University of Cape Town

School of Management Studies

(Actuarial Science)

Date: October 28, 2017

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY

Michael Kateregga

ENTITLED

Stable Processes: Theory and Applications in Finance

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN MATHEMATICAL FINANCE

Thesis Advisor
Dr. Sure Mataramvura

Thesis Co-Advisor
Prof. David Taylor

Head of Department
Prof. Anton Schlechter

Stable Processes: Theory and Applications in Finance

By

Michael Kateregga

Dissertation

Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Mathematical Finance
in the Actuarial Science Department at the
University of Cape Town, 2016

Cape Town, South Africa

“The man who begins to speculate in stocks with the intention of making a fortune usually goes broke, whereas the man who trades with a view of getting good interest on his money sometimes gets rich.” Charles Henry Dow

“Stocks are bought on expectations, not facts.” Gerald Loeb

“One of the funny things about the stock market is that every time one person buys, another sells, and both think they are astute”. William A. Feather

Acknowledgments

I thank God for the blessings and wisdom that enabled me to pursue a PhD. He deserves all the glory and honor for my strength, perseverance and grace throughout the entire process of writing this thesis.

I offer my sincerest gratitude to my supervisor, Dr. Sure Mataramvura for his support and guidance whilst allowing me to work independently. I attribute the level of my PhD thesis to his encouragement, advise and support right from problem-identification and proposal stages to final write-up and without him this thesis would not have been written. In addition, I extend my gratitude to A. Prof. David Taylor for his advise, financial support and allowing me to register as a PhD student in the department. I thank Dr. Farai Mhlanga, Prof. Thaddeus Tarpey and Prof. Martin Keller-Ressel for sparing time out of their busy schedules to read and correct some parts of this thesis.

I also offer gratitude to the entire ACQuFRR research group for allowing me to be part of the group. Special thanks to the Faculty of Commerce and in particular Ms. Joanna Lugutko for always being available to encourage and approve my applications. I also thank all administrative staff for their assistance and guidance.

I would also like to offer my deepest gratitude to all my funders: the National Research Foundation of South Africa (NRF) through the postgraduate funding office, the African Institute for Mathematical Sciences (AIMS), the school of management studies and the African Collaboration for Quantitative Finance and Risk Research (ACQuFRR).

Finally, my dearest and sincere gratitude goes to my wife Estelle Kateregga for her understanding and support during the late and endless hours of hard work every day.

Stable Processes: Theory and Applications in Finance

Michael Kateregga

School of Management Studies
Department of Actuarial Science
University of Cape Town
Cape Town, South Africa
2017

ABSTRACT

This thesis is a study on stable distributions and some of their applications in understanding financial markets. Three broad problems are explored: First, we study a parameter and density estimation problem for stable distributions using commodity market data. We investigate and compare the accuracy of the quantile, logarithmic, maximum likelihood (ML) and empirical characteristic function (ECF) methods. It turns out that the ECF is the most recommendable method, challenging literature that instead suggests the ML. Secondly, we develop an affine theory for subordinated random processes and apply the results to pricing commodity futures in markets where the spot price includes jumps. The jumps are introduced by subordinating Brownian motion in the spot model by an α-stable process, $\alpha \in (0, 1]$ which leads to a new pricing approach for models with latent variables. The third problem is the pricing of general derivatives and risk management based on Malliavin calculus. We derive a Bismut-Elworthy-Li (BEL) representation formula for computing financial Greeks under the framework of subordinated Brownian motion by an inverse α-stable process with $\alpha \in (0, 1]$. This subordination by an inverse α-stable process allows zero returns in the model rendering it fit for illiquid emerging markets. In addition, we demonstrate that the model is best suited for pricing derivatives with irregular payoff functions compared to the traditional Euler methods.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Scope

The financial market has a tendency of deviating from normality as analyzed from historical data. Distributions of empirical asset price returns from commodities, equity, forex, etc. are skewed, exhibit fat tails and they either have high or low peaks compared to normal distribution. Literature suggests ways to capture these features including models with stochastic volatility and/or jumps. The jumps are introduced by a Poisson process or a stable L'evy process. This thesis focuses on the latter to fit stable distributions to market data, model commodity futures, derive derivatives-on-equity prices and compute the Greeks essential for risk management in finance.

We explore the theory underpinning the rich and robust class of stable distributions and their application in understanding financial markets. They exhibit features and properties necessary for capturing market behavior. The thesis tackles three main problems: First, we study a parameter and density estimation problem for stable distributions using commodity market data. We investigate and compare the accuracy of the quantile, logarithmic, maximum likelihood and empirical characteristic function methods. The second problem aims at developing an affine theory for subordinated random processes and applying the results to pricing commodity futures in markets where the spot price includes jumps. The jumps are introduced by subordinating the Brownian motion in the spot model by an α-stable process where $\alpha \in (0, 1]$ to ensure only positive jumps from the subordinator. The meaning of α shall become apparent in the subsequent chapters. Both a one and two factor futures pricing models are developed but we shall put more emphasis on the one factor model to motivate our approach. The difference from existing models is that the price is given as a function of the subordinator as opposed to the underlying spot value. Since the subordinator is not observable, we reduce our model to a latent regression model and apply the techniques usually employed in latent regression models for population dynamics to estimate the model parameters. Seasonality is modeled using a sinusoidal function which is fitted to historical commodity spot prices. The third problem is a risk management problem and employs knowledge from Malliavin calculus to derive a Bismut-Elworthy-Li (BEL) representation formula for computing financial Greeks under the framework of subordinated Brownian motion by an inverse α-stable process. This subordination by an inverse α-stable process allows zero returns in the model rendering it fit for illiquid emerging markets. We discuss these problems in more details and point out the contribution of our research to the broad body of knowledge, in the following section.

1.2 Problems and Contribution

This thesis is aimed at solving three broad problems related to estimation, pricing and hedging in markets with jumps where Gaussianity is not an assumption.

1.2.1 Estimation

The first problem is handled in Chapters 2 and 3, where we explore the theory behind the rich and robust family of α-stable distributions for modeling skewed empirical data. We discuss four parameter estimation techniques for extracting α-stable distribution parameters from data namely: quantiles, logarithmic moments, maximum likelihood and the empirical characteristics function (ECF) method. The contribution of the chapter is two-fold: first, we discuss the four parameter estimation techniques and investigate their performance through error analysis. We show that the empirical characteristic function method is an excellent technique for estimating stable distribution parameters from skewed empirical data more accurately as opposed to the maximum likelihood technique commonly employed in literature. It provides the best precision compared to the other three methods and it can therefore be used to obtain initial input parameters for future and better estimation techniques. The ECF is then applied to skewed empirical commodity data to determine the shape parameter of the data. Secondly, we compare the skewed empirical commodity data to various known distributions including normal distribution to determine their closest distributions.

1.2.2 Pricing

The second problem is handled in Chapter 4 where we develop the theory governing our subordinated affine-structure models for the commodity spot price and obtain explicit price representation for the future price in markets where the spot price includes jumps. The jumps are introduced by subordinating the Brownian motion in the spot model by an α-stable process where $\alpha \in (0, 1]$. The range of α is such as to ensure positive jumps only. Moreover, the stable process is non-decreasing and cádlág. The contribution of this chapter is an extension of the affine-structure models discussed in [KR08, KRST11], [RH15] and [DFS03] to include subordinated processes to provide an alternative, robust and tractable approach of capturing skewness, kurtosis and fat/skinny tails in commodity spot models as opposed to the traditional pure jump Lévy process. The difference between our model and the existing models is that the price in the former is given as a function of the subordinator as opposed to the underlying spot value. In other words, the future price is a function of the business time in addition to the calendar time. Seasonality is modeled using a sinusoidal function which is fitted to historical commodity spot prices.

1.2.3 Sensitivity Analysis

The third problem is handled in Chapter 5. The key objective of the chapter is to extend the results in [FLL+99, CF07] for continuous processes to jump processes based on the Bismut-Elworthy-Li (BEL) formula in [EL94]. We construct a jump process using a subordinated Brownian motion where the subordinator is an inverse α-stable process $(L_t)_{t\geq 0}$ with $\alpha \in (0, 1]$. The results are derived using Malliavin integration by parts formula. We derive representation formulas for computing financial Greeks and show that in the event when $L_t \equiv t$, we retrieve the results in [FLL+99]. The purpose is to by-pass the derivative of an (irregular) payoff function in a jump-type market by introducing a weight term in form of an integral with respect to a subordinated Brownian motion. Using Monte Carlo techniques, we estimate financial Greeks for a digital option and show that the BEL formula still performs better for a discontinuous payoff in a jump asset model setting and that the finite difference methods are better for continuous payoffs in a similar setting. Finally, the contribution of the Chapter will also include deriving much simpler techniques based on the basic Malliavin integration by parts formula to arrive at similar results in related existing literature (such as [Zha12] for instance).

1.3 Literature

The family of stable distributions first introduced by Paul Lévy in his book *Calcul des Probabilitiés*, [Lév25] has been vastly studied and applied in various disciplines in relation to understanding random physical phenomena that do not necessarily follow the normal law. Since Paul Lévy's breakthrough new developments towards the so called Lévy process have been documented including the popular texts of [ST94], [Ber98], [Sat99], [App04] and [BNMR12]. Due to their success, stable distributions provided an active area of research for probability theorists and continue to find applications in the fields of physics, astronomy, economics, communication theory, engineering, statistics, finance etc. The main reason why stable distributions gained popularity is because they generalize the central limit theorem and act as limiting distributions to a wide range of other distributions. In Chapter 2 we shall explore in much more detail, the theory behind the rich and robust family of α-stable distributions.

1.3.1 Literature on Parameter Estimation

In this section we discuss the background literature preceeding Chapter 3. The application of stable distributions in finance is traced way back in the late 50's when [Man59, Man62, Man63] developed a hypothesis that revolutionalized the way economists viewed and interpreted prices in speculative markets such as grains and securities markets. The hypothesis suggested that prices were not Gaussian as it had been previously believed by market participants based on [Bac00]. Mandelbrot's hypothesis was therefore an extension of the widely embraced breakthrough of [Bac00].

In the following years [Zol64] developed integral representations of stable laws and the results have been used to develop their parameter estimation techniques. [Fam63] reviewed the vality of Mandelbrot's hypothesis and came up with statistical tools suitable for dealing with speculative prices. [Dum71] employs this class of distributions in statistical inference for long-tailed data. Graphical representation of their densities and the estimation of their parameters via interpolation appear in [HC73] and in [Kou80] using regression. Parameter estimation methods based on quantile methods are presented in [FR71] for symmetric stable distributions but this approach faces a problem of discontinuity of the traditional location parameter in the asymmetrical cases when the exponent parameter passes unity. A remedy and generalisation of the quantile approach is later introduced by [McC86].

A different parameter estimation technique based on fractional lower order moments (FLOM) appears in [MN95] where the authors develop new methods for estimating parameters in impulsive signal enviroments. However, their methods only cover symmetric stable distributions. There was need to extend the method to asymmetric systems. This came through by [Kur01] where a generalised FLOM method is introduced. Generally, FLOM methods pose a challenge of having to estimate the Sinc function and this in turn affects the accuracy of the results. As a consequence, a better estimation approach referred to as logarithmic moments method (LM) is proposed by [Kur01] to avoid having to compute the *Sinc*.

The third estimation method utilizes the maximum likelihood. It is known that the maximum likelihood (ML) approach is widely favoured in economic and financial applications due to its generality and asymptotic efficiency (see for instance [Yu04]). However, there are cases where the ML method is unreliable especially when the likelihood function is not tractable, or its not bounded over the parameter space or does not have a closed form representation. For instance in our case, the densities considered do not have closed form expressions. However, since there is a one to one correspondence between the density function and its Fourier transform it is worth exploiting the latter since it always exists and it is bounded. This leads us to next estimation method.

The fourth estimation approach is the empirical characteristic function (ECF) method discussed in [Yu04]. Although the likelihood function can be unbounded, its Fourier transform is always bounded and while the likelihood function might not be tractable or might not be of a closed form, the Fourier transform could have a closed form expression. The Fourier transform of the density function is the characteristic function (CF), hence the name empirical characteristic function (ECF) method. In this chapter we aim to show that this approach performs better than all the previously mentioned methods. A useful software package that can be used to estimate stable distributions is provided in [Nol97]. A more theoretical approach to statistical estimation of the parameters of stable laws is extensively discussed in [Zol80]. Readers interested in how to simulate stable process can refer to two excellent literatures of [WW95] and [Zol86].

In Chapter 3 we shall discuss in detail the four parameter estimation methods including the quantiles, logarithmic moments method, maximum likelihood (ML) and the empirical characteristics function (ECF) method. The contribution of the chapter is two-fold: first, we discuss the above parametric approaches and investigate their performance through error analysis. Moreover, we argue that the ECF performs better than the ML over a wide range of shape parameter values, α including values closest to 0 and 2 and that the ECF has a better convergence rate than the ML. Secondly, we compare the t location scale distribution to the general stable distribution and show that the former fails to capture skewness which might exist in the data. This is observed through applying the ECF to commodity futures log-returns data to obtain the stable parameters.

1.3.2 Literature on Commodity Future Pricing

In addition to parameter estimation described before, stable processes can be incorporated in stochastic models to capture observed market features such as jumps, business cycles, zero returns, mean-reversion, stochastic volatility. This will be the main objective of Chapter 4 with a specific focus on developing pricing representations for commodity futures. Trading in commodity markets gained momentum over the years after investors developed interest in diversifying their portfolio risk. It is known historically and in modern times that commodities have tremendous economic impact on nations and people. They play a significant role in the global economy and investors have benefited from economic events in the history of commodity markets. This in turn ignited a wave of various academic models for understanding their market.

Electricity and energy commodities can be suitably modelled using stable distributions due to their erratic price behaviours caused mainly by storage challenges. A similar commodity is bandwidth which depends on capacitated physical networks. Research based on modeling commodity futures and prices of options written on them emerged after the successful results of [Sch97] including the continuous models of [GS90, ELO14, SS00]. Early models suggested a single factor with constant volatility but after poor data fitting, improved multi-factor models were introduced that incorporated stochastic volatility and/or stochastic interest rates, seasonality etc. Jumps were incorporated in [DJX02], [HR98] through a Poisson jump process. Various energy commodity derivative models have been discussed in [Den98]. The author's models can be used to evaluate the generation and transimission capacity of electricity, determine the value of investment opportunities and the threshold value above which a firm should invest. A multi-factor jump-diffusion model which allows prices of long-dated futures contracts to jump by smaller magnitudes than short-dated futures contracts and includes stochastic interest rates is presented in [Cro08]. Stable continuous auto-regressive moving average spot models for Futures pricing in electricity markets are presented in [BKMV12]. The authors present a new model for the electricity spot price dynamics that is able to capture seasonlaity, extreme spikes in the market, low-frequency using a non-stationary independent increments process and large fluctuations using a non-Gaussian stable CARMA process. A more recent model based on affine-structure features appears in [KNPP15]. It is evident that commodity models are easily built using this affine-structure property. The reader is referred to [MS13, KRST11, KR08, DFS03, RH15] for a selected number of references on affine-structure models. Known existing models that incorporate jumps use the Poisson jump-type. In Chapter 4, we extend the models discussed in [KNPP15], introducing jumps using subordinated Brownian motion by a process drawn from a four parameter α-stable distribution family. We consider both a one- and two-factor mean-reverting models for obtaining future prices. However, we shall focus on the former to explain our approach in much more detail including numerical implementation. Our model is explicit, tractable and robust with the future price given as a function of the subordinator as opposed to the spot. Interestingly, the one-factor model can be reduced to a latent regression model usually used in population dynamics whose parameters are estimated using the expectation maximization approach.

1.3.3 Literature on Sensitivity Analysis

In equity, stable processes can be used through subordination of Brownian motion to introduce jumps in asset prices in a similar way as in the previous section for commodity modeling. However, the focus in Chapter 5 is on inverse stable processes. The motivation for the choice of model stems from the nature of emerging markets where liquidity is a challenge. In emerging markets trading is slow and zero returns on two or more consecutive days are a possibility, see for instance [ARSB09] where the author introduces a mixed-stable model to solve the passivity problem in emerging markets of Baltic states. Chapter 5 we wish to explore the sensitivity analysis in emerging markets. We model asset price dynamics by subordinating Brownian motion using an inverse stable process. The objective is to extend the results in [FLL+99, CF07] for continuous processes to jump processes based on the Bismut-Elworthy-Li (BEL) formula in [EL94]. Our approach of capturing jumps is different from the usual method of using a Poisson process, (see for instance [Mhl15]). We construct the jump process using a subordinated Brownian motion where the subordinator is an inverse α-stable process $(L_t)_{t\geq 0}$ with $0 < \alpha \leq 1$. We derive representation formulas for computing financial Greeks and show that in the event when $L_t \equiv t$, we retrieve the results in [FLL+99]. The purpose is to by-pass the derivative of an (irregular) payoff function in a jump-type market by introducing a weight term in form of an integral with respect to subordinated Brownian motion. Using Monte Carlo techniques, we estimate financial Greeks for a digital option and show that the BEL formula still performs better for a discontinuous payoff in a jump asset model setting and that the finite difference methods are better for continuous payoffs in a similar setting.

The problem of computating the Greeks of derivatives with smooth pay-off functions has been extensively studied. The problem where the pay-off function is not necessarily regular poses a different level of difficult and requires a different approach [FLL+99, CF07]. Exisiting and successful techniques avoid differentiating the pay-off function by introducing a weight function. The Bismut-Elworthy-Li (BEL) representation formula [EL94] is one scenerio of such innovations. In Chapter 5 we show that the known relationship between the Malliavin derivative and the first variation process still holds for an alpha-stable subordinated Brownian motion and results in an explicit martingale weight factor. This allows for an extension of the BEL formula to subordinated Brownian motion and as a result, Greeks can be easily computed in jump-type emerging markets. The subordinator belongs to the Lévy family of four parameter α-stable distributions. Price dynamics of almost all instruments in financial markets are observed to deviate from the Gaussian distribution. Various models in literature have been developed to closely estimate the dynamics of these markets. The rich and robust family of α-stable distributions has proven successful over most traditional models in capturing skewed and heavy tailed distributions. An application to estimate the densities of subordinated SDE under the Malliavin framework is discussed in [Kus10]. A rather different approach is discussed in [Wył12] where the authors, in addition to investigating the densities of subordinated Brownian motion, also discuss some properties related to transforms and averaged mean squared displacements of the process. They consider and compare both cases of stable processes and their inverses. In addition they provide some parameter estimation techniques. An intuitive study related to the work in Chapter 5 is [EL94] which is limited to the Delta. The authors derive derivatives of solutions of diffusion equations and demonstrate that they exhibit and allow for the estimation of the diffusion equations' smoothening properties. In addition, they use their results to study the logarithmic gradient of heat kernels. Their results can be extended to derive representation formulas for other Greeks. The work by [Zha12] on the derivative and gradient estimate for SDEs driven by stable processes is another intuitive literature vital to our study. Other related recent work include [Tak10, KKH10, Khe12, Mhl15]. A less related but still interesting work is by [SV03] where properties of a killed subordinated Brownian motion by an $\alpha/2$-stable are compared with those of the α-stable subordinated Brownian motion. They show possible comparability in their killing measures and propose bounds on the Green function and the jumping kernel of the subordinated $(\alpha/2)$ process. There is more existing literature on stable distributions that we cannot exhaustively discuss. Interested readers can refer to [SX14] for a comprehensive literature on stable distributions.

Chapter 2

Stable Distributions: Theory

This chapter is devoted to discussing the theory underpinning the rich and robust class of α-stable distributions which are central in this thesis. We shall adapt some of the definitions and notations of [KMT17].

2.1 Introduction

Stable also known as alpha-stable (or equivalently α-stable) processes belong to a general class of Lévy distributions. They are limiting distributions with a definitive exponent parameter $\alpha \in (0, 2]$ that determines their shape.

2.1.1 Definition and Construction

Definition 1 *Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed random variables and suppose a random variable S defined by*

$$\frac{1}{b_n}\left(\sum_{i=1}^{n} X_i - a_n\right) \longrightarrow S, \tag{2.1}$$

where "$\longrightarrow$" represents weak convergence in distribution, b_n is a positive constant and a_n is real. S follows a stable distribution and the constants a_n and b_n need not to be finite.

Definition 1 allows the modeling of a number of natural phenomena beyond normality using stable distributions. The fact that a_n and b_n do not necessarily have to be finite provides the generalized central limit theorem.

Theorem 1 (Generalised Central Limit Theorem [Rac03]) *Suppose $X_1, X_2 \cdots$ denotes a sequence of independent and identically distributed random variables and let sequences $a_n \in \mathbb{R}$ and $b_n \in \mathbb{R}^+$. Then we can define a sequence*

$$Z_n := \frac{1}{b_n}\left(\sum_{i=1}^{n} X_i - a_n\right) \tag{2.2}$$

of sums Z_n such that their distribution functions weakly converge to some limiting distribution:

$$\mathtt{P}(Z_n < x) \Rightarrow H(x), \quad n \longrightarrow \infty, \tag{2.3}$$

where $H(x)$ is some limiting distribution.

The traditional central limit theorem assumes finite mean $a := \mathtt{E}[X_i]$ and finite variance $\sigma^2 := \mathtt{V}ar[X_i]$ and defines the sequence of sums

$$Z_n := \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} X_i - na\right), \tag{2.4}$$

such that the distribution functions of Z_n weakly converge to $h^{sG}(x)$:

$$\mathtt{P}(x_1 < Z_n < x_2) \Rightarrow \int_{x_1}^{x_2} h^{sG}(x)\,\mathrm{d}x, \quad n \longrightarrow \infty, \tag{2.5}$$

where $h^{sG}(x)$ denotes the standard Gaussian distribution

$$h^{sG}(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2). \tag{2.6}$$

Suppose the independent and identically distributed random variables X_i equal to a positive constant c almost surely and the sequences a_n and b_n in (2.2) are defined by $a_n = (n-1)c$ and $b_n = 1$, then Z_n is also equal to c for all $n > 0$ almost surely. In this case the random variables X_i are mutually independent and as a result, the limiting distribution for the sums Z_n belongs to the stable family of distributions by definition. This is one reason why they are regarded as stable.

2.1.2 Parametrization

Definition 2 *A stable distribution is a four-parameter family denoted by $S(\alpha, \beta, \nu, \mu)$, where $\alpha \in (0, 2]$ is responsible for the shape of the distribution, $\beta \in [-1, 1]$ is responsible for skewness of the distribution, $\nu > 0$ is the scale parameter that narrows or extends the distribution around $\mu \in \mathbb{R}$ and μ is the location parameter that shifts the distribution to the left or the right.*

Suppose a random variable S follows a stable distribution $S(\alpha, \beta, \nu, \mu)$ then the random variable $Z = (S - \mu)/\nu$ has the same-shaped distribution as S but with the location parameter $\mu = 0$ and the scale parameter $\nu = 1$. This is another reason why they are referred to as stable, the shape is maintained after any re-scaling.

Densities of α-stable distributions do not have closed-form representations except for the case of a Gaussian ($\alpha = 2$), Cauchy ($\alpha = 1,\ \beta = 0$) and Inverse Gaussian or Pearson ($\alpha = 0.5,\ \beta = \pm 1$) distributions.

1. Gaussian distribution $N(\mu, \sigma^2)$: $S(2, 0, \frac{\sigma}{\sqrt{2}}, \mu)$.

$$h^G(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right); \quad -\infty < x < \infty.$$

2. Cauchy distribution: $S(1,0,\nu,\mu)$.

$$h^C(x) = \frac{1}{\pi}\frac{\nu}{\nu^2+(1-x)^2}; \quad -\infty < x < \infty.$$

3. Levy distribution (Inverse-Gaussian or Pearson): $S(1/2,1,\nu,\mu)$.

$$h^L(x) = \sqrt{\frac{\nu}{2\pi}}(x-\mu)^{-3/2}\exp\left(-\frac{\nu}{2(x-\mu)}\right); \quad \mu < x < \infty.$$

The densities are generally computed using characteristic functions through transformations such as the Fourier[1]. One can also refer to the work of [Zol64, Zol80, Zol86] for straight-forward and easy-to-compute integral representations of stable distribution and density functions. The distribution functions for the different α values have been tabulated in [Dum71], [FR68] and [HC73].

2.2 Density and Distribution Properties

2.2.1 Special Case

Let $(X_t, t \geq 0)$ denote a Lévy process in $\mathbb{R}$. The characterization of X_t is deduced from the Lévy-Khintchine formula.

Definition 3 (Lévy Measure, [BNMR12]) *Let $(X_t, t \geq 0)$ denote a Lévy process in $\mathbb{R}$, we define a Lévy measure by*

$$m(\mathrm{d}x) = \mathbb{1}_{(0,\infty)}(x)e^{-x}\,\mathrm{d}x.$$

[1]Note that characteristic functions always exist.

Definition 4 (Lévy-Khintchine) *[App04] Let $(X_t, t \geq 0)$ denote a Lévy process in $\mathbb{R}$. There exist $b \in \mathbb{R}$, $\sigma \geq 0$ such that the characteristic function of X is given by*

$$\Phi(t) := \mathtt{E}[e^{itX}] = \exp(itb - \frac{1}{2}\sigma^2 t^2 + \int_{\mathbb{R}-\{0\}} (e^{itx} - 1 - itx\mathbb{1}_{|x|<1})m(\mathrm{d}x)), \tag{2.7}$$

where $\mathbb{1}_{\{\cdot\}}$ is an indicator function and m is a σ-finite measure satisfying the constraint

$$\int_{\mathbb{R}-\{0\}} \min(1, |x|^2)m(\mathrm{d}x) < \infty; \quad \textit{alternatively} \quad \int_{\mathbb{R}-\{0\}} \frac{|x|^2}{1+|x|^2} m(\mathrm{d}x) < \infty. \tag{2.8}$$

Definition 5 (The Lévy-Itô Decomposition [App04]) *If X_t is a Lévy process, there exist $b \in \mathbb{R}$, a Brownian motion $B_\sigma(t)$ with variance $\sigma \in \mathbb{R}^+$ and an independent Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R} - \{0\})$ such that, for each $t \geq 0$,*

$$X_t = bt + B_\sigma(t) + \int_{|x|<1} x\tilde{N}(t, \mathrm{d}x) + \int_{|x|\geq 1} xN(t, \mathrm{d}x), \tag{2.9}$$

where

$$b = \mathtt{E}[X_1 - \int_{|x|\geq 1} xN(1, \mathrm{d}x)]. \tag{2.10}$$

The compensated compound Poisson random measure is defined by $\tilde{N} := N - t\lambda$ to preserve the martingale property where the Lévy measure λ satisfies (2.8).

A stable distribution can be constructed by setting σ to zero in (2.7) or the second term on the right of (2.9) to zero and the Lévy measure in (2.8) to

$$m(\mathrm{d}x) = \frac{C}{|x|^{1+\alpha}} \mathrm{d}x; \quad C > 0, \tag{2.11}$$

This gives a pure jump Lévy process which is a simple example of a stable family of distributions. We discuss a general case in the following.

2.2.2 General Case

In the following, $(S_t)_{t\geq 0}$ will represent a stable process. The characteristic function Φ of S_t is obtained using the domain of attraction of stable random variables (See [GKP+99]) and the Lévy-Khinchine representation formula (See Definition 4 or [App04] for a detailed explanation) i.e.

$$\Phi(\theta) = \mathtt{E}[\exp(i\theta S_t)] = \begin{cases} \exp(-\nu^\alpha|\theta|^\alpha[1 - i\beta\text{sign}(\theta)\tan(\frac{\pi\alpha}{2})] + i\mu\theta); & \text{for } \alpha \neq 1. \\ \exp(-\nu|\theta|[1 + i\beta\text{sign}(\theta)\frac{2}{\pi}\log|\theta|] + i\mu\theta); & \text{for } \alpha = 1. \end{cases} \tag{2.12}$$

Alternative forms of parametrization are discussed in [McC86] for simpler numerical implementation. We expand more on this in Section 2.3.

The density of S_t is computed from (2.12) using the Fourier transform. That is

$$h_S(t,u) = \frac{1}{\pi}\int_0^\infty e^{-iu\theta}\Phi(\theta)\,\mathrm{d}\theta. \tag{2.13}$$

Figure 2.1 shows density graphs of S_t for different exponent parameter values.

The drawback in approximating (2.13) is that elementary techniques such as expressing the integral in terms of simple functions or using infinite polynomial expressions of the density function are not sufficient for meaningful numerical analysis. Some authors propose a standard parameterized integral expression of the density given by (see for instance [AO16])

$$h_{S_t}(\alpha,\beta,\nu,\mu) = \frac{1}{\sigma\pi}\int_0^\infty e^{-t^\alpha}\cdot\cos\left(t\cdot(\frac{s-\mu}{\sigma}) - \beta t^\alpha\tan(\frac{\pi\alpha}{2})\right)\mathrm{d}t. \tag{2.14}$$

However, this representation consists of an oscillating integrand. A much better ap-

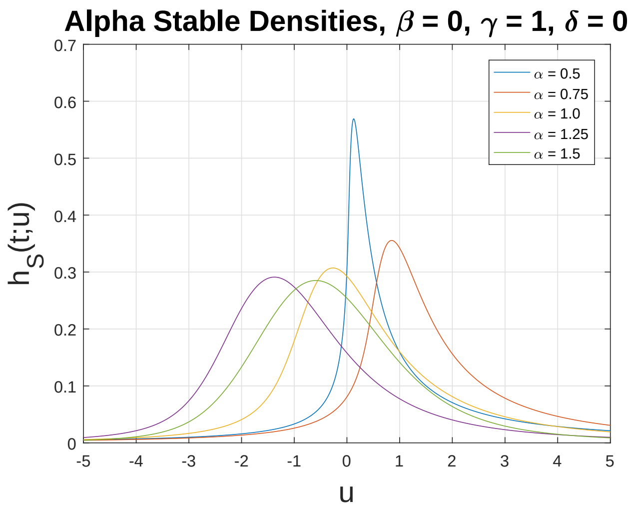


Fig. 2.1: α-Stable densities of S_t for $\alpha \in (0, 2]$.

proach is proposed in [Zol86] where the density of S_t is given by

$$h_{S_t}(\alpha,\beta,\nu,\mu) = \begin{cases} \frac{\alpha|\frac{s-\mu}{\sigma}|^{\frac{1}{\alpha-1}}}{2\sigma|\alpha-1|} \int_{-\theta}^{1} U_\alpha(\varphi,\theta) \exp\left(-|\frac{s-\mu}{\sigma}|^{\frac{\alpha}{\alpha-1}} U_\alpha(\varphi,\theta)\right) \mathrm{d}\varphi; \text{if } s \neq \mu \\ \frac{1}{\pi\sigma} \cdot \Gamma(1+\frac{1}{\alpha}) \cdot \cos\left(\frac{1}{\alpha}\arctan(\beta \cdot \tan(\frac{\pi\alpha}{2}))\right); \text{if } s = \mu, \end{cases} \tag{2.15}$$

$$U_\alpha(\varphi,\vartheta) = \left(\frac{\sin(\frac{\pi}{2}\alpha(\varphi+\vartheta))}{\cos(\frac{\pi\varphi}{2})}\right)^{\frac{\alpha}{1-\alpha}} \cdot \left(\frac{\cos(\frac{\pi}{2}(\alpha-1)\varphi+\alpha\vartheta)}{\cos(\frac{\pi\varphi}{2})}\right), \tag{2.16}$$

where $\theta = \arctan(\beta \tan\frac{\pi\alpha}{2})\frac{2}{\alpha\pi}\mathrm{sign}(s-\mu)$.

Definition 6 *[MS13] The inverse L_t of S_t, $t \in [0,T]$ is defined by*

$$L_s := \begin{cases} \inf\{t : S_t > s\} & \text{if } s \in [0, S_t) \\ T & \text{if } s = S_T. \end{cases} \tag{2.17}$$

Figure 2.2 shows the graphical representations of S_t and L_t.

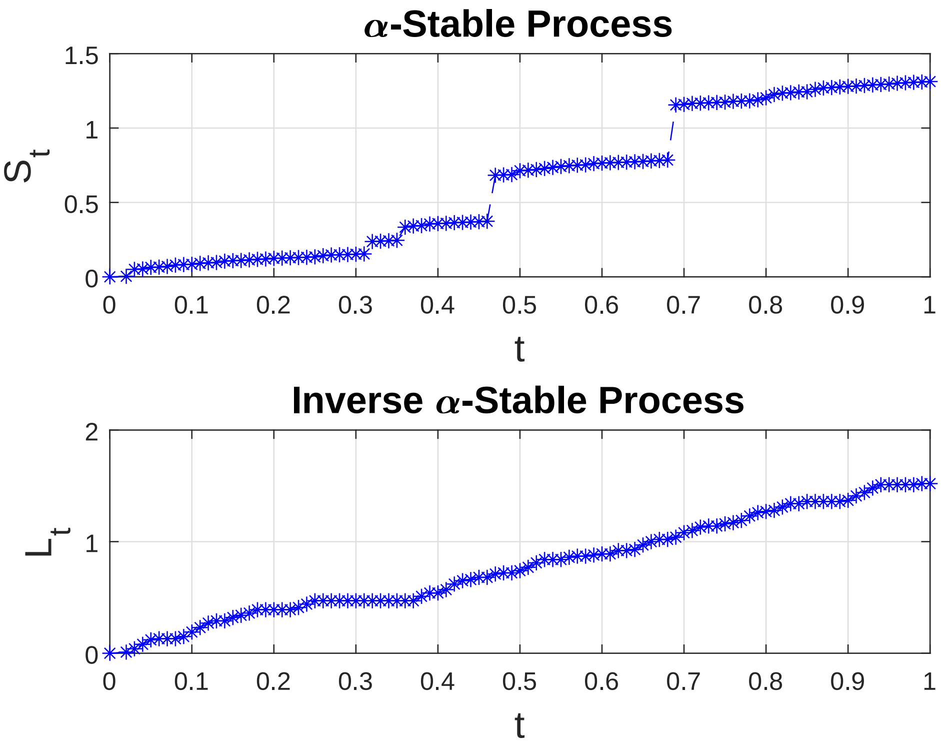


Fig. 2.2: Stable process S_t and the inverse stable process L_t, $\alpha = 0.8$.

For $l \in [0, T]$, it is readily seen that the following equivalence relation holds:

$$S_l < t \Longleftrightarrow L_t \geq l. \tag{2.18}$$

The process L_t is also interpreted as the first passage time of S. Moreover

$$L_{S_t} = t \quad \text{and} \quad S_{L_{t-}} \leq t \leq S_{L_t}. \tag{2.19}$$

Let $h_L(l;t)$ denote the density function of L_t. Using relation (2.18) we deduce $F(t;l) := \mathtt{P}(S_l < t) = \mathtt{P}(L_t \geq l) = \int_l^\infty h_L(\tau;t)\,\mathrm{d}\tau$ which implies

$$h_L(l;t) = -\frac{\partial F(t;l)}{\partial l} = -\frac{\partial}{\partial l}\int_{-\infty}^{t} h_S(u,l)\,\mathrm{d}u. \tag{2.20}$$

We can therefore approximate $h_L(l;t)$ by estimating the integral in (2.20) using the density $h_S(t;l)$ in (2.13) and its characteristic function (2.12).

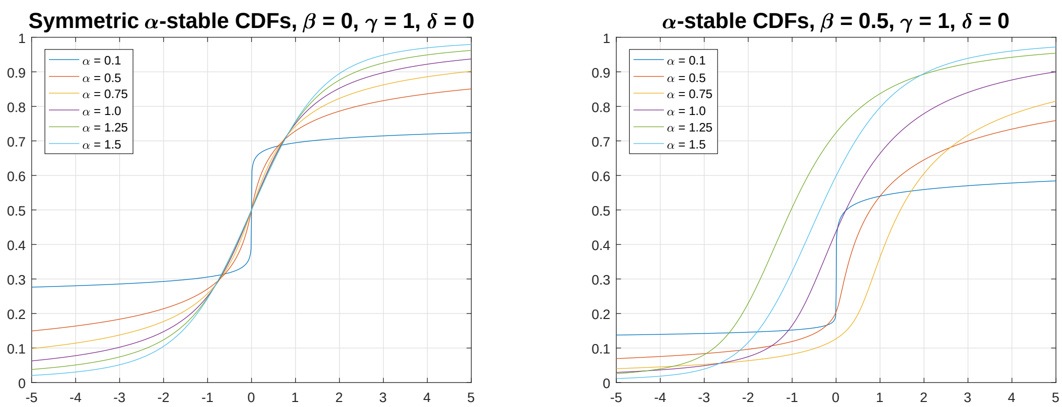
According to [MS13], the density $h(t,u)$ can also be given by

$$h_S(t,u) = u^{-1/\alpha} h(tu^{-1/\alpha}), \tag{2.21}$$

where $h(\tau)$ is the density of a standard stable process with a Laplace transform $\tilde{h}(\tau) = \exp(-\tau^\alpha)$. This follows from the fact that S_u has the same distribution as $u^{1/\alpha} S_1$. As a result, the density of the inverse stable process L_t can be given in terms of the standard stable process by

$$h_L(u,t) = \frac{t}{\alpha} u^{-1-1/\alpha} h(tu^{-1/\alpha}). \tag{2.22}$$

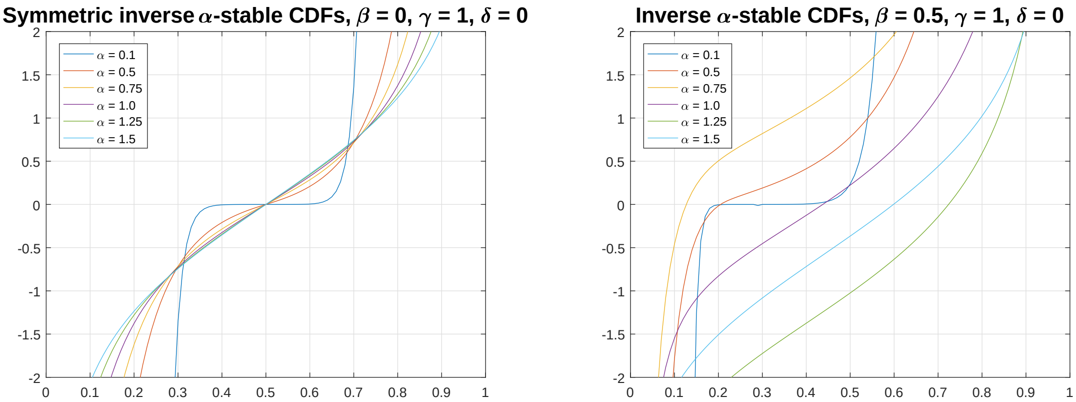
The cumulative distribution functions of both the symmetric and asymmetric stable distributions are plotted in Figure 2.3 and for their inverse counterparts in Figure 2.4



(a) CDF of symmetric Stable

(b) CDF of skewed Stable

Fig. 2.3: Cumulative distribution functions of symmetric and skewed stable distributions.



(a) CDF of inverse symmetric Stable (b) CDF of inverse skewed Stable

Fig. 2.4: Cumulative distribution functions of inverse symmetric and inverse skewed stable distributions.

2.2.3 Some Properties of Stable Distribution Functions

Firstly, recall that for any two admissible sets of parameters of stable distributions we can find two unique numbers $a > 0$ and b such that

$$S(\alpha, \beta, \nu, \mu) \stackrel{d}{=} aS(\alpha, \beta, \nu', \mu'), \tag{2.23}$$

where

$$a = \frac{\nu}{\nu'}, \quad b = \begin{cases} \mu - \mu' \frac{\nu}{\nu'}; & \alpha \neq 1 \\ \mu - \mu' \frac{\nu}{\nu'} + \nu\beta\frac{2}{\pi}\log\frac{\nu}{\nu'}; & \alpha = 1. \end{cases} \tag{2.24}$$

The intuition is that a general stable distribution can be expressed in terms of a standard stable distribution. That is, we can write $S(\alpha, \beta, \nu, \mu) \stackrel{d}{=} aS(\alpha, \beta, 1, 0) + b$ where

$$a = \nu, \quad b = \begin{cases} \mu; & \alpha \neq 1 \\ \mu + \frac{2}{\pi}\beta\nu\log\nu; & \alpha = 1. \end{cases} \tag{2.25}$$

Secondly, suppose h, H and Φ denote the respective probability, cumulative distribution and characteristic functions of a stable random variable, S, where

$$h(s,\alpha,\beta) = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\cos st - i\sin st)\Phi(t,\alpha,\beta)\,\mathrm{d}t,$$

then it is readily seen that the following properties hold:

1. $h(-s,\alpha,\beta) = h(s,\alpha,-\beta)$.

2. $H(-s,\alpha,\beta) = 1 - H(s,\alpha,-\beta)$.

3. $\Phi(-s,\alpha,\beta) = \Phi(s,\alpha,-\beta)$.

The above three relations can be verified by trigonometric properties.

2.3 Simulating α-Stable Random Variables

The two excellent references for simulating stable processes are [Zol86] and [CMS76].

Definition 7 *Suppose S_t is a stable process with parameters $(\alpha, \beta_2, \nu_2, \mu)$, the characteristic function is given by*

$$\ln\Phi(\theta) = \begin{cases} i\mu\theta - \nu_2^{\alpha}|\theta|^{\alpha}\exp(-i\beta_2 sign(\theta))\frac{\pi}{2}K(\alpha)); & \alpha \neq 1, \\ i\mu\theta - \nu_2|\theta|(\frac{\pi}{2} + i\beta_2 sign(\theta))\ln|\theta|); & \alpha = 1, \end{cases} \tag{2.26}$$

where

$$K(\alpha) = \alpha - 1 + sign(1-\alpha) = \begin{cases} \alpha; & \alpha \neq 1 \\ \alpha - 2; & \alpha = 1, \end{cases} \tag{2.27}$$

$$(\beta_2, \nu_2) = \begin{cases} \frac{2}{\pi K(\alpha)} \tan^{-1}(\beta \tan \frac{\pi\alpha}{2}),\ \nu(1+\beta^2 \tan^2 \frac{\pi\alpha}{2})^{\frac{1}{2\alpha}}; & \alpha \neq 1 \\ (\beta, \frac{2}{\pi}\nu); & \alpha = 1. \end{cases} \tag{2.28}$$

Lemma 1 *Let $\gamma \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ be a uniformly distributed random variable and let W be an independent exponential random variable with mean 1. Then*

$$S = \begin{cases} \frac{\sin \alpha(\gamma + \frac{\pi}{2}\beta_2 \frac{K(\alpha)}{\alpha})}{(\cos\gamma)^{\frac{1}{\alpha}}} (\frac{\cos(\gamma - \alpha(\gamma + \frac{\pi}{2}\beta_2 \frac{K(\alpha)}{\alpha}))}{W})^{\frac{1-\alpha}{\alpha}}; & \alpha \neq 1 \\ (\frac{\pi}{2} + \beta_2\gamma)\tan\gamma - \beta_2 \log(\frac{W\cos\gamma}{\frac{\pi}{2}+\beta_2\gamma}); & \alpha = 1 \end{cases} \tag{2.29}$$

is a standard α-stable process with parameters $(\alpha, \beta_2, 1, 0)$.

Proof 1 *See [Zol86].*

A stable random variable can be easily generated using Lemma 1. Programming languages such as R or MATLAB can be utilised to generate a uniformly distributed random variable U on the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ and an independent exponential random variable E with mean 1[2]. Then the stable random variable would be generated by computing

$$S = \begin{cases} A_{\alpha,\beta} \frac{\sin(\alpha(U + B_{\alpha,\beta}))}{(\cos U)^{\frac{1}{\alpha}}} \left(\frac{\cos(U - \alpha(U + B_{\alpha,\beta}))}{E} \right)^{\frac{1-\alpha}{\alpha}}; & \alpha \neq 1 \\ \frac{2}{\pi} \left((\frac{\pi}{2} + \beta U)\tan U - \beta \log(\frac{\frac{\pi}{2} E \cos U}{\frac{\pi}{2} + \beta U}) \right); & \alpha = 1 \end{cases} \tag{2.30}$$

where $A_{\alpha,\beta} = (1 + \beta^2 \tan^2 \frac{\pi\alpha}{2})^{\frac{1}{2\alpha}}$ and $B_{\alpha,\beta} = \frac{\tan^{-1}(\beta \tan \frac{\pi\alpha}{2})}{\alpha}$.

[2]These are easily obtained from in-built functions in MATLAB

2.4 Moments of Stable Processes

Statistical moments $\mathtt{E}[|\cdot|^k]$ of stable distributions are finite only when $k \leq \alpha$. Moreover, for $\alpha < 2$ the variance is infinite, for $\alpha \in (0,1]$ the mean does not exist and the mean is zero when $\alpha \in (1,2)$. This is not always the case for symmetric stable distributions where $\beta = 0$.

2.4.1 Fractional Lower Order Moments

The FLOM is an alternative for computing moments of alpha-stable random variables especially in situations where the mean and/or variance are infinite. FLOM representation formulas are discussed in [MN95] for symmetric stable random data and its generalization to asymmetric stable random data in [Kur01]. In the latter, if $S_i \sim S(\alpha, \beta, \nu, \gamma)$ and $\alpha \neq 1$, then

$$\mathtt{E}[S^{<p>}] = \frac{\Gamma(1-\frac{p}{\alpha})}{\Gamma(1-p)}\left|\frac{\gamma}{\cos\theta}\right|^{\frac{p}{\gamma}} \frac{\sin(\frac{p\theta}{\alpha})}{\sin(\frac{p\pi}{2})}, \quad \text{for} \quad p \in (-2,-1) \cup (-1,\alpha).$$
$$\mathtt{E}[|S|^p] = \frac{\Gamma(1-\frac{p}{\alpha})}{\Gamma(1-p)}\left|\frac{\gamma}{\cos\theta}\right|^{\frac{p}{\gamma}} \frac{\cos(\frac{p\theta}{\alpha})}{\cos(\frac{p\pi}{2})}, \quad \text{for} \quad p \in (-1,\alpha),$$

where $\theta = \arctan(\beta \tan\frac{\alpha\pi}{2})$ and Γ denotes the Gamma function. We also define

$$x^{<p>} = sign(x)|x|^p. \tag{2.31}$$

From the above representations, moments with negative values of p are attainable. This results in the logarithmic moments approach that provides an easier way of estimating stable distribution parameters compared to the FLOM.

2.4.2 Logarithmic Moments

This approach is as a result of the challenges encountered when using the FLOM method which requires computing Gamma functions, the inversion of the *sinc* function and it only works for some p. The current method suggests computing derivatives with respect to the moment order p resulting in moments of the logarithms of the stable process. We illustrate in the following.

Lemma 2 *Let S denote a symmetric stable random variable and let $p \in \mathbb{R}$. Then*

$$M_n := \mathtt{E}[(\log|S|)^n] = \lim_{p\to 0} \frac{\mathrm{d}^n}{\mathrm{d}p^n} \mathtt{E}[|S|^p], \quad n = 1, 2, \cdots. \tag{2.32}$$

The moments follow readily for $n = 1, 2, \cdot$. That is

$$M_1 = \mathtt{E}[\log|S|] = \varphi_0\left(1 - \frac{1}{\alpha}\right) + \frac{1}{\alpha}\log\left|\frac{\nu}{\cos\theta}\right|. \tag{2.33}$$

$$M_2 = \mathtt{E}[(\log|S| - \mathtt{E}[\log|S|])^2] = \varphi_1\left(\frac{1}{2} + \frac{1}{\alpha^2}\right) - \frac{\theta^2}{\alpha^2}. \tag{2.34}$$

$$M_3 = \mathtt{E}[(\log|S| - \mathtt{E}[\log|S|])^3] = \varphi_3\left(1 - \frac{1}{\alpha^3}\right), \tag{2.35}$$

where $\theta = \arctan(\beta\tan\alpha\pi/2)$ and terms φ_k are given by $\varphi_0 = -0.57721566$, $\varphi_1 = \pi^2/6$, $\varphi = 1.2020569$ derived from the polygamma function

$$\varphi_{k-1} = \frac{\mathrm{d}^k}{\mathrm{d}x^k}\log\Gamma(x)|_{x=1}. \tag{2.36}$$

Proof 2 *The proof is provided in [Kur01].*

2.5 Laplace Transforms

Definition 8 *Let X_u be a subordinator. The Laplace transform of X_u is defined by*

$$\mathtt{E}[e^{-\tau X_u}] = e^{-u\phi(\tau)}, \tag{2.37}$$

where ϕ is the Laplace exponent of X_u known as the Bernstein function represented by

$$\phi(\tau) = a + b\tau + \int_{(0,\infty)} (1 - e^{-\tau x})\Pi(\mathrm{d}x). \tag{2.38}$$

where $a, b > 0$ and Π is the Lévy measure on $(0, \infty)$ such that $\int \frac{x}{1+x}\Pi(\mathrm{d}x) < \infty$.

The Laplace tranform of the stable process S_u is given by (see [MS13])

$$\tilde{h}_S(\tau, u) = \int_0^\infty e^{-t\tau} h_S(t, u)\,\mathrm{d}t = \exp(-uC\Gamma(1-\alpha)\tau^\alpha) = \exp(-u((\tau+\beta)^\alpha + \beta^\alpha)), \tag{2.39}$$

where $0 \le \beta \le 1$. For $C = \Gamma(1-\alpha)$ (alternatively $\beta = 0$), the Laplace transform simplifies to that of a standard stable process:

$$\tilde{h}_S(\tau, u) = \mathtt{E}[e^{-\tau S_u}] = \exp(-u\tau^\alpha); \qquad 0 < \alpha < 1. \tag{2.40}$$

The Laplace transform $\tilde{h}_L(u;\tau)$ of the inverse stable process L_t is obtained from (2.20):

$$\begin{aligned}\tilde{h}_L(u, \tau) &= -\frac{\partial}{\partial u}(\tau^{-1}\exp(-u((\tau+\beta)^\alpha + \beta^\alpha))), \\ &= \tau^{-1}((\tau+\beta)^\alpha + \beta^\alpha)\exp(-u((\tau+\beta)^\alpha + \beta^\alpha)),\end{aligned} \tag{2.41}$$

where the Laplace transform of $\int_0^t f(y)\,\mathrm{d}y$ is $\tau^{-1}\tilde{f}(\tau)$ and $h_L(u, \tau) := 0$ for $l < 0$ or $\tau < 0$. Since (2.41) does not have the general form for a Laplace transform of a Lévy process, then L_t is not a Lévy process.

2.6 Moment Generating Function

There is a relationship between a moment-generating function of a random variable and its Laplace transform.

Lemma 3 *Let $M_u(\tau)$ and $\tilde{h}(\tau, u)$ denote the respective moment-generating function and Laplace transform of a random variable then*

$$M_u(\tau) = \tilde{h}(\tau, -u) + \tilde{h}(-\tau, u), \tag{2.42}$$

where

$$\tilde{h}(a, b) = \int_0^\infty e^{-ta} h(t, b)\, \mathrm{d}t.$$

Proof 3 *The relationship is verified in [Mil51].*

As a consequence of Lemma 3 and the explicit Laplace transform given by (2.39), we can deduce the first and second moments of S_t. That is

$$\begin{aligned} M_u(\tau) &= \exp(u((\tau+\beta)^\alpha + \beta^\alpha)) + \exp(-u((-\tau+\beta)^\alpha + \beta^\alpha)). \\ \mathtt{E}[S_t] &= M_u'(0) = \alpha u \beta^{\alpha-1}[e^{2u\beta^\alpha} + e^{-2u\beta^\alpha}]. && (2.43) \\ \mathtt{V}ar[S_t] &= M_u''(0) - (M_u'(0))^2. && (2.44) \\ &= \alpha(\alpha-1)u\beta^{\alpha-2}[e^{2u\beta^\alpha} - e^{-2u\beta^\alpha}] + \alpha^2 u^\alpha \beta^{2\alpha}[\beta^2 e^{2u\beta^\alpha} + e^{-u\beta^\alpha}]. && (2.45) \end{aligned}$$

2.7 Subordination

In this section we investigate moments of subordinated Brownian motion using semigroup properties. Detailed literature on Markov processes, semigroups and infinitesimal generators can be found in [App04] and [SV03]. We denote by B_{S_t} subordinated Brownian motion, where S_t is an α-stable process introduced above with $\alpha \in (0, 1)$.

Definition 9 (Joint Probability Space) *The notation* $(\Omega, \mathcal{F}, \mathtt{P}) = (\Omega^X \times \Omega^S, \mathcal{F}^X \times \mathcal{F}^S, \mu^X \times \mu^S)$ *shall denote a complete joint probability space endowed with a filtration* $(\mathcal{F}_t)_{t\geq 0}$ *such that* $\mathcal{F}_t = \mathcal{F}_t^X \vee \mathcal{F}_t^S$ *where* $\mathcal{F}_t^X$ *and* $\mathcal{F}_t^S$ *are filtrations generated by* X_t *and* S_t *respectively. The process* $\{S_t\}_{t\geq 0}$ *is a* α*-stable subordinator.*

Definition 9 ensures both X_t and S_t are adapted to the filtration $\mathcal{F}_t$.

Figure 2.5 shows three graphs of returns of standard Brownian motion, subordinated Brownian motion by L_t and by S_t.

Fig. 2.5: Brownian motion and Subordinated Brownian motion: $\alpha = 0.8$.

As seen from the bottom graph, the occasional spikes in the subordinated Brownian motion by S_t indicate the process is a good model for capturing jumps. On the other hand, in addition to capturing jumps, subordination by L_t is suitable for modeling illiquid markets where trading is limited and as a result zero returns are a possibility.

Now suppose $X_t = B_t$ is standard Brownian motion then we have the following lemma.

Lemma 4 *Suppose* $B = (B_t, \mathtt{P}^x)$ *is Brownian motion in* $\mathbb{R}$ *with transition density* $p(x, y; t) = p(y - x; t)$ *given by*

$$p(x;t) = \frac{1}{2\sqrt{\pi t}} \exp(-|x|^2/4t), \quad t > 0, x, y \in \mathbb{R}. \tag{2.46}$$

The semigroup $(P_t : t \geq 0)$ *of* B *is given by*

$$P_t f(x;t) = \mathtt{E}_x[f(B_t)] = \int_{\mathbb{R}} p(x, y; t) f(y; t)\,\mathrm{d}y, \tag{2.47}$$

where f *is a non-negative Borel function on* $\mathbb{R}$.

Lemma 4 follows from the fact that B_t is a Markov process whose generator is

$$\mathcal{G}f(x;t) := \lim_{t \downarrow 0} \frac{\mathtt{E}_x[f(x;t)] - f(x;t)}{t} = \lim_{t \downarrow 0} \frac{P_t f(x;t) - f(x;t)}{t}.$$

Lemma 5 *Suppose* $Y_t := B_{S_t}, t \geq 0$ *is a subordinated Brownian motion. Its semigroup* $(Q_t : t \geq 0)$ *is defined by*

$$Q_t f(y,t) = \mathtt{E}_y[f(Y_t, t)] = \mathtt{E}_y[f(B_{S_t})] = \int_0^\infty P_u f(y, u) h_S(t, u)\,\mathrm{d}u. \tag{2.48}$$

Then, the semigroup Q_t has a transition density $q(y,z,t) = q(y-z,t)$ defined by

$$q(y,t) = \int_0^\infty p(y,u)h_S(t,u)\,\mathrm{d}u. \tag{2.49}$$

Since B_{S_t} is a Markov process a similar argument for Lemma 4 applies for Lemma 5.

Lemma 6 *The mean and variance of B_{S_t} are computed as*

$$\mathtt{E}_y[B_{S_t}] = \int_0^\infty \mathtt{E}_y[B_s]h_S(t,u)\,\mathrm{d}u = 0. \tag{2.50}$$

$$\mathtt{E}_y[B_{S_t}^2] = \int_0^\infty \mathtt{E}[B_s^2]h_S(t,u)\,\mathrm{d}u = \int_0^\infty uh_S(t,u)\,\mathrm{d}u = \mathtt{E}_y[S_t], \tag{2.51}$$

The last term shows the variance of subordinated Brownian motion is non-existent for $\alpha \in (0,1)$. However, for $\alpha \geq 1$ the variance of subordinated Brownian motion is equal to the mean of the subordinator.

Proof 4 *Suppose f in Lemma 4 and Lemma 5 is such that $f(z,t) = z$. Using (2.47) and (2.48) and partitioning the time interval $[0,T]$ such that $0 \leq \tau_1 < \cdots < \tau_n \leq T$, where τ_i are the jump times of the process B_{S_t}, we observe that (2.50) and (2.51) hold for every interval $[\tau_i, \tau_{i+1})$. Thus, in the limits, their sums converges respectively to 0 and $\mathtt{E}_x[S_t]$ on $[0,T]$.* □

Note that for $\alpha \in (0,1]$, $\mathtt{E}[L_t]$ exists and can be computed. If $L_t \equiv t$ in (2.50) and (2.51) we recover standard Brownian motion with mean 0 and variance t. In general, the $k-$th moment of L_t is given by

$$\langle L_t^k \rangle = \frac{\Gamma(k+1)t^{k\alpha}}{\Gamma(k\alpha+1)}, \tag{2.52}$$

where $k \geq 1,\ (k \in \mathbb{R})$ and $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}\,\mathrm{d}x$ is the Gamma function.

Lemma 7 *The covariance of B_{L_t} is given by*

$$Cov[B_{L_t}, B_{L_s}] = \min(\mathtt{E}[L_t], \mathtt{E}[L_s]).$$

Proof 5 *Let $s \leq t$ then $L_s \leq L_t$ and we have*

$$B_{L_t} = B_{L_s} + (B_{L_t} - B_{L_s}).$$

$$B_{L_t} B_{L_s} = B_{L_s}^2 + B_{L_s}(B_{L_t} - B_{L_s}).$$

Since for all $t \in \mathbb{R}_+$, B_{L_t} has independent increments with zero mean, we have

$$\begin{aligned}
Cov[B_{L_t}, B_{L_s}] &= \mathtt{E}[B_{L_t} B_{L_s}] \\
&= \mathtt{E}[B_{L_s}^2] + \mathtt{E}[B_{L_s}(B_{L_t} - B_{L_s})] \\
&= \mathtt{E}[B_{L_s}^2] + \mathtt{E}[B_{L_s}]\mathtt{E}[(B_{L_t} - B_{L_s})] \\
&= \mathtt{E}[B_{L_s}^2] \\
&= \mathtt{E}[L_s].
\end{aligned}$$

Similarly for $L_t \leq L_s$ we have the covariance as $\mathtt{E}[L_t]$. Then, we write

$$Cov[B_{L_t}, B_{L_s}] = \min(\mathtt{E}[L_t], \mathtt{E}[L_s]).$$

□

Lemma 8 *[Boc12] Let X be a Lévy process with characteristic exponent Ψ and S an independent subordinator with Laplace exponent Φ. Then the subordinated process X_{S_t} is a Lévy process with characteristic exponent*

$$m(\cdot) = \Phi(\Psi(\cdot)). \tag{2.53}$$

Proof 6 *The proof is given in [Boc12].* □

It is known that any Lévy process X_s, $t < s \leq T$ with drift μ is fully determined by its characteristic function given by (see [FR07])

$$\mathtt{E}[e^{i\lambda X_s}] = e^{\mu\Delta + \Psi(\lambda)\Delta}, \tag{2.54}$$

where $\Delta = s - t$, μ is the drift parameter and $\Psi(\lambda)$ is the characteristic exponent. A typical example is Brownian motion whose characteristic function is given by

$$\mathtt{E}[e^{i\lambda B_s}] = e^{\mu\Delta - \frac{1}{2}\sigma^2\lambda^2\Delta}, \quad \text{where} \quad \Psi(\lambda) = -\frac{1}{2}\sigma^2\lambda^2. \tag{2.55}$$

Therefore the characteristic exponent of subordinated Brownian motion B_{S_t} can be deduced from (2.39), (2.53), (2.55):

$$m(u) = (-\frac{1}{2}\sigma^2 u^2 + \beta)^\alpha + \beta^\alpha. \tag{2.56}$$

Result (2.56) will be very useful in obtaining the results of Chapter 4, see Theorem 4.

Chapter 3

Parameter Estimation of Stable Distributions

3.1 Introduction

This chapter explores the theory behind the rich and robust family of α-stable distributions to estimate parameters from financial asset log-returns data. We discuss four parameter estimation methods including the quantiles, logarithmic moments, maximum likelihood (ML) and the empirical characteristics function (ECF) method. The contribution of the chapter is two-fold: first, we discuss the above parametric approaches and investigate their performance through error analysis. Moreover, we argue that the ECF performs better than the ML over a wide range of shape parameter values, α including values closest to 0 and 2 and that the ECF has a better convergence rate than the ML. Secondly, we compare the t-location scale distribution to the general stable distribution and show that the former fails to capture skewness which might exist in the data. This is observed through applying the ECF to commodity futures log-returns data to obtain the skewness parameter. The study provides useful information for portfolio managers, speculators and hedgers. It is therefore imperative that the most accurate estimation method is established. It is known that in general, market data deviates from the Gaussian distribution, its distribution is either skewed, high or low peaked and/or with fat or skinny tails. The current chapter is geared towards establishing a better parameter estimation method between the commonly known ECF, ML, quantile and logarithm moments methods used in economic and financial analysis for skewed data assumed to follow stable distributions. The results in this chapter have been published (see [KMT17]).

The rest of this Chapter is organized as follows: The following Section 3.2 explains how the four parameter estimation methods discussed in this chapter work and provides an analysis on their accuracy. In Section 3.3 we study and analyze some commodity data and show that the data deviates from the normal distribution hypothesis. We use the ECF to obtain the four stable parameters from the data and in addition fit it to various distributions to determine its closest shape which turns out to be the t-location scale distribution. This distribution is suited for data that is highly peaked and heavily tailed with outliers. However, we propose stable distribution fitting to check for any existing tails. Section 3.4 summarizes.

3.2 Estimation Methods

The four common methods for estimating parameters of stable processes include: quantiles method (see [FR71] and [McC86], [McC96]), the logarithmic moments method (see [Kur01]), the empirical characteristics method (see [Yan12]) and the maximum likelihood method (see [Nol01]). We investigate their accuracy in the following.

3.2.1 The Quantiles Method

The quantile method was pioneered by [FR71] but was much more appreciated through [McC86] after its extension to include asymmetric distributions and for cases where $\alpha \in [0.6, 2]$ unlike the former approach that restricts it to $\alpha \geq 1$.

Suppose $\hat{s}$ is a given data sample then the estimates for α and β are given by $\hat{\alpha} =$

$\Theta_1(\hat{\vartheta}_\alpha, \hat{\vartheta}_\beta)$ and $\hat{\beta} = \Theta_2(\hat{\vartheta}_\alpha, \hat{\vartheta}_\beta)$ where

$$\hat{\vartheta}_\alpha = \frac{\hat{s}_{0.95} - \hat{s}_{0.05}}{\hat{s}_{0.75} - \hat{s}_{0.25}}, \quad \hat{\vartheta}_\beta = \frac{\hat{s}_{0.95} + \hat{s}_{0.05} - 2\hat{s}_{0.05}}{\hat{s}_{0.95} - \hat{s}_{0.05}}. \tag{3.1}$$

The notation $\hat{s}_q$ represents the qth quantile of sample $\hat{s}$ and, $\hat{\alpha}$ and $\hat{\beta}$ are obtained by functions $\Theta_1(\hat{\vartheta}_\alpha, \hat{\vartheta}_\beta)$ and $\Theta_2(\hat{\vartheta}_\alpha, \hat{\vartheta}_\beta)$ given in Tables III and IV in [McC86] through linear interpolation. Consequently, the scale parameter is given by

$$\hat{\nu} = \frac{\hat{s}_{0.75} - \hat{s}_{0.25}}{\Theta_3(\hat{\alpha}, \hat{\beta})}, \tag{3.2}$$

where $\Theta_3(\hat{\alpha}, \hat{\beta})$ is given by Table V in [McC86]. The consistent estimator ν is then obtained through interpolation.

Finally the location parameter μ is estimated through a new parameter defined by

$$\zeta = \begin{cases} \mu + \beta\gamma \tan \frac{\pi\alpha}{2}; \alpha \neq 1 \\ \mu; \alpha = 1. \end{cases} \tag{3.3}$$

Moreover, ζ is estimated by

$$\hat{\zeta} = \hat{s}_{0.5} + \hat{\nu}\Theta_5(\hat{\alpha}, \hat{\beta}), \tag{3.4}$$

where $\Theta_5(\hat{\alpha}, \hat{\beta})$ is obtained from Table VII [McC86] by linear interpolation. The location parameter is estimated consistently by

$$\hat{\mu} = \hat{\zeta} + \hat{\beta}\hat{\nu} \tan \frac{\pi\hat{\alpha}}{2}. \tag{3.5}$$

3.2.2 Empirical Characteristic Function Method

Suppose a set of observable data $\{s_1, s_2, \cdots, s_N\}$ follows a stable distribution. Then we can approximate the characteristic function of this data by applying a basic Monte Carlo approach based on the law of large numbers i.e

$$\Phi(u) = \mathtt{E}[e^{ius_j}] \approx \hat{\Phi}(u) = \frac{1}{N}\sum_{j=1}^{N} e^{ius_j}. \tag{3.6}$$

We can express the characteristic function (2.12) in terms of the cosine and sine function from basic trigonometric principles, i.e.

$$\Phi(u) = e^{-|\nu u|^{\alpha}}(\cos\eta + i\sin\eta), \tag{3.7}$$

where

$$\eta = \nu u - |\nu u|^{\alpha}\beta sign(u)\omega(u,\alpha)$$

$$\omega(u,\alpha) = \begin{cases} \tan\frac{\pi\alpha}{2}, \alpha \neq 1 \\ \frac{2\log|u|}{\pi}, \alpha = 1 \end{cases}$$

As a result, we observe that

$$|\Phi(u)| = e^{-|\nu u|^{\alpha}}. \tag{3.8}$$

The estimated characteristic function relates to the model parameters by

$$\log|\hat{\Phi}(u_k)| = -\nu^{\alpha}|u_k|^{\alpha};\ \text{for } k = 1, 2,\ u_k > 0,\ \alpha \neq 1. \tag{3.9}$$

Solving this system leads to the estimation representation formulas for the stability and variance parameters:

$$\hat{\alpha} = \frac{\log \frac{\log|\hat{\Phi}(u_1)|}{\log|\hat{\Phi}(u_2)|}}{\log|\frac{u_1}{u_2}|}.$$

$$\log \hat{\nu} = \frac{\log|u_1|\log(-\log|\hat{\Phi}(u_2)|) - \log|u_2|\log(-\log|\hat{\Phi}(u_1)|)}{\log|\frac{u_1}{u_2}|}.$$

The real and imaginary parts of the characteristic function (3.7) provide estimates for $\hat{\beta}$ and $\hat{\mu}$:

$$\arctan\frac{Im(\Phi(u))}{Re(\Phi(u))} = \mu u - |\nu u|^{\alpha}\beta sign(u)\omega(u,\alpha). \tag{3.10}$$

Suppose $\Upsilon(u) := \arctan(Im\Phi(u)/Re\Phi(u))$ and choose another set of positive numbers $u_k,\ k = 3,4$ together with $\hat{\alpha}$ and $\hat{\nu}$ then the estimates of the location and skewness parameters are given respectively by

$$\hat{\mu} = \frac{u_4^{\hat{\alpha}}\Upsilon(u_3) - u_3^{\hat{\alpha}}\Upsilon(u_4)}{u_3 u_4^{\hat{\alpha}} - u_3 u_4^{\hat{\alpha}}}. \tag{3.11}$$

$$\hat{\beta} = \frac{u_4\Upsilon(u_3) - u_3\Upsilon(u_4)}{\hat{\nu}^{\hat{\alpha}}\tan\frac{\pi\hat{\alpha}}{2}(u_4 u_3^{\hat{\alpha}} - u_3 u_4^{\hat{\alpha}})}. \tag{3.12}$$

Notice, it can be deduced from equation (3.7) that

$$\log(-\log(|\Phi(u)|^2)) = \log(2\nu^{\alpha}) + \alpha\log(u).$$

This provides an alternative way to envision the regression estimation method:

$$y_k = m + \alpha x_k + \varepsilon_k; \quad k = 1, 2, \cdots, M;$$

where $y_k = \log(-\log|\hat{\Phi}(u_k)|^2)$, $m = \log(2\nu^{\alpha})$, $x_k = \log(u_k)$ and ε_k is an error term. The stability parameter α and the scale parameter ν can be estimated by selecting

$u_k = \frac{\pi k}{25}$, $k = 1, 2, \cdots, M$; of real data (see [Kou80], Table I). The estimates $\hat{\alpha}$ and $\hat{\nu}$ are then used to estimate β and μ using the following relation

$$z_l = \eta_l + \varsigma_l, \quad l = 1, 2, \cdots, Q.$$

where $z_l = \Upsilon_n(u_l) + \pi k_n(u_l)$, $\eta_l = \hat{\nu}_l u - |\hat{\nu}_l u|^{\hat{\alpha}} \beta sign(u) \omega(u, \hat{\alpha})$ and ς_l is some random error. The proposed real data set for Q (see [Kou80], Table II) is $u_l = \frac{\pi l}{50}$, $l = 1, 2, \cdots, Q$.

3.2.3 Logarithmic Moments Method

This approach follows the theory discussed in Section 2.4.2. The key innovation with this method is that there is no need of computing Gamma functions and the *sinc* function as in the FLOM. Secondly, techniques of parameter estimation for symmetric stable random variables (i.e. $\beta = 0$) can be applied to skewed stable random variables (i.e. $\beta \neq 0$) and, techniques of parameter estimation for centered stable random variables (i.e. $\mu = 0$) to non-centered ones (i.e. $\mu \neq 0$) through centro-symmetrization. However, this comes at a cost of losing almost half of the sample data. Therefore to obtain better estimates one has to use large sample data sets.

Centro-Symmetrization of Stable Random Data Sets

Let S_k be a sequence of n independent stable random variables distributed according to

$$S_k \sim S(\alpha, \beta, \nu, \mu).$$

Then the distribution of a weighted sum of the above sequence with weights a_k can be estimated using their characteristic function:

$$Z = \sum_{k=1}^{n} a_k S_k \sim S\left(\alpha, \frac{\sum_{k=1}^{n} a_k^{<\alpha>}}{\sum_{k=1}^{n} |a_k|^\alpha}\beta, \sum_{k=1}^{n} |a_k|^\alpha \nu, \sum_{k=1}^{n} a_k \mu\right), \tag{3.13}$$

where the p^{th} power of a number x is defined by

$$x^{<p>} = sign(x)|x|^p.$$

As a result, it is easy to obtain sequences of independent stable random variables with zero μ, zero β as well as both zero μ and zero β for $\alpha \neq 1$. This yields the centered, deskewed and symmetrized sequences:

$$S_k^C = S_{3k} + S_{3k-1} - 2S_{3k-2} \sim S(\alpha, [\frac{2-2^\alpha}{2+2^\alpha}]\beta, [2+2^\alpha]\nu, 0), \tag{3.14}$$

$$S_k^D = S_{3k} + S_{3k-1} - 2^{1/\alpha} S_{3k-2} \sim S(\alpha, 0, 4\nu, [2-2^{1/\alpha}]\mu), \tag{3.15}$$

$$S_k^S = S_{2k} - S_{2k-1} \sim S(\alpha, 0, 2\nu, 0). \tag{3.16}$$

Parameter Estimation

Suppose S_k is a data set assumed to be drawn from $S(\alpha, \beta, \nu, \mu)$. Then the exponent parameter α is estimated by setting $\theta = 0$ in (2.34), and the log moment M_2 is estimated from the obverted data (3.16). That is

$$\hat{\alpha} = \left(\frac{M_2}{\varphi_1} - \frac{1}{2}\right)^{-1/2}. \tag{3.17}$$

The estimated $\hat{\alpha}$ is used to estimate θ using (2.33) where M_1 is estimated from the obverted data (3.15). That is

$$|\hat{\theta}| = \left((\frac{\varphi_1}{2} - M_2)\hat{\alpha}^2 + \varphi_1 \right)^{1/2}. \tag{3.18}$$

From the definition of θ, $|\beta_0|$ can be estimated by

$$\hat{\beta}_0 = \frac{\tan \hat{\theta}}{\tan \frac{\hat{\alpha}\pi}{2}}. \tag{3.19}$$

Centering (see 3.14) requires $|\hat{\beta}_0|$ to be multiplied by $(2+2^{\alpha})/(2-2^{\alpha})$ to obtain $|\hat{\beta}|$ of the original data where the sign of β is determined by

$$K = sign(|S_{max} - S_{md}| - |S_{min} - S_{md}|), \quad \text{such that} \quad \hat{\beta} = K|\hat{\beta}|.$$

where S_{max}, S_{md}, S_{min} is the maximum, median and minimum of the original data.

Next we estimate the scale parameter $\hat{\nu}_0$ using (2.33) where M_1 is estimated from the obverted data (3.14). That is

$$\hat{\nu}_0 = |\cos \hat{\theta}| \exp((M_1 - \varphi_0)\hat{\alpha} + \varphi_0). \tag{3.20}$$

Again centering (see (3.14)) gives the parameter estimate $\hat{\nu}$ of the original data by $\hat{\nu} = \hat{\nu}_0(2 - 2^{1/\alpha})^{-1}$.

Finally, the location parameter μ is estimated by

$$\hat{\mu} = \hat{\mu}_0(2 - 2^{1/\alpha})^{-1}. \tag{3.21}$$

where μ_0 is the median or mean of the obverted data (3.15).

3.2.4 Maximum Likelihood Method

The maximum likelihood (ML) method is the most favored parameter estimation method in economic and financial applications. The method relies on the density function which in the case of stable distributions poses closed form representation problem. In this case we propose a numerical estimation of the density function. For a vector $s = (s_1, s_2, \cdots, s_n)$ of iid random variables assumed to follow a stable distribution, the ML estimate of the parameter vector $\Theta = (\alpha, \beta, \nu, \mu)$ is obtained by maximizing the log-likelihood function given by

$$L_\Theta(s) = \sum_{i=1}^{n} \ln \tilde{h}(s_i; \Theta), \tag{3.22}$$

where $\tilde{h}(s; \Theta)$ denotes a numerically estimated stable probability density function. It is shown for instance in [MRDC99] that the best algorithms to compute the ML is by using Fast Fourier Transforms (FFT) or by direct integration method as in [Nol01]. The ML algorithms require carefully chosen initial input parameters which in our case can be obtained for example, through the quantiles method described above. The FFT is faster for large data sets and the direct integral approach is suitable for smaller data sets since it can be evaluated at any arbitrary point.

In the following section, we analyze commodities and apply the empirical characteristic functions method to estimate the stable distribution parameters.

It is important to mention the restrictions on the parameters under which the different estimation methods operate.

3.2.5 Error Analysis

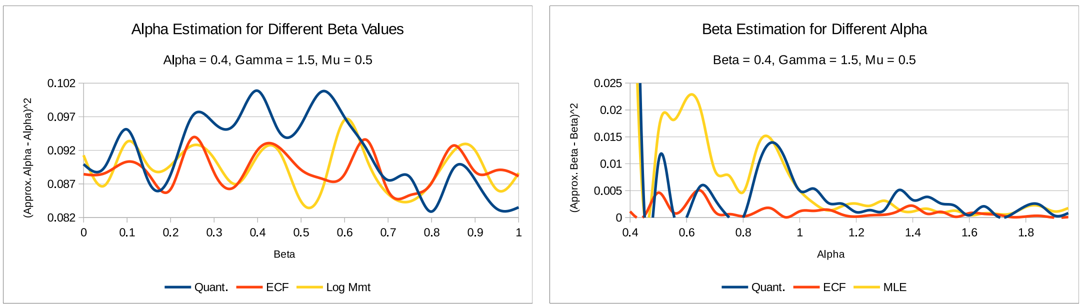
In this section we simulate datasets from the stable family of distributions based on the theory in [CMS76] and [WW95]. Then use the above four methods to retrieve the stable parameters from the simulated data. Our focus is on the α and β but the arguments extend to the other two parameters.

First, it is important to mention that all the four methods perform poorly close to the boundaries i.e. $\alpha \to 0$, $\alpha \to 2$ and $\beta \to \pm 1$. Moreover, [Inc09] shows that the methods operate efficiently under the parameter restrictions in Table 3.1.

Table 3.1: Estimation methods and their parameter restrictions

Estimation Method	Parameter Restrictions
Quantile	$\alpha \geq 0.1$
Logarithm Moments	$\beta = \mu = 0$
Maximum Likelihood	$\alpha \geq 0.4$
Empirical Characteristic Function	$\alpha \geq 0.1$

The ML is the most preferred and used estimation method. However, we observe in our analysis that this method fails for particular parameter ranges and it is not robust. For instance in estimating $0.1 < \alpha < 1.0$ with respect to β, the ML fails to converge and returns huge unrealistic errors. This is why it's not included in Figure 3.1(a).

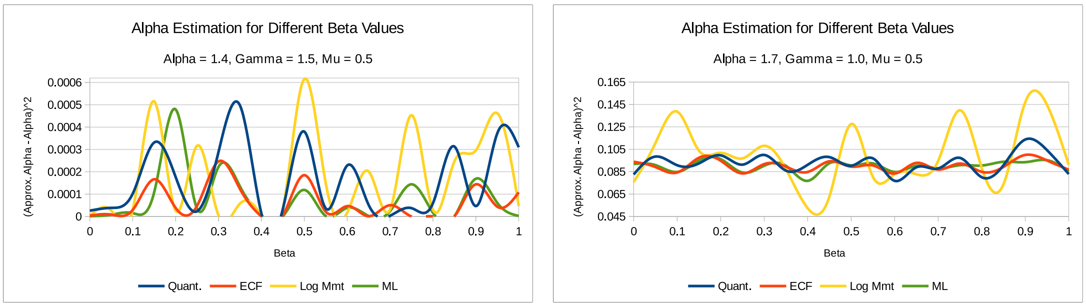


(a) $\alpha = 0.4$ estimation w.r.t β (b) $\beta = 0.4$ estimation w.r.t α

Fig. 3.1: Method comparison for $\alpha = 0.4$ and $\beta = 0.4$ estimation.

Similarly, for $\beta = 0.4$ estimation with respect to α, the logarithm moments method returns either negative or very large β values which is expected according to the constraints in Table 3.1. We omit its graph in Figure 3.1(b). Meanwhile, we notice that in both cases, the quantile and ECF methods work properly with the latter providing relatively the best estimates.

Figure 3.2 shows the error associated with estimating $1.0 < \alpha < 2.0$ for different β.

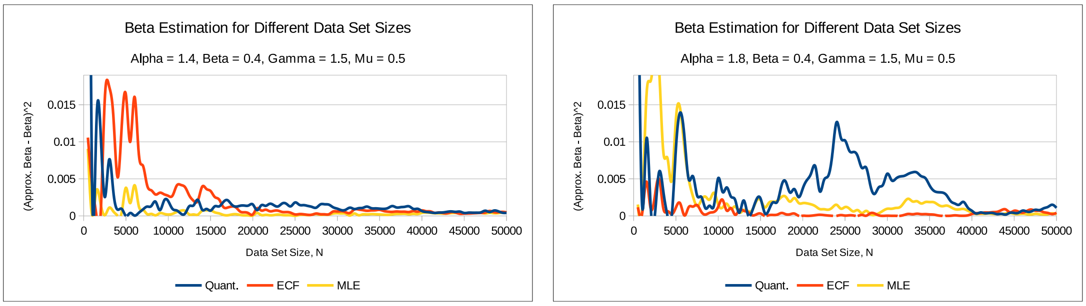


(a) $\alpha = 1.4$ estimation w.r.t β (b) $\alpha = 1.7$ estimation w.r.t β

Fig. 3.2: Method comparison for $\alpha = 1.4$ and $\alpha = 1.7$ estimation.

Note that all the four methods work properly and we still notice the ECF being relatively the most accurate and robust method. Recall that for $\alpha \to 1$ and $\alpha \to 2$ the estimation methods perform poorly. An example is 3.2(a) (for $\alpha = 1.4$) which was the closest for which the ML would converge but for higher $\alpha > 1.4$ values but far less than 2.0 (see for instance 3.2(b) for $\alpha = 1.7$) the methods performed relatively better except for the logarithm moments methods. The graphs in Figure 3.3 illustrate convergence of the quantile, ECF and the ML in estimating $\alpha = 1.4$ and $\alpha = 1.7$.

We simulated 50000 points and divided it into 100 sets starting with a 500-sized set and increasing it by 500 to 50000. The logarithm moments method performed extremely poorly and incomparable to the above three methods. It is not included in Figure 3.3(a) and Figure 3.3(b). The ECF is seen to be performing better than the quantile and ML

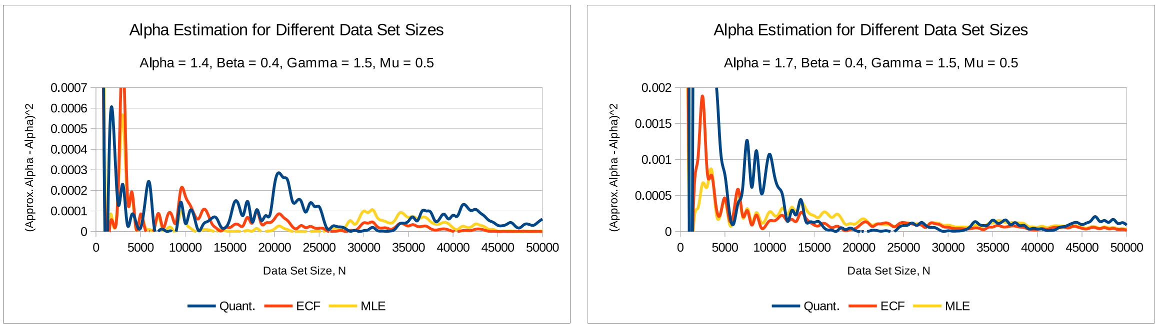


(a) $\beta = 0.4$ estimation for $\alpha = 1.4$ (b) $\beta = 0.4$ estimation for $\alpha = 1.8$

Fig. 3.3: β estimation for differing data set sizes and α values.

methods with a relatively better convergence rate.

Similarly Figure 3.4 shows the convergence rates for the quantile, ECF and ML methods. The ECF still provides a better precision in both cases i.e. 3.4(a) and 3.4(b).



(a) $\alpha = 1.4$ estimation for $\beta = 0.4$ (b) $\alpha = 1.7$ estimation for $\beta = 0.4$

Fig. 3.4: α estimation for differing data set sizes for $\beta = 0.4$ values.

In summary the empirical characteristic function method outperforms all the three other methods discussed in this chapter in the following way:

1. It is robust and can consistently estimate a wide range of α and β parameters.
2. It provides a better precision compared to the quantile, logarithm moments and ML methods for a wide range of α and β parameters.

3. It has a better convergence rate.

Therefore the quantile, logarithm moments or the maximum likelihood methods can be used to provide initial parameters for the empirical characteristics function method. Similarly, the latter can be used to provide initial parameters for future better estimators.

The following section is devoted to extracting stable parameters from log-returns commodity futures data using the empirical characteristics function method.

3.3 Commodity Data

The data sets used here are obtained from Quandl Financial and Economic Data website. The sets differ in sizes and include settled prices of Corn Futures Continuous Contract C#1 from 1959-07-01 to 2017-02-10; Crude Oil Futures Continuous Contract C#1 from 1983-03-30 to 2017-02-10; Gasoline Futures Continuous Contract C#1 from 2005-10-03 to 2017-02-10; Gold Futures Continuous Contract C#1 from 1974-12-31 to 2017-02-10; Natural Gas Futures Continuous Contract C#1 from 1990-04-03 to 2017-02-10; Platinum Futures Continuous Contract C#1 from 1969-01-02 to 2017-02-10; Silver Futures Continuous Contract C#1 from 1963-06-13 to 2017-02-10; Soybeans Futures Continuous Contract C#1 from 1959-07-01 to 2017-02-10; Wheat Futures Continuous Contract C#1 from 1959-07-01 to 2017-02-10. To avoid multi-distributional effects, we work with log-returns of the data sets.

3.3.1 The t-location Scale Distribution

The t-location scale distribution is most suited for modeling data distributions with heavier tails, more prone to outliers than the Gaussian distribution. The distribution uses the following parameters

Parameter	Description	Support
μ	Location parameter	$-\infty < \mu < \infty$
ν	Scale parameter	$\nu > 0$
α^*	Shape parameter	$\alpha^* > 0$

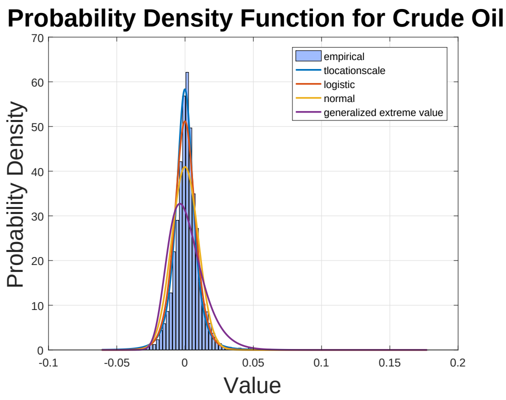
The probability density function (pdf) of the t-location scale distribution is given by

$$h(x) = \frac{\Gamma(\frac{\alpha^*+1}{2})}{\nu\sqrt{\alpha^*\pi}\Gamma(\frac{\alpha^*}{2})}\left(\frac{\alpha^* + (\frac{x-\mu}{\nu})^2}{\alpha^*}\right)^{-(\frac{\alpha^*+1}{2})},$$

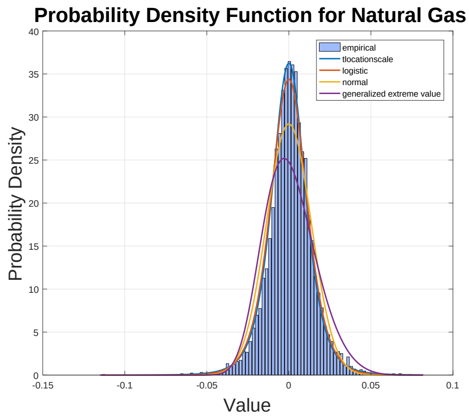
where $\Gamma(\cdot)$ denotes the gamma function. The mean of the t-location scale distribution is μ and it is defined for $\alpha^* > 1$ and undefined otherwise. The variance is given by

$$\mathtt{V}ar = \nu^2 \frac{\alpha^*}{\alpha^* - 2}.$$

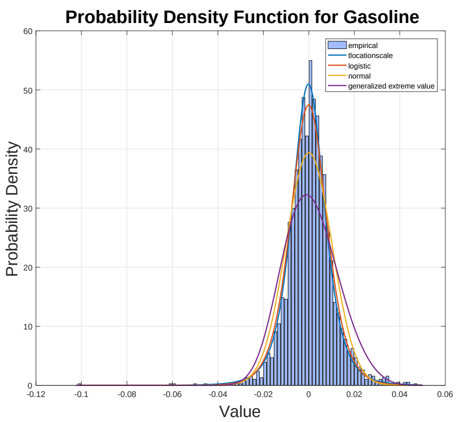
The t-location scale distribution approaches the Gaussian distribution as α^* approaches infinity and smaller values of α^* yield heavier tails. This distribution does not take skewness into consideration and its three parameters are usually estimated using the maximum likelihood estimation method. Using algorithms by [She12] on log-returns commodity futures data we obtain fittings in Figures 3.5-3.7.



(a) Crude Oil

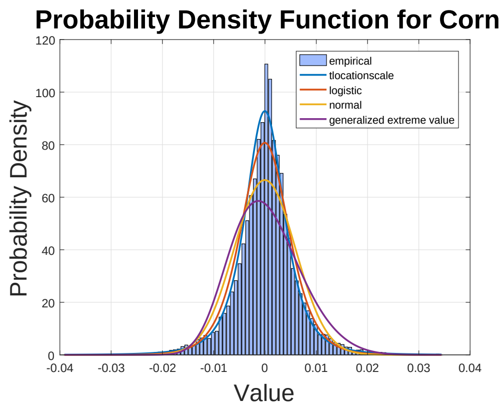


(b) Natural Gas

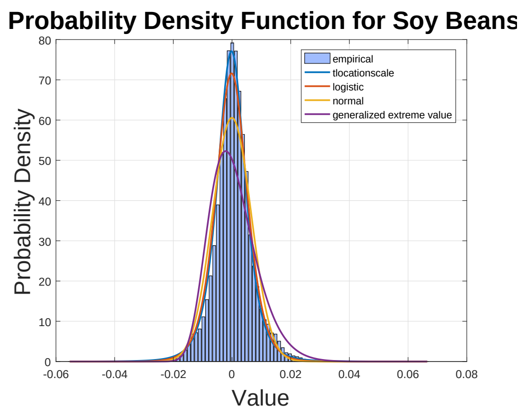


(c) Gasoline

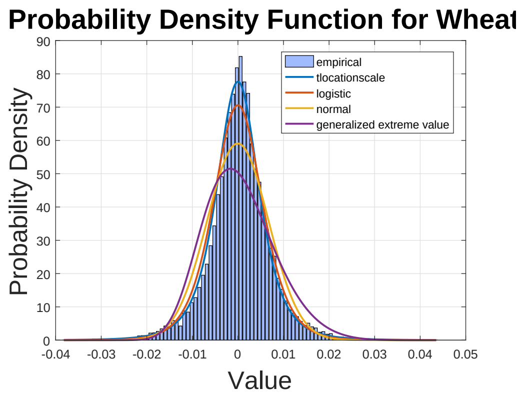
Fig. 3.5: Energy: The data exhibits high peaks and skinny tails.



(a) Corn

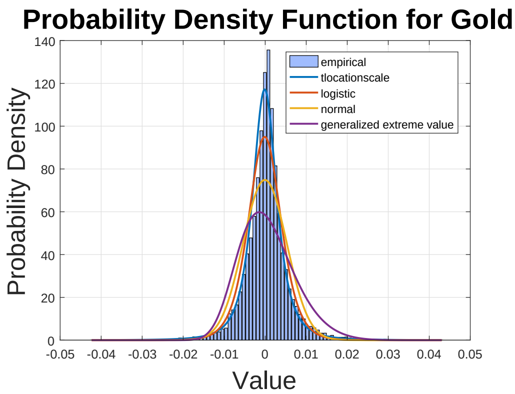


(b) Soy Beans

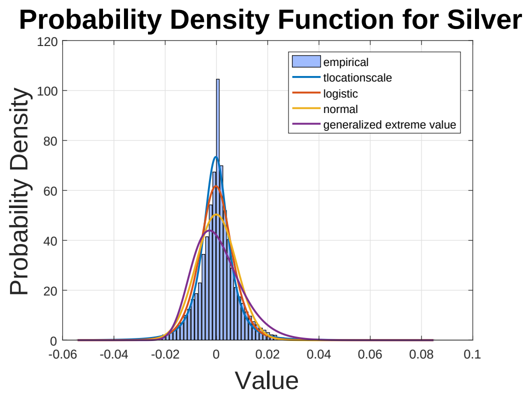


(c) Wheat

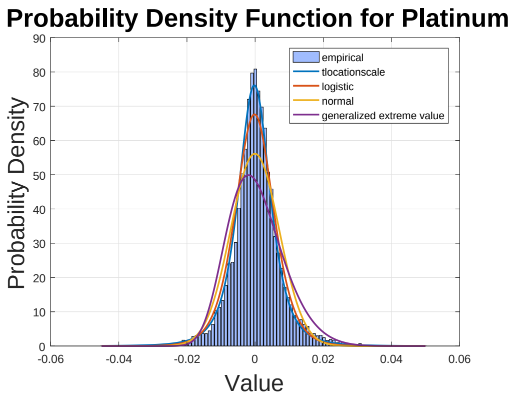
Fig. 3.6: Grains: The data exhibits high peaks and skinny tails.



(a) Gold



(b) Silver



(c) Platinum

Fig. 3.7: Metals: The data exhibits high peaks and skinny tails.

		t-location scale parameters		
		μ	ν	α^*
Energy	Gasoline	0.000905848	0.0193048	4.45922
	Natural Gas	0.000117668	0.0181944	2.50848
	Crude Oil	0.000313708	0.00912801	1.75246
Grains	Corn	$5.52294e-05$	0.00439924	3.03782
	Soy Beans	0.000400308	0.00535733	2.43218
	Wheat	$1.63112e-05$	0.0113397	3.41863
Metals	Gold	0.000905944	0.0108491	2.70553
	Platinum	0.00044219	0.0110761	3.13631
	Silver	$-1.72459e-05$	0.000682631	0.512196

Table 3.2: t- location scale distribution parameters extracted from the log-returns data.

Table 3.2 shows the parameter estimates of the t-location distribution after fitting to log-returns of energy, grains and precious metals commodities.

According to the α^* values, the log-returns data exhibit some tails. To determine the nature of the details one would require to run some QQ plots but this can also be observed directly from the Figures 3.5-3.7.

It is important to mention that QQ plots do not straight away provide conclusive evidence about the nature of the tails. More tests would still need to be made. For instance under the t-location scale it is not obvious to observe any skewness in the data. We however, view this effect when we fit the data to stable distribution as discussed in the following section.

3.3.2 Stable Distribution Fitting

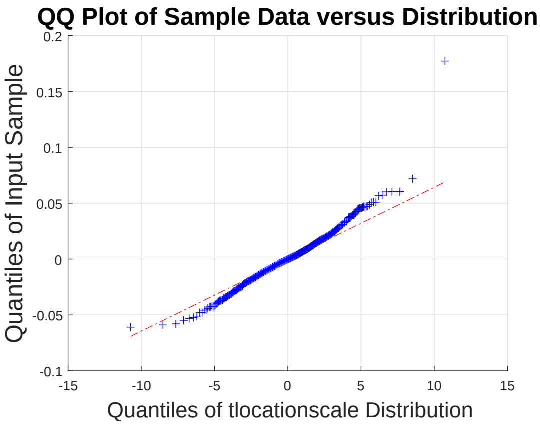
On the other hand, assuming stable distribution for log-returns commodity futures data, we employed the ECF method and obtained the stable parameters in Table 3.3.

		Stable Distribution Parameters			
		α	β	ν	μ
Energy	Gasoline	1.7504	−0.3806	0.0152	−0.0005
	Natural Gas	1.5329	0.0371	0.015	0.0005
	Crude Oil	1.2322	−0.1526	0.0075	−0.0022
Grains	Corn	1.651	0.2117	0.0036	0.0004
	Soy Beans	1.4665	−0.0968	0.0043	0.0001
	Wheat	1.638	0.0929	0.0091	0.0003
Metals	Gold	1.5007	−0.1324	0.0088	0.0001
	Platinum	1.5943	−0.1339	0.0089	−0.0001
	Silver	0.4461	0.0176	0.0011	−0.0001

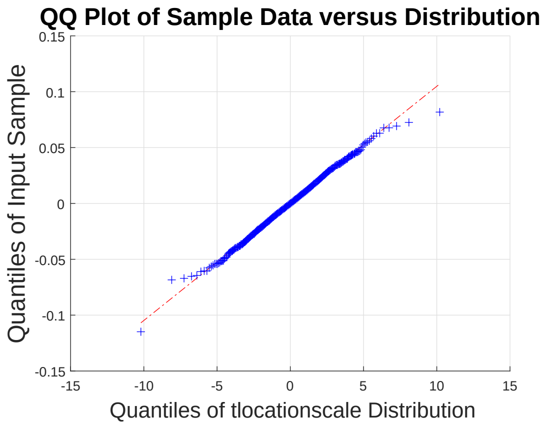
Table 3.3: Stable distribution parameters extracted from the log-returns data.

Table 3.3 shows stable distribution parameters extracted from the log-returns data using the empirical characteristic function parameter estimation method. We notice that the data exhibit a bit of skewness which is not reflected in the t-location scale distribution fitting.

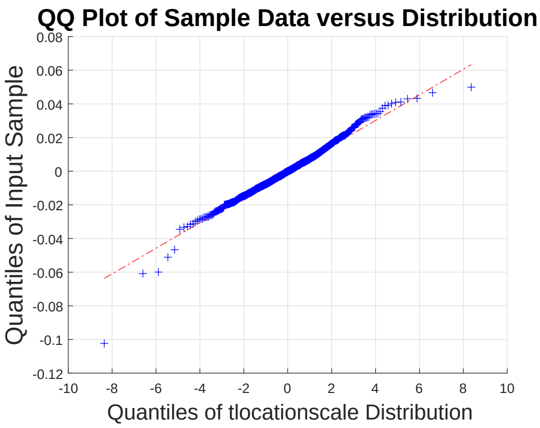
Log-returns of commodity futures are not only high peaked but they also have left and right skinny tails with extreme outliers as observed from the QQ-plots for energy commodities (i.e. Crude oil, Natural gas and Gasoline) in Figure 3.8, the grains commodities in Figure 3.9 and the precious metals in Figure 3.10. The comparative distribution in the QQ-plots is the t-location distribution with location parameter $\mu = 0$, scale parameter $\nu = 1$ and shape parameter $\alpha^* = 5$. Note that other distributions such as Weibull, Gaussian and Extreme value, etc. can also be compared with the data but the t-location turns out to be the closest fit.



(a) Crude Oil vs t-location distribution

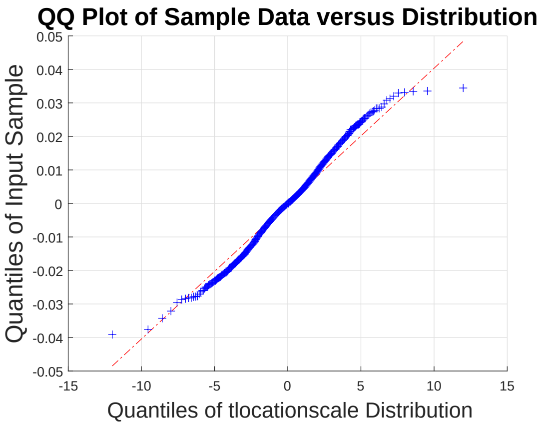


(b) Natural Gas vs t-location distribution

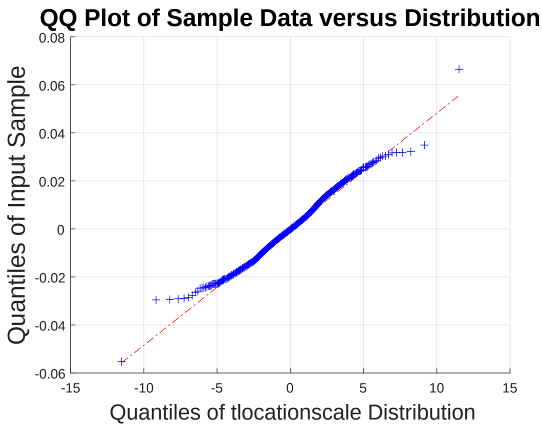


(c) Gasoline vs t-location distribution

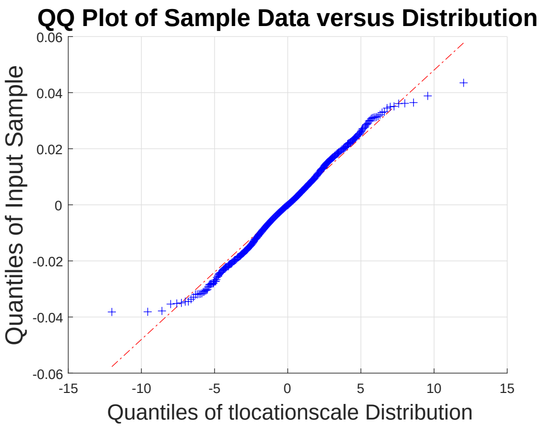
Fig. 3.8: Energy commodities data samples compared with t-location distribution.



(a) Corn vs t-location distribution

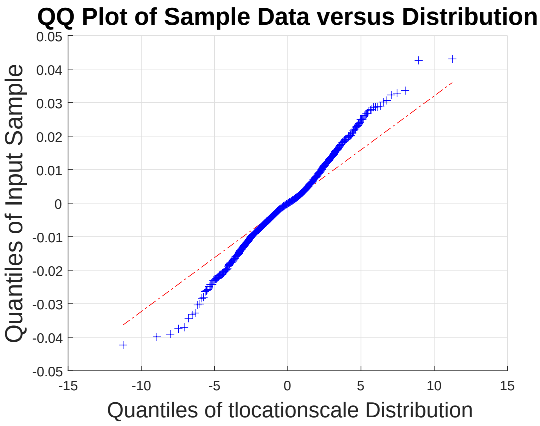


(b) Soy Beans vs t-location distribution

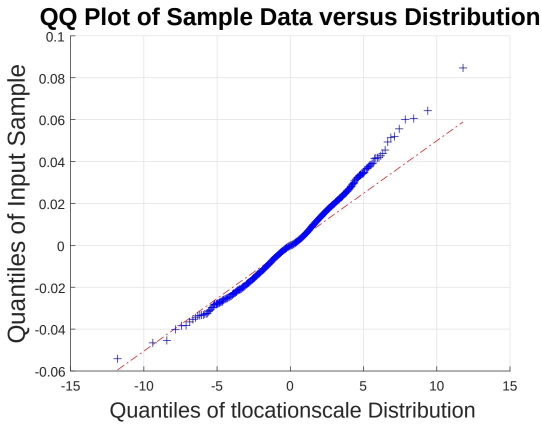


(c) Wheat vs t-location distribution

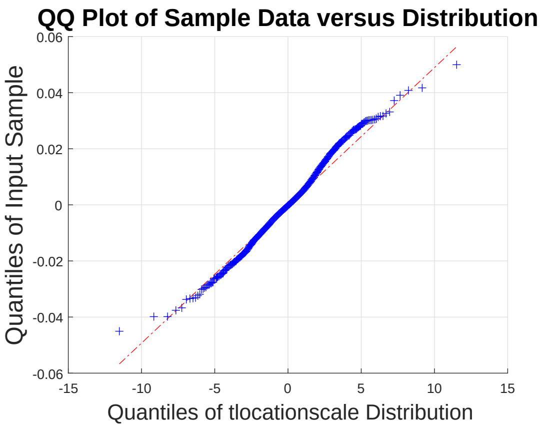
Fig. 3.9: Grain commodities data samples compared with t-location distribution.



(a) Gold vs t-location distribution



(b) Silver vs t-location distribution



(c) Platinum vs t-location distribution

Fig. 3.10: Precious metal commodities data samples compared with t-location distribution.

3.4 Summary

First we showed that the ECF provides the best precision in estimating a wide range of α and β parameters, it is robust and provides better convergence compared to the quantile, ML and the logarithm moments. Secondly, we have illustrated that in general, the distribution of the commodity futures log-returns data is closest to a t-location scale distribution due to its high peaks, skinny tails and extreme outliers. Moreover, by using the ECF estimation method we realise some minor skewness effects not captured in the t-location scale fitting. We recommend the ECF as a suitable approach for estimating parameters of any skewed financial market data and could be used to obtain initial input parameters for future and better estimation techniques.

Chapter 4

Subordinated Affine-Structure Models for Commodity Futures Prices

4.1 Introduction

Commodities exhibit distinctive features a good model should capture. The features include mean-reversion, contango, backwardation and seasonality (see [MZ15]). They also experience extreme volatility and price spikes resulting in heavy-tailed distribution of the returns. Commodity markets are unique compared to other markets such as equity, bond, currency or interest rate markets in the sense that most commodities are real physical assets that are produced, transported, stored and consumed. They are not assets valued on long-lived companies like in equity markets.

Usually, most companies and organizations are in need of large quantities of a particular commodity at a particular point in time implying increased mobility, delivery and storage costs across the world. However, in reality, the liquidation of a contract to deliver a required commodity is not that simple in real time due to a number of factors including seasonality. This is the main reason why spot price data for some commodities is not available and only the futures is. Moreover, futures prices are usually listed on exchanges which means their prices are observable. However, it is important to note that the underlying spot price is crucial for certain market activities such as investment valuation and pricing claims contingent on commodities among other reasons. Usually the futures price closest to maturity is used as a proxy for the spot price, (see [Ken10, Sch97] for a detailed rationale for this proxy). The unobserved spot is usually optimally estimated through recursive estimation procedures that rely on the observed futures prices. One such methods is filtering where it's a common practice to represent a model in a state-space form once there are unobservabed factors involved. See [Aok13] for a good reference on state-space representation. The representation aims at staging the problem appropriate enough for estimation of the proposal distribution of the unoberseved factor. Depending on the linearity and Guassianity of the model, different filtering techniques discussed in [DK12, CK07, DHKS04] can be used as well. For instance particle filtering is recommended for non-linear and non-Gaussian models. In contrast, the current chapter, employs latent regression models usually applied in population dynamics to obtain the futures prices where the key tool used is the expectation maximization.

As indicated in [FF88], commodity pricing can be approached from two perspectives, the theory of storage which explains why high supplies and inventories running at minimum would result into contango, low futures and spot price volatilities, and in turn futures premiums being equivalent to full storage costs. On the other hand, why low supplies and enhanced production inventory levels yield to backwardation, a rise in volatilities of the spot and the nearby futures prices. Another feature explained by this theory is the periodically continuously compounded convenience yield (usually denoted by δ) on inventory which is the benefit of holding a physical commodity as opposed to having a futures contract of its delivery at some future time and secondly, the cost of storage. The futures price motivated by the theory of storage is given by

$$F_{t,T} = S_t e^{(r-c-\delta)(T-t)}, \tag{4.1}$$

where c accounts for the storage costs, r is the periodically continuously compounded interest rate, S_t is the current spot price and T is the maturity date of the future contract.

The second perspective is the theory of expected risk premium discussed in [Key30] and [Hic39]. It asserts that the futures prices are given by the discounted (by the risk premium) expected future spot price:

$$F_{t,T} = \mathtt{E}_t[S_T]e^{-r\gamma[T-t]},$$

where γ is the risk premium and $\mathtt{E}_t[\cdot] = \mathtt{E}[\cdot|\mathcal{F}_t]$, $\mathcal{F}_t$ is the filtration up to time t.

Examples of models based on the latter include Schwartz's common continuous stochastic factor models [Sch97], [SS00] and the jump models of [KNPP15]. We follow a similar perspective in this chapter and introduce a new pricing approach.

The motivation and contribution of this chapter is based on the existing erratic features in electricity and energy markets where jumps are evident resulting in skewed distributions of the spot prices. We consider a subordinated Brownian motion by an α-stable process, $\alpha \in (0, 1)$, as the source of randomness in the underlying asset to model commodity future prices. The stunning feature in our pricing approach is the new simple technique derived from our novel approach for subordinated affine structure models.

We show that the affine property is attainable and applicable to generalised commodity spot models and as an illustration we consider a stochastic differential equation with subordinated Brownian motion as the source of randomness to derive the commodity futures price. It is argued in some existing literature that the likelihood function exists in integrated form for models with singular noise meanwhile for cases of partially observed processes a filtering technique is required (See for instance [DP10], [YLLN14]). However, the work presented in this chapter provides a new approach of pricing commodity futures for models with latent variables using expectation maximization. We show that the commodity future price under a one factor model with a subordinated random source driver can be expressed in terms of the subordinator which can then be reduced to the latent regression models commonly used in population dynamics with their parameters easily estimated using the expectation maximization method. In our case the underlying joint probability distribution is a combination of Gaussian and stable densities.

The rest of this chapter is organized as follows. Section 4.2 reviews the concept of affine models and extend the idea of obtaining Laplace transforms of random processes to subordinated processes. In Section 4.3 we derive our pricing formulas for commodity futures using the results derived in Section 4.2. In Section 4.4 we discuss the numerical implementation of our one factor commodity futures model. Section 4.5 summarizes.

4.2 Affine Models

In this section we provide an overview on affine processes and provide some crucial results in Theorem 2 and Theorem 3. We retain some definitions and notations used in [KR08, KRST11] and [DFS03]:

1. $\mathcal{D} := \mathbb{R}^m_{\geq 0} \times \mathbb{R}^n$.
2. $\mathcal{U} := \left\{u \in \mathbb{C}^d : Re\ u_I \leq 0, Re\ u_J = 0\right\}$,
   where $I := \{1, \cdots, m\}$, $J := \{m+1, \cdots, m+n\}$ and $M := I \cup J = \{1, \cdots, d\}$.
3. $\langle x, y \rangle := \sum_{i=1}^{d} x_i y_i = x \cdot y$.
4. $f_u(x) := \exp(u \cdot x)$ where $u \in \mathbb{C}^d$ and $x \in \mathcal{D}$. From point 2., $f_u(x)$ is bounded.
5. $P_t f(x) = \mathtt{E}^x[f(X_t)]$ for all $x \in \mathcal{D}, t \geq 0$ where P_t is a semigroup operator.

6. $\mathcal{O} := \{(t,u) \in \mathbb{R}_{\geq 0} \times \mathcal{U} : P_s f_u(0) \neq 0 \ \forall s \in [0,t]\}$.

7. $\mathcal{X}$ will denote a closed state space.

8. $\mathtt{P}^x$ will denote the law of a Markov process $(X_t)_{t\geq 0}$ started at $X_0 = x$.

We state the following essential definitions from [KR08].

Definition 10 *A process is stochastically continuous if for any sequence $t_n \to t$ in $\mathbb{R}_{\geq 0}$ the random variables X_{t_n} converge to X_t in probability with respect to $(\mathtt{P}^x)_{x\in\mathcal{D}}$.*

Definition 11 *An affine process is a stochastically continuous time-homogeneous Markov process $(X_t, \mathtt{P}^x)_{t\geq 0, x\in\mathcal{D}}$ whose characteristic function is an exponentially affine function of the state vector such that*

$$\mathtt{E}^x[e^{u\cdot X_t}] = \exp(\psi_0(t,u) + \psi_1(t,u)x),$$

for all $x \in \mathcal{D}$ and for all $(t,u) \in \mathbb{R}_{\geq 0} \times i\mathbb{R}^d$, where $i\mathbb{R}^d$ is a space of purely imaginary numbers in $\mathbb{C}^d$.

The space $\mathcal{O}$ is introduced to cater for points in $\mathcal{U}$ where $P_s f_u(x)$ is 0 with an undefined logarithm. Definition 11 can be extended to $\mathcal{O}$ satisfying the following properties [Prop. 1.3, [KR08]]:

1. ψ_0 maps $\mathcal{O}$ to $\mathtt{C}_-$ where $\mathtt{C}_- := \{u \in \mathbb{C} : \mathit{Re}\ u \leq 0\}$.

2. ψ_1 maps $\mathcal{O}$ to $\mathcal{U}$.

3. $\psi_0(0,u) = 0$ and $\psi_1(0,u) = u$ for all $u \in \mathcal{U}$.

4. ψ_0 and ψ_1 admit the 'semi-flow property':

- $\psi_0(t+s,u) = \psi_0(t,u) + \psi_0(s,\psi_1(t,u))$,
- $\psi_1(t+s,u) = \psi_1(s,\psi_1(t,u))$, for all $t,s \geq 0$ with $(t+s,u) \in \mathcal{O}$.

5. ψ_0 and ψ_1 are jointly continuous on $\mathcal{O}$.
6. With the remaining arguments fixed, $u_I \mapsto \psi_0(t,u)$ and $u_I \mapsto \psi_1(t,u)$ are analytic functions in $\{u_I : Re\ u_I < 0; (t,u) \in \mathcal{O}\}$.
7. Let $(t,u),\ (t,w) \in \mathcal{O}$ with $Re\ u \leq Re\ w$. Then
   - $Re\ \psi_0(t,u) \leq \psi_1(t, Re\ w)$,
   - $Re\ \psi_1(t,u) \leq \psi_1(t, Re\ w)$.

Definition 11 is also known as the affine property and it implies that the PDE

$$\frac{\partial}{\partial t}\mathtt{E}^x[e^{u\cdot X_t}] = \mathcal{A}\mathtt{E}^x[e^{u\cdot X_t}], \quad \mathtt{E}^x[e^{u\cdot X_0}] = \exp(u\cdot x),$$

where $\mathcal{A}$ denotes the infinitesimal generator of X, can be reduced to a system of non-linear ODEs known as generalized Riccati differential equations.

Lemma 9 *An affine process $(X_t)_{t\geq 0}$ is regular if the following right derivatives exist for all $u \in \mathcal{U}$ and are continuous at $u = 0$:*

$$F^{(0)}(u) := \frac{\partial \psi_0}{\partial t}(t,u)\bigg|_{t=0^+}, \quad F^{(1)}(u) := \frac{\partial \psi_1}{\partial t}(t,u)\bigg|_{t=0^+}$$

The regularity condition can be extended to $\mathcal{O}$ for which case the following Riccati equations hold:

$$\begin{aligned}\frac{\partial \psi_0}{\partial t}(t,u) &= F^{(0)}(\psi_1(t,u)), \quad \psi_0(0,u) = 0,\\ \frac{\partial \psi_1}{\partial t}(t,u) &= F^{(1)}(\psi_1(t,u)), \quad \psi_1(0,u) = u.\end{aligned}$$

Proof 7 *See [KR08, KRST11, RH15].*

We are interested in the affine property of the solution to the SDE:

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_{t-})\,\mathrm{d}\mathcal{M}_t, \tag{4.2}$$

where $b : \mathcal{X} \to \mathbb{R}^d$ is continuous, $\sigma : \mathcal{X} \to \mathbb{R}^{d\times d}$ is measurable such that the diffusion matrix $\sigma(x)\sigma^T$ is continuous and $\mathcal{M}_t$ is a d-dimensional standard Lévy process. We also require that b and σ are Lipschitz with linear growth and bounded derivatives to ensure a strong solution X_t.

The following theorem is one of the contributions in this chapter.

Theorem 2 *Suppose X_t is a regular affine solution to (4.2). Then b and σ can be expressed as:*

$$\begin{aligned} b(x) &= K_0 + K_1 x_1 + \cdots + K_d x_d, \quad K_i \in \mathbb{R}^d \\ \sigma(x)\sigma(x)^T &= H_0 + H_1 x_1 + \cdots + H_d x_d, \quad H_i \in \mathbb{R}^d \times \mathbb{R}^d, \end{aligned}$$

where $i = 0, \cdots d$. Moreover, the characteristic function of X_t has a log-linear form

$$\begin{aligned} &\mathtt{E}[e^{iu_1 X_t^{(1)} + \cdots + iu_d X_t^{(d)}}] \\ &= \exp\left(\psi_0(t, u_1, \cdots, u_d) + \psi_1(t, u_1, \cdots, u_d) x_0^{(1)} + \cdots + \psi_d(t, u_1, \cdots, u_d) x_0^{(d)}\right), \end{aligned}$$

where $u_i \in i\mathbb{R}$. The coefficients ψ_i satisfy the system of Riccati equations:

$$F^{(i)}(t, \psi_1, \cdots, \psi_d) = \frac{\partial \psi_i}{\partial t} = K_i^T \eta + \frac{1}{2}\eta^T H_i \eta, \quad i = \{0, \cdots, d\},$$

where $\eta^T = (\psi_1 \cdots \psi_d)$, subject to conditions $\psi_i(0, u_1, \cdots, u_d) = iu_i$ for $i = 1, \cdots, d$ and

$\psi_0(0, u_1, \cdots, u_d) = 0.$

Proof 8 *The proof is a generalisation of the 2-dimensional case in [RH15].*

There is extensive literature on affine processes X_t where $\mathcal{M} := B_t$ or $\mathcal{M} := B_t + \sigma^{-1}\int_0^t \xi_s \, \mathrm{d}s$ with ξ_t, a Poisson jump process. We are interested in the solution to

$$\mathrm{d}Y_t = b(Y_t)\,\mathrm{d}S_t + \sigma(Y_{t-})\,\mathrm{d}B_{S_t}, \tag{4.3}$$

where $\{S_t\}_{t\geq 0}$ is a subordinator. Another contribution in this chapter follows in this following theorem. We show that $Y_t = X_{S_t}$ is affine in the following theorem with $d = 1$.

In the following, we shall recall the definition of the joint filtration $(\mathcal{F}_t)_{t\geq 0}$ introduced in Section 2.7.

Theorem 3 *Let $(\Omega, \mathcal{F}, \mathtt{P}^x, (\mathcal{F}_t)_{t\geq 0})$ denote a joint probability space for $(S_t)_{t\geq 0}$, a non-decreasing affine process taking values in $\mathcal{D}$ and $(X_t)_{t\geq 0}, X_0 = x$, an independent Lévy process. Define a process $Y_t := X_{S_t}, Y_0 = y$ with Lévy exponent $m(w)$ and suppose $(S_t)_{t\geq 0}$ is regular with functional characteristics $F^{(0)}(u), F^{(1)}(u)$. Then $(Y_t)_{t\geq 0}$ is regular affine with functional characteristics $F^{(0)}(m(w)), F_X^{(1)}(m(w))$ and $F_Y^{(1)}(m(w)) = 0$, $u, w \in i\mathbb{R}$ with the characteristic function given by*

$$\mathtt{E}^s[e^{iuY_\tau}] = \exp(\psi_0(\tau - t, m(w)) + \psi_1(\tau - t, m(w))S_t),$$

for some regular functions ψ_0 and ψ_1 and $0 \leq t < \tau$.

Proof 9 *The Markov property of S_t and the definition of its Laplace transform yields*

$$\begin{aligned}\mathtt{E}^s[e^{wY_\tau}|\mathcal{F}_t] = \mathtt{E}^s[e^{wX_{S_\tau}}] &= \mathtt{E}^s\left[\mathtt{E}^s[\exp(wX_{S_\tau})|\sigma(S_s)_{0\le s\le t}]\right]. \\ &= \mathtt{E}^s[\exp(m(w)S_t)]. \\ &= \exp(\psi_0(\tau - t, m(w)) + \psi_1(\tau - t, m(w))S_t).\end{aligned}$$

The last equality follows from the affine property of S_t (see Definition 11).

4.3 Commodity Future Pricing

4.3.1 Introduction

We develop representation formulas for futures prices using the concepts introduced before. The source of randomness in the models developed in this section is Brownian motion subordinated by a non-decreasing α-stable process where $\alpha \in (0, 1)$. The aim is to obtain futures price formulas for commodity spot price models that incorporate stochastic volatility, jumps, seasonality and mean-reversion effects.

4.3.2 One Factor Commodity Spot Model

We consider a one factor commodity spot price model given by

$$z_t = f(t) + e^{Y_t}, \tag{4.4}$$

where Y_t satisfies (4.3) and seasonality is defined according to [KNPP15], as

$$f(t) = \delta_0 t + \delta_1 \sin(\delta_2[t + 2\pi/264]) + \delta_3 \sin(\delta_4[t + 4\pi/264]), \tag{4.5}$$

where $\delta_0, \delta_1, \delta_2, \delta_3, \delta_4$ account for deterministic regularities in the spot price dynamics.

The following theorem presents the first main contribution of this chapter. Specifically, the futures price is given as an exponential function of the subordinator, the source of the random jumps in the spot.

Theorem 4 *Suppose without seasonality (i.e. $f = 0$), the commodity spot price z given by (4.4) satisfies the following stochastic differential equation*

$$\mathrm{d}z_t = \kappa(\theta - \ln z_t)z_t \,\mathrm{d}S_t + \sigma z_t \,\mathrm{d}B_{S_t}, \tag{4.6}$$

where S_t is an independent, non-decreasing stable process with $\alpha \in (0,1)$. Then the future price is

$$\begin{aligned} &F(t,\tau) \\ &= \exp\Bigg([(-\frac{1}{2}\sigma^2+\beta)^\alpha+\beta^\alpha]\gamma(1-e^{-\kappa(\tau-t)}) + \frac{1}{4\kappa}\sigma^2[(-\frac{1}{2}\sigma^2+\beta)^\alpha+\beta^\alpha]^2(1-e^{-2\kappa(\tau-t)}) \\ &\qquad + [(-\frac{1}{2}\sigma^2+\beta)^\alpha+\beta^\alpha]S_t e^{-\kappa(\tau-t)}\Bigg), \end{aligned} \tag{4.7}$$

where $\gamma := \theta - \frac{\sigma^2}{2\kappa}$, $\beta \in [0,1]$ denotes the skewness parameter of the subordinator S_t.

Proof 10 *By applying Itô's formula to $Y_t = \ln z_t$, it is readily seen that*

$$\mathrm{d}Y_t = \kappa(\gamma - Y_t)\,\mathrm{d}S_t + \sigma \,\mathrm{d}B_{S_t}, \tag{4.8}$$

where $\gamma := \theta - \frac{\sigma^2}{2\kappa}$ and the future price with maturity date τ is given by (see [Sch97])

$$F(t,\tau) = \mathtt{E}[z_\tau] = \mathtt{E}[e^{Y_\tau}]. \tag{4.9}$$

Theorem 2 suggests an explicit representation of (4.9) is attainable and it can be deduced by considering first, the continuous case $\mathtt{E}[e^{X_\tau}]$. *Suppose a continuous mean-reverting model given by*

$$\mathrm{d}X_t = \kappa(\gamma - X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t,\ \ X_0 = x. \tag{4.10}$$

The corresponding affine forms of the coefficients according to Theorem 2 yield:

$$K_0 = \kappa\gamma,\ \ K_1 = -\kappa,\ \ K_2 = 0,\ \ H_0 = H_1 = 0,\ \ H_2 = \sigma.$$

Since X_t *is regular affine and* σ *is constant for all* $t \in [0, \tau]$, *then we have*

$$\mathtt{E}[e^{iuX_\tau}] = \exp(\psi_0(\tau - t, u) + \psi_1(\tau - t, u)x + \psi_2(\tau - t, u)\sigma),\ \ u \in i\mathbb{R}, \tag{4.11}$$

where $\psi_0(\tau - t, u)$, $\psi_1(\tau - t, u)$ *and* $\psi_2(\tau - t, u)$ *satisfy the set of Riccati equations:*

$$\begin{aligned}
\frac{\partial \psi_0}{\partial \tau} &= K_0^T\eta + \frac{1}{2}\eta^T H_0 \eta = \kappa\gamma\psi_1; \quad \psi_0(0, u) = 0.\\
\frac{\partial \psi_1}{\partial \tau} &= K_1^T\eta + \frac{1}{2}\eta^T H_1 \eta = -\kappa\psi_1; \quad \psi_1(0, u) = iu.\\
\frac{\partial \psi_2}{\partial \tau} &= K_2^T\eta + \frac{1}{2}\eta^T H_2 \eta = \frac{1}{2}\sigma\psi_1^2; \quad \psi_2(0, u) = 0,
\end{aligned}$$

where $\eta^T = (\psi_1\ \psi_2)$. *The solution set to the system of Riccati equations is given by*

$$\psi_1(\tau - t, u) = iue^{-\kappa(\tau - t)}, \tag{4.12}$$

$$\psi_0(\tau - t, u) = iu\gamma(1 - e^{-\kappa(\tau - t)}). \tag{4.13}$$

$$\psi_2(\tau - t, u) = \frac{\sigma u^2}{4\kappa}(1 - e^{-2\kappa(\tau - t)}). \tag{4.14}$$

Using (4.11) where $u = -i$, *one can easily deduce* $\mathtt{E}[e^{X_\tau}]$ *leading to the price of a one* *factor commodity futures price under a continuous model framework:*

$$F(t,\tau) = \mathtt{E}[e^{X_\tau}] = \exp(\gamma(1-e^{-\kappa(\tau-t)}) + X_t e^{-\kappa(\tau-t)} + \frac{\sigma^2}{4\kappa}(1-e^{-2\kappa(\tau-t)})).$$

Capitalizing on the affine nature of Y_t and Theorem 3 we have the formula for $\mathtt{E}[e^{Y_\tau}]$ as:

$$\mathtt{E}[e^{iuY_\tau}] = \exp(\psi_0(\tau-t,m(u)) + \psi_1(\tau-t,m(u))S_t + \psi_2(\tau-t,m(u))\sigma),$$

where the volatility σ is a constant and the system of Riccati equations takes the form

$$\begin{aligned}
\frac{\partial\psi_0}{\partial\tau}(\tau-t,m(u)) &= \kappa\gamma\psi_1; \quad \psi_0(0,m(u)) = 0.\\
\frac{\partial\psi_1}{\partial\tau}(\tau-t,m(u)) &= -\kappa\psi_1; \quad \psi_1(0,m(u)) = m(iu).\\
\frac{\partial\psi_2}{\partial\tau}(\tau-t,m(u)) &= \frac{1}{2}\sigma\psi_1^2; \quad \psi_2(0,m(u)) = 0.
\end{aligned}$$

Consequently, the solution set is directly deduced from (4.12) - (4.14) to obtain

$$\begin{aligned}
\psi_1(\tau-t,m(u)) &= m(iu)e^{-\kappa(\tau-t)},\\
\psi_0(\tau-t,m(u)) &= m(iu)\gamma(1-e^{-\kappa(\tau-t)}),\\
\psi_2(\tau-t,m(u)) &= \frac{1}{4\kappa}\sigma m(iu)^2(1-e^{-2\kappa(\tau-t)}).
\end{aligned}$$

Setting $u := -i$ yields

$$\mathtt{E}[e^{Y_\tau}] = \exp\left(m(1)\gamma(1-e^{-\kappa(\tau-t)}) + \frac{1}{4\kappa}\sigma^2 m(1)^2(1-e^{-2\kappa(\tau-t)}) + m(1)S_t e^{-\kappa(\tau-t)}\right).$$

The required result follows by substituting the Lévy exponent $m(1)$ from (2.56).

4.3.3 The Two Factor Commodity Spot Model

In the two factor spot model, the volatility is modeled as a stochastic process while retaining jumps in the spot model. The futures price model is given by

$$F(t, X_t, V_t) = f(t) + \mathtt{E}[\exp(X_\tau + V_\tau)].$$

We present the second main contribution of the chapter in the following theorem.

Theorem 5 *Suppose without seasonality (i.e $f = 0$), the commodity spot price X satisfies the set of subordinated stochastic differential equation*

$$\mathrm{d}z_t = \kappa(\theta - z_t)\,\mathrm{d}S_t + \sqrt{V_t}\,\mathrm{d}B^{(1)}_{S_t} \tag{4.15}$$

$$\mathrm{d}V_t = \lambda(\varepsilon - V_t)\,\mathrm{d}t + \upsilon\sqrt{V_t}\,\mathrm{d}B^{(2)}_t, \tag{4.16}$$

such that $\mathrm{d}\,[B^{(1)}_{S_\cdot}, B^{(2)}_\cdot]_t = \rho\,\mathrm{d}A_t$ *where* $A_t = g(t, S_t)$ *some random process. The futures price is:*

$$F = \exp(\psi_0 + \psi_1 z_0 + \psi_2 V_0),$$

where the coefficients ψ_0, ψ_1 and ψ_2 are given as

$$\psi_1(\tau - t, u_1, u_2) = m(iu_1)e^{-\kappa(\tau-t)}.$$

$$\psi_2(\tau - t, u_1, u_2) = -\frac{2\kappa m(iu_1)}{\upsilon^2}e^{-\kappa(\tau-t)}\sum_{j=1}^{\infty} d_j m(iu_1)^j e^{-j\kappa(\tau-t)} + \frac{I_f(\tau-t,u_1)}{C(u_1,u_2) - \frac{1}{2}\upsilon^2 \int_t^\tau I_f(s-t,u_1)\,\mathrm{d}s}.$$

$$\psi_0(\tau - t, u_1, u_2) = \theta m(iu_1)(1 - e^{-\kappa(\tau-t)}) + \frac{-2\kappa\lambda\varepsilon m(iu_1)}{\upsilon^2}\sum_{j=1}^{\infty} d_j m(iu_1)^j \left(\frac{1}{\kappa(1+j)}(1 - e^{-\kappa(\tau-t)(1+j)})\right). + \int_t^\tau \frac{I_f(s-t,u_1)}{C(u_1,u_2) - \frac{1}{2}\upsilon^2 \int_t^s I_f(q-t,u_1)\,\mathrm{d}q}\,\mathrm{d}s,$$

where coefficients $\{d_j\}_{j=1}^{\infty}$ satisfy

$$d_{j+1} = \frac{\sum_{i=1}^{j-1} d_j d_{j-1}\mathbb{1}_{j>1} - \frac{\rho\upsilon}{\kappa}d_j\mathbb{1}_{j>0} + \frac{\upsilon^2}{4\kappa^2}\mathbb{1}_{j=0}}{(j+1-\frac{\kappa-\lambda}{\kappa})}.$$

Note that the values of $u_i \in \mathbb{C},\ i = \{1,2\}$ are chosen carefully to ensure that the futures price is positive real. In this case $u_i = -i$.

The factor $C(u_1, u_2)$ is defined as

$$C(u_1,u_2) = \frac{\exp\left(-\frac{\rho\upsilon m(iu_1)}{\kappa} + 2\sum_{j=1}^{\infty}\frac{d_i}{1+j}m(iu_2)^{j+1}\right)}{m(iu_2) + \frac{2\kappa m(iu_1)}{\upsilon^2}\sum_{j=1}^{\infty} d_j m(iu_1)^j}.$$

Lastly, the integrating factor I_f is such that

$$I_f(\tau - t, u_1) = \exp\left(- \lambda(\tau - t) - \frac{\rho \upsilon m(iu_1)}{\kappa} e^{-\kappa(\tau - t)} + 2e^{-\kappa(\tau - t)} \sum_{j=1}^{\infty} \frac{d_j}{j+1} m(iu_1)^{j+1} \right).$$

$$\int_0^\tau I_f(s - t, u_1)\, \mathrm{d}s = \frac{I_f(\tau - t, u_i)}{(-\lambda + \rho \upsilon m(iu_1) e^{-\kappa(\tau - t)} - 2e^{-\kappa(\tau - t)} \sum_{j=1}^{\infty} \frac{d_j}{j+1} m(iu_1)^{j+1}} - \frac{\exp\left(\frac{-\rho \upsilon m(iu_1)}{\kappa} + 2 \sum_{j=1}^{\infty} \frac{d_j}{j+1} m(iu_1)^{j+1} \right)}{(-\lambda + \rho \upsilon m(iu_1) - 2 \sum_{j=1}^{\infty} \frac{d_j}{j+1} m(iu_1)^{j+1})}.$$

We provide the proof using the following proposition and subsequent lemmas.

Proposition 1 *The mean and variance of the model (4.15) - (4.16) are given by*

$$\mu := \begin{pmatrix} \kappa(\theta - X_t) \\ \lambda(\varepsilon - V_t) \end{pmatrix}, \quad \sigma := \begin{pmatrix} \sqrt{V_t} & 0 \\ \rho \upsilon \sqrt{V_t} & \upsilon \sqrt{1 - \rho^2} \sqrt{V_t} \end{pmatrix}, \quad H := \sigma \sigma^T = \begin{pmatrix} V_t & \rho \upsilon V_t \\ \rho \upsilon V_t & \upsilon^2 V_t \end{pmatrix}.$$

Moreover, their affine forms can be given as linear models of both X and V:

$$\mu = K_0 + K_1 X_t + K_2 V_t,$$

$$H = H_0 + H_1 X_t + H_2 V_t,$$

where

$$K_0 = \begin{pmatrix} \kappa\theta \\ \lambda\varepsilon \end{pmatrix}, \quad K_1 = \begin{pmatrix} -\kappa \\ 0 \end{pmatrix}, \quad K_2 = \begin{pmatrix} 0 \\ -\lambda \end{pmatrix}.$$

$$H_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad H_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & \rho\upsilon \\ \rho\upsilon & \upsilon^2 \end{pmatrix}.$$

As a consequence, we deduce the following system of Riccati equations:

$$\frac{\partial \psi_0}{\partial \tau} = K_0^T \eta + \frac{1}{2}\eta^T H_0 \eta = \kappa\theta\psi_1 + \lambda\varepsilon\psi_2, \tag{4.17}$$
$$\frac{\partial \psi_1}{\partial \tau} = K_1^T \eta + \frac{1}{2}\eta^T H_1 \eta = -\kappa\psi_1 \tag{4.18}$$
$$\frac{\partial \psi_2}{\partial \tau} = K_2^T \eta + \frac{1}{2}\eta^T H_2 \eta = -\lambda\psi_2 + \frac{1}{2}\psi_1^2 + \rho\upsilon\psi_1\psi_2 + \frac{1}{2}\upsilon^2\psi_2^2, \tag{4.19}$$

where $\eta^T = (\psi_1\ \psi_2)$ *with conditions* $\psi_0(0, m(u_1), m(u_2)) = 0$, $\psi_1(0, m(u_1), m(u_2)) = m(iu_1)$, $\psi_2(0, m(u_1), m(u_2)) = m(iu_2)$. *The solutions take the form:*

$$\psi_1(\tau - t, u_1, u_2) = m(iu_1)e^{-\kappa(\tau - t)}. \tag{4.20}$$
$$\psi_0(\tau - t, u_1, u_2) = \theta m(iu_1)(1 - e^{-\kappa(\tau-t)}) + \lambda\varepsilon \int_t^\tau \psi_2(s, u_1, u_2)\, \mathrm{d}s. \tag{4.21}$$

Proof 11 *This follows from the applications of Theorems 2 and 3 and similar steps as in solving the system of Riccati equations in the case of the one factor model.*

To obtain the solution ψ_2 to the Riccati equation (4.19) is not trivial. However, a similar problem has been handled in [KNPP15].

Lemma 10 *Consider Proposition 1 and let* $\zeta(y),\ y \in \mathbb{C} - \{0\}$ *be such that it satisfies*

$$\frac{\mathrm{d}\zeta(y)}{\mathrm{d}y} = \zeta(y)^2 + \left(\frac{\kappa - \lambda}{\kappa y} - \frac{\rho\upsilon}{\kappa}\right)\zeta(y) + \frac{\upsilon^2}{4\kappa^2}, \tag{4.22}$$

then the solution to (4.19) can be expressed by

$$\chi(\tau - t, u_1) = -\frac{2\kappa m(iu_1)}{\upsilon^2} e^{-\kappa(\tau-t)}\zeta(m(iu_1)e^{-\kappa(\tau-t)}). \tag{4.23}$$

Moreover, the general solution to (4.19) takes the form

$$\psi_2 = \chi + \frac{1}{\omega}, \tag{4.24}$$

where ω satisfies

$$\frac{\partial\omega}{\partial\tau}+(-\lambda+\rho\upsilon m(iu_1)e^{-\kappa(\tau-t)}+\upsilon^2\chi)\omega=-\frac{1}{2}\upsilon^2, \tag{4.25}$$

with the general solution given by

$$\omega(\tau-t)=\frac{C-\frac{1}{2}\upsilon^2\int_t^\tau I_f(s-t)\,\mathrm{d}s}{I_f(\tau-t)}, \tag{4.26}$$

where I_f is an integrating factor and C is the constant of integration.

Proof 12 *Claim (4.23) is verified by differentiating with respect to t and relating it to (4.22) and (4.20):*

$$\frac{\partial\chi}{\partial\tau}=-\lambda\chi+\frac{1}{2}\upsilon^2\chi^2+\rho\upsilon\chi\psi_1+\frac{1}{2}\psi_1^2. \tag{4.27}$$

Similarly, (4.24) is verified by substitution into (4.19) and relating it to (4.27) resulting into

$$-\frac{1}{\omega^2}\frac{\partial\omega}{\partial\tau}=-\frac{\lambda}{\omega}+\frac{1}{2}\upsilon^2\left(\frac{2\chi}{\omega}+\frac{1}{\omega^2}\right)+\frac{\rho\upsilon\psi_1}{\omega}, \tag{4.28}$$

from which (4.25) follows. The general solution to (4.25) is obtained using the integrating factor

$$\begin{aligned}&I_f(\tau-t,u_1)\\&:=\exp\left(-\lambda(\tau-t)-\frac{\rho\upsilon m(iu_1)}{\kappa}e^{-\kappa(\tau-t)}-2\kappa m(iu_1)\int e^{-\kappa(s-t)}\zeta(m(iu_1)e^{-\kappa(s-t)})\,\mathrm{d}s\right)\\&=\exp\left(-\lambda(\tau-t)-\frac{\rho\upsilon m(iu_1)}{\kappa}e^{-\kappa(\tau-t)}+2\int_0^{m(iu_1)e^{-\kappa(\tau-t)}}\zeta(y)\,\mathrm{d}y\right).\end{aligned} \tag{4.29}$$

Lemma 11 *A representation of the solution ψ_2 to (4.19) is given by*

$$\psi_2(\tau - t, u_1, u_2) = -\frac{2\kappa m(iu_1)}{\upsilon^2} e^{-\kappa(\tau-t)} \zeta(m(iu_1)e^{-\kappa(\tau-t)}) + \frac{I_f(\tau - t, u_1)}{C(u_1, u_2) - \frac{1}{2}\upsilon^2 \int_t^\tau I_f(s-t, u_1)\,\mathrm{d}s}, \tag{4.30}$$

where the constant of integration C is determined by applying $\psi_2(0, u_1, u_2) = m(iu_2)$:

$$C(u_1, u_2) = \frac{\exp(-\frac{\rho\upsilon m(iu_1)}{\kappa} + 2\int_0^{m(iu_1)} \zeta(y)\,\mathrm{d}y)}{m(iu_2) + \frac{2\kappa m(iu_1)}{\upsilon^2}\zeta(m(iu_1))}. \tag{4.31}$$

Proof 13 *The function ζ can be expressed in the form (see [KNPP15]):*

$$\zeta(y) = \sum_{j=1}^{\infty} d_j y^j,$$

Functions $\psi_2(t, u_1, u_2)$ and $\psi_0(t, u_1, u_2)$ in (4.30) and (4.21) respectively can be rewritten as

$$\psi_2(\tau - t, u_1, u_2) = -\frac{2\kappa m(iu_1)}{\upsilon^2} e^{-\kappa(\tau-t)} \sum_{j=1}^{\infty} d_j m(iu_1)^j e^{-j\kappa(\tau-t)} + \frac{I_f((\tau - t), u_1)}{C(u_1, u_2) - \frac{1}{2}\upsilon^2 \int_t^\tau I_f(s-t, u_1)\,\mathrm{d}s}.$$

$$\psi_0(\tau - t, u_1, u_2) = \theta m(iu_1)(1 - e^{-\kappa(\tau-t)}) + \frac{-2\kappa\lambda\varepsilon m(iu_1)}{\upsilon^2} \sum_{j=1}^{\infty} d_j m(iu_1)^j \left(\frac{1}{\kappa(1+j)}(1 - e^{-\kappa(\tau-t)(1+j)}) \right). + \int_t^\tau \frac{I_f(s-t, u_1)}{C(u_1, u_2) - \frac{1}{2}\upsilon^2 \int_t^s I_f(q-t, u_1)\,\mathrm{d}q}\,\mathrm{d}s,$$

where the integrating factor introduced in (4.29) and its integral are given by

$$I_f(\tau - t, u_1) = \exp\left(- \lambda(\tau - t) - \frac{\rho v m(iu_1)}{\kappa} e^{-\kappa(\tau - t)} + 2e^{-\kappa(\tau-t)} \sum_{j=1}^{\infty} \frac{d_j}{j+1} m(iu_1)^{j+1} \right).$$

$$\int_t^{\tau} I_f(q - t, u_1)\, \mathrm{d}q = \frac{I_f(\tau - t, u_i)}{(-\lambda + \rho v m(iu_1) e^{-\kappa(\tau-t)} - 2e^{-\kappa(\tau-t)} \sum_{j=1}^{\infty} \frac{d_j}{j+1} m(iu_1)^{j+1}} - \frac{\exp\left(\frac{-\rho v m(iu_1)}{\kappa} + 2 \sum_{j=1}^{\infty} \frac{d_j}{j+1} m(iu_1)^{j+1} \right)}{(-\lambda + \rho v m(iu_1) - 2 \sum_{j=1}^{\infty} \frac{d_j}{j+1} m(iu_1)^{j+1})}.$$

Finally, the constant of integration (4.31) can be re-written as

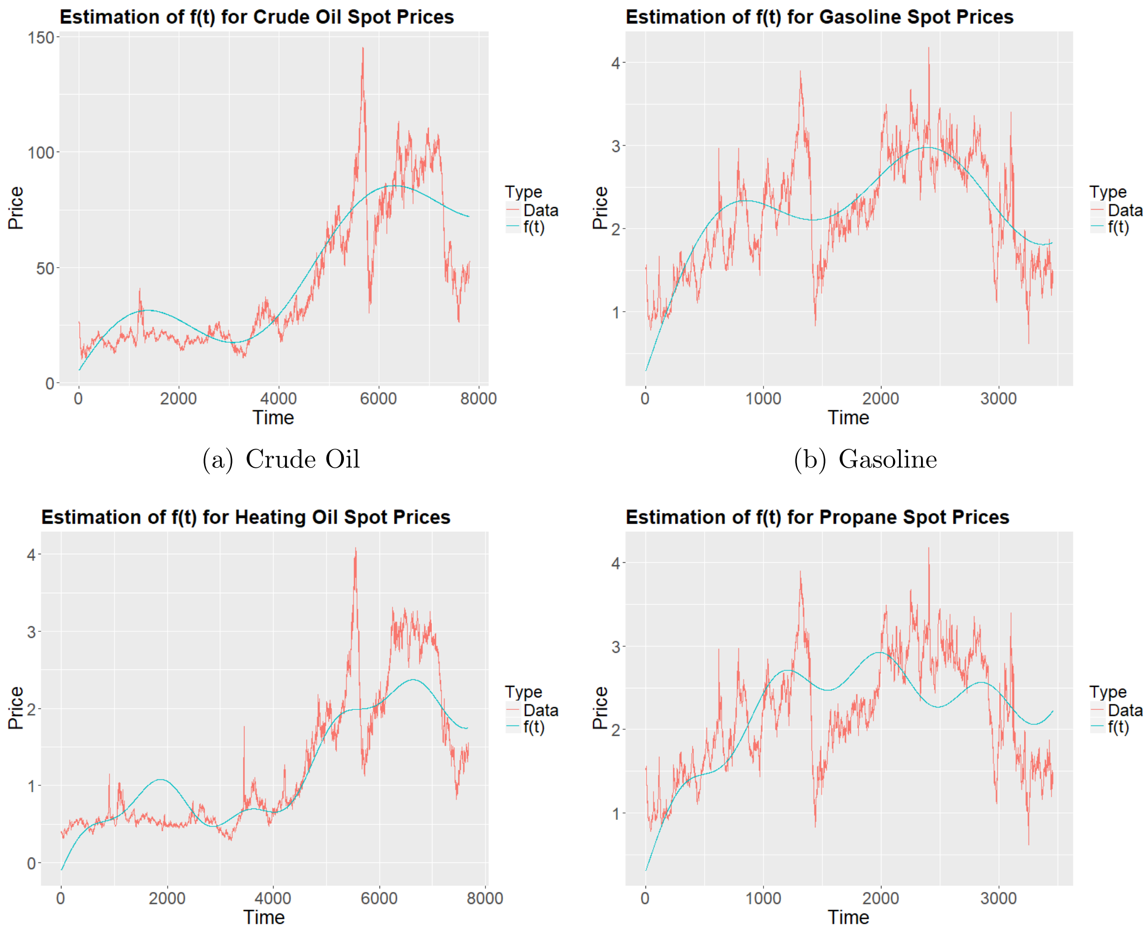
$$C(u_1, u_2) = \frac{\exp\left(- \frac{\rho \eta m(iu_1)}{\kappa} + 2 \sum_{j=1}^{\infty} \frac{d_i}{1+j} m(iu_2)^{j+1} \right)}{m(iu_2) + \frac{2\kappa m(iu_1)}{\eta^2} \sum_{j=1}^{\infty} d_j m(iu_1)^j}.$$

This completes the proof.

4.4 Numerical Implementation

We focus on the one factor model to explain our approach for estimating the model parameters. The data used in this section is obtained from the US Energy Information Administration and includes future prices of Crude Oil (Light-Sweet, Cushing, Oklahoma) from Mar 30, 1983 to Dec 06, 2016 (8452 observations), Reformulated Regular Gasoline (New York Harbor) from Dec 03, 1984 to Oct 31, 2006 (5492 observations), Heating Oil (New York Harbor) from Jan 02, 1980 to Dec 06, 2016 (9262 observations) and Propane (Mont Belvieu, Texas) from Dec 17, 1993 to Sep 18, 2009 (3941 observations).

The parameters in the seasonality function (4.5) are estimated by fitting the function to the historical spot prices. The spot prices used include Crude Oil from Jan 02, 1986 to Dec 12, 2016 (7807 observations), RBOB Regular Gasoline from Mar 11, 2003 to Dec 12, 2016 (3460 observations), No. 2 Heating Oil from Jun 02, 1986 to Dec 12, 2016 (7683 observations) and Propane from Jul 09, 1992 to Dec 12, 2016 (6133 observations).



(a) Crude Oil (b) Gasoline (c) No. 2 Heating Oil (d) Propane

Fig. 4.1: Seasonality estimation from commodity spot prices.

The parameters of the seasonality are estimated from historical data using the `optim()` function in R software. The accuracy of the fitting in Figure 4.1 depends on the choice of the initial values of parameters δ_0, δ_1, δ_2, δ_3 and δ_4.

Table 4.1: Estimation of parameters in the seasonality function

Commodity	δ_0	δ_1	δ_2	δ_3	δ_4
Crude Oil	0.01097129	-0.04210297	13.29908652	18.43819425	6.28446063
Gasoline	0.001018775	-0.515953818	6.27995144	1.252457783	6.28435941
Heating Oil	0.000287309	0.159154792	12.57029401	-0.49422029	6.28182486
Propane	0.000243901	-0.252880065	12.56438012	0.071328671	6.286055367

$$f(t) = \delta_0 t + \delta_1 \sin(\delta_2[t + 2\pi/264]) + \delta_3 \sin(\delta_4[t + 4\pi/264]).$$

4.4.1 Equivalent Latent Regression Model

The deseasonalised one factor future price given by (4.7) can be written as

$$y = a + bx + \varepsilon, \tag{4.32}$$

where $y = \ln F$, $x = S_t$ and a and b are given by

$$a = [(-\frac{1}{2}\sigma^2 + \beta)^\alpha + \beta^\alpha]\gamma(1 - e^{-\kappa(\tau - t)}) + \frac{1}{4\kappa}\sigma^2[(-\frac{1}{2}\sigma^2 + \beta)^\alpha + \beta^\alpha]^2(1 - e^{-2\kappa(\tau - t)}) \tag{4.33}$$

$$b = [(-\frac{1}{2}\sigma^2 + \beta)^\alpha + \beta^\alpha]e^{-\kappa(\tau - t)}, \tag{4.34}$$

and ε is an independent random error distributed as $N(0, \Theta)$ with zero mean and co-variance matrix Θ.

Clearly, (4.32) belongs to the class of latent regression models since $x = S_t$ is not observable.

This kind of problem can be handled using expectation-maximization (EM) algorithms (see [DLR77]) to estimate the model parameters. The latent variable x can be considered binary where in this case the EM algorithm would give estimates for a 2-component normal mixture model. On the other hand, x can be allowed to be continuous between 0 and 1 with a beta distribution as in [TP10]. The EM algorithm for estimating the model parameters in this case is more involved than for the 2-component mixture model and more computationally challenging, but can be done nonetheless. Basically, the latent or unobserved x variables are imputed by their conditional expectation given the outcomes y. We adapt the latter approach through the Dynkin-Lamperti Theorem (see [GN04]) where the unobserved variable follows a stable distribution defined on $(-\infty, \infty)$ with $\alpha \in (0, 2]$ and the observable variable y represents the log-returns of the futures prices. The algorithm is applied to the joint likelihood of the response y. We assume the error ε is independent of the latent predictor x. The joint density for x and y is given by

$$
\begin{aligned}
h(x, y) &= h(y|x; a, b, \Theta) S(x; \alpha, \beta, \nu, \mu). \\
&= N(y; a + bx, \Theta) S(x; \alpha, \beta, \nu, \mu),
\end{aligned} \tag{4.35}
$$

where $S(x; \alpha, \beta, \nu, \mu)$ is the α stable distribution, $N(y; A, B)$ denotes the normal distribution of a random variable y with mean A and variance B; thus, Θ is the variance of the outcome sample data. The marginal density of the response y is

$$
\begin{aligned}
f(y) = \int \frac{1}{\sqrt{2\pi}} \Theta^{-1/2} \exp(-\Theta^{-1}(-y - \beta_0 - \beta_1 x)'(y - \beta_0 - \beta_1 x)/2) \\
\times S(x; \alpha, \beta, \nu, \mu) \, \mathrm{d}x.
\end{aligned} \tag{4.36}
$$

Density (4.36) is an example of infinite mixture models used in ecological statistics ([FNYS15]).

4.4.2 The EM Algorithm

For a data set $(x_1, y_1), \cdots, (x_n, y_n)$ in (4.32), the log-likelihood is derived from (4.35) as

$$
\begin{aligned}
L(\alpha, \beta, \nu, \mu, \Theta, a, b) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log|\Theta| \\
- \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)'(y_i - \beta_0 - \beta_1 x_i)/2 \\
+ \sum_{i=1}^{n}\log(S(x_i; \alpha, \beta, \nu, \mu)). \qquad (4.37)
\end{aligned}
$$

Since x is not observable, the EM algorithm requires maximizing the conditional expectation of the log-likelihood given the response vector y. That is

$$\mathtt{E}[L(\alpha, \beta, \nu, \mu, \Theta, a, b)|y],$$

where at each iteration of the EM algorithm, the above conditional expectation is computed using the current parameter estimates. This current expectation-maximization problem is similar to the problem handled in [TP10] which implies a similar EM algorithm can be applied here. That is, suppose $\rho = \begin{pmatrix} a \\ b \end{pmatrix}$ is a $2 \times p$ dimension coefficient matrix where each of the p columns of ρ provides the intercept and slope regression coefficients for each of the ρ response variables. Then denote by $\mathbf{X}$ the design matrix whose first column consists of ones for the intercept and the second column consists of the latent predictors $x_i,\ i = 1, \cdots, n$. The multivariate regression model follows:

$$\mathbf{Y} = \mathbf{X}\rho + \varepsilon.$$

Moreover, and as indicated in [TP10], the likelihood for the multivariate normal regression model can be given as

$$L(\rho, \Theta) = (2\pi)^{-np/2} |\Theta|^{-1/2} \exp(-tr[\Theta^{-1}(\mathbf{Y} - \mathbf{X}\rho)'(\mathbf{Y} - \mathbf{X}\rho)]/2). \tag{4.38}$$

The EM approach requires that we maximize the expectation of the logarithm of (4.38) conditional on $\mathbf{Y}$ with respect to ρ and Θ. This leads to the following optimal factors:

$$\hat{\rho} = (\widetilde{\mathbf{X}'\mathbf{X}})^{-1} \tilde{\mathbf{X}}'\mathbf{Y}.$$
$$\hat{\Theta} = \mathbf{Y}'\mathbf{Y} - \hat{\rho}'(\widetilde{\mathbf{X}'\mathbf{X}})\hat{\rho},$$

where $\tilde{\mathbf{X}} = \mathtt{E}[\mathbf{X}|\mathbf{Y}]$ and $(\widetilde{\mathbf{X}'\mathbf{X}}) = \mathtt{E}[\mathbf{X}'\mathbf{X}|\mathbf{Y}]$.

To implement the EM method in the R programming language, we first highlight that there are minor differences to bear in mind before implementing the algorithm as we explain in the following.

First, the density of the predictor in our case is from a stable distribution. Recall, in general the densities of stable processes cannot be expressed analytically which makes it difficult to compute the log likelihood. However, with the help of inbuilt packages in R including `stabledist` and `StableEstim`, the log likelihood can be satisfactorily estimated using `Estim()` to obtain the stable parameters of S_t, and `dstable()` for its corresponding stable density.

Secondly, from the log-likelihood expression, we notice that we only require estimates of the conditional expectations of x, x^2 and $\log x$ with respect to the joint probability density given the response vector y.

On the other hand, we retain some of the steps in [TP10]. The initial values for the regression parameters a and b can be obtained from fitting a two-component finite

Table 4.2: Parameters obtained from maximum likelihood method

Parameter Estimation		
Parameter	1st Iteration	2nd Iteration
loglike	30687	31133.2
a	$7.74707e-05$	$1.17456e-05$
b	0.00218499	0.00166723
α	1.6605	1.6605
β	-0.0651915	-0.0651915
ν	0.00576286	0.00576286
μ	0.000415904	0.000415904
Θ	0.0106344	0.0110766
T	5000 days	5000 days

mixture model or by a preliminary search over the parameter space. Initial values for Θ can be obtained using the sample covariance matrix from the raw data.

4.4.3 Data for the EM Algorithm

The data used is stored in a data frame with three columns containing futures log-returns, spot price log-returns and binary data of 1's and 0's representing whether or not a jump has occurred within a given window size (see Table 4.3). Table 4.2 shows the estimated parameters as a result. The parameters were obtained from 5000 data points of crude oil log-returns arranged as in Table 4.3. We have displayed results from only two iterations because for large data sets the code tends to be slow in addition to suffering convergence issues. However, this can be improved and by using faster machines.

The jump occurrence due to S_t is determined by the method discussed in [LM08] (also see [MRD13]). That is: The realized return at any given time is compared to a continuously estimated instantaneous volatility σ_{t_i} to measure local variation arising from the continuous part of the process. The volatility σ_{t_i} is estimated using a modified version

Table 4.3: A snapshot of the structure of the data used

Futures log-returns	Spot log-returns	Jump detection
.	.	.
.	.	.
.	.	.
-0.015646306	-0.020072054	0
0.0322302	0.047966668	0
0	-0.016401471	0
-0.001778683	0.001489862	0
-0.017271048	-0.019778648	1
-0.02611648	-0.022030684	1
-0.043103026	-0.033698961	0
0	-0.010387745	0
0.001808809	0.003969483	0
0.037999099	0.036829588	1
-0.030483308	-0.030057522	0
0.024152859	0.019704761	0
0.009625568	0.008368168	0
.	.	.
.	.	.
.	.	.

of realized bipower variation calculated as the sum of products of consecutive absolute returns in the local window (see [BNS04]). Then, the jump detection statistic $\mathcal{L}_i \in$ testing for jumps in returns occurring at a time t_i within a window size K is calculated as the ratio of realized returns to estimated instantaneous volatility:

$$\mathcal{L}_i \equiv \frac{\log Y_{t_i}/Y_{t_{i-1}}}{\hat{\sigma}_{t_i}},$$

where Y_t at $t \geq 0$ represents the commodity spot price and $\hat{\sigma}_{t_i}$ is estimated by

$$\hat{\sigma}^2_{t_i} \equiv \frac{1}{K-2} \sum_{j=i-K+2}^{i-1} |\log(Y_{t_j}/Y_{t_{j-1}})||\log(Y_{t_{j-1}}/Y_{t_{j-2}})|.$$

Care must be taken in choosing K, it must be large enough to accurately estimate integrated volatility but small enough for the variance to be approximately constant. In other words, K should be large enough but smaller than N, the number of observations so that the effect of jumps on estimating instantaneous volatility disappears. Some authors recommend K to be computed as $K = \sqrt{252 \cdot n}$, where n is the daily number of observations, whereas 252 is the number of days in the (financial) year. Moreover the window size should be such that $K = \mathcal{O}(\triangle t^{\lambda})$ with $-1 < \lambda < -0.5$. For high frequency data, [LM08] recommend, for returns sampled at frequencies of 60, 30, 15 and 5 minutes, the corresponding values of K to be 78, 110, 156 and 270. For our case we shall choose $K = 4$ for crude oil futures prices with returns sampled daily.

Detection of Jumps in the Data

The test statistic $\mathcal{L}$ follows approximately a normal distribution when the data set has no jumps and its value becomes large otherwise. According to [LM08], the region for $\mathcal{L}$ is chosen based on the distribution of its maximum. For instance, suppose a particular interval $(t_{i-1}, t_i]$ has no jumps and the distance between two consecutive observations in this interval is small (i.e. $\triangle \to 0$). Then the maximum should converge to the Gumbel variable:

$$\frac{\max_{i \in \bar{A}_N} |\mathcal{L}_i| - c_N}{s_N} \to \xi, \tag{4.39}$$

where ξ has a cumulative distribution function $\mathtt{P}(\xi \leq x) = \exp(-e^{-x})$, $\bar{A}_N$ is the set of $i \in \{1, 2, \cdots, N\}$ such that there is no jump in $(t_{i-1}, t_i]$ and c_N, s_N are defined as

$$c_N = \frac{(2\log N)^{1/2}}{0.8} - \frac{\log \pi + \log(\log N)}{1.6(2\log N)^{1/2}}.$$
$$s_N = \frac{1}{0.8(2\log N)^{1/2}}.$$

The test is conducted by comparing the standardized maximum of $\mathcal{L}_i$ in (4.39) to the critical values from the Gumbel distribution where the null hypothesis of no jump is rejected when the jump statistic

$$\mathcal{L}_i > G^{-1}(1-\lambda)s_N + c_N,$$

where $G^{-1}(1-\lambda)$ is the $(1-\lambda)$ quantile function of the standard Gumbel distribution. Suppose $\lambda = 0.1$, then we reject the null hypothesis of no jump when $\mathcal{L}_i > s_N\eta^* + c_N$ where η^* is such that $\exp(-e^{-\eta^*}) = 1-\eta^* = 0.9$. That is, $\eta^* = -\log(-\log(0.9)) = 2.25$. Figure 4.3 shows a graph of jumps detected in crude oil futures prices where we have used 1's to record a jump occurrence and 0 for no jump.

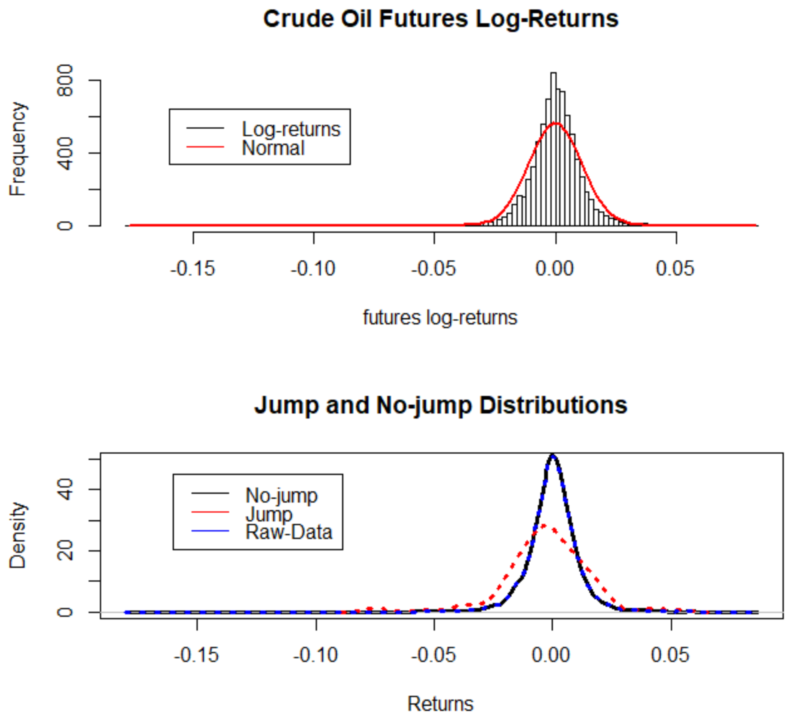


Fig. 4.2: Detection of jumps in crude oil futures prices.

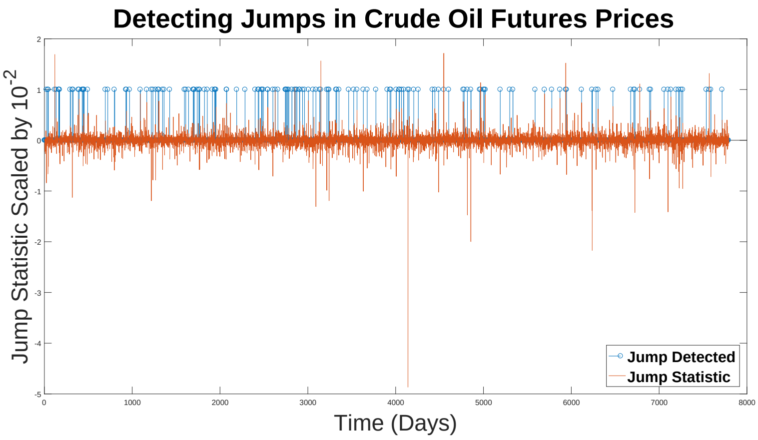


Fig. 4.3: Detection of jumps in crude oil futures prices.

4.5 Summary

We have shown that the affine property is attainable and applicable to generalized spot models. We considered a stochastic differential equation with the source of randomness as subordinated Brownian motion as a specific example to derive the futures price. Moreover, it has been argued in some existing literature that the likelihood function exists in integrated form for models with singular noise meanwhile for cases of partially observed processes a filtering technique is required. However, the work presented in this chapter provided a new approach of pricing commodity futures for models with latent variables using the maximum expectation maximization, without using any filtering. Our approach is easy to implement once the joint probability density is established. The numerical implementation of the two factor model is left for future work.

Chapter 5

Bismut-Elworthy-Li Formula for Subordinated Brownian motion with Application to Hedging

5.1 Introduction

This chapter follows from Section 1.3.3 and utilizes tools from Chapter 2. The contribution of the Chapter will also include deriving much simpler techniques based on the basic Malliavin integration by parts formula to arrive at similar results in related existing literature (such as [Zha12] for instance). It is organized as follows. In Section 5.2 we review common methods in literature for computing the Greeks. Section 5.3 utilizes results from Chapter 1.3.3 and investigates the differential calculus for subordinated Brownian motion. We show the integration by parts formula exists for a subordinated Brownian motion process. By employing results from the preceding sections, Section 5.4 derives the Bismut-Elworthy-Li formula for subordinated Brownian motion which leads to the main results of the chapter. In Section 5.5 we discuss the applications of the BEL formula in computing financial Greeks. Section 5.6 summarizes.

5.2 Sensitivity Analysis

The theory of risk-neutral valuation asserts that given a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathtt{P})$ with a subjective probability measure $\mathtt{P}$ and a filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions, we can construct a payoff function Φ of an option under an equivalent martingale measure whose price is given by

$$V_t = \mathtt{E}[\tilde{\Phi}(X_T)|\mathcal{F}_t], \tag{5.1}$$

where $\tilde{\Phi} = e^{-rT}\Phi$ is the discounted payoff function by a risk free interest rate r and X_T is the value of the underlying at maturity time T. A Greek is a derivative of (5.1) with respect to a certain model parameter p which could be the initial value X_0 of the underlying, the volatility parameter σ, time to maturity $\tau = T - t$, the option strike, E or the interest rate, r, i.e.

$$Greek = \partial_p V_t := \frac{\partial V_t}{\partial p}. \tag{5.2}$$

Clearly, (5.2) poses a problem if $\tilde{\Phi}(X_T)$ is not differentiable. Various methods in literature explore this challenge as we discuss in the following. The Likelihood method [FLL+99] is suitable for known distribution of the underlying. It takes the form

$$\frac{\partial V_t}{\partial p} = \mathtt{E}[\tilde{\Phi}(p)\partial_p \ln(\rho(p))|\mathcal{F}_t], \tag{5.3}$$

where ρ denotes the density of the underlying. Malliavin calculus provides another approach (see [BM06, BN13, DNØP08, Nua95, Ber00, Mhl15]). It eliminates differentiation of the payoff function by introducing a weight factor in terms of the Malliavin derivative or the Ornstein-Uhlenbeck operator, i.e.

$$\frac{\partial V_t}{\partial p} = \mathtt{E}[\tilde{\Phi}(X(p))\mathcal{E}(X(p), G(p))|\mathcal{F}_t], \tag{5.4}$$

where the weight factor $\mathcal{E}$ consists of Malliavin derivatives of random variables X and G belonging to some space with nice properties e.g. L^2. This approach is more flexible than the previous ones in the sense that the distribution of the underlying is irrelevant to computing the Greeks. However, it is computationally expensive. Finally, is the Bismut-Elworthy-Li representation formula [EL94, Tak10, KKH10, Khe12, CF07, BnDMBP15],

$$\frac{\partial V_t}{\partial x} = \mathtt{E}\Big[\Phi(X_T)\int_0^T a_s \frac{\partial X_s}{\partial x}\,\mathrm{d}B_s\Big|\mathcal{F}_t\Big], \tag{5.5}$$

where $x = X_0$ and a_t is some bounded function satisfying

$$\int_0^T a_s\,\mathrm{d}s = 1. \tag{5.6}$$

The Bismut-Elworthy-Li Formula (5.5) applies to continuous diffusion processes but can be adapted to finite (see [CF07]) and infinite (see [CSZ15]) jump processes. The usefulness of this formula is its allowance for an explicit representation of the Delta of a financial derivative.

For instance by employing the theory of Malliavin calculus, the weight $\mathcal{E}$ can be obtained explicitly for different Brownian motion functionals. Malliavin calculus for both continuous and jump diffusion processes has been extensively discussed in literature and there exist enormous applications on the subject (see [Kus10], [BnDMBP15], [EL94], [DNØP08], [BM06], [BN13], [CF07] and [Mhl15]). The focus of this chapter is to compute the Greeks for a wide range of payoffs irrespective of their structure by employing the BEL formula in the framework of subordinated Brownian motion. The subordinator is an $\alpha/2$-stable process where $\alpha \in (0, 2)$ to ensure only positive jumps.

5.3 Malliavin derivative in the direction of jump processes

This section explores the differential calculus of B_{L_t} and to the best of our knowledge, little has been done in this direction, some references include [Kus10, Zha12].

We now replace X by B and S by L in Definition 9 to introduce a new joint probability space corresponding to Brownian motion B_t and the inverse process L_t and define $\Omega := \mathcal{C}([0, L_T])$ endowed with the natural filtrations:

$$\mathcal{F}_t := \sigma\left\{B_{L_\tau} : \tau \leq L_t\right\}, \quad \mathcal{F}_t^B := \sigma\left\{B_\tau : \tau \leq L_t\right\}. \tag{5.7}$$

and introduce a separable Hilbert space

$$\mathcal{H} := \left\{h \in \mathcal{C}(\Omega; \mathbb{R}),\ h \text{ is absolutely continuous and } \dot{h} \in \mathcal{L}^2(\Omega; \mathbb{R})\right\}, \tag{5.8}$$

to obtain a complete abstract probability space $(\mathcal{H}, \Omega, \mathtt{P}, (\mathcal{F}_t)_{t\geq 0})$.

Lemma 12 *Let f be an $(\mathcal{F}_t)$-adapted right-continuous process from (5.7) with left limits satisfying*

$$\mathtt{E}\left[\int_0^T |f(\tau-)|^2 \,\mathrm{d}L_\tau\right] < \infty. \tag{5.9}$$

where the notation $f(\tau-)$ represents value of the function at the left limit. Then its $(\mathcal{F}_t)$-martingale stochastic integral exists, it is well defined and can be expressed as an $(\mathcal{F}_t^B)$-martingale stochastic integral i.e.

$$\int_0^T f(t-)\,\mathrm{d}B_{L_t} = \int_0^{L_T} f(S_\tau-)\,\mathrm{d}B_\tau, \tag{5.10}$$

where S_t is the inverse stable process of L_t. See [Kus10].

Proof 14 *This follows from the standard change of time.* □

Following Definition 9, we denote by D the Malliavin derivative operator defined on $\mathcal{H}$ such that $\dot{h}$ represents differentiation of h with respect to L_t. In addition, we denote by D_h the Malliavin differentiation in direction h.

Lemma 13 *Let B_{L_t} be subordinated Brownian motion associated with $(\mathcal{H}, \Omega, (\mathcal{F}_t)_{t\geq 0}, \mathtt{P})$. Then $\mathrm{D}_h B_{L_t} = h(L_t)$, for all $h \in \mathcal{H}$ and $t \in [0, T]$. For a càdlàg process f we have*

$$\mathrm{D}_h \int_0^T f(t-)\,\mathrm{d}B_{L_t} = \int_0^{L_T} f(S_t-)\,\mathrm{d}h(t), \tag{5.11}$$

where S_t is the inverse stable process of L_t.

Proof 15 *Let $f = B_{L_t}$. It is easy to see that $\mathrm{D}_h B_{L_t} = h(L_t)$ since by definition*

$$\begin{aligned}(\mathrm{D}_h f)(L_t) &= \lim_{\varepsilon\to 0} \frac{f(L_t + \varepsilon h(L_t)) - f(L_t)}{\varepsilon}, \;\; h \in \mathcal{H}, \\ &= \lim_{\varepsilon\to 0} \frac{B_{L_t} + \varepsilon h(L_t) - B_{L_t}}{\varepsilon} \\ &= h(L_t), \end{aligned} \tag{5.12}$$

for $f \in (\mathcal{H}, \Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathtt{P})$, provided the limit exists in $\mathcal{L}^2(\Omega)$.

For the second part of the lemma we notice that since L_t is of bounded variation, the contribution of its small jumps is almost negligible and the number of jumps is finite. We partition $[0, T]$ as $0 = \tau_0 < \tau_1 < \cdots < \tau_{n-1} < \tau_n = T$ where τ_i, $i = 1, \cdots n-1$ are the jump times of B_{L_t} and let $\{t_{i,j}; j = 0, 1, 2, \cdots, N_i\}$ be a partition of $[\tau_{i-1}, \tau_i)$.

Suppose $\Delta := \max_{i,j}(t_{i,j} - t_{i,j-1})$*, we have*

$$\begin{aligned}
\mathrm{D}_h \int_0^T f(t-)\,\mathrm{d}B_{L_t} &= \int_0^T f(t-)\,\mathrm{d}h(L_t). \\
&= \lim_{\Delta\to 0} \sum_{i=1}^{n} \Bigg[\sum_{j=1}^{n-1} f(t_{i,j-1}-)[h(L_{t_{i,j}}) - h(L_{t_{i,j-1}})] \\
&\quad + f(t_{i,n_i-1}-)[h(L_{t_{i,n_i}}-) - h(L_{t_{i,n_i-1}})] \\
&\quad + f(t_{i,n_i}-)[h(L_{t_{i,n_i}}) - h(L_{t_{i,n_i}}-)]\Bigg]. \qquad (5.13)
\end{aligned}$$

Now if we let $u_{i,j} := L_{t_{i,j}}$ *then* $S_{u_{i,j}} = t_{i,j}$ *for* $j = 0, 1, 2, \cdots, N_i - 1$ *and as a consequence we write*

$$\begin{aligned}
\mathrm{D}_h \int_0^T f(t)\,\mathrm{d}B_{L_t} &= \int_0^T f(t)\,\mathrm{d}h(L_t). \\
&= \lim_{\Delta\to 0} \sum_{i=1}^{n} \Bigg[\sum_{j=1}^{n-1} f(S_{u_{i,j-1}}-)[h(u_{i,j}) - h(u_{i,j-1})] \\
&\quad + f(S_{u_{i,n_i-1}}-)[h(u_{i,n_i}-) - h(u_{i,n_i-1})] \\
&\quad + f(S_{u_{i,n_i}}-)[h(u_{i,n_i}) - h(u_{i,n_i}-)]\Bigg]. \\
&= \sum_{i=1}^{n} \int_{L_{\tau_{i-1}}}^{L_{\tau_i-}} f(S_u-)dh(u) + \sum_{i=1}^{n} f(S_{u_{i,n_i}}-)[h(\tau_i) - h(\tau_i-)].
\end{aligned}$$

Note that S_u *is constant on* $u \in [L_{\tau_i-}, L_{\tau_i}]$*. Therefore*

$$f(S_{u_{i,n_i}}-)[h(L_{\tau_i}) - h(L_{\tau}-)] = \int_{L_{\tau_{i-1}}}^{L_{\tau_i-}} f(S_u-)dh(u), \quad i = 1, 2, \cdots, n.$$

The result follows immediately, i.e.

$$\mathrm{D}_h \int_0^T f(t-)\,\mathrm{d}B_{L_t} = \int_0^{L_T} f(S_t-)\,\mathrm{d}h(t), \qquad (5.14)$$

Lemma 14 *Suppose f is a right continuous function with left limits, then*

$$\int_0^T f(\tau-)\,\mathrm{d}L_\tau = \int_0^{L_T} f(S_\tau-)\,\mathrm{d}\tau, \tag{5.15}$$

where S_t is the inverse stable process of L_t.

Proof 16 *This follows from standard change of time computations. Also see [Kus10].*

5.3.1 Discrete Multiple Stochastic Integral

In this section we derive the integration by parts formula associated with the random process B_{L_t}. We denote by ζ_k the time between the k-th and $(k+1)$-th jumps of B_{L_t}.

We start with the triplet $(\mathcal{H}, \Omega, \mathtt{P})$, a joint probability space introduced in Definition 9 with a real seperable Hilbert space $\mathcal{H}$ and a scalar product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. The norm for $g \in \mathcal{H}$ is denoted by $\|g\|_{\mathcal{H}}$. Finally, $\mathtt{P}$ the extension to the Borel $\sigma-$algebra of Ω of a cylindrical measure. We define independent stable random variables $\zeta_k := B_{L_k} - B_{L_k-}$ which are canonical projections from Ω to $\mathbb{R}$. We assume L_k is a form of subordinator introduced in ([JW93], P. 33) which is an $\alpha/2$-stable totally skewed Lévy motion with increasing sample paths $(\alpha \in (0,2), \beta = 1)$. This is a symmetric alpha-stable process $(S\alpha S)$ with positive Poisson jumps. Therefore, B_{L_k} belongs to a class of $S\alpha S$ Lévy motion processes with its jumps only at the jump times of L_k. As a consequence, we use Charlier polynomials to define the multiple stochastic integrals with respect to our process.

Definition 12 *The Charlier polynomials are defined as*

$$C_n^{(\lambda)}(x) = (-1)^n \lambda^{-x} e^{\lambda} \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}(e^{-\lambda}\lambda^x). \tag{5.16}$$

Alternatively, this can be expressed explicitly as

$$C_n^{(\lambda)}(x) = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} (-1)^{n-k}(\lambda)^{-k}(x)_k, \quad x \in \Omega,$$

where $(x)_k := x(x-1)\cdots(x-k+1)$.

The Charlier polynomials form an orthogonal basis of $\mathcal{L}^2(\Omega, \mathcal{F}, \mathtt{P})$ with respect to the Poisson measure $\mu(\mathrm{d}x) = \frac{\lambda^x}{x!}e^{\lambda}\,\mathrm{d}x$. Moreover, we have:

$$\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} C_n^{(\lambda)}(x) C_m^{(\lambda)}(x) = n!\lambda^{-n}e^{\lambda}\delta_{nm}, \ \lambda > 0,$$

where $\delta_{nm} = 0$ when $n \neq m$ and $\delta_{nm} = 1$ for $n = m$. Therefore any function $F \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathtt{P})$ can be uniquely represented as

$$F(x) = \sum_{n \geq 0} f_n C_n^{(\lambda)}(x), \quad f_n \in \mathbb{R}_+.$$

with the corresponding norm given by $\|f\|^2 = \sum_{n \geq 0} |f_n|^2 \lambda^n n!$.

Next, we construct the discrete multiple stochastic integral using the Wick product following similar arguments in [Pri94].

Suppose $\mathcal{P}^*$ is a set of all functionals of the form $Q(\zeta_0, \cdots, \zeta_{n-1})$ where Q is a real polynomial and $n \in \mathbb{N}$, we regard $\mathcal{P}^*$ as an algebra generated by $\left\{C_n^{(\lambda)}(\zeta_k) : k, n \in \mathbb{N}\right\}$ and define the Wick product in the following manner.

Definition 13 *The Wick product of two elements $F, G \in \mathcal{P}^*$ denoted by $F \diamond G$ is defined* *(relaxing λ for simplicity) as:*

$$(C_{n_1}(\zeta_{k_1})\cdots C_{n_d}(\zeta_{k_d}))\diamond(C_{m_1}(\zeta_{k_1})\cdots C_{m_d}(\zeta_{k_d}))$$
$$=\frac{(n+m)!}{n!m!}C_{n_1+m_1}(\zeta_{k_1})\cdots C_{n_d+m_d}(\zeta_{k_d}),$$

where for $a\in\mathbb{N}^d$, $a!=a_1!\cdots a_d!$ and $n=(n_1,\cdots,n_d)$, $m=(m_1,\cdots,m_d)$ and $k_1\neq\cdots\neq k_d$.

Let $H=l^2(\mathbb{N})$ be a space of square-summable sequences. There exists a discrete chaotic decomposition of $\mathcal{L}^2(\Omega,\mathtt{P})$ whose elements F can each be represented as a sum of multiple stochastic integrals of kernels of $H^{\circ n}=l^2(\mathbb{N})^{\circ n}$, that is,

$$F=\sum_{n=0}^{\infty}I_n(f_n), \tag{5.17}$$

where $f_n\in H^{\circ n}$, $n\in\mathbb{N}$ and $I_n(f_n)$ is the discrete multiple stochastic integral of symmetric functions of discrete variable. The stochastic integral of $f\in l^2(\mathbb{N})$ is an isometry from $H=l^2(\mathbb{N})$ to $\mathcal{L}^2(\Omega,\mathcal{F},\mathtt{P})$. This is shown in the following proposition.

Proposition 2 *Let $f\in l^2(\mathbb{N})$ and define $I_1(f):=\int_0^{+\infty}f(B_{L_t})\,\mathrm{d}B_{L_t}$. Then*

$$\mathtt{E}[I_1(f)^2]=\sum_{k,l\in\mathbb{N}}f_kf_l\mathtt{E}[\zeta_k\zeta_l]=\sum_0^{\infty}f_k^2=\|f^2\|_{\mathcal{L}^2([0,L_T],\Omega)}. \tag{5.18}$$

Proof 17 *Consider a partition $\tau_1<\cdots<\tau_{m-1}$ where $\tau_i,\ i=1,\cdots m-1$ are the jump times of B_{L_t} and let $\{t_{i,j};j=0,1,2,\cdots,n_i\}$ be a partition for each $[\tau_{i-1},\tau_i)$. Suppose $\Delta:=\max_{i,j}(t_{i,j}-t_{i,j-1})$, we have*

$$I_1(f)=\lim_{\Delta\to0}\sum_{\substack{i=1\\0\le j\le n_i}}^{i=m-1}f(B_{L_{t_{i,j}}})(B_{L_{t_{i,j+1}}}-B_{L_{t_{i,j}}})+f(B_{L_{t_{n_i}}})(B_{L_1}-B_{L_{t_{n_i}}}),$$

where $n_i = \max\{j : t_j \leq 1\}$. *Consequently*

$$\mathtt{E}[I_1(f)^2] = \lim_{\Delta\to 0} \sum_{\substack{i=1 \\ 0\leq j_1, j_2 \leq n_i}}^{M} \mathtt{E}[f(B_{L_{t_{i,j_1}}}) f(B_{L_{t_{i,j_2}}}) \times (B_{L_{t_{i,j_1+1}}} - B_{L_{t_{i,j_1}}})(B_{L_{t_{i,j_2+1}}} - B_{L_{t_{i,j_2}}})].$$

We notice that for $j_1 < j_2$ *on each* $[\tau_i, \tau_{i+1})$, *we have*

$$\mathtt{E}[f(B_{L_{t_{i,j_1}}}) f(B_{L_{t_{i,j_2}}})(B_{L_{t_{i,j_1+1}}} - B_{L_{t_{i,j_1}}})(B_{L_{t_{i,j_2+1}}} - B_{L_{t_{i,j_2}}})] = 0,$$

which is arrived at by conditioning with respect to $\mathcal{F}_{t_{i,j_2}}$ *and applying the tower property. Meanwhile for* $j_1 = j_2 = j$ *we have*

$$\begin{aligned}
&\mathtt{E}[f^2(B_{L_{t_{i,j}}})(B_{L_{t_{i,j+1}}} - B_{L_{t_{i,j}}})^2]. \\
&= \mathtt{E}[f(B^2_{L_{t_{i,j}}})\mathtt{E}[(B_{L_{t_{i,j+1}}} - B_{L_{t_{i,j}}})^2 | \mathcal{F}_{t_{i,j}}]]. \\
&= \mathtt{E}[f^2(B_{L_{t_{i,j}}})\mathtt{E}[(L_{t_{i,j+1}} - L_{t_{i,j}}) | \mathcal{F}_{t_{i,j}}]. \\
&= \mathtt{E}[f^2(B_{L_{t_{i,j}}})(L_{t_{i,j+1}} - L_{t_{i,j}})].
\end{aligned}$$

The last equation follows from the law of total expectation. The result follows immediately by combining both cases.

The discrete multiple stochastic integral $I_n(f_n)$, f_n symmetric in $l^2(\mathbb{N}^n)$ with finite support can be defined directly using the Wick product.

Definition 14 *The symmetric tensor product* $f_1 \circ \cdots \circ f_n$ *is defined as*

$$f_1 \circ \cdots \circ f_n = \frac{1}{n!} \sum_{\sigma \in \Sigma_n} f_{\sigma(1)} \otimes \cdots \otimes f_{\sigma(n)} \tag{5.19}$$

where $f_1, \cdots, f_n \in \mathcal{H}$ *and* Σ_n *is the set of all permutations of* $\{1, \cdots, n\}$*. Moreover, suppose* $g_1, \cdots, g_n \in l^2(\mathbb{N})$ *with finite supports, we have*

$$I_n(g_n \circ \cdots \circ g_n) = I_1(g_1) \diamond \cdots \diamond I_1(g_n),$$

where $I_1(g_i) = \sum_{k=0}^{\infty} g_i(k) C_1(k)$*,* $1 \leq i \leq n$*.*

The definitions above suggest the results [Prop.2 & Prop.3, [Pri94]] for the Poisson process, also hold for our choice of process B_{L_t} with similar proofs.

Lemma 15 *1. Suppose* $(e_n)_{n \in \mathbb{N}}$ *is a canonical basis in* $l^2(\mathbb{N})^{\circ n}$*. Then*

(a) For $k_1 \neq \cdots \neq k_d$ *and* $n_1 + \cdots + n_d = n$*, we have*

$$I_n(e_{k_1}^{\circ n_1} \circ \cdots \circ e_{k_d}^{\circ n_d}) = n_1! \cdots n_d! C_{n_1}(\zeta_{k_1}) \cdots C_{n_d}(\zeta_{k_d}).$$

(b) Suppose $f = \sum_{k=0}^{\infty} f_k e_k \in l^2(\mathbb{N})$ *has finite support, then*

$$I_n(f^{\circ n}) = n! \sum_{\substack{k_1 \neq \cdots \neq k_d \\ n_1 + \cdots + n_d = n \\ n_1, \cdots, n_d > 0}} f_{k_1}^{n_1} \cdots f_{k_d}^{n_d} C_{n_1}(\zeta_1) \cdots C_{n_d}(\zeta_d).$$

(c) If $f_n \in l^2(\mathbb{N}^n)$*,* $g_m \in l^2(\mathbb{N}^m)$ *are symmetric with finite supports, then*

$$I_n(f_n) \diamond I_m(g_m) = I_{n+m}(f_n \circ g_m).$$

2. Let $\mathcal{D}_n = \{(k_1, \cdots, k_n) \in \mathbb{N}^n : \exists\, i \neq j \text{ such that } k_i = k_j\}$ *which represents the diagonals in* $\mathbb{N}^n$ *and let* $\mathcal{X}_n = \mathbb{N}^n \setminus \mathcal{D}_n$*. Suppose* $f_n \in l^2(\mathbb{N}^n)$ *and* $g_m \in l^2(\mathbb{N}^m)$ *are* *symmetric with finite supports, then*

$$\langle I_n(f_n), I_m(g_m)\rangle_{\mathcal{L}^2(\Omega)} = \begin{cases} n!\langle f_n, g_m\rangle_{l^2(\mathcal{X}_n)} + (n!)^2\langle f_n, g_m\rangle_{l^2(\mathcal{D}_n)} & \text{if } n = m \\ \langle I_n(f_n), I_m(g_m)\rangle_{\mathcal{L}^2(\Omega)} = 0 & \text{if } n \neq m. \end{cases}$$

Proof 18 *See [Pri94].*

Lemma 16 *Let $\mathcal{C}_n = \{I_n(f_n) : f_n \in l^2(\mathbb{N}^{\circ n})\}$ define chaos of order $n \in \mathbb{N}$ in $\mathcal{L}^2(\Omega)$. Then $\mathcal{L}^2(\Omega)$ has a chaotic decomposition:*

$$\mathcal{L}(\Omega, \mathcal{F}, \mathtt{P}) = \bigoplus_{n=0}^{\infty} \mathcal{C}_n.$$

Moreover, if K_n is the tensor product $\mathcal{H}^{\circ n}$, endowed with the norm

$$|||f_n|||_n^2 = n!\langle f_n, f_n\rangle_{l^2(\mathcal{X}_n)} + (n!)^2\langle f_n, f_n\rangle_{l^2(\mathcal{D}_n)},$$

equivalent to $\|\cdot\|_{l^2(\mathbb{N})}$ then the Fock space $\mathcal{F}(\mathcal{H}) := \bigoplus_{n=0}^{\infty} K_n$ is isometrically isomorphic to $\mathcal{L}^2(\Omega)$.

Proof 19 *The $\mathcal{C}_n$'s are orthogonal according to Lemma 15. Secondly, Q is dense in $\mathcal{L}^2(\Omega)$ since the polynomials of $\mathcal{P}^*$ are dense in $\mathcal{L}^2(\mathbb{R}_+, \frac{\lambda^x}{x!}e^{\lambda}\,\mathrm{d}x)$: Suppose $F \in \mathcal{L}^2(\Omega)$ and $\mathtt{E}[FQ_n(\zeta_0)\cdots Q_n(\zeta_n)] = 0,$, for any $Q_0, \cdots, Q_n$, $n \in \mathbb{N}$, then $\mathtt{E}[F|\zeta_0, \cdots, \zeta_n] = 0$, $n \in \mathbb{N}$. This implies*

$$\lim_{n\to\infty} \mathtt{E}[F|\zeta_0, \cdots, \zeta_n] = F \quad \mathtt{P} - a.s.,$$

since $\mathtt{E}[F|\zeta_0, \cdots, \zeta_n]$ is a discrete-time martingale. Therefore $F = 0$, $\mathtt{P} - a.s.$ and $\mathcal{P}^*$ is dense in $\mathcal{L}^2(\Omega)$.

The annihilation operator defined in (5.12) has an equivalent in the discrete chaotic decomposition of $\mathcal{L}^2(\Omega, \mathcal{F}, \mathtt{P})$ given by

$$\mathrm{D}_k(I_n(f_n)) = \sum_{p=0}^{n-1} \frac{n!}{p!} I_p\Big(f_n(*, \underbrace{k, \cdots, k}_{n-p \text{ times}})\Big); \; k \in \mathbb{N}. \tag{5.20}$$

Moreover, the following lemma holds.

Lemma 17 *Suppose $\mathcal{U}$ denotes a dense set of elements of $u \in \mathcal{L}^2(\Omega) \otimes l^2(\mathbb{N})$ such that $u_k = \zeta_k h_k$, $k \in \mathbb{N}$ where $h : \mathbb{N} \to Q$ has finite support in $\mathbb{N}$ and define an operator $\delta : \mathcal{U} \to \mathcal{L}^2(\Omega)$ by*

$$\delta(u) = -\sum_{k=0}^{\infty} (u_k + \mathrm{D}_k\, u_k),$$

Then for any $F \in Dom(\mathrm{D})$ and $u \in Dom(\delta)$, we have

$$\mathtt{E}[(\mathrm{D}F, u)_{l^2(\mathbb{N})}] = \mathtt{E}[F\delta(u)], \tag{5.21}$$

where Dom represents domain and the operators D and δ are also closable and adjoint to each other.

Proof 20 *See [Pri94].*

The above results can be extended from $\mathcal{L}^2(\Omega) \otimes l^2(\mathbb{N})$ to $\mathcal{L}^2(\Omega) \otimes \mathcal{L}^2(\mathbb{R}_+)$ yielding:

$$\int_0^{L_T} F u(B_\tau)\, \mathrm{d}B_\tau = F \int_0^{L_T} u(\tau)\, \mathrm{d}B_\tau - \int_0^{L_T} (\mathrm{D}_t\, F) u(\tau)\, \mathrm{d}\tau, \quad t \leq \tau. \tag{5.22}$$

This observation is explained in [Pri94]. This in turn leads to the following duality relation:

$$\mathtt{E}\left[\int_0^{L_T} (\mathrm{D}_t\, F) u(\tau)\, \mathrm{d}\tau\right] = \mathtt{E}\left[F \int_0^{L_T} u(\tau)\, \mathrm{d}B_\tau\right], \quad t \leq \tau. \tag{5.23}$$

The duality formula is an important tool for finding alternative representations of derivatives of expectations of irregular functions. We discuss this later.

5.3.2 Malliavin Derivative of Solutions to Subordinated SDEs

In the following, K is defined as a Hilbert space, $D^{1,2}(K)$ as a Sobolev space of K-valued functions associated with the H-derivative, $\mathcal{L}_2(H;K)$ as a total set of a K-valued linear operator of Hilbert-Schmidt class on H, $\mathcal{L}^{1,2}(\mathrm{d}B_{L_t};K)$ as the total set of $(\mathcal{F}_t)$-predictable $(\mathbb{R}\times K)$-valued functions σ such that for $\sigma(t,X)\in D^{1,2}(\mathbb{R}\times K)$, $t\in[0,T]$ with

$$\|\sigma\|_{\mathcal{L}(\mathrm{d}B_{L_t};K)} := \mathtt{E}\left[\left|\mathrm{D}_h\,\sigma(t-,X)\right|^2_{\mathcal{L}_2(H;\mathbb{R}\times K)}\,\mathrm{d}L_t\right]^{1/2} < \infty, \tag{5.24}$$

where h is given in (5.8). We denote $\mathcal{L}(\mathrm{d}t;K)$ as the total set of $(\mathcal{F}_t)$-predictable $(\mathbb{R}\times K)$-valued functions b satisfying $b(t,X)\in D^{1,2}(\mathbb{R}\times K)$, $t\in[0,T]$ with

$$\|b\|_{\mathcal{L}(\mathrm{d}t;K)} := \int_0^T \left|\mathrm{D}_h\,b(t-,X)\right|^2_{\mathcal{L}_2(H;K)}\,\mathrm{d}t < \infty. \tag{5.25}$$

Let $(\Omega,\mathcal{F},\mathtt{P})$ represent a joint probability space introduced in the previous sections and consider the following SDE:

$$\mathrm{d}X_t = b(t,X_t)\,\mathrm{d}t + \sigma(t,X_t)\,\mathrm{d}B_{L_t};\quad X_0 = x, \tag{5.26}$$

A scenario of (5.26) with Lipschitz coefficients and standard Brownian motion has been discussed in [FLL+99] and [BN13] to compute financial Greeks. The case of non-Lipschitz coefficients with subordinated Brownian motion is handled in [BM06], [SX14] and much less related in [CF07] and [DNØP08]. For non-Lipschitz with standard Brownian motion, (see for instance [BnDMBP15]). Throughout the current chapter, we shall consider an SDE with coefficients that are continuously differentiable with bounded Lipschitz derivatives and subordinated Brownian motion will be the source of randomness.

Proposition 3 *Suppose in (5.26) the SDE has Lipschitz coefficients with linear growth and assume $\sigma(t, X) \in D^{1,2}(\mathbb{R} \times K)$ and $b(t, X) \in D^{1,2}(\mathbb{R} \times K)$. Then the solution X_t to (5.26) exists, is unique and belongs to $D^{1,2}$ for all $t \in [0, T]$.*

Proof 21 *The proof is based on Picard's succesive approximation and it follows similar steps of [Thm 3.1 and Prop 2.1, [Kus10]] .*

Let $[\cdot, \cdot]$ denote the dot product endowed on $\mathcal{H}$ in (5.8). A representation of the derivative of X_t follows in the following proposition.

Proposition 4 *Denote the directional derivative of X_t by $\mathrm{D}_r X_t[h]$, $r \leq t$ where D_r is the Malliavin derivative operator. Then from (5.26) we have*

$$\mathrm{D}_r X_t[h] = \int_r^t \mathrm{D}_r b(s, x)[h] \, \mathrm{d}s + \int_r^t \mathrm{D}_r \sigma(s-, x)[h] \, \mathrm{d}B_{L_s} + \int_r^t \sigma(s-, x) \, \mathrm{d}h(L_s). \quad (5.27)$$

If we assume $\{h\}$ is a complete orthonormal basis in $\mathcal{H}$, then

$$\begin{aligned} \mathrm{D}_r X_t &= \int_r^t \mathrm{D}_r b(s) \, \mathrm{d}s + \int_r^t \mathrm{D}_r \sigma(s-) \, \mathrm{d}B_{L_s} \\ &\quad + \sum_{i=1}^{\infty} h^i(s) \otimes \int_{L_r}^{L_t} \sigma(S_s-) \dot{h}^i(s) \, \mathrm{d}s. \end{aligned} \quad (5.28)$$

Moreover, if $\sigma \equiv 1$, then

$$\mathrm{D}_r X_t = \int_r^t b'(X_s) \, \mathrm{D}_r X_s \, \mathrm{d}s + \sum_{i=1}^{\infty} h^i(t) \otimes \int_{L_r}^{L_t} \dot{h}^i(s) \, \mathrm{d}s. \quad (5.29)$$

Proof 22 *First, represent (5.26) in its integral form:*

$$X_t = x + \int_0^t b(\tau, X_\tau) d\tau + \int_0^t \sigma(\tau, X_\tau) dB_{L_\tau}. \quad (5.30)$$

Then apply the product rule and (5.12) to the second term on the RHS of (5.30) to obtain (5.27). If $\{h\}$ is an orthogonal basis, we can express σ as

$$\sigma = \sum_{i=1}^{\infty} \frac{[\sigma, h^i]}{\|h^i\|} h^i, \tag{5.31}$$

where $h = \|h\|\hat{h}$ and $\hat{h}$ is a unit vector of h. For h orthonormal gives (5.28).

If $\sigma \equiv 1$ the last term of (5.29) is zero and by the Grönwall's inequality, we have:

$$\mathrm{D}_r X_t = \exp\left(\int_r^t b'(X_s)\, \mathrm{d}s \right). \tag{5.32}$$

Also note from (5.30) that the first variation process can be deduced as:

$$\frac{\partial X_t}{\partial x} = \exp\left(\int_0^t b'(X_s)\, \mathrm{d}s \right). \tag{5.33}$$

Combining (5.32) and (5.33) results into the following useful relation

$$\mathrm{D}_r X_t = \frac{\partial X_t}{\partial x} \exp\left(- \int_0^r b'(X_s)\, \mathrm{d}s \right). \tag{5.34}$$

Alternatively,

$$\frac{\partial X_t}{\partial x} = \mathrm{D}_r X_t \frac{\partial X_r}{\partial x}, \quad r \leq t. \tag{5.35}$$

5.4 BEL Formula for Subordinated Stochastic Differential equations

Bismut-Elworthy-Li formula for general Lévy processes is studied in [CF07]. We derive representations for subordinated Brownian motion based on [EL94].

Proposition 5 *Let X_t be the solution to (5.26) on the horizon $[0, T]$ where $L_t \equiv t$ (see for instance [CF07, BnDMBP15]) and let $\Phi : \mathbb{R} \to \mathbb{R}$ denote some bounded function. Suppose we can define a functional $V_t(X_T)$ of X_T by*

$$V_t(X_T) = \mathtt{E}[\Phi(X_T]; \quad X_0 = x. \tag{5.36}$$

Then the derivative of V with respect to x is given by (5.5).

Proof 23 *Apply the classical chain rule on $\mathtt{E}[\partial_x \Phi(X_T)]$ and use the relation (5.35) followed by the chain rule in the Malliavin sense. Finally apply the duality relation (5.23), in that order. A similar proof is provided in [Stu04] by using the identity $D_t X_T = J_T J_t^{-1} \sigma(t, X_{t-}) 1_{\{t \leq T\}}$, where $J_t := \frac{\partial X_t}{\partial X_0}$, $X_0 = x$.*

Equation (5.5) is Bismut-Elworthy-Li formula for Geometric Brownian motion.

Proposition 6 *Suppose (5.30) has Lipschitz coefficients. Let $R(t, X_t)$ denote the right inverse of $\sigma(t, X_t)$ where $\sigma(t, X_t)$ is elliptic. For any function $\Phi \in C_b^1(\mathbb{R})$ and $h \in \mathbb{R}$, we have (the x argument is relaxed for simplicity)*

$$\mathrm{D}_h \, \mathtt{E}[\Phi(X_t)] = \frac{1}{L_t} \mathtt{E}\left[\Phi(X_t) \int_0^t R(\tau) \cdot \mathrm{D}_h \, X_\tau \, \mathrm{d}B_{L_\tau}\right], \tag{5.37}$$

where

$$\mathrm{D}_h \, X_t = h(t) + \int_0^t b'(s, X_s) \cdot \mathrm{D}_h \, X_s \, \mathrm{d}s. \tag{5.38}$$

Proof 24 *Working backwards and using the results obtained above we have*

$$\mathrm{D}_h \, \mathtt{E}[\Phi(X_T)] \;\; = \;\; \mathtt{E}[\mathrm{D}_r \, \Phi(X_t)[h]] = \mathtt{E}[\Phi'(X_t) \, \mathrm{D}_r \, X_t[h]].$$

Next we apply the duality relation (5.23), equation (5.34), Grönwall's inequality and

(5.29) for some arbitrary h not necessarily an orthonormal basis and $\sigma \geq 1$ in that order, where we have chosen $a_t = T$ for all $t \in [0, T]$..

We provide a detailed analysis on the above result in the following section.

5.4.1 Main Results

This section presents the main results of the chapter. The idea is to extend the results by [FLL+99] to a subordinated stochastic differential equation model by deriving the first and second order derivative representation for the expectation of a function that is not necessarily regular. Specifically, the idea is to by-pass the derivative of the expected (irregular) function by introducing a weight term in the form of an integral with respect to subordinated Brownian motion. The results in this section are employed in the following section to estimate the Greeks using Monte Carlo simulations.

In this section, operators δ and D will be used interchangeably to represent weak derivatives, and J_t shall denote the first variation process given by

$$J_t = \frac{\partial X_t}{\partial x}, \text{ and } J_0 = \frac{\partial X_t}{\partial x}\bigg|_{t=0}. \tag{5.39}$$

$\mathcal{L}$ shall denote a space of bounded integrable functions, $(Q_t : t \geq 0)$ shall denote the semigroup of the solution X_t to (5.30). Lastly, $\{S_t\}_{t \geq 0}$ shall denote a non-decreasing cádlág α-stable process and $\{L_t\}_{t \geq 0}$ its inverse with $\alpha \in (0, 1]$. We require X_t to be complete to enforce some integrability conditions on $\mathrm{D}X_t$. Let $\mathcal{L}(\mathbb{R})$ denote the space of integrable functions on $\mathbb{R}$. Then we have the following corollary.

Corollary 1 *Let $U : \mathbb{R} \to \mathcal{L}(\mathbb{R})$ with bounded first derivative such that $\delta Q_t(U) : \mathbb{R} \to \mathcal{L}(\mathbb{R})$ and define $U_X := (\mathrm{d}f)_X = \mathrm{D}f(X_t)$. Then the weak derivative of Q_t with respect* *to the initial state x is given by*

$$\delta Q_t(U)_x[J_0] = \mathrm{E}[U_X[J_t]] = \mathrm{E}[\mathrm{D}f(X_t)[J_t]], \tag{5.40}$$

provided the last term exists. Moreover

$$(\delta Q_t(\mathrm{d}f))_x[J_0] = \mathrm{d}(Q_t f)_x[J_0]. \tag{5.41}$$

Since X_t is non-explosive, the weak derivative δQ_t is well defined.

Proof 25 *This follows directly from semigroup arguments and the application of a weak chain rule.*

Corollary 2 *Suppose $X_t \in \mathbb{R}$ is non-degenerate and elliptic, there exist an inverse $R(t, X_t)$ of $\sigma(t, X_t)$ smooth in X_t such that $\left|R(X)[Y]\right| \leq \varepsilon |Y|^2$ for all $X, Y \in \mathbb{R}$ for some $\varepsilon > 0$. Moreover, if*

$$\int_0^{L_t} \mathrm{E}[\,|J_{S_\tau}|^2]\,\mathrm{d}\tau < \infty. \tag{5.42}$$

then,

$$\int_0^{L_t} \mathrm{E}[\,\left|R(S_\tau, X_{S_\tau}) J_{S_\tau}\right|\,]\,\mathrm{d}\tau < \infty. \tag{5.43}$$

Proof 26 *Recall that the result holds for the case of continuous processes [CF07, EL94]. We can therefore employ similar arguments of partitioning as in the second part of the Proof of Lemma 13 and apply similar steps as in the continuous case but piece-wise, to arrive at the required result.*

Theorem 6 *Let $\Phi : \mathbb{R} \to \mathbb{R}$ with its first derivative bounded and continuous:*

$$\delta Q_t(\mathrm{d}\Phi) = \mathrm{d}(Q_t \Phi) \ \textit{a.s.}\ t \geq 0, \tag{5.44}$$

Moreover, for $x \in \mathbb{R}$, $T > 0$, the derivative with respect to x is given by

$$\partial_x \mathtt{E}[\Phi(X_T)] = \frac{1}{L_T}\mathtt{E}\left[\Phi(X_T)\int_0^{L_T} R(S_\tau, X_{S_\tau})J_{S_\tau}\,\mathrm{d}B_\tau\right], \tag{5.45}$$

where $\int_0^{L_T} R(S_\tau, X_{S_\tau})J_{S_\tau}\,\mathrm{d}B_\tau,\ T \geq 0$ is a martingale.

Proof 27 *From Corollary 2, let $T > 0$, applying Itô's formula to*

$$(t, X_t) \mapsto Q_{T-t}\Phi(X_t),\ 0 \leq t < T, \tag{5.46}$$

yields

$$Q_{T-t}\Phi(X_t) = Q_T\Phi(x) + \int_0^t \mathrm{d}(Q_{T-\tau}\Phi)_{X_\tau}\sigma(\tau, X_\tau)\,\mathrm{d}B_{L_\tau},\ \textit{for } t \in [0, T]. \tag{5.47}$$

As $t \to T$, and applying the knowledge from Lemmas 14 and 5.10 yields

$$\Phi(X_T) = Q_T\Phi(x) + \int_0^{L_T} \mathrm{d}(Q_{S_T - S_\tau}\Phi)_{X_{S_\tau}}(\sigma(S_\tau, X_{S_\tau})\,\mathrm{d}B_\tau. \tag{5.48}$$

Multiplying (5.48) by a martingale $\int_0^{L_T} R(S_\tau, X_{S_\tau})(J_{S_\tau})\,\mathrm{d}B_\tau$ yields

$$\begin{aligned}
\mathtt{E}\left[\Phi(X_T)\int_0^{L_T} R(S_\tau, X_{S_\tau})(J_{S_\tau})\,\mathrm{d}B_\tau\right] &= \mathtt{E}\left[\int_0^{L_T} \mathrm{d}(Q_{S_T-S_\tau}\Phi)_{X_{S_\tau}} J_{S_\tau}\,\mathrm{d}\tau\right]. \\
&= \mathtt{E}\left[\int_0^{L_T} ((\delta Q_{S_T-S_\tau})(\mathrm{d}\Phi))_{X_{S_\tau}}(J_{S_\tau})\,\mathrm{d}\tau\right]. \\
&= \int_0^{L_T} ((\delta Q_{S_\tau})((\delta Q_{S_T-S_\tau})(\mathrm{d}\Phi)))_x(J_0)\,\mathrm{d}\tau. \\
&= \int_0^{L_T} (\delta Q_{S_T}(\mathrm{d}\Phi))_x(J_0)\,\mathrm{d}\tau. \\
&= L_T\delta Q_T(\mathrm{d}\Phi)_x(J_0).
\end{aligned} \tag{5.49}$$

Since $J_0 = 1$ and $\delta Q_T(\mathrm{d}\Phi)_x = \mathtt{E}[(\mathrm{d}\Phi)_x] = \partial_x\mathtt{E}[\Phi]$. The required result follows.

Corollary 3 *Suppose $X_t \in \mathbb{R}^2$, $t \geq 0$ and indexes $0 \leq j,k \leq m$, then*

$$\int_0^{L_T} \mathtt{E}[\, |\, \mathrm{D}X_{S_\tau,x}(J_0^j)|^2]\, \mathrm{d}\tau \leq \varepsilon |J_0^j|^2,\ \varepsilon > 0. \tag{5.50}$$

and

$$\sup_{0 \leq S_\tau \leq t} \sup_{x \in \mathbb{R}^2} \mathtt{E}[|\, \mathrm{D}^2 X_{S_\tau,x}(J_0^j, J_0^k)|] \leq \varepsilon |J_0^j||J_0^k|, \tag{5.51}$$

and

$$\sup_{0 \leq S_\tau \leq t} \sup_{x \in \mathbb{R}^2} \mathtt{E}[|\, \mathrm{D}X_{S_\tau,x}|] \leq \varepsilon. \tag{5.52}$$

where $|\cdot|$ denotes the Euclidean norm.

Proof 28 *The proof follows directly by applying (5.38) where $h \equiv J_0$.*

Theorem 7 *Let $\Phi : \mathbb{R}^2 \to \mathbb{R}$ be such that its first and second derivatives are bounded and continuous and,*

$$\mathrm{d}(Q_t \Phi) = \delta Q_t(\mathrm{d}\Phi) \ a.s.\ t \geq 0 \tag{5.53}$$

such that for almost all $x_j, x_k \in x$ for $0 \leq j,k \leq 2$,

$$\begin{aligned} \mathrm{D}^2 Q_t \Phi(x)(J_0^j)(J_0^k) &= \mathtt{E}\Big[\mathrm{D}^2 \Phi(X_t)(\mathrm{D}X_{S_t,x} J_0^j, \mathrm{D}X_{S_t,x} J_0^k)\Big] \\ &+ \mathtt{E}\Big[\mathrm{D}\Phi(X_t)\sigma(S_t, X_{S_t})(J_0^j, J_0^k)\Big], \end{aligned} \tag{5.54}$$

where

$$X_0 = x,\ J_0^j = \left.\frac{\partial X_{S_t}}{\partial x_j}\right|_{t=0} \quad \text{and} \quad J_0^k = \left.\frac{\partial X_{S_t}}{\partial x_k}\right|_{t=0}. \tag{5.55}$$

Then

$$\begin{aligned}
\frac{\partial^2}{\partial x_j x_k}\mathtt{E}[\Phi(X_T)] &= \frac{4}{L_T^2}\mathtt{E}\Bigg[\Phi(X_T)\int_{L_T/2}^{L_T}\left(R(S_t,X_{S_t-})\frac{\partial X_{S_t-}}{\partial x_j}\right)\mathrm{d}B_t\\
&\times \int_0^{L_T/2}\left(R(S_t,X_{S_t-})\frac{\partial X_{S_t-}}{\partial x_k}\right)\mathrm{d}B_t\Bigg]\\
&+ \frac{2}{L_T}\mathtt{E}\Bigg[\Phi(X_T)\int_0^{L_T/2}\left(\mathrm{D}R(S_t,X_{S_t-})\frac{\partial X_{S_t-}}{\partial x_j}\frac{\partial X_{S_t-}}{\partial x_k}\right)\mathrm{d}B_t\Bigg]\\
&+ \frac{2}{L_T}\mathtt{E}\Bigg[\Phi(X_T)\int_0^{L_T/2}\left(R(S_t,X_{S_t-})\frac{\partial^2 X_{S_t-}}{\partial x_j x_k}\Big|_{t=0}\right)\mathrm{d}B_t\Bigg].
\end{aligned}$$

Proof 29 *From equation (5.54) we deduce*

$$\begin{aligned}
L_T\Big(\mathrm{D}^2\,Q_T\Phi(x)(J_0^j,J_0^k)\Big) &= \mathtt{E}\Bigg[\mathrm{D}\Phi(X_T)(J_T)\int_0^{L_T}R(S_\tau,X_{S_\tau-})(J_{S_\tau})\,\mathrm{d}B_\tau\Bigg]\\
&- \mathtt{E}\Bigg[\int_0^{L_T}\mathrm{D}(Q_{S_T-S_\tau}\Phi)(X_{S_\tau-})\\
&\times (\mathrm{D}\sigma(S_\tau,X_{S_\tau-})(J_{S_\tau}^k)R(S_\tau,X_{S_\tau-})(J_{S_\tau}^j)\,\mathrm{d}\tau\Bigg]\\
&+ \mathtt{E}\Bigg[\int_0^{L_T}(Q_{S_T-S_\tau}\Phi)(X_{S_t-})(\mathrm{D}^2\,X_{S_t-,x})(J_0^j,J_0^k)\,\mathrm{d}\tau\Bigg].
\end{aligned}$$

Suppose $L_T=\frac{1}{2}L_t$ *and consider* $0\le\tau\le t/2$ *then*

$$\begin{aligned}
\mathrm{D}^2\,Q_{t/2}\Phi(x)(J_0^j,J_0^k) &= \frac{4}{L_t^2}\mathtt{E}\Bigg[\Phi(X_t)\int_{L_t/2}^{L_t}R(S_\tau,X_{S_\tau})J_{S_\tau}^j\,\mathrm{d}B_\tau\\
&\times \int_0^{L_t/2}R(S_\tau,X_{S_\tau-})J_{S_\tau}^k\,\mathrm{d}B_\tau\Bigg]\\
&- \frac{2}{L_t}\mathtt{E}\Bigg[\int_0^{L_t/2}\mathrm{D}(Q_{S_{t/2}-S_\tau}\Phi)(X_{S_\tau-})\\
&\times (\mathrm{D}\sigma(S_\tau,X_{S_\tau-})(J_{S_\tau}^j)(R(S_\tau,X_{S_\tau-})(J_{S_\tau}^k)\,\mathrm{d}\tau\Bigg]\\
&+ \frac{2}{L_t}\mathtt{E}\Bigg[\int_0^{L_t/2}\mathrm{D}(Q_{S_{t/2}-S_\tau}\Phi)(X_{S_\tau})(\mathrm{D}^2\,X_{S_\tau,\tau})(J_0^j,J_0^k)\,\mathrm{d}\tau\Bigg].
\end{aligned}$$

Applying Itô formula to $\{Q_{t-\tau}\Phi(X_\tau) : 0 \le L_\tau < L_t\}$ *at* $L_\tau = L_t/2$*, yields*

$$Q_{t/2}\Phi(X_t) = Q_t\Phi(x) + \int_0^{L_t/2} \mathrm{D}(Q_{S_{t}/2-S_\tau}\Phi)(X_{S_\tau})(\sigma(S_\tau, X_{S_\tau -})\,\mathrm{d}B_\tau. \tag{5.56}$$

Multiply (5.56) by $\int_0^{L_t/2} \mathrm{D}R(S_\tau, X_{S_\tau})J^j_{S_\tau}J^k_{S_\tau}\,\mathrm{d}B\tau$ *and* $\int_0^{L_t/2} R(S_\tau, x)(J^j_0, J^k_0)\,\mathrm{d}B_\tau$*. Next, taking expectations and applying the identity (see [EL94])*

$$\mathrm{D}\sigma J^j R J^k + \sigma\,\mathrm{D}R J^j J^k = 0, \tag{5.57}$$

yields the required result.

5.5 Applications

This section is dedicated to estimating the Delta, Gamma and Vega from two stochastic models of the asset price namely; the subordinated stochastic differential equation (SSDE) and the Geometric Brownian motion (GBm).

5.5.1 Hedging Discontinuous-Payoff type Options in Black-Scholes Framework

We focus only on the digital option but the analysis could be extended to other discontinuous-payoff type or irregular payoff options (see [FLL+99] for instance). The Delta and Gamma follow directly from Theorem 6 and Theorem 7 respectively.

Let X_t satisfy the stochastic differential equation

$$\mathrm{d}X_t = rX_t dL_t + \sigma X_t\,\mathrm{d}B_{L_t}; \quad X_0 = x, \tag{5.58}$$

with the solution given by

$$X_t = x\exp(rL_t + \sigma B_{L_t}). \tag{5.59}$$

Figure 5.1 shows the dynamics of solutions to SSDE and GBm. The Greeks are com-

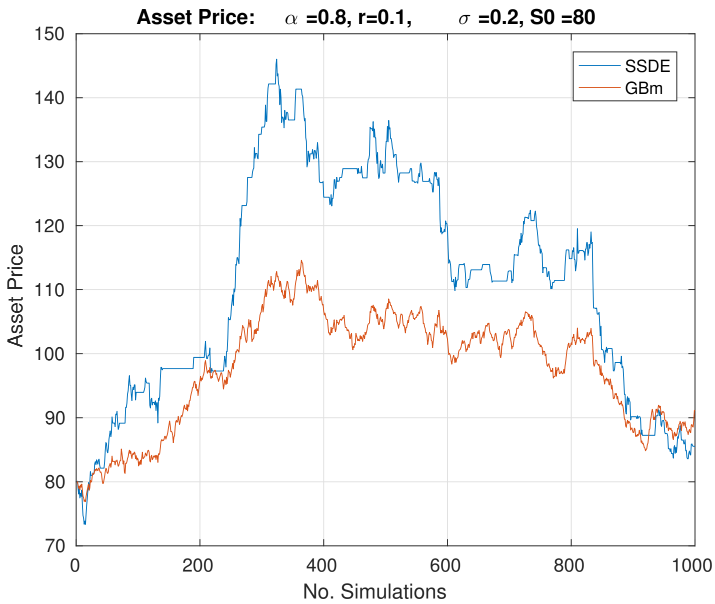


Fig. 5.1: Price evolution from SSDE and GBm models

puted by conditioning on $L_T = \tau$, as follows :

Delta

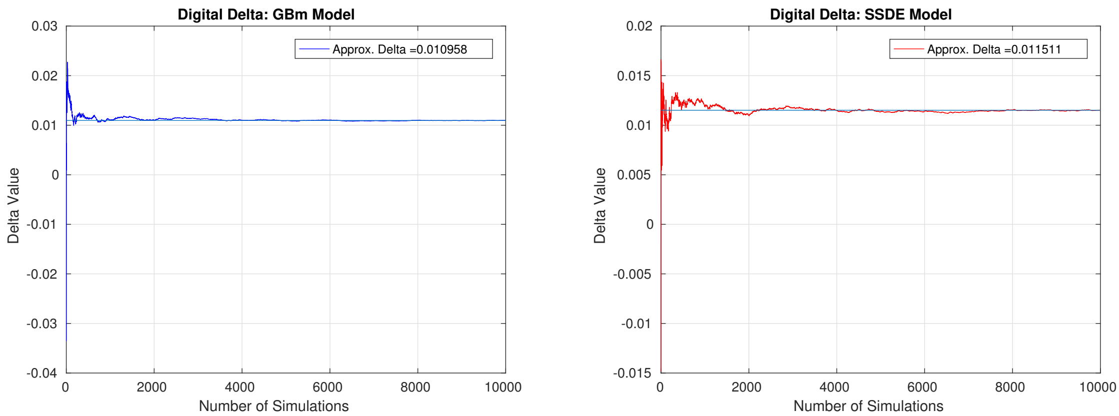
According to Theorem 6

$$\partial_x \mathtt{E}[\Phi(X_T)] = \frac{1}{\tau}\mathtt{E}\left[\Phi(X_T)\int_0^T \frac{1}{\sigma X_t}\frac{\partial X_t}{\partial x}\,\mathrm{d}B_{L_t}\right]. \tag{5.60}$$

Assume a discounted payoff of a digital option i.e. $\Phi(X_T) = e^{-rT}\mathbf{1}_{E<X_T}$ where E is the strike price, then we can express the Delta by

$$\partial_x \mathtt{E}[\Phi(X_T)] = \mathtt{E}\left[e^{-rT}\mathbf{1}_{\{E<X_T\}}\frac{B_{L_T}}{x\sigma\tau}\right]. \tag{5.61}$$

Figure 5.2 shows the digital option delta from both the SSDE and GBm models.



(a) Delta of a digital from GBm

(b) Delta of a digital from SSDE

Fig. 5.2: Digital option delta: $\alpha = 0.8$, $r = 0.1$, $\sigma = 0.2$, $S_0 = 110$, $E = 100$, $T = 1.$.

Observe that the delta from SSDE is slightly higher due to existence of jumps and its convergence is slightly slower.

Gamma

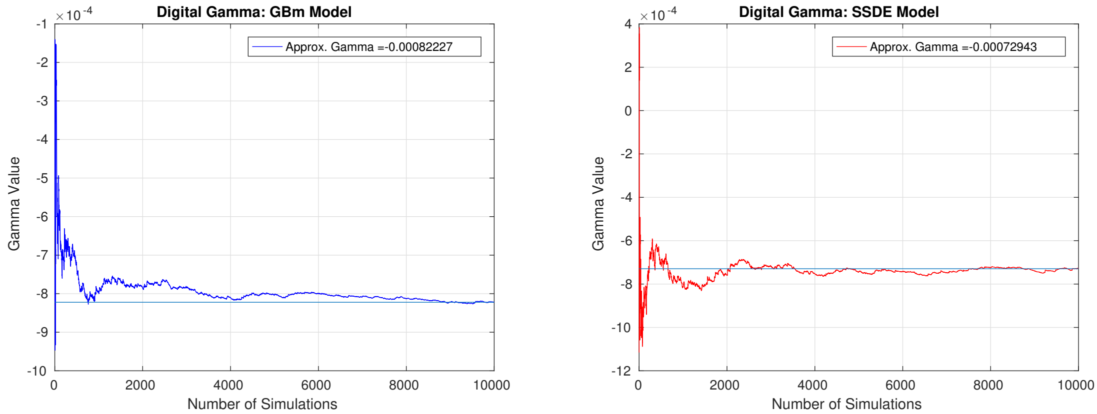
Consider $n = 2$ and let $X_t^1 = X_t^2$ in Theorem 7 then we deduce

$$\begin{aligned}\frac{\partial^2}{\partial x^2}\mathtt{E}[\Phi(X_T] &= \frac{4e^{-rT}}{x^2\sigma^2\tau^2}\mathtt{E}[\Phi(X_T)B_{L_{T/2}}(B_{L_T} - B_{L_{T/2}})] \\ &- \frac{2e^{-rT}}{x^2\sigma\tau}\mathtt{E}[\Phi(X_T)B_{L_{T/2}}] \quad (5.62)\\ &+ \frac{2e^{-rT}}{x^2\sigma\tau}\mathtt{E}[\Phi(X_T)B_{L_{T/2}}]. \quad (5.63)\end{aligned}$$

Suppose $L_T/2 \equiv L_T$. We obtain a similar expression of the Gamma as in the case of continiuious Brownian motion by applying the identity $\mathtt{E}[B_{2L_T} B_{L_T}] = \mathtt{E}[B_{L_T}(B_{2L_T} - B_{L_T}) + B^2_{L_T}]$ and simplying. That is

$$\frac{\partial^2}{\partial x^2}\mathtt{E}[\Phi(X_T)] = e^{-rT}\mathtt{E}\left[\Phi(X_T)\frac{1}{x^2\sigma\tau}\left(\frac{B^2_{L_T}}{\sigma\tau} - B_{L_T} - \frac{1}{\sigma}\right)\right]. \tag{5.64}$$

Figure 5.3 shows digital option gamma from the SSDE and GBm models. Again we



(a) Gamma of a digital from GBm (b) Gamma of a digital from SSDE

Fig. 5.3: Digital option gamma: $\alpha = 0.8$, $r = 0.1$, $\sigma = 0.2$, $S_0 = 110$, $E = 100$, $T = 1$.

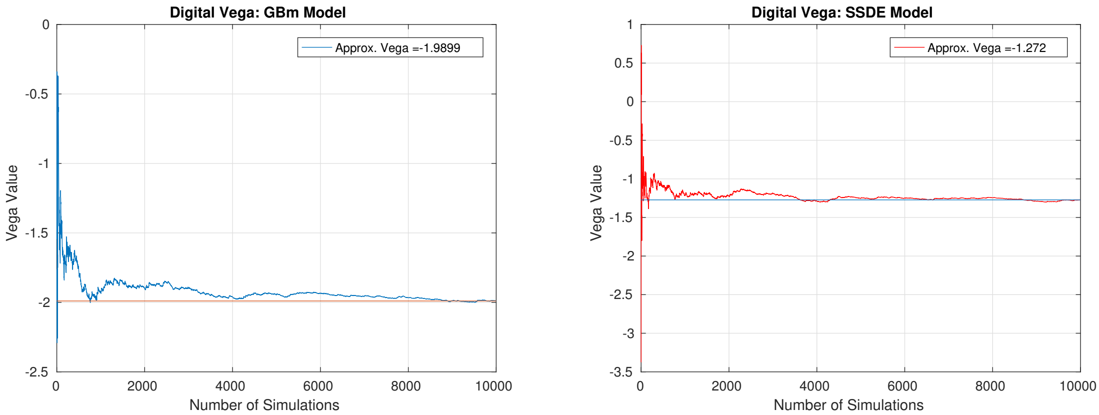
observe that the gamma from SSDE is slightly higher than that from GBm.

5.5.2 Vega

The Vega can be deduced similarly by using integration by parts. That is

$$\frac{\partial}{\partial \sigma}\mathtt{E}[\Phi(X_T)] = e^{-rT}\mathtt{E}\left[\Phi(X_T)\left(\frac{B^2_{L_T}}{\sigma\tau} - B_{L_T} - \frac{1}{\sigma}\right)\right]. \tag{5.65}$$

Figure 5.4 shows the vega from the SSDE and GBm models. Note that the Vega from



(a) Vega of a digital from GBm

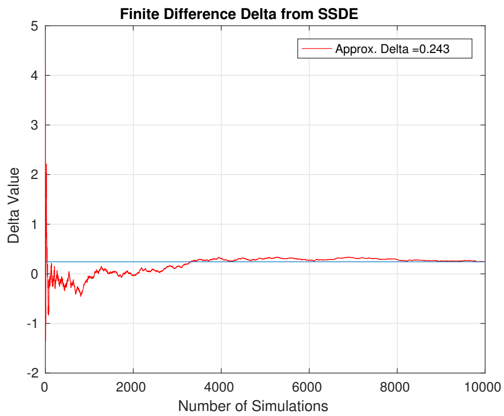
(b) Vega of a digital from SSDE

Fig. 5.4: BEL formula with $\alpha = 0.8,\ r = 0.1,\ \sigma = 0.2,\ S_0 = 110,\ E = 100.$

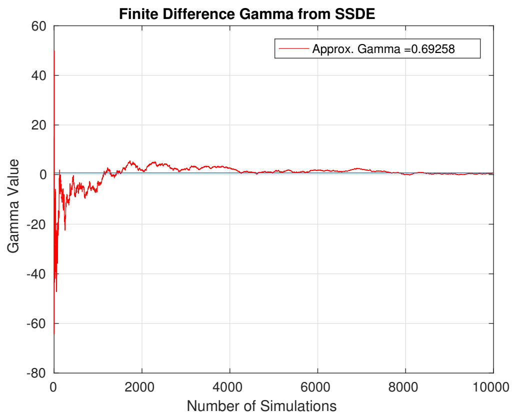
SSDE is slightly higher. Observe that despite fact the SSDE model has jumps, the convergence rate for the estimation of the Greeks from the model is as good as in the GBm model.

As a matter of interest, we apply the finite difference method on the subordinated Brownian motion model to estimate the Greeks for a call option from the SSDE model. Recall from [FLL+99] that the finite difference method is recommended for computing the Greeks from European options compared to the BEL formula, it performs better in this case. Figure 5.5 shows the estimation of the Greeks of a European call option using the SSDE model.

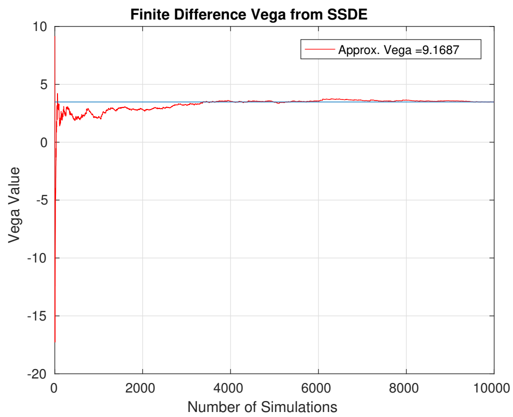
$$\partial_x \mathtt{E}[\Phi(X_T)] = \mathtt{E}\left[e^{-rT}\max(X_T - E, 0)\frac{B_{L_T}}{x\sigma L_T}\right]. \tag{5.66}$$



(a) Call Delta from SSDE



(b) Call Gamma from SSDE



(c) Call Vega from SSDE

Fig. 5.5: Finite difference method for Call Greeks from SSDE with $\alpha = 0.7,\ r = 0.1,\ \sigma = 0.2,\ S_0 = 110,\ E = 100.$

5.6 Summary

We extended the integration by parts formula approach to computing the Greeks of options with discontinuous payoffs presented in [FLL+99] to markets with jumps. As an application, we estimated Greeks from the SSDE model and observed that BEL formula still performs well for SSDE as in the continuous diffusion models. As a concrete practical application, our model can be applied by investors in emerging/illiquid markets to construct hedge portfolios.

Chapter 6

Conclusion

The class of alpha-stable distributions can be used to capture various phenomena including the financial markets. We have managed to show some scenarios in the financial market where this family of distributions plays an important role as summarized below.

In Chapter 3 we showed that the ECF as a parameter estimation method provides the best precision in estimating a wide range of α and β parameters, it is robust and provides better convergence compared to the maximum likelihood method which has been the mostly used in applications according to literature. Secondly, we illustrated that in general, the distribution of the commodity futures log-returns data is closest to a t-location scale distribution due to its high peaks, skinny tails and extreme outliers but the ECF estimation method could be used in addition to capture skewness effects that are not captured in the t-location scale fitting. We recommend the ECF method as the suitable approach for estimating parameters of any skewed financial market data and can be used to obtain initial input parameters for future and better estimation techniques.

In Chapter 4, we showed that the affine property is attainable and applicable to generalized spot models. We considered a stochastic differential equation with the source of randomness as subordinated Brownian motion as a specific example to derive the futures price. Moreover, it has been argued in some existing literature that the likelihood function exists in integrated form for models with singular noise meanwhile for cases of partially observed processes a filtering technique is required. However, the work presented in this chapter provided a new approach of pricing commodity futures for models with latent variables using the maximum expectation maximization, without using traditional filtering methods. Our approach is easy to implement and is robust in the sense that it can accommodate any predictor density but at the expense of computational speed. The numerical routines can be improved for faster computations. This and the numerical implementation of the two factor model are left for future work.

In Chapter 5 we extended the integration by parts formula approach to computing the Greeks of options with discontinuous payoffs presented in [FLL+99] to markets with jumps. In addition, we applied the Bismut-Elworthy-Li (BEL) formula usually used in investigating density regularities for solutions to stochastic differential equations to derive our main results. This in turn enabled the estimation of the Greeks with the subordinated stochastic differential equation (SSDE) as the underlying spot model. We observed that BEL formula performs well for SSDE as in the continuous diffusion models in literature. As a concrete practical application, our model can be applied by investors in emerging/illiquid markets to construct hedge portfolios.

Our work is not exhaustive, there are other interesting applications in finance that were left out but could be interesting to investigate such as modeling foreign exchange markets, investment annuities, zero coupon bonds among others.

Bibliography

[AO16] S. Ament and M. O'Neal. Accurate and efficient numerical calculation of stable densities via optimized quadrature and asymptotics. *arXiv*, 1607.04247v1, 2016. 17

[Aok13] M. Aoki. *State Space Modeling of Time Series.* Universitext. Springer Berlin Heidelberg, Berlin, Germany, 2013. 57

[App04] D. Applebaum. *Lévy Processes and Stochastic Calculus.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2004. 5, 16, 17, 28

[ARSB09] Kabasinskas A., S. Rachev, L. Sakalauskas, and I. Belovas. Alpha-stable paradigm in financial markets. *Journal of Computational Analysis and Applications*, 11(4):641–668, 2009. 9

[Bac00] L. Bachelier. Théorie de la spéculation. *Annales scientifiques de l'É.N.S. 3e série*, 17:21–86, 1900. 5

[Ber98] J. Bertoin. *Lévy Processes.* Cambridge Tracts in Mathematics. Cambridge University Press, 1998. 5

[Ber00] H-. P. Bermin. Hedging Lookback and partial Lookback options using Malliavin calculus. *Applied Mathematical Finance*, 7(2):75–100, 2000. 86

[BKMV12] F. E. Benth, C. Klüppelberg, G. Müller, and L. Vos. Futures pricing in electricity markets based on stable carma spot models. *arXiv*, 1201.1151v1, 2012. 8

[BM06] M-.P. Bavouzet and M. Messaoud. Computation of Greeks using Malliavin's calculus in jump type market models. *Electronic Journal of Probability*, 11(10):276–300, 2006. 86, 87, 98

[BN13] D. Bayazit and C. A. Nolder. Malliavin calculus for Lévy markets and new sensitivities. *Quantitative Finance*, 13(8):1257–1287, 2013. 86, 87, 98

[BnDMBP15] D. R. Baños, S. Duedahl, T. Meyer-Brandis, and F. Proske. Computing Deltas without derivatives. Preprint, 5 2015. 87, 98, 101

[BNMR12] O.E. Barndorff-Nielsen, T. Mikosch, and S.I. Resnick. *Lévy Processes:Theory and Applications*. Birkhäuser Boston, Boston, 2012. 5, 15

[BNS04] O. E. Barndorff-Nielsen and N. Shephard. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37, 2004. 81

[Boc12] S. Bochner. *Harmonic Analysis and the Theory of Probability*. Dover books on mathematics. Dover Publications, 2012. 31, 32

[CF07] T. R. Cass and P. K. Fritz. The Bismut-Elworthy-Li formula for jump-diffusions and applications to Monte Carlo methods in finance. 3 2007. 4, 9, 10, 87, 98, 100, 101, 103

[CK07] J.J.F. Commandeur and S.J. Koopman. *An Introduction to State Space Time Series Analysis*. PRACTICAL ECONOMETRICS SERIES. OUP Oxford, Oxford, 2007. 57

[CMS76] J. M. Chambers, C. L. Mallows, and B . W. Stuck. A method for simulating stable random variables. *Journal of the American Statistical Association*, 71(354):340–344, 1976. 22, 42

[Cro08] J. Crosby. A multi-factor jump-diffusion model for commodities. *Journal of Quantitative Finance*, 8(2):181–200, 2008. 8

[CSZ15] Z.-Q. Chen, R. Song, and X. Zhang. Stochastic flows for Lévy processes with Hölder drifts. *arXiv:1501.04758v1 [math-ph]*, 2015. 87

[Den98] S. Deng. Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes. 10 1998. 8

[DFS03] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053, 2003. 3, 9, 59

[DHKS04] J. Durbin, A. Harvey, S.J. Koopman, and N. Shephard. *State Space and Unobserved Component Models: Theory and Applications*. Cambridge University Press, Cambridge, 2004. 57

[DJX02] S.-J. Deng, W. Jiang, and Z. Xia. Alternative statistical specifications of commodity price distribution with fat tail. *Advanced Modeling and Optimization*, 4(2):1–8, 2002. 8

[DK12] J. Durbin and S.J. Koopman. *Time Series Analysis by State Space Methods: Second Edition*. Oxford Statistical Science Series. OUP Oxford, Cambridge, 2012. 57

[DLR77] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the American Statistical Association*, 39:1–38, 1977. 76

[DNØP08] G. Di Nunno, B. Øksendal, and F. Proske. *Malliavin Calculus for Lévy Processes with Applications to Finance*. Universitext. Springer, 2008. 86, 87, 98

[DP10] P. Date and K. Ponomareva. Linear and nonlinear filtering in mathematical finance: a review. *IMA Journal of Management Mathematics*, 22(3):195–211, 2010. 58

[Dum71] W. H. Dumouchel. Stable distributions in statistical inference. *The Journal of the American Statistical Association*, 78(342):469–477, 1971. 6, 15

[EL94] K. D. Elworthy and X.-M. Li. Formulae for the derivatives of heat semigroups. *Journal of Functional Analysis*, 125:252–286, 1994. 4, 10, 11, 87, 100, 103, 107

[ELO14] P. Erb, D Lüthu, and S. Otziger. Schwartz 1997 two-factor mode. *Technical Document*, (10), 2014. 8

[Fam63] E. F. Fama. Mandelbrot and the stable paretian hypothesis. *The Journal of Business*, 36(4):420–429, 1963. 5

[FF88] E. F. Fama and K. R. French. Permanent and temporary components of stock prices. *The Journal of Political Economy*, 96(2):246–273, 1988. 57

[FLL+99] E. Fournié, J-.M. Lasry, P-.L. Lions, J. Lebuchoux, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3(4):391–412, 1999. 4, 9, 10, 86, 98, 102, 107, 111, 113, 115

[FNYS15] G.A. Fox, S. Negrete-Yankelevich, and V.J. Sosa. *Ecological Statistics: Contemporary Theory and Application*. Oxford University Press, 2015. 77

[FR68] E. F. Fama and R. Roll. Some properties of symmetric stable distributions. *Journal of the American Statistical Association*, 63(323):817–836, 1968. 15

[FR71] E. F. Fama and R. Roll. Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association*, 66:331–338, 1971. 6, 34

[FR07] G. Fusai and A. Roncoroni. *Implementing Models in Quantitative Finance: Methods and Cases*. Springer Finance. Springer Berlin Heidelberg, 2007. 32

[GKP+99] B. Grigelionis, J. Kubilius, V. Paulauskas, V. Statulevicius, and H. Pragarauskas. *Probability Theory and Mathematical Statistics: Proceedings of the Seventh Vilnius Conference (1998), Vilnius, Lithuania, 12-18 August, 1998*. Probability Theory and Mathematical Statistics. TEV, 1999. 17

[GN04] A.K. Gupta and S. Nadarajah. *Handbook of Beta Distribution and Its Applications*. Statistics: A Series of Textbooks and Monographs. Taylor & Francis, 2004. 77

[GS90] R. Gibson and E. S. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3):959–976, 1990. 8

[HC73] D. R. Holt and E. L. Crow. Tables and graphs of the stable probability density function. *J. Res. Nat. Bur. Standards Sect. B*, 77:143–198, 1973. 6, 15

[Hic39] J. R. Hicks. *Value and Capital.* Cambridge: Oxford University Press, Cambridge, 1939. 58

[HR98] J. E. Hilliard and J. A. Reis. Jump processes in commodity futures prices and option pricing. *American Journal of Agricultural Economics*, 81(2):273–286, 1998. 8

[Inc09] Robust Analysis Inc. User manual for stable 5.1 R version, 2009. 42

[JW93] A. Janicki and A. Weron. *Simulation and Chaotic Behavior of Alpha-stable Stochastic Processes.* Chapman & Hall/CRC Pure and Applied Mathematics. Taylor & Francis, 1993. 91

[Ken10] J.F. Kensinger. *Research in Finance.* Number v. 26 in Research in Finance. Emerald Group Publishing Limited, United Kingdom, 2010. 57

[Key30] J. M. Keynes. *A Treatise on Money, Vol II.* Macmillan, London, United Kingdom, 1930. 58

[Khe12] Asma Khedher. Computation of the delta in multidimensional jump-diffusion setting with applications to stochastic volatility models. *Stochastic Analysis and Applications*, 30(3):403–425, 2012. 11, 87

[KKH10] Reiichiro Kawai and Arturo Kohatsu-Higa. Computation of greeks and multidimensional density estimation for asset price models with time-changed brownian motion. *Applied Mathematical Finance*, 17(4):301–321, 2010. 11, 87

[KMT17] M. Kateregga, S. Mataramvura, and D. Taylor. Parameter estimation for stable distributions with application to commodity futures log-returns. *Cogent Economics & Finance*, 5(1):1318813, 2017. 12, 34

[KNPP15] I. Kyriakou, N. K. Nomikos, P. K. Pouliasis, and N. C. Papapostolou. Affine-structure models and the pricing of energy commodity derivatives. *European Financial Management*, pages n/a–n/a, 2015. 9, 58, 64, 71, 73

[Kou80] I. A. Koutrouvelis. Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, 75(372):918–928, 1980. 6, 38

[KR08] M Keller-Ressel. *Affine Processes - Theory and Applications in Finance.* PhD thesis, Technischen Universität Wien, Vienna, Austria, 12 2008. . 3, 9, 59, 60, 62

[KRST11] Martin Keller-Ressel, Walter Schachermayer, and Josef Teichmann. Affine processes are regular. *Probability Theory and Related Fields*, 151(3):591–611, 2011. 3, 9, 59, 62

[Kur01] E. E. Kuruoğlu. Density parameter estimation of skewed alpha-stable distributions. *IEEE Transactions on signal processing*, 49(10):2192–2201, 2001. 6, 24, 25, 34

[Kus10] S. Kusuoka. Malliavin calculus for stochastic differential equations driven by subordinated Brownian motions. *Kyoto Journal of Mathematics*, 50(3):469–644, 2010. 11, 87, 88, 89, 91, 99

[Lév25] P. Lévy. *Calcul des probabilités.* PCMI collection. Gauthier-Villars, 1925. 5

[LM08] S. Lee and P. A. Mykland. Jumps in financial markets: A new nonparametric test and jump clustering. *Review of Financial Studies*, 20, 2008. 80, 82

[Man59] B. Mandelbrot. Variables et processus stochastiques de pareto-Lévy et la répartition des revenus. *Comptes Rendus de l'Académie des Sciences*, 249:613–615, 2153–2155, 1959. 5

[Man62] B. Mandelbrot. Paretian distributions and income maximization. *The Quarterly Journal of Economics*, 76(1):57–85, 1962. 5

[Man63] B. Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419, 1963. 5

[McC86] J. H. McCulloch. Simple consistent estimators of stable distribution parameters. *Communications in Statistics - Simulation and Computation*, 15(4):1109–1136, 1986. 6, 17, 34, 35

[McC96] J. H. McCulloch. 13 financial applications of stable distributions. *Statistical Methods in Finance*, 14:393–425, 1996. 34

[Mhl15] F. J. Mhlanga. Calculations of greeks for jump diffusion processes. *Mediterranean Journal of Mathematics*, 12(3):1141–1160, 2015. 10, 11, 86, 87

[Mil51] J. F. Miller. Moment-generating functions and laplace transforms. *Journal of the Arkansas Academy of Science*, 4(1):16, 1951. 27

[MN95] X. Y. Ma and C. L. Nikias. Parameter estimation and blind channel identification in impulsive signal environments. *IEEE Transactions on signal processing*, 43(12):2884–2897, 1995. 6, 24

[MRD13] Svetlana Maslyuka, Kristian Rotarub, and Alexander Dokumentovc. Price Discontinuities in Energy Spot and Futures Prices. Monash Economics Working Papers 33-13, Monash University, Department of Economics, July 2013. 80

[MRDC99] S. Mittnik, S. T. Rachev, T. Doganoglu, and D. Chenyao. Maximum likelihood estimation of stable paretian models. *Mathematical and Computer Modelling*, 9:275–293, 1999. 41

[MS13] M. M. Meerschaert and P. Straka. Inverse stable subordinators. *Math Model Nat Phenom.*, 8(2):1–16, 2013. 9, 18, 20, 26

[MZ15] A.G. Malliaris and W.T. Ziemba. *The World Scientific Handbook of Futures Markets*. World Scientific Handbook in Financial Economics Series. World Scientific Publishing Company, 2015. 56

[Nol97] J. P. Nolan. Numerical calculation of stable densities and distribution functions. *Communications in Statistics, Stochastic Models*, 13(4):795–774, 1997. 7

[Nol01] John P. Nolan. *Maximum Likelihood Estimation and Diagnostics for Stable Distributions*, pages 379–400. Birkhäuser Boston, Boston, MA, 2001. 34, 41

[Nua95] D. Nualart. *Malliavin Calculus*. Probability and Its Applications. Springer-Verlag, 1995. 86

[Pri94] N. Privault. Chaotic and variational calculus in discrete and continuous time for the poisson process. *Stochastics and Stochastics Reports*, 51:83–109, 1994. 92, 95, 96, 97

[Rac03] S.T. Rachev. *Handbook of Heavy Tailed Distributions in Finance: Handbooks in Finance*. Handbooks in Finance. Elsevier Science, 2003. 13

[RH15] F.D. Rouah and S.L. Heston. *The Heston Model and Its Extensions in VBA*. Wiley Finance. Wiley, 2015. 3, 9, 62, 63

[Sat99] K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1999. 5

[Sch97] E. S. Schwartz. The stochastic behaviour of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3):923–973, 1997. 8, 57, 58, 65

[She12] M. Sheppard. Fit all valid parametric probability distributions to data. Technical Report , Kennesaw State University Department of Mathematics, Kennesaw State, February 2012. ALLFITDIST Matlab code. 46

[SS00] E. S. Schwartz and J. E. Smith. Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7), 2000. 8, 58

[ST94] G. Samoradnitsky and M.S. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Stochastic Modeling Series. Taylor & Francis, United Kingdom, 1994. 5

[Stu04] S. Sturm. Calculation of the greeks by Malliavin calculus. Technical report, Institut für Matematik, Universität Wien, 2 2004. Diplomarbeit zur Erlangung des Akademischen Grades. 101

[SV03] R. Song and Z. Vondraček. Potential theory of subordinate killed Brownian motion in a domain. *Probab. Theory Relat. Fields*, 125:578–592, 2003. 11, 28

[SX14] X. Sun and Y. Xie. Smooth densities for SDEs driven by degenerate subordinated Brownian motion with state-dependent switching. *arXiv:1410.5913v1 [math.PR]*, 2014. 11, 98

[Tak10] A Takeuchi. Bismut-Elworthy-Li-type formulae for stochastic differential equations with jumps. *Journal of Theoretical Probability*, 23(2):576–604, 2010. 11, 87

[TP10] T Tarpey and E. Petkova. Latent regression analysis. *Statistical Modelling*, 10(2):133–158, 2010. 76, 78, 79

[WW95] A. Weron and R. Weron. Computer simulation of Levy alpha-stable variables and processes. *Lecture Notes in Physics*, 457:379–392, 1995. 7, 42

[Wył12] A. Wyłomańska. Arithmetic Brownian motion subordinated by tempered stable and inverse tempered stable processes. *arXiv:1203.0892v1 [math-ph]*, 2012. 11

[Yan12] Y. Yang. Option pricing with non-gaussian distribution-numerical approach. Technical Report , Stony Brook University, Department of Applied Mathematics and Statistics, New York, September 2012. . 34

[YLLN14] J. Yang, B. Lin, W. Luk, and T. Nahar. Particle filtering-based maximum likelihood estimation for financial parameter estimation. In *2014 24th International Conference on Field Programmable Logic and Applications (FPL)*, pages 1–4, Sept 2014. 58

[Yu04] J. Yu. Empirical characteristic function estimation and its applications. *Econometric Reviews*, 23(2):93–123, 2004. 6, 7

[Zha12] X. Zhang. Derivative formula and gradient estimate for SDEs driven by α-stable processes. 4 2012. 4, 11, 85, 88

[Zol64] V. M. Zolotarev. On the representation of stable laws by integrals. *Trudy Mat. Inst. Steklov.*, 71:46–50, 1964. 5, 15

[Zol80] V. M. Zolotarev. Statistical estimates of the parameters of stable laws. *Banach Center Publications*, 6(1):359–376, 1980. 7, 15

[Zol86] V.M. Zolotarev. *One-dimensional Stable Distributions*. Translations of mathematical monographs. American Mathematical Society, Providence, Rhode Island, United States, 1986. 7, 15, 18, 22, 23