PRICING ENERGY CONTRACTS UNDER REGIME SWITCHING TIME-CHANGED LEVY PROCESSES

by

Konrad Gajewski

Bachelor of Music, University of Toronto, 2016

A thesis presented to Ryerson University

in partial fulfilment of the

requirements for the degree of

Master of Science

in the program of

Applied Mathematics

Toronto, Ontario, Canada, 2018

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of the other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

PRICING ENERGY CONTRACTS UNDER REGIME SWITCHING TIME-CHANGED LEVY PROCESSES

Konrad Gajewski

Master of Science, 2018

Applied Mathematics

Ryerson University

Abstract

The failures of the popular Black-Scholes-Merton (BSM) model led to an interest in new, robust models which could more accurately model the behavior of historical prices. We consider one such model, the regime switching time-changed Levy process, which builds upon the BSM model by incorporating jumps through a random clock, as well as randomly varying parameters according to a continuous-time Markov chain. We develop the characteristic function as well as two methods for pricing European call options. Finally, we estimate the parameters of the model by incorporating historic energy data and option quotes using a variety of methods.

Acknowledgements

Dedication

To my family

for their endless love, support, encouragement and sacrifice

- I could never have done it without you.

Table of Contents

List of Tables

List of Figures

List of Appendices

Introduction

The famous Black-Scholes-Merton (BSM) model assumes that the market
consists of a risky asset such as a stock, and a riskless asset such as a bond
or bank account. It can be seen as the continuous limit of the binomial
tree model. Furthermore, the model implies that the market is complete
and assumes that the market admits no arbitrage opportunities. A market
is complete if every payoff is attainable and every derivative asset can be
perfectly hedged. A market admits arbitrage if it is possible to make money
without taking on any risk. Under the BSM model, the rate of return on
the riskless asset is assumed to be constant and hence is called the risk-free
interest rate. The underlying risky assets are distributed under a Gaussian
probability distribution with constant volatility. Hence, the price movements
are continuous. This assumption is not supported by the empirical evidence.
In particular for energy markets such as oil and electricity futures, more
complex movements are observed, such as mean reversion, sudden oscilla-
tions and price jumps. The failures of the BSM model can be seen through
the implied volatility skew. Given the pricing formula and a set of historical
option prices, one can solve for the implied volatility, which would then cor-

rectly price each contract. The limits of the BMS model become apparent when volatilities are compared at different maturity times and strike prices - such volatilies vary substantially, implying that diffusion models are not robust enough to capture the movements of the market [5].

While there is overwhelming evidence from the historical data that the volatility should depend on time, there is no consensus as to how this volatility should be modelled [20]. Examples include local volatility models where the volatility term varies deterministically in time and stochastic volatility models where the volatility term is itself a stochastic process. Another approach is to consider time driven by a non-negative stochastic process called a subordinator, leading to the class of Levy time changed processes.

In option pricing theory, it is a consequence of the model that the market does not admit arbitrage opportunities. This means that the discounted asset prices are martingales under some probability measure $\mathbb{Q}$, called the risk-neutral measure. By introducing stochastic volatility or jumps, this measure $\mathbb{Q}$ loses uniqueness and the existence of infinitely many risk-neutral measures is referred to as market incompleteness, meaning it is not possible to perfectly hedge every derivative asset.

Many financial time series including commodity futures seem to exhibit dramatic breaks in their behaviour, for example in the events of political changes or financial crises. Different intervals sharing similar characteristic

can be grouped together under a single regime. One class of models that describe such behaviour are regime switching Levy models. Under such a model, the process switches randomly between different Levy processes according to an unobservable Markov chain. We modify the regime switching model by introducing a random clock, which introduces random jumps in the model. The regime switching time-changed Levy process is a pure jump process which captures two key features of the market: the existence of regimes and price jumps.

We make a few more assumptions on the market and on the assets. For the market, we assume that it is possible to buy and sell any real number of the risky or riskless assets. This includes fractional amounts of a single share, and negative values which correspond to shorting. The transactions happen instantaneously and at zero cost. Finally, the principle of no arbitrage holds. For the assets, we assume that the risk-free interest rate is constant and doesn't change between regimes and that the risky asset pays no dividends.

The objectives of the thesis are to model regime switching Levy processes, price European call options, and estimate the parameters using option quotes and historical prices of comodities under a variety of methods. Although time-changed Levy processes and switching models have been studied separately, an integrated model has not been considered in connection to pricing and parameter estimation. Consequently, Monte Carlo simulation and Fourier Cosine pricing methods have not been implemented so far.

In Chapter 1 we introduce the relevant background terminology that will be used throughout the thesis. In Chapter 2 we define the regime switching time-changed Levy process, derive its characteristic function under Gamma and Inverse Gaussian subordinators and give two pricing methods: Monte Carlo and Fourier Cosine method. We compare prices under the regime switching model to those given by the Black-Scholes equation and show that the prices agree when the switching model is reduced to the BSM model. For Monte Carlo, we develop an algorithm for simulating trajectories of the regime switching Levy process, as well as for pricing European call options by simulating many regime switching processes simultaneously. In Chapter 3 we use calibration and various estimation methods to estimate values of the parameters using option quotes and historical prices of oil and electricity commodities. Relevant code and rejection methods for simulating random variables are given in the Appendix.

Chapter 1

Background Material

We give several definitions and theorems related to Levy processes and regime switching processes. We begin with an overview of stochastic processes and build up to include time-changed Levy processes, characteristic functions and continuous-time Markov chains.

Definition 1.0.1. (Stochastic Process) [18][27].

A stochastic process is a collection of random variables $X = \{X_t\}_{t \in I}$ defined on a common filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in I}, \mathbb{P})$ and indexed by some totally ordered set I, usually denoting some interval of time. If I is countable, then X is called a discrete-time process, and if I is uncountable then X is a continuous-time process. We will usually take $I = [0, T]$ for some maturity time $T \in \mathbb{R}^+ \cup \{\infty\}$. The probability space consists of:

(1) The nonempty set Ω of all possible paths ω.

(2) The σ-algebra $\mathcal{F}$, which denotes the set of all events. Formally $\mathcal{F}$ consists of the empty set ϕ and the entire set Ω as well as a set of subsets of Ω such that $\mathcal{F}$ is closed under complements and countable unions. The pair consisting of a set with its σ-algebra $(\Omega, \mathcal{F})$ is called a measurable space and the elements of $\mathcal{F}$ are called measurable sets.

(3) The probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$.

(4) A filtration $\{\mathcal{F}_t\}_{t \in I}$, a non-decreasing sequence of σ-algebras, such that

$$\mathcal{F}_0 = \{\phi, \Omega\} \quad \text{and} \quad \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F} \quad \text{for} \quad s \leq t. \qquad (1.1)$$

A measurable space is complete if all subsets of a set of measure zero are also in the σ-algebra. The filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in I}, \mathbb{P})$ is $\mathbb{P}$-complete if $\mathcal{F}_t$ contains all subsets of sets of measure zero under $\mathbb{P}$ for each $t \in I$, and $\{\mathcal{F}_t\}_{t \in I}$ is right continuous i.e.

$$\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s.$$

A Borel set is a set in a topological space that can be formed by complements and countable unions of open sets defined by the topology. The collection of all Borel sets forms the Borel σ-algebra - the smallest σ-algebra containing all open sets. If we consider $\mathbb{R}$ with the standard topology, the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is the smallest σ-algebra containing all the open intervals and the pair $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ form a measurable space. A function between measurable spaces is called a measurable function if the preimage of a measurable set is

measureable. In particular, continuous functions between Borel σ-algebras are measurable because the preimage of every open set under a continuous function is open.

A random variable:

$$X_t : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

is an $\mathcal{F}_t$-measurable function if for each set $B \subseteq \mathcal{B}(\mathbb{R})$,

$$X_t^{-1}(B) = \{\omega \in \Omega : X_t(\omega) \in B\} \in \mathcal{F}_t.$$

Intuitively, $\mathcal{F}_t$ represents the information available at time t, and the filtration $\{\mathcal{F}_t\}_{t \in I}$ represents the information flow evolving over time. Mathematically,

$$\mathcal{F}_t = \{\sigma(X_s), s \in [0, t]\}$$

where the σ-algebra generated by the random variables X_s is defined as:

$$\sigma(X_s) = \{X_s^{-1}(B) | B \in \mathcal{B}(\mathbb{R})\}.$$

A process X is cadlag if it is right continuous with left limits. X is adapted (or non-anticipating) to the filtration $\{\mathcal{F}_t\}_{t \in I}$ if X_t is F_t-measurable for each $t \in I$. The image of $E \subset \mathbb{R}$ of X_t is called the state space [15]. Finally, the conditional expectation of a random variable X_t given an event $A \in \mathcal{F}$, is defined as:

$$\mathbb{E}[X_t | A] = \frac{\mathbb{E}[\mathbb{1}_A X_t]}{\mathbb{P}(A)},$$

where $\mathbb{1}_A$ is the indicator function on a set A.

Definition 1.0.2. (Stopping Time) [8][24].

A random time is a positive random variable $\tau : \Omega \to \mathbb{R}_+$ which represents the time at which an event takes place, for example a change of parameters. Given an information flow $\{\mathcal{F}_t\}_{t \geq 0}$, if we can determine at time t, given the information $\mathcal{F}_t$ whether an event has occured, the random time τ is called a stopping time.

Formally, a stopping time with respect to a filtration $\{F_t\}_{t \geq 0}$ is a random variable τ such that,

$$\{\omega \in \Omega : \tau(\omega) \leq t\} \in F_t, \quad \forall t \geq 0.$$

If τ_1 and τ_2 are stopping times, then $\tau_1 \wedge \tau_2 \equiv \inf\{\tau_1, \tau_2\}$ is also a stopping time. Given a stopping time τ and an adapted process $\{X_t\}_{t \geq 0}$, we can define a new process $X_{t \wedge \tau}$, the process X stopped at τ as:

$$X_{t \wedge \tau} = \begin{cases} X_t, & \text{if } t < \tau \\ X_\tau, & \text{if } t \geq \tau \end{cases}$$

Given an adapted process $\{X_t\}_{t \geq 0}$, the hitting time τ_A of an open set $A \subset E$ is defined as the first time that X hits A:

$$\tau_A = \inf\{t \geq 0 : X_t \in A\}$$

The set of hitting times is a subset of the set of stopping times.

Definition 1.0.3. (Martingales) [23][24].

A cadlag process $\{X_t\}_{t\geq 0}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in I}, \mathbb{P})$ is called a martingale if the best prediction of the future value of a process is its current value:

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \quad \text{for} \quad t > s \quad \text{and} \quad \mathbb{E}[|X_t|] < \infty.$$

A local martingale is a process $\{X_t\}_{t\geq 0}$ for which there exists a sequence of bounded stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ increasing to infinity for which $\{X_{t\wedge\tau_n}\}_{t\geq 0}$ is a martingale for each n.

A semimartingale $\{Y_t\}_{t\geq 0}$ is a process which can be decomposed as:

$$Y_t = M_t + A_t$$

where $M = \{M_t\}_{t\geq 0}$ is a local martingale and $A = \{A_t\}_{t\geq 0}$ is a cadlag adapted process with locally bounded variation. A process $\{A_t\}_{t\geq 0}$ is said to be of bounded variation if its total variation $V(A)$ is finite. The total variation is given by:

$$V(A) = \sup_{P\in\mathcal{P}} \sum_{i=0}^{n_P-1} |A_{t_{i+1}} - A_{t_i}| < \infty.$$

where

$$\mathcal{P} = \left\{P = \{t_0, ..., t_{n_P}\} \text{ where } P \text{ is a partition over } [0, T]\right\}.$$

Definition 1.0.4. (Continuous-Time Markov Chain) [2][19][28].

A Markov process is a stochastic process with the additional Markov property; the future value is independent of the past given the present value. A

$(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued stochastic process $\{X_t\}_{t \in I}$ adapted to the filtration $\{\mathcal{F}_t\}_{t \in I}$ has the Markov property if for each $B \in \mathcal{B}(\mathbb{R})$:

$$\mathbb{P}(X_t \in B | \mathcal{F}_s) = \mathbb{P}(X_s \in B | X_s) \quad \text{when} \quad s < t.$$

A Markov process with a finite countable state space is called a Markov chain. Additionally, if $I = \mathbb{N}$, the Markov process is called a discrete-time Markov chain. The Markov chain is said to be in regime i at time t if $X_t = i$ for each i in the state space. Each discrete-time Markov chain has an associated one-step transition matrix $P = P_{ij}$, $i, j \in E$ where P_{ij} is the probability of moving from regime i to regime j at time $t = 0$.

If $I = [0, \infty)$ and the state space is finite countable, the Markov process is called a continuous-time Markov chain. In the continuous case, there is a transition matrix $P(t) = P_{ij}(t)$ for every $t \geq 0$. As in the discrete case, the distribution of the future X_t, given the present X_t, does not depend upon the present time t, but only on the present state $X_t = i \in E$:

$$P_{ij}(t) = \mathbb{P}(X_t = j | X_0 = i)$$

In order to satisfy the Markov property, the amount of time spent between transitions must not depend on how much time has already elapsed. This is the property of memorylessness. A random variable M is memoryless if :

$$\mathbb{P}(M > t + h | M > t) = \mathbb{P}(M > h) \quad \forall t, h > 0$$

The only continuous probability distribution with this property is the exponential distribution.

The random times at which the continuous-time Markov chain X does transition between regimes are called switching times. The random time interval between switching times during which X does not transition is called a holding-time (or interarrival time), which is an exponentially distributed random variable. The holding time is equal to the duration of remaining in a single regime.

Given a continuous-time Markov chain, we can define its embedded discrete Markov chain, which is the discrete-time process which keeps track of the jumps of the continuous-time Markov chain. The embedded Markov chain then has its own one-step transition matrix $P = P_{ij}$. For each continuous-time Markov chain, there is an associated infinitesimal generator matrix Q, with the same dimensions as P defined element-wise:

$$Q = q_{ij} \equiv \lim_{t \to 0^+} \frac{P_{ij}(t) - \delta_{ij}}{t} = P'_{ij}(0), \tag{1.2}$$

where δ_{ij} are the elements of the identity matrix, $P'_{ij}(0) = \lambda^i P_{ij}$ for $i \neq j$ is the transition rate from regime i to regime j when the process is in regime i, and $P'_{ii}(0) = -\sum_{i \neq j} q_{ij}$. The parameter λ^i is called the holding-time rate.

The elements q_{ij} of the infinitesimal generator Q can now be used to define the infinitesimal transition probabilities:

$$\mathbb{P}(X_{t+h} = j | X_t = i) = q_{ij} h + o(h), \quad i \neq j$$

$$\mathbb{P}(X_{t+h} = i | X_t = i) = 1 + q_{ii} h + o(h) \quad \text{otherwise,}$$

for some arbitrarily small time increment $h > 0$. The term $o(h)$ denotes any function $f(h)$ with the property that $\lim_{h \to 0} f(h)/h = 0$.

Definition 1.0.5. (Brownian Motion) [10].

One of the most import examples of a stochastic process is Brownian motion (also called the Weiner process), denoted $B = \{B_t\}_{t \geq 0}$. The process has independent and stationary increments, i.e. given $0 \leq s < t$ the increment $B_t - B_s$ is independent of $\mathcal{F}_s$, and $B_{t+s} - B_t \sim B_s$ respectively. The increments $B_t - B_s$ are normally distributed random variables with mean 0 and variance $t - s$. The process is called Standard Brownian motion if $B_0 = 0$. Given the set of partitions:

$$\mathcal{P} = \left\{ P = \{t_0, ..., t_{n_P}\} \text{ where } P \text{ is a a partition over } [0, T] \right\},$$

Brownian motion has infinite total variation:

$$V(B) = \sup_{P \in \mathcal{P}} \sum_{i=0}^{n_P - 1} |B_{t_{i+1}} - B_{t_i}| = \infty,$$

and finite quadratic variation at each time $t \in (0, T]$:

$$< B >_t = \lim_{\substack{\|P\| \to 0 \\ P \in \mathcal{P}}} \sum_{i=0}^{n_P - 1} |B_{t_{i+1}} - B_{t_i}|^2 = t < \infty.$$

Here the norm of a partition $\|P\|$ is defined as $\|P\| = \max\{(t_i - t_{i-1}) : i = 1, .., n_P\}$. Brownian motion is known to not be differentiable anywhere, however, it is almost surely everywhere continuous.

Definition 1.0.6. (Poisson Process and Random Measures) [26].

A counting process is a non-decreasing pure jump stochastic process with a state space consisting of non-negative integers. The Poisson process is

the simplest example of a counting process. If $\{\tau_i\}_{i \geq 1}$ is a sequence of identical, independent exponential random variables with parameter λ and $T_n = \sum_{i=1}^{n} \tau_i$, then the process $\{N_t\}_{t \geq 0}$ defined by:

$$N_t = \sum_{n \geq 1} \mathbb{1}_{t \geq T_n}$$

is called a Poisson process with intensity λ. The process at time t counts the number of random times $\{T_n, n \geq 1\}$ occurring in the interval $[0, t]$. For any path ω, the sample path $\omega \mapsto N_t(\omega)$ is piece-wise constant, increases by jumps of unit size, and the conditional expectation of future states depends only on the present state (this is the Markov property). For any $t > 0$, the random variable N_t follows a Poisson distribution with parameter λt:

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad \forall n \in \mathbb{N},$$

and has expectation $\mathbb{E}[N_t] = \lambda t$. The increments $N_t - N_s$ are Poisson distributed with parameter $\lambda(t - s)$ for $0 \leq s < t$, and are independent of $\mathcal{F}_s$. With each Poisson process we can define an associated Poisson random measure M, such that for a given realization ω and $A \subset \mathbb{R}^+$ [8],

$$M(\omega, A) = \|i \geq 1 | T_i(\omega) \in A\|$$

For any set A, $M(\omega, A)$ follows the Poisson distribution with intensity $\lambda \|A\|$ where $\|A\|$ is the Lebesgue measure of A. M is called a random measure because it depends on the random path ω. For each path ω, the Poisson process is related to its random measure by:

$$N_t(\omega) = \int_{s \in [0, t]} M(\omega, ds).$$

13

Given Poisson Process with intensity λ, we can define a "Compensated" Poisson process $\{\tilde{N}_t\}_{t \geq 0}$ as:

$$\tilde{N}_t = N_t - \lambda t$$

The term λt is called the compensator, and is the quantity subtracted from N_t in order to recover the martingale property. The Compensated Poisson Process is no longer integer valued and can be negative, and so the compensated random measure:

$$\tilde{M}(\omega, A) = M(\omega, A) - \lambda \|A\|$$

is now a signed measure, meaning it can take on negative values.

Definition 1.0.7. (Compound Poisson Processes).

A process $Y = \{Y_t\}_{t \geq 0}$ is called a Compound Poisson process with parameter λ is defined as:

$$Y_t = \sum_{j=1}^{N_t} D_j$$

where $\{N_t\}_{t \geq 0}$ is a Poisson process with intensity λ, and $\{D_j, j \geq 1\}$ is a set of identically distributed random variables, all independent of each other and of N_t for each t. We note that if D_j is the constant unit function the Compound Poisson process reverts to an ordinary Poisson process.

For each Compound Poisson process, we can associate a random measure on $\mathbb{R} \times [0, \infty)$ describing the jumps of Y. For any measurable set $A \times [t_1, t_2] \subset \mathbb{R} \times [0, \infty)$, define:

$$J_Y(A \times [t_1, t_2]) = \#\{t \in [t_1, t_2] \quad \text{such that} \quad \Delta Y_t \neq 0, \quad \Delta Y_t \in A\}, \quad (1.3)$$

14

where $\Delta Y_t = Y_t - \lim_{s \to t^-} Y_s$. Intuitively, J_Y counts the number of jump times between t_1 and t_2 whose jump sizes are in A. Jumps of zero size are not allowed, and negative sizes values correspond to descending jumps. If $\mathbb{E}[D_j] = \kappa < \infty$ for each j, then $\mathbb{E}[Y_t] = \lambda t \kappa$. Just like in the ordinary Poisson process, the Compound Poisson process has a compensated version $\{\tilde{Y}_t\}_{t \geq 0}$ which is a martingale. It it defined as:

$$\tilde{Y}_t = \sum_{j=1}^{N_t} D_j - \lambda t \kappa.$$

Definition 1.0.8. (Levy Processes) [22][23].

A cadlag, adapted, real-valued stochastic process $X = \{X_t\}_{t \geq 0}$ is called a Levy Process if:

(1) $X_0 = 0$ almost surely.

(2) $\forall t_1, t_2, ..., t_n$ such that $0 \leq t_1 < t_2 < ... < t_n < \infty$, the increments $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, ..., X_{t_n} - X_{t_{n-1}}$ are independent.

(3) The increments are stationary, i.e $\forall \ t \geq 0, h > 0, X_{t+h} - X_t$ is equal in distribution to X_h, meaning the distribution of the increments depends only on the length of elapsed time h.

(4) X is stochastically continuous (continuous in probability): $\forall \epsilon > 0$ and $t \geq 0$,

$$\lim_{h \to 0} \mathbb{P}(\{\omega \in \Omega : |X_{t+h}(\omega) - X_t(\omega)| > \epsilon\}) = 0$$

15

Every Levy process has the Markov property. Important examples of Levy processes include Brownian motion and Compound Poisson processes.

Definition 1.0.9. (Levy Triplet and Levy-Ito Decomposition) [8].

Every Levy Process $X = \{X_t\}_{t \geq 0}$ is completely determined by three parameters called the Levy Triplet (μ, σ^2, ν), where $\mu \in \mathbb{R}$ is the drift, $\sigma > 0$ is the volatility and ν is the Levy measure:

$$\nu : \mathcal{B}(\mathbb{R}) \to \mathbb{R}^+.$$

Given any Borel set $B \in \mathcal{B}(\mathbb{R})$, $\nu(B)$ is the expected number of jumps whose size belongs to B, per unit time.

Formally:

$$\nu(B) = \mathbb{E}\Big[\#\{t \in [0,1] : \Delta X_t \neq 0, \ \Delta X_t \in B\}\Big],$$

where ΔX_t is the jump size. If $\nu(\mathbb{R}) < \infty$, the process has a finite number of jumps on every compact interval. If $\nu(\mathbb{R}) = \infty$, the process jumps an infinite number of times on every compact interval and we say that the process has infinite activity. The measure ν also satisfies the following condition:

$$\nu(\{0\}) = 0 \qquad \text{and} \qquad \int_{\mathbb{R}\setminus\{0\}} \min\{1, |x|^2\}\nu(dx) < \infty, \qquad (1.4)$$

Every Levy Process X with an associated Levy Triplet (μ, σ^2, ν) can be written as a sum of the processes X^1, X^2, X^3, X^4 such that X converges to

$X^1 + X^2 + X^3 + \lim_{\epsilon \to 0} X^4$ almost surely. The four processes are defined:

$$X_t^1 = \mu t$$

$$X_t^2 = \sigma B_t$$

$$X_t^3 = \int_{|x| \geq 1} \int_{s \in [0,t]} x J_X(ds \times dx)$$

$$X_t^4 = \int_{\epsilon \leq |x| < 1} \int_{s \in [0,t]} x \{ J_X(ds \times dx) - \nu(dx)ds \}.$$

This decomposition shows that X is a combination of a deterministic component, Brownian motion and a pure jump process. For the pure jump process, we differentiate between "large" jumps and "small" jumps. X^3 is a compound Poisson process whose jump sizes are greater or equal to 1:

$$X_t^3 = \sum_{0 \leq s \leq t}^{|\Delta X_s| \geq 1} \Delta X_s,$$

while X^4 is a Compensated Compound Poisson process with jump sizes between $\epsilon > 0$ and 1:

$$X_t^4 = \sum_{0 \leq s \leq t}^{\epsilon \leq |\Delta X_s| < 1} \Delta X_s. \tag{1.5}$$

J_X is the Poisson random measure defined by equation (1.3). Equation (1.4) requires that there are finite number of jumps of size greater or equal to 1, while allowing an infinite number of jumps smaller than 1 providing that the sum of the squared values of the jumps converges to a finite value.

Since jumps can only be of nonzero size, this requires X^4 to be dependent on ϵ which cannot be taken to zero directly. The threshold separating the "large" jumps from the "small" jumps is taken to be $|\Delta X| = 1$, however this is arbitrary, and any jump size greater than zero can be used.

17

From the Levy-Ito Decomposition, we can see that if $\nu = 0$, the process is almost surely continuous. If instead $\sigma = 0$ and $\int_{|x|1} |x| \nu(dx) < \infty$, then almost all paths have finite total variation. If $\sigma^2 \neq 0$ or $\int_{|x|1} |x| \nu(dx) = \infty$ (or both) then almost all paths have infinite total variation. Finally, if X has the triplet $(0, 0, \nu)$, then X is a pure jump process.

There are many ways of constructing martingales from Levy processes. Given any Levy process X with $\mathbb{E}[|X_0|] < \infty$, the process $\{\tilde{X}_t\}_{t \geq 0}$ defined as

$$\tilde{X}_t = X_t - \mathbb{E}[X_t],$$

is a martingale. X is also a martingale if and only if

$$\int_{|x| \geq 1} |x| \nu(dx) < \infty \quad \text{and} \quad \mu + \int_{|x| \geq 1} x \nu(dx) = 0.$$

The exponential of a Levy process $Y = \{Y_t\}_{t \geq 0}$ defined as $Y_t = \exp X_t$ is martingale if and only if:

$$\int_{|x| \geq 1} \exp(x) \nu(dx) < \infty$$

and

$$\mu + \frac{\sigma^2}{2} + \int_{\mathbb{R}} (\exp(x) - 1 - x \mathbb{1}_{|x| \leq 1} \nu(dx) = 0.$$

Finally the process $\{M_t\}_{t \geq 0}$ defined by

$$M_t = \frac{\exp(\zeta X_t)}{\mathbb{E}[\exp(\zeta X_t)]}$$

is a martingale for $\zeta \in \mathbb{R}$ providing that $\mathbb{E}[\exp(\zeta X_t)] < \infty$.

Definition 1.0.10. (Time-Changed Levy Processes) [22][4][6].

An almost surely increasing Levy process $L = \{L_t\}_{t\geq 0}$ is called a Sub-ordinator. Its Levy triplet $(\mu, 0, \nu)$ must satisfy $\nu((-\infty, 0)) = 0$ as well as equation (1.4). It can be modelled as a nondecreasing semi-martingale:

$$L_t = \mu t + \int_{x>0} \int_{s\in[0,t]} x J_L(dx \times ds).$$

Financially, t is called the calendar time or operational time while the subordinator L is called the business time. If $\{X_t\}_{t\geq 0}$ is any Levy process, we define the time-changed Levy process as $\{Y_t\}_{t\geq 0}$:

$$Y_t \equiv X_{L_t}.$$

The process $\{X_t\}_{t\geq 0}$ is then called the base process. If $X = \{X_t\}_{t\geq 0}$ is a Levy process with a Levy triplet (μ_X, σ_X, ν_X) and $L = \{L_t\}_{t\geq 0}$ is a subordinator with Levy triplet $(b, 0, \rho)$, then the time-changed process $\{Y_t\}_{t\geq 0}$ is again a Levy process with Levy triplet (μ_Y, σ_Y, ν_Y) where:

$$\mu_Y = \mu_X b + \int_0^\infty \rho(ds) \int_{|x|\leq 1} x p_s^X(dx)$$

$$\sigma_Y = b\sigma_X$$

$$\nu_Y(B) = b\nu_X(B) + \int_0^\infty p_s^X(B)\rho(ds), \quad \forall B \in \mathcal{B}(\mathbb{R}), \tag{1.6}$$

where $p_s^X(B)$, is the probability of B under the probability measure of X_s. While equation 1.0.10 is useful for computing Levy triplets of new Levy processes, it is difficult to find closed form formulas in most cases.

Every local martingale $M = \{M_s\}_{s\geq 0}$ can be written as time-changed Brownian motion $\{B_{<M>_s}\}_{s\geq 0}$, where $< M >_s$ is the quadtratic variation of

19

M. If the process $\{X_t\}_{t\geq 0}$ defined as:

$$X_t = \mu t + \sigma B_t,$$

is subordinated by a pure jump process $\{L_t\}_{t\geq 0}$, the resulting process $\{X_{L_t}\}_{t\geq 0}$ defined as:

$$X_{L_t} = \mu L_t + \sigma B_{L_t}, \tag{1.7}$$

is a pure jump process, and is distributed as:

$$X_{L_t}|L_t \sim N(\mu L_t, \sigma^2 L_t) \sim \mu L_t + \sigma\sqrt{L_t}N(0,1),$$

where $N(0,1)$ is a random variable with standard normal distribution.

We will focus on two particular subordinators: Gamma processes and Inverse Gaussian processes. When the subordinator $\{L_t\}_{t\geq 0}$ is a Gamma process, the increments $L_{t+\Delta t} - L_t$ have Gamma distribution with parameters $\alpha\Delta t$ and β. In particular, if $\alpha = \beta = 1/\gamma$ for some $\gamma > 0$, the the process (1.7) is called a variance gamma process with parameters μ, σ, γ. When the subordinator $\{L_t\}_{t\geq 0}$ is an Inverse Gaussian process, the increments $L_{t+\Delta t} - L_t$ have Inverse Gaussian distribution with parameters $\alpha\Delta t$ and β. In particular, if $\alpha = 1$ and $\beta = \delta\sqrt{a^2 - b}$ for $a > 0, -a < b < a$ and $\delta > 0$; then the process

$$Y_t = b\delta^2 L_t + \delta B_{L_t},$$

is a Normal Inverse Gaussian process with parameters a, b, δ. While the variance gamma and Normal Inverse Gaussian processes have been thoroughly studied, we will focus on the more general Gamma and Inverse Gaussian processes.

20

Definition 1.0.11. (Characteristic function) [4][27].

Given a Levy Process $X = \{X_t\}_{t \geq 0}$ with Levy triplet (μ, σ, ν), the characteristic function is given by the Levy- Khinchine formula:

$$\psi_{X_t}(u) = \mathbb{E}[\exp(iuX_t)] = \exp(-t\phi_X(u))$$

where the characteristic exponent (or Levy symbol) is defined as follows:

$$\phi_X(u) = -\mathrm{i}\mu u + \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} \left(1 - e^{iu\xi} + iu\xi \mathbb{1}_{|\xi|<1}\right)\nu(d\xi) \qquad (1.8)$$

where $u \in \mathbb{C}$. The characteristic exponent is completely determined by the Levy triplet of X. In particular, $\nu(d\xi)$ is called the Levy density and is known in closed form in the case of Gamma and Inverse Gaussian processes. The Levy density of a Gamma process with parameters α and β is given by:

$$\nu(d\xi) = \frac{\alpha}{\xi}\exp(-\beta\xi)\mathbb{1}_{\xi>0}d\xi.$$

The Levy density of an Inverse Gaussian process with parameters α and β is given by:

$$\nu(d\xi) = \frac{\alpha}{\sqrt{2\pi\xi^3}}\exp(\beta^2\xi/2)\mathbb{1}_{\xi>0}d\xi.$$

If $Y_t = X_{L_t}$ is a Time-Changed process, where X and L are independent Levy processes, we can use the tower rule to find a closed for solution to its characteristic function:

$$
\begin{aligned}
\psi_{Y_t}(u) &= \mathbb{E}[\exp(iuX_{L_t})] \\
&= \mathbb{E}[\mathbb{E}[\exp(iuX_{L_t})|L_t]] \quad \text{(tower rule)} \\
&= \mathbb{E}[\exp(-L_t\phi_X(u))] \\
&= \mathcal{L}_{L_t}(\phi_X(u)),
\end{aligned}
$$

where $\mathcal{L}_{L_t}$ is the Laplace Transform of the probability density function of the process L at time t, evaluated at the characteristic exponent of X. The characteristic exponent of L can then be recovered by:

$$\phi_L(u) = t^{-1} \log \mathcal{L}_{L_t}(\phi_X(u)). \tag{1.9}$$

The Laplace Transform of the density function of any random variable M_t, evaluated at $s \in \mathbb{C}$, is defined as $\mathcal{L}_{M_t}(s) \equiv \mathbb{E}[\exp(-sM_t)]$.

Definition 1.0.12. (Switching Levy Process) [5].

Assume the filtered probability space $(\Omega, \mathcal{F}, \{F_t\}_{t \in I}, \mathbb{P})$. Consider a collection of independent Levy processes $X^j = \{X_t\}_{t \geq 0}^j$, for $j = 1, ..., N$, where each process X^j has Levy triplet $(\mu^j, (\sigma^2)^j, \nu^j)$, as well as a continuous-time Markov chain $s = \{s_t\}_{t \geq 0}$ with state space $E = \{1, 2, ..., N\}$, and infinitesimal generator matrix Q. We can then define a new process $Z = \{Z_t\}_{t \geq 0}$ where:

$$Z_t = X_t^{s_t}.$$

The process Z switches between the different Levy processes according to the continuous Markov chain s, which is governed by the transition rates in Q. The Switching Levy process is in general not Levy because the increments are no longer stationary. The process s is generally not observable, and must be inferred from empirical data.

In the next section, we introduce the regime switching time-changed

model where the Levy processes are now subordinated by pure jump processes.

Chapter 2

Pricing Methods

In this chapter we define our model - the regime switching time-changed Levy process and give two pricing methods to find the fair value of European call options, namely pricing by Monte Carlo simulation, and Fourier-Cosine pricing. We define the characteristic function for our model, which will be used when computing the price using the Fourier-Cosine method and also find the conditions under which our model, conditional on each regime, is a martingale.

2.1 The Regime Switching Time-Changed Levy Process

Let $V = V(T, K)$ be the price of a European call option under the risk neutral measure $\mathbb{Q}$. The original Black-Scholes partial differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

is derived from the Feymann-Kac formula which relates partial differential equations with stochastic processes. This equation can in turn be solved to obtain the Black-Scholes formula. The Black Scholes PDE is derived from the stochastic process $\{S_t\}_{t \geq 0}$ where:

$$S_t = S_0 \exp(X_t) \quad \text{where} \quad X_t = \mu t + \sigma B_t. \tag{2.1}$$

The process $\{S_t\}_{t \geq 0}$ represents the price movements while $\{X_t\}_{t \geq 0}$ represents the logarithmic prices. S_t is the asset price at time t, μ is the constant mean return, also called the drift and σ is the volatility. Under this model, price movements are assumed to follow geometric Brownian motion, and therefore follow continuous trajectories. Moreover, the parameters of the process (2.1) are assumed to remain constant with respect to time. We sacrifice both properties in the interest of introducing a more realistic model. This is done by introducing price discontinuities in the form of randomly changing regimes, as well as through time-changed Levy processes.

We fix a maturity time $T < \infty$ and define a continuous-time Markov chain $\{s_t\}_{0 \leq t \leq T}$ driving the changes between regimes. We assume our model will

25

jump between two states, and therefore fix the state space of the continuous-time Markov chain to be $E = \{1, 2\}$. The switching times are defined as the random variables $\tau_1, \tau_2, .. \in (0, T]$ such that:

$$\lim_{t \to \tau_k^-} s^t \neq s^{\tau_k} \quad \text{for} \quad k = 1, 2, ...$$

With each state $j \in E$ there is an associated parameter $\lambda^j > 0$ which controls the jump rate between states. For ease of notation and because we are working with two states, we are denoting λ^j to be the parameter which controls the jump time from state j to state $(j \mod (2) + 1)$, meaning λ^1 controls the jump rate from state 1 to state 2 and λ^2 controls the jump rate from state 2 back to state 1.

The infinitesimal generator matrix of the continuous-time Markov chain is given by:

$$Q = \begin{bmatrix} -1/\lambda^1 & 1/\lambda^2 \\ 1/\lambda^1 & -1/\lambda^2 \end{bmatrix}. \tag{2.2}$$

We also assume a constant risk-free interest rate which remains the same between the two states i.e. $r^1 = r^2 = r \in \mathbb{R}^+$.

Consider a collection of independent geometric Brownian motion processes $X^j = \{X_t^j\}_{0 \leq t \leq T}$ for $j \in E$, where the j'th process is given by

$$X_t^j = \mu^j t + \sigma^j B_t,$$

where $\mu^j \in \mathbb{R}$ and $\sigma^j > 0$, as well as a collection of independent subordinators $L^j = \{L_t^j\}_{0 \leq t \leq T}$ for $j \in E$ where each subordinator L^j is also independent of each process X^i, for $i, j \in E$. We consider two different subordinators:

26

Gamma processes and Inverse Gaussian processes. Each is characterized by two parameters $\alpha^j, \beta^j > 0$ which change between states. Each process L^j is a pure jump process and each process X^j has almost surely continuous paths, therefore the time-changed Levy process, where each random variable X_t^j is indexed by the random variable L_t^j, is jump process. This can be seen by computing the new Levy measure from equation(1.0.10).

We define the collection of time-changed Levy processes $Y^j = \{Y_t^j\}_{0 \leq t \leq T}$ where

$$Y_t^j = X_{L_t^j}^j = \mu^j L_t^j + \sigma^j B_{L_t^j}. \tag{2.3}$$

The economic interpretation for evaluating Brownian motion at random times is that the relevant trading times are random [6].

We now define the regime switching time-changed Levy process $Z = \{Z_t\}_{0 \leq t \leq T}$ as:

$$Z_t \equiv Y_t^{s_t} \quad \text{where} \quad Y_t^{s_t} = \mu^{s_t} L_t^{s_t} + \sigma^{s_t} B_{L_t^{s_t}}. \tag{2.4}$$

The regime switching time-changed Levy process Z is assumed to be the log-price process of the underlying asset and the stochastic process of the asset price itself $\{S_t\}_{0 \leq t \leq T}$ is defined as:

$$S_t = S_0 \exp(Z_t). \tag{2.5}$$

2.2 Simulating a Single Trajectory

One popular method for pricing options is by randomly sampling payoffs at maturity and then computing the sample mean or confidence interval. The

class of computational algorithms that rely on random sampling are called Monte Carlo simulation. This method has the advantage of being able to price many different types of financial derivatives which rely on complicated models, (which usually don't have closed form solutions such as in the case of the Black Scholes equation) as well as when the derivative is both path dependent and independent. It is in general very difficult to derive partial differential equations from regime switching models, so Monte Carlo simulations are a popular method for pricing derivatives. Monte Carlo simulations also have the benefit of being able to handle multivariate models (for example if we wanted to simulate a regime process with a larger number of regimes, or with multiple stocks) with low additional computational cost. By contrast, numerical solutions to partial differential equations become computationally expensive in higher dimensional spaces.

Monte Carlo simulation relies on the Law of Large numbers. The strong version states that, given a sample of independent and identically distributed copies $X_1, X_2, ..., X_n$ of a random variable X, the sample mean $\bar{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$ converges to $\mathbb{E}[X]$ almost surely [11]. Equivalently:

$$\mathbb{P}(\lim_{n \to \infty} \bar{X}_n = \mathbb{E}[X]) = 1.$$

We begin by giving a method for simulating the values of the log-price process Z given by equation (2.4) on an partitioned time interval $t \in [0, T]$ where $T < \infty$ is the maturity time. The switching time-changed Levy process is more revealing if written in differential form:

$$dZ_t \equiv dY_t^{s_t} \quad \text{where} \quad dY_t^{s_t} = \mu^{s_t} dL_t^{s_t} + \sigma^{s_t} dB_{L_t^{s_t}}. \tag{2.6}$$

Equation 2.6 is equivalent to (2.4) because conditional on each regime, the continuous-time Markov chain $\{s_t\}_{0 \le t \le T}$ is constant.

The simulation is done in two main steps: simulating the continuous-time Markov chain, and simulating the regime switching process. The continuous-time Markov chain spends exponentially distributed random time between transitions. To find the switching times $\tau_1, \tau_2, \dots$ we simulate exponentially distributed random variables $\Delta\tau_1, \Delta\tau_2, \dots$ and the k'th switching time is given by:

$$\tau_k = \sum_{i=1}^{k} \Delta\tau_i$$

The parameters of the exponential random variables alternate between λ^1 and λ^2, which causes the continuous-time Markov chain to remain in one regime longer than the other, on average. The number of switching times cannot be predicted and varies randomly with each simulation.

The differential equation (2.6) can be approximated numerically through finite differences using the Euler Method, in a way similar to the method used in numerical solutions of ordinary differential equations. Conditional on each regime $j \in E$, the increment of the process Z during the time interval Δt is given by:

$$\Delta Z_t^j = \mu^j \Delta L_t^j + \sigma^j \sqrt{\Delta L_t^j} N.$$

Here we used the fact that conditional on ΔL_t^j, the increment $\Delta B_{L_t^j}$ is Brownian motion with variance $\sqrt{\Delta L_t^j}$ [3]:

$$\Delta B_{L_t^j} | L_t^j \sim \sqrt{\Delta L_t^j} N$$

where N is a normally distributed random variable with mean 0 and variance 1. The random variable ΔL_t^j is a Gamma (or an Inverse Gaussian) distributed random variable with parameters $\alpha^j \Delta t$ and β^j. The value of Z_t at time t is a summation of the increments:

$$Z_t = \sum_{i=0}^{t} \Delta Z_i.$$

Computations for simulating a single trajectory of the switching Levy process on a partitioned interval are summarized in the following algorithm, where the subordinator is a Gamma process. The algorithm is almost identical when the subordinator is an Inverse Gaussian process.

1. Initialize:

   the log-price process $Z_0 = 0$

   regime state $j = 1$

   index $k = 1$

   first stopping time $\tau_k \sim exp(\lambda^j)$

2. Recursively define the continuous-time Markov chain:

   (a) while $\tau_k < T$,

   switch regime $j = j \mod 2 + 1$

   $k = k + 1$

   simulate random variable $e \sim exp(\lambda^j)$

   define $\tau_k = \tau_{k-1} + e$,

   (b) otherwise stop.

3. Define $\tau = [\tau_1, ..., \tau_{k-1}, T]$ as the set of stopping times. (Note τ_k was removed because $\tau_k > T$)

4. Initialize a partition of the domain into N subintervals of width $\Delta t = T/N : t = [0, t_1, .., t_n, ..., t_{N-1}, T]$ where $t_n = n\Delta t$

5. Define $\mathcal{T}$ to be the sorted union of τ and t in ascending order. Let $\mathcal{T}_n$ denote the n'th element of $\mathcal{T}$

6. To simulate the trajectory of the log-returns: Reinitialize regime state $j = 1$

7. for $1 \leq n \leq \text{size}(\mathcal{T})$ (for each element in $\mathcal{T}$)

    (a) initialize $\Delta Z = 0$

    (b) simulate a gamma random variable $\Delta L \sim Gamma(\alpha^j(\mathcal{T}_n - \mathcal{T}_{n-1}), \beta^j)$

    (c) simulate standard normal variable $\mathcal{N} \sim N(0, 1)$

    (d) $\Delta Z = \Delta Z + \mu^j \Delta L + \sigma^j \sqrt{\Delta L} N$

        i. if $\mathcal{T}$ is an element of the partitioned time interval (if $\mathcal{T}_n \in t$)

           $Z_n = Z_{n-1} + \Delta Z$

           $\Delta Z = 0$

        ii. if $\mathcal{T}$ is a switching time (if $\mathcal{T}_n \in \tau$)

           switch regimes $j = j \mod 2 + 1$
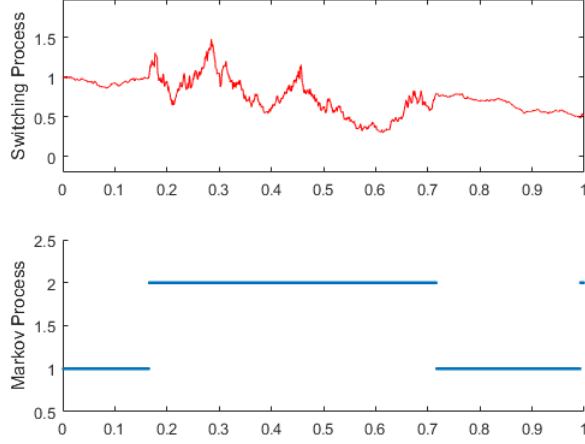
Figure 2.1: Parameters: $T = 1, \mu^1 = 0.01, \mu^2 = -0.1, \sigma^1 = 1, \sigma^2 = 5, \alpha^1 = \alpha^2 = 0.1, \beta^1 = 0.1, \beta^2 = 10, \lambda^1 = 0.2, \lambda^2 = 0.5$.
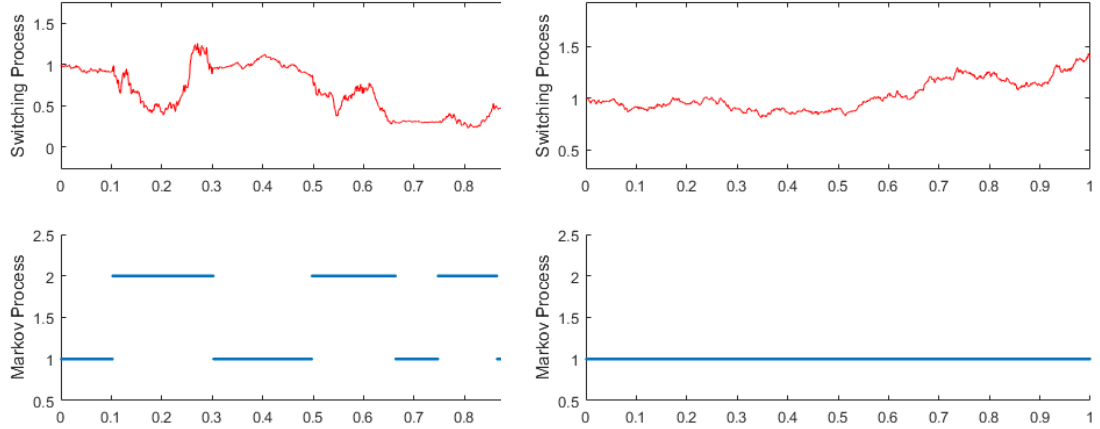
(a) $\lambda^1 = \lambda^2 = 0.25$ (b) $\lambda^1 = 10, \lambda^2 = 0.25$

Figure 2.2: Same parameters as Figure 2.1 except for the holding-time rates λ^1 and λ^2.

Figures 2.1 and 2.2 show a single realization of a switching time-changed

Levy process (top) as well as the underlying continuous-time Markov chain (bottom), with time domain $[0, 1]$ partitioned into 1000 intervals. Figure (3.2b) is a time-changed Levy process, meaning no regime change has occured. If we have $\lambda^1 \to \infty$ the process will remain in regime 1 indefinitely and therefore it reduces to a time-changed process. The average length of a regime is directly proportional to the holding rate parameter.

2.3 Pricing European Options under Monte Carlo Simulation

For all option pricing methods, we assume that the present value of an option is equal to the discounted expectation of its payoff at maturity time, i.e. we will work under an equivalent martingale measure (EMM) $\mathbb{Q}$. A measure $\mathbb{Q}$ is an EMM if:

1. $\mathbb{Q}$ and the real world probability measure $\mathbb{P}$ are equivalent i.e. they share the same null sets;

2. the discounted price process is a martingale under $\mathbb{Q}$.

There are several criteria to change the probability to a risk-netural setting. See for example [12] for a change based on an Esscher transform or [8] for a minimum entropy criterion. In this thesis we will not address this problem; we simply change the drift leaving the remaining parameters constant, as explained in section 2.5. At time t, the value of a European option is given

by [8]:

$$V(t) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[H(S_T)|\mathcal{F}_t] \quad \text{for} \quad 0 \le t \le T.$$

Here H is the payoff function, which for European vanilla options is a function of the asset price at maturity given by:

$$H(S_T) = \max\{\gamma(S_T - K), 0\}$$

where $K > 0$ is the strike price, and $\gamma = +1$ for call options or $\gamma = -1$ for put options.

If the payoff of the option is path independent and the underlying model of the log-return process is an ordinary Levy process, the payoff at time T can be simulated in one step because of the property of stationary increments. When the log-return process is a switching Levy process, the increments are independent only conditionally on each regime.

We devise an algorithm which computes the payoff of a European Call option by simulating N independent realizations of the process Z simultaneously and then calculating the payoff according to:

$$C(T) = \exp(-rT)\frac{1}{N}\sum_{n=1}^{N}\{(S_0 \exp(Z_{n,1}) - K), 0\}$$

where $Z_{n,1}$ is the n'th element of the N-by-1 array $Z_{:,1}$ is at maturity T, and the expected value is approximated by the sample mean. The simulation of all N values at each step simultaneously allows for a much smaller computation time than simulating a single process at a time.

The switching times are now contained in a (N-by-k) matrix τ where $\tau_{m,n}$ is the n'th switching time of the m'th simulation. The number of columns

of τ changes randomly with each simulation. The n'th column of τ (denoted as $\tau_{:,n}$) is generated by simulating an (N-by-1) array of exponential random variables (all with the same parameter) and adding it component-wise to the $(n-1)'$th column of τ :

$$\tau_{:,n} = \tau_{:n-1} + [e_1, ..., e_N]^T \quad \text{where} \quad e_i \sim \exp(\lambda^j), \quad 1 \leq i \leq N,$$

and $j \in \{1, 2\}$. The λ^j parameter alternates between λ^1 and λ^2 after each step. The process of appending columns to τ continues until each element of the final column is greater or equal to T.

Then, for each pair of indices m, n such that $\tau_{m,n} > T$ we redefine that element to equal T, because the stochastic process $\{Z_t\}_{0 \leq t \leq T}$ is defined only up to T.

This is not the most efficient method as we have to simulate more random variables than is actually needed. This is because each of the processes require a different amount of steps to arrive at maturity. However, because we need to keep the number of steps the same for all paths, we add 0 instead of exponential random variables for all paths that finish early, and this amounts to replacing any elements greater than T by T. Because of this caveat, we also extend the domain of the parameters in the case of the subordinator to allow for the first parameter to equal 0. For example: if $\Delta t = 0$ and $\alpha, \beta > 0$, then $\Delta L_t \sim \text{Gamma}(\alpha \cdot 0, \beta) \equiv 0$. The same holds for the Inverse Gaussian process. Nevertheless, as seen in Table (2.1), simulating through vectorization is much more efficient than simulating one process at a time.

Table 2.1: Comparison of European call payoffs under vectorizatrion vs. simulating one trajectory at a time, with parameters $T = 1, \lambda^1 = 1, \lambda^2 = 0.5$

Number of Simulations	Computational time under vectorization (sec.)	Computational time under one by one simulation (sec.)
N=10	0.001915	0.012886
N=10^4	0.085080	8.238712
N=10^6	9.617767	879.354925

We note that computational time is very sensitive to the size of the matrix τ. In particular, if $\lambda^1, \lambda^2 \ll T$, the matrix τ has around $T/(\lambda^1 + \lambda^2)$ columns on average.

Computing the switching Levy process is more straightforward; instead of simulating a value at each point in a partition, we simulate the increments in each regime. At each step, Δt holds the duration at which each process remains in a single regime. ΔL is an array of random variables depending on the subordinator with each element having a different first parameter.

Computations for the payoff of a European call option are summarized in the following algorithm, where the subordinator is a Gamma process. A similar algorithm holds when the subordinator is an Inverse Gaussian process.

1. Initialize:

    (a) regime state $j = 1$

    (b) index $k = 2$

(c) first stopping time (N-by-1 matrix) given by $\tau_{:,k} = [e_1, ..., e_N]^T$ where $e_n \sim exp(\lambda^j)$ for $n = 1, ..., N$

2. Recursively define the continuous-time Markov Chain:

(a) while there exists at least one indice n such that $\tau_{n,k} < T$,

- redefine $k = k + 1$

- switch regime $j = j \mod 2 + 1$

- define the k'th column (N-by-1) array given by $\tau_{:,k} = \tau_{:,k-1} + [e_1, ..., e_N]^T$ where each $e_n \sim exp(\lambda^j)$ for $n = 1, ..., N$

(b) otherwise stop.

(c) Define (N-by-1) array $\tau_{:,1} = [0, ..., 0]^T$ and redefine the (N-by-1) array $\tau_{:,k} = [T, ..., T]^T$

(d) Define the (N-by-k) matrix $\tau = [\tau_{:,1}, ..., \tau_{:,k-1}, \tau_{:,k}]$

(e) For each pair of indices m, n such that $\tau_{m,n} > T$, redefine $\tau_{m,n} = T$

3. To simulate the switching process:

(a) Reinitialize initial regime state $j = 1$

(b) Initialize (N-by-1 array) $Z_{:,1} = [0, ..., 0]^T$

4. For $2 \leq k \leq length(\tau_{1,:})$

(a) define $\Delta t_{:,1} = \tau_{:,k} - \tau_{:,k-1}$

(b) simulate (N-by-1) gamma random variable $\Delta L_{:,1} \sim Gamma(\alpha^j \Delta t_{:,1}, \beta^j)$

(c) simulate (N-by-1) array of standard normal variables $\mathcal{N}_{:,1}$

(d) $\Delta Z_{:,1} = \mu^j \Delta L_{:,1} + \sigma^j \sqrt{\Delta L_{:,1}} \mathcal{N}_{:,1}$

(multiplication of column arrays is done component-wise)

(e) $Z_{:,1} = Z_{:,1} + \Delta Z_{:,1}$

5. Expected Call Option Payoff

C= $\exp(-rT) \frac{1}{N} \sum_{n=1}^{N} ((S_0 \exp(Z_{n,1}) - K)_+.$
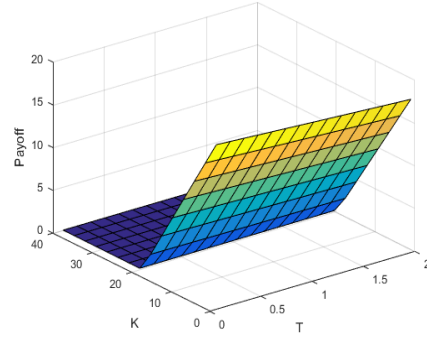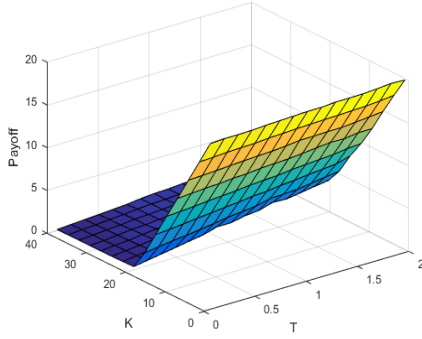
(a) Inverse Gaussian subordinator,
$\mu^1 = 0.3204, \mu^2 = 0.6450$

(b) Gamma subordinator,
$\mu^1 = -0.2316, \mu^2 = 0.0541$

Figure 2.3: European call payoff as a function of T and K under two different subordinators with risk neutral drift (see section 2.5). The other parameters are identical for each figure: $\sigma^1 = 0.03, \sigma^2 = 0.7, \alpha^1 = \alpha^2 = 0.1, \beta^1 = 1, \beta^2 = 1.2, \lambda^1 = 0.4, \lambda^2 = 1$.

Figures (2.3a) and (2.3b) demonstrate the behaviour of the payoff of a European call option under the regime switching time-changed Levy process model under Inverse Gaussian and Gamma subordinators. Both payoff models are monotone in T and K. Moreover as T increases, the expected payoff increases. For $K >> S_0$ the probability that $S_T \geq K$ is very small, therefore

38

the payoff is close to 0.

We can also estimate the price using confidence intervals. The confidence interval is useful because it provides a range of values that are likely to contain the population mean. At the $\gamma = 95\%$ confidence level, the endpoints of the confidence interval are given by:

$$\bar{x} \pm z_\gamma \frac{s}{\sqrt{N}},$$

where $\bar{x}$ is the sample mean, s is the standard deviation of the sample, N is the sample size and $z_{0.95} = 1.96$ is the z-score at the 95% confidence level.
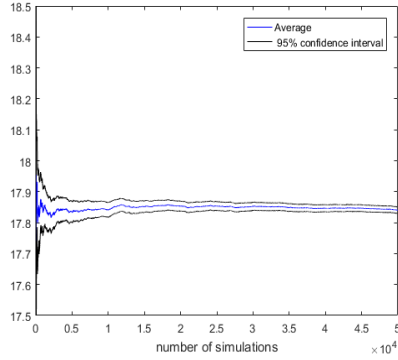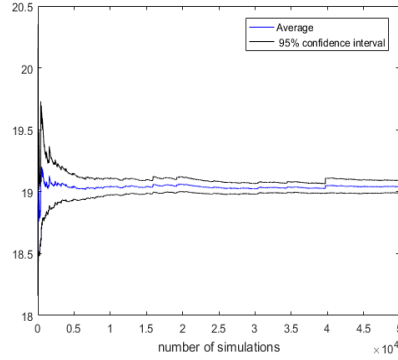
(a) Gamma subordinator   (b) Inverse Gaussian subordinator

Figure 2.4: Comparison of long term behaviour of the 95% confidence interval under different subordinators with identical parameters, including $\beta^1 = 0.1$, $\beta^2 = 0.01$.

The confidence interval decreases as the number of simulations approaches infinity, however in the case of the Inverse Gaussian subordinator, the interval is larger at each simulation because Inverse Gaussian random variables have a larger variance than Gamma random variables (see section 3.10) when

39

$\beta < 1$. Figures (2.4a) and (2.4b) show the behaviour of the payoff as different parameters are varied. At $N = 10^4$, the confidence interval when the subordinator is Inverse Gaussian is: $[18.7353, 19.1168]$. When the subordinator is a Gamma process, the confidence level is: $[\,17.8768, 17.9136]$.
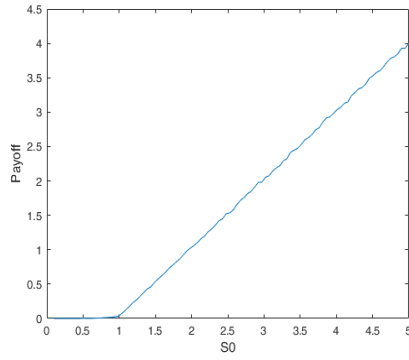
(a) Payoff as a function of the initial prices when $K = 1$.
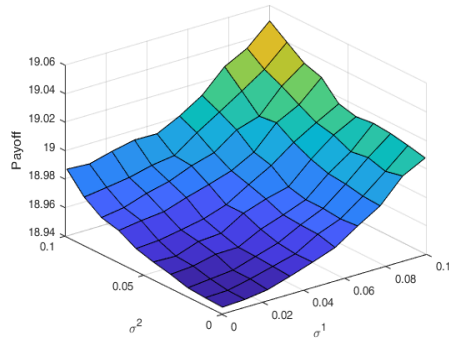
(b) Payoff as a function of σ^1 and σ^2

Figure 2.5: Behaviour of the payoffs when the parameters of the subordinator are varied.
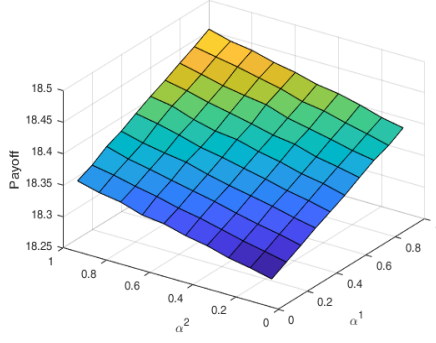
(a) Payoff as a function of parameters α^1, α^2
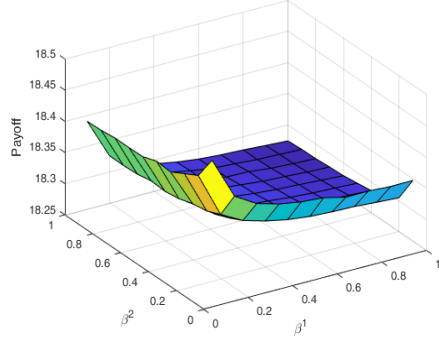
(b) Payoff as a function of parameters β^1, β^2

Figure 2.6: Behaviour of the payoffs when the parameters of the subordinator are varied.
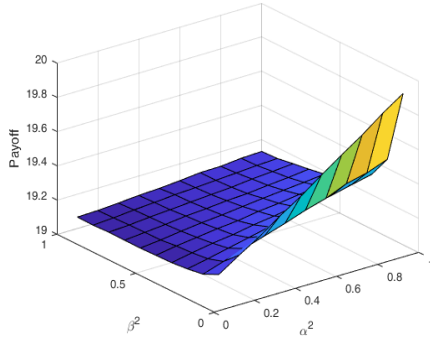
(a) Payoff as a function of parameters α^2, β^2
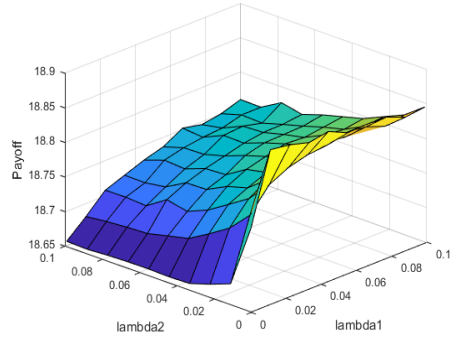
(b) Payoff as a function of parameters λ^1, λ^2

Figure 2.7: Behaviour of the payoffs when varying different parameters.

In (almost) each case, parameters between states were held constant un-

less they were being varied: $\mu^1 = \mu^2 = 0.01$, $\alpha^1 = \alpha^2 = \beta^1 = \beta^2 = 0.1$, $\sigma^1 = \sigma^2 = 0.01$, $\lambda^1 = \lambda^2 = 0.25$, $r = 0.04$, $T = 1$, $S_0 = 20$ and $K = 1$. Setting the parameters $\lambda^1 = \lambda^2$ implies the process spends an equal amount of time in each regime, on average. The only exception made is in Figure (2.7a), where $\lambda^1 = 0.001$ and $\lambda^2 = 10$ so that the process would spend a majority of the time in the second regime; this makes the process an approximation to a time-changed Levy process. When plotting Figure (2.7b), the domain had to be relatively small as the payoff varied greatly with changes in λ^1, λ^2.

The payoff of the European call option is monotone with respect to every parameter. In Figure (2.5b), the payoff increases as σ^1 and σ^2 increase. For changes in β^1, β^2, the payoff approaches an asymptote because Gamma and Inverse Gaussian random variables both have mean α/β, which approaches infinity for $\beta \to 0^+$. In Figure (2.7b), it is shown that the payoff decreases when $\lambda^1 \to 0^+$, because the process then spends more time in regime 2 on average, and we let $\mu^2 = -\mu^1 < 0$.

By comparing the payoff at different maturities Table (2.2), we demonstrate that the regime switching time-changed Levy process has more variability compared to the payoff under Black Scholes under identical parameters. This variability can be adjusted at will because the switching model has two additional types of parameters comparing to Black Scholes, namely the parameters characterizing the subordintor: α and β. The difference in the payoff between subordinators can be seen by the fact that, even though the expected value is the same in both cases, namely α/β, the variance (and

Table 2.2: Comparison of various European call payoffs under different strike prices and maturities. We compare Black Scholes to the switching model under Gamma and Inverse Gaussian subordinators. We set $r = 0.04, \sigma^1 = \sigma^2, \alpha^1 = \alpha^2, \beta^1 = \beta^2, \mu^1 = \mu^2$ where μ^1 is the Black Scholes risk neutral drift.

(T, K)	Gamma subordinator	Inverse Gaussian subordinator	Black Scholes model
$(1, 1)$	18.3364	18.2022	19.0368
$(2, 1)$	18.5638	19.6440	19.0768
$(1, 2)$	17.0453	16.0378	18.0784
$(2, 2)$	17.5120	16.7613	18.1537

other moments) differs. In the case of a Gamma subordinator, the variance is equal to α/β^2, and in the case of an Inverse Gaussian subordinator, the variance is equal to α/β^3.

2.4 Characteristic function of the Regime Switching time-changed Levy Model process

For any regime switching Levy process $M = \{M_t\}_{t \geq 0}$ which switches between n Levy processes $X^1, ..., X^n$ according to the continuous-time Markov chain $\{s_t\}_{t \geq 0}$ with generator matrix Q, the characteristic function is given by Choudarski [5]:

$$\mathbb{E}^{(x, \mathbf{\Pi})} \exp(iuM_t) = \exp(iux) \cdot [\mathbf{1}^T \cdot \exp(t \cdot \mathbf{\Phi}(u)) \cdot \mathbf{\Pi}], \qquad (2.7)$$

where $\mathbf{\Pi}$ is the initial regime distribution - an $n \times 1$ array where the j'th element is the probability that Z will begin in state j, i.e. $s_0 = j$. The term $x \equiv \log S_0$ is the initial log-price of the underlying asset, $\mathbf{1} = [1, 1, ..., 1]^T$ is a $n \times 1$ vector and $\mathbf{\Phi}(u)$ is an $n \times n$ matrix with elements:

$$[\mathbf{\Phi}(u)]_{i,j} = \begin{cases} q(i,i) + \phi_{X^i}(u), & \text{if } j = i \\ q(j,i), & \text{otherwise.} \end{cases}$$

Here $q(i,j)$ is the i,j'th element of the infinitesimal generator matrix Q and $\phi_{X^i}(u)$ is the characteristic exponent of the i'th base process X^i. The exponential term refers to the exponential of a matrix. If $\mathbf{\Phi}$ is a $n \times n$ matrix:

$$e^{\mathbf{\Phi}} = \sum_{k=1}^{\infty} \frac{\mathbf{\Phi}^k}{k!}.$$

We attempted to simplify the exponential by using singular value decomposition however, this resulted in eigenvalues and eigenvectors that did not simplify the computations. Instead, the method used to numerically compute the exponential is the scaling and squaring algorithm [14]. The method is based on the following approximation:

$$e^{\mathbf{\Phi}} = (e^{2^{-s}\mathbf{\Phi}})^{2^s} \approx r_m(2^{-s}\mathbf{\Phi})^{2^s},$$

where $r_m(x)$ is the $[m/m]$ Pade approximant of e^x and the nonnegative integers m and s are chosen in such a way as to achieve minimum error at minimal cost. A table of errors as a function of s and m is given in [1]. The $[k/m]$ Pade approximant for the exponential function is given by

$$r_{km(x)} = p_{km}(x)/q_{km}(x),$$

44

where

$$p_{km}(x) = \sum_{j=0}^{k} \frac{(k+m-j)!k!}{(k+m)!(k-j)!} \frac{x^j}{j!}, \quad q_{km}(x) = \sum_{j=0}^{m} \frac{(k+m-j)!m!}{(k+m)!(m-j)!} \frac{(-x)^j}{j!}.$$

We develop the characteristic function for the regime switching time-change Levy process where conditional on each regime, the Levy process is Brownian motion with a subordinator.

Given the process Z defined in equation (2.4) along with a continuous-time Markov chain with the infinitesimal generator matrix Q defined as equation (2.2), the characteristic function is given by equation (2.7) where:

$$[\mathbf{\Phi}(u)]_{i,j} = \begin{cases} q(i,i) + t^{-1} \log \mathcal{L}_{L_t^i}(\phi_{X^i}(u)), & \text{if } j = i \\ q(j,i), & \text{otherwise.} \end{cases}$$

Here the term $t^{-1} \log \mathcal{L}_{L_t^i}(\phi_{X^i}(u))$ is the characteristic exponent of the i'th time-changed Levy process $L^i = \{L_t^i\}_{t \geq 0}$.

We will now find the characteristic function for switching time-changed Levy processes under different subordinators. The set of regimes is again $E = \{1,2\}$ and we assume that the process will always start out in state 1 with probability 1, and hence the initial probability distribution is given by $\mathbf{\Pi} = [1,0]^T$. Working conditionally on each regime $j \in E$, the characteristic exponent of the j'th base process $\{X_t^j\}_{t \in [0,T]}$ where $X_t^j = \mu^j dt + \sigma^j B_t$ with j'th Levy triplet $(\mu^j, (\sigma^j)^2, 0)$ is derived from equation (1.8):

$$\phi_{X^j}(u) = -\mathrm{i}\mu^j u + (\sigma^j)^2 u^2/2.$$

The associated Laplace transform $\mathcal{L}_{L_t^j}(s)$ of the density function of the random variable L_t^j is derived from the characteristic function via a change of

variables when the characteristic function is known in closed form. When the subordinator L^j is an Inverse Gaussian process with shape parameter α^j and rate parameter β^j, we have:

$$\mathcal{L}_{L_t^j}(s) = \exp(\alpha^j t\sqrt{2s + (\beta^j)^2} - \beta^j).$$

Therefore the characteristic function of the j'th time-changed process $Y^j = \{X_{L_t}^j\}_{t\geq 0}$ under Inverse Gaussian subordinator is:

$$\psi_{Y_t^j}(u) \equiv \mathbb{E}[\exp(iuX_{L_t^j})] = \exp(\alpha^j t\sqrt{2(-i\mu^j u + (\sigma^j)^2 u^2/2) + (\beta^j)^2} - \beta^j)$$

(2.8)
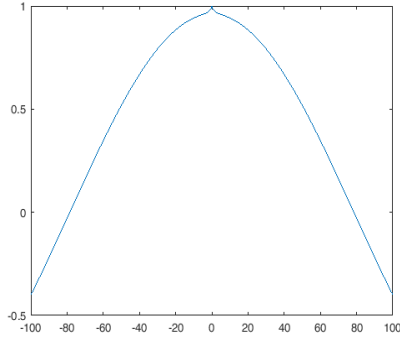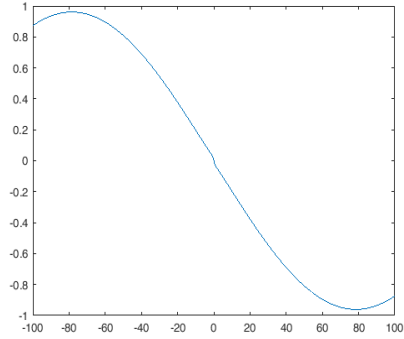
(a) The real part of $\psi_{Y_t}(u)$ (b) The imaginary part of $\psi_{Y_t}(u)$

Figure 2.8: The Characteristic function $\psi_{Y_t}(u)$ with parameters $\alpha = 0.1, \beta = 0.1, \mu = 0.01, \sigma = 0.01, t = 1/250$ where the subordinator L has increments with Inverse Gaussian distribution.

We now give the characteristic function of $Z = \{Z_t\}_{t\geq 0}$ where $Z_t = Y_t^{s_t}$ and the subordinator is an Inverse Gaussian process:

46

Corollary 0.1. The characteristic function of Z_t under an Inverse Gaussian subordinator is given by:

$$\psi_{Z_t}(u) = \mathbb{E}[\exp(iuZ_t)] = \exp(iux) \cdot [\mathbf{1}^T \cdot \exp(t \cdot \boldsymbol{\Phi}(u)) \cdot \boldsymbol{\Pi}], \qquad (2.9)$$

where

$$[\boldsymbol{\Phi}(u)]_{i,j} = \begin{cases} -1/\lambda^j + \alpha^j(\sqrt{2(-i\mu^j u + (\sigma^j)^2 u^2/2) + (\beta^j)^2} - \beta^j), & \text{if } j = i \\ 1/\lambda^j, & \text{otherwise.} \end{cases}$$
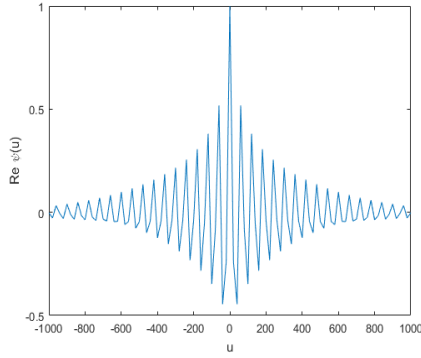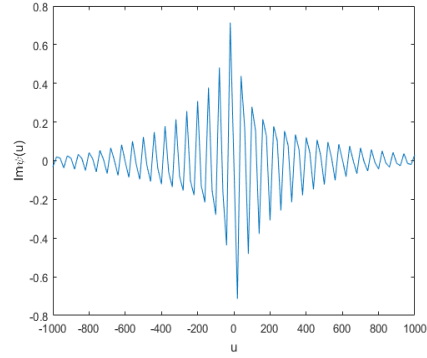
(a) The real part of $\psi_{Z_t}(u)$ (b) The imaginary part of $\psi_{Z_t}(u)$

Figure 2.9: The Characteristic function $\psi_{Z_t}(u)$ with parameters $\alpha^1 = 0.1 = \alpha^2, \beta^1 = 0.1, \beta^2 = 0.01, \mu^1 = 0.01, \mu^2 = -0.01, \sigma^1 = 0.01, \sigma^2 = 0.05, t = 1/250$ where conditionally on each regime, the subordinator L has increments with Inverse Gaussian distribution.

When the subordinator L^j is a Gamma process with shape parameter α^j and rate parameter β^j, we have:

$$\mathcal{L}_{L_t^j}(s) = \left(1 + \frac{s}{\beta^j}\right)^{-\alpha^j t}$$

47

Then the characteristic function of the j'th time-changed process $Y^j = X_L^j$ is:

$$\psi_{Y_t^j}(u) \equiv \mathbb{E}[\exp(iuX_{L_t^j})] = \left(1 + \frac{-i\mu^j u + (\sigma^j)^2 u^2/2}{\beta^j}\right)^{-\alpha^j t}. \qquad (2.10)$$
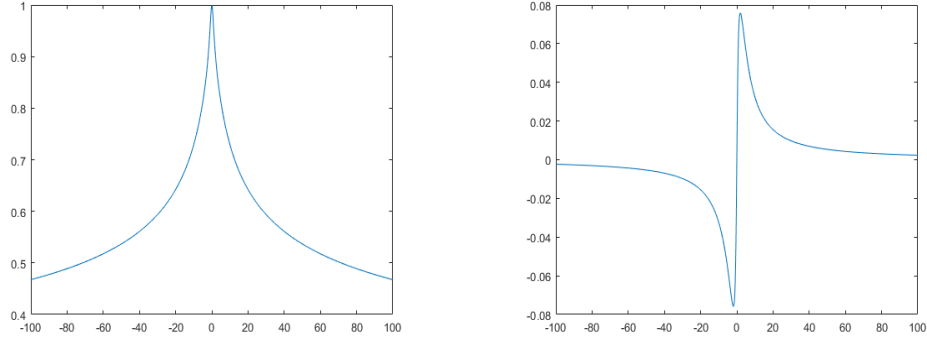
(a) The real part of $\psi_{Y_t}(u)$ (b) The imaginary part of $\psi_{Y_t}(u)$

Figure 2.10: The characteristic function $\psi_{Y_t}(u)$ with parameters $\alpha = 0.1, \beta = 0.1, \mu = 0.01, \sigma = 0.01, t = 1/250$ where the subordinator L has increments with Gamma distribution.

We now give the characteristic function of $Z = \{Z_t\}_{t\geq 0}$ where $Z_t = X_{L_t}$ and the subordinator is an Inverse Gaussian process:

Corollary 0.2. The characteristic function of $\{Z_t\}_{t\geq 0}$ under a Gamma subordinator is given by equation (2.9) where:

$$[\mathbf{\Phi}(u)]_{i,j} = \begin{cases} -1/\lambda^j - \alpha^j \log(1 + (-i\mu^j u + (\sigma^j u)^2/2)/\beta^j), & \text{if } j = i \\ 1/\lambda^j, & \text{otherwise.} \end{cases}$$

48
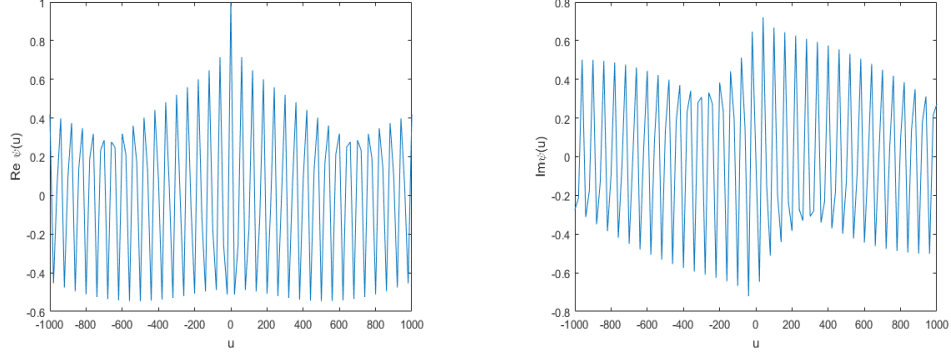
(a) The real part of $\psi_{Z_t}(u)$ (b) The imaginary part of $\psi_{Z_t}(u)$

Figure 2.11: The characteristic function $\psi_{Z_t}(u)$ with identical parameters as 2.11, but now conditionally on each regime, the subordinator L has increments with Gamma distribution. Note the thicker tails.

2.5   Pricing under a Risk Neutral Measure

Consider a Market model consisting of one riskless asset $B = \{B_t\}_{t \geq 0}$ where $B_t = \exp(rt)$ for some $r \geq 0$, and exponential Levy processes $S^j = \{S^j_t\}_{t \geq 0}$ where $S^j_t = S_0 \exp(Y^j_t)$. The discounted price processes $\tilde{S}^j = \{S^j_t\}_{t \geq 0}$ are defined as

$$\tilde{S}^j_t = \exp(-rt)S^j_t \quad \text{where} \quad j = 1, 2.$$

It is convenient to work under a risk neutral measure, also called an equivalent martingale measure (EMM) $\mathbb{Q}$ when pricing contingent claims on a stock, as discounted prices are $\mathbb{Q}$-martingales. If the risky asset is modelled under an underlying Levy process with jumps, then the market is incomplete, and there are infinitely many EMMs. This in turn implies that the parameters,

under which the underlying process is a martingale under $\mathbb{Q}$, are not unique.

From [23] we have that under an EMM $\mathbb{Q}$, the discounted price process $\tilde{S}^j$ is a martingale under $\mathbb{Q}$ if and only if the following equation is satisfied:

$$\phi_{Y^j}(-i) = r, \tag{2.11}$$

where $\phi_{Y^j}(u)$ is the characteristic exponent of the Levy process Y^j. The process Y^j is time-changed Brownian motion 2.3; when the subordinator is the Gamma process with parameters α^j, β^j, the characteristic function of the process is given by equation (2.10) and hence in each state $j \in E$ we solve:

$$\psi^j_{Gamma}(-i) = t^{-1} \log\left[\left(1 + \frac{i\mu^j u - \frac{(\sigma^j)^2 u^2}{2}}{\beta^j}\right)^{-\alpha^j t}\right]\bigg|_{u=-i}$$

$$\Longleftrightarrow r = -\alpha^j \log(1 + \frac{-\mu^j + (\sigma^j)^2/2}{\beta^j})$$

$$\mu^j = -\beta^j(\exp(-\frac{r}{\alpha^j}) - 1) + \frac{(\sigma^j)^2}{2}. \tag{2.12}$$

Holding all the other parameters constant, the drift verifies equation (2.11). The value of μ^j is such that the j'th discounted price process when the subordinator is a Gamma process is a martingale.

When the process Y^j is a time-changed process subordinated by an Inverse Gaussian process with parameters α^j, β^j, the characteristic function is given by equation (2.8) and therefore for each state $j \in E$ we solve for μ^j:

$$\psi^j_{IG}(-i) = t^{-1} \log\left[\exp(-\alpha^j t\left(\sqrt{2(-i\mu^j u + \frac{(\sigma^j)^2 u^2}{2}) + (\beta^j)^2} - \beta^j\right))\right]\bigg|_{u=-i}$$

$$\Longleftrightarrow r = -\alpha^j\left(\sqrt{2(\mu^j - \frac{(\sigma^j)^2}{2}) + (\beta^j)^2} - \beta^j\right)$$

$$\mu^j = \frac{1}{2}[(\beta^j - \frac{r}{\alpha^j})^2 + (\beta^j)^2] + \frac{(\sigma^j)^2}{2}. \tag{2.13}$$

Holding all the other parameters constant, the drift verifies equation (2.11). The value of μ^j is such that the j'th discounted price process when the subordinator is an Inverse Gaussian process is a martingale.

We can reduce the regime switching Levy process (2.4) to the Black Scholes model (2.1) by defining the subordinator $L_t = t$ and setting the parameters $\mu, \sigma, \alpha, \beta$ equal across all regimes:

$$\mu^1 = \mu^2, \quad \sigma^1 = \sigma^2, \quad \alpha^1 = \alpha^2, \quad \beta^1 = \beta^2.$$

Setting the drift such that the discounted Black Scholes price process is a martingale:

$$\mu^1 = r - \frac{1}{2}(\sigma^1)^2,$$

we find that the expected European call payoff is close to the payoff given by the famous Black Scholes formula:

Table 2.3: European call option payoff comparison between the Black Scholes formula and Monte Carlo simulation of the reduced switching Levy process at different parameters

Parameters (T, K, r, σ)	Reduced Switching Levy Model (# simulations $N = 10^6$)	Black Scholes formula
(1,1,0.04,0.5)	19.0463	19.0392
(3,1,0.1,1)	19.2955	19.3139
(2,30,0.5,0.001)	8.963608	8.96361

2.6 Fourier-Cosine Method

An alternative method of finding the fair price of options is via a Fourier-Cosine Series expansion. The derivation of the following method is given by [9]. For many stochastic processes, the probability density function is not known, however, the characteristic function might be known explicitly, as in the case of Levy processes where the characteristic function is computed using the Levy-Khintchine formula (1.0.11).

The risk-neutral valuation for European options is as follows:

$$v(x, t_0) = e^{-r\Delta t}\mathbb{E}^{\mathbb{Q}}[v(y, T)|x] = e^{-r\Delta t}\int_{\mathbb{R}} v(y, T)f(y|x)dy, \qquad (2.14)$$

where $v(x, t_0)$ is the payoff of the option as a function of the log-asset price $x := \ln(S_{t_0}/K)$ and initial time t_0 and $v(y, T)$ is the payoff at maturity time T and log-asset price $y := \ln(S_T/K)$. S_0 and S_T are the prices of the underlying asset at t_0 and T respectively. $\Delta t = T - t_0$ is the interval between the initial time and maturity time, K is the strike price and r is the risk-neutral interest rate. $\mathbb{E}^{\mathbb{Q}}$ is the expectation value operator with respect to some risk-neutral probability measure $\mathbb{Q}$ and $f(y|x)$ is the probability density function of y given x. There are different ways of computing the integral in (2.14) such as using an Inverse Fast Fourier Transform method, or by approximating the expected value using Monte Carlo simulation. In this section, the solution to (2.14) is given by truncating the integral appropriately and replacing the unknown density function f with the characteristic function ψ which was derived for regime switching time-changed Levy processes under Gamma and Inverse Gaussian subordinators in the previous section.

52

The domain of integration can be truncated to a finite interval for the purposes of numerical integration, because in order for the integral in (2.14) to be finite, the integrand must converge to 0 at $\pm\infty$. Fourier-Cosine expansions have been shown to give an optimal approximation of functions with finite support, while the usual Fourier series expansion is optimal when the function is periodic[17]. By choosing an interval $[a, b] \subset \mathbb{R}$ appropriately, the density function f can be approximated by a function f' which is equal to f on $[a, b]$ and 0 otherwise. We denote approximations to f and later A_k, ψ and v by f', A'_k, ψ' and v' respectively.

Every density function f has an associated characteristic function ψ which is the Fourier Transform, and similarly for f':

$$\psi(u) = \int_{\mathbb{R}} e^{iux} f(x) dx \quad \text{and} \quad \psi'(u) = \int_a^b e^{iux} f'(x) dx,$$

Given the function $f'(y|x)$ compactly supported on $[a, b]$, its Fourier-cosine expansion is given by:

$$f'(y|x) = \sum_{k=0}^{\infty} A'_k(x) \cdot \cos\left(k\pi \frac{y - a}{b - a}\right), \tag{2.15}$$

where the first term of the summation is weighted by one-half. The coefficients of f are denoted by A_k and the coefficients of f_1 are denoted by A'_k. The terms $A'_k(x)$ are given by:

$$A'_k(x) = \frac{2}{b - a} \int_a^b f_1(y|x) \cdot \cos\left(k\pi \frac{y - a}{b - a}\right) dy. \tag{2.16}$$

The coefficients A'_k written above can be rewritten in terms of the characteristic function:

$$A'_k(x) = \frac{2}{b - a} \text{Re}\left\{\psi'\left(\frac{k\pi}{b - a}; x\right) \cdot \exp\left(-ik\pi \frac{a}{b - a}\right)\right\}.$$

If the interval of support $[a, b]$ is chosen so that the integral on $[a, b]$ approximates an infinite integral appropriately, we can then approximate ψ_1 with ψ which is known, and A_k with A'_k:

$$A'_k(x) \approx A_k(x) = \frac{2}{b-a} \mathrm{Re} \left\{ \psi \left(\frac{k\pi}{b-a}; x \right) \cdot \exp \left(-\mathrm{i} k\pi \frac{a}{b-a} \right) \right\}.$$

The payoff can be approximated as:

$$v(x, t_0) \approx v_1(x, t_0) = e^{-r\Delta t} \int_a^b v(y, T) \sum_{k=0}^{\infty} A_k(x) \cos \left(k\pi \frac{y-a}{b-a} \right) dy. \quad (2.17)$$

If we define V_k to be the Fourier coefficients of $v(y, T)$:

$$V_k := \frac{2}{b-a} \int_a^b v(y, T) \cos \left(k\pi \frac{y-a}{b-a} \right) dy \quad (2.18)$$

we can rewrite

$$v_1(x, t_0) = \frac{1}{2}(b-a) e^{-r\Delta t} \sum_{k=0}^{\infty} A_k(x) V_k.$$

Our pricing formula is now in terms of A_k and V_k, which are the Fourier coefficients of $v(y, T)$ and $f(y|x)$ respectively. Putting everything in and truncating the infinite series to N terms, we get the generic payoff formula:

$$v(x, t_0) \approx e^{-r\Delta t} \sum_{k=0}^{N-1} \mathrm{Re} \left\{ \psi \left(\frac{k\pi}{b-a}; x \right) e^{-\mathrm{i} k\pi \frac{a}{b-a}} \right\} V_k. \quad (2.19)$$

The payoff of an option v can be computed from of the characteristic function instead of the density function. The characteristic function is computed in section 2.4 for Inverse Gaussian and Gamma subordinators. Fourier-cosine series of entire functions converges exponentially, so N need not be very large for good approximations [9].

The terms V_k defined in equation (2.18) depend on the option payoff formula. For European call and put options the expression for V_k can be obtained analytically. The payoff for European Vanilla options as a function of the log-asset price is:

$$v(y, T) \equiv \left[\eta \cdot K(e^y - 1) \right]^+ \text{ with } \eta = \begin{cases} 1 & \text{for a call option} \\ -1 & \text{for a put option.} \end{cases}$$

For the call option:

$$V_k^{Call} = \frac{2}{b-a} \int_0^b K(e^y - 1) \cos\left(k\pi\frac{y-a}{b-a}\right) dy,$$

and the solution is found by using integration by parts:

$$V_k = \frac{2}{b-a} K(\chi_k(0,b) - \psi_k(0,b)) \tag{2.20}$$

where:

$$\chi_k(c,d) = \frac{1}{1 + \left(\frac{k\pi}{b-a}\right)^2} \left[\cos\left(k\pi\frac{d-a}{b-a}\right)e^d - \cos\left(k\pi\frac{c-a}{b-a}\right)e^c \right.$$
$$\left. + \frac{k\pi}{b-a} \sin\left(k\pi\frac{d-a}{b-a}\right)e^d - \frac{k\pi}{b-a} \sin\left(k\pi\frac{c-a}{b-a}\right)e^c \right].$$

and

$$\psi_k(c,d) \equiv \begin{cases} \left[\sin\left(k\pi\frac{d-a}{b-a}\right) - \sin\left(k\pi\frac{c-a}{b-a}\right) \right]\frac{b-a}{k\pi} & k \neq 0 \\ d - c & k = 0. \end{cases}$$

For the European put option, the Fourier coefficients are found to be:

$$V_k^{Put} = \frac{2}{b-a} K(-\chi_k(a,0) + \psi_k(a,0)) \tag{2.21}$$

55

In pricing European call options, it was found that the payoff was not accurate and extremely sensitive to the values of b. It was also found that for large values of b, V_k^{Call} diverged to $\pm\infty$, while V_k^{Put} converged quickly and varied little with changes in a. We therefore rely on the put-call parity which allows for the computation of the European call option using the put option:

$$v^{Call}(x, t_0) = v^{Put}(x, t_0) + S_0 \exp(-qT) - K \exp(-rT),$$

where q is the number of dividends. We always take $q = 0$. It was found that the payoff under the call-put parity varied very little for large a and b, so it was arbitrarily chosen for $a = 0$ and $b = 10$.

2.7 Pricing European Call Options under the Fourier-Cosine Method

We now give an algorithm for pricing European call options using Fourier-Cosine Pricing (2.19) :

1. Initialize appropriate boundary points a, b such that the truncated integral will appropriately approximate the integral over $\mathbb{R}$, constant risk-neutral interest rate r, initial stock price S_0, strike price K, with $x := S_0/K$ time increment $\Delta t = T$, number of steps N

2. Initialize $N \times 1$ array of payoffs v^{Put} and v^{Call}

3. for $k = 0$ to $N - 1$

(a) Define the k'th element of v^{Put} to be:

$$v^{Put}(k) = e^{-rT}\text{Real}\left\{\psi_{Z_{\Delta t}}(k\pi/(b-a); x)\exp(-ik\pi(a/(b-a)))\right\}V_k^{Put};$$

where $\psi_{Z_{\Delta t}}$ depends on the choice of the subordinator, equation (2.9) for Inverse Gaussian processes or equation (2.10) for Gamma processes respectively. V_k^{Put} is defined in (2.21)

(b) $v^{Call}(k) = v^{Put}(k) + S_0 - Ke^{-rT}$ (put-call parity)

4. $v_{final}^{Call} = \frac{1}{2}v(1) + \sum_{k=1}^{N-1} v(k)$ (Summation)

We compare the European call payoff and running time under Monte Carlo simulation and Fourier-Cosine pricing in table 2.4 : It was found that

Table 2.4: Comparison of European Call option Payoffs using Monte Carlo Simulation and Fourier-Cosine Pricing, as well as their computational times.

(T, K)	Monte Carlo	Running Time (sec.)	Fourier-Cosine	Running time (sec.)
$(1,1)$	18.9554	9.31	19.0401	0.0234
$(2,1)$	19.9612	17.54	20.3456	0.1433
$(1,2)$	17.9942	10.01	18.2164	0.1339
$(2,2)$	19.0166	11.40	18.5523	0.193

the difference between pricing European call options using Monte Carlo and Fourier-Cosine pricing remained constant for different strike prices. The error however grows linearly for increasing maturity times.
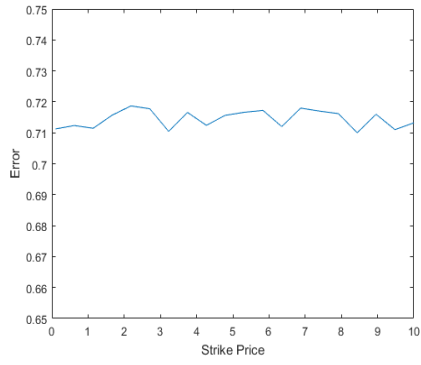
(a) Error as a function of K (b) Error as a function of T
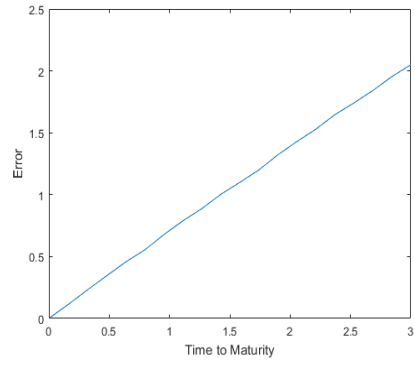
Figure 2.12: The difference between Fourier-Cosine Pricing and Monte Carlo, as a function of the strike price K and time to maturity K.

Chapter 3

Parameter Estimation and Calibration

In this chapter we estimate the value of the parameters using historical data. We use 4 different techniques: calibration using option data, method of moments, minimum distance estimation and maximum likelihood method. The data used was option and historical data of energy commodities - namely oil and electricity - because energy prices often exhibit sudden, drastic changes which make them suitable candidates for study under regime switching models.

We use two approaches of fitting the parameters of the underlying model to financial historical data: calibration and estimation. In calibration, the parameters are estimated by minimizing the error between option payoffs obtained numerically and option quotes. The option quotes are taken at a

variety of strike prices and times to maturity from Bloomberg. In parameter estimation, we use a variety of techniques based on historic asset prices: method of moments, minimum distance method and maximum likelihood estimation.

For the estimation methods we use daily historical NYMEX West Texas Intermediate (WTI) crude oil futures (11-16-2012 to 06-05-2018) and IESO Ontario Zone 24H electricity average spot prices (06-06-2008 to 06-05-2018). For calibration, the parameters are estimated using European call option quotes of WTI crude oil. For convenience we assume that there are 250 trading days in a year with each trading day corresponding to $\Delta t = 1/250$ of unit time. The codes for numerical computation of the methods are found in the appendix.

Figure 3.1 plots the historic futures and average spot prices, as well as the log-returns. For historic electricity spot prices, the price sometimes moves below zero, implying a surplus of electricity produced during low demand. Because electricity produced by power suppliers must be consumed immediately, the supplier pays wholesale customers to buy the surplus energy [25]. All negative prices were modified to CAD $0.01 for estimation purposes.

We also compare the empirical density function of the log-returns of each commodity to the normal distribution under the same mean and variance parameters as the historical log-return process. The density kernel method is summarized in 3.4.

Table 3.1 summarizes common moments for each commodity, as well as
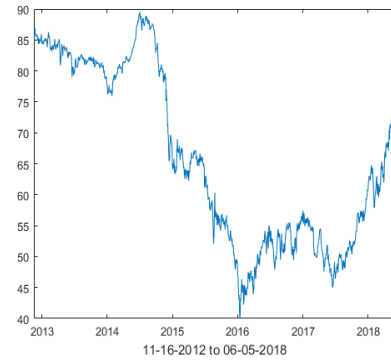
(a) WTI futures
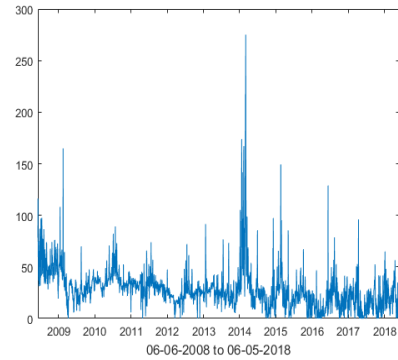
(b) Average electricity spot price
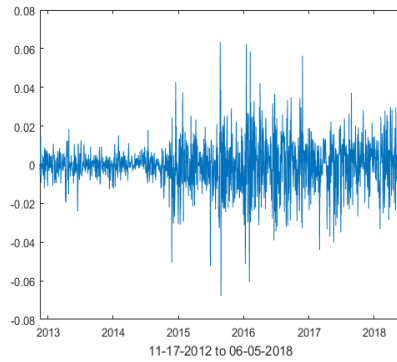
(c) Log-returns of WTI futures
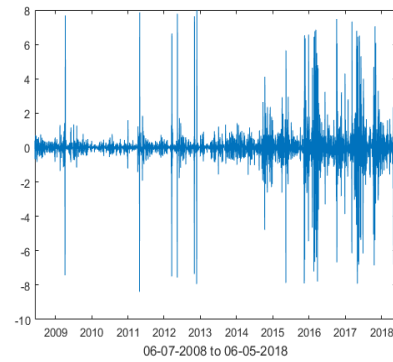
(d) Log-returns for Electricity spot price

Figure 3.1: Price Process and Log-returns for two different commodities

Source: Blooomberg Terminal, April 2018

for the log-return prices of each commodity.

The kurtosis for both log-returns is much larger than that of the normal distribution, implying heavier tails and high risk assets. The variance of the log-returns of electricity spot prices is several orders of magnitude larger than the variance of oil futures and this difference can be seen qualitatively by comparing the difference of the range of the log-return processes.

61
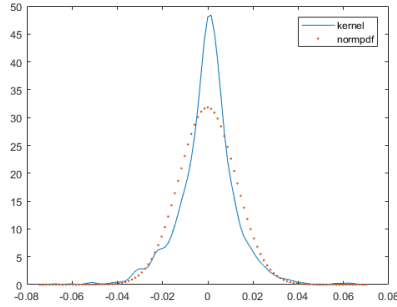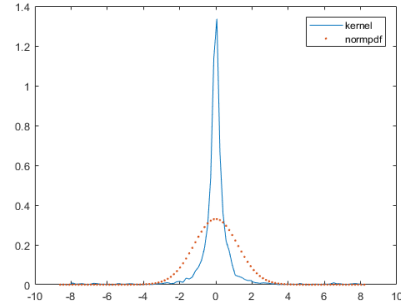
(a) oil log-returns

(b) electricity log-returns

Figure 3.2: Empirical density functions vs. normal distrubtions

Table 3.1: Standard statistics of commodities

Commodity	Mean	Variance	Skewness	Kurtosis
WTI crude oil futures	65.777	202.11	0.1750	1.4890
WTI crude oil log-return	-1.8577e-4	1.5707e-4	-0.1314	6.3383
Electricity spot price	28.011	380.45	3.2243	27.368
Electricity spot price log-return	-01.2039e-3	1.4581	-0.1377	23.947

Source: Bloomberg Terminal, April 2018

In our model 2.4, the regime switching time-changed Levy process can be in one of two states, therefore the parameters to be estimated are the holding-rate parameters λ^j and the four parameters of each time-changed Levy process $\theta^j = (\mu^j, \sigma^j, \alpha^j, \beta^j)$ for $j = 1, 2$. The locations of the regime changes are estimated by inspecting changes in the statistics of different regions of the historical log-return price processes. The random holding times are assumed to have exponential distribution. If X is an exponentially

distributed random variable with parameter λ, the expected value is given by $\mathbb{E}[X] = \lambda$.

We assume that the duration of the j'th observed historic regime is the most probable value i.e. it is equal to the expectation value λ^j. For each commodity, the j'th holding-rate parameter is given by:

$$\lambda^j = \frac{\text{total number of days in regime } j}{\text{number of occurrences of regime } j}, \tag{3.1}$$

where it is assumed that there are 250 trading days every year.

By inspecting the log-return process data of oil futures (3.1c), we define the process to be in regime 1 between 11-16-2012 and 11-16-2014 as well as between 02-06-2017 and 06-05-2018; otherwise, we assume that the process is in regime 2.

In the case of the log-returns of electricity spot prices; we define the process be in regime 2 whenever the absolute value of the log-returns exceeds 3 and in regime 1 otherwise.

We also include the variance of the log-returns within each regime; the different orders of magnitude between regimes justifies the use of a switching model.

By having defined the location of the regime changes and therefore estimated the values of λ^1, λ^2, the historic log-returns are separated into two sets of data, one containing all the data points for each regime. Each of these two new processes is assumed to follow a time-changed Levy process and we are left with estimating $\theta^j = (\mu^j, \sigma^j, \alpha^j, \beta^j)$ in each regime.

Commodity	$\hat{\lambda}^1$	$\hat{\lambda}^2$	Variance (regime 1)	Variance (regime 2)
Oil	0.900	3.80	2.6669e-05	2.1891e-04
Electricity	0.2618	0.0081	0.3624	37.3146

Table 3.2: Holding-rate parameters estimation for each commodity as well as the variance in each regime

The two states follow different parameters and are assumed to be uncorrelated. Define $\Theta^j \subset \mathbb{R}^4$ to be the set of all feasible parameters for θ^j. We assume that the two sets of parameters belong in different parameter spaces i.e. $\Theta^1 \neq \Theta^2$. The two parameters of the subordinator and the diffusion coefficient are required to be positive, therefore we add the following natural constraints: $\sigma^j, \alpha^j, \beta^j > 0$.

3.1 Calibration

In calibration we are given a set of option quotes with maturity times at a future date and therefore we need to estimate the future regimes using our estimated holding-rate parameters. We assume that the sequence and duration of future regimes follow the most probable path, therefore in the case of oil prices, the estimated holding-rate parameters imply that the continuous-time Markov chain remains in regime 1 for $0.9 \cdot 250 = 225$ trading days and then switches to regime 2 for $3.8 \cdot 250 = 950$ trading days. Table 3.2 gives the estimated holding-time rates for each commodity.

To estimate the parameters $\theta^j = (\mu^j, \sigma^j, \alpha^j, \beta^j)$ for $j = 1, 2$, within each

regime we minimize the root mean square error between the numeric option payoffs and European call option quotes. The numeric option payoff is simulated using Monte Carlo simulation (section 2.3), because the Fourier Cosine method exhibits significant error when the option is out of the money. We define the j'th objective function to be the root mean square error between the numeric payoffs $Call_{numeric}^{j}(\theta; T, K)$ - primarily a function of the parameters, and n option quotes $Call_{historical}^{j}(T, K)$, taken over a range of strike prices K and maturity times T which reside in regime j:

$$J^{j}(\theta) = \sqrt{\frac{1}{n} \sum_{T,K} (Call_{numeric}^{j}(\theta; T, K) - Call_{historical}^{j}(T, K))^2}, \quad j \in \{1, 2\}.$$

(3.2)

For each j, we define $\hat{\theta}^{j} = (\hat{\mu}^{j}, \hat{\sigma}^{j}, \hat{\alpha}^{j}, \hat{\beta}^{j})$ to be the estimate for the true parameter $\theta^{j} = (\mu^{j}, \sigma^{j}, \alpha^{j}, \beta^{j})$:

$$\hat{\theta}^{j} = \arg\min_{\theta \in \Theta^{j}} J^{j}(\theta).$$

(3.3)

We compute equation (3.3) numerically using the gradient method [30]. The gradient descent algorithm approximates $\hat{\theta}$ using the following algorithm:

1. Initialize the iterative step $t = 0$, the step size $\eta > 0$ and a starting point θ^{0}.

   Define some stopping criteria.

2. Repeat until stopping criteria is reached:

   (a) Redefine $\theta^{t+1} = \theta^{t} - \eta \nabla J(\theta^{t})$

65

(b) Iterate $t = t + 1$.

Here ∇J is the gradient of the cost function, $\eta > 0$ is the step size chosen accordingly. The `fmincon` function in MATLAB uses a variant of the gradient method. The particular variant we use is the Trust Region Reflective algorithm: instead of directly minimizing the cost function $J(\theta)$, the algorithm minimizes the first two terms of the Taylor series of J in a spherical neighborhood N called the trust region:

$$\min_{s \in N} q(s) \quad \text{where} \quad q(s) = \frac{1}{2}s^T \nabla^2 J s + s^T \nabla J \quad \text{and} \quad N = \{\|Ds\| \leq r\} \quad (3.4)$$

where $\nabla^2 J$ is the Hessian matrix of J, D is a diagonal matrix, $\| \cdot \|$ is the 2-norm and r is some positive scalar. The trial step s which minimizes (3.4) can be solved in a variety of ways such as a variant of Newton's method [16].

The modified algorithm is then given by:

1. Define D, r, initial θ and s.

2. Repear until stopping criteria is reached:

   Find the trial step s such that s minimizes (3.4)

   (a) If $J(\theta + s) < J(\theta)$, define $\theta = \theta + s$

   (b) Else : Decrease r (This corresponds to shrinking the neighborhood N)

   and repeat the process

We apply the above algorithm once for each regime. The final set of estimated parameters gives the best fit of the regime switching time-changed

66

Levy process to the option quotes - on average.

Table 3.3: Parameter Calibration using Root Mean Square Error

Commodity (subordinator)	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Oil log-return Regime 1 (Gamma)	-0.03387	0.0030	2.640710	1.007e-8
Oil log-return Regime 2 (Gamma)	-0.01445	1.116184	2.56567e-5	10.32441
Oil log-return 1 (Inverse Gaussian)	-0.04976	0.130011	0.24788	92.6926
Oil log-return 2 (Inverse Gaussian)	-0.04950	0.515891	8.531e-4	8.43091

The stopping criteria is taken to be step tolerance, taken to be equal to 1e-10. The step tolerance is a lower bound on the size of the step $(\theta^t - \theta^{t-1})$. The solver stops if the stopping criteria is reached, or if the maximum number of iterations (fixed to 1000 steps) is exceeded. One payoff of the Monte Carlo Simulation is computed using 1e6 simulations of the switch process and different initial starting points were found to give similar estimation of the parameters. Table 3.3 gives the estimated calibration in the case when the subordinator is a Gamma process and an Inverse Gaussian process. As expected, $\hat{\sigma}^2 > \hat{\sigma}^1$ in the case of both subordinators.

3.2   Method of Moments

The method of moments is a method used to estimate the parameters (again, possibly vector valued) θ from the historic log-return prices $x = (x_1, ..., x_n)$ by solving a system of equations [31]. Suppose each element of the sample is a

realization of some unknown random variable X with unknown distribution $f(x; \theta)$. Furthermore assume that the first d moments $\mu_1, \ldots, \mu_d$ can be written as functions of the parameters:

$$\mu_1 = \mathbb{E}[X] = g_1(\theta)$$
$$\vdots \qquad \qquad \vdots$$
$$\mu_d = \mathbb{E}[X^d] = g_d(\theta)$$

The idea is to find the j sample moments for each $j = 1, \ldots, d$ defined as:

$$\hat{\mu}_j = \frac{1}{n} \sum_{i=1}^{n} x_i^j \tag{3.5}$$

and then to solve the following system of equations:

$$\hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} x_i = g_1(\hat{\theta})$$
$$\vdots \qquad \qquad \vdots$$
$$\hat{\mu}_d = \frac{1}{n} \sum_{i=1}^{n} x_i^d = g_d(\hat{\theta})$$

for the vector valued estimator $\hat{\theta}$. The method of moments is based on the law of large numbers, namely that the sample moments converge to the moments of the distribution with probability 1 as the number of observations increase.

The moments can be computed from the derivatives of characteristic function, which are known for time-changed Levy processes:

$$\mathbb{E}[X^n] = i^{-n} \frac{d^n}{du^n} \psi(u) \Big|_{u=0} \tag{3.6}$$

This results in a system of nonlinear equations, which needs to be solved numerically.

When the subordinator L_t has Inverse Gaussian distribution with α, β, the resulting increments of length Δt have characteristic function:

$$\psi_{L_{\Delta t}}(u) = \exp(-\alpha \Delta t(\sqrt{2(-\mathrm{i}\mu u + (\sigma u)^2/2) + \beta^2)} - \beta)), \qquad (3.7)$$

where again $\mu \in \mathbb{R}$ and $\sigma, \alpha, \beta > 0$ and $\Delta t = 1/250$. Computing the first four moments results in a system of nonlinear equations:

$$0 = -\hat{\mu}_1 + (\alpha \mu \Delta t)/\beta$$

$$0 = -\hat{\mu}_2 + (\alpha \Delta t(\mu^2 + \beta^2 \sigma^2 + \alpha \beta \mu^2 \Delta t))/\beta^3$$

$$0 = -\hat{\mu}_3 + (\alpha \mu \Delta t(3\mu^2 + 3\beta^2 \sigma^2 + 3\alpha \beta \mu^2 \Delta t + \alpha^2 \beta^2 \mu^2 \Delta t^2 + 3\alpha \beta^3 \sigma^2 \Delta t))/\beta^5$$

$$0 = -\hat{\mu}_4 + (\alpha \Delta t(15\mu^4 + 3\beta^4 \sigma^4 + 18\beta^2 \mu^2 \sigma^2 + 15\alpha \beta \mu^4 \Delta t +$$

$$6\alpha^2 \beta^2 \mu^4 \Delta t^2 + \alpha^3 \beta^3 \mu^4 \Delta t^3 + 3\alpha \beta^5 \sigma^4 \Delta t + 6\alpha^2 \beta^4 \mu^2 \sigma^2 \Delta t^2 + 18\alpha \beta^3 \mu^2 \sigma^2 \Delta t))/\beta^7.$$

The system of equations is solved separately for each regime, therefore for two commodities there are a total of 4 systems of equations. The solution must at minimum satisfy the natural constraints: $\sigma, \alpha, \beta > 0$. We used the function `fsolve`, which solves a system of equations numerically using the trust region alogrithm, similarly to the function `fmincon`. The results are summarized in Table 3.5.

For convenience, we re-parametrize the Gamma subordinator. When the subordinator L_t is Gamma distributed with new parameters ρ, β where $\rho = \alpha/\beta > 0$, the resulting increments of length Δt have characteristic function:

$$\psi_{L_{\Delta t}}(u) = (1 + \frac{i\mu u + (\sigma u)^2/2}{\beta})^{-\rho\beta\Delta t}, \qquad (3.8)$$

where again $\mu \in \mathbb{R}$ and $\sigma, \alpha, \beta > 0$ and $\Delta t = 1/250$. Notice that ρ is the expected value of the Gamma random variable (see 3.10). Computing the first four moments results in a system of nonlinear equations:

$$0 = -\hat{\mu}_1 + \mu\rho\Delta t$$

$$0 = -\hat{\mu}_2 + \rho\sigma^2\Delta t + (\mu^2\rho\Delta t(\beta\rho\Delta t + 1))/\beta$$

$$0 = -\hat{\mu}_3 + (\mu\rho\Delta t(\beta\rho\Delta t + 1)(3\beta\sigma^2 + 2\mu^2 + \beta\mu^2\rho\Delta t))/\beta^2$$

$$0 = -\hat{\mu}_4 + (\rho\Delta t(\beta\rho\Delta t + 1)(6\mu^4 + 3\beta^2\sigma^4 + 12\beta\mu^2\sigma^2 + 5\beta\mu^4\rho\Delta t +$$

$$\beta^2\mu^4\rho^2\Delta t^2 + 6\beta^2\mu^2\rho\sigma^2\Delta t))/\beta^3.$$

Again we solve 4 systems of equations and the results are summarized in Table 3.4.

Table 3.4: Parameter Estimation using Method of Moments under Gamma Subordinator

Commodity	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Oil log-return Regime 1	0.0874	0.4979	2.3103	2.5059
Oil log-return Regime 2	0.0728	7.5945	3.125e-4	3.4075
Electricity log-return 1	0.1277	3.2767	1.0792	1.1205
Electricity log-return 2	-0.0328	13.3053	6.3567	1.1376

The method had trouble finding a global minimum in the case where the empirical moments were calculated using electricity log-return prices.

Table 3.5: Parameter Estimation using Method of Moments under Inverse Gaussian Subordinator

Commodity	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Oil log-return Regime 1	0.1624	0.7213	0.3238	1.6971
Oil log-return Regime 2	-0.0354	1.3402	0.0400	1.9584
Electricity log-return 1	0.1111	3.1233	28.4386	3.4862
Electricity log-return 2	-3.7405	19.5346	0.0132	0.3539

Changing the initial starting points resulted in varying results. Solving the system of equations using different algorithms and stopping criteria did not solve the problem. For the results in Table (3.4) and Table (3.5), we used as initial starting points the results from the minimum distance method (see Table (3.7) and Table (3.6)). In each case, the σ in regime 2 is much larger than that of regime 1.

3.3   Minimum distance estimation

Another approach to estimate the parameters $\theta^j = (\mu^j, \sigma^j, \alpha^j, \beta^j)$ of regime j is by minimizing the difference between the theoretical and empirical characteristic functions. While the Maximum Likelihood function can be unbounded, its Fourier Transform is necessarily bounded and therefore minimum distance estimation can work in cases where Maximum Likelihood method doesn't. Consider the sample $x_1, ...x_n$ of a random variable X with unknown cumulative distribution $F(x; \theta)$ depending on the parameters θ.

The characteristic function of X is defined as:

$$\psi(u;\theta) = \mathbb{E}[\exp(iuX)] = \int_{-\infty}^{\infty} \exp(iux)dF(x;\theta) \tag{3.9}$$

and the empirical characteristic function associated with the sample is defined as [32]:

$$\psi_n(u) = \frac{1}{n}\sum_{k=1}^{n} \exp(iux_k) = \int_{-\infty}^{\infty} \exp(iux)dF_n(x), \tag{3.10}$$

where $F_n(x)$ is the empirical cumulative distribution function:

$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{x_j \leq x}. \tag{3.11}$$

Define $\mathscr{F}$ to be the class of distribution functions. Define a non-negative distance function $d : \mathscr{F} \times \mathscr{F} \to \mathbb{R}$ such that:

$$d[\psi(u;\theta), \psi_n(u)] = \left(\int_{-\infty}^{\infty} |\psi(u;\theta) - \psi_n(u)|^p w(x)dx\right)^{1/p}, \tag{3.12}$$

where $w : \mathbb{R} \to (0,\infty)$ is a weight function [7]. We take $w(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$ and $p = 2$. Then $\hat{\theta}$ is the minimum distance estimate of θ if

$$d[\psi(u;\hat{\theta}), \psi_N(u)] = \inf_{\theta \in \Theta}\{d[\psi(u;\theta), \psi_N(u)]\}. \tag{3.13}$$

Again, we apply the algorithm to each regime separately. The empirical characteristic function is given by (3.10) and the analytic function (equation 3.9) is given by (3.8) when the subordinator is a Gamma process and (3.7) when the subordinator is an Inverse Gaussian process.

The integral is computed numerically using a global adaptive quadrature algorithm, where the interval of integration is subdivided and the integration

takes place on each subdivided interval. Intervals are further subdivided if the algorithm determines that the integral was not computed to sufficient accuracy.

Table 3.6: Parameter Estimation using Minimum Distance Method under Gamma subordinator

Commodity	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Oil log-return Regime 1	-0.2018	0.02975	0.421	10.039
Oil log-return Regime 2	0.19952	2.1951	0.00128	13.346
Electricity log-return 1	-0.494	0.00198	32.4815	0.0019
Electricity log-return 2	-0.1205	2.24089	0.0012	16.2528

Table 3.7: Parameter Estimation using Minimum Distance Method under Inverse Gaussian subordinator

Commodity	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Oil log-return Regime 1	0.01736	0.11675	31.648	8.0554
Oil log-return Regime 2	-0.4956	2.0078	2.2260	10.141
Electricity log-return 1	0.00813	02.0139	67.456	0.00154
Electricity log-return 2	5.7435	4.48714	76.004	6.871e-4

3.4 Maximum Likelihood Estimation

Given a random sample $x = (x_1, ..., x_n)$ of a random variable X with an associated density function $f(x; \theta)$ of the data x under the real world and

unknown parameters θ, maximum likelihood estimation (MLE) is a method used to estimate the vector valued parameter θ of the model by maximizing the likelihood function [29] :

$$\mathcal{L}(\theta; x) = \prod_{k=1}^{n} f(x_k; \theta); \quad \theta \in \Theta, \tag{3.14}$$

with respect to θ. The value of θ is constrained to $\Theta \subset \mathbb{R}^4$, the space of all feasible values of the parameters. The maximum likelihood function $\mathcal{L}$ is primarily a function of the unknown parameters θ. The maximum likelihood estimator is given by:

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta; x). \tag{3.15}$$

Finding the maximum (3.15) involves taking the derivative with respect to θ, or using numerical methods such as the gradient descent method.

To find the likelihood function $\mathcal{L}$ we require a sample and some density function. We derive the sample of daily log-returns by simulating $n = 10^6$ time-changed Levy process up to time $T = 1/250$. We approximate the density function by an empirical density function.

We use the kernel smoothing technique to approximate an empirical probability density function given the historical data [13]. A kernel smoothing function K has the following properties:

$$K(-x) = K(x), \quad K(x) \geq 0, \quad \text{and} \quad \int_{\mathbb{R}} K(x)dx = 1, \tag{3.16}$$

and has an associated scaled version K_h, given by:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right), \quad \text{where} \quad h > 0. \tag{3.17}$$

Assuming the log-return of the historical data $x_1, x_2, ...x_n$ are independent and identically distributed, the kernel density estimate of the probability distribution is given by:

$$\hat{f}(x) = \frac{1}{n} \sum_{j=1}^{n} K_h(x - x_j) = \frac{1}{nh} \sum_{j=1}^{n} K\left(\frac{x - x_j}{h}\right). \tag{3.18}$$

There are many possible choices for the Kernel K including the uniform, triangle and parabolic kernels; we have decided to use the Gaussian kernel:

$$K(x - x_i) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{(x - x_i)^2}{2}\right) \tag{3.19}$$

The value of h, is chosen to equal Silverman's value $h = 1.06\sigma n^{-1/5}$ where σ is the standard deviation of the data [21].

In each of the four cases, the values of σ were found to be higher in the

Table 3.8: Parameter Estimation using Maximum Likelihood Method under Gamma subordinator

Commodity	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Oil log-return Regime 1	0.0023	0.0431	42.928	11.9960
Oil log-return Regime 2	-0.372	0.52851	17.3008	88.556
Electricity log-return 1	5.844e-3	1.5002	93.271	2.1903
Electricity log-return 2	-0.0148	7.543	90.5900	0.01770

second regime, hence justifying the use of a regime switching model. In nearly every method, the value of $|\mu|$ was found to be very small, which is expected as the long term deterministic contribution to the process is expected to be near zero.

75

Table 3.9: Parameter Estimation using Maximum Likelihood Method under Inverse Gaussian subordinator

Commodity	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Oil log-return Regime 1	-0.4883	0.5058	0.64603	63.709
Oil log-return Regime 2	0.1201	2.9707	0.00014	9.993
Electricity log-return 1	-0.1781	0.25873	0.91878	20.0860
Electricity log-return 2	-0.0191	4.9752	5.512e-5	11.016

In choosing constraints, we set the lower bound of σ, α, β to be some small number $\epsilon = 10^{-6}$. We set the upper bound of σ to 5 as the diffusion is expected to be smaller than 1 and for α, β, we set the upper bound to be 100, as the expected value of both Inverse Gaussian and Gamma random variables depends on the ratio α/β rather than any particular value for α and β. The drift μ is expected to be small, so in most cases, it was constrained to the set $[-1, 1]$.

Conclusion

The goal of this thesis was to develop a new model by combining the regime switching Levy process, and the time-changed Levy process. We found the characteristic equation under Gamma and Inverse Gaussian subordinators, priced European call options under two pricing methods and estimated the parameters. Future work may include: using machine learning to rigorously estimate the location of the regimes based on characteristics of the historic data, as well as generate the most likely sequence of regimes into the future. The density function in the maximum likelihood method can be computed numerically in other ways, for example using the Inverse Fast Fourier Transform method. Finally, the number of regimes can be extended, multivariate processes can be explored, and pricing can be done on more exotic options.

Appendix A: Simulating Random Variables

Gamma Random Variables

We simulate Gamma Random random variables using Berman's Gamma Generator [27]. Given the shape parameter $\alpha > 0$ and rate parameter $\beta > 0$, the probability density function of the Gamma distribution is given as:

$$f_{\text{gamma}}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}. \tag{3.20}$$

Sometimes the gamma distribution is parametrized with the shape parameter $k > 0$ and scale parameter $\theta > 0$, which are related by $k = \alpha$ and $\theta = 1/\beta$.

If $X \sim \text{Gamma}(\alpha, \beta)$ then $X/c \sim \text{Gamma}(\alpha, \beta c)$, therefore all that is needed is to generate $\text{Gamma}(\alpha, 1)$ random variables. If $\alpha = 1$ the gamma distribution reduces to the exponential distribution with parameter β. We use Berman's Gamma Generator when $0 < \alpha \leq 1$:

1. 1 Simulate two independent random variables $u_1 \sim U(0, 1)$ and $u_2 \sim$

$U(0, 1)$

2. Set $x_1 = u_1^{1/\alpha}$ and $x_2 = u_2^{1/(1-\alpha)}$

3. If $x_1 + x_2 \leq 1$ move on to step 4, otherwise go back to step 1.

4. Simulate two independent random variables $v_1 \sim U(0, 1)$ and $v_2 \sim U(0, 1)$

5. Return the number $g = -\log(v_1 v_2)$, which is a Gamma$(\alpha, 1)$.

If $\alpha > 1$ we use Best's rejection method [8] to generate Gamma$(\alpha, 1)$:

1. Define $c = \alpha - 1$ and $d = 3\alpha - 3/4$.

2. Simulate two independent, identically distributed standard uniform random variables u, v.

3. Define $w = u(1 - u)$, $y = \sqrt{d/w}(u - 1/2)$ and $x = c + y$.

4. If $x < 0$, go to step 2.

5. Define $z = 64 \log w^3 v^3$

6. If $\log z > 2(c \log (x/c) - y)$, got to step 2.

7. Return x, which is a Gamma$(\alpha, 1)$ distributed random variable

Inverse Gaussian Random Variables

We use the rejection method of Michael, Schucaney and Haas to generate Inverse Gaussian random variables [27]. Given the parameters $\alpha > 0$ and

$\beta > 0$, the probability density function of the Inverse Gaussian distribution is given by:

$$f_{IG}(x) = \frac{\alpha}{2\pi} \exp(\alpha\beta) x^{-3/2} \exp(-\frac{1}{2}(\alpha^2 x^{-1} + \beta^2 x)), \quad x > 0 \qquad (3.21)$$

A random variable with the above density function can be simulated in the following way:

1. Simulate a standard normal variable z and let $v = z^2$.

2. Set $x = \alpha/\beta + v/(2\beta^2) - \sqrt{4\alpha\beta v + v^2}/(2\beta^2)$.

3. Generate a standard uniform random variable u.

4. If $u \leq \alpha/(\alpha + x\beta)$, then x is an Inverse Gaussian random variable with parameters α, β, else, $\alpha^2/(\beta^2 x)$ is the Inverse Gaussian random variable with parameters α, β.

Inverse Gaussian random variables can also be parametrized by the mean and shape parameters $\mu > 0, \lambda > 0$ respectively: as $\mu = \alpha/\beta$ and $\lambda = \alpha^2$.

The following table summarizes important moments for the Gamma and Inverse Gaussian distribution under the particular parametrization which is used throughout the thesis [27].

Common Moments Moments	Gamma distribution	Inverse Gaussian distribution
Mean	α/β	α/β
Variance	α/β^2	α/β^3
Skewness	$2(\alpha)^{-1/2}$	$3(\alpha\beta)^{-1/2}$
Kurtosis	$3 + 6(\alpha)^{-1}$	$3 + 15(\alpha\beta)^{-1}$

Table 3.10: Important moments of the Inverse Gaussian and Gamma distributions.

Appendix B: MATLAB Code

Simulating a Single Trajectory Regime Switching Time-Changed Levy Process

```
function Switching_Levy_Trajectory(T,N)
%Plots the trajectory of regime switching process Z

%T=Maturity time (in years)
%N=number of increments
%EX: Switching_Levy_2(1,250)
%Inverse Gaussian Subordinator

%drift coefficient
mu1=0.01; mu2=-0.1; mu=[mu1,mu2];

%diffusion coefficient
sigma1=1; sigma2=5; sigma=[sigma1,sigma2];
```

```
%parameter for exponential distribution i.e. holding times
lambda1=0.5; lambda2=0.25; lambda=[lambda1, lambda2];


%s_0=1 Continuous-time Markov chain begins in regime 1 with probability 1
state=1;


%Set up a single switching time:
Tau=exprnd(lambda(state));


%temp_state is only used to generate switching times
temp_state=state;


%these arrays are only used to plot continuous-time Markov chain
markov_t=[]; markov_st=[];


%initialize regime switching process
Z=zeros(N,1);


%[0,dt,2dt,3dt,..., Ndt ] where dt=T/N
time_vector=(0:N-1) * (T/N);


%generate stopping times
while Tau(end)<T
```

```
        Tau=[Tau, Tau(end)+exprnd(lambda(temp_state))];

        temp_state=mod(temp_state,2)+1;

end


for t=2:N

    %DELTA=[t-1,stopping times which occur between t-1 and t, t]
    DELTA=[time_vector(t-1), Tau(Tau>=time_vector(t-1) &&...
        Tau<time_vector(t)),time_vector(t)];


    dZ=0;


    %dZ is the Z increment betweem t-1 and t



    for i=2:length(DELTA)
        L=Lt(DELTA(i)-DELTA(i-1),state);
       dZ=dZ+mu(state)*L+sigma(state)*...
            sqrt(L)*randn;
       if i !=length(DELTA)
            state=mod(state,2)+1;  %change states
        end
    end
```

```matlab
    %Z=sum dZ
    Z(t)=Z(t-1)+dZ;


end
end
%    %    %    %    %    %    %    %    %    %
function L=Lt(dt,state)    %Subordinator increments


a1=0.1; b1=0.1;
a2=0.1; b2=0.01;
a=[a1,a2];
b=[b1,b2];
dist=makedist('InverseGaussian',a(state)*dt,b(state));
L=random(dist);
end
```

Monte Carlo Simulation for European Call Options by simulating N Processes Simultaneously

```matlab
function Expected_Price=MatrixLevy(T,K,N)
%T=maturity, K=strike price, N=#of independent processes


global alpha, global beta, global string, global state


string='Gamma';


%All other  necessary variables are sored in file Variables.m
run Variables



%generate a column of the first stopping time
ExpMatrix=zeros(N,1);
state=1;
while sum(ExpMatrix(:,end)<T)>0  %while there exist some final elements<T

    %add columns until all elements are >=T
    ExpMatrix=[ExpMatrix,ExpMatrix(:,end)+exprnd(lambda(state),N,1)];
    state=mod(state,2)+1;
end
```

```
%Replace all the elements > T by T
ExpMatrix(ExpMatrix>T)=T;


state=1;
Z_end=zeros(N,1);


% Final value of Z_end is a N by 1 column vector of price values at Maturity
for t=2:length(ExpMatrix(1,:))


    L=Lt(ExpMatrix(:,t)-ExpMatrix(:,t-1),N);
    dZ=Mu(state)*L+Sigma(state)*sqrt(L).*randn(N,1);
    Z_end=Z_end+dZ;
end
    Price=S0*exp(Z_end);
    EU_CALL=exp(-r*T)*max(Price-K*ones(N,1),0);
    Expected_Price=mean(EU_CALL);
end



%    %    %    %    %    %    %    %    %    %
%creates distribution for subordinator
function L=Lt(dt,N)    %subordinator


global alpha,global beta,global state
```

```
%alpha,beta are stored in Variables.m
%RVsimualtionMatrix generates Gamma and Inverse Gaussian random variables
%given by the rejection method in the apprendix
L=RVsimulationMatrix(alpha(state)*dt,beta(state),N);


end
```

Fourier-Cosine Pricing

```
function Price=COS_Pricing(T,K)


global dt, global string, global alpha, global beta, global Mu,
global Sigma, global lambda


N=2^10;


string='Gamma';
run Variables %All other variables are stored in this file


x=log(S0/K);
dt=(T-t0);
a=-5; b=5;
```

```
v3=zeros(N,1);

k=0:N-1;

Psi=zeros(N,1);

V_k=zeros(N,1);


for n=1:N
    %calls Characteristic, given below
    Psi(n)=Characteristic(k(n)*pi/(b-a),x);
    V_k(n)=V(a,b,k(n),K,'Put');
end


Re=real(Psi.*exp(-1i.*k*pi*a/(b-a))');


%Put-Call Parity
v3=exp(-r*dt).*Re.*V_k+S0-K*exp(-r*T);


Price=v3(1)*1/2 + sum(v3(2:end));


end


%   %   %   %   %   %   %   %   %   %


function Char=Characteristic(u,x)
%Characteristic function of Switching time-changed process
```

```
%u is the variable, x is initial price/strike price


global dt,global lambda,global Mu,global Sigma,
global alpha, global beta,global string



%infinitesimal generator matrix
Q=[-lambda(1),lambda(2);lambda(1),-lambda(2)];


if strcmp(string,'Gamma')
PHI=[-alpha(1)*log(1+(-1i*Mu(1).*u+(Sigma(1).*u)^2/2)/beta(1))),0;...
    0,-alpha(2)*log(1+(-1i*Mu(2).*u+(Sigma(2).*u)^2/2)/beta(2))];


elseif strcmp(string,'InverseGaussian')
PHI=[-alpha(1)*(sqrt(2*(-1i*Mu(1).*u+(Sigma(1).*u)^2/2)+beta(1)^2)...
    -beta(1))),0;0,-alpha(2)*(sqrt(2*(-1i*Mu(2).*u+...
    (Sigma(2).*u)^2/2)+beta(2)^2)-beta(2))];
end



  %matrix exponential expm
Char=exp(1i*u*x)*([1,1]*expm(dt*(Q+PHI))*[1,0]');
end
```

```
%    %    %    %    %    %    %    %    %    %


function Chi=Chi(a,b,c,d,k)
    w=k*pi/(b-a);
    Chi=(1/(1+w^2)) * (cos(w*(d-a))*exp(d) - cos(w*(c-a))*exp(c) + ...
    w*sin(w*(d-a))*exp(d) - w*sin(w*(c-a))*exp(c));
end


%    %    %    %    %    %    %    %    %    %


function Psi=Psi(a,b,c,d,k)
if k==0
    Psi=d-c;
else
    w=k*pi/(b-a);
    Psi=(w^-1) * (sin(w*(d-a))-sin(w*(c-a)));
end
end


%    %    %    %    %    %    %    %    %    %


function Vk=V(a,b,k,K,option) %Vk coefficients for Call and Put
```

```
if strcmp(option,'Call')

    Vk=(2/(b-a))*K*(Chi(a,b,0,b,k)-Psi(a,b,0,b,k));

elseif strcmp(option,'Put')

    Vk=(2/(b-a))*K*(-Chi(a,b,a,0,k)+Psi(a,b,a,0,k));

end

end
```

Minimum Distance Estimation Function

```
function Distance=Distance_minimizing_Estimation_function(theta)

%theta=[mu,sigma,alpha,beta]

global data,global string


string='InverseGaussian';

p=2;


%data=log-returns of historical data

Integrand_function=@(u)abs(Analytic_characteristic_function(u,...

    theta,string)-Empirical_characteristic_function(u,...

    data)).^p.*(1/sqrt(2*pi)).*exp(-u.^2/2);


%global adaptive quadrature with default tolearance

Distance=integral(Integrand_function,-1e6,1e6)^(1/p);
```

```matlab
end



%    %    %    %    %    %    %    %    %    %



function ECF=Empirical_characteristic_function(u,data)

%u is an array of arguments

ECF=0;



for j=1:length(u)

    ECF(j)=sum(exp(1i*u(j)*data));

end

ECF=ECF/length(data);



end

%    %    %    %    %    %    %    %    %    %



function CHAR=Analytic_characteristic_function(u,theta,string)

dt=1/250;

if strcmp(string,'Gamma')

    Mu=theta(1);

    Sigma=theta(2);

    rho=theta(3);

    beta=theta(4);
```

```matlab
        CHAR=(1+(1i*u*Mu-(u*Sigma).^2/2)./-beta).^(-rho*dt/beta);


elseif strcmp(string,'InverseGaussian')

    Mu=theta(1);

    Sigma=theta(2);

    alpha=theta(3);

    beta=theta(4);



    CHAR=exp(-alpha.*dt.*(sqrt(2*(1i.*u.*Mu-(u.*Sigma).^2/2)+beta^2)-beta));
end
end
```

Maximum Likelihood Estimation

```matlab
function MLE=MLE_function(x)


%data is log-returns of historical data
global string, global data
string='Gamma';


N=1e6;
dt=1/250;
```

```
Mu=x(1);

Sigma=x(2);

alpha=x(3);

beta=x(4);


%LevyMLE defined below

dY=Mu*LevyMLE(dt*ones(N,1),N,x)+...

        Sigma*sqrt(LevyMLE(dt*ones(N,1),N,x)).*randn(N,1);

[f_MLE,xi] = ksdensity(dY);


MLE=1;

for k=1:length(data)

%interpolation

    MLE=MLE*interp1(xi,f_MLE,data(k));

end

MLE=-MLE; %Because we will want to MAXIMIZE using fmincon

end


%   %   %   %   %   %   %   %   %   %


function L=LevyMLE(dt,N,x)

Mu=x(1);

Sigma=x(2);

alpha=x(3);
```

```
beta=x(4);

L=RVsimulationMatrix(alpha*dt,beta,N);
end
```

Simulating Gamma and Inverse Gaussian Random Variables

```
function arrayX=RVsimulationMatrix(alpha,beta,N)
global subordinator
Array=zeros(N,1);

for n=1:N

    if strcmp(subordinator,'Gamma') %To simulate Gamma R.V

    if alpha(n)==0
        X=0;
    elseif alpha(n)<=1 && alpha(n)>0     %Berman Gamma Generator
        x=rand()^(1/alpha(n));y=rand()^(1/(1-alpha(n)));

        while x+y>1
            x=rand()^(1/alpha(n));y=rand()^(1/(1-alpha(n)));
```

```
        end
    X=-x*log(rand()*rand())/beta;


elseif alpha(n)>1
    b=alpha(n)-1; c=3*alpha(n)-3/4;
    U=rand();V=rand();
    W=U*(1-U);
    Y=sqrt(c/W)*(U-1/2);
    x=b+Y;
    Z=64*W^3 * V^3;
    i=0;
    while 1==1
        i=i+1;
        U=rand();V=rand();
        W=U*(1-U);
        Y=sqrt(c/W)*(U-1/2);
        x=b+Y;
        if x>0

            Z=64*W^3 * V^3;

            if log(Z)<=2*(b*log(x/b)-Y)
                X=x/beta;
```

```
                    break
                end
            end
        end
    end




    elseif strcmp(subordinator,'InverseGaussian')
    %Simulate Inverse Gaussian  R.V
        if alpha(n)==0
            X=0;
        else
        y=randn()^2;
        x=(alpha(n)/beta)+y/(2*beta^2)-sqrt(4*alpha(n)*beta*y+y^2)/(2*beta^2);
            if rand()<=alpha(n)/(alpha(n)+x*beta)
            X=x;
            else
            X=alpha(n)^2/(beta^2*x);
            end
        end
    end

Array(n)=X;
```

```
end

arrayX=Array;

end
```

Bibliography

[1] Awad H. Al-Mohy and Nicholas J. Higham. A New Scaling and Squaring Algorithm for the Matrix Exponential. *SIAM J. Appl.*, 31(3):970–989, 2009.

[2] Gidi Amir. *Chapter 6: Continuous Time Markov Chains*.

[3] Catherine Beauchemin. *Computational Methods in Medical Physics*. 2014.

[4] Peter Carrr and Liuren Wu. Time-changed levy processes and option pricing. 2002.

[5] Kyriakos Choudarski. Switching lévy models in continuous time: Finite distributions and option pricing. *Center for Computational Finance and Economic Agents*, 2005.

[6] Kyriakos Choudarski. Multinomial method for option pricing under variance gamma. *Center for Applied Mathematics and Economics - ISEG*, 2018.

[7] Andreas N. Philippou Constantine A. Drossos. A note on minimuim distance estimates.

[8] Rama Cont and Peter Tankov. *Financial Modelling with Jump Processes.* CRC Press, London, United Kingdom, 2003.

[9] F.Fang and C.W. Oosterlee. A novel pricing method for european options based on fourier-cosine series expansions. *Journal on Scientific Computing*, 2008.

[10] Roman N. Makarov Giuseppe Campolieti. *Financial Mathematics: A Comprehensive Treatment.* CRC Press, Boca Raton, United States, 2014.

[11] David Williams Grimmett-Stirzaker. *Probability with Martingales.* Press Syndicate of the University of Cambridge, Cambridge, 1991.

[12] Elias S.W. Shiu Hans U. Gerber. Option pricing by esscher transforms. *Transactions of Society of Actuaries, Vol. 46*, 1994.

[13] Nathaniel E. Helwig. *Density and Distribution Estimation.*

[14] Nicholas J. Higham. The scaling and squaring method for the matrix exponential revisited. *SIAM J. Appl.*, 26(4):11791193, 2005.

[15] Magdalena Hyksova. *Stochastic Processses and their Classification.*

[16] Mor J.J. and D.C. Sorensen. Computing a trust region step. *SIAM Journal on Scientific and Statistical Computing*, 3:55357211791193, 1982.

[17] Boyd J.P. *Chebyshev Fourier spectral methods*. Springer Verlag, Berlin, 1989.

[18] Oliver Knill. *Probability Theory and Stochastic Processes with Applications*. Overseas Press, Daryanganj, New Delhi, 2009.

[19] Kevin Lin. *Notes on Memoryless Random Variables*.

[20] Ronnie Sircar Matthew Lorig. Stochastic volatility: Modeling and asymptotic approaches to option pricing and portfolio selection. pages 1–3, 2014.

[21] A.I. McLeod and B. Quenneville. *Mean Likelihood Estimators*. Statistics and Computing 11, 57-65, 2001.

[22] Antonis Papapantoleon. An introduction to levy processes with applications in finance. 2000.

[23] Andrea Pascucci. *PDE and Martingale Methods in Option Pricing*. Bocconi University Press, 2011.

[24] Philip Protter. *Martingales and Local Martingales*.

[25] Stefan Trck Rafa Weron, Michael Bierbrauer. *Modeling electricity prices: Jump diffusion and Regime Switching*. Physica A, Hugo Steinhaus Center, Wroclaw University of Technology, 2004.

[26] Sheldon M. Ross. *Stochastic Processes*. John Wiley and Sons, Inc., United States, 1996.

[27] Wim Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives*. John Wiley Sons, Ltd., London, United Kingdom, 2003.

[28] Karl Sigman. *Continuous-Time Markov Chains*.

[29] Kyle Smith. Estimation methods of reference evapotranspiration at unsampled locations. *New Mexico Institute of Mining and Technology*.

[30] M. Soleymani. Gradient descent - machine learning.

[31] J. Watkins. Method of moments.

[32] Jun Yu. Empirical characteristic function estimation and its applications.