A Practical Guide to Pricing and Hedging with Lévy Processes

Ariel Eliasov

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science.

October 14, 2010

Programme in Advanced Mathematics of Finance,
University of the Witwatersrand,
Johannesburg.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

_____

October 14, 2010

Abstract

This dissertation provides an accessible framework for simulating, pricing and hedging options contracts, using Lévy processes that offers several insights and advantages when compared to conventional techniques (particularly when short dated options or options near maturity are considered). A minimal review of Lévy processes is provided for those unacquainted with the subject, presenting fundamental theorems central to the subsequent analysis. A flexible subclass (the Variance Gamma 'family') is then selected and studied, characterizing relevant properties, with which simulation algorithms are then developed. The significance of risk neutral prices in such a market made incomplete by jumps is described and an efficient method for calculating such prices is given. In the final sections, several avenues for mitigating risk by hedging in the underlying asset are explored and a numerical 'hedging race' is conducted and analyzed to compare the performance of several techniques on realistic data.

Acknowledgments

I would like to thank my supervisor Tom McWalter for his guidance and mentorship. Tom's tireless and painstaking editing improved this document immeasurably. I would also like to thank my parents, the National Research Foundation[1] and the Department of Advanced Mathematics of Finance at the University of Witwatersrand, Johannesburg for funding and materials provided for this research.

[1] Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The canonical Black-Scholes option pricing model [3] assumes that the log-returns process of the underlying security may be approximated by a Wiener process. This choice of stochastic process, results in sample paths which are continuous and have log-normal distributions, providing tractable closed-form expressions for many option pricing problems. However, this model fails to describe reality as observed in the markets, leading to significant inaccuracies when pricing and hedging contingent claims.

For instance, empirical studies [6, 33] show that extreme returns are more likely than predicted by the normal distribution. Also, the distribution of returns has been found to have significant skewness — extreme negative returns (crashes) are more likely than large gains. As a consequence, the Black-Scholes model under-prices out of the money options (particularly puts).

To correct for these effects, most traders 'fudge' the volatility parameter in the model, leading to the phenomenon of a volatility smile or skew for different option strikes and maturities [20]. Furthermore, studies on realized volatility indicate that, far from being stationary, asset price volatilities evolve in a random way, exhibiting such phenomena as volatility clustering (autocorrelation in the 'volatility process') and the leverage effect (negative returns cause the volatility to increase).

The drawbacks of the Black-Scholes model, however, are not merely distributional or arise from non-stationarity — in contrast to the continuous paths generated by models based on diffusions, discontinuities, or 'jumps', account for much of the price fluctuation on short timescales. As we shall see throughout the docu-

ment, the fine structure of the price process has important implications for market completeness and the optimum choice of hedging strategy.

This dissertation studies an alternate model for the underlying asset based on Lévy processes. This class of processes is broad enough to encompass a wide range of distributional properties, but more importantly can reproduce the jumps seen on fine timescales, while still retaining much of the tractability of geometric Brownian motion.

In contrast to local volatility models (proposed by Dupire [13]), which are merely able to reproduce the observed volatility surface, using a bizarre profile of unrealistic and state-dependent 'local volatilities' (that predict a flat future surface [11]); a realistic skew or smile *is a natural consequence* of choosing a heavy-tailed Lévy process for the underlying. Moreover, such a choice provides an intuitive explanation of the smile phenomenon, which may be seen as incorporating additional compensation for the risk of unfavorable (unhedgeable) large jumps.

As noted, stochastic volatility effects (which incorporate non-stationarity into the return series) are significant drivers of price dynamics in the long term. However, modeling such effects offers a poor description of fine-scale price behavior, with all diffusion based models (such as the Heston model [21]) requiring excessive values of 'volatility of volatility' to reproduce 'discontinuous-like' movements. Although, there are several models that incorporate stochastic volatility and jumps, in the interests of usability and developing intuition such processes are not considered here.

1.1 Document Structure

Each chapter is used to build up an overarching framework, utilizing Lévy processes to generate pricing and hedging rules and risk statistics that are relevant to market practitioners. Hence, this dissertation is pitched at a level appropriate for market practitioners with strong quantitative backgrounds. Most of the document provides a broad overview, dealing only with subject matter deemed relevant, and for the most part ignores the subtleties of stochastic calculus. Hence, this review is neither rigorous nor comprehensive (such reviews may be found elsewhere). A bird's-eye

view of each chapter is now presented.

The second chapter provides a minimal review of Lévy processes that attempts to extract intuition from the properties and limitations of such processes. Fundamental results, that are central to the rest of the document, are developed in this chapter.

The third chapter focuses on a particular sub-class of exponential Lévy processes based on the variance gamma process introduced by Madan and Seneta [31]. The relevance of this family of processes in the context of modeling short dated options is explored and compared to other techniques. An understanding of the parameter set describing the model is developed and several properties are derived that facilitate the simulation of this class of processes.

Every aspect of the simulation algorithms for such processes are presented in a self contained fashion in the subsequent chapter, of interest are recently developed techniques for simulating the CGMY process, that utilize a representation which may be easily extended to multiple assets. Naturally, such techniques facilitate the use of Monte-Carlo methods for estimating the price of exotic options.

The fifth chapter discusses an efficient fast Fourier transform approach to generating risk neutral prices for vanilla European options and more importantly, the relevance and meaning of such prices.

The sixth chapter examines strategies for hedging in a market made incomplete by jumps, using a portfolio consisting only of holdings in the underlying asset and a riskless bank account. The approaches considered include super-replication, utility maximization and quadratic hedging.

The final chapter attempts to gauge the efficacy of some of the aforementioned hedging approaches by comparing them to each other and benchmark techniques such as the Black-Scholes methodology through a series of numerical 'hedging races'. The results provide several insights into the applicability of Lévy process models and hedging in general.

Chapter 2

Review of Lévy Processes

In this chapter, a basic review of Lévy processes is provided. This will form the basis of most results in the subsequent chapters. Modest prerequisites are needed to understand the mathematical reasoning in this chapter — a rudimentary understanding of set theory, probability theory, random variables and a basic grasp of stochastic processes are advised. It is also assumed that the reader is familiar with the essentials of option pricing. Most complicated or difficult proofs are not reproduced here, instead reference will be made to a reliable source in each case. This chapter draws extensively from the work of Cont and Tankov [7].

The chapter is structured as follows — first, some preliminary definitions are presented. Thereafter, Lévy processes are defined. Next, the basic properties of such processes are discussed. Then, fundamental examples are explored. Finally, various representation theorems for Lévy Processes are presented.

2.1 Preliminaries

The following concepts and definitions are useful in the sequel:

2.1.1 Càdlàg Processes

Definition 2.1 (Right (or Left) Continuity)**.** Let $(X_t)_{t\in[0,T]}$ be a stochastic process. If for some $t_0 \in [0,T]$,

$$X_{t_{0+}} \equiv \lim_{t\downarrow t_0} X_t = X_{t_0} \quad \left(\text{or } X_{t_{0-}} \equiv \lim_{t\uparrow t_0} X_t = X_{t_0}\right), \tag{2.1}$$

X_t is said to be *right continuous* (or *left continuous*) at t_0.

Obviously, continuous processes are both right and left continuous at all times.

Definition 2.2 (Discontinuities of the First and Second Kind)**.** Let $(X_t)_{t\in[0,T]}$ be a stochastic process. A time point t is called a *discontinuity of the first kind* if both X_{t+} and X_{t-} exist but are not equal. Any other discontinuity is said to be of the *second kind*.

The class of processes which are right-continuous with left limits (possess at worst discontinuities of the first kind) is abbreviated *càdlàg* in French. This family is limited to having at most a countable number of discontinuities (jumps with absolute value smaller than ϵ in the limit as $\epsilon \downarrow 0$) and a finite number of large jumps (with magnitude larger than ϵ for all $\epsilon > 0$) on a given time interval [28, pg. 3][15]. The nature of discontinuities has important implications when defining stochastic integrals.

We can define all jumps of a *càdlàg* process by:

Definition 2.3 (Process of Jumps)**.** Let $(X_t)_{t\in[0,T]}$ be a càdlàg stochastic process. The *process of jumps* of X_t is defined as:

$$\Delta X_t \equiv X_t - X_{t-}. \tag{2.2}$$

2.1.2 Finite Variation

Definition 2.4 (Finite Variation)**.** A real valued process X is said to be of finite variation if for every t:

$$S_t(X) \equiv \sup_{\pi\in\Pi} \sum_{i=0}^{n-1} \left|X_{t_{i+1}} - X_{t_i}\right| = \lim_{\delta_n\to 0} \sum_{i=0}^{n-1} \left|X_{t_{i+1}} - X_{t_i}\right| < +\infty, \tag{2.3}$$

where Π is the set of all partitions π of the interval $[0,t]$ with $0 = t_0 < t_1 < ... < t_n = t$ and

$$\delta_n = \max_{0\le i\le n-1}(t_{i+1} - t_i).$$

The function $S_t(X)$, known as the total variation, is positive and non-decreasing in t. All bounded real-valued monotone processes have finite variation.

Higher order variations, of the form:

$$S_t^{(k)}(X) \equiv \lim_{\delta_n \to 0} \sum_{i=0}^{n-1} \left| X_{t_{i+1}} - X_{t_i} \right|^k, \tag{2.4}$$

are also studied. A continuous process of finite variation will have zero higher order variation [28, Theorem 1.10].

2.1.3 Admissible Trading Strategies

A stochastic process is termed *predictable* or *previsible* if its current value may only be determined from information available at preceding instants:

Definition 2.5 (Predictable Processes [7])**.** The *predictable* σ-algebra is the σ-algebra $\mathcal{P}$ generated on $[0,T] \times \Omega$ (where Ω is the set of all scenarios) by all adapted left-continuous processes. A mapping $X : [0,T] \times \Omega \mapsto \mathbb{R}^d$ which is measurable with respect to $\mathcal{P}$ is called a *predictable process*.

All realistic trading strategies should be predictable — an investor can only make trading decisions based on information available immediately prior to the transaction. Permitting càdlàg trading strategies, amounts to allowing infinitesimal clairvoyance.

However, the class of predictable processes is too broad to ensure that a trading strategy may be implemented in practice. Consequently, trading strategies will be constrained to a countable number of rebalancings:

Definition 2.6 (Simple Predictable Processes)**.** Let $(\phi_t)_{t\in[0,T]}$ be a stochastic process. The process ϕ_t is called a *simple predictable process* if it can be represented as:

$$\phi_t = \phi_0 \mathbb{1}_{t=0}(t) + \sum_{i=0}^{n} \phi_i \mathbb{1}_{(T_i, T_{i+1}]}(t),$$

where $T_0 = 0 < T_1 < T_2 < \cdots < T_{n+1} = T$ are a sequence of non-anticipating (adapted) random transaction times and each ϕ_i is bounded and $\mathcal{F}_{T_i}$-measurable.

It is possible to define an integral for such a predictable process in the following manner:

$$\int_0^t \phi_u \, dS_u = \phi_0 S_0 + \sum_{i=0}^{n} \phi_i \left(S_{T_{i+1} \wedge t} - S_{T_i \wedge t} \right), \tag{2.5}$$

where $x \wedge y \equiv \inf(x, y)$. This integral corresponds to the gain resulting from adopting a trading strategy ϕ_t with the asset S_t.

In addition, admissible strategies will be limited to a finite borrowing capacity in the risk-free asset (eliminating doubling strategies and their ilk). In summary:

Definition 2.7 (Admissible Trading Strategy)**.** A trading strategy ϕ, with holdings $(\phi^0, \ldots, \phi^{d-1})$ in the d tradeable assets and value process $V_t(\phi) \equiv \sum_{k=0}^{d-1} \phi_t^k S_t^k$ is called *admissible* if:

- Each holding may be approximated by a simple predictable process given in Definition 2.6.
- The strategy is square-integrable:

$$\int_0^T |\phi_s^k|^2 \, ds < \infty \quad \text{for each } k \in (0, \ldots, d-1).$$

- The strategy is self-financing:

$$V_t(\phi) = V_0 + \sum_{k=0}^{d-1} \int_0^t \phi_t^k \, dS_t^{(k)}.$$

  In other words, changes in the portfolio value V_t result only from fluctuations in the asset values, not from new cash flows.

- The value process is almost surely bounded from below (i.e. $V_t(\phi) > C$ for every $t \in [0, T]$, for some $C > -\infty$).

2.1.4 Compensation

In the context of no-arbitrage pricing, constructing a price process which is a martingale under a risk-neutral measure is central. Moreover, many beneficial theorems (such as the Khinchin and Kolmogorov variance criterion [26, Lemma 4.16]) are only valid for martingales. As a result, the concept of a *compensator* is useful: a compensator is a *predictable* stochastic process, that when subtracted from another stochastic process, will produce a martingale.

For a processes with independent increments, $(X_t)_{t\in[0,T]}$, it is trivial to show that $M_t = X_t - \mathbb{E}[X_t]$ is a martingale (provided $\mathbb{E}[X_t] < \infty$, $\forall t \geq 0$):

$$\mathbb{E}\left[X_t - \mathbb{E}[X_t]\,|\mathcal{F}_s\right] = \mathbb{E}\left[X_t - X_s + X_s - \mathbb{E}[X_t]\,|\mathcal{F}_s\right], \quad t > s.$$

The term $X_t - X_s$ is independent of the filtration, while X_s is measurable, hence using the linearity of the expectation operator:

$$\begin{aligned} &= \mathbb{E}\left[X_t - X_s\right] + X_s - \mathbb{E}[X_t] \\ &= X_s - \mathbb{E}[X_s]. \end{aligned}$$

In this case, $\mathbb{E}[X_t]$ is the compensator of X_t.

2.2 Definition of a Lévy Process

Lévy processes are a family of stochastic processes that have independent and stationary increments, or formally:

Definition 2.8 (Lévy Process)**.** A *càdlàg* stochastic process $(X_t)_{t\in[0,T]}$ on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$[1] is called a *Lévy process* if:

1. $X_0 = 0$ by convention.
2. For every increasing sequence of times $t_0, ..., t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, ..., X_{t_n} - X_{t_{n-1}}$ are independent.
3. For any increment length h, the law of $X_{t+h} - X_t$ does not depend on t.
4. For all $\epsilon > 0$, $\lim_{h\to 0} \mathbb{P}\left(|X_{t+h} - X_t| \geq \epsilon\right) = 0$.

Property number two requires that all non-overlapping increments of the process are independent. The penultimate requirement enforces stationarity — the law of an increment does not depend on the time of observation, only the length of the increment. In reality, the statistical nature of an asset return series can change drastically over time (the behavior in a 'crash' can be markedly different to the

[1] All subsequent definitions, theorems, lemmas and propositions are stated with respect to this probability space.

dynamics during a 'boom' period). Consequently, any option pricing model that makes an assumption of stationarity will generally only be valid for a short time frame (if at all). The final property preserves stochastic continuity by excluding fixed (non-random) jumps. The above definition is valid for multi-dimensional processes, in such a case X_t takes values from $\mathbb{R}^d$, where d is the number of dimensions.

2.3 Properties

2.3.1 Infinite Divisibility

From Definition 2.8 it is clear that we can express any Lévy process as a sum of n independent and identically distributed (i.i.d.) random variables:

$$X_t = \sum_{i=0}^{n-1} X_{(i+1)t/n} - X_{it/n}. \tag{2.6}$$

This means that, the distributions of all Lévy processes are *infinitely divisible*:

Definition 2.9 (Infinite Divisibility)**.** A probability distribution F is *infinitely divisible* if, for all $n \geq 2$, there exists a sequence of i.i.d. random variables $Y_1, ..., Y_n$ such that $Y_1 + ... + Y_n$ has the distribution F.

Only those distributions whose n-th convolution root is still a probability distribution will yield an associated Lévy Process (this excludes the uniform distribution). Note that the distribution of the increments $(Y_1, ..., Y_n)$ need not be of the same type as F (this will occur if F does not belong the class of convolution-closed distributions). For instance, if the law of a Lévy process $(X_t)_{t\in[0,T]}$ is given by a log-normal distribution at time one, at time two the process will not be log-normally distributed.

2.3.2 The Characteristic Function

Definition 2.10 (Characteristic Function)**.** Let X_t be a random variable on $\mathbb{R}^d$ then

$$\Phi_t(z) \equiv \Phi_{X_t}(z) \equiv \mathbb{E}\left[e^{iz.X_t}\right], \quad z \in \mathbb{R}^d \tag{2.7}$$

is known as the *characteristic function* of X_t.

If a random variable admits a density function, then the characteristic function is its dual, in the sense that they are related by the Fourier transform. Using Definition 2.8, it can be shown that Lévy processes possess the following multiplicative property of their characteristic functions:

$$\begin{aligned} X_{t+s} =& X_s + (X_{t+s} - X_s) \\ \Rightarrow \Phi_{t+s}(z) \equiv \Phi_{X_{t+s}}(z) =& \Phi_{X_s}(z)\Phi_{X_{t+s}-X_s}(z) \\ =& \Phi_{X_s}(z)\Phi_{X_t}(z) = \Phi_s(z)\Phi_t(z), \end{aligned}$$

since the addition of two independent random variables is equivalent to multiplying their characteristic functions. Furthermore, from property 4 of Definition 2.8, it can be shown that $\Phi_s(z) \to \Phi_t(z)$ as $s \to t$. As a result, $\Phi_t(z)$ must be a continuous function in time. This implies that the characteristic function of a Lévy process must have the following exponential form:

Theorem 2.11 (Characteristic Function of a Lévy Process)**.** *Let $(X_t)_{t\in[0,T]}$ be a Lévy process on $\mathbb{R}^d$, then there exists a continuous function $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ called the* characteristic exponent *of X, such that:*

$$\mathbb{E}\left[e^{iz.X_t}\right] = e^{t\psi(z)}, \quad z \in \mathbb{R}^d.$$

Hence, the law of X_t may be determined from knowledge of X at any specific time (e.g. X_1). This limits one's ability to specify the distribution of a Lévy process across time scales.

2.3.3 Behavior Across Timescales

As an illustration of the behavior of a Lévy process across timescales, observe that the cumulant generating function of a Lévy process is given by $\Psi(z) \equiv \log(\Phi(z)) = t\psi(z)$. The n-th cumulant of a process is defined by:

Definition 2.12 (Cumulants)**.** Let X be a random variable. The *n-th cumulant*, C_n, of X is given by:

$$C_n(X) = \frac{1}{i^n}\frac{\partial^n \Psi_X}{\partial z^n}(0), \tag{2.8}$$

where Ψ_X is the cumulant generating function of X.

For Lévy processes, this form reduces to:

$$C_n(X_t) = \frac{t}{i^n}\psi^{(n)}(0). \tag{2.9}$$

The variance, skewness and excess kurtosis (if they exist) of an increment $X_{t+h} - X_t \stackrel{d}{=} X_h$ may then be computed in terms of their values at time one:

$$\mathrm{VAR}[X_h] \equiv C_2(X_h) = \frac{h}{i^2}\psi''(0) = h\mathrm{VAR}[X_1] \tag{2.10}$$

$$S(X_h) \equiv \frac{C_3(X_h)}{C_2(X_h)^{3/2}} = \frac{ih\psi^{(3)}(0)}{[-h\psi''(0)]^{3/2}} = \frac{S(X_1)}{\sqrt{h}} \tag{2.11}$$

$$\kappa(X_h) \equiv \frac{C_4(X_h)}{C_2(X_h)^2} = \frac{h\psi^{(4)}(0)}{[h\psi''(0)]^2} = \frac{\kappa(X_1)}{h}. \tag{2.12}$$

Clearly, both the excess kurtosis and skewness (if there is any) decay in a fixed way to normality as the interval length is increased, limiting the behavior of Lévy processes across timescales. Practically, this means that an asset price model based on an ordinary Lévy process cannot be calibrated to a realistic term structure of volatilities.

Furthermore, since the variance increases linearly with the the interval size, the autocorrelation of most Lévy processes behaves as follows:

Lemma 2.13 (Autocorrelation of a Lévy Process)**.** *Let X_t be a Lévy process, with a finite second moment. The autocorrelation of the process, defined as:*

$$\rho_{(X_s,X_t)} \equiv \frac{\mathbb{E}\left[(X_t - \mathbb{E}[X_t])\,(X_s - \mathbb{E}[X_s])\right]}{\sigma_{X_t}\sigma_{X_s}}, \tag{2.13}$$

for $t > s \geq 0$, where $\sigma_X \equiv \sqrt{\mathbb{E}[X^2] - \mathbb{E}[X]^2}$ is the standard deviation of X, is equal to $\sqrt{\frac{s}{t}}$.

Proof. Manipulating the definition of autocorrelation:

$$\begin{aligned}\rho_{(X_s,X_t)} &= \frac{\mathbb{E}[X_tX_s] - \mathbb{E}[X_t]\mathbb{E}[X_s]}{\sigma_{X_t}\sigma_{X_s}} \\ &= \frac{\mathbb{E}[(X_t - X_s + X_s)X_s] - \mathbb{E}[X_t]\mathbb{E}[X_s]}{\sigma_{X_t}\sigma_{X_s}}\end{aligned}$$

Using the linearity of the expectation operator and the independence of increments:

$$\begin{aligned}\rho_{(X_s,X_t)} &= \frac{\mathbb{E}[X_t - X_s]\mathbb{E}[X_s] + \mathbb{E}[X_s^2] - \mathbb{E}[X_t]\mathbb{E}[X_s]}{\sigma_{X_t}\sigma_{X_s}} \\ &= \frac{\mathbb{E}[X_s^2] - \mathbb{E}[X_s]^2}{\sigma_{X_t}\sigma_{X_s}}\end{aligned}$$

From the definition of standard deviation and (2.10):

$$\rho_{(X_s,X_t)} = \frac{\sigma_{X_s}}{\sigma_{X_t}} = \sqrt{\frac{s}{t}}.$$

■

Remarkably, such behavior is the same for all Lévy processes (provided the moments in question exist) and is independent of the parameters governing the process.

2.4 Examples

As we shall see in Section 2.5.1, both Brownian motion and the Poisson process can be seen as the building blocks of all Lévy processes. Owing to their fundamental nature, a brief discussion of their properties is now presented.

2.4.1 Brownian Motion

Brownian Motion (which we denote B_t) belongs to the class of the Lévy Processes. In addition to the properties outlined in Definition 2.8 it has the following characteristics that distinguish it from the general class:

- The sample paths of Brownian motion are continuous (and nowhere differentiable).
- $B_t - B_s$ is normally distributed with zero mean and variance $t-s$ (i.e. $B_t - B_s \sim \mathcal{N}(0, t-s)$).
- Brownian motion displays the following remarkable statistical self-similarity:

$$(B_{at})_{t\geq 0} \overset{d}{=} \left(\sqrt{a}B_t\right)_{t\geq 0} \quad \text{for all} \quad a > 0.$$

  In other words, a Brownian motion is an example of a fractal curve — the statistical properties of the process are invariant across timescales, provided that it is scaled appropriately.

Brownian motion may also be thought of as the limit of a random walk (in a similar vein to the central limit theorem). Consider, the following modified simple random walk:

$$X_t^{(n)} = \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nt \rfloor} Z_i,$$

where the sequence of random variables $Z_1, Z_2, ...$ is i.i.d. with $\mathbb{P}(Z = 1) = 0.5$ and $\mathbb{P}(Z = -1) = 0.5$, $\lfloor n \rfloor$ denotes the `floor` of n. In other words, this simple random walk either jumps up or down by $\frac{1}{\sqrt{n}}$ every $\frac{1}{n}$ units of time. As $n \to \infty$, $X_t^{(n)}$ behaves like a Brownian motion. This is true of many random walks:

Proposition 2.14 (Donsker's Theorem)**.** *Let $(S_n)_{n\geq 0}$ be a random walk with $\mathbb{E}[S_1^2] < \infty$. Then, as $n \to \infty$,*

$$\frac{S_{\lfloor nt \rfloor} - \mathbb{E}[S_{\lfloor nt \rfloor}]}{\sqrt{\text{VAR}[S_{\lfloor nt \rfloor}]}} \to B_t \quad \textit{locally uniformly in distribution.}$$

Proof. See Dudley [12, Chapter 1] ■

2.4.2 The Poisson Process

An exponential distribution has the following probability density:

$$f(x; \lambda) = \lambda e^{-\lambda x} \mathbb{1}_{x \geq 0} \tag{2.14}$$

The exponential distribution has a unique property called the 'absence of memory':

Proposition 2.15 (The Absence of Memory)**.** *If $T \geq 0$ is a non-zero random variable, such that for any $t, s > 0$:*

$$\mathbb{P}\left(T > t + s | T > t\right) = \mathbb{P}\left(T > s\right), \tag{2.15}$$

then T is exponentially distributed.

Proof. See Cont and Tankov [7, pg. 45-46]. ■

Therefore, it is understandable that exponential variables will play an important part in generating Markovian processes with jumps.

A Poisson process is a *counting* process that counts the arrivals of independent exponentially distributed random variables. It can only have non-negative integer values at any time. A basic Poisson process can be expressed as:

Definition 2.16 (Poisson Process)**.** Let $(\tau_i)_{i\geq 1}$ be a sequence of independent exponential random variables with rate parameter λ and $S_n = \sum_{i=1}^{n} \tau_i$. The process $(N_t)_{t\in[0,T]}$ defined by:

$$N_t = \sum_{n\geq 1} \mathbb{1}_{t\geq S_n} \tag{2.16}$$

is called a *Poisson process* with *intensity* λ.

A new Poisson processes may be created by thinning (rejecting) the jumps of a Poisson process, with the following effect on the intensity of the new process:

Proposition 2.17 (Thinning Property)**.** *Let $(N_t)_{t\in[0,T]}$ be a simple Poisson process with intensity λ. If the jumps of N_t are rejected with probability $1-p$, then the resulting process M_t is a simple Poisson process with intensity $p\lambda$.*

Proof. See Gut [18, Chapter 8, Section 6] ■

A simple extension of the basic Poisson process is to replace the jump sizes (which were constantly one) with random variables. This gives rise to a compound Poisson process:

Definition 2.18 (Compound Poisson Process)**.** Let $(N_t)_{t\in[0,T]}$ be a Poisson process with intensity λ and $(Y_i)_{i\geq 1}$ be an independent sequence of jump sizes with distribution f, then

$$X_t = \sum_{i=1}^{N_t} Y_i \tag{2.17}$$

is a *compound Poisson process.*

It can be shown that the compound Poisson process is the only piecewise constant process with independent and stationary increments [7, Proposition 3.3]. Therefore, every piecewise constant Lévy Process is a compound Poisson process.

2.5 Representation

2.5.1 The Lévy-Itô Decomposition

In order to study the jump behavior of a process, the following random measure is introduced:

Definition 2.19 (Jump Measure)**.** For any measurable set $B \subset [0,T) \times \mathbb{R}^d \backslash \{0\}$

$$J_X(B) = \#\{(t, \Delta X_t) \in B\} \equiv \sum \mathbb{1}_{(t,\Delta X_t)\in B}$$

is the *jump measure* of a Lévy process $(X_t)_{t\in[0,T]}$ on $\mathbb{R}^d$.

Hence, $J_X([t_1,t_2] \times A)$ counts the number of jumps with sizes in A that occur in a process X between t_1 and t_2.

On the other hand, $J_X([t_1,t_2] \times A)$ can be thought of as a standard Poisson process at time $t_2 - t_1$, where jumps of the original compound Poisson process, X, are accepted (counted) with probability $f(A)$ (the probability of a jump size occurring from the set A). From the 'thinning' property 2.17, it is evident that $J_X(dx \times dt)$ is a Poisson distributed random variable with intensity $\lambda f(dx)dt$, where f is the jump size distribution and λ is the expected number of jumps per unit time (intensity) of the original process.

Alternatively, this measure may be expressed in terms of the jumps of the Lévy process and the times at which they occur, using the Dirac delta function:

$$J_X(\omega, .) = \sum_{n\geq 1} \delta_{(T_n(\omega),Y_n(\omega))} = \sum_{t\in[0,T],\Delta X_t(\omega)\neq 0} \delta_{(t,\Delta X_t(\omega))}, \quad \text{for any } \omega \in \Omega, \quad (2.18)$$

where the series of jumps Y_n, known as 'marks', take values from $\mathbb{R}^d\backslash\{0\}$. Each jump is revealed at a series of non-anticipating random times T_n on $[0,T]$, with $T_n \to T$ almost surely as $n \to \infty$ (i.e. there are a countable number of jumps on $[0,T]$ and Y_n is $\mathcal{F}_{T_n}$ measurable for all n). The sequence (T_n, Y_n) is known as the marked point process of X_t.

All information about the jumps of the process can be recovered by integrating various functions of the jump measure. For instance, a compound Poisson process X_t, may be expressed as the sum of its jumps:

$$X_t = \sum_{s\in[0,t]} \Delta X_s = \int_{[0,t]\times\mathbb{R}} x J_X(ds \times dx). \quad (2.19)$$

From Section 2.1.1, the process can have at most a finite number of jumps on $[0,t]$. Consequently, there are no convergence problems with the integral representation. However, the jump measure contains no information about the continuous part of X_t. Consequently, for continuous processes $J_X = 0$.

The jump measure defined previously can be normalized to a single unit of time under expectation, providing a unique and convenient mechanism for describing the jump behavior of every Lévy process:

Definition 2.20 (Lévy Measure)**.** Let $(X_t)_{t\in[0,T]}$ be a Lévy Process on $\mathbb{R}^d$. The *Lévy measure* ν of X is defined as:

$$\nu(A) = \mathbb{E}\left[\#\left\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in A\right\}\right]. \tag{2.20}$$

This measure represents the expected number of jumps with sizes in A per unit time. For any compact set A, such that $0 \notin A$, $\nu(A)$ is almost surely finite — in other words, X_t can have at most a finite number of 'large' jumps (see Section 2.1.1). However, the integral of ν may not necessarily be finite, particularly in the region of zero — X may have an infinite number of small jumps on a finite interval, whose sum may not converge.

With this measure in hand, Lévy [29] and Ito [24] were able to show that every Lévy Process may be decomposed as follows:

Theorem 2.21 (The Lévy-Itô Decomposition)**.** *Let $(X_t)_{t\in[0,T]}$ be a Lévy process on $\mathbb{R}^d$. The Lévy measure ν of X satisfies the following conditions:*

$$\int_{|x|\leq 1} |x|^2 \, \nu(dx) < \infty \text{ and } \int_{|x|\geq 1} \nu(dx) < \infty.$$

The process X_t may be expressed as:

$$X_t = \gamma t + B_t + X_t^l + \lim_{\epsilon \downarrow 0} \tilde{X}_t^\epsilon, \tag{2.21}$$

where

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

and

$$\begin{aligned} \tilde{X}_t^\epsilon &= \int_{\epsilon\leq|x|\leq 1, s\in[0,t]} x \left\{J_X(ds \times dx) - \nu(dx)ds\right\} \\ &\equiv \int_{\epsilon\leq|x|\leq 1, s\in[0,t]} x \left\{\tilde{J}_X(ds \times dx)\right\}, \end{aligned}$$

where B_t is a Brownian motion on $\mathbb{R}^d$ with covariance matrix A and $\gamma \in \mathbb{R}^d$ is a drift vector. The terms in (2.21) *are independent.*

Proof. See Cont and Tankov [7, pg. 81-82]. ■

The significance of each term in (2.21) will now be explained. The term $\gamma t + B_t$ is the continuous Gaussian part of the Lévy process. The jumps of the process with absolute value larger than one are described by the compound Poisson process X_t^l. The X_t^l term will be almost surely finite, since $\int_{|x|\geq 1} \nu(dx) < \infty$. The small jumps are accounted for by $\tilde{X}_t^\epsilon$, which may be viewed as a possibly infinite superposition of centered compound Poisson processes[2]. Since ν may have a singularity at zero, we cannot take ϵ directly to zero. Furthermore, in order to prove convergence using the Khinchin and Kolmogorov variance criterion [26, Lemma 4.16], the jump integral term is *compensated*.

2.5.2 The Lévy-Khinchin Representation

Using the Lévy-Itô decomposition, it is relatively simple to derive the following representation of the characteristic function of a Lévy process:

Theorem 2.22 (The Lévy-Khinchin Representation)**.** *Let $(X_t)_{t\in[0,T]}$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (A, ν, γ). Then*

$$\mathbb{E}[e^{iz.X_t}] = e^{t\psi(z)},\ z \in \mathbb{R}^d, \tag{2.22}$$

with

$$\psi(z) = -\frac{1}{2} z.Az + i\gamma.z + \int_{\mathbb{R}^d} (e^{iz.x} - 1 - iz.x\mathbb{1}_{|x|<1})\,\nu(dx).$$

Proof. From the Lévy-Itô decomposition:

$$X_t = \gamma t + B_t + X_t^l + \lim_{\epsilon \downarrow 0} \tilde{X}_t^\epsilon.$$

Since the terms are independent we can write the characteristic function as follows:

$$\mathbb{E}[e^{iz.X_t}] \equiv \Phi_{X_t}(z) = \Phi_{\gamma t + B_t}(z)\Phi_{X_t^l}(z)\Phi_{\tilde{X}_t^\epsilon}(z).$$

[2] Compound Poisson processes have a finite number of jumps, hence infinitely many of them may be required to match an infinite activity Lévy process.

It can be shown [7, pg. 74] that the characteristic function of a compound Poisson process is given by:

$$\mathbb{E}\left[\exp(iu.X_t)\right] \equiv \Phi_{X_t}(u) = \exp\left\{t\int_{\mathbb{R}^d}\left(e^{iu.x}-1\right)\,\nu(dx)\right\} \tag{2.23}$$

and the characteristic function of a Gaussian process is:

$$\mathbb{E}\left[\exp(iz.B_t)\right] \equiv \Phi_{B_t}(z) = \exp\left\{t\left(i\mu.z-\frac{1}{2}z.\sigma z\right)\right\}, \tag{2.24}$$

where μ is the drift vector and σ is the covariance matrix. Consequently,

$$\Phi_{\gamma t+B_t}(z) = \exp\left\{t\left(i\gamma.z-\frac{1}{2}z.Az\right)\right\},$$
$$\Phi_{X_t^l}(z) = \exp\left\{t\int_{|x|\geq 1}\left(e^{iz.x}-1\right)\,\nu(dx)\right\} \text{ and}$$
$$\Phi_{\tilde{X}_t^\epsilon}(z) = \exp\left\{t\int_{\epsilon<|x|<1}\left(e^{iz.x}-1-iz.x\right)\,\nu(dx)\right\},$$

which leads to the final form of the Lévy-Khinchin formula. ■

Accordingly, the characteristic triplet (A,ν,γ) may be used to *uniquely* specify a Lévy process.

Dividing the process into two components, one with jumps smaller than one and another with jumps bigger than one, is arbitrary. In general, the process may be truncated by any bounded measurable function $g:\mathbb{R}^d\to\mathbb{R}$ satisfying $\int|e^{izx}-1-izxg(x)|\,\nu(dx)<\infty$ [36], replacing the function $g(x)=\mathbb{1}_{|x|\leq 1}$ used in this instance. The choice of truncation function does not affect the Lévy measure or covariance matrix, only the 'drift' γ is modified [7, pg. 83]. Notationally, γ will denote the drift associated with the standard truncation function $g(x)=\mathbb{1}_{|x|\leq 1}$, while γ_g will denote the drift associated with a truncation function g.

Often, we may wish to specify the jump structure of a Lévy process, using the Lévy measure. As a result, the Lévy-Khinchin representation is indispensable for computing the characteristic function, which in general will have a closed form solution. On the other hand, parsimonious Lévy densities may produce not only intractable but sometimes not even expressible probability densities.

2.5.3 The Predictable Representation Property

In the subsequent discussion, the following power jump processes are useful:

Definition 2.23 (Power Jump Processes)**.** Let $(Z_t)_{t\in[0,T]}$ be a Lèvy process, define the *ith-power jump process* as:

$$Z_s^{(i)} := \sum_{0<t\leq s} (\Delta Z_t)^i, \quad i \geq 2. \tag{2.25}$$

For convenience, define $Z_t^{(1)} \equiv Z_t$. Power jump processes have jumps in the same places as the original Lèvy process. However, the jump sizes of these new processes are the ith-power of the jump sizes of the original Lèvy process.

It has been shown by Nualart and Schoutens [35], that under some weak moment assumptions, any square integrable random variable F has the following version of the predictable representation property:

Theorem 2.24 (Predictable Representation Property (PRP))**.** *Let F be a square integrable, $\mathcal{F}_T$–measurable random variable, where $(\mathcal{F}_t)_{t\in[0,T]}$ is the* natural filtration *generated by the process Z_t. Provided the Lévy measure ν of Z_t satisfies:*

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

for some $\epsilon > 0$ and $\lambda > 0$. Then F may be expressed as follows:

$$F = \mathbb{E}[F] + \sum_{i=1}^{\infty} \int_0^T a_s^{(i)}\, d\left(Z_s^{(i)} - \mathbb{E}\left[Z_s^{(i)}\right]\right), \tag{2.26}$$

where $a_t^{(i)}$ is predictable *and $Z^{(i)}$ is the* power jump process *of order i.*

Proof. See Nualart and Schoutens [35, pg. 118] ■

The integrability requirement in the region of zero, ensures that the Lévy density is analytic in the region of zero and that moments of all orders exist.

In the case of a Wiener process, the sample paths are continuous, hence $B_s^{(i)} = 0$ for all $i \geq 2$ [40]. Consequently, a Wiener process leads to the following well-known PRP:

$$F = \mathbb{E}[F] + \int_0^T a_s\, dB_s^{(1)}. \tag{2.27}$$

Since a Poisson process only has jumps of size one, $N_s^{(i)} = N_s^{(1)}$ for all $i \geq 2$. Consequently, when the filtration is given by the natural filtration generated by a Poisson process, the predictable representation may be reduced to:

$$F = \mathbb{E}[F] + \int_0^T a_s \, d\left(N_s^{(1)} - \lambda s\right). \tag{2.28}$$

Only in such cases, where the infinite sum in (2.26) may be simplified, will the market be *complete* [9][19]. However, in most realistic stock price models, such a neat simplification cannot be made. As a result, in general, perfect replication of options will not be possible.

2.5.4 Stochastic Time Changes

Any semi-martingale may be written as a Brownian motion (possibly defined on some adequately extended probability space) evaluated at a random time change[3] [17]. Such a time change may be defined in general as follows: if a random process $(X_t)_{t\in[0,T]}$ is viewed on a new random time scale given by $(Z_t)_{t\in[0,T]}$, then the resulting process may be written as $Y_t = X_{Z_t}$, which is defined for each $\omega \in \Omega$ by $Y(t,\omega) = X(Z(t,\omega),\omega)$. All non-decreasing processes constitute valid time changes. However, a *subordinator* specifically refers to a non-decreasing Lévy process (such as the Poisson process).

When a Lévy process is subordinated, the following effect on the distribution of the newly created process is observed:

Theorem 2.25 (Subordination)**.** *Let $(X_t)_{t\in[0,T]}$ be a Lévy process with characteristic exponent $\psi(u)$ and $(Z_t)_{t\in[0,T]}$ be a subordinator with Laplace exponent[4] $l(u)$ then the process $Y_t = X_{Z_t}$, has the following characteristic function:*

$$\mathbb{E}\left[e^{iz.Y_t}\right] = e^{tl(\psi(z))}$$

[3] The time change process need not be independent from the Brownian motion or a Lévy process.

[4] A Laplace exponent is defined similarly to the characteristic exponent as follows:

$$\mathbb{E}\left[e^{uS_t}\right] = e^{tl(u)}.$$

Proof. See Cont and Tankov [7, pg. 108] ■

A Lévy process meeting the following conditions may be expressed as Brownian motion subordinated by an independent Lévy process:

Theorem 2.26 (Brownian Motion Subordination Requirements)**.** *Let $(X_t)_{t\in[0,T]}$ be a Lévy process with Lévy measure ν. The process X_t may be expressed as $X_t = B(Z_t) + \mu Z_t$, where $(Z_t)_{t\in[0,T]}$ is some subordinator and $(B_t)_{t\in[0,T]}$ is a Brownian motion independent of Z and $\mu \in \mathbb{R}$, if and only if the following conditions are met:*

1. *ν is absolutely continuous with density $\nu(x)$.*
2. *$\nu(x)e^{-\mu x} = \nu(-x)e^{\mu x}$ for all x.*
3. *$\nu(\sqrt{u})e^{-\mu\sqrt{u}}$ is a completely monotonic[5] function on $(0,\infty)$*

Proof. See Cont and Tankov [7, pg. 114]. ■

In such an instance, the Lévy measure of the resulting processes may be determined from:

Proposition 2.27 (Brownian Subordination)**.** *Let $(Z_t)_{t\in[0,T]}$ be a non-decreasing Lévy process with Lévy measure $\nu(dx)$ and $(B_t)_{t\in[0,T]}$ be an independent standard Brownian motion. The process $(X_t)_{t\in[0,T]}$ defined by:*

$$X_t = \theta Z_t + B(Z_t) \tag{2.29}$$

will have the following Lévy measure:

$$\mu(dx) = dx \int_0^\infty \frac{\exp\left(-\frac{(x-\theta y)^2}{2y}\right)}{\sqrt{2\pi y}}\,\nu(dy).$$

Proof. See Sato [39, Theorem 30.1] or [32]. ■

This representation of Lévy processes provides a convenient mechanism to introduce a dependence structure between multiple copies of the same process. This may be accomplished by incorporating dependence into the normally distributed increments of the Brownian motions and co-ordinating the jumps through a common

[5] A function is completely monotonic if all its derivatives exist and $(-1)^k \frac{d^k f(u)}{du^k} > 0$ for all $k \geq 1$.

subordinator [7, Section 5.1]. Moreover, it permits the expression of a complicated Lévy process (such as the variance gamma process) as the subordination of a simpler one (such as a Brownian motion subordinated by a gamma process).

Interestingly, such representations indicate that returns may in fact be normally distributed, not in calendar time, but in some other 'economically relevant' time-scale [17]. Indeed, Ané and Geman [1] show that high frequency returns of liquid technology stocks deviate significantly from normality when viewed in calender time. However, when a stochastic time change is employed normality may be recovered.

Chapter 3

The Variance Gamma Family of Processes

Since the seminal work of Black and Scholes [3], models based on geometric Brownian motion (GBM) have dominated the field of no-arbitrage pricing, owing to their tractability. Over the years, the original model has been augmented as needed. For instance, the class of jump diffusions, first studied by Merton [34], introduces an additional compound Poisson component to the process. In this arrangement, the diffusion captures the randomness of the returns process during the normal course of events, while uncommon extreme events or 'shocks' (such as a worse-than-expected earnings announcement) are accounted for by the unrelated (independent) jumps of the Poisson process. Other models vary the parameters (typically the variance) of the geometric Brownian motion over time in order to capture observed effects, such as: the smile phenomenon, autocorrelation of the volatility series (clustering) and the leverage effect (whereby volatility is negatively correlated with asset returns, implying that a crash leads to highly unpredictable trading behavior)[6].

However, as with all models, diffusions are at best an approximation of reality — practical constraints preclude the possibility of 'continuous trading', indicating that the observed price process cannot be continuous or of infinite variation (an inherent property of all Brownian motion based models, requiring an infinite number of changes on a finite time interval). Indeed, in reality all price changes arrive as jumps or 'shocks' (as opposed to infinitesimal adjustments), it can be inferred from this that pure jump models provide a more natural description of asset returns

than diffusions. For instance, in pricing problems where the fine structure of asset returns is of interest, such as short-dated options, diffusion based models perform comparatively poorly and require extreme parameter choices, since the principal mechanism of price change on such scales is continuous [7, Chapter 1].

As an alternative, pure-jump processes based on the variance gamma (VG) process have been used extensively in modeling asset returns since the early 1990s [4, 30, 31]. This class of processes and its extensions has been found to be flexible enough to capture a wide range of realistic distributional and stochastic properties, while remaining relatively tractable. Thus, this class of model will form the basis of the numerical investigation of Section 7.

Concordantly, this chapter is devoted to describing and deriving the properties of such processes. The chapter follows the chronological sequence in which the processes where introduced: the first section introduces the variance gamma process and the gamma process that underpins it. Next, the CGMY process[1], an extension of the VG process, and its intricacies are elaborated upon. The techniques for simulating such processes are differed to Chapter 4.

3.1 The Variance Gamma Process

3.1.1 Properties

The variance gamma process is a purely discontinuous, *infinite activity*[2] process of *finite variation* with real–valued parameters $(\sigma, \theta, \nu > 0)$. Its characteristic function is given by:

$$\Phi_{X_t}(z) = \left(1 - iz\theta\nu + z^2\sigma^2\nu/2\right)^{-t/\nu}. \tag{3.1}$$

Using the characteristic function, the following key properties of the VG process may be derived:

[1] The process is named after the authors of the paper in which it was introduced — **C**arr, **G**eman, **M**adan and **Y**or.

[2] The Lévy measure of the VG process does not integrate to a finite value in the region of zero. Consequently, the VG process will have infinitely many small jumps.

Name	Symbol	Value
Characteristic Function	$\Phi_{X_t}(z)$	$\left(1 - iz\theta\nu + z^2\sigma^2\nu/2\right)^{-t/\nu}$
Lévy Density	$\nu(x)$	$\frac{1}{\nu\lvert x\rvert}\exp\left\{\frac{\theta}{\sigma^2}x - \frac{1}{\sigma}\sqrt{\frac{2}{\nu}+\frac{\theta^2}{\sigma^2}}\lvert x\rvert\right\}$
Probability Density	$p_{X_t}(x)$	$C\lvert x\rvert^{\frac{t}{\nu}-\frac{1}{2}}\exp\left(\frac{\theta}{\sigma^2}x\right)K_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma}\sqrt{\frac{2}{\nu}+\frac{\theta^2}{\sigma^2}}\lvert x\rvert\right)$ with $C = \sqrt{\frac{\nu\sigma^2}{2\pi}}\frac{\left(\theta^2\nu^2+2\sigma^2\nu\right)^{\frac{1}{4}-\frac{t}{2\nu}}}{\Gamma(t/\nu)}$ [7, Table 4.5 — Erratum]
Expectation	$\mathbb{E}[X_t]$	θt
Variance	$\mathrm{VAR}[X_t]$	$\sigma^2 t + \nu\theta^2 t$
Skewness	$S[X_t]$	$\frac{\theta\nu}{\sqrt{t}}\left(\frac{3\sigma^2+2\nu\theta^2}{(\sigma^2+\nu\theta^2)^{3/2}}\right)$
Kurtosis	$\kappa[X_t]$	$\frac{3\nu}{t}\left(2 - \frac{\sigma^4}{(\sigma^2+\nu\theta^2)^2}\right)$

Tab. 3.1: A summary of the properties of the variance gamma process. Skewness and kurtosis are defined in the same way as (2.11) and (2.12), while K denotes the modified Bessel function of the second kind.

The flexibility of three parameters permits one to produce a VG process with almost any variance, skewness and kurtosis. In the special case when $\nu = 0$, the variance gamma process reduces to a Wiener process[3]. Otherwise, the sign of θ determines the sign of the skewness. When $\theta = 0$, the VG process is symmetric. Similar to observed returns, the VG process is always leptokurtic, with kurtosis that decays at a rate of $1/t$ to normality (as with all Lévy processes see Section 2.3.2). The significance of each of the three parameters is transparent when an alternate parameterization of the Lévy density (3.8) is considered, further explanation will be provided there.

As we shall see, the VG process is formulated in terms of the gamma process. Accordingly, a brief discussion of the gamma process is now introduced.

[3] This property is hinted at by the moments in Table 3.1. Moreover, as we shall see in Section 3.1.4, for $\nu = 0$, the subordinator of the VG process has no variance and a drift equal to t (identical to the predictable process t). Hence, subordination (given by (2.29)) results in a Wiener process with mean θt and variance $\sigma^2 t$.

3.1.2 The Gamma Process

Definition 3.1 (The Gamma Process)**.** *The gamma process*, denoted $\Gamma_t(\mu, \nu)$ with mean $\mu \in \mathbb{R}^+$ and variance $\nu \in \mathbb{R}^+$ per unit time, is the Lévy process with gamma distributed increments — each disjoint interval is independent with probability density:

$$f_{\Gamma_h}(x; \gamma, \lambda) = \frac{(\lambda x)^{\gamma h} e^{-\lambda x}}{x\Gamma(\gamma h)} \mathbb{1}_{x>0}, \tag{3.2}$$

with $\gamma = \mu^2/\nu$ and $\lambda = \mu/\nu$ [40, Chapter 9], where h is the interval length. The characteristic function is given by:

$$\Phi_{\Gamma_h}(z; \gamma, \lambda) = \left(1 - \frac{iz}{\lambda}\right)^{-\gamma h}. \tag{3.3}$$

Properties

Since the density of each increment is concentrated on the positive half-axis, the process is monotonically increasing and possesses finite variation (all monotonic processes are of finite variation [28, pg. 5]). Scaling the gamma process has the following effect:

Lemma 3.2 (Scaling Property of the Gamma Process)**.** *Let $\Gamma_t(\gamma, \lambda)$ be a gamma process with parameters $\gamma \in \mathbb{R}^+$ and $\lambda \in \mathbb{R}^+$, then $c\Gamma_t(\gamma, \lambda)$ is equal in distribution to $\Gamma_t(\gamma, \lambda/c)$ for all $t \in \mathbb{R}^+$ and $c \in \mathbb{R}/\{0\}$.*

Proof. From the definition of the characteristic function (2.7):

$$\begin{aligned}
\Phi_{c\Gamma_t(\gamma,\lambda)}(z) &\equiv \mathbb{E}\left[\exp\left(izc\Gamma_t(\gamma, \lambda)\right)\right] \\
&= \left(1 - \frac{icz}{\lambda}\right)^{-\gamma t} = \Phi_{\Gamma_t(\gamma,\lambda/c)}(z) \\
\Rightarrow c\Gamma_t(\gamma, \lambda) &\overset{d}{=} \Gamma_t(\gamma, \lambda/c)
\end{aligned}$$

because equivalence of characteristic functions implies equivalence in distribution. ∎

Recall the expression of the above parameter set in terms of the mean μ and variance ν per unit time (given in Definition 3.1). From the definition, the mean

per unit time of the scaled process is $\mu = \frac{c\gamma}{\lambda}$, while the variance is $\nu = \frac{c^2\gamma}{\lambda^2}$. Scaling the process increases the drift linearly and the variance in a quadratic fashion.

Adding independent gamma processes has the following effect on the parameter set of the resulting process:

Lemma 3.3 (Addition of Gamma Processes)**.** *Let* $\Gamma_t^{(1)}(\gamma_1, \lambda)$ *and* $\Gamma_t^{(2)}(\gamma_2, \lambda)$ *be independent gamma processes. Then* $\Gamma_t^{(1)}(\gamma_1, \lambda) + \Gamma_t^{(2)}(\gamma_2, \lambda)$ *is equal in distribution to* $\Gamma_t(\gamma_1 + \gamma_2, \lambda)$ *for all* $t \in \mathbb{R}^+$.

Proof. Adding independent random variables is equivalent to multiplying their characteristic functions, hence:

$$\begin{aligned}
\Phi_{\Gamma_t^{(1)}+\Gamma_t^{(2)}}(z) &= \Phi_{\Gamma_t^{(1)}}(z)\Phi_{\Gamma_t^{(2)}}(z) \\
&= \left(1 - \frac{icz}{\lambda}\right)^{-\gamma_1 t}\left(1 - \frac{icz}{\lambda}\right)^{-\gamma_2 t} \\
&= \Phi_{\Gamma_t(\gamma_1+\gamma_2,\lambda)}(z) \\
\Rightarrow \Gamma_t^{(1)}(\gamma_1, \lambda) + \Gamma_t^{(2)}(\gamma_2, \lambda) &\stackrel{d}{=} \Gamma_t(\gamma_1 + \gamma_2, \lambda),
\end{aligned}$$

since equivalence of characteristic functions implies equivalence in distribution. ■

Again, considering the mean and variance, it becomes clear that adding gamma processes increases the mean and variance relative to λ and λ^2 respectively.

To determine the form of the Lévy density of the gamma process, write the characteristic function as follows:

$$\Phi_{\Gamma_t}(z) = \exp\left[-\gamma t \log\left(1 - \frac{iz}{\lambda}\right)\right].$$

Using the expansion $\log(1 - \frac{iz}{\lambda}) = -\sum_{n=1}^{\infty} \frac{(iz/\lambda)^n}{n}$:

$$\begin{aligned}
\Phi_{\Gamma_t}(z) &= \exp\left[\gamma t \sum_{n=1}^{\infty} \frac{(iz/\lambda)^n}{n}\right] = \exp\left[\gamma t \sum_{n=1}^{\infty} \frac{(iz/\lambda)^n}{n!}\Gamma(n)\right] \\
&= \exp\left[\gamma t \sum_{n=1}^{\infty} \frac{(iz/\lambda)^n}{n!} \int_{\mathbb{R}^+} y^{n-1}e^{-y}dy\right].
\end{aligned}$$

Choose $x = \frac{y}{\lambda}$:

$$\Phi_{\Gamma_t}(z) = \exp\left[\gamma t \sum_{n=1}^{\infty} \frac{(iz)^n}{n!} \int_{\mathbb{R}^+} x^{n-1}e^{-\lambda x}dx\right].$$

Using the expansion $e^{izx} = 1 + \sum_{n=1}^{\infty} \frac{(izx)^n}{n!}$:

$$\Phi_{\Gamma_t}(z) = \exp\left[\gamma t \int_{\mathbb{R}^+} (e^{izx} - 1)\frac{e^{-\lambda x}}{x} dx\right].$$

Since the gamma process is of finite variation, the truncation function $g(x) = 0$ may be used [7, Corollary 3.1]. Hence the Lévy-Khinchin representation with 'drift' equal to zero, implies that:

$$\nu_{\Gamma_t}(x) = \frac{\gamma e^{-\lambda x}}{x} \mathbb{1}_{x>0}. \tag{3.4}$$

Various representations, expressed in terms of the gamma process, for the VG process are now considered.

3.1.3 The Difference of Two Gamma Processes

Since the VG process is of finite variation, the Jordan decomposition theorem [28, Theorem 1.6] dictates that it may be written as the difference of two increasing processes. As the name 'variance gamma' suggests, these two processes are gamma processes.

Consider the difference of two independent gamma processes:

$$X_t = \Gamma_t(\gamma_p, \lambda_p) - \Gamma_t(\gamma_n, \lambda_n), \tag{3.5}$$

with

$$\gamma_p = \frac{1}{\nu}, \quad \lambda_p = \frac{2}{\nu\left(\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \theta\right)}, \tag{3.6}$$

$$\gamma_n = \frac{1}{\nu}, \quad \lambda_n = \frac{2}{\nu\left(\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \theta\right)}, \tag{3.7}$$

in terms of the parametrization given in (3.1). Examining the effect on the characteristic functions implied by (3.5):

$$\Phi_{X_t}(z) = \Phi_{\Gamma_t(\gamma_p,\lambda_p)}(z)\Phi_{-\Gamma_t(\gamma_n,\lambda_n)}(z) = \Phi_{\Gamma_t(\gamma_p,\lambda_p)}(z)\Phi_{\Gamma_t(\gamma_n,-\lambda_n)}(z)$$

from Lemma 3.2.

$$\Rightarrow \Phi_{X_t}(z) = \left[\left(1 - \frac{i\nu z}{2}\left(\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \theta\right)\right)\left(1 + \frac{i\nu z}{2}\left(\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \theta\right)\right)\right]^{-\frac{t}{\nu}}$$
$$= \left[1 - i\nu\theta z + \frac{\nu\theta^2 z^2}{2}\right]^{-\frac{t}{\nu}} = \Phi_{\mathrm{VG}}(z).$$

Hence, the difference of two gamma processes with the same γ parameter is a VG process.

Once expressed in this way, it is trivial to show that the Lévy measure of the VG process may take following form:

$$\nu_{\mathrm{VG}}(x) = \begin{cases} \frac{Ce^{-Mx}}{x} & x > 0, \\ \frac{Ce^{Gx}}{|x|} & x < 0, \end{cases} \tag{3.8}$$

for $C = \frac{1}{\nu} > 0$, $G = -\lambda_n \geq 0$, $M = \lambda_p \geq 0$, since both the positive and negative components arise independently from each of the gamma processes. Examining this parameterization, the interpretation of each parameter becomes clear: the level of activity, C, scales the expected number of jumps of all sizes (Figure 3.1 illustrates the effect of altering this parameter). It also acts as a time scale parameter — two VG processes, one with a C parameter twice as large the other, will be distributionally identical if time were measured at half the rate for the more active process (this fact arises upon examination of the exponent of the characteristic function). Hence, increasing C, can be thought of as squeezing more of the 'same' activity into a shorter time frame. The parameters M and G control the positive and negative tails of the Lévy density respectively, increasing these parameters reduces the number of large (positive or negative) jumps that are expected per unit time (see Figures 3.2 and 3.3).

3.1.4 The VG Process as a Time Changed Brownian Motion

It will be proven in Section 4.2, that the VG process satisfies the requirements for Brownian motion subordination (Theorem 2.26). Unsurprisingly, the desired time-change (see Section 2.5.4) is a gamma process with mean $\mu = 1$ and variance $\nu = \nu$ per unit time:

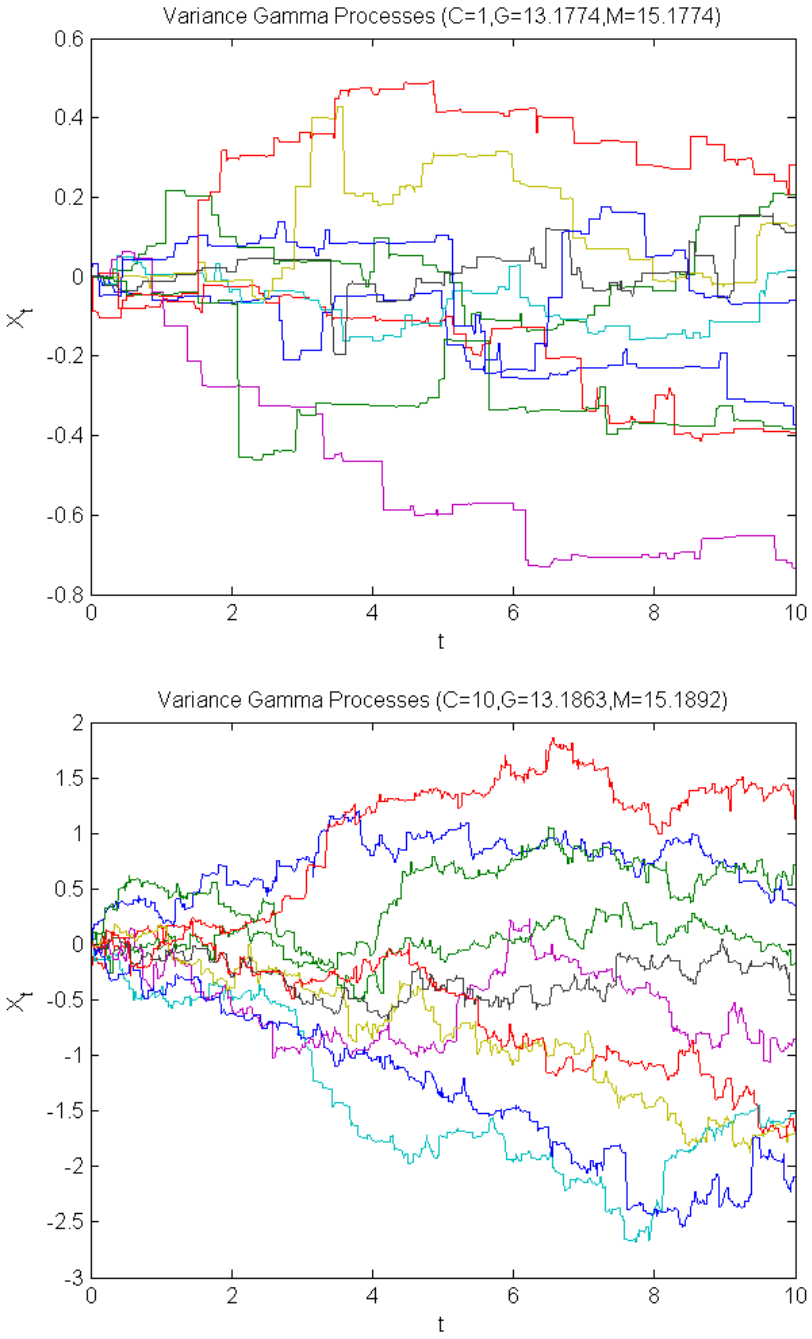

Fig. 3.1: Ten realizations per graph of the VG process with parameters as indicated. Note the larger axis for the higher value of C, indicating that there is much more jump activity.

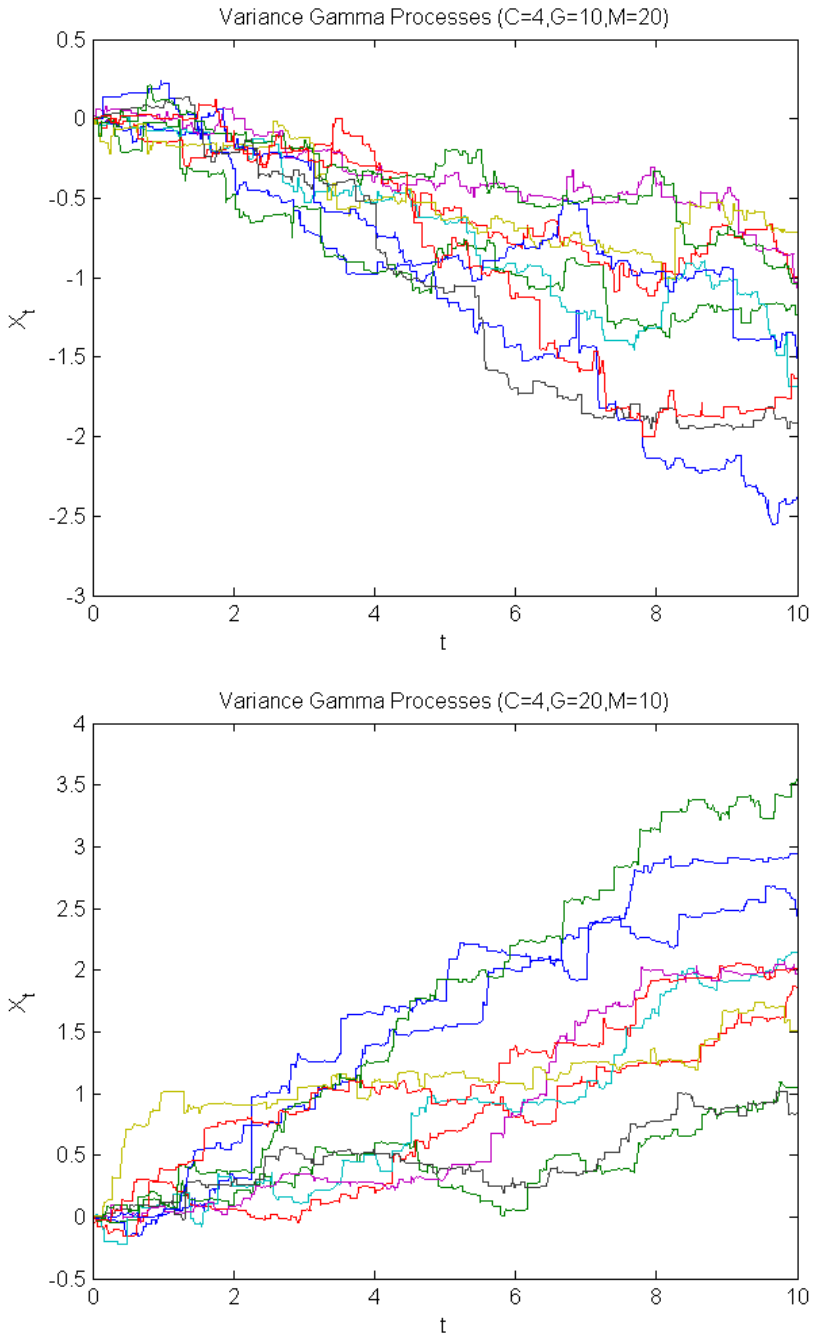

Fig. 3.2: Variance Gamma paths in which first positive then negative jumps are curtailed by increasing M then G.

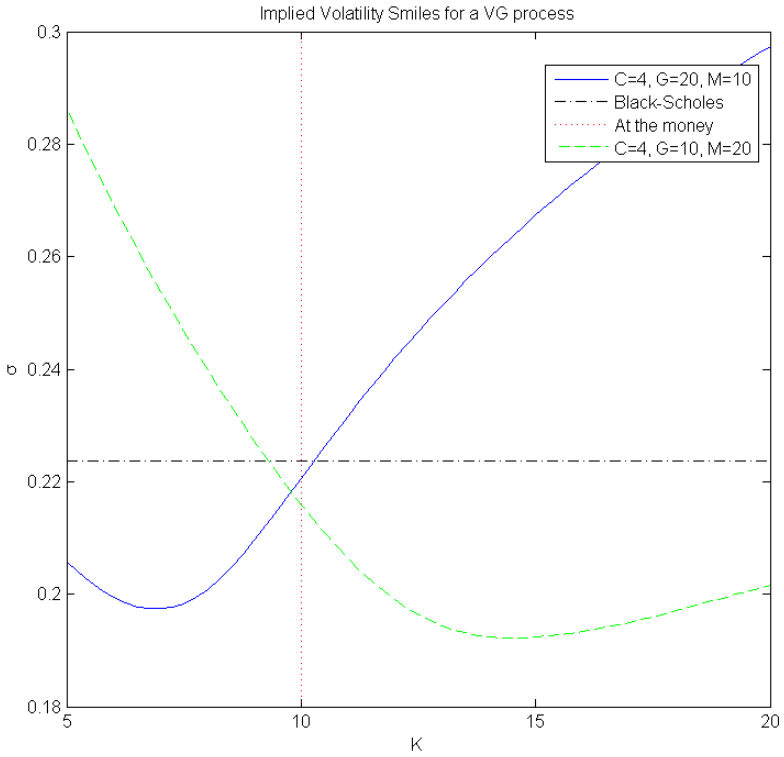

Fig. 3.3: Implied volatility smiles, for an European option maturing at $T = 1$, generated for the parameters indicated in the legend. The dotted and dashed black line is the standard deviation of the process (Black-Scholes model volatility), while the dotted red line is the at-the-money strike. The methods for generating this graph are provided in Section 5.2. In each case, the risk neutral measure was constructed by modifying the drift only, i.e. using $\mu = r$ in (5.2).

Theorem 3.4 (The VG Process as a Time Changed Brownian Motion)**.** *Let $B(t),\ t \in [0,T]$ be a Brownian motion and $\Gamma_t(\gamma,\lambda)$ be a gamma process (with representation (3.3)). Then the process*

$$X_t = \sigma B\left(\Gamma_t\left(\frac{1}{\nu},\frac{1}{\nu}\right)\right) + \theta t\,, \tag{3.9}$$

is a VG process with parameters (σ,θ,ν).

Proof. From Theorem 2.25:

$$\Phi_{\sigma B\left(\Gamma_t\left(\frac{1}{\nu},\frac{1}{\nu}\right)\right)+\theta t}(z) = e^{tl(\psi(z))}$$

with

$$\psi(z) = i\theta z - \frac{1}{2}\sigma^2 z^2, \quad \text{from (2.24), and}$$
$$l(u) = -\frac{1}{\nu}\log\left(1-\nu u\right), \quad \text{from (3.3).}$$

$$\Rightarrow \Phi_{\sigma B\left(\Gamma_t\left(\frac{1}{\nu},\frac{1}{\nu}\right)\right)+\theta t}(z) = \exp\left[-\frac{t}{\nu}\log\left(1 - iz\theta\nu + \frac{z^2\sigma^2\nu}{2}\right)\right] = \Phi_{\mathrm{VG}}(z).$$

■

As discussed in Section 2.5.4, the above representation suggests that returns are normally distributed on a random time scale, which, in this case, is given by the gamma process.

3.1.5 An Improved VG Process

This subordinator representation suggests a possible improvement to the VG model: while the gamma process may adequately capture the random behavior of 'trading time', it does not take into account several observed effects. Namely, that trading occurs at irregularly spaced 'business hours', outside of which information relevant to the process may accumulate. Moreover, during the trading day, volumes (activity) are highest at market open and close and taper off towards midday [42]. As a result, the following is proposed as a refinement of the gamma subordinator:

$$Z_t = \Gamma_t\left(\frac{1}{\nu(t)},\frac{1}{\nu(t)}\right) + \sum_i a_i \mathbb{1}_{t>t_i}. \tag{3.10}$$

The time variable t records time while the market is open. The jump times $t_0, t_1, \ldots$ represent known times at which the market opens. The jump sizes, $a_1, a_2, \ldots$, take into account the amount of price information that will arrive while the market is closed and may be modeled as random variables or deterministic constants related to the length of the recess. Note that ν is now time dependent, allowing the level of activity to be adjusted at various times during the day. This results in the following process for returns:

$$\begin{aligned} X_t &= \sigma B\left(Z_t\right) + \theta t \\ &= \sigma B\left(\Gamma_t\left(\frac{1}{\nu(t)}, \frac{1}{\nu(t)}\right) + \sum_i a_i \mathbb{1}_{t>t_i}\right) + \theta t \\ &= X_t^{\mathrm{VG}}(\sigma, \theta, \nu(t)) + \sum_i N_i \mathbb{1}_{t>t_i}. \end{aligned} \tag{3.11}$$

The summation term forms a Poisson-like process with jumps fixed in time, each N_i is independently distributed $\mathcal{N}(0, a_i)$.

Inevitably, such changes reduce the tractability of the resulting process — the process is no longer a Lévy process, since it is not stationary and depends on a time-varying parameter $\nu(t)$. Furthermore, in contravention of property four of Definition 2.8, such a process exhibits jumps at fixed times. At this stage the above model is mere conjecture and is unsupported by empirical investigation (since it is not the subject of this dissertation).

3.2 CGMY

To extend the VG process, Carr et. al. [4] add a $Y < 2$ parameter to the model, redefining the Lévy density as follows:

Definition 3.5 (CGMY process)**.** The *CGMY process* is the Lévy process with Lévy density:

$$\nu_{\mathrm{CGMY}}(x) = \begin{cases} \frac{Ce^{-Mx}}{x^{1+Y}} & x > 0, \\ \frac{Ce^{Gx}}{|x|^{1+Y}} & x < 0. \end{cases} \tag{3.12}$$

3.2.1 Properties

The Y parameter provides control over the integrability of the Lévy density in the region of zero, dictating the fine structure of the process by controlling the behavior of small jumps:

For $Y < -1$, the Lévy density is not monotonic, causing small jumps to arrive less frequently than large jumps. In this case, the exponent of the $|x|$ factor becomes positive, forcing the density to zero in the limit as $x \to 0$ and creating turning points at $-\frac{Y+1}{M}$, $\frac{Y+1}{G}$ in each tail respectively. This implies that infinitesimally small jumps almost never occur, while any range of larger jumps is far more likely. Recall from Section 2.5.1 that the integral of the Lévy density describes the expected frequency of jumps over the domain of integration. For $Y = -1$, the process reduces to the difference of two Poisson processes with exponentially distributed jumps.

When $1 < Y < 0$, the density is completely monotonic and integrates to a finite value:

$$\int_{\mathbb{R}/\{0\}} \nu(dx) = C\left[\int_{\mathbb{R}^+} \frac{e^{-Mx}}{x^{1+Y}}dx + \int_{\mathbb{R}^-} \frac{e^{Gx}}{(-x)^{1+Y}}dx\right]. \tag{3.13}$$

Performing a change of variables $y = Mx$ and $z = -Gx$:

$$\int_{\mathbb{R}/\{0\}} \nu(dx) = C\left[\int_{\mathbb{R}^+} \frac{M^{1+Y}e^{-y}}{y^{1+Y}}\frac{dy}{M} + \int_{\mathbb{R}^+} \frac{G^{1+Y}e^{-z}}{z^{1+Y}}\frac{dz}{G}\right] \tag{3.14}$$

and using the definition of the gamma function for arguments with positive real parts:

$$\int_{\mathbb{R}/\{0\}} \nu(dx) = C\Gamma(-Y)\left(M^Y + G^Y\right) \tag{3.15}$$

this indicates that the process has a finite number of jumps of all sizes (aggregate arrival rate) and consequently, has finite activity for $Y < 0$ (and by implication finite variation). When $Y = 0$, the VG process results.

When Y exceeds zero, the above integral (3.13) becomes divergent in the region of zero, resulting in infinitely many small jumps. However, for $Y < 1$ the sum of all expected jump sizes (i.e. the total variation, see [7, Proposition 3.9]) is still expected to be finite, since

$$\int_{\mathbb{R}/\{0\}} |x|\nu(dx) = C\Gamma(-Y+1)\left(M^{Y-1} + G^{Y-1}\right) < \infty. \tag{3.16}$$

If $Y > 1$, then the process has infinite total variation but finite quadratic variation ($\int_{\mathbb{R}/\{0\}} x^2 \nu(dx) = C\Gamma(-Y+2)\left(M^{Y-2} + G^{Y-2}\right) < \infty$ almost surely). The parameter Y is restricted to be smaller than two, since valid Lévy densities must integrate $|x|^2$ in the region of zero (see Theorem 2.21). The effect of modifying the Y parameter is illustrated in Figures 3.4 and 3.5.

Both the C and Y parameters are invariant under an equivalent measure change [4]. Hence, regardless of investor preference, the basic form of the asset price process dictates the nature of the risk-neutral process.

To compute the characteristic function for all values of Y, it is simplest to employ the Lévy-Khinchin relation with a centered truncation function $g_c(x) = 1$, see Section 2.5.2 [7, pg. 121]. In this case, the γ_c of the Lévy triplet corresponds to the drift. If no truncation is used (as in the proof in [4]), several integrals will diverge for $Y > 1$, invalidating the simplification to the gamma function in (3.17).

$$\begin{aligned}\Phi_{\text{CGMY}}(z) &= \exp\left[i\gamma_c z + t\int_{\mathbb{R}/\{0\}} (e^{izx} - 1 - izx)\nu_{\text{CGMY}}(dx)\right] \\ &= \exp\left[i\gamma_c z + tC\int_{\mathbb{R}/\{0\}} \left(e^{izx} - 1 - izx\right)\left(\frac{e^{-Mx}}{x^{1+Y}}\mathbb{1}_{x>0} + \frac{e^{Gx}}{|x|^{1+Y}}\mathbb{1}_{x<0}\right)dx\right],\end{aligned}$$

using the series $e^{izx} = 1 + izx + \sum_{n=2}^{\infty}\frac{(izx)^n}{n!}$:

$$\begin{aligned}\Phi_{\text{CGMY}}(z) = \exp\Bigg[i\gamma_c z + tC\sum_{n=2}^{\infty}\frac{(iz)^n}{n!}\Bigg(&\int_{\mathbb{R}^+} x^{n-1-Y}e^{-Mx}dx \\ &+ \int_{\mathbb{R}^-}(-1)^n(-x)^{n-1-Y}e^{Gx}dx\Bigg)\Bigg].\end{aligned}$$

Choose $w = Mx$, $v = -Gx$:

$$\begin{aligned}\Phi_{\text{CGMY}}(z) &= \exp\Bigg[i\gamma_c z + tC\sum_{n=2}^{\infty}\frac{(iz)^n}{n!}\Bigg(M^{Y-n}\int_{\mathbb{R}^+} w^{n-1-Y}e^{-w}dw \\ &\qquad +(-1)^n G^{Y-n}\int_{\mathbb{R}^+} v^{n-1-Y}e^{-v}dv\Bigg)\Bigg] \\ &= \exp\Bigg[i\gamma_c z + tC\Bigg(\sum_{n=2}^{\infty}\frac{(iz)^n}{n!}M^{Y-n}\Gamma(n-Y) \\ &\qquad + \sum_{n=2}^{\infty}\frac{(-iz)^n}{n!}G^{Y-n}\Gamma(n-Y)\Bigg)\Bigg],\end{aligned} \tag{3.17}$$

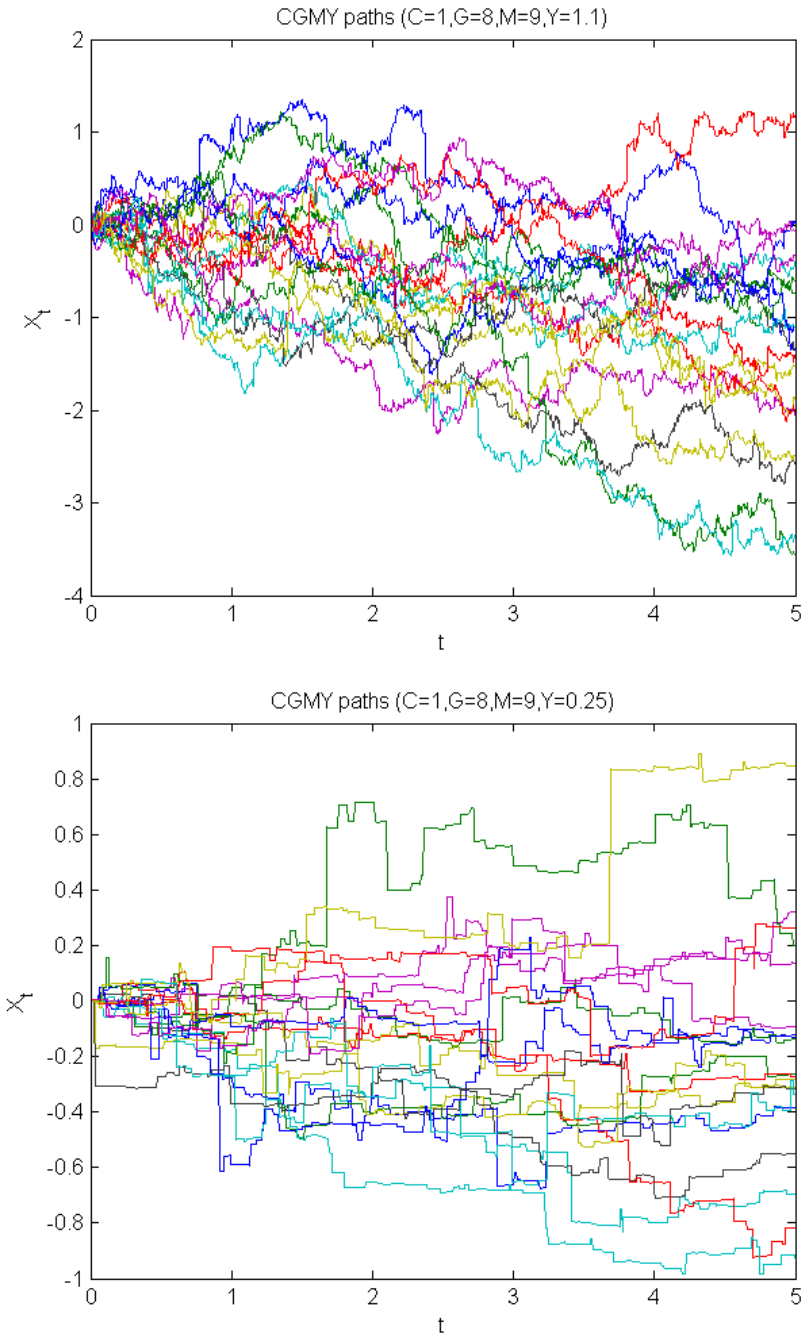

Fig. 3.4: CGMY paths with parameters as indicated. A larger value of Y leads to a proliferation of small jumps and greater dispersion.

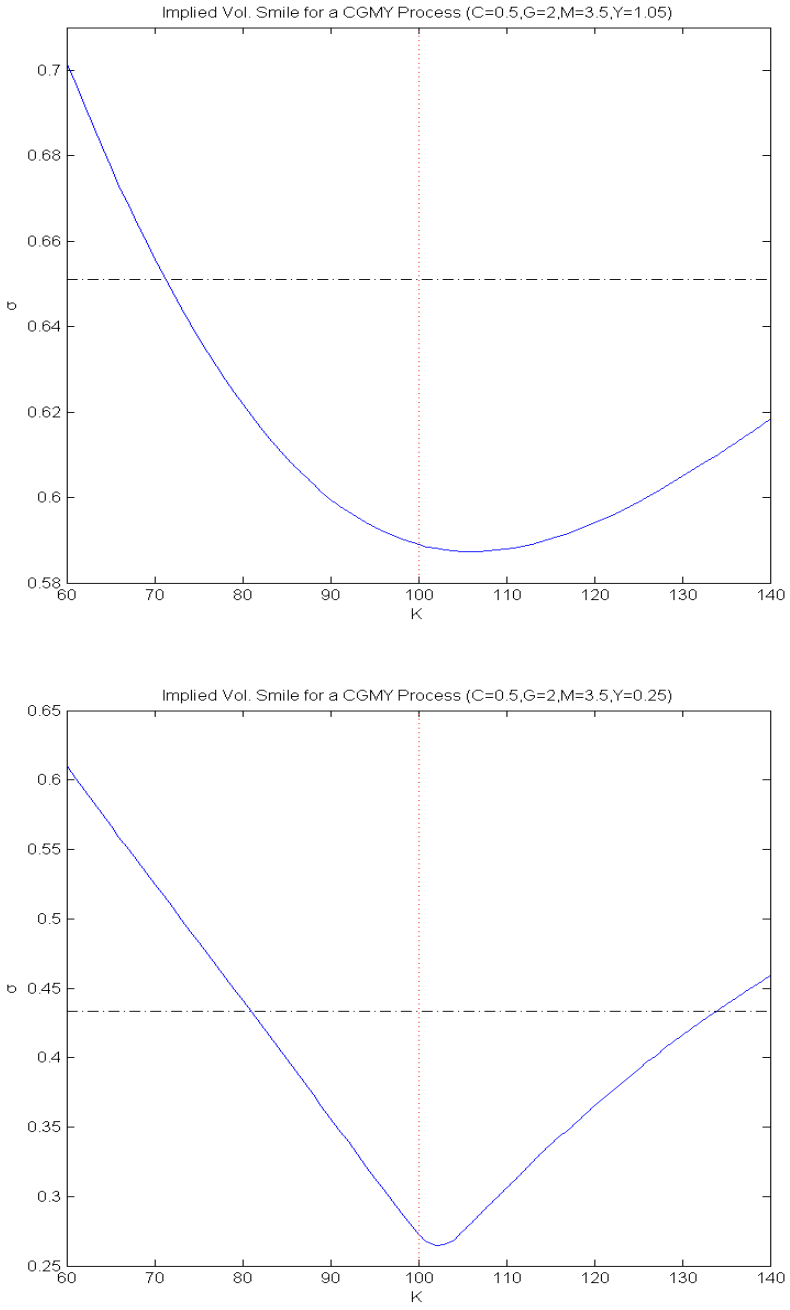

Fig. 3.5: Implied volatility smiles for European options with maturity $T = 0.25$. Increasing the Y parameter introduces more small jumps, amplifying the volatility (translating the first smile upwards) and reducing kurtosis (flattening the first curve, since the process is more like a Brownian motion.) This is evident from the moments given in Table 3.2, as $Y \to 2$, $\Gamma(2-Y) \to \infty \;\Rightarrow\; \text{VAR} = \infty,\ \kappa = 0$.

since $n > Y$. Using the recursion relation $\Gamma(p+1) = p\Gamma(p)$:

$$\Phi_{\text{CGMY}}(z) = \exp\left[i\gamma_c z + tC\Gamma(-Y)\left(M^Y \sum_{n=2}^{\infty} \frac{(-iz/M)^n}{n!} \prod_{i=0}^{n-1}(Y-i) + G^Y \sum_{n=2}^{\infty} \frac{(iz/G)^n}{n!} \prod_{i=0}^{n-1}(Y-i)\right)\right].$$

Using the series $(1+x)^\alpha = 1 + \alpha x + \sum_{n=2}^{\infty} \frac{x^n \prod_{i=0}^{n-1}(\alpha - i)}{n!}$:

$$\Phi_{\text{CGMY}}(z) = \exp\left[i\gamma_c z + tC\Gamma(-Y)\left((M-iz)^Y - M^Y + (G+iz)^Y - G^Y + izY\left(M^{Y-1} - G^{Y-1}\right)\right)\right].$$

For $\gamma_c = tC\Gamma(1-Y)(M^{Y-1} - G^{Y-1})$:

$$\Phi_{\text{CGMY}}(z) = \exp\left[tC\Gamma(-Y)\left((M-iz)^Y - M^Y + (G+iz)^Y - G^Y\right)\right]$$

Using the characteristic function above and the recursion relation for the gamma function $\Gamma(p+1) = p\Gamma(p)$, the cumulants given in Table 3.2 may be computed.

Name	**Symbol**	**Value**
Characteristic Function	$\Phi_{X_t}(z)$	$\exp\left[tC\Gamma(-Y)\left((M-iz)^Y - M^Y + (iz+G)^Y - G^Y\right)\right]$
Lévy Density	$\nu(x)$	$\frac{Ce^{Ax-B\lvert x\rvert}}{\lvert x\rvert^{1+Y}} \quad A = \frac{G-M}{2},\ B = \frac{G+M}{2}$
Expectation	$\mathbb{E}[X_t]$	$tC\Gamma(1-Y)(M^{Y-1} - G^{Y-1})$
Variance	$\text{VAR}[X_t]$	$tC\Gamma(2-Y)(M^{Y-2} + G^{Y-2})$
Skewness	$S[X_t]$	$\frac{C\Gamma(3-Y)(M^{Y-3} - G^{Y-3})}{\sqrt{t}[C\Gamma(2-Y)(M^{Y-2} + G^{Y-2})]^{3/2}}$
Kurtosis	$\kappa[X_t]$	$\frac{C\Gamma(4-Y)(M^{Y-4} + G^{Y-4})}{t[C\Gamma(2-Y)(M^{Y-2} + G^{Y-2})]^2}$

Tab. 3.2: A summary of the properties of the CGMY process. Skewness and kurtosis are defined in the same way as (2.11) and (2.12)

The originators of the CGMY process used it to investigate the stochastic properties of returns [4]. They concluded that observed returns, particularly those of indices, can be described by a pure jump process of finite variation and infinite activity (such as the VG process). Moreover, they noted that the Brownian component of the returns process, whilst present to a small degree in single stocks, is notably

absent in indices, leading the authors to conjecture (based on option price data) that this continuous component may be diversified away and hence should not form part of the risk neutral process.

3.2.2 The Tempered Stable Process

The tempered stable process further generalizes the CGMY process, allowing for different levels of activity (and even different stochastic properties) for positive and negative jumps, using the Lévy measure:

$$\nu(x) = \begin{cases} \frac{c_+ e^{-\lambda_+ x}}{x^{1+\alpha_+}} & x > 0, \\ \frac{c_- e^{\lambda_- x}}{|x|^{1+\alpha_-}} & x < 0, \end{cases} \tag{3.18}$$

where $c_+, c_-, \lambda_+, \lambda_- \in \mathbb{R}^+$ and $\alpha_+, \alpha_- < 2$ are constants. Despite its generality, this process is not considered in the numerical investigation, since it cannot be expressed as an independent, subordinated Brownian motion [7, Proposition 4.1] (introducing complications when generating correlated paths). A thorough treatment of the generalized tempered stable process may be found in [7, Section 4.5].

Chapter 4

Simulation

Monte-Carlo simulation provides one of the most flexible and general methods for estimating risk-neutral prices and other statistical quantities (such as profit and loss distributions, which encapsulate risk measures such as VaR and expected shortfall). Furthermore, such techniques have been shown to be more computationally efficient than most traditional quadrature techniques in evaluating multi-dimensional problems [25] (such as those commonly encountered when dealing with exotic options, particularly when an appropriate variance reduction technique is applied). In many instances, when a problem is too complex to solve using classical methods, Monte-Carlo simulation provides the only approach to estimate a solution.

To facilitate the use of such methods, various techniques for approximating Lévy process paths are now presented. Although some of the methods presented here may be applicable to other Lévy processes, particular attention is paid to the variance gamma family of processes which were presented in Chapter 3.

4.1 The Variance Gamma Process

Regrettably, the density of the variance gamma process cannot be expressed in a simple form. Consequently, it is difficult to generate VG sample paths using the inverse of the cumulative density or similar methods. However, it was shown in Section 3.1, that the VG process can be expressed as the difference of two independent gamma processes or as a Brownian motion subordinated by a gamma process. Hence, the gamma process is central to simulating the VG process. In order that this

document should provide a self-contained reference, two rejection based methods, based on Cont and Tankov [7, Chapter 6], for sampling from a gamma distribution are now presented:

Algorithm 4.1 (Johnk's Generator)**.** *To simulate a* $\Gamma(a,1)$ *distributed random variable with* $a \leq 1$*:*

Generate an exponential r.v. E
IF $(a = 1)$
    RETURN E
ELSE
    Set $X, Y = 1$
    WHILE $(X + Y > 1)$
        Generate i.i.d. uniform [0,1] r.v.s U,V
        Set $X = U^{\frac{1}{a}}$, $Y = V^{\frac{1}{1-a}}$
    END WHILE
END IF
RETURN $\frac{XE}{X+Y}$

Algorithm 4.2 (Best's Generator)**.** *To simulate a* $\Gamma(a,1)$ *distributed random variable with* $a > 1$*:*

Set $b = a - 1, c = 3a - \frac{3}{4}$
Set $X = 0$
WHILE $\log(64W^3V^3) > 2\left(b\log\left(\frac{X}{b}\right) - Y\right)$ OR $X \leq 0$
    Generate i.i.d. uniform [0,1] r.v.s U,V
    Set $W = U(1-U), Y = \sqrt{\frac{c}{W}}(U - 1/2), X = b + Y$
END WHILE
RETURN X

With the ability to simulate gamma random variables, we can perform the Brownian subordination on a fixed time grid $t_1, ..., t_n$:

Algorithm 4.3 (The VG Process as a Subordination)**.** *To generate a variance gamma process with parameters* (σ, ν, θ)*:*

Generate n independent gamma increments $\Delta S_1, ..., \Delta S_n$
with parameters $\left(\frac{t_1}{\nu}, 1\right), \left(\frac{t_2 - t_1}{\nu}, 1\right), ..., \left(\frac{t_n - t_{n-1}}{\nu}, 1\right)$.
Set $\Delta S_i = \nu \Delta S_i$ for all i
Generate n i.i.d. standard normal r.v.s $N_1, ..., N_n$
Set $\Delta X_i = \sigma N_i \sqrt{\Delta S_i} + \theta \Delta S_i$ for all i.
The trajectory is given by $X(t_i) = \sum_{k=1}^{i} \Delta X_k$

Alternately, the process may be generated as the difference of two independent gamma process:

Algorithm 4.4 (The VG Process as the Difference of Two Gamma processes)**.** *To generate a variance gamma process with parameters* (σ, ν, θ)*:*

Set $\nu_p = \nu \left(\frac{1}{2} \sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2} \right)$, $\nu_n = \nu \left(\frac{1}{2} \sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2} \right)$.
Generate two independent sets of n gamma distributed increments:
$\Delta S_1^{(1)}, ..., \Delta S_n^{(1)}$ and $\Delta S_1^{(2)}, ..., \Delta S_n^{(2)}$ with parameters $\left(\frac{t_1}{\nu}, 1\right), \left(\frac{t_2 - t_1}{\nu}, 1\right), ..., \left(\frac{t_n - t_{n-1}}{\nu}, 1\right)$.
Set $\Delta S_i^{(1)} = \nu_p \Delta S_i^{(1)}$ for all i
Set $\Delta S_i^{(2)} = \nu_n \Delta S_i^{(2)}$ for all i
The trajectory is given by $X(t_i) = \sum_{k=1}^{i} \Delta S_k^{(1)} - \Delta S_m^{(2)}$

4.2 CGMY Simulation

Like the variance gamma process, the more general CGMY process does not have a simple analytical expression for its density, disqualifying simple simulation techniques. However, unlike the VG process, the form of the subordinator of the CGMY process is not obvious, presenting several challenges for simulation. Naturally, the process can be approximated by a compound Poisson process. However, this method of simulation would make it difficult to incorporate a dependence structure in multi-dimensional problems. Fortunately, Madan and Yor [32] were able to compute the Lévy density of the CGMY subordinator and show that it is *absolutely continuous*[1]

[1] If μ and ν are two measures on the same measurable space Ω, then μ is said to be *absolutely continuous* with respect to ν if $\mu(A) = 0$ for every set $A \subseteq \Omega$ for which $\nu(A) = 0$. This is different to equivalence, since μ may have a larger null space than ν. However, if both measures are absolutely

with respect to an α-stable subordinator. This allows one to use rejection techniques to approximate the subordinator and hence the process. In addition, dependence may be incorporated through the normal random variables of the subordinated process.

4.2.1 The Stable Process

As discussed, the stable process is central to simulating CGMY paths. In order to clarify the subsequent analysis, the stable process is now introduced.

An α-stable process with parameters $(\alpha \in (0,2], \beta \in [-1,1], \sigma \geq 0)$ is a pure jump Lévy process. The process is so-named because its distribution is stable under addition: if $X^{(1)}, \ldots, X^{(n)}$ are independent copies of a stable random variable X, then

$$X^{(1)} + \cdots + X^{(n)} \stackrel{d}{=} cX + d \tag{4.1}$$

$$\Rightarrow \Phi_X(z)^n = \Phi_X(cz)e^{idz} \tag{4.2}$$

for $c > 0$ and $d \in \mathbb{R}$.

Owing to the stability of their distributions, the family of stable processes exhaust the class of Lévy processes that are *self-similar* under translation [7, pg. 94]. This is evident, since iff X_t is a stable Lévy process,

$$X_{at} = \sum_{i=0}^{n-1} X_{(i+1)at/n} - X_{iat/n}.$$

For any Lévy process, the increments are i.i.d:

$$\Rightarrow X_{at} \stackrel{d}{=} \sum_{i=0}^{n-1} X^{(i)}_{at/n}.$$

Applying the stability property (4.1):

$$\Rightarrow X_{at} \stackrel{d}{=} mX_{at/n} + k.$$

This suggests (based on several corollaries in [38, Section 2.1]) that:

$$(X_{at})_{t\in[0,T]} \stackrel{d}{=} \left(a^{\frac{1}{\alpha}} X_t + ct\right)_{t\in[0,T]}, \tag{4.3}$$

continuous with respect to each other, then they are equivalent.

for any $a > 0$ and some $c \in \mathbb{R}$.

It has been shown [39, pg. 77-80], that a stable processes must have a Lévy measure of the form:

$$k(x) = \frac{c_p}{x^{1+\alpha}} \mathbb{1}_{x>0} + \frac{c_n}{|x|^{1+\alpha}} \mathbb{1}_{x<0}. \tag{4.4}$$

Using the Lévy-Khinchin representation, the following standard form for the characteristic function may be derived:

$$\Phi(z) = \begin{cases} \exp\left\{-\sigma^\alpha |z|^\alpha t \left(1 - i\beta \text{sign}(z) \tan\left(\frac{\pi\alpha}{2}\right)\right)\right\}, & \alpha \neq 1 \\ \exp\left\{-\sigma|z|t \left(1 + i\beta \text{sign}(z) \frac{2}{\pi} \log|z|\right)\right\}, & \alpha = 1 \end{cases} \tag{4.5}$$

with

$$\beta = \frac{c_p - c_n}{c_p + c_n} \tag{4.6}$$

and

$$\sigma = \left[\frac{c_p + c_n}{2} \frac{\Gamma\left(\frac{\alpha}{2}\right)\Gamma\left(1 - \frac{\alpha}{2}\right)}{\Gamma(1+\alpha)}\right]. \tag{4.7}$$

The following parametric special cases are noted:

- If $\beta = 0$, then the process is symmetric.
- If $\beta = 1$, then the Lévy density is concentrated on the positive half-axis, i.e. in the case of finite variation ($\alpha < 1$), the process is monotonically increasing.
- When $\alpha = 1$, a translated, Cauchy distributed, process results.
- When $\alpha = 2$, a Wiener process is recovered.

4.2.2 The CGMY Subordinator

If one examines the Brownian subordination requirements given in Theorem 2.26, it becomes clear that the CGMY process satisfies all the conditions:

1. Its density is absolutely continuous with respect to the Lebesgue measure (since $\nu_{\text{CGMY}}(0^+) = \nu_{\text{CGMY}}(0^-)$) and may be expressed as follows:

$$\nu_{\text{CGMY}}(x) = \frac{Ce^{ax-b|x|}}{|x|^{1+Y}}, \quad a = \frac{G-M}{2}, \quad b = \frac{G+M}{2}. \tag{4.8}$$

2. For $\mu = \frac{G-M}{2}$:

$$\begin{aligned}
&\nu_{\text{CGMY}}(-x)e^{\mu x} \\
&= \left[\frac{Ce^{Mx}}{|x|^{1+Y}}\mathbb{1}_{x<0} + \frac{Ce^{-Gx}}{|x|^{1+Y}}\mathbb{1}_{x>0}\right]\exp\left(\frac{(G-M)x}{2}\right) \\
&= \frac{Ce^{\frac{(G+M)x}{2}}}{|x|^{1+Y}}\mathbb{1}_{x<0} + \frac{Ce^{\frac{-(G+M)x}{2}}}{|x|^{1+Y}}\mathbb{1}_{x>0} \\
&= \left[\frac{Ce^{Gx}}{|x|^{1+Y}}\mathbb{1}_{x<0} + \frac{Ce^{-Mx}}{|x|^{1+Y}}\mathbb{1}_{x>0}\right]\exp\left(\frac{(M-G)x}{2}\right) \\
&= \nu_{\text{CGMY}}(x)e^{-\mu x}.
\end{aligned}$$

3. For $\mu = -M$:

$$\begin{aligned}
&\nu_{\text{CGMY}}(\sqrt{u})e^{-\mu\sqrt{u}} \\
&= \frac{Ce^{-M\sqrt{u}}}{u^{\frac{1+Y}{2}}}e^{M\sqrt{u}} \\
&= Cu^{-\frac{1+Y}{2}} \quad \text{which is completely monotonic for } Y > -1.
\end{aligned}$$

By employing Proposition 2.27 and relating the Lévy density of the CGMY process to the Lévy density of a one-sided stable subordinator, Madan and Yor [32] were able to derive the exact form of the subordinator:

Proposition 4.5 (The Lévy Density of the CGMY subordinator)**.** *Let $(Z_t)_{t\in[0,T]}$ be a subordinator with Lévy density:*

$$\nu_{Z_t}(x) = \frac{Ce^{\frac{x}{2}(a^2-\frac{1}{2}b^2)}D_{-Y}(b\sqrt{x})}{|x|^{1+\frac{Y}{2}}}, \tag{4.9}$$

where D is the parabolic cylinder function[2]*, with $a = \frac{G-M}{2}$, $b = \frac{G+M}{2}$. Then the process*

$$X_t = aZ_t + B(Z_t), \tag{4.11}$$

where $B(t) : t \in [0,T]$ is a standard Brownian motion, is a CGMY process.

[2] The parabolic cylinder function $D_\nu(z)$ is defined as the solution to the Weber differential equation [43, pg. 347]:

$$\frac{d^2D_\nu(z)}{dz^2} + \left(\nu + \frac{1}{2} - \frac{1}{4}z^2\right)D_\nu(z) = 0. \tag{4.10}$$

This form (found in [36]) is a slight simplification of the original form produced by Madan and Yor [32].

Proof. See Poirot and Tankov [36]. ■

Furthermore, Madan and Yor [32] were able to express the Lévy density of the subordinator in terms of the Lévy density of an $\frac{\alpha}{2}$-stable process and a continuous function f bounded by one:

Theorem 4.6 (The relationship between CGMY and stable subordinators)**.** *Let ν_{Z_t} be the Lévy density of the CGMY subordinator. Let $\nu_{\frac{\alpha}{2}}$ be the Lévy density of the $\frac{\alpha}{2}$-stable subordinator. The two densities are related as follows:*

$$\nu_{Z_t}(x) = f(x)\nu_{\frac{\alpha}{2}}(x), \tag{4.12}$$

with

$$f(x) = \frac{2^{\frac{Y}{2}}\Gamma\left(\frac{Y}{2}+\frac{1}{2}\right)e^{\frac{a^2}{2}x-\frac{b^2}{4}x}}{\sqrt{\pi}} D_{-Y}(b\sqrt{x}), \tag{4.13}$$

where $a = \frac{G-M}{2}$, $b = \frac{G+M}{2}$; and

$$\nu_{\frac{\alpha}{2}}(x) = \frac{C\sqrt{\pi}}{2^{\frac{Y}{2}}\Gamma\left(\frac{Y}{2}+\frac{1}{2}\right)}\frac{\mathbb{1}_{x>0}}{|x|^{1+\frac{Y}{2}}} \equiv \frac{K}{|x|^{1+\frac{Y}{2}}}\mathbb{1}_{x>0}, \tag{4.14}$$

where $f(x) \leq 1$.

Proof. See Madan and Yor [32] ■

This result implies that the CGMY subordinator is absolutely continuous with respect to the stable subordinator [32].

4.2.3 Rosiński Rejection

Since the subordinator may be expressed in this way (form 4.12), we may apply a rejection technique developed by Rosiński [37] to simulate it using a stable subordinator:

Theorem 4.7 (Rosiński Rejection)**.** *Let $X_0(t) : t \in [0,T]$ be a Lévy process on $\mathbb{R}^d$ with Lévy measure Q_0 and $X(t) : t \in [0,T]$ be a related Lévy process with Lévy measure Q such that:*

$$\frac{dQ}{dQ_0} \leq 1.$$

Let X_0^* be the point process of jumps (see Section 2.5.1) associated with X_0 admitting representation:

$$X_0^* = \sum_{n\geq 1} \delta_{(T_n(\omega), J_n^0(\omega))}, \tag{4.15}$$

where $(T_n)_{n\geq 1}$ is an increasing sequence of non-anticipating times and $(J_n^0)_{n\geq 1}$ is an independent adapted sequence of jumps. Let $(U_n)_{n\geq 1}$ be an i.i.d sequence of uniform [0,1] random variates, define:

$$J_n = \begin{cases} J_n^0 & \text{if } \frac{dQ}{dQ_0}\left(J_n^0\right) > U_n \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$X^* = \sum_{n\geq 1} \delta_{(T_n(\omega), J_n(\omega))}, \tag{4.16}$$

where X^* is the marked point process associated with X.

Proof. See Rosiński [37]. ■

Compound Poisson processes are among the class of processes that may be represented using (4.15). Hence, as will be shown, we must first approximate the stable subordinator (to which the rejection method will be applied) as a compound Poisson process.

4.2.4 Compound Poisson Approximation

Several techniques exist for simulating the stable process directly. However, the stable process may have infinitely many jumps. As a result, in order to make use of the rejection method to generate the desired subordinator (and with it the CGMY process), the $\frac{\alpha}{2}$-stable subordinator must be approximated by a compound Poisson process with drift.

This is achieved by ignoring all jumps smaller than ϵ for some $\epsilon > 0$ and noting that the Lévy density of a compound Poisson process is given by $\lambda g(x)$, where $g(x)$ is the jump size distribution and λ is the intensity of the process (see Section 2.5.1). Consequently, we must choose λ and g such that the Lévy measure of the approximating compound Poisson process corresponds to the truncated Lévy measure of the stable subordinator.

In the particular case under consideration, the truncated Lévy density is given by:

$$\nu_\epsilon(x) = \frac{C\sqrt{\pi}}{2^{\frac{Y}{2}}\Gamma\left(\frac{Y}{2}+\frac{1}{2}\right)}\frac{\mathbb{1}_{x>\epsilon}}{|x|^{1+\frac{Y}{2}}} \equiv \frac{K}{|x|^{1+\frac{Y}{2}}}\mathbb{1}_{x>\epsilon}. \tag{4.17}$$

from (4.14). This density must be normalized to obtain a valid probability density g.

$$\begin{aligned}\int_{-\infty}^{\infty}\nu_\epsilon(x)\,dx &= \int_\epsilon^\infty \frac{K}{x^{1+\frac{Y}{2}}}\,dx = \left[-\frac{2K}{Yx^{\frac{Y}{2}}}\right]_\epsilon^\infty = \frac{2K}{Y\epsilon^{\frac{Y}{2}}} \quad (Y>0)\\ \Rightarrow g(x) &= \frac{\nu_\epsilon(x)}{\int_\epsilon^\infty \nu_\epsilon(x)\,dx} = \frac{Y\epsilon^{\frac{Y}{2}}}{2x^{1+\frac{Y}{2}}}\mathbb{1}_{x>\epsilon},\end{aligned} \tag{4.18}$$

which has cumulative density G.

$$\begin{aligned}G(x) &= \int_{-\infty}^x g(y)\,dy = \int_\epsilon^x \frac{Y\epsilon^{\frac{Y}{2}}}{2y^{1+\frac{Y}{2}}}\,dy = \left[-\frac{\epsilon^{\frac{Y}{2}}}{y^{\frac{Y}{2}}}\right]_\epsilon^x = 1-\frac{\epsilon^{\frac{Y}{2}}}{x^{\frac{Y}{2}}}\\ \Rightarrow G^{-1}(x) &= \frac{\epsilon}{(1-x)^{\frac{2}{Y}}} \quad (x>\epsilon>0).\end{aligned} \tag{4.19}$$

Hence, jump sizes for the approximating process may be generated using the inverse of the cumulative distribution (using $\epsilon/U^{\frac{2}{Y}}$, where U is uniformly distributed and ϵ is vanishingly small). The expected arrival rate of jumps λ must then be chosen so that $\nu_\epsilon(x) = \lambda g(x)$, hence:

$$\lambda = \frac{2K}{Y\epsilon^{\frac{Y}{2}}}. \tag{4.20}$$

To improve the accuracy of the approximation, the small jumps, which were lost during truncation, are replaced with their expected drift per unit time:

$$\begin{aligned}\int_0^\epsilon x\nu_{\frac{Y}{2}}(x)\,dx &= \int_0^\epsilon \frac{K}{x^{\frac{Y}{2}}}\,dx\\ &= \left[\frac{Kx^{1-\frac{Y}{2}}}{1-\frac{Y}{2}}\right]_0^\epsilon = \frac{K\epsilon^{1-\frac{Y}{2}}}{1-\frac{Y}{2}} \equiv d,\end{aligned} \tag{4.21}$$

since $Y<2$.

It should be apparent, that the truncation will introduce an error into the approximation, particularly as $Y \to 2$ which causes the small jumps to become a significant driver of the process dynamics. Inconveniently, the implications of this error on the final process are difficult to quantify [36]. Consequently, small values of ϵ are advised of the order of 10^{-4} or smaller. However, as ϵ decreases, the expected computation time on a fixed time interval increases proportional to $O(\lambda) = O(\epsilon^{-\frac{Y}{2}})$.

4.2.5 A CGMY Simulation Algorithm

Finally, with these building blocks in place an algorithm to simulate the CGMY may be described:

Algorithm 4.8 (CGMY Simulation)**.** *To simulate a CGMY process with parameters* (C, G, M, Y)*:*

Simulate the Subordinator:
    Choose ϵ
    Set $i = 0,\ t(0) = 0,\ Z(0) = 0$
    Generate an exponential r.v. E with rate λ
    WHILE $t(i) + E < T$
        Set $t(i+1) = t(i) + E$
        Generate independent uniform [0,1] r.v.s U, V
        IF $f\left(G^{-1}(U)\right) > V$
            Set $Z(i+1) = Z(i) + G^{-1}(U) + d(t(i+1) - t(i))$
        ELSE
            Set $Z(i+1) = Z(i) + d(t(i+1) - t(i))$
        END IF
        i=i+1
        Generate an exponential r.v. E with rate λ
    END WHILE
Subordinate the Brownian motion:
    Set $j = 0,\ X(0) = 0$
    FOR EACH element of Z
        Set $\Delta Z = Z(j+1) - Z(j)$
        Generate an independent standard normal r.v. N
        Set $X(j+1) = X(j) + A\Delta Z + N\sqrt{\Delta Z}$
        j=j+1
    END FOR
    RETURN t, X

Chapter 5

Pricing

This chapter is focused on pricing options where the underlying follows an exponential Lévy process. Particular attention is paid to the class of processes introduced in Chapter 3. Moreover, this chapter is dependent on results from Chapter 2. Once again, it is assumed that the reader is acquainted with the fundamentals of no-arbitrage pricing theory.

The chapter is laid out as follows: in the first section, the notion of price is discussed. Then, Fourier transform techniques, developed by Carr and Madan [5], are presented for European options where the characteristic function of the underlying is known.

Techniques for generating paths from the variance gamma family of processes are discussed in Chapter 4. Naturally, such techniques facilitate the use of Monte-Carlo methods to estimate the price of most options.

5.1 The Significance of Risk Neutral Prices

In the Black-Scholes model world, the notion of a price has a simple interpretation: the price is the cost of setting up a self-financing hedging strategy that will replicate the option payoff almost surely. As a result, trading at any other price will present an arbitrage opportunity. In this setting, using the second fundamental theorem of asset pricing, the price of an asset H may be expressed as the discounted expectation of its payoff under the unique equivalent martingale measure (EMM) $\mathbb{Q}$ [19]:

$$H_t = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[H_T|\mathcal{F}_t] \tag{5.1}$$

In reality, however, option payoffs cannot be reproduced with certainty and many equivalent martingale measures may exist (this justifies the existence of a market in contingent claims — options are not redundant assets as the Black-Scholes model would suggest).

Here we consider exponential Lévy models of the form:

$$S_t = S_0 \exp\left((\mu + \omega)t + X_t\right), \tag{5.2}$$

where S_0 is the initial price, X_t is some Lévy process, μ is a drift and

$$\omega = -\log\left(\mathbb{E}[e^{X_t}]\right) = -\log(\Phi_{X_t}(-i)), \tag{5.3}$$

adjusts the drift to construct a martingale for $\mu = r$ (see Section 6.3.3). Such models are arbitrage free (provided that the random component is neither non-increasing nor non-decreasing [7, Proposition 9.9]) and all realistic models are incomplete (see Section 2.5.3). Heuristically speaking, this results in a plethora of pricing rules which are arbitrage free (glossing over the subtleties of the general semi-martingale statement of the first fundamental theorem of asset pricing given in [10]).

As we shall see in Section 6.1, for a general exponential Lévy model, an unrealistically wide range of prices may be arbitrage free. This indicates that, to a market participant, the absence of arbitrage is not a sufficient basis for determining a price. It is merely a broad prerequisite. Indeed, attempts to dynamically hedge the option at such arbitrage free prices may result in profit and loss distributions which include massive undesirable losses. It can be argued, that the profit and loss distribution of hedging the option is far more instructive in determining bids/ask spreads than no arbitrage prices.

It should be apparent that the price of an option is intimately linked to the choice of hedging strategy, hence the price can be thought of as: *The cost of setting up a self-financing*[1] *replicating portfolio that will reproduce the option payoff with an acceptable profit and loss distribution.* The acceptability of the profit and loss distribution is subjective and may be determined from its variance or some utility function as we shall see in Chapter 6.

[1] The strategy may be self-financing *under expectation* for more details see Section 6.3.

Taking this into consideration, methods for determining risk-neutral prices are now presented.

5.2 Fourier Transform Techniques

In many instances, when a Lévy process is defined by means of the Lévy measure, the probability density of the increments cannot be expressed in a closed form. Owing to the intractability of this density, it is not possible to compute prices directly using equation (5.1):

$$C_t(S) = e^{-r(T-t)} \int_{\mathbb{R}^+} C_T(S) Q_{(t,T)}(S)\, dS, \tag{5.4}$$

where $(C_t(S))_{t\in[0,T]}$ is some contingent claim written on an underlying S and $Q_{(t,T)}$ is a risk neutral density for the asset price on a time interval $[t, T]$.

However, in the context of Lévy processes, an analytical form for the characteristic function arises naturally from the Lévy-Khinchin representation (see Section 2.5.2). This suggests that Fourier transform techniques may be used to express the price of the option in terms of the characteristic function (see Definition 2.10) of the underlying.

5.2.1 The Method of Carr and Madan

If one considers the price of a European call, with strike K, expressed in terms of the log of the strike, k, and the log of the asset price, s, (5.1) becomes:

$$C_t(k;T) = e^{-r(T-t)} \int_{\mathbb{R}} (e^s - e^k) \mathbb{1}_{s>k} q_{(t,T)}(s)\, ds, \tag{5.5}$$

where $q_{(t,T)}$ is the risk-neutral density of the log price.

Unfortunately, $C_t(k)$ is not square integrable since the negative tail tends to a finite value as k tends to negative infinity ($C_t(k) \to S_t$ as $k \to -\infty$). Consequently, one cannot apply the Fourier transform directly to the option price. To obtain a square integrable function, a damped version of the call price is considered:

$$c_t(k;T) = \exp(\alpha k) C_t(k;T) \tag{5.6}$$

for some $\alpha \in \mathbb{R},\ \alpha > 0$.

The algorithm is fairly robust to the choice of the parameter α. However, selecting an overly large value for α, whilst rapidly damping the negative tail, will interfere with the integrability of the positive tail of the modified call value. Carr and Madan [5, pg. 64], show that a sufficient condition to ensure integrability is:

$$\mathbb{E}[S_T^{\alpha+1}] < \infty. \tag{5.7}$$

Moreover, the authors recommend choosing α as a quarter of the upper bound.

An analytical expression, in terms of the characteristic function of the underlying, may be computed for the Fourier transform of this modified call price: first, the Fourier transform of this function is defined:

$$\Psi_{c_t}(v) \equiv \int_{\mathbb{R}} e^{ivk} c_t(k;T)\, dk. \tag{5.8}$$

The desired result is arrived at by combining this definition and equations (5.5) and (5.6):

$$\Psi_{c_t}(v) = e^{-r(T-t)} \int_{-\infty}^{\infty} e^{(\alpha+iv)k} \int_{-\infty}^{\infty} (e^s - e^k)\mathbb{1}_{s>k} q_{(t,T)}(s)\, dsdk,$$

changing the order of integration yields:

$$\begin{aligned}
\Psi_{c_t}(v) &= e^{-r(T-t)} \int_{-\infty}^{\infty} q_{(t,T)}(s) \int_{-\infty}^{\infty} e^{(\alpha+iv)k}(e^s - e^k)\mathbb{1}_{s>k}\, dkds \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} q_{(t,T)}(s) \int_{-\infty}^{s} (e^s e^{k(iv+\alpha)} - e^{k(1+iv+\alpha)})\, dkds \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} q_{(t,T)}(s) \left[\frac{e^{s(1+iv+\alpha)}}{iv+\alpha} - \frac{e^{s(1+iv+\alpha)}}{1+iv+\alpha}\right] ds \\
&= \frac{e^{-r(T-t)}}{(iv+\alpha)(1+iv+\alpha)} \int_{-\infty}^{\infty} q_{(t,T)}(s) e^{is(v-i(\alpha+1))} \left[1 + iv + \alpha - iv - \alpha\right] ds.
\end{aligned}$$

From the definition of the characteristic function (2.10):

$$\Psi_{c_t}(v) = \frac{e^{-r(T-t)} \Phi_{s_{(t,T)}}(v - i(\alpha+1))}{(iv+\alpha)(1+iv+\alpha)}. \tag{5.9}$$

A similar line of reasoning can be used to show that:

$$\begin{aligned}
\Psi_{c_t^*}(v) &= Ae^{-r(T-t)} \int_{-\infty}^{\infty} q_{(t,T)}(s) \int_{-\infty}^{s} e^{(\alpha+iv)k}\, dkds \\
&= \frac{Ae^{-r(T-t)}}{\alpha+iv} \int_{-\infty}^{\infty} q_{(t,T)}(s) e^{is(v-\alpha i)}\, ds. \\
&= \frac{Ae^{-r(T-t)} \Phi_{s_{(t,T)}}(v - i\alpha)}{\alpha+iv},
\end{aligned} \tag{5.10}$$

for a cash-or-nothing call C_t^*, potentially paying A at maturity.

If the asset price dynamics under a particular risk neutral measure are given by an exponential Lévy model (5.2), then the characteristic function of the log asset price, $\Phi_{s_{(t,T)}}$, may be replaced with:

$$\begin{aligned} s_{(t,T)} = s_{T-t} &= \log(S_t) + (r+\omega)(T-t) + X_{T-t} \\ \Rightarrow \Phi_{s_{T-t}}(z) &= \exp\left\{iz\left(\log(S_t) + (r+\omega)(T-t)\right)\right\}\Phi_{X_{T-t}}(z) \end{aligned} \tag{5.11}$$

in terms of $\Phi_{X_{T-t}}$, the characteristic function of the martingale component of S.

Finally, using the inverse of the Fourier transform, an expression for the original call price may be obtained:

$$C_t(k;T) = \frac{\exp(-\alpha k)}{2\pi}\int_{-\infty}^{\infty} e^{-ivk}\Psi_{c_t}(v)\,dv = \frac{\exp(-\alpha k)}{\pi}\Re\left\{\int_0^{\infty} e^{-ivk}\Psi_{c_t}(v)\,dv\right\}. \tag{5.12}$$

The second equality holds because the call price is real, which implies that its Fourier transform is odd in its imaginary part and even in its real part.

Put prices, for the same martingale measure, may be obtained using the following parity relationship derived from elementary no-arbitrage arguments:

$$\begin{aligned} P_T(K;T) &= C_T(K;T) - S_T + K \\ \Rightarrow e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[P_T(K;T)] &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[C_T(K;T)] - e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[S_T] + Ke^{-r(T-t)} \\ \Rightarrow P_t(K;T) &= C_t(K;T) - S_t + Ke^{-r(T-t)}. \end{aligned} \tag{5.13}$$

5.2.2 The Fast Fourier Transform

In order to harness the tremendous computational efficiencies[2] of the fast Fourier transform (FFT), expression (5.12) can be modified to conform to the specifications of the algorithm:

The FFT algorithm can be used to rapidly evaluate the sum:

$$w(k) = \sum_{j=1}^{N}\exp\left(-i\frac{2\pi}{N}(j-1)(k-1)\right)x(j) \equiv \mathrm{FFT}(x(j)), \tag{5.14}$$

[2] Computing the Fourier transform naïvely at N points, by evaluating the integral at N points, will have computational complexity proportional to $O(N^2)$. The FFT improves the complexity substantially to $O\left(N\log_2(N)\right)$.

for $k = 1, \ldots, N$, where N is typically a power of 2.

The call price (5.12) may be approximated for a given strike using the trapezoid rule, with spacing η, as follows:

$$C_t(k;T) \approx \frac{\exp(-\alpha k)}{\pi} \Re \left\{ \sum_{j=1}^{N} e^{-iv_j k} \Psi_{c_t}(v_j) \eta \right\}, \tag{5.15}$$

where

$$v_j = \eta(j-1). \tag{5.16}$$

Principally, we are concerned with values of $C_t(k;T)$ that are near the at-the-money point. A regular spacing of λ is thus employed to place the strikes in the desired region of $k = 0$, as follows:

$$k_u = -b + \lambda(u-1), \tag{5.17}$$

for $u = 1, \ldots, N$, yielding strikes on $[-b; b]$ with

$$b = \frac{(N-1)\lambda}{2}. \tag{5.18}$$

Combining equation (5.15) with (5.16) and (5.17):

$$C_t(k_u;T) \approx \frac{\exp(-\alpha k_u)}{\pi} \Re \left\{ \sum_{j=1}^{N} e^{-i\lambda\eta(j-1)(u-1)} e^{ibv_j} \Psi_{c_t}(v_j) \eta \right\},$$

for $u = 1, \ldots, N$ and incorporating Simpson's rule weightings into this equation:

$$C_t(k_u;T) \approx \frac{\exp(-\alpha k_u)}{\pi} \Re \left\{ \sum_{j=1}^{N} e^{-i\lambda\eta(j-1)(u-1)} e^{ibv_j} \Psi_{c_t}(v_j) \frac{\eta}{3} \left(3 + (-1)^j - \delta_{j-1}\right) \right\}, \tag{5.19}$$

for $u = 1, \ldots, N$, where δ_i is the Kronecker delta function[3], results in a form consistent with (5.14). All that is required, is to choose $\lambda = \frac{2\pi}{N\eta}$, finally:

$$C_t(k_u;T) = \frac{\exp(-\alpha k_u)}{\pi} \Re \left\{ \text{FFT} \left(e^{ibv_j} \Psi_{c_t}(v_j) \frac{\eta}{3} \left(3 + (-1)^j - \delta_{j-1}\right) \right) \right\}. \tag{5.20}$$

Unfortunately, due to the exponential nature of the strikes, a very low proportion of option prices will be computed in the desired zone. This means that either a large number of points must be used or η can be increased, reducing the accuracy of the integral.

[3] The Kronecker delta function, δ_i, is one for $i = 0$ and zero for all other values of i.

Chapter 6

Hedging

As discussed in Section 2.5.3, introducing jumps with random sizes to the model of the underlying asset, precludes exact hedging of option contracts (even without relaxing the assumption of continuous trading). The following caricature of a market illustrates this.

Example 6.1 (A simple incomplete market)**.** *Consider an asset S_t with a single jump at $t = 1$ after which it takes one of three states depicted in Figure 6.1. Also,*

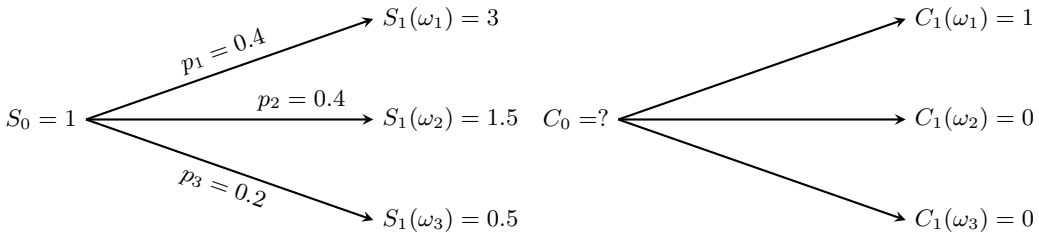

Fig. 6.1: A trinomial tree illustrating market incompleteness.

consider a digital option C_t written on S paying 1 in the uppermost state. Clearly, one cannot construct a portfolio comprising the underlying asset and a risk free asset that will reproduce the payoff in all states (to replicate the lower two states requires no holdings (in either asset), which is incompatible with replicating the upper state).

Now imagine the impossibility of finding a hedge in a realistic market, where the asset may take a continuum of values and jumps occur at unknown times[1]!

[1] In the case of GBM, the market will be complete despite the fact that the underlying asset can take a continuum of values. Heuristically, this can be explained by the fact that GBM may be

Since exact replication is impossible in such markets, options possess a residual risk that cannot be hedged away. The price of an option is no longer the cost of setting up a hedging portfolio — an additional premium (dependent on the investor's preferences) must be charged to compensate for unhedgeable risk. Depending on how this risk is viewed, different hedging strategies will be most efficient at mitigating it.

This chapter examines three methods for quantifying the residual risk and the hedging methods that result from each. All approaches in this dissertation are restricted to trading in the underlying asset and a risk-free account exclusively. All other instruments are assumed to be illiquid (as is often the case in the South African context), hence hedging with options [8] is not considered. It is assumed that the reader is acquainted with the material presented in Chapters 2 and 5. This chapter is principally based on Cont and Tankov [7] and Schweizer [41].

6.1 Super-Hedging

Super-hedging aims to find the cheapest portfolio that will almost surely exceed the option payoff. The cost of the option then corresponds to the total cost of setting up and maintaining the hedging strategy. Such prices are linear and preference free, dictating the choice of martingale measure — pricing occurs under the least favorable measure (for a particular option) that is absolutely continuous (see Section 4.2) with respect to the real world measure $\mathbb{P}$ [7][Proposition 10.1]. Under this measure, the maximum permissible probability is concentrated in the 'worst-case scenarios', typically leaving favorable scenarios with zero probability (in general, this 'super-hedging measure' is not equivalent to the real world measure and hence not arbitrage free).

In the contrived example, the problem reduces to a classic linear programming optimization: the objective function is simply the initial price of the replicating portfolio ($x_1 + x_2$, where x_1 and x_2 are the initial holdings in the stock and the risk-free asset respectively), while each state provides a linear constraint (e.g. $3x_1 + e^{rt}x_2 \geq 1$,

viewed as the limit of a random walk with two states (Section 2.4.1). Hence, if hedging occurs often enough (continuously), then the replicating portfolio will converge to the option value.

where $r < \log(3)$ is the continuously compounded risk-free rate). Alternately, the problem may be elegantly restated through a measure change: First the range of martingale measures that are absolutely continuous with respect to the real world measure is determined:

$$S_0 = e^{-r}\mathbb{E}^{\mathbb{Q}}\left[S_1\right]$$

for some martingale measure $\mathbb{Q}$.

$$\Rightarrow e^r = 3q_1 + 1.5q_2 + 0.5q_3$$

where q_1, q_2 and q_3 are the risk-neutral probabilities of each state. For valid probabilities:

$$\begin{aligned} q_1 + q_2 + q_3 = 1, \quad q_1, q_2, q_3 \geq 0 \\ \Rightarrow q_1 = \frac{1}{5}\left(2e^r - 1 - 2q_2\right) \end{aligned}$$

For a writer of the option, the uppermost state is the least favorable, so we must maximize q_1 by choosing $q_2 = 0$.

$$\Rightarrow C_0 = e^{-r}\mathbb{E}^{\mathbb{Q}}\left[C_1\right] = \frac{2 - e^{-r}}{5}$$

The hedging strategy then involves holding $\frac{2}{5}$ units of the underlying asset and $-\frac{e^{-r}}{5}$ units of the risk-free asset.

For realistic models, this approach is unrealistically conservative, yielding excessive bounds for the option price when a single stand-alone option (as opposed to a portfolio) is considered:

Proposition 6.2 (The Range of Option Prices)**.** *Let $S_t \equiv S_0 e^{X_t}$ be an exponential Lévy price process, where $(X_t)_{t\in[0,T]}$ is a purely discontinuous, infinite variation Lévy process with some positive jumps and the potential to produce negative jumps of any size (i.e. $\int_{-\infty}^{a} \nu_X(dx) > 0$, for all $a \in \mathbb{R}$ and $\int_{-1}^{0} |x|\, \nu_X(dx) = \int_0^1 x\, \nu_X(dx) = \infty$, where ν_X is the Lévy density of X_t). The range of arbitrage-free prices for a European call written on S with strike K and maturity T is* dense *on the interval:*

$$\left(\left(S_0 - Ke^{-rT}\right)^+, S_0\right).$$

Proof. See Eberlein and Jacod [14, Theorem 2] ■

Clearly, super-hedging regards *any* risk as unacceptable, effectively providing no-arbitrage bounds for trading in the option. For a writer of a European call, this entails buying and holding one unit of the underlying asset (with cost S_0). From a buyer's perspective, the option price must always be positive, since it offers an exclusively positive payoff. Furthermore, a forward, with an identical strike to the option (and initial cost $S_0 - Ke^{-rT}$), may be used to hedge the option, since its payoff is dominated.

In reality, however, very little trade would occur even close to these extremes. Rather, by partially hedging the contract (with favorable returns under expectation), market makers must take on some risk to provide attractive option prices. To this end, the next section examines methods for finding the optimum hedge depending on the investor's attitude to risk.

6.2 Utility Maximization

Super-hedging generates infeasible prices because it places undue weight on unlikely outcomes. In contrast, in determining the optimum hedging portfolio, utility maximization weights each scenario by severity and likelihood as viewed through the prism of an investor's preferences. In other words, it chooses from a set of payoffs $\mathcal{Z}$ based on the criteria:

$$\max_{Z \in \mathcal{Z}} \mathbb{E}^{\mathbb{P}} \left[U(Z) \right]. \tag{6.1}$$

The measure $\mathbb{P}$ may either be interpreted as an objective real-world measure, or the investor's view of events. The utility function $U(x)$ is now described.

6.2.1 Utility Functions

Utility functions measure the perceived value, attached by an investor, to a payoff. A utility function, usually denoted $U(x)$, is therefore a mapping from numeraire units x to a subjective 'value'. Based on econometric observation, realistic utility functions have the following properties:

- **Non-Satiation**: Utility functions must be monotonically increasing, since investors always prefer more wealth to less.

- **Continuity**: Utility functions should be continuous — a minute change in wealth should not result in a drastic change in utility.

- **Law of Diminishing Marginal Utility**: The first derivative ought to be monotonically decreasing. Intuitively, this is because an extra dollar is more valuable to a pauper than a millionaire (i.e. $U(x) - U(x - \Delta x) \geq U(x + \Delta x) - U(x)$). This leads to:

- **Risk-Aversion**: Where possible, economic agents seek to avoid risk — most people prefer a certain \$50 to the possibility of winning \$100 depending on the outcome of coin toss (i.e. $U(\mathbb{E}[x]) \geq \mathbb{E}[U(x)]$). This is only true for concave utility functions.

From the last point, is it clear that an agent's attitude to risk is embedded in the utility function.

Typical utility functions include the logarithmic utility $U(x;\alpha) = \log(\alpha x + 1)$, the exponential utility $U(x;\alpha) = 1 - e^{-\alpha x},\ \alpha > 0$ and the power utility $U(x;\gamma) = \frac{x^\gamma - 1}{\gamma},\ \gamma < 1$.

6.2.2 Utility Indifference Pricing

In economic theory, a *certainty equivalent*, $\pi_U(x, H)$, is the amount of cash that must be added to an agents initial wealth, x, in order to match the expected utility gained from an uncertain payoff, H, or:

$$U(x + \pi_U(x, H)) = \mathbb{E}[U(x + H)]. \tag{6.2}$$

The investor is then impartial to being paid $\pi_U(x, H)$ upfront or taking a risk with payoff H. For a risk-averse investor, the quantity $\pi_U(x, H)$ will incorporate compensation for the uncertainty of the payoff.

In the context of option valuation, an investor is afforded the opportunity to increase the expected utility of holding the option by trading in the underlying

asset: a-buy and-hold strategy may be foolhardy when one can hedge dynamically. Concordantly, we define

$$u(x,H) \equiv \sup_{\phi\in\mathcal{S}} \mathbb{E}^{\mathbb{P}}\left[U(x+H+\int_0^T \phi\, dS_t)\right] \tag{6.3}$$

as the best expected utility resulting from a portfolio in which an option, with payoff H and maturity T, is hedged using a self-financing strategy ϕ, from the space of all admissible strategies which may be approximated by simple predictable processes, $\mathcal{S}$ (see Section 2.1.3). Using this definition, Hodges and Neuberger [22] introduced the concept of a utility indifference price $\pi_U(x,H)$; that is, the price at which one is indifferent between buying the option and continuing to trade without it:

$$u(x,0) = u(x-\pi_U(x,H),H). \tag{6.4}$$

6.2.3 Example

Once again, we return to the toy example (Example 6.1). To apply utility indifference pricing, the maximum expected utility for the portfolio without the option must be found. Under the real world probabilities ($p_1 = p_2 = 0.4,\ p_3 = 0.2$) one may compute:

$$u(x,0) = \sup_{x_1+x_2=x} \mathbb{E}^{\mathbb{P}}\left[U(x_1S_1 + e^r x_2)\right] \tag{6.5}$$

from (6.3). As before, x_1 and x_2 give the holdings in the stock and the risk-free asset. For an exponential utility function with risk aversion parameter α:

$$u(x,0) = \sup_{x_1+x_2=x}\ 1 - \frac{1}{4}\left(e^{-\alpha(3x_1+e^r x_2)} + e^{-\alpha(1.5x_1+e^r x_2)} + 2e^{-\alpha(0.5x_1+e^r x_2)}\right).$$

The result is not dependent on the initial wealth [7][Section 10.3.3], so choose $x = 0$:

$$u(0,0) = \sup_{x_1}\ 1 - \frac{1}{4}\left(e^{-\alpha x_1(3-e^r)} + e^{-\alpha x_1(1.5-e^r)} + 2e^{-\alpha x_1(0.5-e^r)}\right). \tag{6.6}$$

Finding the maxima:

$$(3-e^r)e^{-\alpha\breve{x}_1(3-e^r)} + (1.5-e^r)e^{-\alpha\breve{x}_1(1.5-e^r)} + (1-2e^r)e^{-\alpha\breve{x}_1(0.5-e^r)} = 0 \tag{6.7}$$

which may be solved numerically to find the optimal stock holding without the option $\breve{x}_1$; and with it, the expected utility associated with such a portfolio. Next,

the optimal portfolio for a writer of the option is found:

$$u(\pi_U(C), -C) = \sup_{x_1+x_2=\pi_U(C)} \mathbb{E}^{\mathbb{P}}\left[U(x_1 S_1 + e^r x_2 - C_1)\right] \tag{6.8}$$

from (6.3).

$$\begin{aligned} u(\pi_U(C), -C) = \sup_{x_1+x_2=\pi_U(C)} & 1 - \frac{1}{4}\Big(e^{-\alpha(3x_1+e^r x_2-1)} + e^{-\alpha(1.5x_1+e^r x_2)} \\ & +2e^{-\alpha(0.5x_1+e^r x_2)}\Big) \\ = \sup_{x_1} & 1 - \frac{\exp\left(-\alpha e^r \pi_U(C)\right)}{4}\Big(e^{-\alpha((3-e^r)x_1-1)} + e^{-\alpha x_1(1.5-e^r)} \\ & +2e^{-\alpha x_1(0.5-e^r)}\Big). \end{aligned} \tag{6.9}$$

The optimal stock holding for the portfolio with the option, $\bar{x}_1$, satisfies:

$$(3-e^r)e^{-\alpha(\bar{x}_1(3-e^r)-1)} + (1.5-e^r)e^{-\alpha\bar{x}_1(1.5-e^r)} + (1-2e^r)e^{-\alpha\bar{x}_1(0.5-e^r)} = 0 \tag{6.10}$$

independent of $\pi_U(C)$. Using the optimal stock holding, it is now possible to compute the utility indifference price using (6.4) and the utility found using (6.6):

$$u(\pi_U(C), -C) = u(0, 0) \tag{6.11}$$

For $r = 0$, the optimal parameter sets for several values of α are tabulated in Table 6.1 and compared to super-hedging and the original option payoff. As $\alpha \to \infty$, the strategy matches super-hedging. This may be explained as follows: since α provides a measure of risk aversion (it corresponds to the coefficient of absolute risk aversion $-U''(x)/U'(x)$); when it reaches infinity, a strategy in which no risk is acceptable results. For low values of α, the investor will have large existing positions in the underlying asset ($\breve{x}_1$), with very little of the portfolio attributable to the option. This indicates that utility indifference pricing is more concerned with optimal portfolio allocation for profit than reproducing the payoff! Indeed, the terminal value of the 'hedging' portfolio very seldom approaches the option payoff.

6.2.4 The Applicability of Indifference Pricing

Several factors call into question the usefulness and amenability of indifference pricing. Counter-intuitively, certainty equivalence and utility indifference results in a

Method	α	**Price**	$\breve{x}_1$	$\bar{x}_1$	$\phi(\omega_1)$	$\phi(\omega_2)$	$\phi(\omega_3)$
Utility Max.	0.01	0.059	120.0	120.3	240.6	60.21	-60.09
	0.1	0.060	12.00	12.26	24.57	6.188	-6.068
	1	0.072	1.200	1.476	3.024	0.810	-0.666
	5	0.119	0.240	0.576	1.271	0.407	-0.169
	10	0.152	0.120	0.484	1.120	0.394	-0.090
	100	0.195	0.012	0.408	1.012	0.399	-0.009
Super-Hedging	∞	0.2	0	0.4	1	0.4	0
Quadratic	—	$-\frac{1}{3}$	0	$\frac{2}{3}$	1	0	$-\frac{2}{3}$

Tab. 6.1: A comparison of hedging strategies in the simple three state market of Example 6.1. The value of the hedging portfolio in a state ω is given by $\phi(\omega)$, the rest of the symbols are defined in the text. The example was constructed to illustrate that the variance-optimal initial capital may be negative (see [23] for a further example), emphasizing many of the points made in Section 5.1.

valuation method that is non-linear — two option contracts are not twice the value of one, or in general

$$\pi_U(x, aH_1 + bH_2) \neq a\pi_U(x, H_1) + b\pi_U(x, H_2), \ a, b \in \mathbb{R}. \tag{6.12}$$

Furthermore, unless exponential utility is used, the value of an option contract is dependent on the investor's initial wealth.

In addition, as illustrated by the example, embedded in the utility maximization approach is a complicated portfolio optimization problem, which requires specification of all relevant joint and individual asset dynamics under the real-world measure $\mathbb{P}$.

Moreover, under exponential utility, the α parameter merely interpolates between the super-hedging cost and the price under the minimum entropy measure (to be discussed in 6.3.3)[2, Proposition 3.2]. This calls into question the validity of this approach for pricing — effectively, one arbitrarily chooses a price, from a broad range, based on a sensitive unobservable parameter α, for which one then

computes an associated optimal hedging strategy. To circumvent these difficulties, a preference free method of approximating the option payoff as closely as possible is now presented.

6.3 Quadratic Hedging

As suggested, quadratic hedging results in a linear, preference free method for option valuation. Quadratic hedging may be categorized into two varieties: local and global risk minimization. Typically, the global strategy is difficult to compute but has a tidy economic interpretation; while the local strategy is easier to find but may result in undesirable intermediate cashflows. Once again, a broad overview is provided, omitting most technical proofs (which may be found in Schweizer [41]).

6.3.1 Local Risk Minimization

A first approach, is to relax the self-financing constraint in order to replicate the option payoff exactly. The simplest example of such a strategy is to 'wait and pay'; this involves a hedging portfolio with no holdings until maturity, at which point the payoff is matched by a holding in cash. Clearly, such an approach is highly risky, since one is liable for the full unhedged payoff at maturity. Rather, the 'best' strategy in this context should also minimize the variability of the cash-flows needed to maintain itself. Such a strategy is found by minimizing a *risk process* defined by

$$R_t(\phi) \equiv \mathbb{E}^{\mathbb{P}}\left[(C_T(\phi) - C_t(\phi))^2 \middle| \mathcal{F}_t \right], \tag{6.13}$$

in continuous time, where the cumulative *cost process* is defined as

$$C_t(\phi) \equiv V_t(\phi) - \int_0^t \nu_t \, d\hat{S}_t. \tag{6.14}$$

where $\hat{S}_t$ is the discounted asset price. Here the *value process* is given by

$$V_t(\phi) = \nu_t S_t + \eta_t, \tag{6.15}$$

where ϕ is an admissible trading strategy with holdings η_t and ν_t in the risk-free asset and the underlying asset respectively at time t.

The cost process is the cumulative amount of money injected into the strategy up until time t, including the initial cost.

To illustrate the central properties of the local risk minimization strategy without getting distracted by the technicalities arising from continuity, the analysis now proceeds in the discrete case, closely following Schweizer [41, Section 3]. Considering discrete trading times $k = 0, 1, \ldots, N \in \mathbb{N}$, the incremental cost of the strategy ϕ is given by:

$$C_{k+1}(\phi) - C_k(\phi) = V_{k+1}(\phi) - V_k(\phi) - \nu_k \left(S_{k+1} - S_k\right)$$

from (6.14) and (2.5). Consequently,

$$\Delta C_{k+1}(\phi) = \Delta V_{k+1} - \nu_k \Delta S_{k+1}, \tag{6.16}$$

where the Delta operator Δ is defined as $\Delta U_k \equiv U_k - U_{k-1}$ for any stochastic process $(U_k)_{k\in[0,1,\ldots,N]}$. Now consider the incremental risk process

$$\begin{aligned}\mathbb{E}\left[\left(\Delta C_{k+1}(\phi)\right)^2 \middle| \mathcal{F}_k\right] =& \mathrm{VAR}\left[\Delta V_{k+1}(\phi) - \nu_k \Delta S_{k+1} \middle| \mathcal{F}_k\right] \\ &+ \mathbb{E}\left[\Delta V_{k+1}(\phi) - \nu_k \Delta S_{k+1} \middle| \mathcal{F}_k\right]^2,\end{aligned}$$

from the definition of variance and Equation (6.16). When conditioning is applied:

$$\begin{aligned}\mathbb{E}\left[\left(\Delta C_{k+1}(\phi)\right)^2 \middle| \mathcal{F}_k\right] =& \mathrm{VAR}\left[V_{k+1}(\phi) - \nu_k \Delta S_{k+1} \middle| \mathcal{F}_k\right] \\ &+ \left(\mathbb{E}\left[V_{k+1}(\phi) - \nu_k \Delta S_{k+1} \middle| \mathcal{F}_k\right] - V_k(\phi)\right)^2. \end{aligned} \tag{6.17}$$

The first term does not depend on η_k, hence we are free to chose η_k at each time such that

$$V_k(\phi) = \mathbb{E}\left[V_{k+1}(\phi) - \nu_k \Delta S_{k+1} \middle| \mathcal{F}_k\right] \tag{6.18}$$

in order to minimize the local risk $\mathbb{E}\left[\left(\Delta C_{k+1}(\phi)\right)^2 \middle| \mathcal{F}_k\right]$. In this case

$$\mathbb{E}\left[\Delta C_{k+1}(\phi) \middle| \mathcal{F}_k\right] = 0 \tag{6.19}$$

$$\Rightarrow \mathbb{E}\left[C_{k+1}(\phi) \middle| \mathcal{F}_k\right] = C_k(\phi), \tag{6.20}$$

implying (if one ignores integrability requirements) that the cost process C_t of the local risk minimizing hedge must be a $\mathbb{P}$-martingale (and hence self-financing under

expectation, since self-financing strategies have constant cost and $C_0 = \mathbb{E}[C_t]$ for any martingale C_t).

It only remains to minimize the variance term of (6.17), using an appropriate choice of ν_k. Schweizer [41] shows that this occurs when the cost process is orthogonal[2] to the martingale component M_t (in the Doob-Meyer decomposition) of the price process S_t.

Under several technical assumptions beyond the scope of this document (such as a structure condition outlined in [41, Section 3]), the above results may be generalized to the continuous case. Moreover, in this case, the optimal local risk minimizing hedge ξ_t satisfies

$$H = H_0 + \int_0^T \xi_t^H \, d\hat{S}_t + L_T^H \qquad \mathbb{P}\text{-a.s.}, \tag{6.21}$$

for a contingent claim with payoff H, where L_T^H is a $\mathbb{P}$-martingale orthogonal to M_t [41, Proposition 3.5]. This *Föllmer-Schweizer decomposition* has the following interpretation: an optimal trading strategy, with initial capital H_0 and holdings ξ_t in the underlying asset at each time t, may be used to reproduce the payoff H such that the random error of the scheme, L_T^H, is unhedgeable (orthogonal to the space of attainable claims produced by every valid choice of ξ). When $L_T^H = 0$, the claim H is termed *attainable* since it may be replicated almost surely. However, for realistic models of the underlying asset, this only occurs if H is affine (i.e. a combination of forward contracts)[7, pg. 338]. The above decomposition may be viewed as a more general semi-martingale statement of the classical Galtchouk-Kunita-Watanabe decomposition [41].

6.3.2 Global Risk Minimization

Mean-variance hedging (or global risk minimization) provides a self-financing strategy that reproduces the option payoff with minimum expected terminal error. In other words, it punishes any deviation from the payoff, profit of loss, finding a hedging strategy ϕ that satisfies:

$$\inf_{\phi \in \mathcal{S}} \mathbb{E}^{\mathbb{P}} \left[|\hat{V}_T(\phi) - \hat{H}|^2 \right], \tag{6.22}$$

[2] Two random variables X and Y are said to be orthogonal if $\mathbb{E}[XY] = 0$

where $\hat{H}$ is the discounted option payoff and $\hat{V}_T(\phi)$ is the terminal value of the trading strategy ϕ. For a self-financing strategy, the value process obeys [41, Definition (1.1)]:

$$\hat{V}_t(\phi) = V_0 + \int_0^t \phi_u \, d\hat{S}_u, \qquad \phi \in \mathcal{S} \tag{6.23}$$

where $\hat{S}_u$ is the discounted price process of the underlying asset. The set of admissible strategies $\mathcal{S}$ is discussed in Section 2.1.3.

It should be apparent from Section 2.5.3 that, for realistic models, the absence of a (first-order) predictable representation implies that it is impossible to find a self-financing strategy equating $\hat{V}_T(\phi)$ and $\hat{H}$ (which would naturally solve the minimization problem (6.22)), unless H may be expressed as a combination of forward contracts.

Mean-variance hedging is equivalent to hedging using the 'utility function' $U(x) = -|x|^2$. However, since investors are averse to profits resulting from the hedge, this utility function does not satisfy all of the axioms discussed in Section 6.2.1. In this context, as in market practice, risk is defined as the variance of the hedged payoff rather than just the potential for losses.

The Method of Hubalek, Kallsen and Krawczyk [23]

Typically, the mean-variance strategy is found by applying a path dependent feedback adjustment to the local risk minimizing strategy (given in [41, Theorem 4.6]). Often an explicit computation of this strategy is difficult find; however, Hubalek, Kallsen and Krawczyk [23] were able find the to the Föllmer-Schweizer decomposition and with it compute the following discrete-time[3] variance-optimal hedge for Lévy processes (the method was later generalized [27] to processes with non-stationary independent increments, including models such as the Heston model):

Theorem 6.3 (Variance Optimal Hedging for Exponential Lévy Processes)**.** *For a*

[3] A continuous-time expression was also derived in the same paper. However, the discrete expression is more relevant to the hedging races conducted in Chapter 7, where hedging occurs at discrete time points.

contingent claim, with terminal payoff which may be expressed as

$$H_N = \int_{R-i\infty}^{R+i\infty} \hat{S}_N^z \Pi(z) dz \tag{6.24}$$

for some complex density $\Pi(z)$,[4] *with* $R \in \mathbb{R}$, *written on a discounted underlying asset* $\hat{S}_n \equiv S_0 e^{X_n}$, *where* $(X_n)_{n=0,1,\ldots,N}$ *is a Lévy process, viewed on equally spaced time points* $0 = t_0 < t_1 < \cdots < t_N = T$; *the variance optimal hedge is given by the following recursive expression:*

$$\phi_n = \xi_n + \frac{\lambda}{\hat{S}_n} \left(H_n - V_0 - G_n(\phi) \right), \tag{6.25}$$

for some initial capital V_0 *(*$V_0 = H_0$ *for optimality) and*

$$\begin{aligned}
m(u) &\equiv \mathbb{E}[e^{uX_1}], \\
g(z) &\equiv \frac{m(z+1) - m(1)m(z)}{m(2) - m(1)^2}, \\
h(z) &\equiv m(z) - (m(1) - 1)\, g(z), \\
\lambda &\equiv \frac{m(1) - 1}{m(2) - 2m(1) + 1}, && (6.26) \\
H_n &\equiv \int_{R-i\infty}^{R+i\infty} \hat{S}_n^z h(z)^{N-n} \Pi(z)\, dz, && (6.27) \\
\xi_n &\equiv \int_{R-i\infty}^{R+i\infty} \hat{S}_n^{z-1} g(z) h(z)^{N-n+1} \Pi(z)\, dz, && (6.28) \\
G_n(\phi) &= G_{n-1}(\phi) + \phi_{n-1} \left(\hat{S}_n - \hat{S}_{n-1} \right), \quad G_0(\phi) = 0. && (6.29)
\end{aligned}$$

Proof. See Hubalek, Kallsen and Krawczyk [23, Section 2.] ■

Note that ξ is the local risk minimizing strategy.

Unfortunately, this method involves expressing the option payoff as an inverse Mellin transform, which may be difficult to obtain in general. However, functional forms for the complex density, $\Pi(z)$, are generously provided in [23, Section 4] for common options. Of interest,

$$\Pi(z) = \frac{1}{2\pi i} \frac{K^{1-z}}{z(z-1)}, \tag{6.30}$$

[4] Complex measures are not rigorously treated in this text; however, it will suffice to think of the density as a function with a complex argument z in this context.

for vanilla puts and calls with strike K, with some real-valued $R > 1$ for a call and $R < 0$ for a put in (6.24).

Clearly, the method considers information from the whole past stock path when selecting the best hedge. If the replication is working perfectly (i.e. the profit from the hedge portfolio ($V_0 + \sum_{i=1}^{n} \phi_n \left(\hat{S}_n - \hat{S}_{n-1}\right)$) matches the current optimal capital[5] (H_n) exactly), then the hedge is given by ξ_n. When the hedge performs better than the option, the local risk minimizing hedge is relaxed proportional to $\lambda/\hat{S}_{n-1}$ in an attempt to steer the hedge closer to the option value. Similarly, when the replicating portfolio falls behind, the strategy 'over-hedges' to try catch up.

6.3.3 The Method of Cont and Tankov

In a novel approach, Cont and Tankov [7, Section 10.4] apply the quadratic risk minimizing criteria to the process dynamics *under a martingale measure*, in effect determining:

$$\inf_{\phi \in \mathcal{S}} \mathbb{E}^{\mathbb{Q}} \left[|\hat{V}_T(\phi) - \hat{H}|^2 \right], \tag{6.31}$$

for some equivalent martingale measure $\mathbb{Q}$.

Theorem 6.4. *For any contingent claim H with maturity T, satisfying the Lipschitz property:*

$$|H(x) - H(y)| \leq L\,|x - y|\,, \tag{6.32}$$

for some $L > 0$; written on an underlying asset that follows an exponential Lévy process $S_t = S_0 e^{rt + X_t}$, where X_t is a Lévy process and r is the continuously compounded risk-free rate; the risk minimizing holding in the underlying asset, under the equivalent martingale-measure $\mathbb{Q}$, is given by:

$$\phi_t(S) = \frac{\sigma^2 \frac{\partial C}{\partial S}(t,S) + \frac{1}{S}\int_{\mathbb{R}} (e^x - 1)\left[C(t, Se^x) - C(t,S)\right] \nu_X(x)\,dx}{\sigma^2 + \int (e^x - 1)^2\, \nu_X(x)\,dx} \tag{6.33}$$

where $C(t,S) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[H(S_T)\,|S_t = S]$ is the 'price' at time t, given $S_t = S$, $\nu_X(x)$ is the Lévy density of X and σ is standard deviation of the Brownian component of X_t.

[5] This quantity should not be confused with a price, since, in general, hedging at this value will have a high probability of incurring a loss and, as noted in Example 6.1, may even be negative.

Proof. See Cont and Tankov [7, Proposition 10.5]. ■

It is evident that, in the case of a pure jump process, the 'ideal' holding in the underlying asset is a normalized aggregation of all the price changes for a given jump size and likelihood, rather than an attempt to match price changes for infinitesimal movements in the underlying asset only, using $\partial C/\partial S$ from a geometric Brownian motion framework (where prices evolve incrementally through minute movements. It will be shown in Section 7 that this naive 'delta' approach is a poor at hedging the option in the presence of larger jumps).

The above expression (6.33) is slightly less complex than the approach of Hubalek *et al.* However, as the authors note, minimizing the error under an arbitrary martingale measure may be inadequate — market participants are concerned with the distribution of profits and losses *in reality*, not under another measure that provides a pricing rule. Indeed, outcomes with small weights under the equivalent martingale measure may be likely in reality, for instance, in Example 6.1, $q_1 < 0.2$ for $r = 0$, but $p_1 = 0.4$! As a result, unless the underlying process is a $\mathbb{P}$-martingale, for this approach to yield reasonable results, the martingale measure under which it is applied must be chosen carefully.

Choosing a Measure

Apart from pointing out this shortcoming and showing that the global and local risk minimization[6] strategies may be difficult to compute (or may not even exist), Cont and Tankov provide little guidance in constructing an appropriate measure for this strategy. However, a later discussion on 'optimal measures' [7, Section 10.5] is elucidating — barring a better alternative, for the minimization under the martingale measure to closely reflect the real-world outcome, one should choose the equivalent martingale measure that is 'most similar' to the real world measure. How this 'similarity' is best described in the context of the problem at hand, is now explored.

[6] The method of Cont and Tankov will yield a locally risk minimizing strategy, if an EMM that preserves the orthogonality of the error term in Equation (6.21) is used.

The concept of relative entropy (quantified using metrics such as Kullback-Leibner distance[7]) provides a measure of the 'informational' distance between two measures. A low-relative entropy implies that both measures assign similar weights to the same events. In cognizance of this, the Cont and Tankov method may be applied under the *minimal entropy measure* (the equivalent martingale measure with lowest entropy with respect to the real world measure), where it exists (for examples where it does not exist see [16, Example 3.3]), in the hope that minimizing the error under this measure will be most similar to a minimization under the real-world measure.

In the case of an exponential Lévy process, the desired measure change may be determined explicitly using the following theorem [16]:

Theorem 6.5 (The Minimum Entropy Measure for Exponential Levy Processes)**.** *Let $S_t = S_0 e^{X_t}$ be a one dimensional exponential (or geometric) Lévy process, where $(X_t)_{t\in[0,T]}$ is pure jump Lévy process with characteristic triplet $(0, \nu(dx), \gamma)$. The minimum entropy measure, characterized by its Radon-Nikodym derivative[8] with respect to the real world measure, is given by the following Esscher Transform:*

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \frac{e^{b^\star \hat{X}_t}}{\mathbb{E}^{\mathbb{P}}\left[e^{b^\star \hat{X}_t}\right]}, \tag{6.34}$$

where $\hat{X}_t$ is a Lévy process such that $\mathcal{E}(\hat{X})_t = \exp(X_t)$,[9] while the constant $b^\star$

[7] For two measures $\mathbb{P}$ and $\mathbb{Q}$ on the same space, the Kullback-Leibner distance between them is defined as:

$$H(\mathbb{Q},\mathbb{P}) = \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\log\frac{d\mathbb{Q}}{d\mathbb{P}}\right],$$

if $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ and ∞ otherwise.

[8] For a measure $\mathbb{Q}$ that is absolutely continuous with respect to another measure $\mathbb{P}$ on the same space, the Radon-Nikodym derivative $\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)$ is unique random variable, allowing translation from one measure to the other and is chosen such that:

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\mathbb{1}_A\right]$$

for any measurable set A [28, Theorem 10.6].

[9] The stochastic (or Doléans-Dade) exponential of a pure jump process X_t, denoted $\mathcal{E}(X)_t$ is given by:

$$\mathcal{E}(X)_t = e^{X_t - X_0}\prod_{0\le s\le t}(1+\Delta X_s)\,e^{-\Delta X_s},$$

satisfies

$$\int_{|x|>1} e^x e^{b^\star(e^x-1)}\nu(dx) < \infty \tag{6.35}$$

and

$$\gamma + \int \left\{ \left(e^x - 1\right) e^{b^\star(e^x-1)} - x\mathbb{1}_{|x|\leq 1} \right\} \nu(dx) = r, \tag{6.36}$$

where r is the risk-free rate. Under this measure change, the dynamics of X_t may be characterized by the new Lévy triplet:

$$\left(0, e^{b^\star(e^x-1)}\nu(dx), \gamma + \int_{|x|\leq 1} x\left(e^{b^\star(e^x-1)} - 1\right)\nu(dx)\right).$$

Proof. See Fujiwara and Miyahara [16]. ■

Obviously, this Esscher transform modifies the Lévy density of the process. As a consequence, the characteristic function, central to most pricing techniques, may become difficult to compute analytically, eliminating much of the tractability of the framework. For instance, if the original process was a variance gamma process, it may no longer be characterized as such under the minimum entropy measure.

A computationally simpler heuristic approach (common to other applications such as pricing [5, Section 5.] and estimating risk neutral densities [4]) constructs a martingale price process by directly adjusting the drift. Observe that the exponential Lévy process $\frac{e^{X_t}}{\mathbb{E}[e^{X_t}]}$ is a martingale, if $\mathbb{E}[e^{X_t}] < \infty$:

$$\begin{aligned}\mathbb{E}\left[\frac{e^{X_t}}{\mathbb{E}[e^{X_t}]}\middle|\mathcal{F}_s\right] &= \mathbb{E}\left[\frac{e^{X_s}e^{X_t-X_s}}{\mathbb{E}[e^{X_s}]\mathbb{E}[e^{X_t-X_s}]}\middle|\mathcal{F}_s\right], \quad t > s > 0\\ &= \frac{e^{X_s}\mathbb{E}[e^{X_t-X_s}]}{\mathbb{E}[e^{X_s}]\mathbb{E}[e^{X_t-X_s}]} = \frac{e^{X_s}}{\mathbb{E}[e^{X_s}]},\end{aligned}$$

since X_s is adapted and $X_t - X_s$ is independent of $\mathcal{F}_s$. In this case, the structure of the jumps and the characteristic function are preserved (up to a factor, see Equation 5.11). Provided this shift in the probability distribution of the exponent of the underlying asset is small enough, the resulting probability distribution should be similar to the original measure (and hence should have a low entropy).

which is the solution ($Z_t = \mathcal{E}(X)_t$) of the stochastic differential equation $dZ_t = Z_t dX_t$ [28, Theorem 8.33].

In instances where the real world drift is roughly similar to the 'drift' used to construct a martingale[10]; this approach is applied with great effect, producing a similar variance in the final payoff to the global optimum given by the method of Hubalek *et al.* [23] (see Chapter 7).

[10] One would expect the 'corrected' process drift and the risk-free rate to be fairly similar in reality, since excessive imbalances between them would lead to inefficiencies that allow for profitable trading strategies in the underlying asset with little risk of a loss.

Chapter 7

Numerical Results

This chapter, gauges the performance of the aforementioned quadratic hedging methods and compares them to benchmark hedging strategies such as the Black-Scholes methodology and naive 'Delta' hedging.

First, the structure of the optimal holdings of each method is compared, providing insight into the differences between each method. Then a numerical 'hedging race' is conducted whereby realistic exponential CGMY stock price paths are generated, using the algorithm described in Section 4.2 for a wide range of parameter values and stochastic properties. Each method is then applied to hedge a vanilla European call written on these paths on a fixed time grid. The resulting profit and loss distributions for each of the hedging methods is then computed for each Monte-Carlo sample. Relevant metrics such as the expectation, variance and 95%-VAR (over the lifetime of the option) are then estimated to assess the performance of each method. The implementation is discussed in greater detail in the next section.

7.1 Method

7.1.1 Parameters

A CGMY process (see Section 3.2) was used in all simulations. The following parameters were assumed to have little impact and for simplicity remain constant in all simulations unless otherwise stated: the risk-free rate $r = 0$, the maturity of the option $T = 1$ represents a ten day option (as discussed in Section 3.1.3, despite the static maturity, the C parameter may be used to reproduce behavior over longer

time horizons), the truncation level of the CGMY subordinator $\epsilon = 10^{-4}$ (see Section 4.2), the standard initial stock price $S_0 = 10$, the strike of the standard option $K = 10$, the standard 'excess corrected drift' from (5.2) $\mu = 10\%$ anually[1], the number of paths $N = 2.5 \times 10^4$, the number of portfolio rebalancings over the life-span of the option $p = 500$.

For a representative Monte-Carlo sample, it is infeasible to recalculate the optimal holdings at each time point for each path. As a result, the optimal holdings (as well as ξ and H for the method of Hubalek *et al.* (HKK)) were calculated on a time and stock price grid with 75 by 250 points on the interval $[0, T] \times [0, 4K]$. The desired holdings for each path at a given time were then calculated using a two-dimensional linear interpolation, since each surface is smooth. The following section describes the techniques used to generate specific sets of optimal holdings.

7.1.2 Calculating Optimal Holdings

As described in Section 6.3.3, the Cont and Tankov (CT) optimal holdings in the underlying asset were calculated under the equivalent martingale measure that preserves the jump structure (using (5.2) with $\mu = r$).

All strategies were provided with an initial endowment equal to the price under this measure. However, the HKK strategy was permitted to use the variance optimal initial holding for comparison.

For lack of a better measure choice, the 'naive Delta' hedging strategy was calculated by partially differentiating prices under this measure since the initial endowment used by the strategy arises from this measure and an abundance of option prices are available through the FFT method (leading to higher numerical accuracy and smoothness of this derivative).

To achieve a fine grid for the Cont and Tankov strategy within reasonable time constraints, the fast-Fourier transform method described in Section 5.2 was applied (with $N = 4 \times 10^6$ points and a grid spacing of $\eta = 0.005$) to expeditiously generate the prices needed to evaluate the upper integral in (6.33). However, the FFT

[1] This may be considered large and was chosen to illustrate the robustness of the Cont and Tankov method.

method produces a vector of prices varying in strike for a given stock price, hence the following relation was used to convert this vector into the desired form required by (6.33):

$$\begin{aligned} C(t, K_i|S_0) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_0 e^{X_T} - K_i)^+ \big| \mathcal{F}_t], \\ &= \frac{K_i}{K} e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_0^i e^{X_T} - K)^+ \big| \mathcal{F}_t], \end{aligned}$$

where i indexes each element in the vector, with

$$\begin{aligned} S_0^i &\equiv \frac{K}{K_i} S_0. \\ \Rightarrow C(t, S_0^i|K) &= \frac{K}{K_i} C(t, K_i|S_0). \end{aligned}$$

The integrals in (6.33) were found to be numerically unstable in the region of zero for processes of infinite variation (such as the CGMY process with $1 < Y < 2$). Consequently, in this calculation, the small jumps were truncated at a level of $\tilde{\epsilon} = 10^{-4}$ and approximated by their drift (in a similar vein to Section 4.2.4).

7.2 Results and Analysis

7.2.1 Scenarios

A number of scenarios (summarized in the first two columns of Table 7.1) were chosen to capture the essence of hedging when taking into account jumps. As one would expect, the benefits of hedging when considering jumps become more apparent as the large jumps become more pronounced relative to small movements (typically on short time-scales or with unusual price dynamics). An in depth comparison of the various methods is now presented, with an accompanying analysis of each scenario.

7.2.2 Comparison of Optimal Holdings and Prices

Studying Figures 7.1 and 7.2, it is evident that prices under the structure preserving measure are almost identical to the HKK optimal capital in almost any state (except at the money near maturity, where the HKK method utilizes slightly more money). This implies that the minimum variance measure (which generates the optimal initial

	Params		Method	P_0	Mean	Variance	95%-VAR
Base Case	C	0.41	CT	0.1698	-0.002	0.0058	-0.1353
	G	52	HKK	0.1649	-0.007	0.0058	-0.1389
	M	54.2	B-S	0.1698	-0.002	0.0060	-0.1405
	Y	0.5	Naive	0.1698	-0.002	0.0068	-0.1506
Large Jumps	C	0.41	CT	0.5557	0.000	0.1151	-0.5793
	G	11	HKK	0.5526	-0.003	0.1151	-0.5808
	M	10	B-S	0.5557	0.000	0.1301	-0.6188
	Y	0.5	Naive	0.5557	-0.001	0.1423	-0.6678
Out-of-the-Money ($S_0 = 9$)	C	0.41	CT	0.1942	0.002	0.0820	-0.4272
	G	11	HKK	0.1918	0.000	0.0820	-0.4295
	M	10	B-S	0.1942	0.024	0.0974	-0.5038
	Y	0.5	Naive	0.1942	0.024	0.1081	-0.5534
Long Maturity ($T = 5$)	C	0.41	CT	0.3858	-0.008	0.0072	-0.1598
	G	52	HKK	0.3828	-0.012	0.0072	-0.1608
	M	54.2	B-S	0.3858	-0.009	0.0073	-0.1627
	Y	0.5	Naive	0.3858	-0.009	0.0077	-0.1673
Infinite Variation	C	0.41	CT	0.7772	-0.076	0.0070	-0.2309
	G	52	HKK	0.7934	-0.060	0.0068	-0.2090
	M	54.2	B-S	0.7772	-0.076	0.0070	-0.2288
	Y	1.2	Naive	0.7772	-0.076	0.0071	-0.2312

Tab. 7.1: A comparison of the hedging strategies in several scenarios with parameters given in Section 7.1.1. The initial capital for each method is denoted P_0. All statistical quantities apply to the profit and loss distribution.

capital needed for mean-variance hedging) has a fairly low relative entropy with respect to the structure preserving measure, in states where the option is in the money, in this instance.

Each method has sensible limits when hedging a vanilla option — holding one unit of the underlying when the option is far in-the-money and holds nothing when it is far out-the-money, with this transition becoming more extreme near maturity

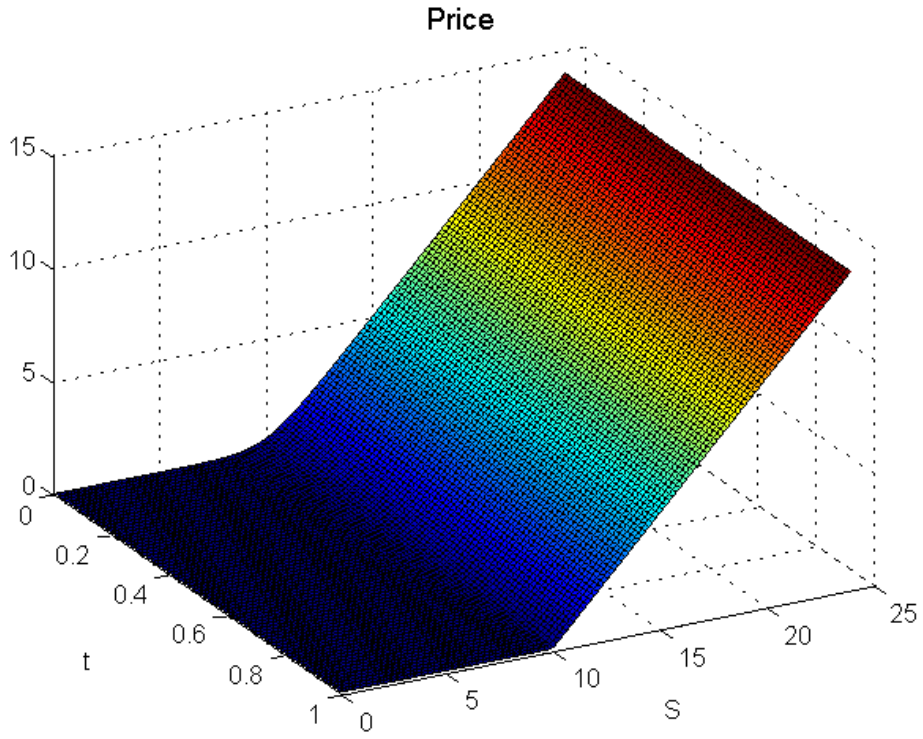

Fig. 7.1: The 'price' under the structure preserving measure for the 'large jumps' scenario.

(as illustrated by Figure 7.3). However, the significant differences, apparent near the at-the-money value, will now be explored.

Noting the scale of Figure 7.4, it is clear that there is little difference between the local risk minimizing component of the HKK hedge (ξ) and the CT method for the standard drift (apart from the numerical artifacts introduced by the oscillatory nature of the integrand in (6.28)), suggesting a similar performance for each method.

However, the benchmark 'delta' methods over-hedge in-the-money calls (underestimating the likelihood of jumping out-of-the-money) while failing to hedge out-of-the-money calls enough (understating the probability of a move into the money), in a similar fashion (see Figures 7.5 and 7.6). This effect is clearly exacerbated around the at-the-money point near maturity, where jumps have a greater effect on the price, and tapers off over time as the process decays to normality.

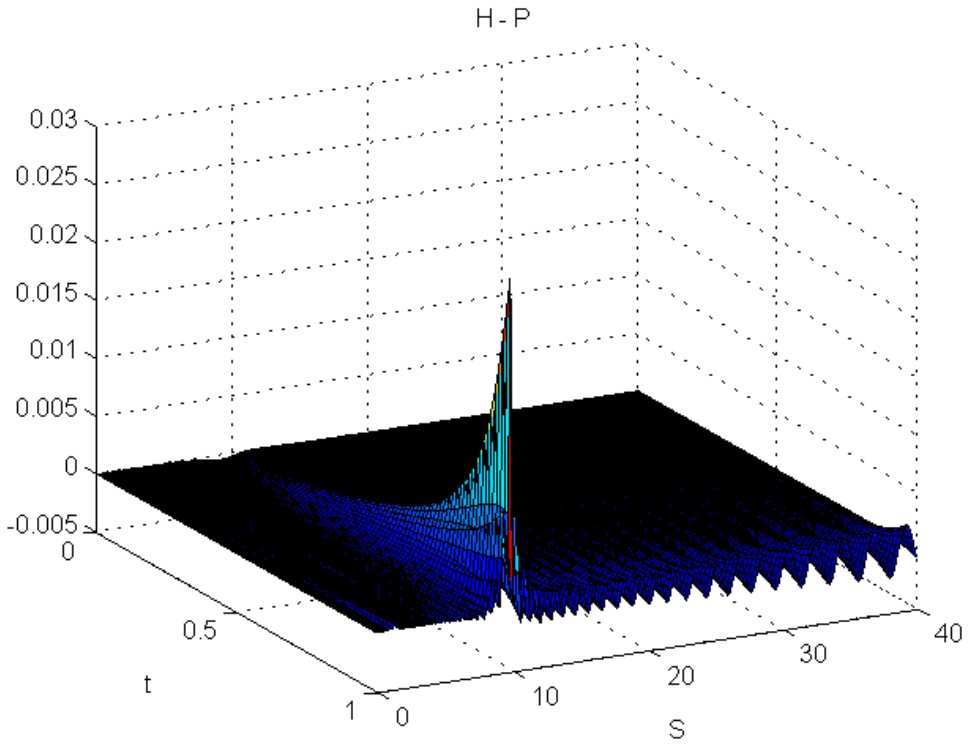

Fig. 7.2: The difference between the HKK optimal capital H and the 'price' under the structure preserving measure P for different values of the underlying asset and times to maturity in the 'large jumps' scenario.

7.2.3 Base Case

The following figures (Figures 7.7 and 7.8) serve as a base case against which the other parameter sets are compared. The parameters were chosen such that they roughly match the moments of actual daily data from the S&P500 given in [40, Table 4.1], with an annualized volatility of $\sigma = 21.68\%$ (assuming 250 business days in a year), daily return skewness of $S = -0.1067$ and daily excess kurtosis of $\kappa = 7.0940$.

As discussed, the optimal initial capital is clearly not a price, since when hedging at this price, one will break even under expectation. Rather, a premium, dependent on the market makers attitude to risk, must be charged to ensure profitability on average. It is for this reason, that it may be better to calibrate the hedging model to

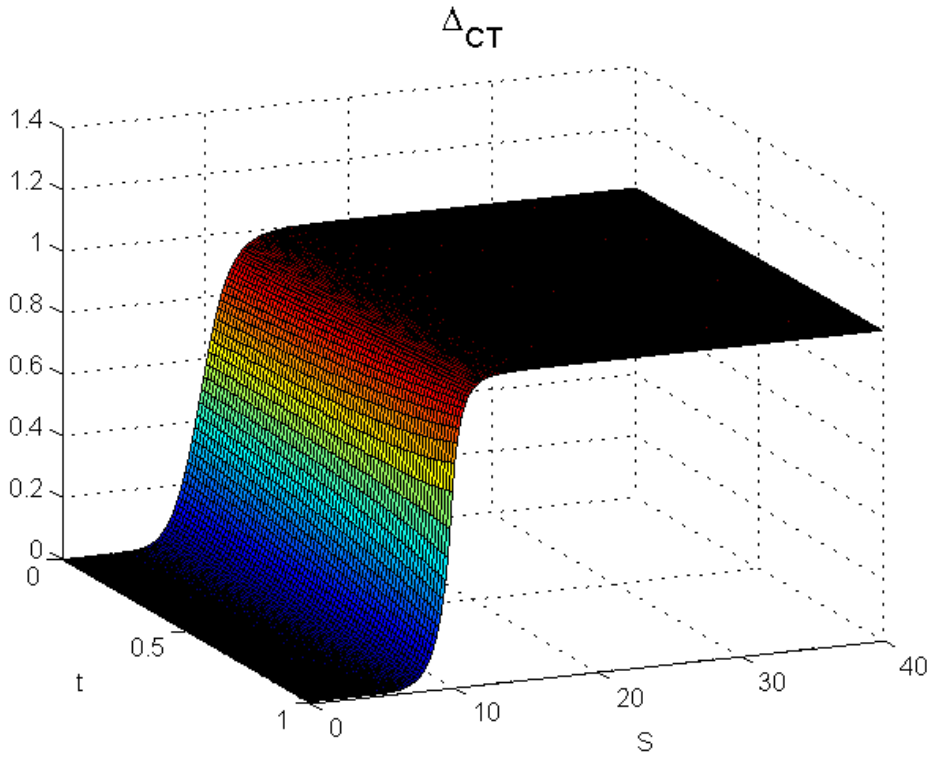

Fig. 7.3: The optimal holdings in the underlying asset according to the CT method in the 'Large Jumps' scenario, for different values of the underlying asset and time to maturity ($T = 1$).

the observed return series of the underlying asset. In general, calibration to market option prices will result in a hedging approach that is suboptimal, since such prices arise from market makers preferences not real-world price dynamics[2] and all the hedging methods under consideration are sensitive to the real-world specification of the price process (not the risk-neutral measure chosen by the market). The drawback, however, is that such a calibration would assume stationarity (that the model parameters do not change with time). Thankfully, this is not a terrible assumption, since large 'regime changes' in price dynamics are unlikely over short option lifespans.

[2] Market makers typically sell an excess of out-of-the-money put protection to pension funds, leading to an unrealistic volatility skew for low strikes not evident in actual prices dynamics [4, Section VI].

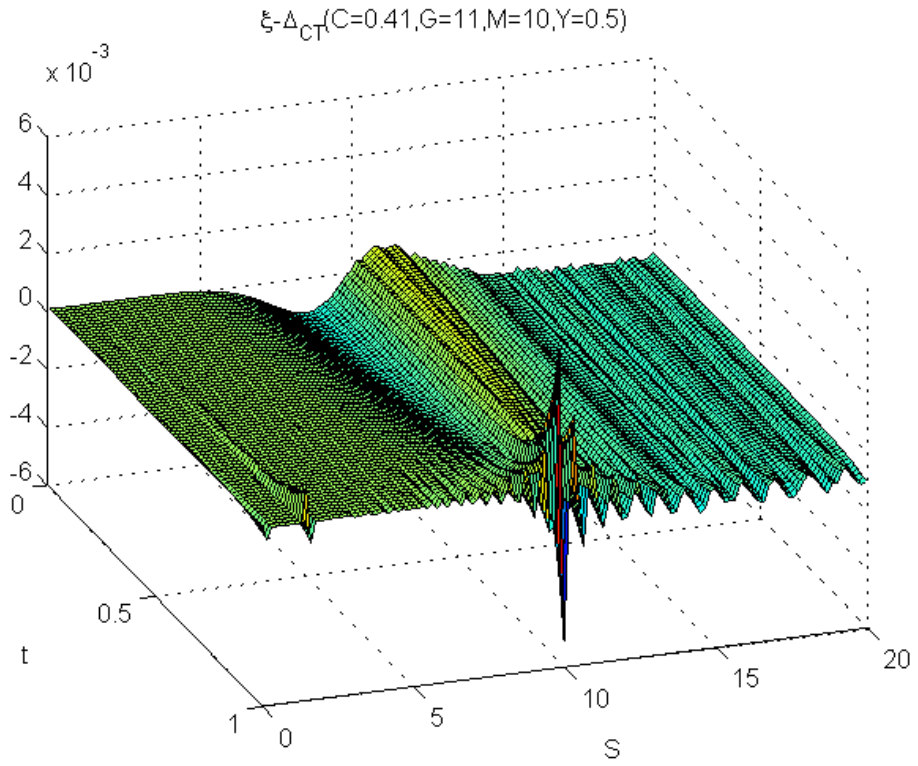

Fig. 7.4: The difference between the 'ideal' HKK hedge (ξ) and the CT optimal hedge, for the parameter set of Figure 7.10, for different values of the underlying and time to maturity ($T = 1$).

The long negative tail (with losses far exceeding the initial hedge capital P_0) and similar 95%-VAR for all methods, provides insight in the inherent riskiness of making a market in options — in uncommon 'disasters', when conditions suddenly move against the position, all hedging strategies in the underlying asset only, regardless of their optimality, offer scant protection.

Comparing the performance of the methods, it is evident that even Black-Scholes is more robust to large jumps than the Naive delta! While the CT approach only offers a variance reduction of about 5% over Black-Scholes (in many ways a validation of the enduring popularity of the method and a testament to its crude effectiveness). The HKK method offers slightly improved performance, but with a reduced initial capital H_0, as one would expect from a globally optimal method.

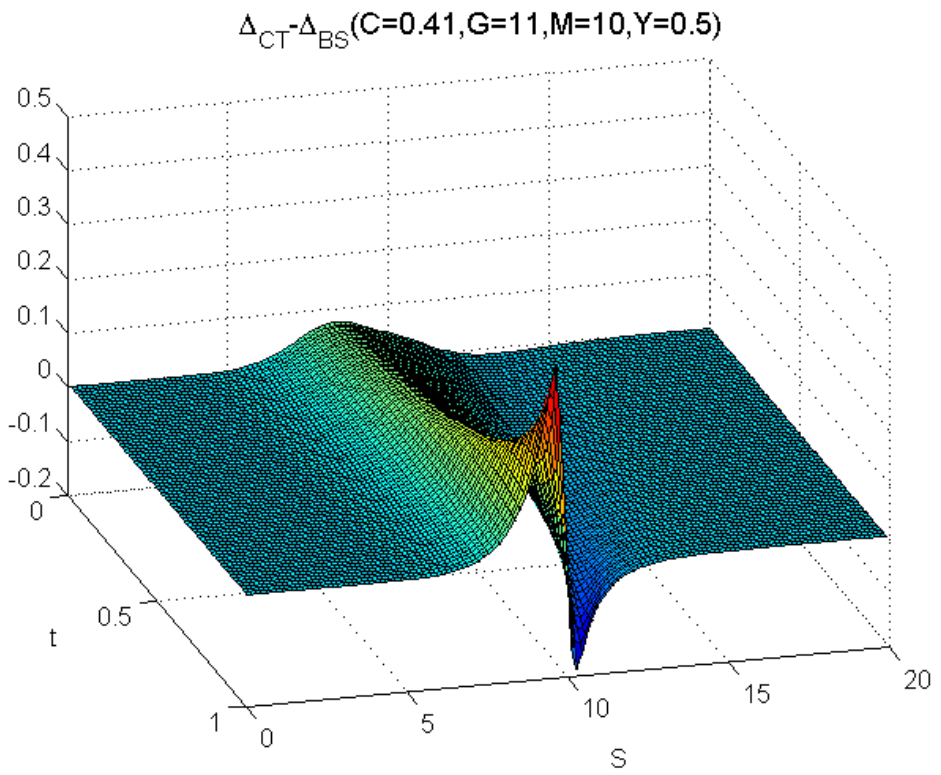

Fig. 7.5: The difference between the CT hedge and the Black-Scholes Delta, for the 'Large Jumps' scenario, for different values of the underlying and time to maturity ($T = 1$).

7.2.4 Large Unfavorable Jumps

When large unfavorable jumps (positive jumps in the case of a call) are common in the underlying price, hedging methods that take into account such structure offer modest advantages (see Figures 7.9 and 7.10). In this instance, the variance of the HKK method is about 12% less than the Black-Scholes approach and the 95%-VAR is improved by about 6%, offering a significant tightening in bid/ask spreads. Naturally, these advantages would be even greater if the large jumps were more pronounced, but it is not clear that such parameter sets are realistic judging by the parameters observed in the study by Carr *et al.* [4, Table 2].

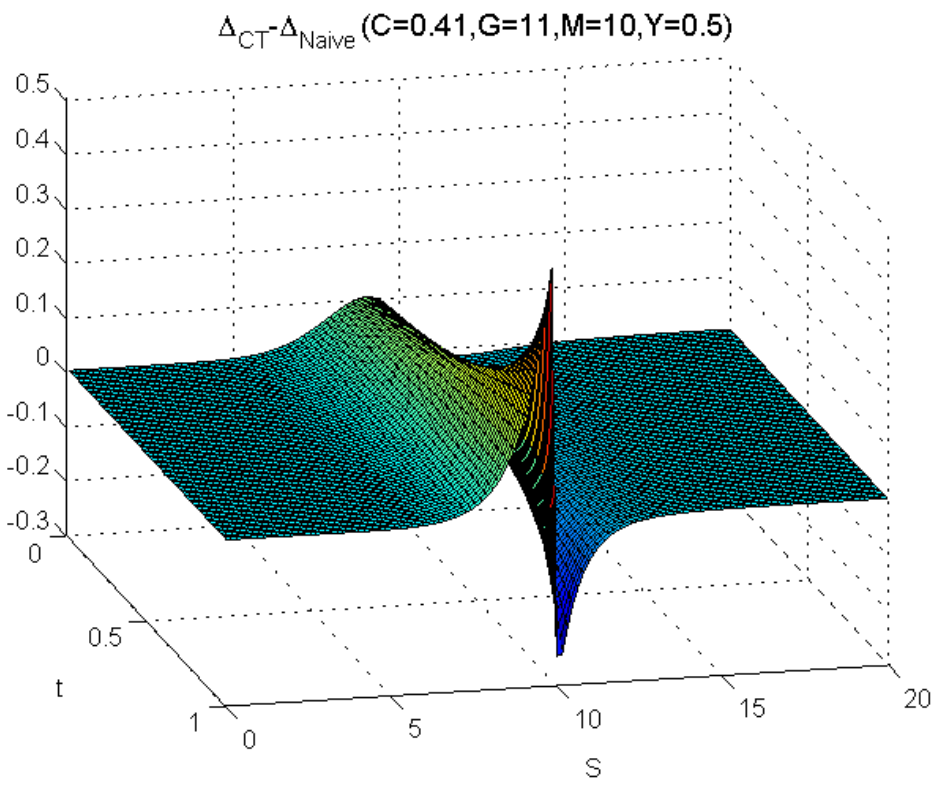

Fig. 7.6: The difference between the CT hedge and the 'Naive' Delta ($\partial C/\partial S$), for the 'Large Jumps' scenario, for different values of the underlying and time to maturity ($T = 1$).

7.2.5 Out-of-the-Money Options

The benefits of optimal hedging, using Lévy processes, are increased even more, if one considers an out-of-the-money option ($S_0 = 9, K = 10$) written on the same underlying process as the previous section (see Figure 7.11 and 7.12). Here, the VAR is improved by over 15%, while the variance is reduced by 16%.

7.2.6 High Activity/ Long Maturity

When hedging a longer duration option written on a process with the same parameters as the base case, the variance of the hedge is greatly reduced when compared to the initial capital (see Figures 7.13 and 7.14). This is to be expected since all Lévy processes aggregate to normality over time (see Section 2.3.3), reducing the impact

of Kurtosis on these timescales. Moreover, as the process begins to resemble geometric Brownian motion, the relative performance of the 'delta' based approaches is greatly enhanced.

This indicates that, over a longer timespan (where large jumps are no longer significant relative to the timescale), models with jumps offer little advantage with unnecessary complexity and may be dispensed with. In this situation, models introducing non-stationarity may provide a more relevant reflection of price behavior.

7.2.7 Infinite Variation

In the case of an infinite variation process (Figures 7.15 and 7.16), the dynamics are dominated by a preponderance small jumps (see Figure 3.4). Once again, as one would expect, there is little or no advantage to using methods that are optimized to deal with jumps — the large jumps are heavily damped by the G and M parameters and play little part in common price changes.

The slight negative expectation, evident in all figures, may arise from inaccuracies in the simulation algorithm resulting from the truncation of small jumps (which becomes significant as $Y \to 2$, see Section 4.2.4).

7.3 Conclusion

Based on the above simulations, it is evident that models based Lévy processes provide insight into pricing and hedging contingent claims under realistic market conditions, and are still somewhat tractable, lending themselves to several applications. Firstly, they show that dynamic replication is a risky business, verifying an intuition evident in market prices (the smile reflects an additional jump risk premium for out-of-the-money options). Moreover, such modeling techniques advise market participants on the range of prices at which it is sensible (profitable) to trade (as opposed to an otherwise ad-hoc smile adjustment). Furthermore, it was shown that optimal hedging techniques, based on the structure of jumps, may be used to achieve modest improvements in mitigating option risk as compared to

the Black-Scholes model, when large jumps are significant (typically on short time horizons). In addition, apart from pricing and hedging, such models may be used in stress testing scenarios for risk management and capital requirements, since the tails of the returns distribution may be constructed to be arbitrarily heavy.

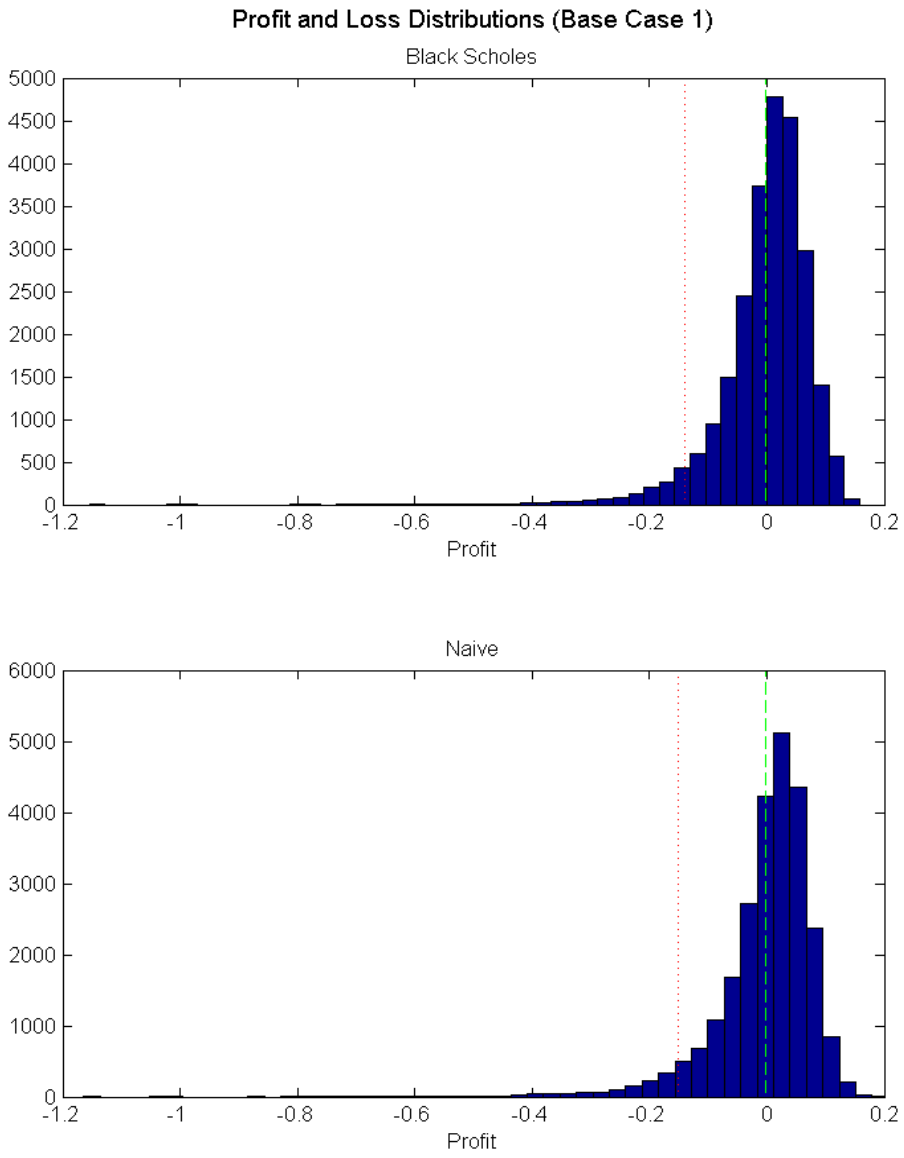

Fig. 7.7: A histogram of profits and losses when hedging a CGMY process with the parameters given in Table 7.1, using first the Black-Scholes delta [BS], then the Naive delta $[\partial C/\partial S]$, given an initial endowment P_0. All other parameters relevant to the simulation were defined in Section 7.1.1. The dotted red line demarcates the 95%-VAR, while dashed green line shows the expectation.

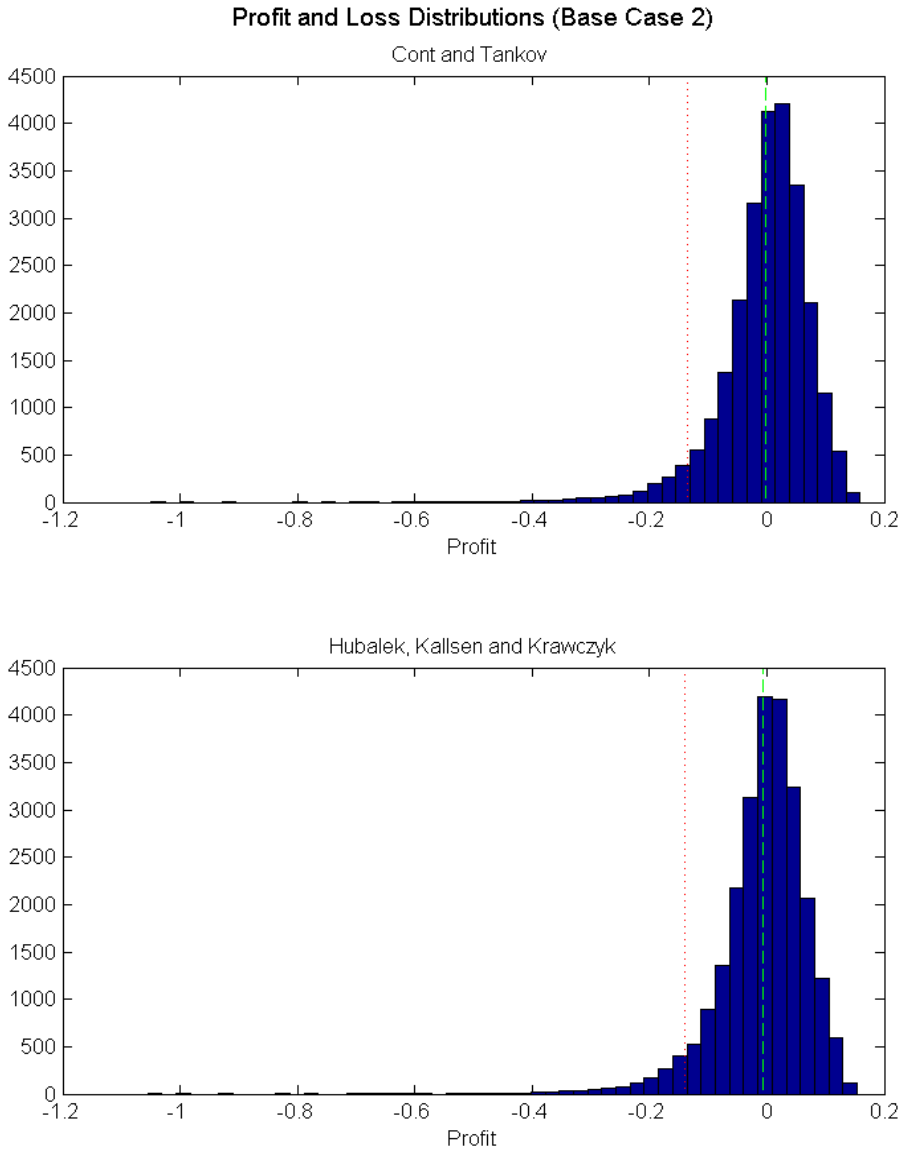

Fig. 7.8: A histogram of profits and losses on the same paths as in Figure 7.7, using first the Cont and Tankov [CT] method, then the approach of Hubalek *et al.* [HKK].

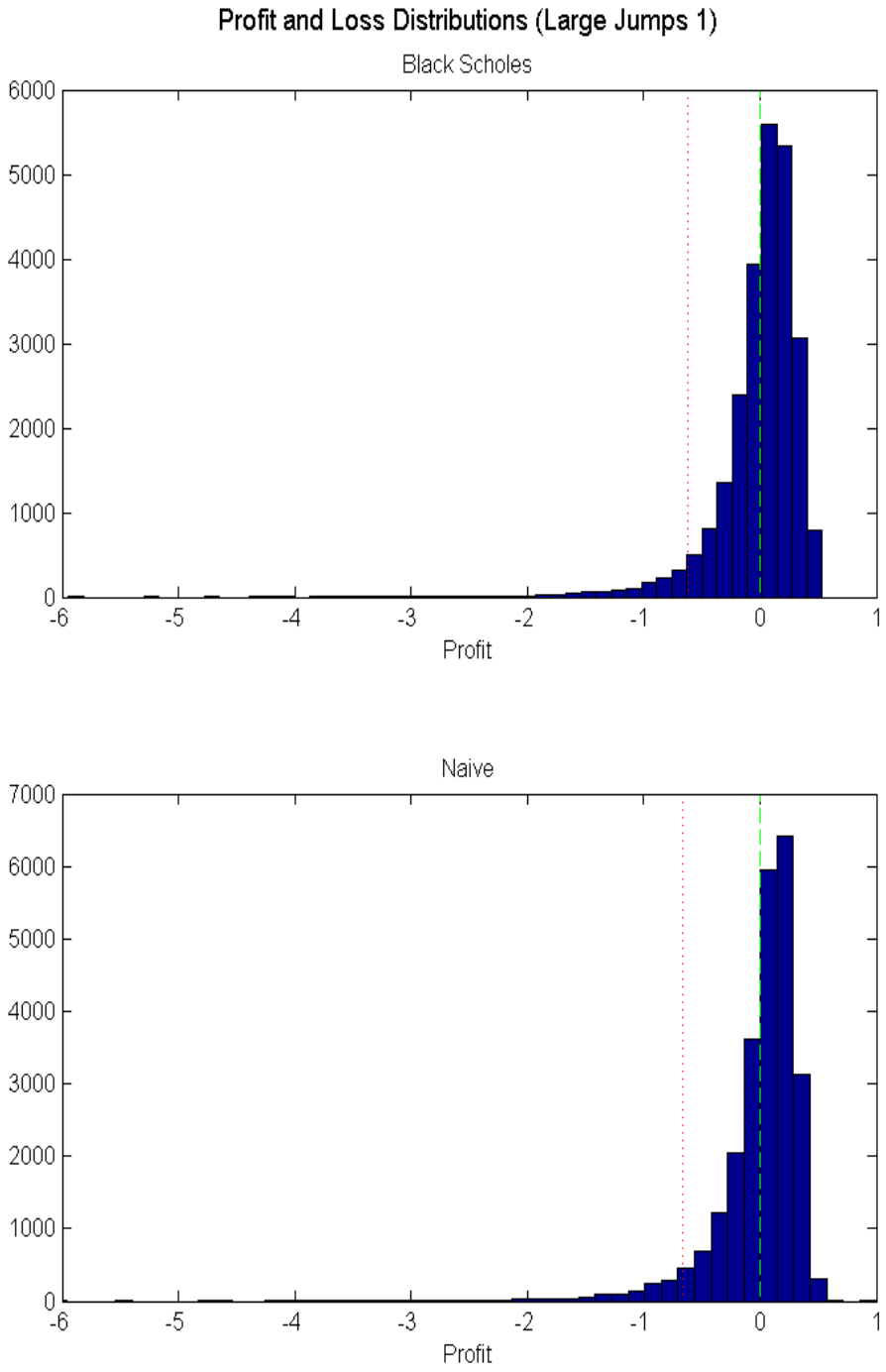

Fig. 7.9: A histogram of profits and losses in the presence of larger jumps (created by lower G and M parameters), using the BS and Naive approaches.

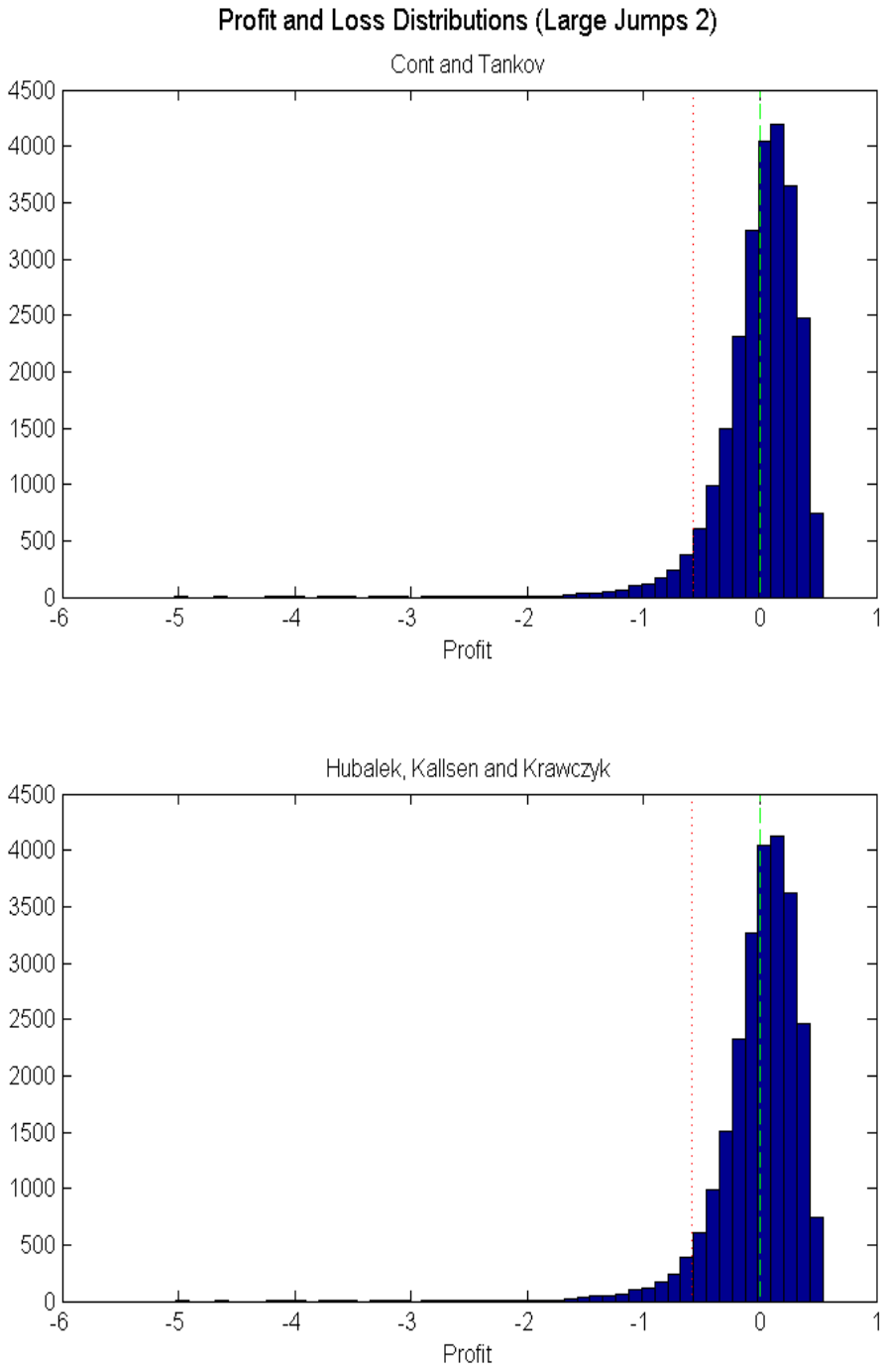

Fig. 7.10: A histogram of profits and losses in the presence of larger jumps (created by lower G and M parameters), using the CT and HKK approaches.

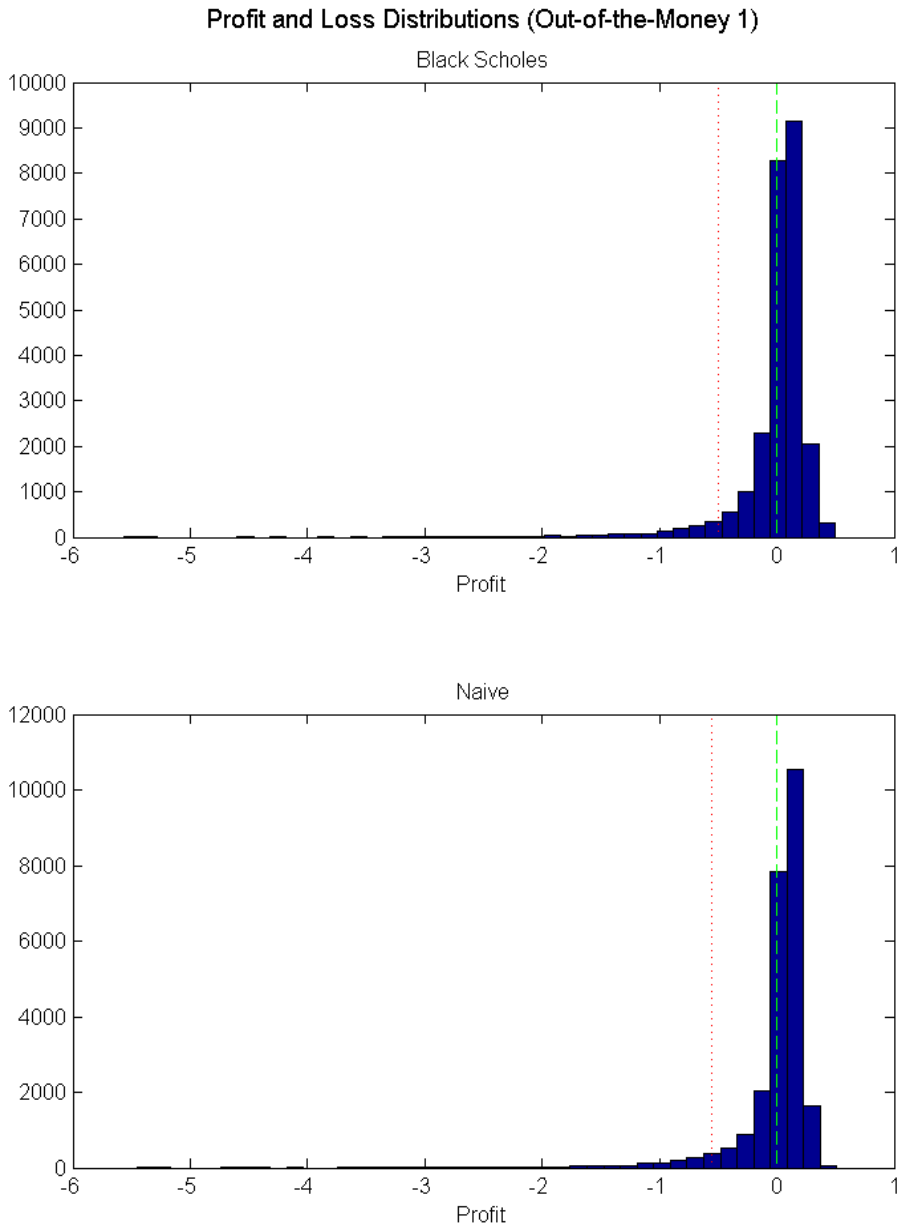

Fig. 7.11: A histogram of profits and losses for an out-of-the-money option ($S_0 = 9$ and $K = 10$) in the presence of larger jumps (created by lower G and M parameters), using the BS and Naive approaches.

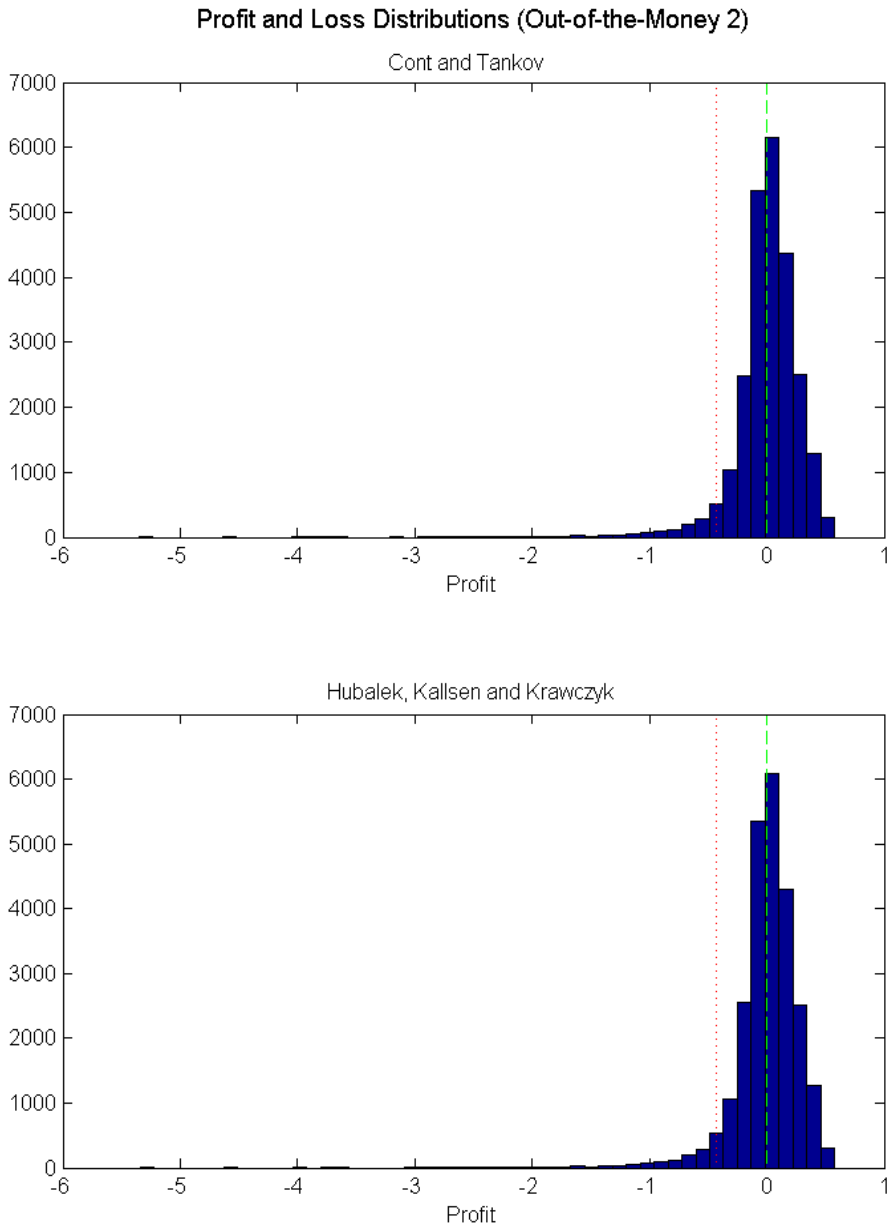

Fig. 7.12: A histogram of profits and losses for an out-of-the-money option ($S_0 = 9$ and $K = 10$) in the presence of larger jumps (created by lower G and M parameters), using the CT and HKK approaches.

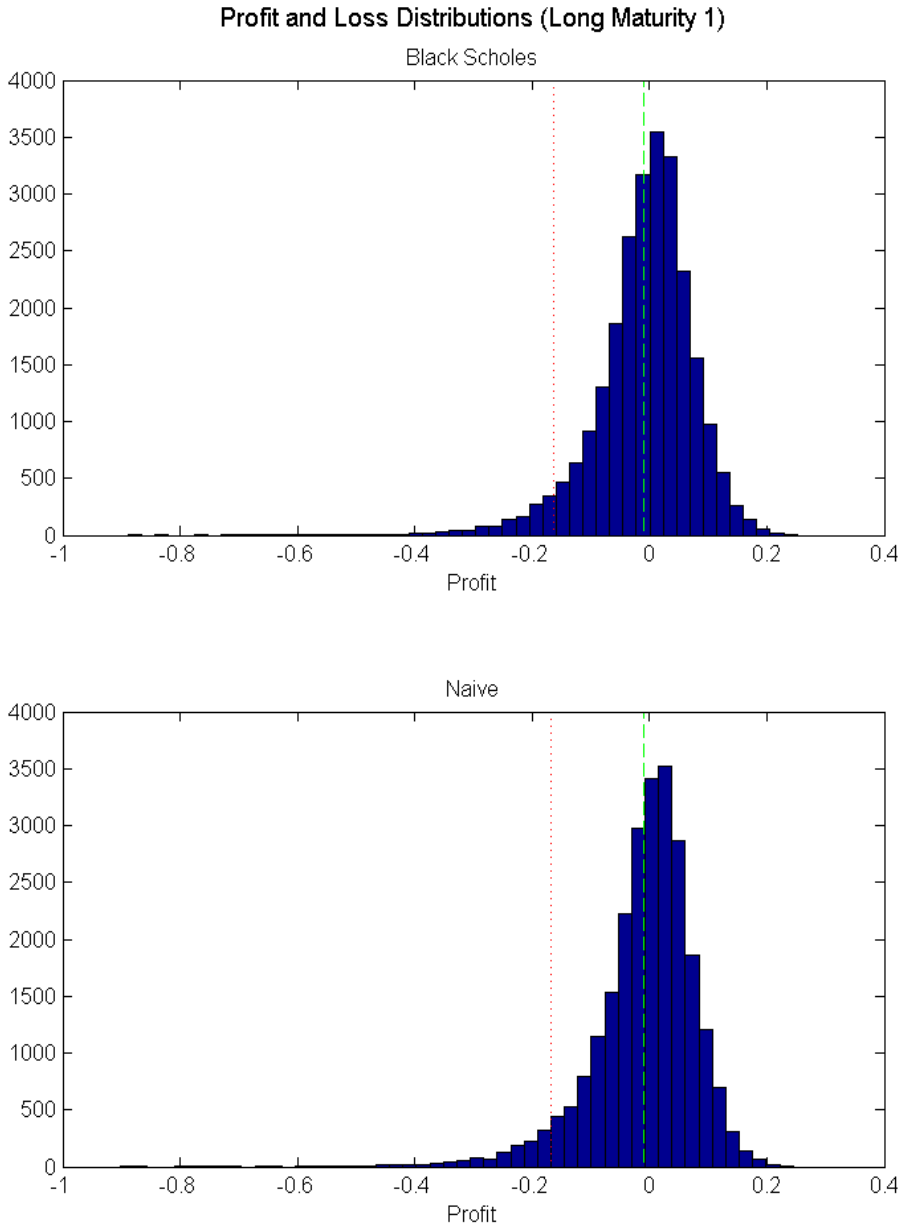

Fig. 7.13: A histogram of profits and losses when hedging over a longer time frame ($T = 5$), using the BS and Naive approaches.

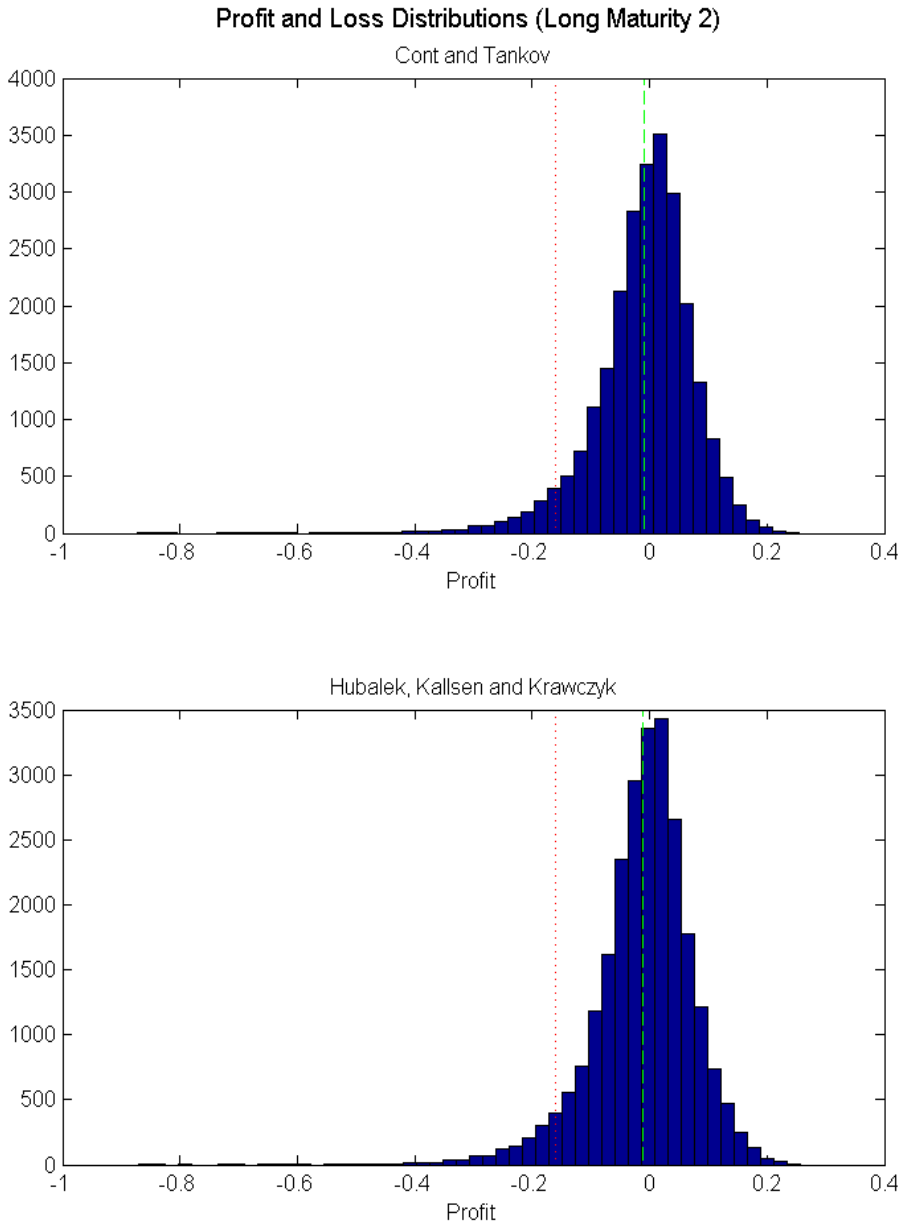

Fig. 7.14: A histogram of profits and losses when hedging over a longer time frame ($T = 5$), using the CT and HKK approaches.

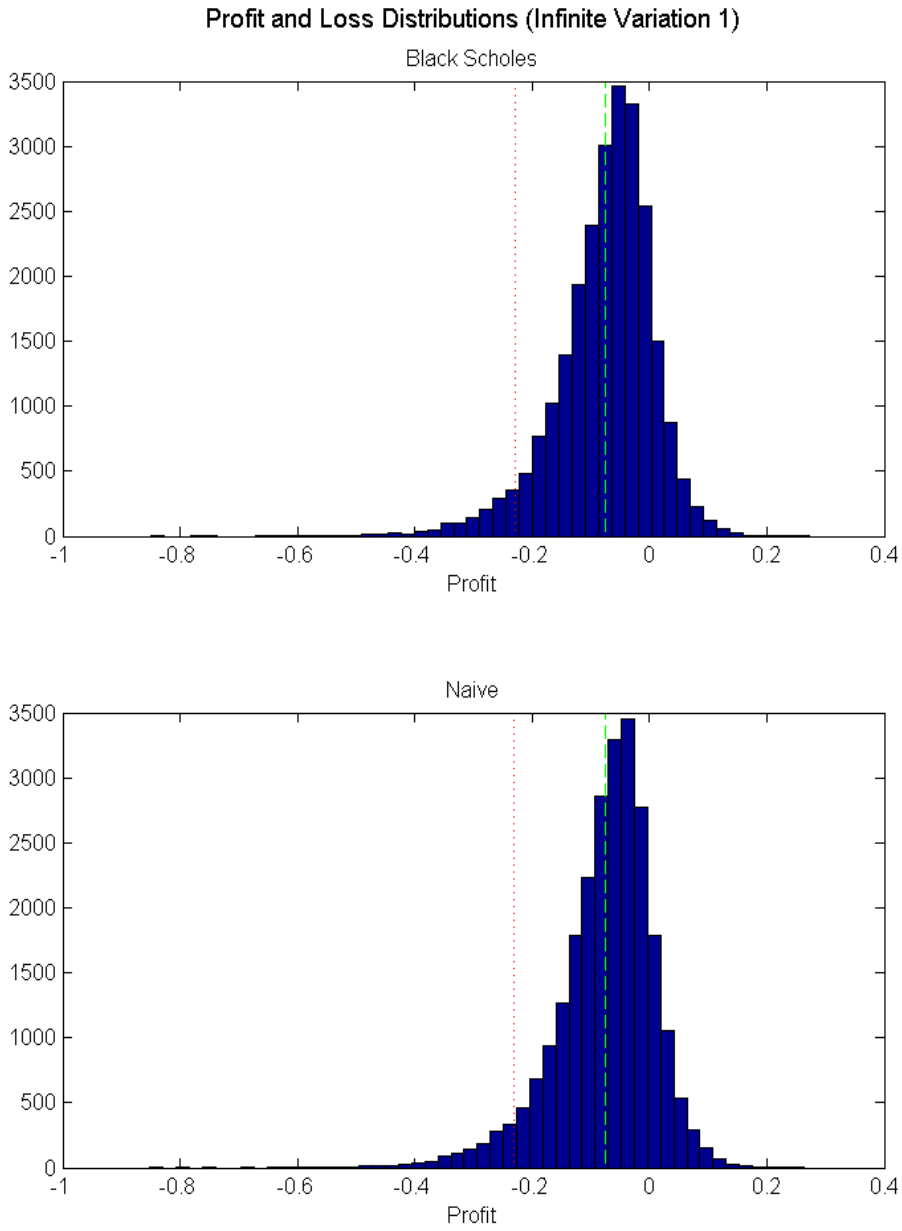

Fig. 7.15: A histogram of profits and losses when hedging an infinite variation CGMY process, using the BS and Naive approaches.

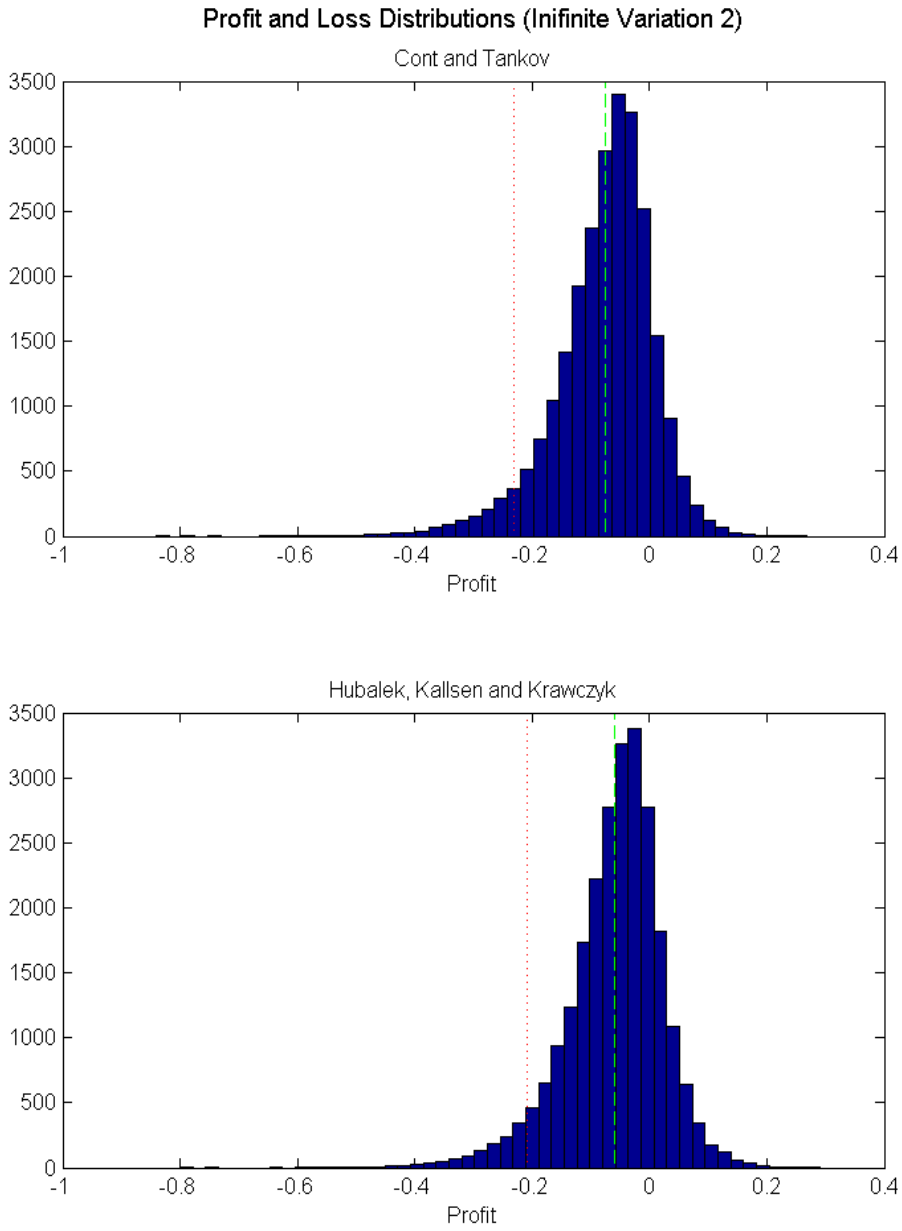

Fig. 7.16: A histogram of profits and losses when hedging an infinite variation CGMY process, using the CT and HKK approaches.

Appendix A

Selected MATLAB Implementations

In this appendix, MATLAB implementations of several of the algorithms discussed in the text are provided.

A.1 FFT Pricing

```
1  % European Call pricing, using the FFT method of
2  % Carr and Madan [5]
3  % OUTPUTS:
4  % P     = a vector of prices with corresponding strikes K
5  % INPUTS:
6  % So    = the initial stock price
7  % r     = the risk free rate
8  % T     = the option maturity
9  % N     = the number of points used in the FFT
10 % eta   = the grid step size
11 % model= the model of the underlying with characteristic function
12 %          implemented by CharFun
13 % varargin = a variable number of model parameters
14 function [P,K]=LevyPrice4(So,r,T,N,eta,model,varargin)
15
16 if So==0
17     P=0;
18     K=[];
```

```
19 else
20     switch upper(model)
21         case ('VG')
22             % Get the parameters
23             sig=varargin{1};
24             theta=varargin{2};
25             nu=varargin{3};
26
27             %Implement the risk neutral characterisitc function
28             %(see (5.11))
29             w=(1/nu)*log(1-theta*nu-0.5*sig^2*nu);
30             CharFun= @(u) exp(1i*u.*(log(So)+(r+w)*T)).*...
31                 (1-1i*u*theta*nu+0.5*sig^2*nu*u.^2).^(-T/nu);
32         case ('CGMY')
33             % Get the parameters
34             C=varargin{1};
35             G=varargin{2};
36             M=varargin{3};
37             Y=varargin{4};
38
39             %Implement the risk neutral characterisitc function
40             %(see (5.11))
41             switch Y
42                 case 0
43                     nu=1/C;
44                     theta=C*(1/M-1/G);
45                     sig=sqrt(0.5*((1/G+1/M)^2-theta^2*nu^2)/nu);
46
47                     w=(1/nu)*log(1-theta*nu-0.5*sig^2*nu);
48                     CharFun= @(u) exp(1i*u.*(log(So)+(r+w)*T)).*...
49                             (1-1i*u*theta*nu+0.5*sig^2*nu*u.^2).^(-T/nu);
50                 case 1
51                     error('This value of Y is unsupported')
52                 otherwise
53                     w=-C*gamma(-Y)*((M-1)^Y-M^Y+(G+1)^Y-G^Y);
54                     CharFun= @(u) exp(1i*u.*(log(So)+(r+w)*T)).*...
55                     exp(T*C*gamma(-Y)*((M-1i*u).^Y-M^Y+(G+1i*u).^Y-G^Y));
56             end
```

```
57          otherwise
58              error('Invalid model')
59      end
60
61      alpha=1;
62      % For a vanilla call. For a binary call replace with (5.10)
63      PSI= @(v)    exp(-r*T).*CharFun(v-(alpha+1)*1i)...
64                   ./(alpha^2+alpha-v.^2+1i*(2*alpha+1)*v);
65
66      v=eta*[0:N-1];                  %equation (5.16)
67      lambda=(2*pi)/(N*eta);
68      b=(N-1)*lambda/2;               %equation (5.18)
69      k=-b+lambda*[0:N-1];            %equation (5.17)
70
71      % The Kronecker Delta
72      delta=zeros(1,N);
73      delta(1)=1;
74      FUN = exp(1i*b*v).*PSI(v)*(eta/3).*...
75              (3+(-1).^[1:N]-delta);
76      P=exp(-alpha*k).*real(fft(FUN))/pi; %equation (5.20)%
77      K=exp(k);
78      %remove large and small entries:
79      K=K((~(P>So|P<So*1E-6))&isfinite(P));
80      P=P((~(P>So|P<So*1E-6))&isfinite(P));
81  end
```

A.2 CGMY Simulation

```
1  % Produces CGMY paths, using a subordinator representation
2  % based on [32] and [36]
3  % OUTPUTS:
4  % Xout = the process values
5  % t    = a corresponding fixed interval time grid
6  % INPUTS:
7  % C,G,M,Y = the process parameters
8  % n       = the number of paths
9  % p       = the number of timesteps
10 % T       = the final time
```

```
% t0       = the initial time
function [Xout,t]=CGMYprocess2(C,G,M,Y,n,p,T,t0)

if nargin <8
    t0=0;
end

A=(G-M)/2;
B=(G+M)/2;

epsilon=1E-4;          %The jump truncation level
K=C*2^(-Y/2)*sqrt(pi)/gamma(Y/2+0.5); %from Equation (4.14)

d=K*epsilon^(1-Y/2)/(1-Y/2);      %Equation (4.21)
lambda=2*K*epsilon^(-Y/2)/Y;      %Equation (4.20)

%Preallocate memory:
Xout=zeros(n,p);

for i=1:n
    %The jump times:
    tj=t0;
    while tj(end)<T
        U2=rand(1,2*round(lambda*(T-t0)));
        int=-log(U2)/lambda;
        tj=[tj tj(end)+cumsum(int)];
    end
    tj=tj(tj<T);

     %Applying (4.19) for the jump sizes:
    U1=rand(1,length(tj)-1);
    yj=[0,epsilon./(U1).^(2/Y)];

    %Applying the rejection (Theorem 4.7):
    U3=rand(size(yj));
    Zt=d*tj+cumsum(yj.*(f(yj)>U3));

    % Performing the subordination:
```

```
49      dZ=Zt(2:end)-Zt(1:end-1);
50      W=[0 cumsum(sqrt(dZ).*randn(size(dZ)))];
51      X=A*Zt+W;
52
53      %A MEX-C file, confining the process to a fixed time grid:
54      [Xout(i,:),t]= GridTime(X,tj,p,T);
55      clc
56      disp(['Path Generation ' num2str(100*i/n) '% Complete'])
57  end
58      % Equation (4.13)
59      function out=f(t)
60          out=2^(Y/2)*gamma(Y/2+0.5)*exp(A^2*t/2-B^2*t/4)...
61              .*D(-Y,B*sqrt(t))/sqrt(pi);
62      end
63
64  end
65
66  %Implementation of f(x) in term of confluent hypergeometric functions:
67  function out=D(nu,z)
68      %A parabolic cylinder function:
69      out=2^(nu/2).*U(-nu/2,0.5,z.^2/2).*exp(-z.^2/4);
70  end
71
72  function out=U(a,b,z)
73      out=zeros(size(z));
74      %A confluent hypergeometric function of the second kind:
75      cut=10;
76      out(z<cut)=(pi/sin(pi*b))*(HYPERGEOM(a,b,z(z<cut))/...
77          (gamma(a-b+1)*gamma(b))-z(z<cut).^(1-b)...
78          .*HYPERGEOM(a-b+1,2-b,z(z<cut))/(gamma(a)*gamma(2-b)));
79      %An alternate implementation with better convergence
80      %for large values of z:
81      Z=z(z>=cut);
82      temp=zeros(size(Z));
83      for i=1:length(Z)
84          fun= @(t) exp(-Z(i)*t).*t.^(a-1).*(1+t).^(b-a-1);
85          temp(i)= quad(fun,0,1E6,1E-4)/gamma(a);
86      end
```

```
87        out(z>=cut)=temp;
88    end
89
90    function out=HYPERGEOM(a,b,z,tol)
91    %A confluent hypergeometric function of the first kind:
92    if nargin<4
93        tol=1E-3;
94    end
95
96    out=1;
97    term=ones(size(z));
98
99    n=1;
100   while max(term)>tol||n<100
101       term=term.*((a+n-1)*z/(n*(b+n-1)));
102       out=out+term;
103       n=n+1;
104   end
105
106   end
```

Bibliography

[1] T. Ané and H. Geman, *Order flow, transaction clock, and normality of asset returns*, The Journal of Finance **55** (2000), no. 5, 2259–2284.

[2] D. Becherer, *Rational hedging and valuation of integrated risks under constant absolute risk aversion*, Insurance Mathematics and Economics **33** (2003), no. 1, 1–28.

[3] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy **81** (1973), 637–659.

[4] P. Carr, H. Geman, D. B. Madan, and M. Yor, *The fine structure of asset returns: An empirical investigation*, Journal of Business **75** (2002), no. 2, 305–332.

[5] P. Carr and D. B. Madan, *Option valuation using the fast fourier transform*, Journal of Computational Finance **2** (1999), no. 4, 61–73.

[6] R. Cont, *Empirical properties of asset returns: stylised facts and statistical issues*, Quantative Finance **1** (2001), 223–236.

[7] R. Cont and P. Tankov, *Financial modeling with jump processes*, Chapman & Hall\CRC, 2004.

[8] R. Cont, P. Tankov, and E. Voltchkova, *Hedging with options in models with jumps*, Stochastic analysis and applications (F. Espen Benth, G. Di Nunno, T. Lindstrøm, B. Øksendal, and T. Zhang, eds.), The Abel Symposium, 2005, pp. 197–217.

[9] J. M. Corcueray, D. Nualart, and W. Schoutens, *Completion of a Lévy market by power-jump assets*, Finance and Stochasics **9** (2005), no. 1, 109–127.

[10] F. Delbaen and W. Schachermayer, *The fundamental theorem of asset pricing for unbounded stochastic processes*, Mathematische Annalen **312** (1998), no. 2, 215–250.

[11] E. Derman, *Laughter in the dark—the problem of the volatility smile*, Euronext Options Conference (Amsterdam), May 2003.

[12] R. M. Dudley, *Uniform central limit theorems*, Cambridge University Press, 1999.

[13] B. Dupire, *Pricing with a smile*, RISK **7** (1994), 18–20.

[14] E. Eberlein and J. Jacod, *On the range of options prices*, Finance and Stochastics **1** (1997), no. 2, 131–140.

[15] B. E. Fristedt and L. F. Gray, *A modern approach to probability theory*, Birkhäuser, 1996.

[16] T. Fujiwara and Y. Miyahara, *The minimal entropy martingale measures for geometric levy processes*, Finance and Stochastics **7** (2003), no. 4, 509–531.

[17] H. Geman, D. B. Madan, and M. Yor, *Time changes for Lévy processes*, Mathematical Finance **11** (2001), no. 1, 79–96.

[18] A. Gut, *An intermediate course in probability*, 2nd ed., Springer, 2009.

[19] J. M. Harrison and S. R. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic Processes and their Applications **11** (1981), no. 3, 215–260.

[20] E. G. Haug and N. N. Taleb, *Why we have never used the Black-Scholes-Merton option pricing formula*, Available from SSRN: http://ssrn.com/abstract=1012075, February 2009.

[21] S. L. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies **6** (1993), no. 2, 327–343.

[22] S. D. Hodges and A. Neuberger, *Optimal replication of contingent claims under proportional transaction costs*, Review of Futures Markets **8** (1989), 222–239.

[23] F. Hubalek, J. Kallsen, and L. Krawczyk, *Variance-optimal hedging for processes with stationary independent increments*, The Annals of Applied Probability **16** (2006), no. 2, 853–885.

[24] K. Ito, *On stochastic processes (infinitely divisible laws of probability)*, Kiyosi Itô: selected papers (S. R. S. Varadhan and D. W. Stroock, eds.), Springer, 1987, pp. 261–301.

[25] F. James, *Monte Carlo theory and practice*, Reports on Progress in Physics **43** (1980), 1145–1189.

[26] O. Kallenberg, *Foundations of modern probability*, 2nd ed., Springer, 2002.

[27] J. Kallsen and R. Vierthauer, *Quadratic hedging in affine stochastic volatility models*, Review of Derivatives Research **12** (2009), no. 1, 3–27.

[28] F. C. Klebaner, *Stochastic calculus with applications*, Imperial College Press, 2005.

[29] P. Lévy, *Sur les intégrales dont les éléments sont des variables aléatoires indépendantes*, Annali della Reale Scuola Normale Superiore di Pisa **3** (1934), no. 2, 337–366.

[30] D. B. Madan, P. Carr, and E. C. Chang, *The variance gamma process and option pricing*, European Finance Review **2** (1998), 79105.

[31] D. B. Madan and E. Seneta, *The variance gamma (V.G.) model for share market returns*, The Journal of Business **63** (1990), no. 4, 511–524.

[32] D. B Madan and M. Yor, *CGMY and meixner subordinators are absolutely continuous with respect to one sided stable subordinators.*, Prépublication du Laboratoire de Probabilités et Modèles Aléatoires,, 2005.

[33] B. Mandelbrot, *The variation of certain speculative prices*, The Journal of Business **36** (1963), no. 4, 394–419.

[34] R. C. Merton, *Option pricing when underlying stock returns are discontinuous*, Journal of Financial Economics **3** (1976), 125–144.

[35] D. Nualart and W. Schoutens, *Chaotic and predictable representations for Lévy processes*, Stochastic Processes and their Applications **90** (2000), 109–122.

[36] J. Poirot and P. Tankov, *Monte carlo option pricing for tempered stable (CGMY) processes*, Asia-Pacific Financial Markets **13** (2006), no. 4, 327–344.

[37] J. Rosiński, *Series representations of Lévy processes from the perspective of point processes*, Lévy processes: theory and applications (O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick, eds.), Birkhauser, 2001, pp. 401–415.

[38] G. Samorodnitsky and M. S. Taqqu, *Stable non-gaussian random processes: stochastic models with infinite variance*, Chapman & Hall/CRC, 1994.

[39] K.-I. Sato, *Lévy processes and infinitely divisible distributions*, Cambridge University Press, 1999.

[40] W. Schoutens, *Lévy processes in finance: Pricing financial derivatives*, John Wiley and Sons, 2003.

[41] M. Schweizer, *A guided tour through quadratic hedging approaches*, Option pricing, interest rates and risk management (E. Jouini, J. Cvitanić, and M. Musiela, eds.), Cambridge University Press, 2001, pp. 548–574.

[42] J. A. Stephan and R. E. Whaley, *Intraday price change and trading volume relations in the stock and stock option market*, The Journal of Finance **XLV** (1990), no. 1, 191–220.

[43] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, 4^{th} ed., Cambridge University Press.