monte Carlo Result	95% Confidence Interval	Binomial Tree Result
1	27.7756	(27.503, 28.0481)	27.5725
2	53.157	(52.7227, 53.5913)	52.7001
3	77.0065	(76.4524, 77.5606)	75.8713
4	98.9669	(98.3136, 99.6202)	97.7845
5	120.5241	(119.7097, 121.3385)	118.7301
6	140.3239	(139.4026, 141.2451)	138.3465
7	158.4772	(157.4191, 159.5353)	156.6262
8	174.4729	(173.2385, 175.7073)	173.8667
9	190.687	(189.303, 192.0709)	190.1361
10	206.8155	(205.3053, 208.3256)	205.1836

Table 6.3: Results of Swing Put Option Pricing Under Model 1F

Figure 6.9 shows the graph of the swing put option prices.

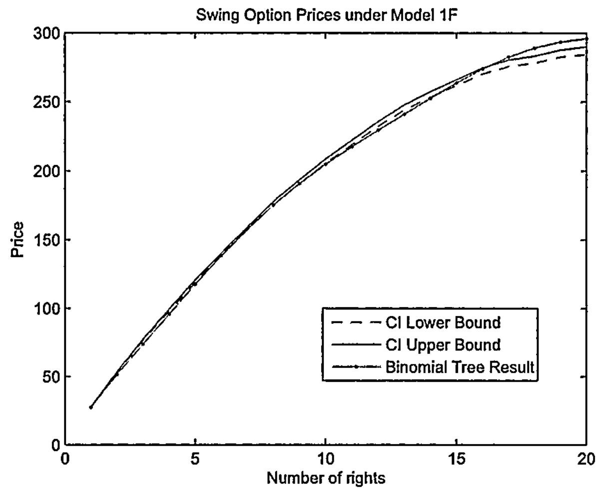

Figure 6.9: Swing Put Option Prices Under Model 1F

6.5.2 Swing Put Option Pricing under Model 1FJ

In this section, we use model 1FJ for asset price, with form

$$dS_t = \alpha(L - S_t)dt + \sigma S_t dZ_t + U_t S_t dP_t.$$

Here we assume that U_t follows the lognormal distribution, i.e.,

$$\log(1 + U_t) \sim N(\mu_j, \sigma_j).$$

The parameters are given in table 6.4.

Parameter	Value
α	10
L	45
σ	1.5
η	0.5
μ_j	0
σ_j	1
S_0	40
K	50

Table 6.4: Parameters for Swing Put Option Pricing Under Model 1FJ

The assumption of the swing contract is the same as in model 1F. We assume that the length of the contract is 24 months, with 30 days for each month. The holder has the right to exercise at the first day of each month. The risk free interest rate r is 0.

The results from Monte Carlo simulation and binomial tree are given in table 6.5.

Rights	Monte Carlo Result	95% Confidence Interval	Binomial Tree Result
1	28.221	(27.9231, 28.5189)	27.5725
2	54.1416	(53.7044, 54.5789)	52.7001
3	78.5583	(77.9668, 79.1499)	75.8713
4	100.7976	(100.0637, 101.5314)	97.7845
5	121.7151	(120.856, 122.5742)	118.7301
6	141.4391	(140.4671, 142.411)	138.3465
7	160.2973	(159.185, 161.4097)	156.6262
8	177.8913	(176.6118, 179.1709)	173.8667
9	193.2613	(191.872, 194.6507)	190.1361
10	207.517	(206.0043, 209.0298)	205.1836

Table 6.5: Results of Swing Put Option Pricing Under Model 1F

Figure 6.10 shows the swing put option prices with different number of rights.

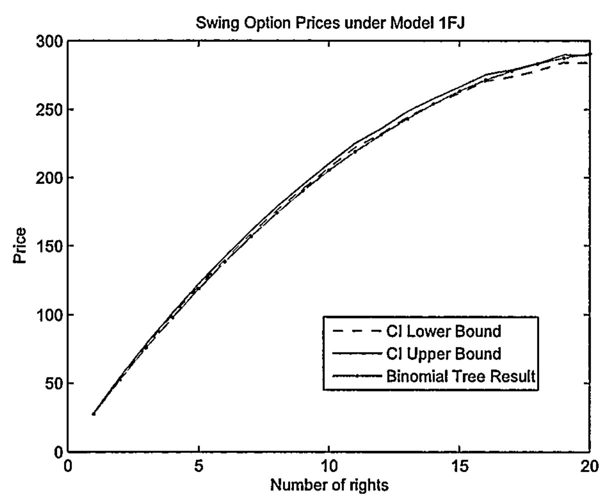

Figure 6.10: Swing Put Option Prices Under Model 1FJ

Chapter 7

Conclusion

Mean reverting models are widely used in mathematical finance research to describe the random walk of spot prices. In energy markets, the price behavior is different from the behavior in money markets, and mean reverting models are used to describe the price of energy assets.

Pilipoviô's model are mean reverting models to describe the spot price in energy markets. It is assumed that there is a long-term equilibrium price which affects the spot price. In one-factor model, the equilibrium price is defined as a constant, while in two-factor model it follows a geometric Brownian motion.

In Pilipoviô's model, the spot prices are assumed to be continuous. However, in the real world, the price may be discontinuous at certain times, and we will discover a jump in the spot price curve. The price may jump up to a higher price, or jump down to a lower price. To make the model be more realistic, it is necessary to consider the jump component in the model.

Options are very popular investments in financial investment. European option is one of the most famous option in finance markets, and the pricing of European options is a popular area in finance research. However, in mean-reverting world, it is always difficult to solve the accurate explicit formula of the option price. Therefore, many kinds of approximation methods are provided to achieve approximate results.

Swing options provide more flexible choices than European options, and are widely used in energy markets. However, since there could be many different conditions in a swing option contract, it has much more complicated payoff function than European option.

In the thesis, the Pilipović's mean-reverting models are used to model the underlying asset, including the one factor model, the two factor model, and the extended one factor model with jumps. The moments of the no jump models are calculated, and the European option pricing formula for one factor case is given in integral form. The Monte Carlo method, and the binomial and multinomial tree method are implemented in the pricing of swing options.

The Monte Carlo method used in the thesis is based on the assumption that the continuation value function can be regressed by Longstaff-Schwartz method, and the prices are achieved through the method by N. Meinshausen and B. M. Hambly. The binomial tree method is used to calculate swing option price under no jump models, based on the paper by A. Lari-lavassani, M. Simchi and A. Ware. The multinomial tree method is an extension of binomial tree method using K. I. Amin's jump-diffusion scheme, which can be applied in jump cases.

There are still some more work to do when applying this thesis on swing option pricing. First, modelling the price behavior is a fundamentally problem in mathematical finance, especially in energy markets when seasonality components and jump components are included. There is no best models in price behavior. Second, the explicit formula of European option price provided in the thesis is of an integral form, and numerical methods are still required to achieve the value of European option prices. Furthermore, the swing option pricing under a general swing option contract is a difficult and complicated work. What is remaining is still a long way and we need to make full effort in solving mathematical finance problems.

Appendix A

Probability Background

A.1 Stochastic Process, Stochastic Differential Equations

It is always difficult to make accurate prediction of the prices in a market, since there is randomness in the behavior of market prices. We can assume that the price follows an underlying stochastic process, and use the process to describe the price behavior. The empirical data of the prices can be seen as a path of practice under the stochastic process.

A.1.1 Stochastic Process

In this subsection, we will introduce some probability definitions regarding stochastic process. These definitions can be seen in most probability theory and stochastic process books. See [5], [20].

Definition A.1-1. (Probability Space) A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a measure space with a probability measure $\mathbb{P}$. Here Ω is the sample space, which is a nonempty set that consists all the possible states, and $\mathcal{F}$ is a σ-algebra of Ω, which contains all the events.

Definition A.1-2. (Stochastic Process) Consider an index set $T \subseteq \overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$. Fix an probability space $(\Omega, \mathcal{A}, \mathbb{P})$ in the index set T. A stochastic process $X = \{X_t\}_{t \in T}$ is a collection of random variables in $(\Omega, \mathcal{A}, \mathbb{P})$, indexed by $t \in T$.

Definition A.1-3. (Filtration) For given index set $T \subseteq \overline{\mathbb{R}}$ and probability space $(\Omega, \mathcal{A}, \mathbb{P})$, the filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t \in T}$ is a family of nondecreasing σ-fields $\mathcal{F}_t$, such that for all $t \in T$, $\mathcal{F}_t \subseteq \mathcal{A}$. Furthermore, consider a stochastic process $X = \{X_t\}_{t \in T}$ in $(\Omega, \mathcal{A}, \mathbb{P})$, we say that X is adapted to $\mathcal{F}$ if X_t is $\mathcal{F}_t$-measurable for all $t \in T$. The smallest filtration of

X is the collection of generated σ-fields, namely, $\mathcal{F}_t = \sigma\{X_s | s \le t\}$. This is called the induced filtration or generated filtration.

Definition A.1-4. (Conditional Expectation) Consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and a random variable X. Let $\mathcal{F}$ be a σ-field such that $\mathcal{F} \subseteq \mathcal{A}$. Then there exists a unique $\mathcal{F}$-measurable, integrable random variable Y such that for all $G \in \mathcal{G}$, $\mathbb{E}[Y|G] = \mathbb{E}[X|G]$. We say Y is a conditional expectation of X conditional on $\mathcal{G}$, and write $Y = \mathbb{E}[X|\mathcal{G}]$.

The filtration of X can be seen as the information of all the possible moves of the process X. In mathematical finance, we treat the spot price $S = \{S_t\}_{t \in T}$ as a stochastic process, and use the induced filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t \in T}$, $\mathcal{F}_t = \sigma\{S_s | s \le t\}$. The filtration $\mathcal{F}$ can be treated as all the information of the price, i.e., $\mathcal{F}_t$ is all the information of the prices before time t. The index set T is normally $\mathbb{R}^+$ when we use continuous models, and in discrete valuation we can let $T = \mathbb{Z}^+$.

The conditional expectations can be seen as the estimate of the random variable under give information. For example, consider the stochastic process S of spot prices, and the induced filtration $\mathcal{F}$, $\mathbb{E}[S_t|\mathcal{F}_s]$ is the estimate of the price at time t, given all the prices before time s.

A.1.2 Brownian Motion

Brownian motions are fundamental components in stochastic calculus. The definition of a standard Brownian motion is as below (see [36]).

Definition A.1-5. A standard Brownian motion (or a standard Wiener process) is a stochastic process $\{W_t\}_{t \in \mathbb{R}^+}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the following 4 properties:

- $W_0 = 0$.
- The function $t \to W_t$ is continuous in $\mathbb{R}^+$ with probability 1.

- The process $\{W_t\}_{t\in\mathbb{R}^+}$ has stationary and independent increments.

- For any $t, s \in \mathbb{R}^+$, $W_{t+s} - W_s \sim N(0, t)$.

There are several properties of the standard Brownian motion W_t. Below we provide some propositions of W_t.

- $-W_t$ is also a standard Brownian motion.

- For any fixed $s \in \mathbb{R}^+$, the process $\{W_{t+s} - W_s\}_{t\in\mathbb{R}^+}$ is a standard Brownain motion.

- For any $a \in \mathbb{R}$, the process $\{aW_{t/a^2}\}_{t\in\mathbb{R}^+}$ is a standard Brownian motion.

A.1.3 Itô Integral

Give a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, consider a measurable, adapted and L^2-integrable function $f(t, \omega)$ where $\omega \in \Omega$. Let W_t be a standard Wiener process. A Itô integral can be defined as the following (see [30]):

Definition A.1-6. (Itô Integral) For the given probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and function $f(t, \omega)$ where $\omega \in \Omega$, and given W_t a standard Wiener process, in interval $[0, t]$, we split $[0, t]$ to n partitions using $0 = u_0 \leq u_1 \leq \ldots \leq u_n = t$, and then define

$$\int_0^t f(u, \omega) dW_u = \lim_{n\to\infty} \sum_{i=1}^{n} f(u_i^*, \cdot\,)[W_{u_{i+1}} - W_{u_i}], \tag{A.1}$$

where $u_i^* = u_i$, and the limit is taken in L^2 sense. When the limit exists, the integral exists, and we say $\int_0^t f(u, \omega) dW_u$ is an Itô Integral.

Note that in definition A.1-6, we choose $u_i^* = u_i$, which is different from the choice in Riemann-Stieltjes integral. In Riemann-Stieltjes integral, we choose middle points $u_i^* = \frac{u_{i+1} - u_i}{2}$, and actually the choice of u_i^* will not change the value of integral as long as $u_i^* \in [u_i, u_{i+1}]$. However, in stochastic integrals, the choice of u_i^* makes difference in the values of integrals. If we choose $u_i^* = \frac{u_{i+1} - u_i}{2}$ in definition A.1-6, the integral is named the Stratonovich integral.

After the definition of Itô Integral, we can introduce stochastic differential equations. When we have

$$X_t = X_s + \int_s^t f(u, \omega) dW_u,$$

we can write it in stochastic differential equation form

$$dX_t = f(t, \omega) dW_t.$$

A.1.4 Stochastic Differential Equations

stochastic differential equations are widely used to describe the price behavior of a market, and there are several methods developed to price derivatives with asset price under stochastic differential equations. Different from regular differential equations, which have deterministic solutions, in stochastic differential equation, there is one or more terms, which is stochastic process. Thus, the solution of a stochastic differential equation is a stochastic process. Most definitions and fundamental theorems can seen found in [30].

Normally, the stochastic differential equations used in mathematical finance are diffusion equations, and in the form

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

where W_t is the Wiener process. Here in the equation, $\mu(X_t, t)$ is the drift term, and $\sigma(X_t, t)$ is the volatility.

The process X_t in equation (A.1.4) is called a diffusion process, and it is usually a Markov process, i.e.

$$\mathbb{P}(X_{t+h} = y | X_s = x(s),\ \forall s \leq t) = \mathbb{P}(X_{t+h} = y | X_t = x(t)).$$

The equation (A.1.4) can also be written in integral form as

$$X_{t+s} - X_s = \int_s^{t+s} \mu(X_u, u)du + \int_s^{t+s} \sigma(X_u, u)dW_u,$$

where the second integral $\int_s^{t+s} \sigma(X_u, u)dW_u$ is the Itô integral.

The most important property of stochastic differential equations is Itô's lemma. The lemma is first proved by K. Itô ([16]). Here we use the description in [29].

Theorem A.1-7. *(Itô's Lemma) Given $f(X_t, t)$ a twice-differentiable function of t and X_t, where X_t is a stochastic process following the stochastic differential equation*

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t, \tag{A.2}$$

with well behaved drift and volatility functions, $\mu(X_t, t)$ and $\sigma(X_t, t)dW_t$, then we have

$$df(X_t, t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X_t}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}\sigma^2(X_t, t)dt, \tag{A.3}$$

or, after substituting dX_t,

$$df(X_t, t) = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial X_t}\mu(X_t, dt) + \frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}\sigma^2(X_t, t)\right) dt + \frac{\partial f}{\partial X_t}\sigma(X_t, dt)dW_t. \tag{A.4}$$

Itô's lemma can be treated as the product rule in stochastic calculus. In stochastic calculus, the derivative rules are changed. For example, the product rule has an additional term.

Theorem A.1-8. *(Itô's Product Rule) Given that X_t and Y_t are two Itô stochastic process, then*

$$d(X_tY_t) = Y_tdX_t + X_tdY_t + dX_tdY_t. \tag{A.5}$$

A.1.5 Stochastic Differential Equations with Jumps

A stochastic differential equation with jumps has the form

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t + \nu(X_t, t)dN_t, \tag{A.6}$$

with an extra jump component $\nu(X_t, t)dN_t$. In the jump component, $\nu(X_t, t)$ is the jump amplitude, and N_t is the distribution of the number of jumps until time t.

We can write equation (A.8) in an integral form as

$$X_{t+s} - X_s = \int_s^{t+s} \mu(X_u, u)du + \int_s^{t+s} \sigma(X_u, u)dW_u + \sum_{i=1}^{N_t} \nu(X_{\tau_i}, \tau_i), \tag{A.7}$$

where τ_i are the time when a jump occurred.

The Itô's lemma for equation (A.8) is (see [2])

Theorem A.1-9. *(Itô's lemma with jumps) Given $f(X_t, t)$ a twice-differentiable function of t and X_t, where X_t is a stochastic process following the stochastic differential equation*

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t + \nu(X_t, t)dN_t, , \tag{A.8}$$

then we have

$$\begin{aligned} f(X_{t+s}) &= f(X_s) + \textstyle\int_s^{t+s} f'(X_{u^-})dX_u + \frac{1}{2}\int_s^{t+s} f''(X_{u^-})\sigma^2(X_u, u)du \\ &\quad - \textstyle\sum_{i=1}^{N_t} f'(X_{\tau_i^-})(X_{\tau_i} - X_{\tau_i^-}) + \sum_{i=1}^{N_t}(f(X_{\tau_i}) - f(X_{\tau_i^-})), \end{aligned} \tag{A.9}$$

or in stochastic differential equation form

$$\begin{aligned} df(X_t) &= f'(X_{t^-})dX_t + \tfrac{1}{2}f''(X_{t^-})\sigma^2(X_t, t)dt \\ &\quad - f'(X_{\tau_i^-})(X_{\tau_i} - X_{\tau_i^-})dN_t + (f(X_{\tau_i}) - f(X_{\tau_i^-}))dN_t. \end{aligned} \tag{A.10}$$

Here in τ_i denotes the jump time, and τ_i^- denotes the time infinitesimally before the jump time τ_i.

A.1.6 Martingales

Martingales are stochastic processes, with the following property (see [20]):

Definition A.1-10. (Martingale) Given probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and an index set $T \subseteq \overline{\mathbb{R}}$, and a filtration $\mathcal{F}$ on T. We say a stochastic process $M = \{M_t\}_{t\in T}$ is a martingale if it is an integrable, $\mathcal{F}$-adapted process such that

$$\mathbb{E}\left[M_t | F_s\right] = M_s, \text{ for all } s, t \in T \text{ such that } s < t.$$

Martingales are important in the non-arbitrage pricing of financial derivatives. For example, the discounted stock price is assumed to be a martingale in Black-Scholes models. For applications of martingales in mathematical finance, see [14], [28], [29], [37], etc.

A.1.7 Statistics of Random Variables

For a given random variable x, the moments of x are used to describe statistical properties of x. The mean value, the variance, the skewness and the kurtosis of x are related to the first 4 moments of x, which are mostly used (see [26]).

- Mean: Mean(x)=$\mathbb{E}[x]$.
- Variance: $\mathrm{Var}(x) = \mathbb{E}[(x-\mu)^2]$, where $\mu = \mathbb{E}[x]$.
- Skewness: $\mathrm{Skew}(x) = \frac{\mathbb{E}\left[(x-\mu)^3\right]}{\sigma^3}$, where $\mu = \mathbb{E}[x]$, and $\sigma^2 = \mathrm{Var}(x)$.

  Skewness is a measure of the asymmetry of the probability distribution of x (see figure A.1).

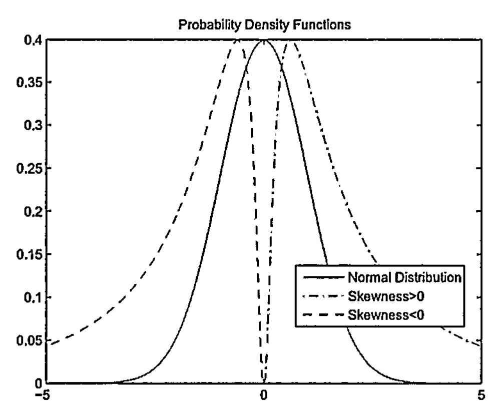

Figure A.1: The Probability Density Function of Distributions with Different Skewness

- Kurtosis: $\mathrm{Kurto}(x) = \frac{\mathbb{E}\left[(x-\mu)^4\right]}{\sigma^4} - 3$, where $\mu = \mathbb{E}[x]$, and $\sigma^2 = \mathrm{Var}(x)$.

  Kurtosis is a measure of the "peakedness" of the probability distribution of x (see figure A.2).

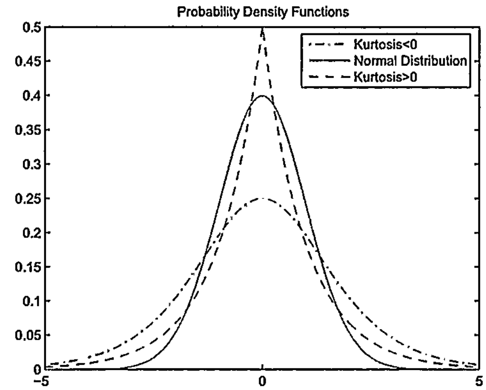

Figure A.2: The Probability Density Function of Distributions with Different Kurtosis

The skewness and kurtosis of a standard normal distribution are both 0.

The mean, variance, skewness and kurtosis can describe the approximate shape of the samples of random variable x.

Appendix B

Proofs of Theorems

B.1 Proof of Theorem 2.3-2

For model 2F, we have

$$\begin{cases} dS_t &= \alpha(L_t - S_t)dt + \sigma S_t dZ_t, \\ dL_t &= \mu L_t dt + \xi L_t dW_t. \end{cases}$$

We want to show that for any $t > 0$, for any positive integer n and positive integer l,

$$\begin{aligned} &\mathbb{E}\left[S_t^n L_t^l\right] \\ = &\sum_{k=0}^{n}\sum_{i=0}^{k} \frac{(-1)^n n! \alpha^{n-k+i} S_0^{k-i} L_0^{n+l-k+i} e^{f(k,n+l-k)t}}{(k-i)! \prod_{j=k-i, j\neq k}^{n} (f(j, n+l-j) - f(k, n+l-k))}, \end{aligned} \tag{B.1}$$

where

$$f(n,l) = \frac{1}{2}n(n-1)\sigma^2 - n\alpha + \mu l + \frac{1}{2}l(l-1)\xi^2. \tag{B.2}$$

First, note that L_t follows a geometric Brownian motion, and therefore

$$L_t = L_0 e^{(\mu - \frac{\xi^2}{2})t + \xi W_t}. \tag{B.3}$$

Now as $W_t \sim N(0,t)$, and therefore for any integer $n > 0$, for any $t > 0$,

$$\begin{aligned} &\mathbb{E}\left[L_t^n\right] = L_0^n e^{n(\mu - \frac{\xi^2}{2})t} \mathbb{E}\left[e^{n\xi W_t}\right] \\ = &L_0^n e^{n(\mu - \frac{\xi^2}{2})t} \int_{-\infty}^{\infty} e^{n\xi x} e^{-\frac{x^2}{2t}} \frac{dx}{\sqrt{2\pi t}} \\ = &L_0^n e^{n(\mu - \frac{\xi^2}{2})t} \int_{-\infty}^{\infty} e^{-\frac{(x-n\xi t)^2}{2t} + \frac{n^2\xi^2 t}{2}} \frac{dx}{\sqrt{2\pi t}} \\ = &L_0^n e^{n(\mu - \frac{\xi^2}{2})t} e^{\frac{n^2\xi^2 t}{2}} \\ = &L_0^n e^{(n\mu + \frac{1}{2}n(n-1)\xi^2)t}. \end{aligned} \tag{B.4}$$

Furthermore, by Itô's lemma,

$$\begin{aligned} dS_t^n &= nS_t^{n-1}dS_t + \tfrac{1}{2}n(n-1)S_t^{n-2}(\sigma^2 S_t^2)dt \\ &= (n\alpha L_t S_t^{n-1} - n\alpha S_t^n + \tfrac{1}{2}n(n-1)\sigma^2 S_t^n)dt + n\sigma S_t^n dZ_t, \end{aligned} \tag{B.5}$$

$$\begin{aligned} dL_t^l &= lL_t^{l-1}dL_t + \tfrac{1}{2}l(l-1)L_t^{l-2}(\xi^2 L_t^2)dt \\ &= (l\mu L_t^l + \tfrac{1}{2}l(l-1)\xi^2 L_t^l)dt + l\xi L_t^l dW_t, \end{aligned} \tag{B.6}$$

and therefore by Itô's product rule,

$$\begin{aligned} & dS_t^n L_t^l \\ = \ & L_t^l dS_t^n + S_t^n dL_t^l + dS_t^n dL_t^l \\ = \ & \left(n\alpha S_t^{n-1}L_t^{l+1} - n\alpha S_t^n L_t^l + \tfrac{1}{2}n(n-1)\sigma^2 S_t^n L_t^l + l\mu S_t^n L_t^l + \tfrac{1}{2}l(l-1)\xi^2 S_t^n L_t^l\right) dt \\ & + n\sigma S_t^n L_t^l dZ_t + l\xi S_t^n L_t^l dW_t + nl\sigma\xi S_t^n L_t^l dZ_t dW_t. \end{aligned} \tag{B.7}$$

By taking expectation on equation (B.1), we can get

$$\begin{aligned} & d\mathbb{E}\left[S_t^n L_t^l\right] \\ = \ & \left(n\alpha\mathbb{E}\left[S_t^{n-1}L_t^{l+1}\right] - n\alpha\mathbb{E}\left[S_t^n L_t^l\right] + \tfrac{1}{2}n(n-1)\sigma^2\mathbb{E}\left[S_t^n L_t^l\right]\right. \\ & \left. + l\mu\mathbb{E}\left[S_t^n L_t^l\right] + \tfrac{1}{2}l(l-1)\xi^2\mathbb{E}\left[S_t^n L_t^l\right]\right) dt, \end{aligned} \tag{B.8}$$

or

$$\frac{d\mathbb{E}\left[S_t^n L_t^l\right]}{dt} = n\alpha\mathbb{E}\left[S_t^{n-1}L_t^{l+1}\right] + f(n,l)\mathbb{E}\left[S_t^n L_t^l\right], \tag{B.9}$$

where $f(n,l)$ was defined in equation (B.2).

Equation (B.9) is actually a linear first-order ordinary differential equation. If we define

$$h(t) = \mathbb{E}\left[S_t^{n-1}L_t^{l+1}\right],$$

the solution of equation (B.9) will be

$$\mathbb{E}\left[S_t^n L_t^l\right] = e^{f(n,l)t}\left(\int n\alpha e^{-f(n,l)t}h(t)dt + C\right) \tag{B.10}$$

To prove theorem 2.3-2, we can use the principle of mathematical induction.

First, let $n+l=M$, where M is a fixed positive integer. Now for any fixed M, we show that theorem 2.3-2 is correct for $n=0,1,\ldots,M$.

Now when $n=0$, by equation (B.1),

$$\mathbb{E}\left[L_t^l\right] = L_0^l e^{(l\mu+\frac{1}{2}l(l-1)\xi^2)t}. \tag{B.11}$$

From equation (B.1) and equation (B.2), we also have

$$\mathbb{E}\left[L_t^l\right] = L_0^l e^{f(0,l)t} = L_0^l e^{(l\mu+\frac{1}{2}l(l-1)\xi^2)t}. \tag{B.12}$$

When $n=1$, $l=M-1$. by equation (B.1) and equation (B.10),

$$\begin{aligned} & \mathbb{E}\left[S_t L_t^{M-1}\right] \\ = \;& e^{f(1,M-1)t}\left(\int \alpha e^{-f(1,M-1)t} L_0^M e^{(M\mu+\frac{1}{2}M(M-1)\xi^2)t} dt + C\right) \\ = \;& e^{f(1,M-1)t}\left(\int \alpha L_0^M e^{-f(1,M-1)t+f(0,M)t} dt + C\right) \\ = \;& \tfrac{\alpha L_0^M e^{f(0,M)t}}{f(0,M)-f(1,M-1)} + C e^{f(1,M-1)t}. \end{aligned} \tag{B.13}$$

As $\mathbb{E}\left[S_t L_t^{M-1}\right] = S_0 L_0^{M-1}$ when $t=0$, we can get

$$C = S_0 L_0^{M-1} - \frac{\alpha L_0^M}{f(0,M)-f(1,M-1)}, \tag{B.14}$$

and therefore

$$\begin{aligned} & \mathbb{E}\left[S_t L_t^{M-1}\right] \\ = \;& S_0 L_0^{M-1} e^{f(1,M-1),t} + \tfrac{\alpha L_0^M e^{f(0,M)t}}{f(0,M)-f(1,M-1)} - \tfrac{\alpha L_0^M e^{f(1,M-1)t}}{f(0,M)-f(1,M-1)}, \end{aligned} \tag{B.15}$$

which is the same as the form of equation (B.1) when $n=1$, $l=M-1$.

Now suppose that when $n=N<M$, equation (B.1) is satisfied. Consider the case when $n=N+1$, $l=M-N-1$, as the expression of $\mathbb{E}\left[S_t^{n-1} L_t^{l+1}\right]$ is already known, we have

$$
\begin{aligned}
&\mathbb{E}\left[S_t^n L_t^l\right] \\
= \;& e^{f(n,l)t}\Big(\int n\alpha e^{-f(n,l)t}\mathbb{E}\left[S_t^{n-1}L_t^{l+1}\right]dt + C\Big) \\
= \;& e^{f(n,l)t}\int \sum_{k=0}^{n-1}\sum_{i=0}^{k}\frac{e^{-f(n,l)t}(-1)^{n-1}n!\alpha^{n-k+i}S_0^{k-i}L_0^{n+l-k+i}e^{f(k,n+l-k)t}}{(k-i)!\prod_{j=k-i,j\neq k}^{n-1}(f(j,n+l-j)-f(k,n+l-k))}dt \\
& + Ce^{f(n,l)t} \\
= \;& e^{f(n,l)t}\sum_{k=0}^{n-1}\sum_{i=0}^{k}\frac{(-1)^{n-1}n!\alpha^{n-k+i}S_0^{k-i}L_0^{n+l-k+i}\int e^{f(k,n+l-k)t-f(n,l)t}dt}{(k-i)!\prod_{j=k-i,j\neq k}^{n-1}(f(j,n+l-j)-f(k,n+l-k))} \\
& + Ce^{f(n,l)t} \\
= \;& \sum_{k=0}^{n-1}\sum_{i=0}^{k}\frac{(-1)^{n}n!\alpha^{n-k+i}S_0^{k-i}L_0^{n+l-k+i}e^{f(k,n+l-k)t}}{(k-i)!\prod_{j=k-i,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))} \\
& + Ce^{f(n,l)t}.
\end{aligned}
$$

Since $\mathbb{E}\left[S_t^n L_t^l\right] = S_0^n L_0^l$ when $t = 0$, we can get

$$
C = S_0^n L_0^l - \sum_{k=0}^{n-1}\sum_{i=0}^{k}\frac{(-1)^{n}n!\alpha^{n-k+i}S_0^{k-i}L_0^{n+l-k+i}}{(k-i)!\prod_{j=k-i,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))}.
$$

Therefore,

$$
\begin{aligned}
&\mathbb{E}\left[S_t^n L_t^l\right] \\
= \;& \sum_{k=0}^{n-1}\sum_{i=0}^{k}\frac{(-1)^{n}n!\alpha^{n-k+i}S_0^{k-i}L_0^{n+l-k+i}e^{f(k,n+l-k)t}}{(k-i)!\prod_{j=k-i,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))}dt \\
& + S_0^n L_0^l e^{f(n,l)t} - \sum_{k=0}^{n-1}\sum_{i=0}^{k}\frac{(-1)^{n}n!\alpha^{n-k+i}S_0^{k-i}L_0^{n+l-k+i}e^{f(n,l)t}}{(k-i)!\prod_{j=k-i,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))}.
\end{aligned}
$$

We need

$$
\begin{aligned}
&\mathbb{E}\left[S_t^n L_t^l\right] \\
= \;& \sum_{k=0}^{n}\sum_{i=0}^{k}\frac{(-1)^{n}n!\alpha^{n-k+i}S_0^{k-i}L_0^{n+l-k+i}e^{f(k,n+l-k)t}}{(k-i)!\prod_{j=k-i,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))},
\end{aligned}
$$

i.e., we need to prove that

$$
S_0^n L_0^l e^{f(n,l)t} - \sum_{k=0}^{n-1}\sum_{i=0}^{k}\frac{(-1)^{n}n!\alpha^{n-k+i}S_0^{k-i}L_0^{n+l-k+i}e^{f(n,l)t}}{(k-i)!\prod_{j=k-i,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))}
$$

$$= \sum_{i=0}^{n} \frac{(-1)^n n! \alpha^i S_0^{n-i} L_0^{l+i} e^{f(n,l)t}}{(n-i)! \prod_{j=n-i,j\neq n}^{n} (f(j,n+l-j) - f(n,l))}. \tag{B.16}$$

Note that when $i = 0$,

$$S_0^n L_0^l e^{f(n,l)t} = \frac{(-1)^n n! \alpha^i S_0^{n-i} L_0^{l+i} e^{f(n,l)t}}{(n-i)! \prod_{j=n-i,j\neq n}^{n} (f(j,n+l-j) - f(n,l))}, \tag{B.17}$$

and

$$\begin{aligned} & \sum_{k=0}^{n-1} \sum_{i=0}^{k} \frac{(-1)^n n! \alpha^{n-k+i} S_0^{k-i} L_0^{n+l-k+i} e^{f(n,l)t}}{(k-i)! \prod_{j=k-i,j\neq k}^{n} (f(j,n+l-j) - f(k,n+l-k))} \\ = & \sum_{k=0}^{n-1} \sum_{i=n-k}^{n} \frac{(-1)^n n! \alpha^i S_0^{n-i} L_0^{l+i} e^{f(n,l)t}}{(n-i)! \prod_{j=n-i,j\neq k}^{n} (f(j,n+l-j) - f(k,n+l-k))} \\ = & \sum_{i=1}^{n} \sum_{k=n-i}^{n-1} \frac{(-1)^n n! \alpha^i S_0^{n-i} L_0^{l+i} e^{f(n,l)t}}{(n-i)! \prod_{j=n-i,j\neq k}^{n} (f(j,n+l-j) - f(k,n+l-k))}, \end{aligned} \tag{B.18}$$

By equation (B.17) and (B.18), the equation (B.16) can be simplified as

$$\begin{aligned} & -\sum_{i=1}^{n} \sum_{k=n-i}^{n-1} \frac{(-1)^n n! \alpha^i S_0^{n-i} L_0^{l+i} e^{f(n,l)t}}{(n-i)! \prod_{j=n-i,j\neq k}^{n} (f(j,n+l-j) - f(k,n+l-k))} \\ = & \sum_{i=1}^{n} \frac{(-1)^n n! \alpha^i S_0^{n-i} L_0^{l+i} e^{f(n,l)t}}{(n-i)! \prod_{j=n-i,j\neq n}^{n} (f(j,n+l-j) - f(n,l))} \\ = & \sum_{i=1}^{n} \frac{(-1)^n n! \alpha^i S_0^{n-i} L_0^{l+i} e^{f(n,l)t}}{(n-i)! \prod_{j=n-i}^{n-1} (f(j,n+l-j) - f(n,l))} \end{aligned}$$

or

$$\begin{aligned} & -\sum_{i=1}^{n} \sum_{k=n-i}^{n-1} \frac{1}{\prod_{j=n-i,j\neq k}^{n} (f(j,n+l-j) - f(k,n+l-k))} \\ = & \sum_{i=1}^{n} \frac{1}{\prod_{j=n-i}^{n-1} (f(j,n+l-j) - f(n,l))}. \end{aligned} \tag{B.19}$$

If equation (B.19) is correct, then by the principle of mathematical induction, equation (B.1) is correct, and thus theorem 2.3-2 is proved.

We can again use the principle of mathematical induction to prove that for any fixed n, for any $i = 1, 2, \ldots, n$,

$$-\sum_{k=n-i}^{n-1} \frac{1}{\prod_{j=n-i,j\neq k}^{n} (f(j,n+l-j) - f(k,n+l-k))}$$

$$= \frac{1}{\prod_{j=n-i}^{n-1}(f(j,n+l-j)-f(n,l))}. \tag{B.20}$$

First, when $i=1$, the left hand side of equation (B.20) will be

$$\begin{aligned} & -\sum_{k=n-1}^{n-1} \frac{1}{\prod_{j=n-1,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))} \\ = & -\frac{1}{f(n,l)-f(n-1,l+1)}, \end{aligned}$$

and the right hand side will be

$$\begin{aligned} & \frac{1}{\prod_{j=n-1}^{n-1}(f(j,n+l-j)-f(n,l))} \\ = & \frac{1}{f(n-1,l+1)-f(n,l)}, \end{aligned}$$

therefore, equation (B.20) is correct when $i=1$.

Now when $i=2$, we have

$$\begin{aligned} & -\sum_{k=n-2}^{n-1} \frac{1}{\prod_{j=n-2,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))} \\ = & -\frac{1}{(f(n-1,l+1)-f(n-2,l+2))(f(n,l)-f(n-2,l+2))} \\ & -\frac{1}{(f(n-2,l+2)-f(n-1,l+1))(f(n,l)-f(n-1,l+1))} \\ = & \frac{f(n,l)-f(n-1,l+1)-f(n,l)+f(n-2,l+2)}{(f(n-2,l+2)-f(n-1,l+1))(f(n,l)-f(n-2,l+2))(f(n,l)-f(n-1,l+1))} \\ = & \frac{1}{(f(n,l)-f(n-2,l+2))(f(n,l)-f(n-1,l+1))} \\ = & \frac{1}{\prod_{j=n-2}^{n-1}(f(j,n+l-j)-f(n,l))}, \end{aligned}$$

so equation (B.20) is satisfied.

From above we can see that when $n=1$ or $n=2$, equation (B.20) is correct. Assume that it is correct until $n-1$, we can prove the case of n.

For the fixed n, first notice that the cases when $i=1$ or $i=2$ are proved above. Now assume that when $i=h<n$,

$$-\sum_{k=n-h}^{n-1} \frac{1}{\prod_{j=n-h,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))}$$

$$= \frac{1}{\prod_{j=n-h}^{n-1}(f(j,n+l-j)-f(n,l))}. \tag{B.21}$$

Then for the case when $i = h+1$, for convenience in expression, we define

$$g(j) = f(j, n+l-j)$$

for any integer j, and

$$\begin{aligned}
& -\sum_{k=n-h-1}^{n-1} \frac{1}{\prod_{j=n-h-1,j\neq k}^{n}(f(j,n+l-j)-f(k,n+l-k))} \\
& -\frac{1}{\prod_{j=n-h-1}^{n-1}(f(j,n+l-j)-f(n,l))} \\
= & -\sum_{k=n-h-1}^{n-1} \frac{1}{\prod_{j=n-h-1,j\neq k}^{n}(g(j)-g(k))} - \frac{1}{\prod_{j=n-h-1}^{n-1}(g(j)-g(n))} \\
= & -\sum_{k=n-h}^{n-1} \frac{1}{\prod_{j=n-h-1,j\neq k}^{n}(g(j)-g(k))} - \frac{1}{\prod_{j=n-h}^{n}(g(j)-g(n-h-1))} \\
& -\frac{1}{\prod_{j=n-h-1}^{n-1}(g(j)-g(n))} \\
= & -\sum_{k=n-h}^{n-1} \frac{1}{\left(\prod_{j=n-h,j\neq k}^{n}(g(j)-g(k))\right)(g(n-h-1)-g(k))} \\
& -\frac{1}{\prod_{j=n-h}^{n}(g(j)-g(n-h-1))} \\
& +\frac{1}{(g(n-h-1)-g(n))} \sum_{k=n-h}^{n-1} \frac{1}{\prod_{j=n-h,j\neq k}^{n}(g(j)-g(k))} \\
= & \sum_{k=n-h}^{n-1} \frac{1}{\prod_{j=n-h,j\neq k}^{n}(g(j)-g(k))} \left(\frac{1}{g(n-h-1)-g(n)} - \frac{1}{g(n-h-1)-g(k)}\right) \\
& -\frac{1}{\prod_{j=n-h}^{n}(g(j)-g(n-h-1))} \\
= & \sum_{k=n-h}^{n-1} \frac{1}{\prod_{j=n-h,j\neq k}^{n}(g(j)-g(k))} \left(\frac{g(n)-g(k)}{(g(n-h-1)-g(n))(g(n-h-1)-g(k))}\right) \\
& -\frac{1}{\prod_{j=n-h}^{n}(g(j)-g(n-h-1))} \\
= & \sum_{k=n-h}^{n-1} \frac{1}{\prod_{j=n-h-1,j\neq k}^{n-1}(g(j)-g(k))} \left(\frac{1}{g(n-h-1)-g(n)}\right)
\end{aligned}$$

$$
\begin{aligned}
& - \frac{1}{\prod_{j=n-h}^{n}(g(j) - g(n-h-1))} \\
= & \frac{1}{\prod_{j=n-h}^{n}(g(j) - g(n-h-1))} - \frac{1}{\prod_{j=n-h}^{n}(g(j) - g(n-h-1))} \\
= & 0.
\end{aligned}
$$

Therefore equation (B.20) is correct in the case $i = h + 1$, and thus it is correct for case of $n+1$. From all above, by the principle of mathematical induction, equation (B.20) is correct, and therefore theorem 2.3-2 is proved.

B.2 Proof of Theorem 2.4-1

For model 1FJ, S_t satisfies

$$dS_t = \alpha(L - S_t)dt + \sigma S_t dZ_t + U_t S_t dP_t.$$

The equation (2.15) is derived by multiplying jump amplitudes at the jump times. The equation (2.16) gives the expression of $\mathbb{E}[S_t]$, which follows

$$\mathbb{E}[S_t] = e^{-\alpha t + \eta t \mathbb{E}[U_t]}(S_0 - L) + Le^{-\alpha t}\mathbb{E}[U_t] J_t + L, \tag{B.22}$$

where

$$J_t = \sum_{n=0}^{\infty} \frac{e^{-\eta t}(\eta t)^n}{n!} \mathbb{E}\left[\sum_{i=1}^{n} (1 + \mathbb{E}[U_t])^{n-i} e^{\alpha \tau_i} | P_t = n \right],$$

and τ_i are the jump times.

To prove equation (B.22), first we show that

$$
\begin{aligned}
& \mathbb{E}[S_t | P_t = n, 0 \leq \tau_1 < \tau_2 < \ldots < \tau_n < t] \\
= \; & e^{-\alpha t}(S_0 - L)(1 + \mathbb{E}[U_t])^n + Le^{-\alpha t}\mathbb{E}[U_t] \sum_{i=1}^{n} (1 + \mathbb{E}[U_t])^{n-i} e^{\alpha \tau_i} + L. \quad \text{(B.23)}
\end{aligned}
$$

Equation (B.23) can be proved by principle of mathematical induction.

First, when $n = 0$,

$$\mathbb{E}\left[S_t | P_t = 0\right] = e^{-\alpha t}(S_0 - L) + L,$$

so equation (B.23) is correct.

When $n = 1$,

$$\begin{aligned}
& \mathbb{E}\left[S_t | P_t = 1, 0 \leq \tau_1 < t\right] \\
= \ & \mathbb{E}\left[\mathbb{E}\left[S_t | S_{\tau_1}\right] | P_t = 1, 0 \leq \tau_1 < t\right] \\
= \ & \mathbb{E}\left[e^{-\alpha(t-\tau_1)}(S_{\tau_1} - L) + L | P_t = 1, 0 \leq \tau_1 < t\right] \\
= \ & \mathbb{E}\left[e^{-\alpha(t-\tau_1)}((1 + U_t)S_{\tau_1^-} - L) + L | P_t = 1, 0 \leq \tau_1 < t\right] \\
= \ & e^{-\alpha(t-\tau_1)}(\mathbb{E}\left[1 + U_t\right] \mathbb{E}\left[S_{\tau_1^-} | P_t = 1, 0 \leq \tau_1 < t\right] - L) + L \\
= \ & e^{-\alpha(t-\tau_1)}\left(\mathbb{E}\left[1 + U_t\right] e^{-\alpha\tau_1}(S_0 - L) + \mathbb{E}\left[1 + U_t\right] L - L\right) + L \\
= \ & e^{-\alpha t}\mathbb{E}\left[1 + U_t\right](S_0 - L) + e^{-\alpha(t-\tau_1)}\mathbb{E}\left[U_t\right] L + L.
\end{aligned}$$

Now suppose that equation (B.23) is right when $n = m - 1$. Then for the case $n = m$,

$$\begin{aligned}
& \mathbb{E}\left[S_t | P_t = m, 0 \leq \tau_1 < \tau_2 < \ldots < \tau_m < t\right] \\
= \ & \mathbb{E}\left[\mathbb{E}\left[S_t | S_{\tau_1}, P_t = m - 1, \tau_1 \leq \tau_2 < \tau_2 < \ldots < \tau_m < t\right] | \tau_1\right] + L \\
= \ & \mathbb{E}\left[e^{-\alpha(t-\tau_1)}(S_{\tau_1} - L)(1 + \mathbb{E}\left[U_t\right])^{m-1} | \tau_1\right] \\
& + Le^{-\alpha t}\mathbb{E}\left[U_t\right] \sum_{i=1}^{m-1}(1 + \mathbb{E}\left[U_t\right])^{m-i-1} e^{\alpha\tau_{i+1}} + L \\
= \ & e^{-\alpha(t-\tau_1)}(\mathbb{E}\left[1 + U_t\right] \mathbb{E}\left[S_{\tau_1^-} | \tau_1\right] - L)(1 + \mathbb{E}\left[U_t\right])^{m-1} \\
& + Le^{-\alpha t}\mathbb{E}\left[U_t\right] \sum_{i=1}^{m-1}(1 + \mathbb{E}\left[U_t\right])^{m-i-1} e^{\alpha\tau_{i+1}} + L \\
= \ & e^{-\alpha t}(S_0 - L)(1 + \mathbb{E}\left[U_t\right])^m + Le^{-\alpha(t-\tau_1)}\mathbb{E}\left[U_t\right](1 + \mathbb{E}\left[U_t\right])^{m-1} \\
& + Le^{-\alpha t}\mathbb{E}\left[U_t\right] \sum_{i=1}^{m-1}(1 + \mathbb{E}\left[U_t\right])^{m-i-1} e^{\alpha\tau_{i+1}} + L \\
= \ & e^{-\alpha t}(S_0 - L)(1 + \mathbb{E}\left[U_t\right])^m + Le^{-\alpha t}\mathbb{E}\left[U_t\right] \sum_{i=1}^{m}(1 + \mathbb{E}\left[U_t\right])^{m-i} e^{\alpha\tau_i} + L.
\end{aligned}$$

From all above, equation (B.23) is proved.

Finally,

$$
\begin{aligned}
&\mathbb{E}\left[S_t\right] \\
= &\sum_{n=0}^{\infty} \frac{e^{-\eta t}(\eta t)^n}{n!} \mathbb{E}\left[S_t | P_t = n, 0 \leq \tau_1 < \tau_2 < \ldots < \tau_n < t\right] \\
= &\sum_{n=0}^{\infty} \frac{e^{-\eta t}(\eta t)^n}{n!} e^{-\alpha t}(S_0 - L)(1 + \mathbb{E}\left[U_t\right])^n \\
&+ \sum_{n=0}^{\infty} \frac{e^{-\eta t}(\eta t)^n}{n!} L e^{-\alpha t} \mathbb{E}\left[U_t\right] \sum_{i=1}^{n} (1 + \mathbb{E}\left[U_t\right])^{n-i} e^{\alpha \tau_i} + L \\
= &e^{-\alpha t + \eta t \mathbb{E}[U_t]}(S_0 - L) + L e^{-\alpha t} \mathbb{E}\left[U_t\right] J_t + L,
\end{aligned}
$$

where

$$
J_t = \sum_{n=0}^{\infty} \frac{e^{-\eta t}(\eta t)^n}{n!} \mathbb{E}\left[\sum_{i=1}^{n} (1 + \mathbb{E}\left[U_t\right])^{n-i} e^{\alpha \tau_i} | P_t = n\right].
$$

Thus theorem 2.4-1 is proved.

Bibliography

[1] K. Aase, 1988, *Contingent Claims Valuation When the Security Price is a Combination of an Ito Process and a Random Point Process*, Stochastic Processes and their Applications 28, 185-220

[2] E. Alison, 2002, *A course in financial calculus*, Cambridge University Press

[3] K. I. Amin, 1993, *Jump Diffusion Option Valuation in Discrete Time*, The Journal of Finance, Vol. 48, No. 5(1993), pp. 1833-1863

[4] M. Barlow, Y. Gusev, M. Lai, 2004, *Calibration of Multifactor Models in Electricity Markets*, International Journal of Theorectical and Applied Finance, Vol. 7, No. 2(2004), 101-120

[5] P. Billingsley, 1995, *Probability and Measure*, 3rd Edition, Wiley-Interscience

[6] L. P. Bos, A. F. Ware, B. S. Pavlov, 2002, *On a semi-spectral method for pricing an option on a mean-reverting asset*, Quantitative Finance, Vol. 2(2002), 337-345

[7] P. Bratley, B. L. Fox, L. E. Schrage, 1983, *A guide to simulation*, Springer-Verlag New York Inc.

[8] P. Carr, M. Schröder, 2004, *Bessel processes, the integral of geometric Brownian motion, and Asian options*, Theory of Probabilty & Applications, Vol. 48(2004), No. 3, 400-425

[9] L. Clewlow, C. Strickland, 1998, *Implementing Derivatives Models*, John Wiley & Sons

[10] L. Clewlow, C. Strickland, 2000, *Energy derivatives: Pricing and risk management*, Lacima Publications

[11] M. Drmota, R. F. Tichy, 1997, *Sequences, discrepancies and applications*, Springer, Berlin

[12] R. Elliott, A. Cadenillas, 2000, *Pricing swing options*

[13] F. Garvan, 2001, *The Maple Book*, Chapman & Hall/CRC

[14] H. Geman, 2005, *Commodities and commodity derivatives: modelling and pricing for agriculturals, metals, and energy*, John Wiley & Sons

[15] J. C. Hull, 2005, *Options, futures, and Other Derivatives*, Prentice Hall

[16] K. Itô, 1951, *On stochastic differential equations*, Memoirs, American Mathematical Society, Vol 4, pp. 1-51.

[17] P. Jaillet, E. I. Ronn, S. Tompaidis, *Valuation of commodity-based swing options*, 2004, Management Science, Vol. 50, Issue 7 (July 2004), pp 909-921

[18] R. Jarrow, S. Turnbull, 1996, *Derivative Securities*, 1st Edition, International Thomson Publishing, pp. 213

[19] S. J. Julier, J. K. Uhlmann, 1997, *A New Extension of the Kalman Filter to nonlinear Systems.* In The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing,Simulation and Controls, Multi Sensor Fusion, Tracking and Resource Management II, SPIE.

[20] O. Kallenberg, 2001, *Foundations of Modern Probability*, 2nd Edition, Springer

[21] P. E. Kloaden, E. Platen, 1999. *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag

[22] A. Lari-Lavassani, A. A. Sadeghi, A. Ware, 2001, *Mean Reverting Models For Energy Option Pricing*, Electronic Publications of the International Energy Credit Association (www.ieca.net).

[23] A. Lari-lavassani, M. Simchi, A. Ware, 2001, *A Discrete Valuation of Swing Options*, Canadian Applied Mathematics Quarterly, Vol. 9, No. 1, pp. 35-73

[24] C. Lemieux, 2004, *Randomized Quasi-Monte Carlo: A Tool for Improving the Efficiency of Simulations in Finance*, Proceedings of the 2004 Winter Simulation Conference

[25] F. A. Longstaff, E. S. Schwartz, 2001, *Valuing American Options by Simulation: A Simple Least-Squares Approach*, The Review of Financial Studies, Spring 2001, Vol. 14, No. 1, pp. 113-147

[26] J. T. McClave, T. Sincich, 2006, *Statistics*, 10th Edition, Pearson Prentice Hall

[27] N. Meinshausen, B. M. Hambly, 2004, *Monte Carlo Methods for the Valuation of Multiple-Exercise Options*, Mathematical Finance, Vol. 14, No. 4(2004), pp. 557-583

[28] R. C. Merton, 1992, *Continuous-Time Finance*, Blackwell Publishers Inc.

[29] S. N. Neftci, 1996, *An Introduction to the Mathematics of Financial Derivatives*, Academic Press

[30] B. Øksendal, 2003, *Stochastic Differential Equations: An Introduction with Applications*, Springer

[31] G. Peskir, A. Shiryaev, 2006, *Optimal stopping and free-boundary problems*, Birkhäuser Verlag.

[32] D. Pilipovic, 1998, *Energy Risk: Valuing and Managing Energy Derivatives*, McGraw-Hill

[33] W. Rudin, 1986, *Real and Complex Analysis*, New York: McGraw-Hill, pp. 116-134.

[34] E. S. Schwartz, *The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging*, The Journal Of Finance, Vol. LII, No. 3(1997), pp. 922-973

[35] A. N. Shiryaev, 1995, *Probability*, Springer

[36] H. Stark, J. W. Woods, 2002, *Probability and Random Processes with Applications to Signal Processing*, 3rd Edition, Prentice Hall, New Jersey

[37] J. M. Steele, 2000, *Stochastic Calculus and Financial Applications*, Springer

[38] A. Swishchuk, 2006, preprint, *Explicit Option Pricing Formula for a Mean-Reverting Asset*, E-Yellow Series, No. 850, Department of Mathematics and Statistics, University of Calgary

[39] O. Vasicek, 1977, *An Equilibrium Characterisation of the Term Structure.* Journal of Financial Economics, Vol. 5, pp. 177-188.

[40] Z. Xu, 2004, *Stochastic models for gas prices*, M.Sc. thesis, University of Calgary

[41] M. Yor, 1992, *On some exponential functions of Brownian motion*, Advances in Applied Probability, Vol. 24, No. 3(1992), 509-531

[42] M. Yor, H. Matsumoto, 2005, *Exponential functionals of Brownian motion, I: Probability laws at fixed time*, Probability Surveys, Vol 2(2005), 312-347