NTNU

Norwegian University of
Science and Technology

Pricing Asian Options by Numerical Path Integration with Advanced Lévy Dynamics

Markus Brabrand Urfjell

Master of Science in Physics and Mathematics
Submission date: July 2018
Supervisor: Arvid Næss, IMF

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

Applying a numerical Path Integration method which utilises the Fast Fourier transform to price fixed strike Asian options under Normal Inverse-Gaussian and Variance Gamma dynamics.

Preface

This master thesis concludes my M.Sc degree in Industrial Mathematics at the Department of Applied Physics and Mathematics at The Norwegian University of Science and Technology (NTNU). The work was performed during the spring 2018 and is an extension of my Project thesis [64]. I would like to thank my supervisor Arvid Næss from the Department of Mathematical Sciences for inspiration, support and guidance in financial mathematics, modelling stochastic processes and stochastic differential equations. I would also like to thank Trond Nilsen and Jan Richard Wilkens for excellent education in mathematics and general motivation. In *one* very specific way, I am like Albert Einstein in that I have been standing on the shoulders of giants.

Special thanks to my family as well for 23 years of love and support so far.

2018-07-10

Abstract

In this thesis I combine the strengths of the Path Integration method and the Fast Fourier transform to price discretely monitored, path dependent, fixed strike Asian options in a fast and accurate manner. The presented method can be used to accurately price various types of exotic options with greatly improved computation speed compared to the frequently used Monte Carlo simulations.
The method, in the form it is presented here, is implemented for underlying assets modelled by advanced Lévy processes; namely the Normal Inverse-Gaussian process and the Variance Gamma process in addition to the simpler Geometric Brownian Motion for comparison. Some interesting characteristics of these processes are uncovered and discussed.

Sammendrag

I denne artikkelen kombinerer jeg styrkene til veiintegrasjonsmethoden og den raske Fourier transformen til å prise diskrétobserverte, veiavhengige, låst strike asiatiske opsjoner på en rask of presis måte.
Den presenterte metoden kan bli brukt til å prise et bredt spekter av eksotiske opsjoner med langt raskere kjøretider enn de mye brukte Monte Carlo-simuleringene. Metoden, som den er presentert her, blir implementert for underliggende verdiprosesser modellert med avanserte Lévyprosesser: nærmere bestemt Normal Inverse-Gaussian prosessen og Variance Gamma prosessen i tillegg til den enklere geometrisk Brownske bevegelse brukt til sammenligning. Noen interessante egenskaper ved disse prosessene avdekkes og diskuteres.

Contents

List of Figures

List of Tables

1 Introduction

The main goal of this master thesis is to find useful and flexible models for time series describing financial price processes and adapt a Path Integration method with the Fast Fourier transform to these models. Additionally, I aim to demonstrate the use of this method by applying it to exotic option pricing by pricing arithmetic Asian options. I will start with a brief introduction to financial markets and the Geometric Brownian motion (GBM) used in Black & Scholes-analysis [21].

From there, I will look at some more advanced models which improve on the shortcomings of the GBM; namely the Normal-Inverse Gaussian model (NIG) and the Variance Gamma model (VG).

With obvious applications, financial markets are heavily studied through the years. The pricing of financial derivatives is one of the most open and active fields in financial theory today. In this thesis I will study some models for emulating the financial market in order to price financial derivatives; more specifically *options*.

A financial option is an agreement between two parts which gives the holder the right, but not obligation to buy (call option) or sell (put option) a financial object from/to the issuer at some agreed-upon time in the future called the "maturity time" for an agreed-upon price called the "strike". There is a vast pool of option types (see Section 2.3) on a variety of financial objects such as assets, interest rates and counterparty credit risk.

The introduction of Black & Scholes analysis in 1973 marked a majour step in financial analysis. The traditional Black & Scholes analysis, however, uses the Geometric Brownian motion (GBM) as model for asset dynamics. Empiric studies have shown this model to be too simplistic. Thus began the search for more sophisticated models. The models I will consider in this thesis are the Geometric Brownian Motion, the Negative Inverse Gaussian model (NIG) and the Variance Gamma model (VG).

The pricing methods I will present is the popular Monte Carlo Simulations and a more sophisticated method based on Path Integration (PI) with the Fast Fourier transform (FFT). I will refer to the more sophisticated method as the PIFFT-method. The PIFFT-method was successfully implemented for fixed strike Asian options under GBM-dynamics in my project thesis [64]. This thesis will extend that work by adopting the method to work under Advanced Lévy dynamics. The NIG and VG distribution are both examples of advanced Lévy distributions.

Asian options consider the average asset price in some period before maturity as well as the instantaneous price. This makes Asian options of higher dimension than the equivalent European option which only considers the instantaneous price. Furthermore, Asian options also offer some extra numerical difficulties which will

be discussed later. These extra challenges are why I have chosen to consider Asian options in particular. The presented method is applicable to a range of other options as well. Examples include lookback options (worked on by Østreim in 2009 [50]) and a range of more simple options such as barrier and spread options (explored in [65] for GBM-dynamics).

Furthermore, the computational cost of the commonly used Monte Carlo simulations prove particularly troublesome for Asian options. PI also conveniently tracks the evolution of the asset price and is therefore particularly useful for pricing path dependent options. Combining Path Integration with the Fast Fourier transform reduces computation times even further while retaining the required accuracy. This motivates my research into Path Integration as an alternative pricing technique for *path dependent* options.

The main idea behind the Path Integration method is to numerically propagate the probability density function of the stochastic process of the asset price forward in time until maturity. When the density of the asset price at maturity is known, one can easily price the considered option by applying the risk-neutral pricing formula. The propagation requires the evaluation of a sequence of convolution integrals. This evaluation will be done by the Fast Fourier transform. The result is a pricing method which is faster than Monte Carlo simulations and at the same time fully accurate with respect to the chosen model of asset dynamics.

Requirements for using Path Integration with the Fast Fourier transform is that the characteristic function of the transition probability of the asset price is known in closed form and that the stochastic process possesses the Markov property[1]. These requirements are met by all Lévy processes[2][60] which make up a large class of commonly used models of asset dynamics.

The concept behind Path Integration dates back to Norbert Wiener[25] before being developed further by P.A.M. Dirac in 1933[18]. The full method was introduced by Feynman in 1948 for use in the world of quantum mechanics. It was applied to the field of finance by Dash in 1989[17] and later by Linetsky [44] and more recently by Skaug and Næss[12]. Kleppe [41] and Østreim [50] were the first to successfully apply the Fast Fourier transform to Path Integration thus improving computation times. As far as I know, this is the first time PIFFT is successfully applied to Asian options under NIG and VG dynamics [3].
Path Integration belongs to the class of pricing methods which are based on transform and quadrature theory. Methods from this class have been increas-

[1]The next value of a stochastic process is only dependent on the current value. i.e. A stochastic process X_t possess the Markov property if and only if $\mathbb{P}(X_{t+1}|X_t, X_{t-1}, \cdots, X_0) = \mathbb{P}(X_{t+1}|X_t)$.[57]

[2]A stochastic process with independent increments that are statistically identical over time intervals of the same length. Examples: The Wiener process and the Poisson process.

[3]As far as I know, it was also the first time PIFFT was successfully applied to pricing Asian options under GBM-dynamics in [64]

ingly researched in recent years and are overrepresented in the pool of today's state-of-the-art methods. A common feature for these state-of-the-art option pricing methods are that they rely on transforms to the frequency domain (More on different pricing methods in Section 2.4).

2 Theory

In this section I will start by introducing the fundamental concepts behind the pricing rules applied in this thesis. Then, I will describe a selection of financial option and review some of the most important pricing methods we have today before presenting some fundamental probability theory.

I begin by stating some notation before introducing the concept of arbitrage and equivalent martingale measure and use this to arrive at the risk neutral pricing formula [54].

2.1 Intro to Pricing Rules

Consider a market evolving in time $t \in [0, T]$ with Ω the set of scenarios of possible realisations of the market in this period. Let $(\mathcal{F}_t)_{t\in[0,T]}$ be the filtration containing all information in the history of the market up to time t. An underlying asset can be described as a *non-anticipating* process with respect to the filtration $(\mathcal{F}_t)$. Let $S_t^i(w)$ denote the value of asset i in a market with $d+1$ assets under the scenario $w \in \Omega$ at time t. That is

$$\begin{aligned} S : [0,T] \times \Omega &\to \mathbb{R}^{d+1} \\ (t, w) &\to (S_t^0(w), S_t^1(w), ..., S_t^d(w)). \end{aligned} \tag{2.1}$$

Here $S_t^0(w)$ is a *numeraire* which denominates all assets to make the S_t^i comparable. A common choice of $S_t^0(w)$ is a cash account with constant risk-free interest rate r; that is

$$S_t^0 = e^{rt}. \tag{2.2}$$

This choice of S_t^0 is called the *discounted price process*[73] and discounting is done by numeraire (2.2).

A contingent claim[4] i is fully described by its payoff function $H^i(w)$ defined for $t = T$ (at maturity). Section 2.3 provides some examples of payoff functions. A pricing rule, Π, then maps a value $\Pi_t(H)$ to payoff function H at time t.

Such rules must satisfy the following requirements:

1. Non-anticipation:
   $\Pi_t(H)$ is only dependent on the history of the market up to time point t.

$$\Pi_t(H|S^0_{t\in[0,T]}, ..., S^{d+1}_{t\in[0,T]}) = \Pi_t(H|\mathcal{F}_t)$$

[4]Common expression for a derivative with a payout that is dependent on the realisation of some uncertain future event. Financial options are examples of contingent claims. Other examples include, but are not limited to swaps, forward and future contracts[39].

2. Positiveness:
   The value of an option with a non-negative payoff function is non-negative.
$$H \geq 0 \Rightarrow \Pi_t(H) \geq 0 \tag{2.3}$$

3. Linearity:
   The value of a portfolio of contingent claims is given by the value of each claim.
$$\Pi_t(\sum_{j\in J} H^j) = \sum_{j\in J} \Pi_t(H^j). \tag{2.4}$$

Consider now an Arrow security 1_A. That is a contingent claim which pays 1 if an event $A \in \mathcal{F}$ occurs and 0 otherwise. 1_Ω is then a zero coupon bond paying 1 at maturity with certainty. By the definition of the discounted price process, (2.2), the value of the zero coupon bond is just the discount factor. That is
$$\Pi_t(1_\Omega) = \frac{S^t_{1_\Omega}}{S^t_0} = \frac{1}{e^{r(T-t)}} = e^{-r(T-t)}.$$
Furthermore, let $\mathbb{Q}$ be defined as follows
$$\mathbb{Q}(A) = \frac{\Pi_0(1_A)}{\Pi_0(1_\Omega)} = e^{rt}\Pi_0(1_A) \tag{2.5}$$
A reasonable demand on Π is that $\Pi_t(1_A) \leq \Pi_t(1_\Omega)$. Additionally, by the positiveness requirement (2.3), $\mathbb{Q}(A) \in [0, 1]$.
Consider two disjoint events, A and B. The linearity requirement, (2.4), yield
$$\mathbb{Q}(A \cup B) = \frac{\Pi_0(1_{A\cup B})}{\Pi_0(1_\Omega)} = \frac{\Pi_0(1_A + 1_B)}{\Pi_0(1_\Omega)} = e^{rT}(\Pi_0(1_A) + \Pi_0(1_B)) = \mathbb{Q}(A) + \mathbb{Q}(B).$$
By letting the number of disjoint events go to infinity, $\mathbb{Q}$ is just a probability measure on the scenario space $(\Omega, \mathcal{F})$. Furthermore, consider a portfolio of a linear combination of Arrow securities, $H = \sum_i c_i 1_{A_i}$. Any random payoff H can be replicated by a linear combination of Arrow securities. The *risk-neutral pricing formula* can then be defined as follows:

Definition 1: Risk-neutral Pricing

Consider a payoff H in the set of all possible payoffs, $\mathcal{H}$ on which a dominated convergence theorem[a] holds. H is then valued by the risk-neutral pricing formula for a probability measure $\mathbb{Q}$ as
$$\Pi_{t=0}(H) = e^{-rT}\mathbb{E}^{\mathbb{Q}}(H) \tag{2.6}$$

[a] Necessary to give sense to the Expectation $\mathbb{E}^{\mathbb{Q}}(H)$ [68][47].

It is important to note that the probability measure $\mathbb{Q}$ does not represent the real world probability of occurrences of market scenarios. $\mathbb{Q}$ is rather a tangible probability measure for convenient representation of pricing rules. More on this in section 2.2

2.2 Arbitrage Pricing Theory

As described in section 1, a financial option is an agreement between two parts which gives the holder the right, but not obligation to buy (*call option*) or sell (*put option*) an asset from/to the issuer at some time in the future for an agreed-upon price called the strike or exercise price, K. To ensure a fair deal for both parties, it is necessary to price such options correctly. To this end I will apply arbitrage pricing theory.

Arbitrage pricing theory, introduced by Ross in 1976[67], is a pricing model which uses arbitrage arguments to derive a pricing relation. The foundation of the pricing rules used in this paper is the demand of no arbitrage opportunities.
An arbitrage opportunity is an opportunity to make money without taking risk. There should be no such opportunities in a fair market. Definition 2 offers a more rigid definition[8].

Definition 2: Arbitrage

An arbitrage opportunity is a value process of a self financing portfolio, θ, such that

$$V_t(\theta) = 0, \quad t = 0,$$
$$\mathbb{P}(V_t(\theta) \geq 0) = 1, \quad t = T > 0$$

where $\mathbb{P}$ is some real world probability measure containing the probabilistic views of investors regarding future scenarios.

The concept of martingales and the equivalent martingale measure are central tools of arbitrage pricing theory. Before I can give a rigid definition of a martingale according to [71], I need to define an *adapted* process.

Definition 3: Adapted process

A process $X = (X_n : n \geq 0)$ is called *adapted* (to the filtration $\{\mathcal{F}_n\}$) if for each n, X_n is $\mathcal{F}_n$-measurable.

Williams[71] offers the intuitive interpretation of Definition 3 that the value of the process X is known to us at time n if and only if X is adapted to $\mathcal{F}_n$.

Definition 4: Martingale

A process X is called a *martingale* (relative to the filtration $\{\mathcal{F}_n\}$ and the probability measure $\mathbb{Q}$) if

(i) X is adapted,

(ii) $\mathbb{E}^{\mathbb{Q}}(|X_n|) < \infty, \forall n$

(iii) $\mathbb{E}^{\mathbb{Q}}[X_n|\mathcal{F}_{n-1}] = X_{n-1}$ a.s. $(n \geq 1)$.

Intuitively, this means that a martingale process has no expected growth.

Definition 5: Equivalent Martingale Measure

A probability measure $\mathbb{Q}$ is an equivalent martingale measure to $\mathbb{P}$ if

1. $\mathbb{Q}$ is equivalent to $\mathbb{P}$ if they define the same (im)possible events, i.e. $\mathbb{Q} \sim \mathbb{P} : \forall A \in \mathcal{F} : \mathbb{P}(A) = 0 \iff \mathbb{Q}(A) = 0$.
2. The discounted stock values $e^{-rt}S_T^i$ are martingales with respect to the probability measure $\mathbb{Q}$, i.e. $\mathbb{E}^{\mathbb{Q}}\left[e^{-rt}S_T^i|\mathcal{F}_t\right] = e^{-rt}S_t^i$, $\forall i \in D$.

The following property holds for pricing rules on probability measures as defined above.

Theorem 1: Risk-Neutral Pricing

In a market governed by a probability measure $\mathbb{P}$, any arbitrage free pricing rule, Π_t, can be represented as

$$\Pi_t(H) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[H|\mathcal{F}_t],$$

where $\mathbb{Q}$ is an equivalent martingale measure to $\mathbb{P}$.

So far, I have assumed that an equivalent martingale measure exists and that this implies that the market is arbitrage free. The converse result is more difficult to prove and is called the First Fundamental Theorem of Asset Pricing:

Theorem 2: The First Fundamental Theorem of Asset Pricing

A market similar to (2.1) is arbitrage-free if and only if there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that the discounted assets S_t^i are martingales with respect to $\mathbb{Q}$.

Thus, when considering a model for the underlying asset, one has to make sure that

the stochastic variable being modelled is a martingale under the chosen market model. More on this in Section 3.1, 3.2.1 and 3.3.1.

2.3 Options

This section will briefly introduce some common types of financial options all of which can be priced by numerical Path Integration with the Fast Fourier transform. I will denote the price process of the underlying asset or stock by S_t, the time at maturity T and the strike price K.

2.3.1 European Options

European options are the simplest type of option. A European call or put option gives the holder the right to respectively buy or sell an asset for some agreed upon strike price at some agreed upon maturity date.

The payoff of a European call option, $C(K, S_T)$ and a European put option, $P(K, S_T)$ is

$$H^{EU_{call}} = C(K, S_T) = \max\{S_T - K, 0\}$$
$$H^{EU_{put}} = P(K, S_T) = \max\{K - S_T, 0\}$$

The value of a European option at some initial time, $t = 0$, before maturity can be expressed by the risk-neutral expectation of the option payoff:

$$\text{Call: } \Pi_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}\Big[\max\{S_T - K, 0\}\Big]$$
$$\text{Put: } \Pi_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}\Big[\max\{K - S_T, 0\}\Big]$$

This is easily generalised from the time period $t \in [0, T]$ to $t \in [t_a, t_b]$. I will use $t \in [0, T]$ because of personal preference.

2.3.2 American Options

An American option is a generalisation of the European option where the holder can choose to exercise the option at any time in the interval $t \in [0, T]$. There are several ways of valuing American options, e.g. solving the Black & Scholes equation with finite differences [53]. I will restrict myself to claiming that the value of an American call option must satisfy

$$\Pi_0 \geq e^{-rT}\mathbb{E}^{\mathbb{Q}}\Big[\max\{S_T - K, 0\}\Big]$$

and refer to [53] for more details.

2.3.3 European Barrier Options

A barrier option deploys an additional condition where the option only has value if the asset price is above (*upper barrier options*) or below (*lower barrier options*)

some agreed upon barrier. Let $\mathcal{I}$ be an indicator function with value 1 if the option is active, i.e. has value, and 0 otherwise. The value of a European barrier call option becomes:

$$\Pi_{t=0} = e^{-rT}\mathbb{E}^{\mathbb{Q}}\Big[\max\{S(T) - K, 0\} \times \mathcal{I}\Big]$$

European barrier options are among the most commonly traded exotic options. They are popular on the over-the-counter (OTC) market and traded with high volume in currency, interest rates and commodity markets.

2.3.4 Spread Options

Spread options have a payoff equal to the difference between the difference of two assets and some agreed upon strike. The spread could for example be induced by differences in prices over different markets. In currency, for example, a certain currency might be priced differently in two different countries. Other examples include energy markets and commodities which may be priced differently according to location and calendar. Spread options are used in market speculation and as a risk management tool and are mostly traded OTC.
The price of a spread option is given by

$$\Pi_{t=0} = e^{-rT}\mathbb{E}^{\mathbb{Q}}[\max(S_1(T) - S_2(T) - K, 0)] \tag{2.7}$$

2.3.5 Lookback Options

The lookback option considers the *optimal value* of the underlying stock in the lifetime of the option [50]. Clearly, this makes lookback options path dependent. There are two main types of lookback options: Fixed strike lookback options and floating strike lookback options.

For fixed strike lookback options, *"Optimal value"* means that the holder of the option can choose the value of the underlying in $t \in [0, T]$ which results in the largest payoff. This has an obvious appeal to investors. However, lookback options are often expensive.

The value of a fixed strike lookback call option is

$$\Pi_{t=0} = e^{-rT}\mathbb{E}^{\mathbb{Q}}\Big[\max\Big\{\max\{S_{t\in[0,T]}\} - K, 0\Big\}\Big]$$

For the equivalent put option, one would select the minimum price of the underlying in its lifetime, $\min\{S_{t\in[0,T]}\}$ instead of the maximum.

A floating strike lookback call (or put) option has the same payoff as its fixed strike equivalent, except it is the strike that is set at maturity as the lowest (or highest) value of the underlying in the options lifetime.

2.3.6 Asian Options

Asian options consider the average price of the underlying over some period before maturity, T, $t \in [t_a, t_b]$, $0 \leq t_a < t_b \leq T$. For simplicity, I choose $t \in [0, T]$ for this paper. All theory is easily generalised to $t \in [t_a, t_b]$. In contrast to European and American options, Asian options are clearly path dependent as they consider the history of the asset price over a time interval as opposed to only at maturity. The average price of the underlying is almost always[5] less volatile than the price of the underlying at maturity. Thus, Asian options cost less than the equivalent European option. Asian options are also less prone to market manipulation as it is more difficult to manipulate the average price than the price at a single point in time.
Asian options appeal to contractors who wish to be less exposed to market volatility. They are commonly found in commodity and energy markets as well as currency markets.
The payoff, $V(T)$, equal to the value of the option at maturity, T, is on the form

$$V(T) = \max\left\{\bar{S}_{t\in[t_a,t_b]} - K, 0\right\}, \tag{2.8}$$

where $\bar{S}_{t\in[t_a,t_b]}$ is some representation of the average asset price over the period $t \in [t_a, t_b]$. The price of an Asian option is thus given by

$$\Pi_{t=0} = e^{-rT}\mathbb{E}^{\mathbb{Q}}\Big[\max\left\{\bar{S}_{t\in[t_a,t_b]} - K, 0\right\}\Big]$$

There are several traits distinguishing different types of Asian options. For example the period of averaging (the values for t_a and t_b), the type of averaging (see Appendix C), the weighting of the average (one could for example give greater weights to more recent prices), the monitoring of the asset price (discretely or continuous) and whether the strike is fixed or floating as discussed for lookback options. I will consider arithmetic, uniformly weighted, discretely monitored options averaging over the time interval $t \in [0, T]$.

2.4 Review of Pricing Methods

There is a vast pool of pricing methods as this is a field which has received a lot of attention. Many of these methods become vastly complicated or unapplicable by the introduction of multi-asset options, path dependency or other types of higher dimensional options. Closed form solutions for example are not always available. The flexibility and easy implementation of Monte Carlo simulations makes this one of the most popular methods for pricing exotic options. Unfortunately, Monte Carlo simulations are, as mentioned, computationally expensive; especially in higher dimension. This is a major source of motivation for my research into faster methods which uphold the required accuracy demands.

This section divides pricing methods into rough classes and gives a brief description. The classes are Monte Carlo methods, partial differential equation based methods, tree methods and transform and quadrature based methods.

[5] Except in cases of measure zero

Monte Carlo Methods

Monte Carlo simulation is a widely used tool and was first applied to finance by Hertz in 1964 [29] and to option pricing by Boyle in 1977[10]. The easy implementation and flexibility with respect to the model of asset dynamics, payoff structures and higher dimensions make Monte Carlo simulations a popular method among practitioners today.
Monte Carlo simulations are based on producing many realisations of the chosen model by drawing random numbers and then averaging to compute the expected payoff at maturity. The expected payoff is then discounted back to the initial point in time to find the value of the option today. The computational cost of Monte Carlo simulations is considered the major drawback of the method where one must keep producing realisations until achieving a desired level of accuracy. Naturally, this becomes increasingly demanding with higher dimensions.

PDE Based Methods

The approach developed by Black and Scholes [21] introduced partial differential equations for modelling the asset dynamics and price simple options. For more complex options and/or models, where these equations cannot be solved analytically one applies numerical techniques; collectively named Partial Differential Equation Based Methods. The Finite Difference method is the most common of these where the asset value is predicted at the modes on a specified grid. Infinitesimal derivatives are approximated by finite differences. The resulting difference equations are then solved iteratively. Other methods in this class include Finite Element methods [72] and Finite Volume methods [66]. One drawback of the PDE Based Method is their often comprehensive requirements e.g. closed form expressions of volatility and drift. More recent developments in this class of methods is associated with the development of spectral methods which introduce global basis functions in the pricing schemes. Several master thesis have been written on the spectral method at NTNU in collaboration with Espen Jakobsen; most recently Frida Bruun in the Autumn 2017 and Spring 2018[6].

Tree Methods

The use of binomial trees for pricing options was presented by Cox and Ross in 1979 [40]. The idea is to discretise the asset dynamics as well as the time such that the asset price may either do an upward jump of rate u or downward jump of rate d at each time step with a specific probability p and $1-p$ respectively. The price of the option today can then be calculated recursively from the prices at all nodes at maturity. In the limit where the number of time steps goes to infinity, tree methods converge to the continuous random walk model [49]. Handling dividends, early exercise and path dependency is easy with tree methods. A further development of tree methods is the introduction of trinomial trees [9] where the price has three different outcomes for each time step.

[6]Bruun's work is yet unpublished

Transform and Quadrature Methods

The methods in this class aims to numerically calculate the risk-neutral expectation of the option payoff and discounting the payoff to present time. If the probability density function of the payoff is known and available in closed form, one can calculate this expectation by numerical integration.
However, closed form distributions are often unavailable for many asset price models. Stein and Stein [19] introduced the use of the Fourier transform to hurdle this difficulty when deriving closed-form approximate expressions for the distribution of the asset process where the volatility follows an Ornstein-Uhlenbeck process. In 1993, Heston [30] introduced a method for pricing European Options with stochastic volatility by use of the inverse Fourier transform of the characteristic function[7] A drawback of Heston's method is that it cannot use the Fast Fourier transform and thus exploit its computational efficiency. In 1999, Carr and Madan [52] introduced a method based on the Fast Fourier transform. The method aims to find an expression for the Fourier transform of the risk-neutral expectation of the options payoff in terms of the characteristic function. The option value is then found by applying the inverse Fourier transform to this expression by the Fast Fourier transform. The method was demonstrated on European options where the asset dynamics were modelled by the Variance Gamma process.

The pricing of more exotic derivatives, for example path dependent options, require more sophisticated methods. Sullivan [61] applied a Gaussian quadrature method for pricing American options[8] in the Black & Scholes framework in 2000. Andricopoulos et al. [1] extended Sullivan's work with the development of the QUAD method in 2003 which is applicable to high-dimensional and path dependent options in the Black & Scholes framework. In 2014, Chen et al. [14] aimed to generalise the QUAD method to make it applicable to other models than the Black & Scholes model. They proposed a quadrature based method using an approximation of the transition density of the asset price.

Moreover, O'Sullivan [51] combined the work of Carr & Madan and Andricopoulos et al. in 2005 and introduced one of the first methods which combine the Fast Fourier transform and quadrature schemes. O'Sullivans innovation is based on the existence of closed-form expressions of the Characteristic functions for many useful asset dynamic models where the probability density function is not available in closed form. The probability density function is then available for numerical integration schemes as the inverse Fourier transform of the characteristic function, computed by the Fast Fourier transform.

In 2000, Reiner [56] made the observation that the risk-neutral valuation formula can be written as a convolution. Lord et al. [55] used this observation in combination with Carr and Madan's work to extend on O'Sullivan's work in 2008 by the

[7] More on the characteristic function in Section 2.5.1.

[8] An American option is a generalisation of a European option which allows for early exercise at any time between initiation and maturity. More on this is section 2.3.2

introduction of the CONV method. Lord et al. used the Fast Fourier transform in the evaluation of the associated convolution integrals thus reducing computational costs considerably. The CONV method is also suitable for pricing options with the possibility of early exercise, for example Bermudan options[9]. Fang and Oosterlee [22] achieved even faster pricing evaluation by using the COS-method where the transitional density function is replaced by its Fourier-cosine expansion. The idea is to exploit the relationship between this Fourier-cosine expansion and the characteristic function of the driving density. An advantage of the COS-method is that it can be applied to the Heston Stochastic volatility model.

2.5 Probability Theory

In this section, I follow [60] to define skewness and kurtosis before stating some fundamental theory.

The *skewness* of a distribution determines to which degree the distribution is asymmetric. Skewness is defined as the third moment about the mean divided by the third power of the standard deviation.

Definition 6: Skewness

The skewness of the distribution of a random variable X with mean μ_X and standard deviation σ_X is defined as

$$\frac{\mathbb{E}[(X-\mu_X)^3]}{\sigma_X^3}.$$

Note that symmetric distributions such as the normal distribution has zero skewness because they have $\mathbb{E}(X) = \mu_X$.

The *kurtosis* of a distribution determines how fat the tails are. That is the likelihood of extreme events. Kurtosis is defined as the fourth moment around the mean divided by the fourth power of the standard deviation (or equivalently the second power of the variance).

Definition 7: Kurtosis

The kurtosis of the distribution of a random variable X with mean μ_X and standard deviation σ_X is defined as

$$\frac{\mathbb{E}[(X-\mu_X)^4]}{\sigma_X^4}$$

[9] Informally, Bermudan options are a discrete form of American options where early exercise is allowed, but only at a set of predetermined dates

The normal distribution has a kurtosis of 3 and is called *mesokurtic*. Distributions with kurtosis less than 3 are called *platykurtic* and have a flatter behaviour around the mean. The uniform distribution is an example of a platykurtic distribution. Distributions with kurtosis greater than 3 are *leptokurtic* with a high, concentrated peak at the mean. Examples include the Student-T distribution. The amount of kurtosis above 3, if any, is called *excess kurtosis*.
The following follows[63]:

Definition 8: Stable Distribution

A random variable Y has a stable distribution if, for all $n \geq 1$, it satisfies

$$Y_1 + ... + Y_n \stackrel{d}{=} a_n Y + b_n$$

where $Y_1, ..., Y_n$ are independently, identically distributed (iid) random variables with the same distribution as Y, $a_n > 0$ and $b_n \in \mathbb{R}$.

Definition 9: Random Walk

S_n is a random walk if, for some iid Z_k for $k \geq 1$,

$$S_n = \sum_{k=1}^{n} Z_k, \qquad n \in \mathbb{N}.$$

Theorem 3

Let S_n be a random walk. A random variable Z is said to have a stable distribution if and only if, for every n, there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$a_n S_n + b_n \stackrel{d}{\to} Z.$$

Definition 10: Infinite Divisability

A random variable Y is said to have an infinite divisible distribution if, for every $m \geq 1$,

$$Y \stackrel{d}{=} Y_1^{(m)} + ... + Y_m^{(m)}$$

for some iid random variables $Y_1^{(m)}, ..., Y_m^{(m)}$.

Note that stable distributions are infinitely divisible.

2.5.1 Characteristic Function

Being able to describe the distribution of a random variable is imperative when working with stochastic differential equations. For many, the probability density

function is the preferable way to do this. However, a closed form expressions for the probability density functions are seldom available.

The characteristic function is a central aspect in probability theory as it provides an alternative way of describing a random variable which also completely defines the probability density function of the variable. The characteristic function exists and is uniquely defined for all distributions [63].

Definition 11: Characteristic Function

The characteristic function of a random variable, X, with probability density function, f_X, is given by

$$\phi_X(u) = \int_{-\infty}^{\infty} \exp(iux) f_X(x) dx = \mathbb{E}[\exp(iuX)], \qquad u \in \mathbb{R} \tag{2.9}$$

One recognises (2.9) as the Fourier transform of the probability density function of X (See Appendix A for more on Fourier Analysis). Thus, for a given, ϕ_X, one can compute f_X by taking the inverse Fourier transform.

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-iux) \phi_X(u) du. \tag{2.10}$$

In Section 5, it will be useful to evaluate convolutions of a function $g(x)$ and a distribution $f_X(x)$. Combining (2.9) with the convolution theorem (Theorem 8) yields

$$(f_X * g)(x) = \mathscr{F}^{-1}\left[\phi_X(u) \mathscr{F}(g(u))\right]. \tag{2.11}$$

The characteristic function can be expanded to higher dimensional random variables. In particular, let $\boldsymbol{Z}_t$ denote a vector of the logarithmic returns of two GBM driven stochastic variables. In this case

$$\phi_{\boldsymbol{Z}_t}(\boldsymbol{u}) = \exp\left[i\boldsymbol{u}(\boldsymbol{\mu e} t - \frac{1}{2}\boldsymbol{\sigma}^2)' - \frac{1}{2}\boldsymbol{u}\boldsymbol{\Sigma}\boldsymbol{u}' t\right], \quad \text{where} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix}.$$

Here $\boldsymbol{e} = (1,1)$, $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ is the mean vector, $\boldsymbol{\Sigma}$ is the covariance matrix, $\boldsymbol{\sigma^2} = (\sigma_1^2, \sigma_2^2)'$ and ρ is the correlation between the two stochastic variables.

Skipping details, ([63, p. 32-33] and [59, Theorem 14.15]), we have that:

Theorem 4

The distribution of a random variable, Y, is stable if and only if its characteristic function can be written as

$$\phi_Y(u) = \exp\left(iu\eta - c|u|^\alpha(1 - i\beta sgn(u)g(u))\right) \tag{2.12}$$

with

$$g(u) = \begin{cases} \tan(\pi\alpha/2) & \text{for } \alpha \in (0,1) \cup (1,2], \\ 2/\pi \log|u| & \text{for } \alpha = 1 \end{cases}$$

for $\alpha \in (0, 2]$, $\beta \in [-1, 1]$, $c > 0$ and $\eta \in \mathbb{R}$.

In Theorem 4, c is a scale parameter, η is the location parameter, α is the *index* of the distribution which determines the shape and β is called the *skewness parameter*.

3 Modelling Financial Markets

The theory in this chapter is mainly a combination of the theory from [63] and [60].

The common characteristic of the models used in this thesis is based on the *"Efficient Market Hypothesis"*. This states that all information which is useful for predicting future price changes is reflected in the current state of the price process. This implies that past information cannot improve the prediction of future prices. Thus, the Efficient Market Hypothesis assumes prices to be *Markov processes*.

Definition 12: Markov Process

A process $X = (X_t)_{t\geq 0}$ is called a Markov process if for all $n \in \mathbb{N} \setminus 0$

$$\mathbb{P}[X_{t_{n+1}}|X_{t_n}, X_{t_{n-1}}, ..., X_{t_0}] = \mathbb{P}[X_{t_{n+1}}|X_{t_n}]$$

The underlying assets I will consider in this thesis will be driven by Lévy processes. Lévy processes make up a large group of models used in finance and are defined as follows [63].

Definition 13: Lévy Process

A real-valued (or $\mathbb{R}^d$-valued) stochastic process $X = (X_t)_{t\geq 0}$ is called a Lévy process if

1. it has **independent increments**, i.e. the random variables $X_{t_0}, X_{t_1} - X_{t_0}, ..., X_{t_n} - X_{t_{n-1}}$ are independent for all $n \geq 1$ and $0 \leq t_0 < t_1 < ... < t_n$.

2. is has **stationary increments**, i.e. $X_{t+h} - X_t$ has the same distribution as X_h for all h, $t \geq 0$.

3. it is stochastically continuous: for every $t \geq 0$ and $\epsilon > 0$

$$\lim_{s\to t} \mathbb{P}\left[|X_s - X_t| > \epsilon\right] = 0.$$

4. the paths $t \to X_t$ are right-continuos with left limits with probability 1.

Note that an immediate consequence of point 1 in Definition 13 is that a Lévy process is a Markov process.

An immediate consequence of point 2 is that $\mathbb{P}[X_{t_0} = 0] = 1$. Inserting this

into (2.9) yields $\phi_{X_{t_0}}(u) = 1$. Furthermore, (2.9) obviously implies that for two independent random variables Δ_1 and Δ_2, $\phi_{\Delta_1+\Delta_2}(u) = \phi_{\Delta_1}(u)\phi_{\Delta_2}(u)$. One understands from this that the process X_t can be expressed as a finite or infinite product of characteristic functions. This characteristic is key for the Path Integration algorithm presented in Section 5.

Moreover, a Lévy process in $\mathbb{R}^d$ can be parameterised by its *Lévy triplet*, (γ, Σ, ν) [63]. γ is a vector in $\mathbb{R}^d$ of drift coefficients, Σ is a matrix in $\mathbb{R}^{d\times d}$ of Brownian coefficients and ν is the *Lévy measure*[10].

Theorem 5: Lévy-Khinchine Representation

Let $X = \{X_t\}_{t\geq 0}$ be a Lévy process in $\mathbb{R}^d$. Then, there are parameters $\gamma \in \mathbb{R}^d$, a positive definite matrix $\Sigma \in \mathbb{R}^{d\times d}$ and a locally finite measure ν on $\mathbb{R}^d \setminus \{0\}$ with $\int_{\mathbb{R}^d\setminus\{0\}} \min(1, |x|^2)\nu(dx) < \infty$ such that

$$\phi_{X_t}(u) = \mathbb{E}\left[e^{iuX_t}\right] = e^{-t\psi(u)} \tag{3.1}$$

where the *characteristic exponent* $\psi(u)$ is given by

$$\psi(u) = -i\gamma^\top u + \frac{1}{2}u^\top \Sigma u - \int_{\mathbb{R}^d} \left(e^{iu^\top x} - 1 - iu^\top x \mathbb{I}_{\{|x|\leq 1\}}\right)\nu(dx), \quad u \in \mathbb{R}^d, \tag{3.2}$$

where $\mathbb{I}$ is an indicator function.

The Lévy measure is on the form $\nu(dx) = u(x)dx$ where $u(x)$ is called the *Lévy density* [60]. The Lévy density is required to have 0 mass at the origin and it does not have to be integrable. Beyond those two requirements, it behaves like a probability density function.
Examining the characteristic exponent (3.2) one recognises the first two terms as a Brownian motion with constant drift. The third term represents the jumps of the process.

In one dimension, $\nu(dx)$ can be interpreted as follows when $\nu(\mathbb{R}) < \infty$. The Lévy process, X_t, makes jumps of size $Y \in \mathbb{R}$ according to a Poisson process with intensity $\nu(Y)$ [60].

The processes I will consider in this thesis will be written on the form

$$S_t = S_0 \exp(X_t)$$

where X_t is a specific Lévy process. The log-returns for these processes will be denoted Z_t and are given by

$$Z_t = \ln \frac{S_t}{S_0} \sim X_t. \tag{3.3}$$

[10]Theorem 5 is a generalisation of Theorem 2.2.9 from [63] from one dimensional Lévy processes to $\mathbb{R}^d$

Some choices for X_t will be thoroughly discussed in this section.
The PIFFT method presented in this thesis is applicable only on Lévy models. The rest of this section is devoted to present some important Lévy models and reflect on their ability to model financial assets.

3.1 Geometric Brownian Motion

The Geometric Brownian Motion (GBM)[60] is the underlying model in the Black and Scholes framework and is therefore widely used. The GBM uses the normal distribution to model log-returns. That is, the choice for X_t, as discussed in the introduction to this section, is a Lévy process with normally distribution increments. We write

$$S_t = S_0 \exp(X_t^{\text{normal}})$$

Definition 14: Geometric Brownian Motion

A stochastic process, S_t, follows a Geometric Brownian Motion if it satisfies the stochastic differential equation

$$dS_t = S_t(\mu dt + \sigma dW_t), \qquad S_0 > 0, \tag{3.4}$$

where μ is the drift, σ is the volatility and W_t is a standard Brownian motion with mean 0 and variance dt.

Note that the GBM satisfies the weak form of the Efficient Market Hypothesis[67] as the current price recursively reflects all past prices.

Applying Ito's Lemma [53] to (3.4) yields the unique solution:

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right).$$

Thus, the log-returns of the underlying for the GBM will follow

$$Z_t \sim \mathcal{N}\left(\left(\mu - \frac{1}{2}\sigma^2\right)t, \sigma^2 t\right). \tag{3.5}$$

Which means the underlying itself is log-normally distributed;

$$S_t \sim \text{lognormal}\left(\left(\mu - \frac{1}{2}\sigma^2\right)t, \sigma^2 t\right).$$

The characteristic function of the normally distributed logarithmic returns, Z_t, in the GBM model is given by

$$\phi_{Z_t}(u) = \exp\left(iu\left(\mu - \frac{1}{2}\sigma^2\right)t - \frac{1}{2}u^2\sigma^2 t\right). \tag{3.6}$$

A major advantage of the GBM is that it results in a convenient closed-form expression of the prices characteristic function (see section 2.5.1).

Comparing (3.6) with (3.1), we observe that the Lévy triplet of the GBM is $\left(\mu - \frac{1}{2}\sigma^2, \sigma^2, 0\right)$. Note that the GBM is a continuous process as it has no jumps.

As discussed in Section 2.2, it is convenient in applications to be able to impose the risk-neutral formulation of the market model. By applying Girsanov's theorem [27], we can see that the risk-neutral formulation of the GBM-model is easily obtained by fixing the drift μ to $r - q$ where r is the risk free rate and q is the continuous dividend yield [46]. (3.4) then becomes

$$dS_t = S_t \left((r - q)dt + \sigma dW_t\right), \quad S_0 > 0 \tag{3.7}$$

and (3.5) becomes

$$Z_t \sim \mathcal{N}\left(\left(r - q - \frac{1}{2}\sigma^2\right)t, \sigma^2 t\right). \tag{3.8}$$

Options with payoffs depending on a stochastic variable evolving according to (3.7), can therefore be priced using the risk-neutral formula.

Alternatively, instead of using Girsanov's theorem one could arrive at the same conclusion by considering a risk free portfolio constructed through "delta hedging" [53]. Arbitrage arguments imply that such a portfolio must have the expected return rate equal to the risk free rate r. Note that delta hedging is only possible for continuous models for the underlying like the GBM.

3.1.1 Weaknesses in Describing Asset Dynamics

The weaknesses of the GBM and the Black-Scholes framework is widely discussed and is a subject in itself. I will only give a very brief indication of where the shoe pinches.

Empirical studies show that the GBM is too simplistic to accurately emulate movements in asset prices because of some rough assumptions. For example the assumption of constant drift and volatility and log-normal asset prices. The tails of the normal distribution are found to be too light to accurately model log-returns. Observations show that large price swings happen more frequently in reality[38] and thus a leptokurtic distribution is preferable. It is also more common with large price swings downwards than upwards, meaning that the real life distribution of log-returns in the financial market is skewed to the left.

Awareness and communication of the weaknesses of a chosen model is imperative. This was experienced under the financial crisis of 2007-2008 where under-communication of model weaknesses led to incorrect pricing of risk[70].

The search for better models gained momentum in the late 1980s and 1990s. One required more sophisticated Lévy processes based on more general distributions than the normal distribution which could handle for example excess kurtosis and

skewness. Resulting models include the Normal Inverse Gaussian (NIG) model and the Variance Gamma (VG) model.

3.2 The Normal Inverse-Gaussian Process

The Normal Inverse-Gaussian (NIG) distribution was initially introduced in 1977 by Barndorff-Nielsen [5]. In the search for more flexible distributions than the normal distribution for modelling financial objects, Barndorff-Nielsen introduced the NIG distribution to finance in 1997 [6]. Unlike the normal distribution, the NIG distribution is flexible in regard to kurtosis and skewness.

There are several parameterisations of the NIG distribution. For continuity with the code used in this thesis, I will follow [7] to define the four parameter NIG distribution, NIG$(\alpha, \beta, \delta, \mu)$:

Definition 15: Normal Inverse-Gaussian Distribution

The normal inverse-Gaussian distribution is defined as the distribution on the whole real line with probability density function

$$f_X(x;\alpha,\beta,\delta,\mu) = \frac{\alpha\delta}{\pi}\exp\left(\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)\right)\frac{K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\sqrt{\delta^2+(x-\mu)^2}}, \tag{3.9}$$

where K_1 is the modified, third order Bessel function with index 1 and $0 \leq |\beta| \leq \alpha$, $\mu \in \mathbb{R}$ and $\delta > 0$.

The modified Bessel function of the third kind with index λ is given by

$$K_\lambda(x) = \frac{1}{2}\int_0^\infty u^{\lambda-1}\exp\left(-\frac{1}{2}x\left(u+u^{-1}\right)\right)du, \quad x > 0. \tag{3.10}$$

For more details on Bessel functions, see [2]. μ is a location parameter (μ is the mean when there is zero skewness), δ is a scale parameter, β is a skewness parameter ($\beta = 0$ implies zero skewness) and α affects the steepness of the probability density function; thus affecting kurtosis.

The NIG distribution has the characteristics:

	$NIG(\alpha, \beta, \mu, \delta)$
mean	$\mu + \delta\beta/\sqrt{\alpha^2 - \beta^2}$
variance	$\alpha^2\delta(\alpha^2 - \beta^2)^{-3/2}$
skewness	$3\frac{\beta}{\alpha\sqrt{\delta}(\alpha^2-\beta^2)^{1/4}}$
kurtosis	$3\left(1 + \frac{\alpha^2+4\beta^2}{\delta\alpha^2\sqrt{\alpha^2-\beta^2}}\right)$

Table 1: NIG-characteristics.

The characteristic function of a NIG distribution with $\mu = 0$ is:

$$\tilde{\phi}_{NIG}(u; \alpha, \beta, \delta) = \exp\left(-\delta\left(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2}\right)\right)$$

The introduction of the drift parameter μ to the characteristic function is achieved by [60]

$$\phi_{NIG}(u; \alpha, \beta, \delta, \mu) = \tilde{\phi}_{NIG}(u; \alpha, \beta, \delta)\exp(iu\mu) \tag{3.11}$$

and thus

$$\phi_{NIG}(u; \alpha, \beta, \delta, \mu) = \exp\left(-\delta\left(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2}\right) + iu\mu\right) \tag{3.12}$$

Barndorff-Nielsen proven the NIG distribution to be infinitely divisible. Thus, one can define the NIG-process

$$X^{NIG} = \left\{X_t^{NIG}, t \geq 0\right\} \tag{3.13}$$

where $X_0^{NIG} = 0$ and with stationary, independent increments $X_{s+t}^{NIG} - X_s^{NIG} \sim NIG(\alpha, \beta, \delta t, \mu t)$.

The Lévy measure for the NIG process is

$$\nu_{NIG}(dx) = \frac{\delta\alpha}{\pi}\frac{\exp(\beta x)K_1(\alpha|x|)}{|x|}dx$$

and γ is given by

$$\gamma = \frac{2\delta\alpha}{\pi}\int_0^1 \sinh(\beta x)K_1(\alpha x)dx.$$

The Lévy triplet of the NIG process becomes $(\gamma, 0, \nu_{NIG})$. Note that the NIG process has no Brownian component and thus the stochastic behaviour of the NIG process is given by a pure jump process.
In theory books one often describes the NIG-process as

$$X_t^{NIG} = \beta\delta^2 I_t + \delta W_{I_t} \tag{3.14}$$

where I_t is an *Inverse-Gaussian process*, $\mathrm{IG}(a, b)$, with $a = 1$ and $b = \delta\sqrt{\alpha^2 - \beta^2}$ and W_t is a standard Brownian motion. The Inverse-Gaussian distribution is the

distribution of the time it takes a a standard Brownian motion with drift b (that is $W_s + bs, s \geq 0$) to reach the positive level $a > 0$.

The IG-distribution has the characteristic function

$$\phi_{IG}(u; a, b) = \exp\Big(- a(\sqrt{-2iu + b^2} - b)\Big)$$

and the explicitly known probability density function

$$f_{IG}(x; a, b) = \frac{a}{\sqrt{2\pi}} \exp(ab) x^{-3/2} \exp\left(-\frac{1}{2}(a^2 x^{-1} + b^2 x)\right), \quad x > 0.$$

The Lévy measure of the IG law is

$$\nu_{IG}(dx) = (2\pi)^{-1/2} a x^{-3/2} \exp\left(-\frac{1}{2} b^2 x\right) \boldsymbol{I}_{(x>0)} dx$$

and

$$\gamma = \frac{a}{b}\left(2\mathbb{N}(b) - 1\right),$$

where $\mathbb{N}(x)$ is the cumulative probability density function of the standard Normal distribution.

3.2.1 Fit to Market

As mentioned, the NIG-distribution provides a more flexible fit to the market than the normal distribution. Defining a NIG market model works similarly as for the GBM. I propose the model

$$S_t = S_0 \exp(X_t^{NIG})$$

for some initial value $S_0 > 0$. Thus, increments are now NIG-distributed

$$\ln S_{s+t} - \ln S_s = X_{s+t}^{NIG} - X_s^{NIG} \sim NIG(\alpha, \beta, \delta t, \mu t).$$

As discussed in Section 2.1, one is often interested in the risk-neutral formulation of the model with the associated risk-neutral probability measure. For the GBM, this was achieved by fixing $\mu = r - q$ in accordance with Girsanov's theorem.
A requirement for applying Girsanov's theorem is that the process is continuous. The NIG-process, however, allows for jumps and is therefore non-continuous. Therefore it is slightly more complicated to transform to the risk-neutral formulation of the NIG-process. There is, in fact no unique transformation.

One option is to use the Esscher transform. This transform is not always applicable and may give more than one possible transform. Another alternative is to use the mean-correcting measure change by manipulating the drift term, μ, according to [60] such that:

$$S_t = S_0 \exp\left(X_t^{NIG}(\alpha, \beta, \delta t, \tilde{\mu} t)\right), \tag{3.15}$$

where $\tilde{\mu} = (r - q) + \omega$ and

$$\omega = \delta \left(\sqrt{\alpha^2 - (\beta + 1)^2} - \sqrt{\alpha^2 - \beta^2} \right).$$

I reintroduce the variable Z_t, the log-returns of the underlying and reformulate (3.15) to

$$Z_t \sim NIG(\alpha, \beta, \delta t, \tilde{\mu} t).$$

Under the mean-correcting measure change, the probability density function of Z_t becomes

$$f^{\mathbb{Q}}_{NIG}(z; t, \alpha, \beta, \delta, \tilde{\mu}) = \frac{\alpha \delta t}{\pi} \exp\left(\delta t \sqrt{\alpha^2 - \beta^2} + \beta(z - \tilde{\mu} t)\right) \frac{K_1(\alpha \sqrt{(\delta t)^2 + (z - \tilde{\mu} t)^2})}{\sqrt{(\delta t)^2 + (z - \tilde{\mu} t)^2})} \tag{3.16}$$

and the associated characteristic function

$$\phi^{\mathbb{Q}}_{NIG}(u; t, \alpha, \beta, \delta, \tilde{\mu}) = \exp\left(-\delta t \left(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2} \right) + iu\tilde{\mu} t \right) \tag{3.17}$$

3.3 The Variance Gamma Process

Introduced by Madan, Carr and Chang in 1998 [13], the three parameter Variance Gamma (VG) process is obtained by evaluating a Brownian motion with constant drift and volatility at a random time change given by a gamma process.

In this section, I will introduce the VG process based on the theory from [60], although with the slight tweak of including a location parameter μ. The location parameter will be integrated into the VG distribution by replacing the variable of the distribution by the Mahalanobis distance [11][15].

I will denote the Variance Gamma distribution by $VG(\sigma, \nu, \theta)$ with σ being the volatility of the Brownian motion, ν the variance rate of the time change and θ the drift of the Brownian motion.
As for the NIG distribution, the Variance Gamma distribution offers more flexibility than the normal distribution by allowing both skewness and excess kurtosis.

I will begin by introducing the Gamma process [13].

Definition 16: Gamma Distribution

The Gamma distribution, $Gamma(a, b)$, with shape a and rate b is defined as the distribution with the probability density function

$$f_{gamma}(t; a, b) = \frac{b^a}{\Gamma(a)} t^{a-1} \exp(-tb), \quad t > 0.$$

[11] Mahalanobis distance: $Q(x) = (x - \mu)' \Sigma^{-1} (x - \mu)$, for some covariance matrix Σ. I will consider the univariate VG distribution. Therefore, the Mahalanobis distance is reduced to $\frac{1}{\sigma^2}(x - \mu)^2$ as seen in (3.20).

The associated Gamma process, $G(t; a, b) = \{G_t, t \geq 0\}$ with $G_0 = 0$, has stationary, independently Gamma distributed increments such that

$$(G_{s+t} - G_s) \sim Gamma(ta, b).$$

The Variance Gamma process can now be defined in terms of a Brownian motion and a Gamma process.

Definition 17: Variance Gamma Process

The Variance Gamma process, $X^{VG} = \{X_t^{VG}, t \geq 0\}$ with $X_0^{VG} = 0$ is defined in terms of a Brownian motion with drift, $b(t; \theta, \sigma) = \theta t + \sigma W_t$ and a Gamma process with unit mean rate, $G(t; 1, \nu)$ as

$$X^{VG}(t; \sigma, \nu, \theta) = b(G(t; 1, \nu); \theta, \sigma), \tag{3.18}$$

where W_t is a standard Brownian motion.

The process has stationary, independent increments

$$X_{s+t}^{VG} - X_s^{VG} \sim VG(\sigma\sqrt{t}, \nu/t, \theta t)$$

The characteristic function of the Variance Gamma distribution is given by

$$\phi_{VG}(u; \sigma, \nu, \theta) = (1 - i\theta\nu u + (\sigma^2\nu/2)u^2)^{-1/\nu} \tag{3.19}$$

and the probability density function is available in closed form as

$$f_{VG}(x; \sigma, \nu, \theta, \mu) = \frac{2\exp((x-\mu)\theta/\sigma^2)}{\nu^{1/\nu}\sqrt{2\pi}\sigma\Gamma(\frac{1}{\nu})} \frac{K_{\frac{1}{\nu}-\frac{1}{2}}\left(\sigma^{-2}\sqrt{(x-\mu)^2\left(\frac{2\sigma^2}{\nu} + \theta^2\right)}\right)}{\left((x-\mu)^2/\left(\frac{2\sigma^2}{\nu} + \theta^2\right)\right)^{\frac{1}{4}-\frac{1}{2\nu}}}, \tag{3.20}$$

Note that a location parameter μ has been introduced as mentioned. Furthermore, $K_{\frac{1}{\nu}-\frac{1}{2}}$ is the modified Bessel function of the third order with index $\frac{1}{\nu} - \frac{1}{2}$. That is

$$K_{\frac{1}{\nu}-\frac{1}{2}}(x) = \frac{1}{2}\int_0^\infty u^{\frac{1}{\nu}-\frac{3}{2}}\exp(-\frac{1}{2}x(u + u^{-1}))du.$$

The characteristics of the Variance Gamma distribution, $VG(\sigma, \nu, \theta)$, are summarised in Table 2.

	$VG(\sigma, \nu, \theta)$
mean	$\mu + \theta$
variance	$\sigma^2 + \nu\theta^2$
skewness	$\theta\nu(3\sigma^2 + 2\nu\theta^2)/(\sigma^2 + \nu\theta^2)^{3/2}$
kurtosis	$3(1 + 2\nu - \nu\sigma^4(\sigma^2 + \nu\theta^2)^{-2})$

Table 2: VG-characteristics.

To facilitate the programming in this thesis, I will consider another parameterisation of the probability density function of the VG distribution. This parameterisation is the one used in the 'ghyp'-package in R. This package delivers functionality for the generalised hyperbolic distribution. The VG and the NIG distribution are both special cases of the generalised hyperbolic distribution. The fitting of distributions to data is performed in this parameterisation. After fitting, I will switch to the parameterisation used in (3.20). The probability density function in (3.21) is in the standard parameterisation of the generalised hyperbolic distribution (See [15] for details).

$$f_{VG}(x;\lambda,\mu,\gamma,\tilde{\sigma}) = \frac{2\lambda^{\lambda}(2\lambda+\gamma^2/\tilde{\sigma}^2)^{\frac{1}{2}-\lambda}}{\sqrt{2\pi}\tilde{\sigma}\Gamma(\lambda)} \exp\left((x-\mu)\gamma/\tilde{\sigma}^2\right) \frac{K_{\lambda-\frac{1}{2}}\left(\sigma^{-2}\sqrt{(x-\mu)^2(2\lambda\sigma^2+\gamma^2)}\right)}{((x-\mu)^2(2\lambda\sigma^2+\gamma^2))^{\frac{1}{2}\left(\frac{1}{2}-\lambda\right)}} \tag{3.21}$$

When switching from the $(\lambda,\mu,\gamma,\tilde{\sigma})$-parameterisation to the (σ,ν,θ,μ)-parameterisation, μ remains the same while the rest of the parameters are given by

$$\sigma \leftarrow \tilde{\sigma} \tag{3.22}$$

$$\theta \leftarrow \gamma \tag{3.23}$$

$$\nu \leftarrow \frac{1}{\lambda}. \tag{3.24}$$

In applications, two alternative formulations of the Variance Gamma process are also often useful.
The first is that the Variance Gamma process can be expressed as the difference between two independent, increasing Gamma processes (See [13] for details):

$$X^{VG}(t;\sigma,\nu,\theta) = G_p(t;\mu_p,\nu_p) - G_n(t;\mu_n,\nu_n), \tag{3.25}$$

where we have the relations to the parameters in (3.18):

$$\mu_p = \frac{1}{2}\sqrt{\theta^2+\frac{2\sigma^2}{\nu}} + \frac{\theta}{2}$$

$$\mu_n = \frac{1}{2}\sqrt{\theta^2+\frac{2\sigma^2}{\nu}} - \frac{\theta}{2}$$

$$\nu_p = \left(\frac{1}{2}\sqrt{\theta^2+\frac{2\sigma^2}{\nu}} + \frac{\theta}{2}\right)^2 \nu$$

$$\nu_n = \left(\frac{1}{2}\sqrt{\theta^2+\frac{2\sigma^2}{\nu}} - \frac{\theta}{2}\right)^2 \nu$$

In the parameterisation with the parameters from (3.25), the Lévy measure is given by

$$\nu_{VG}(x)dx = \begin{cases} \frac{\mu_n^2}{\nu_n}\frac{\exp\left(-\frac{\mu_n}{\nu_n}|x|\right)}{|x|}dx & \text{for } x<0 \\ \frac{\mu_p^2}{\nu_p}\frac{\exp\left(-\frac{\mu_p}{\nu_p}x\right)}{x}dx & \text{for } x>0 \end{cases} \tag{3.26}$$

With the parameters from (3.18), the Lévy measure of the VG process is given by

$$\nu_{VG}(x)dx = \frac{\exp(\theta x/\sigma^2)}{\nu|x|}\exp\left(-\frac{\sqrt{\frac{2}{\nu}+\frac{\theta^2}{\sigma^2}}}{\sigma}|x|\right)dx \tag{3.27}$$

One observes from (3.27) that $\theta = 0$ eliminates the only non-symmetric part of the Lévy measure thus resulting in a symmetric VG process. Negative θ yields higher relative probability for negative x-values than positive x-values. Skewness is therefore controlled by θ where negative θ corresponds to negative skewness. Furthermore, increasing ν results in a decreasing exponential decay rate of the Lévy measure around zero. Therefore, kurtosis is controlled by ν with which it has a positive correlation.

The other mentioned useful formulation comes directly from (3.18). Carr, Madan and Chang [13] showed that the process

$$X_t^{VG} = \theta G\left(t\frac{1}{\nu}, \frac{1}{\nu}\right) + \sigma W_{G\left(t\frac{1}{\nu},\frac{1}{\nu}\right)} \tag{3.28}$$

follows a VG process with the parameters σ, ν and θ.

3.3.1 Fit to Market

As for the NIG-distribution, the VG-distribution provides a more flexible fit to the market than the normal distribution. I propose the model

$$S_t = S_0\exp(X_t^{VG})$$

for some initial value $S_0 > 0$ such that

$$\ln\frac{S_{s+t}}{S_s} \sim VG(\sigma\sqrt{t}, \nu/t, \theta t, \mu t)$$

or equivalently in the parameterisation from (3.21):

$$\ln\frac{S_{s+t}}{S_s} \sim VG(\tilde{\sigma}\sqrt{t}, \lambda t, \gamma t, \mu t).$$

Furthermore, we have that

$$\begin{aligned}\phi_{VG}(u;\sigma\sqrt{t},\nu/t,\theta t,\mu t) &= (\phi_{VG}(u;\sigma,\nu,\theta))^t\exp(iu\mu t)\\ &= (1-i\theta\nu u+(\sigma^2\nu/2)u^2)^{-t/\nu}\exp(iu\mu t)\end{aligned}$$

To impose the risk-neutral measure I introduce the mean-correcting measure change for the VG-process by manipulating the drift, μ.

$$\tilde{\mu} = r - q + \omega_{VG},$$

where

$$\omega_{VG} = \frac{1}{\nu}\ln\left(1-\frac{1}{2}\sigma^2\nu-\theta\nu\right).$$

The risk neutral formulation of the probability density function of the log returns $Z_t = \ln \frac{S_{s+t}}{S_s}$ of the VG-process is then achieved by inserting the time scaled parameters and making the change from μ to $\tilde{\mu}$ in (3.20).
Its equivalent in the parameterisation from (3.21), after some rearranging, is

$$f_{VG}(z;t,\sigma,\lambda,\gamma,\mu) = \frac{2\lambda^{\lambda t}\exp((z-\tilde{\mu})\gamma/\sigma^2)}{\sqrt{2\pi}\sigma\Gamma(\lambda t)} \frac{K_{\lambda t-\frac{1}{2}}\left(\sigma^{-2}\sqrt{(z-\tilde{\mu}t)^2\left(2\lambda\sigma^2+\gamma^2\right)}\right)}{\left((z-\tilde{\mu}t)^2/\left(2\lambda\sigma^2+\gamma^2\right)\right)^{\frac{1}{4}-\frac{1}{2}\lambda t}} \tag{3.29}$$

and the characteristic function

$$\phi_{VG}(u;t,\lambda,\tilde{\mu},\gamma,\tilde{\sigma}) = \left(1-\frac{i\gamma}{\lambda}u+\frac{\tilde{\sigma}^2}{2\lambda}u^2\right)^{-t\lambda}\exp(iu\tilde{\mu}t) \tag{3.30}$$

4 Testing Fit to Market

As seen in Section 3, the NIG and VG distribution offers more flexibility when fitting a model to data than the normal distribution. One expects therefore the fit of the NIG and VG to outperform the fit of the GBM. This section aims to test this hypothesis and provide a first glance at the NIG and VG distribution in action. I will do this by fitting models to some selected markets and evaluating the fit.

4.1 Parameter Estimation

To set parameters for the fitted distributions, I will use built-in R-functions which find the maximum likelihood estimators of the parameters of the respective distributions. I refer the reader to [57] for theory on maximum likelihood estimators. The functions I use are from the "ghyp"-package in R as mentioned in Section 3.3 [16], [15]. The code for fitting and evaluating the models are included in Appendix E.2.2 and E.2.3.
The built-in functions return fitted parameters in the "alpha-bar"-parameterisation. A drawback of this parameterisation is that it is non-existent for the case of $\bar{\alpha} = 0$ and $\lambda \in [-1, 0]$. However, this corresponds to a student-t distribution with non-existing variance which is an unlikely case considering the data I am fitting to. Additionally, the program has a try-and-catch for this rare case and applies a limit if necessary.

The alpha-bar parameterisation consists of the parameters $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$. Their relations to the parameters in (3.29) are included in the ghyp-package. Their relations to the parameters of the parameterisation used by Barndorff-Nielsen for the NIG distribution (and indeed in this thesis in (3.9)), $(\alpha, \mu, \delta, \beta)$, are found by little work (See Appendix B) and are summarised in Table 3.

	$(\alpha, \mu, \delta, \beta)$	$\leftrightarrows$	$(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$
Shape:	α	$=$	$\bar{\alpha} + \gamma$
Location:	μ	$=$	μ
Shape:	δ	$=$	$\sqrt{\Sigma}$
Skewness:	β	$=$	γ

Table 3: Relations between parameterisations for the NIG distribution.

4.1.1 OSEBX

The first market I will consider is the Oslo Stock Exchange Benchmark Index (OSEBX) from 29.04.2013 to 25.04.2018 found at [35]. The OSEBX consists of 82 stocks from mainly major Norwegian companies. Fitting parameters by maximum likelihood estimation yielded the parameters in Table 4.

Normal	$\mu = 0.00045$	$\sigma = 0.00975$		
NIG	$\alpha = 102.4702$	$\mu = 0.00129$	$\delta = 0.00965$	$\beta = -8.86944$
VG	$\lambda = 1.474771$	$\mu = 0.00120$	$\Sigma = 0.00963$	$\gamma = -0.00075$

Table 4: Fitted parameters to the logreturns of the OSEBX index for the NIG, VG and normal distribution.

Table 5 evaluates the fit of the models using the Akaike information criterion and the Log Likelihood. Additionally, Figures 1 and 2 illustrates the fitted distributions and displays a qq-plot.

	AIC	Log likelihood
Normal	-8036.873	4020.437
NIG	-8161.59	4084.795
VG	-8158.093	4083.047

Table 5: Evaluation of fitted distributions for OSEBX market.

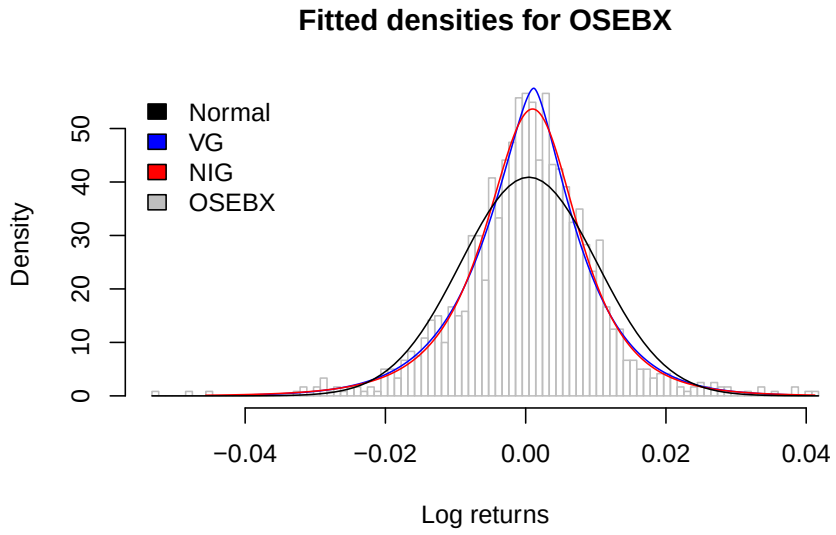


Figure 1: Densities of fitted distributions against a histogram of actual data from OSEBX.

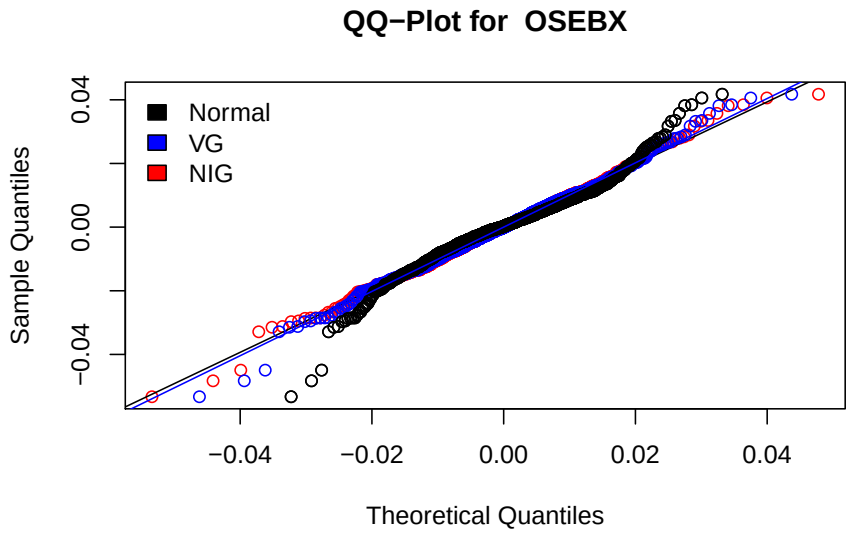


Figure 2: QQ-plot of the fitted distributions for OSEBX.

4.1.2 DAX

The Deutscher Aktienindex (DAX) consists of 30 major German companies [31]. I do the same calculations as for the OSEBX.

Normal	$\mu = 0.00036$	$\sigma = 0.01128$		
NIG	$\alpha = 77.94258$	$\mu = 0.001231431$	$\delta = 0.01011801$	$\beta = -6.739988$
VG	$\lambda = 1.10743$	$\mu = 0.00142$	$\Sigma = 0.01134$	$\gamma = -0.00106$

Table 6: Fitted parameters to the logreturns of the DAX index for the NIG, VG and normal distribution.

	AIC	Log likelihood
Normal	-7747.313	3875.657
NIG	-7864.102	3936.051
VG	-7879.517	3943.758

Table 7: Evaluation of fitted distributions for DAX market.

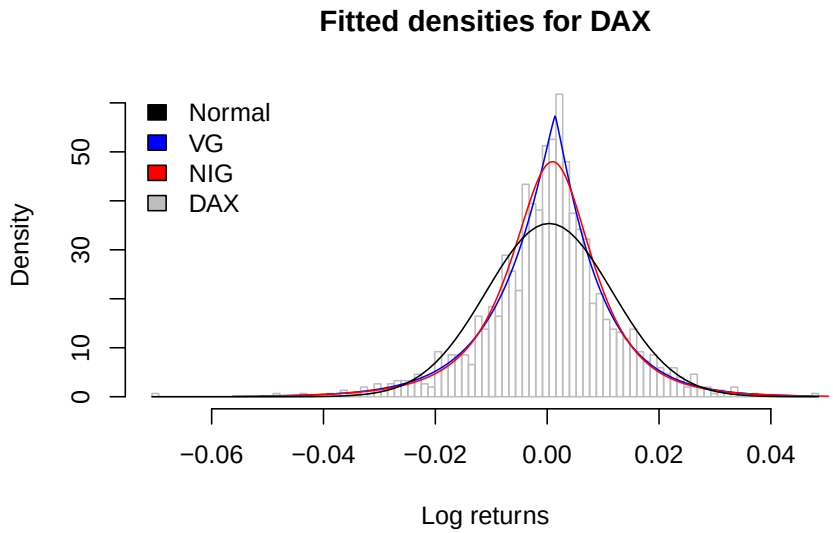


Figure 3: Densities of fitted distributions against a histogram of actual data from DAX.

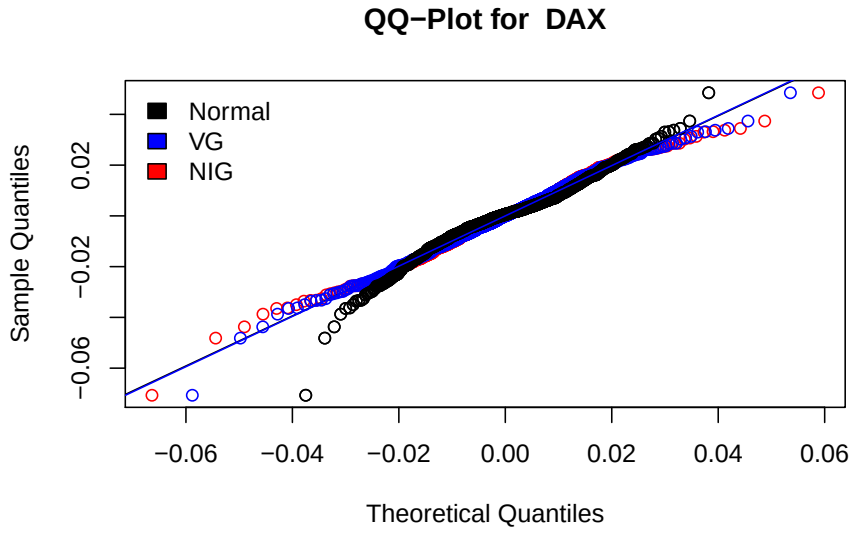


Figure 4: QQ-plot of the fitted distributions for DAX.

4.1.3 Nasdaq 100

The Nasdaq 100 (NDX) [32] consists of 100 major American companies.

Normal	$\mu = 0.00065$	$\sigma = 0.00957$		
NIG	$\alpha = 82.97017$	$\mu = 0.00177$	$\delta = 0.00758$	$\beta = -12.05833$
VG	$\lambda = 1.08930$	$\mu = 0.00150$	$\Sigma = 0.00945$	$\gamma = -0.00085$

Table 8: Fitted parameters to the logreturns of the NDX index for the NIG, VG and normal distribution.

	AIC	Log likelihood
Normal	-8115.697	4059.848
NIG	-8300.992	4154.496
VG	-8299.825	4153.913

Table 9: Evaluation of fitted distributions for NDX market.

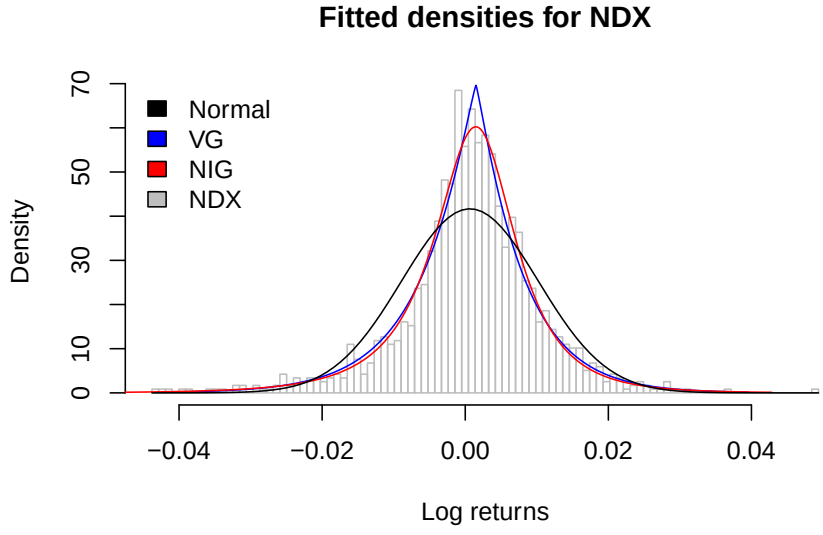


Figure 5: Densities of fitted distributions against a histogram of actual data from NDX.

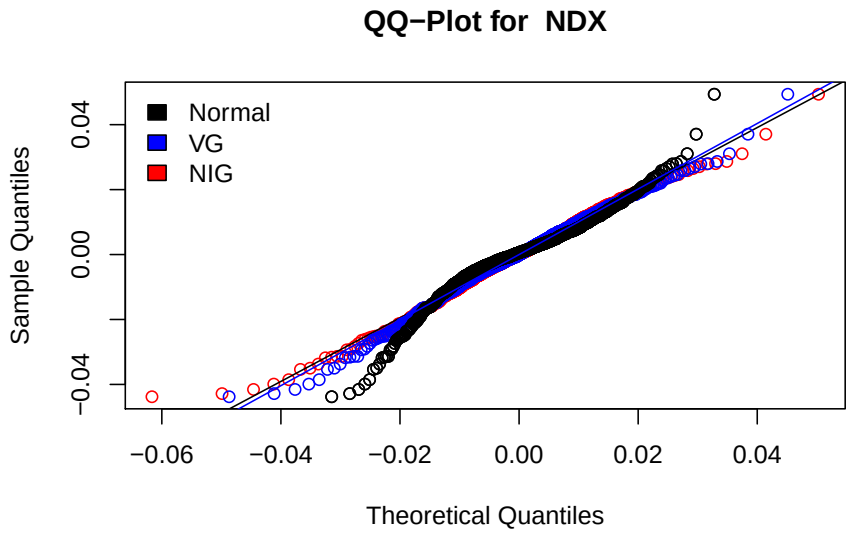


Figure 6: QQ-plot of the fitted distributions for NDX.

4.1.4 OMX Copenhagen 20

The OMX Copenhagen 20 (OMXC20) [33] consists of 20 major Danish companies.

Normal	$\mu = 0.00046$	$\sigma = 0.01086$		
NIG	$\alpha = 95.51966$	$\mu = 0.00136$	$\delta = 0.01116$	$\beta = -7.66917$
VG	$\lambda = 1.50567$	$\mu = 0.00127$	$\Sigma = 0.01075$	$\gamma = -0.00081$

Table 10: Fitted parameters to the logreturns of the OMXC20 index for the NIG, VG and normal distribution.

	AIC	Log likelihood
Normal	-7737.798	3870.899
NIG	-7849.837	3928.919
VG	-7848.505	3928.252

Table 11: Evaluation of fitted distributions for OMXC20 market.

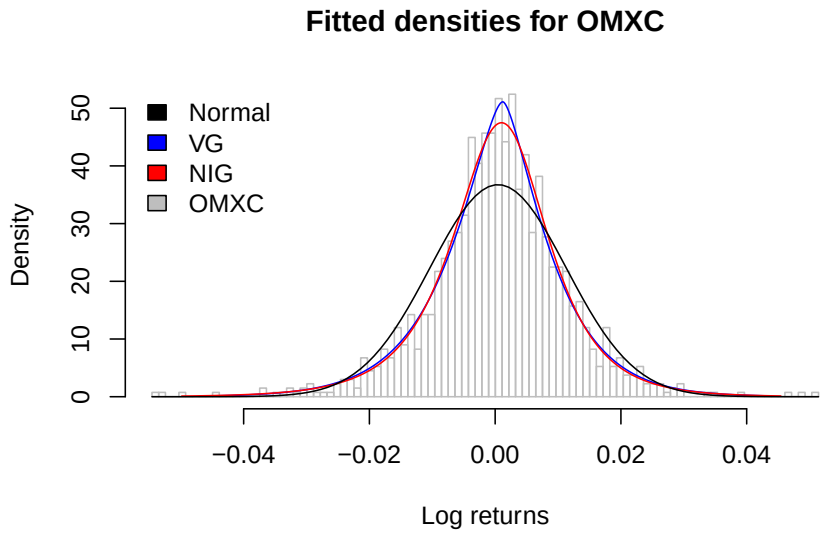


Figure 7: Densities of fitted distributions against a histogram of actual data from OMXC20.

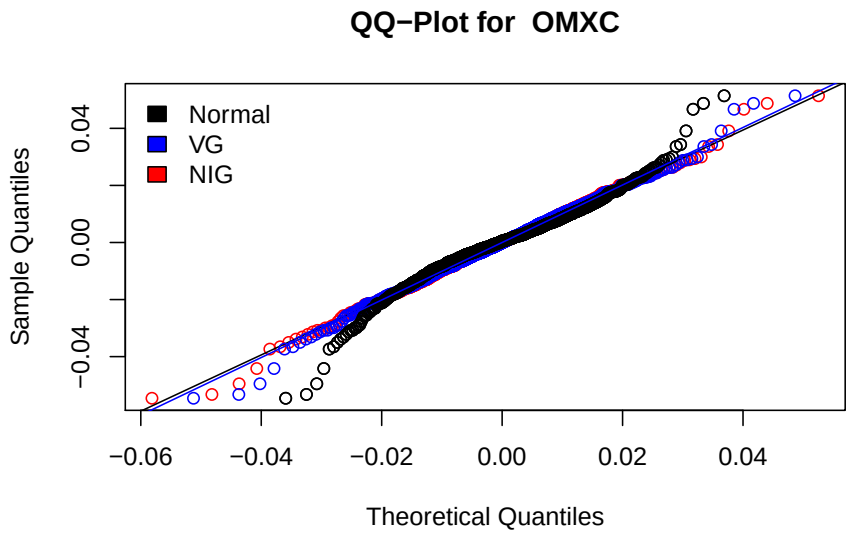


Figure 8: QQ-plot of the fitted distributions for OMXC20.

4.1.5 OMX Stockholm 30

The OMX Stockholm 30 (OMXS30) [34] consists of 30 major Swedish companies.

Normal	$\mu = 0.00021$	$\sigma = 0.01036$		
NIG	$\alpha = 103.5103$	$\mu = 0.00093$	$\delta = 0.01087$	$\beta = -6.85945$
VG	$\lambda = 1.63512$	$\mu = 0.00131$	$\Sigma = 0.01016$	$\gamma = -0.00110$

Table 12: Fitted parameters to the logreturns of the OMXS30 index for the NIG, VG and normal distribution.

	AIC	Log likelihood
Normal	-7885.803	3944.901
NIG	-8012.135	4010.068
VG	-8007.279	4007.64

Table 13: Evaluation of fitted distributions for OMXS30 market.

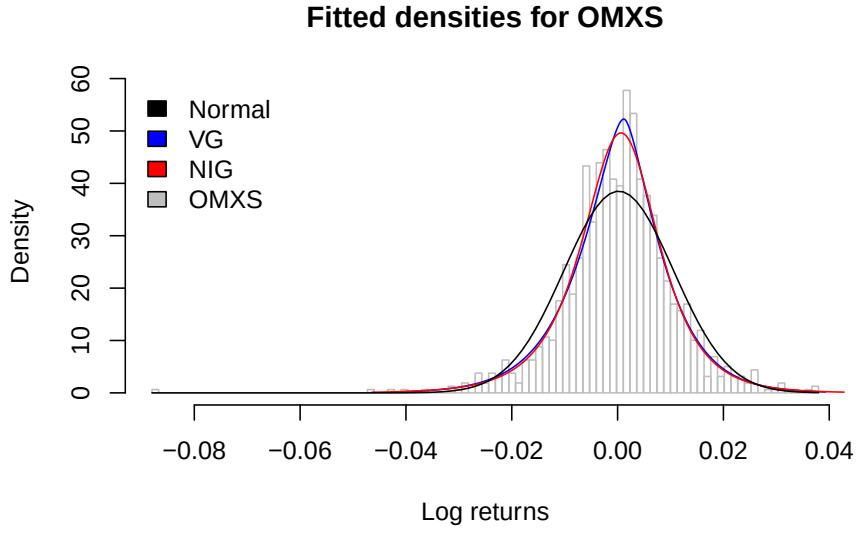


Figure 9: Densities of fitted distributions against a histogram of actual data from OMXS30.

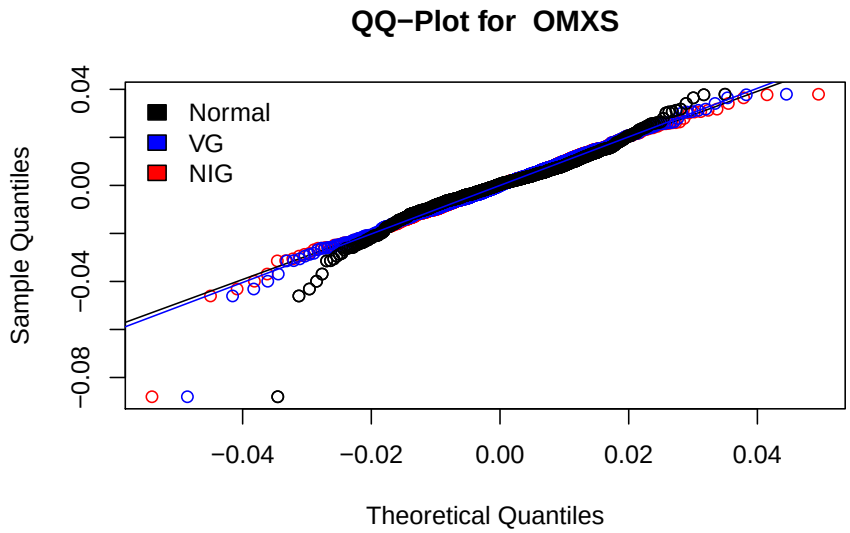


Figure 10: QQ-plot of the fitted distributions for OMXS30.

4.1.6 S&P500

The Standard & Poor's 500 (SPX) [36] consists of 500 large American companies.

Normal	$\mu = 0.00040$	$\sigma = 0.00789$		
NIG	$\alpha = 92.50443$	$\mu = 0.00096$	$\delta = 0.00590$	$\beta = -8.76121$
VG	$\lambda = 0.95276$	$\mu = 0.00055$	$\Sigma = 0.00786$	$\gamma = -0.00016$

Table 14: Fitted parameters to the logreturns of the SPX index for the NIG, VG and normal distribution.

	AIC	Log likelihood
Normal	-8601.168	4302.584
NIG		
VG	-8802.772	4405.386

Table 15: Evaluation of fitted distributions for SPX market.

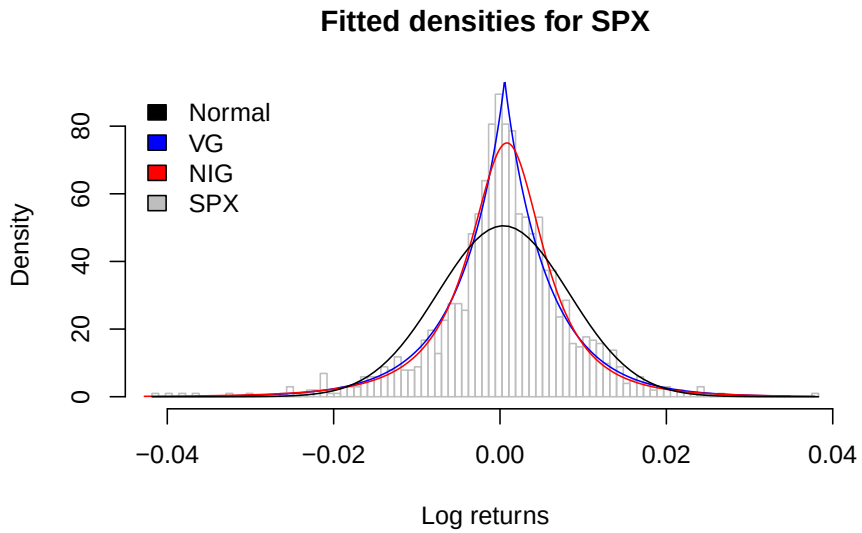


Figure 11: Densities of fitted distributions against a histogram of actual data from SPX.

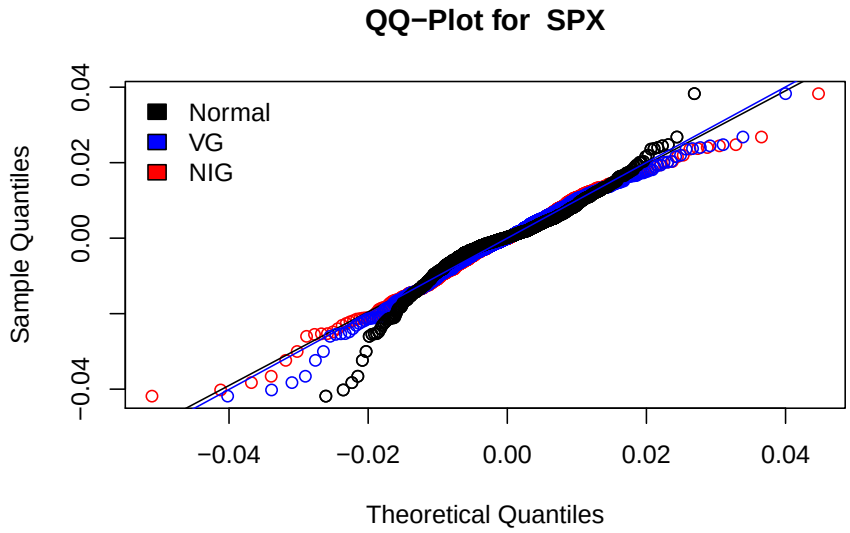


Figure 12: QQ-plot of the fitted distributions for SPX.

Discussion

The density plots clearly indicate that the fit of both the NIG and VG distribution outperforms the fit of the Normal distribution. The QQ-plots provides further evidence. The parameters of the fitted NIG and VG distribution reveals that all the considered markets are slightly skewed. However, the possibility of describing excess kurtosis looks to be more important. This is also reflected by the QQ-plots

where the normal distribution is outperformed in the tails. Therefore, analyses using the NIG and VG distribution will take into account the risk of large price swings better than the normal distribution.

It is difficult to decide which fits the data best of the fitted NIG and VG distribution. One may notice, however, that the NIG distribution performs better than the VG distribution in markets where there is less excess kurtosis (OSEBX, OMXC, OMXS). For the Markets with high kurtosis (DAX, NDX, SPX), the VG distribution looks like the best fit.

Furthermore, all distributions struggles in the left tail for the OMXS market which is the most skewed of the considered markets.

4.2 Goodness-of-Fit

In addition to the statistics and visual evaluations presented above, I will apply one of the most common tests for the *goodness-of-fit* of a distribution to data; namely the Kolmogorov-Smirnov test [23].

The Kolmogorov-Smirnov test considers the largest difference between the cumulative distribution of the fitted distribution and the observed cumulative step function.

Let $S_n(x)$ denote the observed cumulative step function of n data observations. That is

$$S_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathcal{I}(X_i \leq x),$$

where $\mathcal{I}(\cdot)$ is the indicator function. Let $F_0(x)$ denote the cumulative distribution of the distribution fitted to the data. Kolmogorov argues that if $F_0(x)$ truly is the cumulative distribution of the data, then it must be fairly close to $S_n(x)$. The null hypothesis becomes: $H_0 : S_n(x) = F_0(x)$.

For these functions, Kolmogorov provided the following theorem with proof in 1933 [42] which was later proved in a less intricate way by Feller in 1948 [24].

Theorem 6: Kalmogorov Thm. 1

Suppose that $F_0(x)$ is continuous and define the random variable D_n by

$$D_n = \sup_x |S_n(x) - F_0(x)|.$$

Then for every fixed $z \geq 0$ as $n \to \infty$,

$$\mathbb{P}(\sqrt{n}D_n \leq z) \to L(z) \tag{4.1}$$

where $L(z)$, often called the *Kolmogorov distribution function*, is the cumulative distribution function which for $z > 0$ is given by

$$L(z) = 1 - 2\sum_{\nu=1}(-1)^{\nu-1}e^{-\nu^2 z^2}. \tag{4.2}$$

Kolmogorov turns to this theorem for $n > 35$ in [23] which allows me to confidently apply it to my market data with $n > 1250$. The quantiles of $L(z)$ is given in Table 16.

α	0.8	0.9	0.95	0.975	0.99	0.999
K_α	1.0727	1.2238	1.3581	1.4802	1.6276	1.9495

Table 16: Quantiles for the Kolmogorov distribution function.

The null hypothesis is rejected with significance level α if $\sqrt{n}D_n > K_\alpha$; in accordance with (4.1). I perform this test for the fitted NIG, VG and normal distribution. Their values for $\sqrt{n}D_n$ are presented in Table 17.

Market	NIG	VG	Normal
DAX	0.6311	0.4313	2.4052
NDX	0.4939	0.7515	2.9443
OMXC	0.4883	0.3815	1.9466
OMXS	0.5604	0.6882	1.9174
OSEBX	0.4899	0.6245	1.9491
SPX	0.8800	0.8794	3.2049

Table 17: Values for $\sqrt{n}D_n$ for the fitted distributions.

Comparing Table 16 and Table 17, I observe that the normal distribution is rejected for all sensible significance levels, while both the VG and NIG distribution is accepted for all considered significance levels. Based on the test, it is difficult to decide which of the NIG and VG distribution fits the data best as they outperform each other for different markets. However, based on the test, one would conclude that both of them are well suited for modelling financial markets. Plots of the cumulative distributions used in the Kolmogorov-Smirnov test is included in Appendix D.

It is worth noting that drawing conclusions from the Kolmogorov-Smirnov Test about which provides the best fit of the NIG and the VG distributions, lead to different conclusions than those made solely on the density and QQ-plots. This indicates the NIG and VG distribution is similarly suitable for describing log-returns in the financial market.

An alternative, possibly interesting test would be the Anderson-Darling test [62]. The Anderson-Darling test gives more weights to the tails of the distributions which would highlight the value of allowing excess kurtosis. The Anderson-Darling test would therefore probably emphasise the dominance of the NIG and VG distributions over the Normal distribution in modelling financial markets. The disadvantage however, is that one would have to calculate the critical values for the tests for the NIG and VG distribution as these are not available today [28].

4.3 Note on the effect of time

The goal of the presented Lévy processes is to accurately model the price path of some underlying asset. I would like to make a note of an easy-to-miss potential pitfall with regard to this goal. I will consider the DAX as an illustrative example and use the parameters fitted to the DAX. I choose to consider the DAX in particular because it is previously tested and concluded by Rydberg [58] that the NIG process should fit frequently traded stocks (as those included in the DAX) very well compared to less traded stocks for example like those from the Danish market (OMXC).

In academic work in finance, it is common to choose $T = 1$ and $dt = 1/250$ corresponding to a year as time unit and daily measurements (there are approximately 250 business days in a year). I introduce the time step $dt = 1/250$ in the respective fitted distributions and the mean correcting drift term.
The density function of the log returns over the time step dt in the risk-neutral measure is shown in Figure 13 for the NIG and GBM model (the VG model behaves similarly to the NIG model, but is not plotted here because of some numerical difficulties close to the mean for the parameterisation used here[12]).

Note that the NIG (and VG) process obtains a very large excess kurtosis compared to in Figure 3. This is because of the effect of dt on the characteristics of the models as shown in Figure 14.

[12]This is only a problem for the probability density function. There is no problem with drawing VG-distributed numbers or with the characteristic function of the VG distribution. See [15] and [50] for more on this

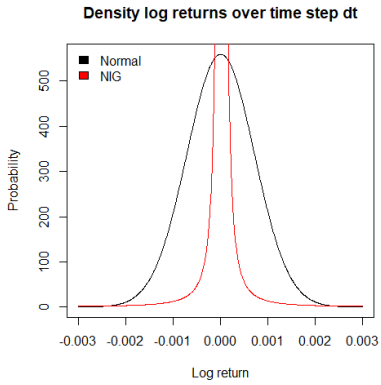
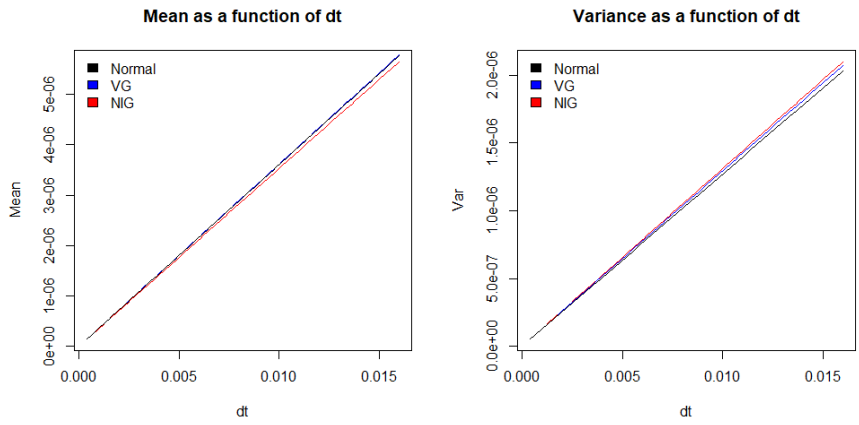
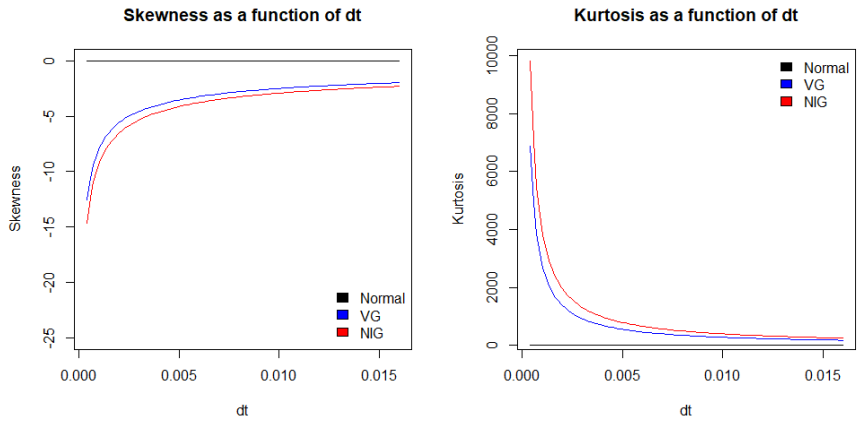


Figure 13: Density of log returns over time step $dt = 1/250$ with fitted DAX-parameters.



(a) Mean

(b) Variance



(c) Skewness

(d) Kurtosis

Figure 14: Model characteristics as a function of the time step dt for the DAX with fitted parameters.

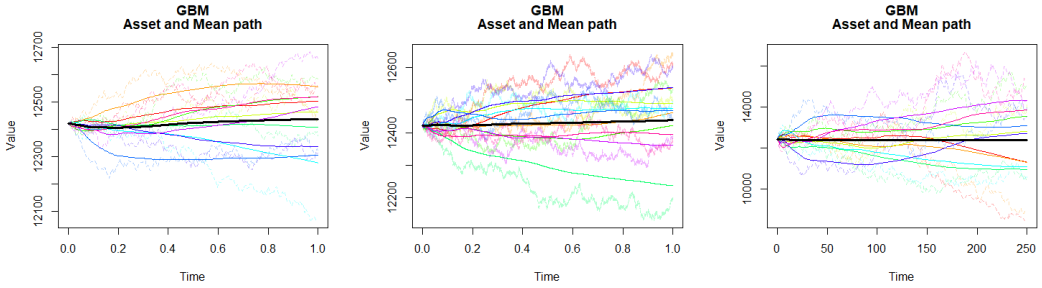
One recognises the large kurtosis when studying individual Monte Carlo simulated price paths; illustrated in Figure 15. Observe that the NIG and VG processes make a lot of very small, practically insignificant jumps before making a very large jump. Consequently, the price paths become too "discrete" to accurately mirror stock movements from the DAX market. Furthermore, this strange behaviour persists when increasing the fineness of the time grid (by decreasing dt by a factor 40). Note, however, that the shape of the price paths from the GBM looks similar for any choices of T and dt as long as the ratio $\frac{T}{dt}$ remains the same.

The strange behaviour of the NIG and VG model for $T = 1$ is of practical importance, but has been overlooked as far as I have seen in other literature. The result of the overly "discrete" behaviour (stemming from overly discrete random time governed by an IG and Gamma process respectively), is that some of the simulated price paths have not "converged" into meaningful, random price paths yet. In fact, some have not yet experienced any significant jumps at all. Consequently, the variance of the NIG and VG process for very small maturity times becomes artificially small. Remember that the price of an option increases with the volatility of the underlying. Consequently, the price of an Asian option with $T = 1$ becomes significantly cheaper under the NIG and VG model compared to the GBM model[13].

The explanation for this is simple. I have fitted distributions to data of daily measurements of the DAX. The time unit in the data is therefore one day and the time step of the data dt_D is one. The choice $T = 1$ therefore corresponds to a maturity time of one day and $dt = 1/250$ corresponds to measurements approximately every five minutes. Clearly, the NIG and VG model is not well suited for options with maturity times so close to the time step of the data. They simply need more time to make enough random jumps to converge into realistic price paths. Keeping this in mind, I accept the time unit of the data and let $dt = 1$ and for example $T = 250$. This would now correspond to an option with maturity time of one year and with daily measurements[14]. We see that all models produce sensible price paths when the maturity time is much larger than the time difference in the data. The prices of Asian options under the different models reflect this as well. The relative difference between the prices is much smaller and the price under the NIG and VG model is slightly higher than under the GBM model. This makes sense because there is a larger probability for large payoffs under the NIG and VG model because of the excess kurtosis.

[13]The option price from the cases with $T = 1$ in Figure 15 was found to be approximately 43 with the NIG and VG model and 46 with the GBM model when doing one million Monte Carlo simulations. The volatility of the Monte Carlo simulated prices were around 0.05.

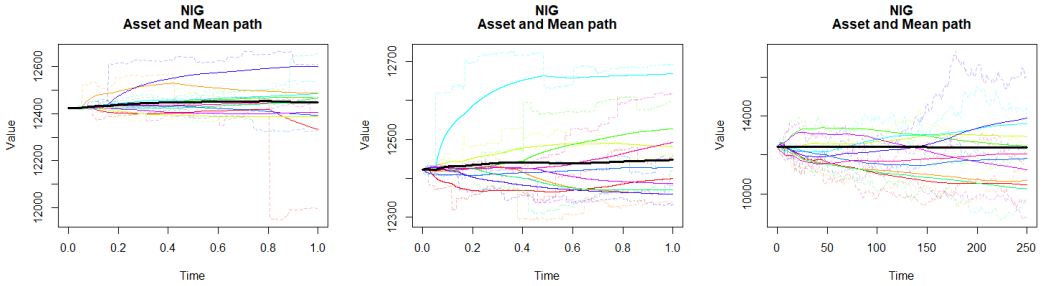
[14]It is important to also remember to scale the risk-free rates accordingly. I have found interest rates with time unit of one year [20]. To scale them to the time unit of one day I divide the rates by 250.



(a) GBM with $dt = 0.004, T = 1$.

(b) GBM with $dt = 1 \cdot 10^{-4}$, $T = 1$.

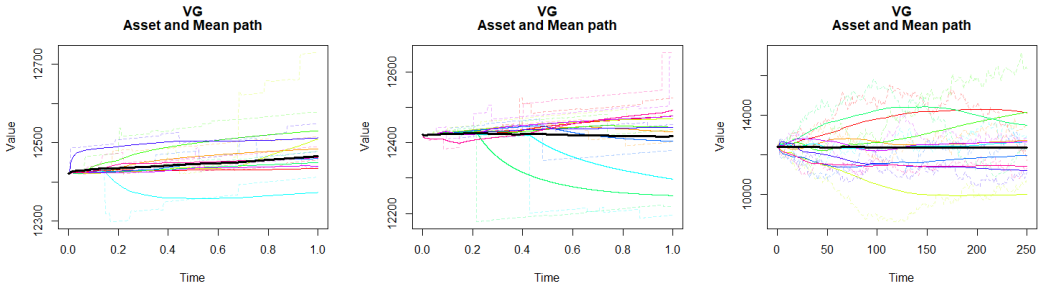
(c) GBM with $dt = 1, T = 250$.



(d) NIG with $dt = 0.004, T = 1$.

(e) NIG with $dt = 1 \cdot 10^{-4}, T = 1$.

(f) NIG with $dt = 1, T = 250$.



(g) GBM with $dt = 0.004, T = 1$.

(h) GBM with $dt = 1 \cdot 10^{-4}, T = 1$.

(i) GBM with $dt = 1, T = 250$.

Figure 15: Price paths for the GBM, NIG and VG model for different choices of T and dt. Dashed lines show price paths, solid lines show running means.

5 Path Integration Method

In this section I will present the Path Integration method and state a Path Integration algorithm using the Fast Fourier transform.

5.1 Path Integration

A path integral is defined as a limit of a sequence of finite-dimensional integrals. Path integration can be thought of as the integral version of the Riemann integral because a Riemann integral is defined as a limit of a sequence of finite sums. Path integrals is a frequently used tool in quantum physics for which it was originally introduced by Richard Feynman in 1942 [25]. In quantum mechanics, a dynamical system can evolve from an initial state to a final state along many different paths. The expected path (average time evolution) is thus of interest in quantum mechanics.

A financial system is similar in many ways. The dynamical asset price can evolve along many different paths from an initial value to some value at maturity. According to the risk-neutral pricing formula, I aim to find the expected price path. To achieve this, I integrate over all possible realisations over the risk-neutral probability measure.
The concept of Path Integration is to propagate the probability density function of an asset price as a continuous stochastic process, $S = \{S_t \geq 0, t \geq 0\}$, forward in time. I will use the law of total probability to find the probability density function of S_{t+1} given the probability density function of S_t thus iteratively calculating the distribution of S_t for a set of equidistant time points $0 < t_1 < \cdots < t_m = T$. In this manner, the probability density function, g, of $S_{t_{i+1}}$ given S_{t_i} is calculated by

$$g_{i+1}(s_{i+1}) = \int_0^\infty p_{i+1|i}(s_{i+1}|s_i) g_i(s_i) ds_i, \text{ for } i = 1, ..., m, \tag{5.1}$$

where $p_{i+1|i}(s_{i+1}|s_i)$ is a known transition probability density of $S_{t_{i+1}}$ given $S_{t_i} = s_i$ and $\Delta t = t_{i+1} - t_i$. Note that if one knows the initial distribution $g_1(s_1)$, one can propagate this distribution forward in time iteratively as was desired. In this way one finds the distribution of S_{t_m} to be

$$g_m(s_m) = \int_0^\infty \cdots \int_0^\infty p_{m|m-1}(s_m|s_{m-1}) \cdots p_{2|1}(s_2|s_1) g_1(s_1) ds_{m-1} \cdots ds_1. \tag{5.2}$$

With the distributions of S_{t_m} known one can apply Theorem 1 to calculate the risk-neutral expectation.

5.2 Path Integration with Fast Fourier Transformation

In order to calculate the distribution of S_{t_m} numerically one must evaluate $m - 1$ integrals. For this task I opt to apply Fourier Analysis[15].

[15]There are numerous alternative approches e.g. the Trapezodial rule, Simpson's method and Gaussian quadrature.

I start by rewriting equation (5.1) as a convolution integral:

$$g_{i+1}(s_{i+1}) = \int_0^\infty p_{i+1|i}(s_{i+1}|s_i)g_i(s_i)ds_i = \int_0^\infty \tilde{p}(s_{i+1} - s_i)g_i(s_i)ds_i \tag{5.3}$$

with $\tilde{p}$ representing an altered density function such that (5.3) is satisfied.
I then consider the logarithmic returns of the asset price and use equation (3.3) to show that:

$$\ln\left(\frac{s_{i+1}}{s_i}\right) = \ln\left(\frac{\frac{s_{i+1}}{s_0}}{\frac{s_i}{s_0}}\right) = \ln\left(\frac{s_{i+1}}{s_0}\right) - \ln\left(\frac{s_i}{s_0}\right) \overset{(3.3)}{=} z_{i+1} - z_i. \tag{5.4}$$

Because of the independent increments of the Geometric Brownian motion, I can introduce the following notation

$$\begin{aligned} f(z_1) &= p_{1|0}(z_1|z_0) \\ f(z_{i+1} - z_i) &= p_{i+1|i}(z_{i+1}|z_i) \\ z_0 &= 0. \end{aligned} \tag{5.5}$$

With this notation, one can write an expression for the distribution of the cumulative logarithmic return of the asset process, S_t, up to maturity, t_m by

$$g_m(z_m) = \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty f(z_m - z_{m-1}) \cdots f(z_2 - z_1)g_1(z_1)dz_{m-1} \cdots dz_1 \tag{5.6}$$

which one recognises as a series of convolution integrals.
Note that implicit in (5.3) and (5.6) is the assumption that the asset process, S, possess the Markov property. That is, the increments $S_{t_{i+1}} - S_{t_i}$ are independent of S_{t_j} for $j \neq i, i+1$. The approach to evaluation of integrals made in this paper is only applicable for asset processes with the Markov property. Many models modelling asset behaviour possess the Markov property which justifies the acceptance of this limitation.

For a path-dependent option, the forward recursive expression (5.6) can be calculated if the initial distribution $g_1(z_1)$ and the transition density $f(z_i - z_{i-1})$ are known. Evaluation of the density $g_m(s_m)$ is made by iterative deployment of the Convolution Theorem (Theorem 8). Thus, one arrive at the following algorithm:
Here ϕ is the characteristic function (see section 2.5.1) of the transition probability density.

For path-dependent options one must perform the inverse transformation in each time step to accord for the dependency structure. For path-independent options, like European- and American options, one can do all iterations in the frequency domain before doing the inverse transformation after completing the loop. Managing the calculation of the g_i with the help of Fourier transforms greatly reduce the computational time required by reducing integration in the price domain to multiplication in the frequency domain. Additionally, as discussed in section A.1, the Fast Fourier transform will reduce the computational time required by algorithm 1 even further.

Algorithm 1: Calculation of g_m

Initialise g_m
for $i = 2 : m$ **do**
  $\mathscr{F}(g_i) = \hat{g}_i(u_i)$
  $\hat{g}_{i+1} = \hat{g}_i \phi$
  $g_{i+1} = \mathscr{F}^{-1}(\hat{g}_{i+1})$
end
Return g_m

Extention to 2D

The same procedure can be applied to bi-variate distributions. The two dimensional equivalent to (5.1) becomes

$$\begin{aligned} g_{i+1}(s_{i+1}, s_{i+2}) &= \int_0^\infty \int_0^\infty p_{i+1|i}(s_{1,i+1}, s_{2,i+1}|s_{1,i}, s_{2,i}) g_i(s_{1,i}, s_{2,i}) ds_{1,i} ds_{2,i} \\ &= \int_0^\infty \int_0^\infty \tilde{p}(s_{1,i+1} - s_{1,i}, s_{2,i+1} - s_{2,i}) g_i(s_{1,i}, s_{2,i}) ds_{1,i} ds_{2,i} \end{aligned}$$

I do the transformation for $s_j = (s_1, s_2)$,

$$\ln\left(\frac{s_{j,i+1} - s_{j,i}}{s_{j,i}}\right) = \ln\left(\frac{s_{j,i+1}}{s_{j,0}}\right) - \ln\left(\frac{s_{j,i}}{s_{j,0}}\right) = z_{j,i+1} - z_{j,i}$$

and arrive at the following expression for the distribution at maturity

$$\begin{aligned} g_m(z_{1,m}, z_{2,m}) = \int_0^\infty \int_0^\infty \cdots \int_0^\infty \int_0^\infty f(z_{1,m} - z_{1,m-1}, z_{2,m} - z_{2,m-1}) \cdots \\ f(z_{1,2} - z_{1,1}, z_{2,2} - z_{2,1}) g_1(z_{1,1}, z_{2,1}) dz_{1,m-1} dz_{2,m-1} \ldots dz_{1,1} dz_{2,1} \end{aligned} \tag{5.7}$$

6 Implementation

When pricing Asian options one must consider both the evolution of the asset price and the running mean. To keep track of the evolution of the asset price I consider the log-returns, $Z_i = \ln \frac{S_i}{S_0}$, for $i = 1, ..., m$. I let A_i denote the average at time t_i. Using the joint density, f_{Z_m, A_m}, the price of a fixed strike Asian option can be expressed as

$$\Pi_{t=0} = e^{-rT} \int_0^\infty \int_{-\infty}^\infty \max(y_m - K, 0) f_{Z_m, A_m}(z_m, y_m) dz_m dy_m. \tag{6.1}$$

The arithmetic average at time t_i of some asset, S_j for $j = 0, ..., i$ is given by

$$A_i = \frac{1}{i+1} \sum_{k=0}^{i} S_k = \frac{S_0}{i+1} \sum_{k=0}^{i} e^{Z_k}.$$

This can be expressed as a running average as

$$A_i = \frac{i-1}{i} A_{i-1} + \frac{S_0}{i} e^{Z_i} \text{ for } i = 2, ..., m.$$

With $A_0 = S_0$ and $A_1 = \frac{S_0}{2}\left(1 + e^{Z_1}\right)$. That is, A_i is given by a deterministic relation to Z_i. Consequently, I can define the transition probability for the process $(Z_i, A_i)_{i=0}^m$ as

$$f_{Z_i, A_i | Z_{i-1}, A_{i-1}}(z_i, y_i | z_{i-1}, y_{i-1}) = \delta\left(y_i - \frac{i-1}{i} y_{i-1} - \frac{S_0}{i} e^{z_i}\right) p_{i|i-1}(z_i | z_{i-1}),$$

where δ denotes the Dirac measure[16] and $p_{i|i-1}(z_i|z_{i-1})$ is the transition probability for $(Z_i)_{i=0}^m$ as defined in (5.5). By the law of total probability, the joint density is thus given by

$$\begin{aligned} f_{Z_i, A_i}(z_i, y_i) = \int_0^\infty \int_{-\infty}^\infty \delta\left(y_i - \frac{i-1}{i} y_{i-1} - \frac{S_0}{i} e^{z_i}\right) p_{i|i-1}(z_i | z_{i-1}) \\ f_{Z_{i-1}, A_{i-1}}(z_{i-1}, y_{i-1}) dz_{i-1} dy_{i-1} \end{aligned} \tag{6.3}$$

I can now perform the integration in the y-dimension to simplify the expression of the joint probability from a double integral to a single integral. Integrating the y-dependent terms in (6.3) yields

$$\begin{aligned} &\int_0^\infty \delta\left(y_i - \frac{i-1}{i} y_{i-1} - \frac{S_0}{i} e^{z_i}\right) f_{Z_{i-1}, A_{i-1}}(z_{i-1}, y_{i-1}) dy_{i-1} \\ &\qquad = \frac{i}{i-1} f_{Z_{i-1}, A_{i-1}}(z_{i-1}, \frac{i}{i-1} y_i - \frac{S_0}{i-1} e^{z_i}) \end{aligned}$$

[16] Dirac measure: For any set B,

$$\delta(B) = \begin{cases} 1, & \text{if } 0 \in B \\ 0, & \text{if } 0 \notin B \end{cases} \tag{6.2}$$

Consequently, the joint density can be expressed as

$$f_{Z_i,A_i}(z_i,y_i) = \frac{i}{i-1}\int_{-\infty}^{\infty} f(z_i - z_{i-1}) f_{Z_{i-1},A_{i-1}}(z_{i-1}, \frac{i}{i-1}y_i - \frac{S_0}{i-1}e^{z_i})dz_{i-1} \quad (6.4)$$

Observe that (6.4) is a recursive expression which propagates the probability density one step forward in time and yields the joint density at t_i as a convolution integral over z_{i-1}. In (6.4), I have used the notation from (5.5) to write $p_{i|i-1}(z_i|z_{i-1})$ as $f(z_i - z_{i-1})$. The transition density $f(z_i - z_{i-1})$ is model specific and available in closed form for the GBM, NIG and VG model. They are easily obtained by setting $t = \Delta t = t_i - t_{i-1}$ in (3.8), (3.16) and (3.29) respectively. For simplicity, Δt is assumed to be constant for all i.

Algorithm 1 requires the characteristic functions of the transition densities to be known in closed form. These are all found by setting $t = \Delta t$ in (3.6), (3.17) and (3.30):

$$\begin{aligned}
\Phi^{GBM}_{z_i - z_{i-1}}(u) &= \exp\left(iu\left(r - q - \frac{1}{2}\sigma^2\right)\Delta t - \frac{1}{2}u^2\sigma^2\Delta t\right), \\
\Phi^{NIG}_{z_i - z_{i-1}}(u) &= \exp\left(-\delta\Delta t\left(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2}\right) + iu\tilde{\mu}\Delta t\right), \\
\Phi^{VG}_{z_i - z_{i-1}}(u) &= \left(1 - \frac{i\gamma}{\lambda}u + \frac{\tilde{\sigma}^2}{2\lambda}u^2\right)^{-\Delta t\lambda} \exp(iu\tilde{\mu}\Delta t).
\end{aligned}$$

It is necessary to determine some initial joint density for Z_i and A_i. Given initial conditions $z_0 = 0$ and $y_0 = S_0$ one obtain the initial joint density

$$f_{Z_0,A_0}(z_0,y_0) = \delta(z_0)\delta(y_0)$$

This distribution is clearly singular which is inconvenient numerically. Inspired by Skaug and Næss [11], I aim to derive a non-singular initial density. Note that because of the deterministic relationship between Z_1 and A_1, the joint density $f_{Z_1,A_1}(z_1,y_1)$ is singular in the y-dimension. Therefore, I consider the joint density after two time steps which will be non-singular. In the following, I derive an expression for $f_{Z_2,A_2}(z_2,y_2)$ and plan to use this as the starting point for the recursions in (6.4). The law of total probability yields that

$$f_{Z_2,A_2}(z_2,y_2) = f_{A_2|Z_2}(y_2|z_2) f_{Z_2}(z_2), \quad (6.5)$$

where

$$f_{Z_2}(z_2) = \int_{-\infty}^{\infty} f_{Z_2|Z_1}(z_2|z_1) f_{Z_1}(z_1) dz_1 \quad (6.6)$$

and the conditional density $f_{A_2|Z_2}(y_2|z_2)$ can be obtained by a variable transformation. Because A_2 is dependent on Z_1 and Z_2 only, there exists some function $w(Z_1)$ such that $A_2|Z_2 = w(Z_1)$.
The average after two time steps, A_2, is given by

$$A_2 = \frac{S_0}{3}(1 + e^{Z_1} + e^{Z_2}) \quad (6.7)$$

It is obvious from (6.7) that the function $w(Z_1)$ is one-to-one which is true if and only if the inverse of $w(Z_1)$ exists. That is, $Z_1 = w^{-1}(A_2|Z_2)$. I denote $w^{-1}(A_2|Z_2)$ as $u(A_2|Z_2)$. Rearranging terms in (6.7) yields

$$Z_1 = u(A_2|Z_2) = \ln\left(\frac{3A_2}{S_0} - 1 - e^{Z_2}\right) \tag{6.8}$$

Consequently, one has that

$$z_1 = u(y_2|z_2) = \ln\left(\frac{3y_2}{S_0} - 1 - e^{z_2}\right)$$

and

$$\frac{\delta u(y_2|z_2)}{\delta y_2} = \frac{3}{3y_2 - S_0 - S_0 e^{z_2}}.$$

Using the well-known transformation formula now yields

$$f_{A_2|Z_2}(z_2|y_2) = f_{Z_1}\Big(u(y_2|z_2)\Big)\left|\frac{\delta u(y_2|z_2)}{\delta y_2}\right|, \tag{6.9}$$

where $f_{Z_1}(z_1) \sim \mathcal{N}\Big(\mu_{Z_1} = (\mu - \frac{1}{2}\sigma^2)\Delta t, \sigma_{Z_1} = \sigma^2\Delta t\Big)$ for the GBM.

In order to state the joint density $f_{Z_2,A_2}(z_2, y_2)$, I also need to express (6.6) by y_2 and z_2 only. The independent increments of the Lévy models imply that $f_{Z_2|Z_1}(z_2|z_1) = f_{Z_1}(z_2 - z_1)$ (see (5.5)). Equation (6.6) becomes

$$f_{Z_2}(z_2) = \int_{-\infty}^{\infty} f_{Z_1}(z_2 - z_1) f_{Z_1}(z_1) dz_1. \tag{6.10}$$

Now, using (6.8) to substitute[17] z_1 with $u(y_2|z_2)$ in $f_{Z_1}(z_2 - z_1)$ yields

$$\begin{aligned} f_{Z_2}(z_2) &= f_{Z_1}(z_2 - u(y_2|z_2)) \int_{-\infty}^{\infty} f_{Z_1}(z_1) dz_1 \\ &= f_{Z_1}(z_2 - u(y_2|z_2)) \\ &= \frac{1}{\sqrt{2\pi}\sigma_{Z_1}} \exp\left(-\frac{1}{2}\frac{z_2 - u(y_2|z_2) - \mu_{Z_1}}{\sigma^2_{Z_1}}\right) \end{aligned}$$

which now only contain z_2 and y_2 as variables.
The joint density, $f_{Z_2,A_2}(z_2, y_2)$, from (6.5) is now given by

$$f_{Z_2,A_2}(z_2, y_2) = f_{Z_1}\left(\ln\left(\frac{3y_2}{S_0} - 1 - e^{z_2}\right)\right)\left|\frac{3}{3y_2 - S_0 - S_0 e^{z_2}}\right| f_{Z_1}\left(z_2 - \ln\left(\frac{3y_2}{S_0} - 1 - e^{z_2}\right)\right). \tag{6.11}$$

Using (6.11) as starting point for the recursions in (6.4) one can use Algorithm 1 to find the joint density at maturity. At this point, one can apply any numerical integration scheme to (6.1) to price the option. I will use a built in function for the trapezoidal rule to evaluate the integration in (6.1).

[17] Note that this is neither a transformation of random variables (I still consider f_{Z_1})) or a substitution of integral variable (I still integrate over dz_1).

7 Numerical Results

In addition to implementing the pricing method described in section 6, I implemented a pricing method based on Monte Carlo simulations for verification of the "PIFFT-prices". Both codes were written in both MATLAB and R. Distributions are fitted to data in R where the "ghyp"-package is available. The PIFFT-pricing requires interpolation for which I have used the built-in function "interp2". This function has been translated from MATLAB to R, but the translation is much slower than the original. Therefore, the pricing is performed in MATLAB. The majority of the codes are included in Appendix E. Implementation in a lower-level language like C# and Java would yield much faster results. However, the aim is to derive and implement the PIFFT method for GBM, NIG and VG dynamics and compare it to the conventional Monte Carlo simulations. This is achieved in MATLAB and R. The functions were run on a computer with Intel Core i7-6600U CPU @ 2.60-2.8 GHz with 16 GB RAM.

I have used the fitted parameters for the DAX and considered strike values equal to $N = \{0.99S_0, S_0, 1.01S_0\}$. The numerical representations of the probability density functions and the transition probabilities are set on an $N \times N$-grid. The fast Fourier transform is most efficient on grids where N is power of 2. With the current code, my computing equipment limits my choices of N to $\{256, 512, 1024, 2048, 4096\}$ (or smaller grids). I performed both 1 million and 5 million Monte Carlo simulations to ensure converged results and to compare runtimes for different accuracy levels. The option price from the Monte Carlo simulations will have some standard deviation. Therefore, I also computed a 95 % confidence interval (CI) of the Monte Carlo prices by performing the pricing procedure 100 times. The results are summarised in the following tables.

Strike \N	128	256	512	1024	2048	4096
12298,1	2321,4213	293,8632	293,8396	293,8518	293,8504	293,8508
12422,3	2214,7912	227,2636	227,2524	227,2590	227,2577	227,259
12546,5	2111,3409	171,5270	171,5286	171,5296	171,5306	171,5307
Time (sec)	0,871	1,919	8,316	37,160	218,467	570,475

Table 18: Values of Asian options for the given parameters under the GBM model.

Strike \N	MC	CI	
12298,1	293.4762	292.0418	294.9105
12422,3	226.8964	225.6124	228.1804
12546,5	171.1875	170.0637	172.3112
Time (sec)	50.2277		

Table 19: Values of Asian options for the given parameters under the GBM model found by Monte Carlo simulations.

Strike \N	MC	CI	
12298.1	293.5583	293.2630	293.8537
12422.3	226.8827	226.7630	227.3023
12546.5	171.3337	171.0897	171.5777
Time (sec)	390.700		

Table 20: Values of Asian options for the given parameters under the GBM model found by 5 million Monte Carlo simulations.

Strike \N	256	512	1024	2048	4096
12298,1	374,5544	296,3830	296,3861	296,3846	296,3847
12422,3	306,1695	229,4202	229,4175	229,4162	229,4171
12546,5	248,3562	173,3334	173,3249	173,3259	173,3257
Time (sec)	1,113	4,643	20,573	111,007	595,022

Table 21: Values of Asian options for the given parameters under the NIG model.

Strike \N	MC	CI	
12298,1	296.1317	294.7474	297.5160
12422,3	229.1588	227.9190	230.3986
12546,5	173.0742	172.0118	174.1366
Time (sec)	53.158		

Table 22: Values of Asian options for the given parameters under the NIG model found by Monte Carlo simulations.

Strike \N	MC	CI	
12298.1	296.1354	295.7614	296.5095
12422.3	229.1594	228.8199	229.4989
12546.5	173.0722	172.7716	173.3728
Time (sec)	401.080		

Table 23: Values of Asian options for the given parameters under the NIG model found by 5 million Monte Carlo simulations.

Strike \N	256	512	1024	2048	4096
12298,1	574,2738	294,8053	295,5703	295,3597	295,3487
12422,3	504,4907	228,8755	228,6386	228,4125	228,4025
12546,5	444,1492	173,6401	172,5962	172,3612	172,3501
Time (sec)	1,113	4,643	20,573	111,007	595,022

Table 24: Values of Asian options for the given parameters under the VG model.

Strike \N	MC	CI	
12298,1	295.0672	293.5469	296.5875
12422,3	228,1431	227,4306	228,8556
12546,5	172.0797	170.9097	173.2496
Time (sec)	57,233		

Table 25: Values of Asian options for the given parameters under the VG model found by Monte Carlo simulations.

Strike \N	MC	CI	
12298.1	295.0282	294.6411	295.4154
12422.3	228.0854	227.7428	228.4280
12546.5	172.0482	171.7739	172.3626
Time (sec)	423.020		

Table 26: Values of Asian options for the given parameters under the VG model found by 5 million Monte Carlo simulations.

7.1 Experimentation

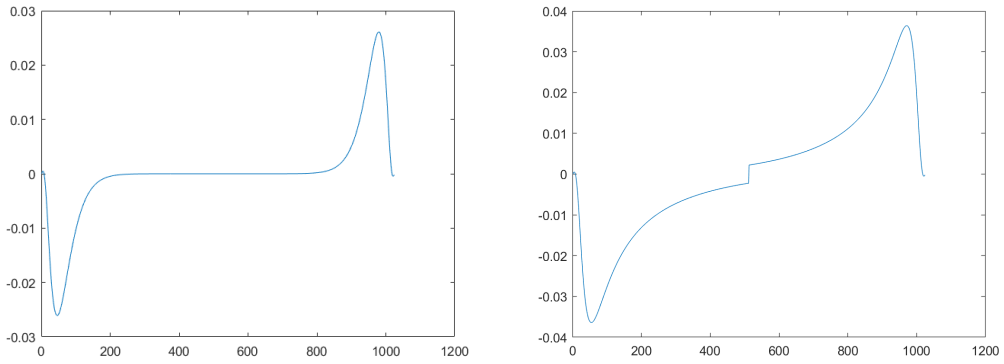
The results presented above are the conclusive results obtained after some numerical tricks to achieve convergence. Initially, there were some problems with pricing under the VG model. Numerical difficulties with the VG model have been experienced in several theses and articles in the past - most often resulting in applying only the NIG model. I too experienced some difficulties with the VG model.

The initial results, without any numerical tricks, under the VG model are given in Table 27. There is clearly no convergence.

Strike \N	256	512	1024	2048	4096
12298,1	2762,7405	2549,1258	2482,7738	2623,5155	2281,6870
12422,3	2649,7535	2445,4906	2375,3518	2508,9662	2178,6244
12546,5	2537,4608	2342,4221	2268,8926	2395,5219	2076,7844
Time (sec)	1,088	4,617	20,268	114,496	577,765

Table 27: Results under the VG model without numerical tricks.

Searching for the source of the numerical issues with the VG model, I compared the characteristic functions of the VG and NIG distribution with the fitted DAX parameters. The imaginary part of the characteristic functions are plotted in Figure 16.

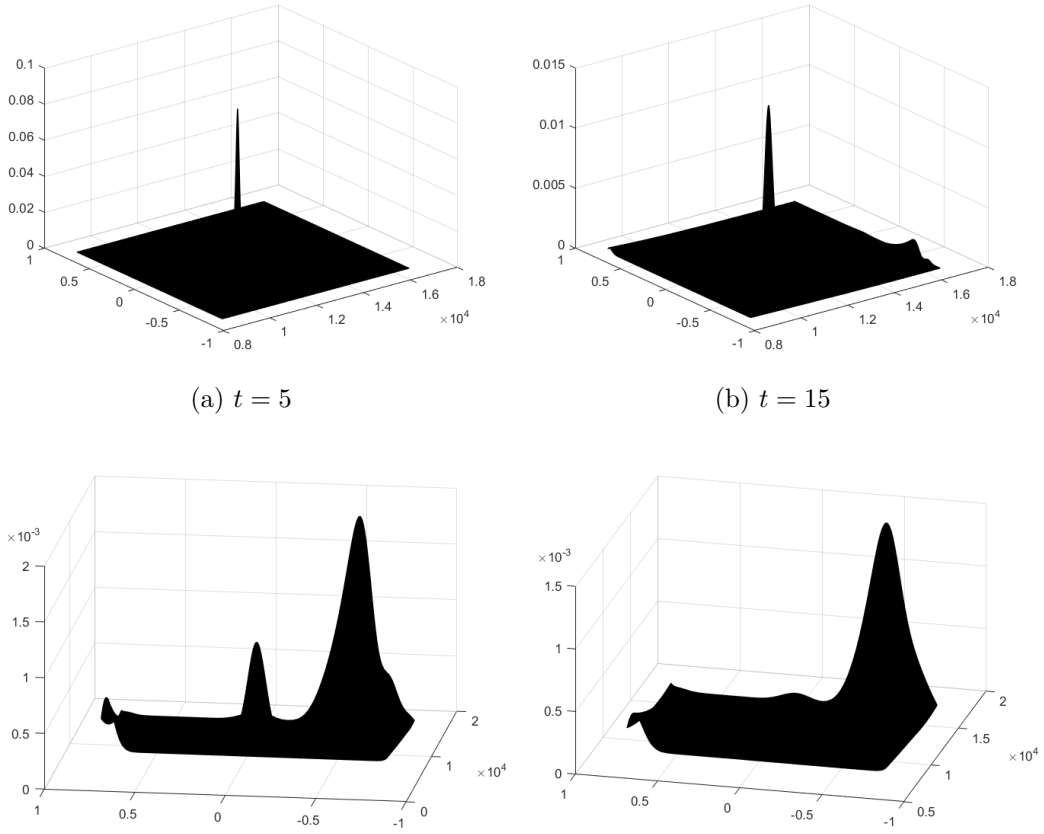


(a) NIG (b) VG

Figure 16: $\mathcal{I}m(\Phi)$ for the NIG and VG distribution with the fitted DAX parameters and $N = 1024$.

The imaginary part of the characteristic function of the VG distribution clearly has a discontinuity which may cause numerical problems. Although reluctant to tamper with the distribution, several experimental alterations were tested. For example Local Regression Smoothing [37] around the discontinuity. The smoothing somewhat improved the results. However, the most effective solution turned out to be to set an area of the probability density function far out in the tails in the z-direction to zero.

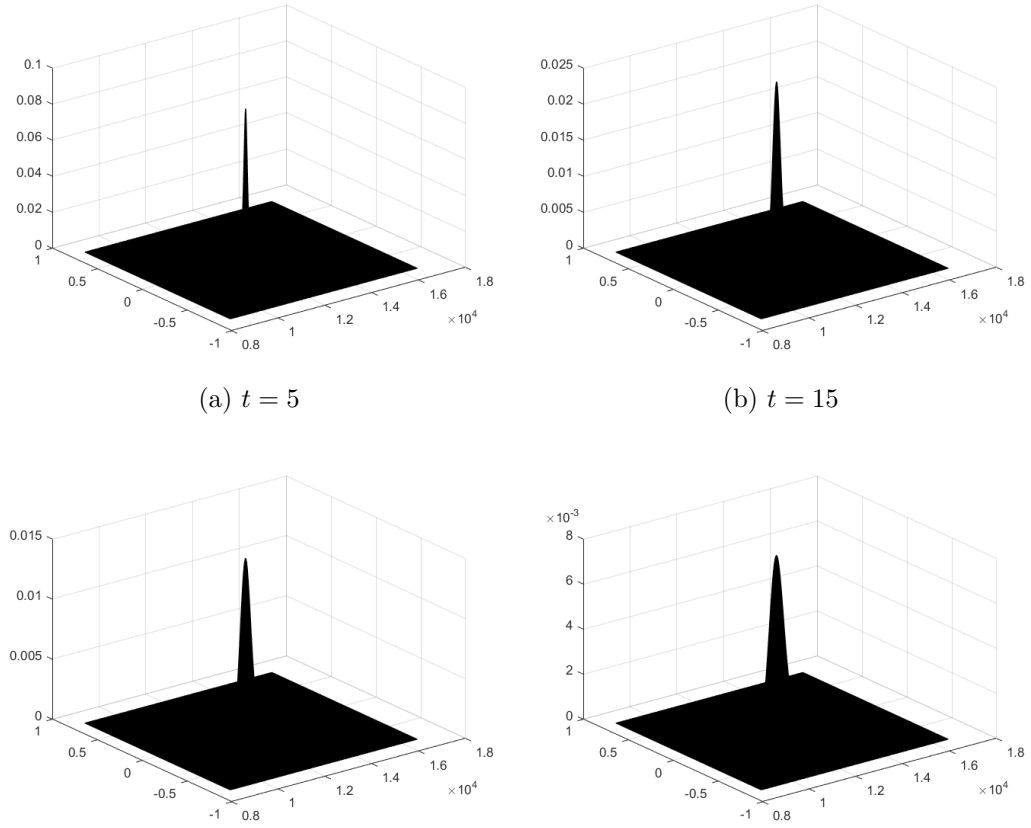
This alteration was motivated from observing the propagated probability density function after each time step. Figure 17 displays the propagated probability density function at some key time steps. One observes that the numerical problems start in the tails at the corners of the grid and that the problems increase with each time step. Numerical problems at the edges of a grid is a common problem in numerical mathematics.
Controlling the tails after each time step stops the problems from increasing before they become significant. Special care was taken to ensure that the controlled area was far enough out in the tails to be statistically insignificant for the purpose of discovering the mean of the distribution. Figure 18 shows the corresponding controlled propagation to Figure 17.



(a) $t = 5$ (b) $t = 15$

(c) $t = 25$ (d) $t = 45$

Figure 17: Propagated probability density function at some key time steps (DAX parameters, $N = 1024$).



(a) $t = 5$ (b) $t = 15$

(c) $t = 25$ (d) $t = 45$

Figure 18: Propagated probability density function at some key time steps with controlled tails (DAX parameters, $N = 1024$).

As one would expect, option prices when controlling the tails are slightly lower than when not controlling the tails because option prices increase with the possibility of large returns. However, this effect can be made as small as desired by expanding the grid and moving the controlled area further out in the tails. By testing the control alternation on the NIG-prices where I get meaningful values with and without controlling the tails, I find its effect on the price to be around 0.037 with the current grid size. The calculated prices when controlling the tails stay within the 95 % confidence interval found by Monte Carlo simulations. Furthermore, when controlling the tails of the NIG distribution, I also achieved converged prices for smaller grids. Specifically, for $N = 512$. The prices when controlling are given in Table 21. When not controlling, the price under NIG dynamics for $N = 512$ had not converged (and was calculated to 309.4829). Interestingly, similarly as for the VG distribution, the characteristic function of the NIG distribution features a discontinuity for N equal to or smaller than 512. Under these conditions, the propagation of the probability density function of the NIG behaves similarly as in Figure 17. Note that the discontinuity in the characteristic function does not disappear for the VG distribution even for $N = 8192$. The GBM distribution first gets a discontinuity in the characteristic function for

$N = 128$. However, meaningful prices under GBM dynamics were achieved for $N = 256$ only when controlling the tails.

Lastly, I would like to add that in order to test the robustness of the PIFFT method under GBM, NIG and VG, the method was also tested for different maturity times and time steps and different markets with equaly satisfying results. It is worth mentioning that the effect of small maturity times, as discussed in Section 4.3, was also observed with the PIFFT method.

8 Discussion

The results in this project thesis undoubtedly present Path Integration combined with the Fast Fourier transform as a worthy of mentioning alternative to Monte Carlo simulations for pricing Asian options. Furthermore, the NIG and VG model is found to fit log returns of stocks prices much better than the traditional GBM.

When comparing Monte Carlo simulations to PIFFT, any PIFFT-prices within the confidence intervals from the MC simulations should be accepted. Therefore, PIFFT under GBM dynamics prices Asian options on the DAX over 26 times faster than 1 million Monte Carlo simulations achieve. Under NIG dynamics, the PIFFT method is over 11 times faster than 1 million Monte Carlo simulations. PIFFT under VG dynamics is about 2.8 times faster than 1 million Monte Carlo simulations.
When considering the Monte Carlo results after 5 million simulations, PIFFT calculated acceptable prices over 10 times faster than the simulations under GBM dynamics. Under NIG dynamics, PIFFT was over 86 times faster and under VG dynamics, PIFFT was about 4 times faster. The phenomenal performance of the PIFFT method under NIG dynamics compared to Monte Carlo simulations is probably partly because of luck with the limits of the confidence interval. The results in Table 21 suggests that $N \geq 1024$ is required to accurately calculate option prices under the given parameters. In this case, PIFFT is around 20 times faster than 5 million Monte Carlo simulations under NIG dynamics.
The size of the confidence interval of the option prices found by 1 million Monte Carlo simulations is arguably too large. The results when performing 5 million simulations are more acceptable. Therefore, the comparison with 5 million Monte Carlo simulation is the most informative.

Furthermore, an additional advantage of the presented method is how intuitive it is and easy to grasp under the assumption of basic understanding of the Fourier transformation. The recursions in (6.4) for example, are merely consecutive applications of the Law of Total Probability while I have also used the transformation formula, the definition of an expected value and some probability theory. Although the required statistical theory is basic, Path Integration with the Fast Fourier transform also requires some numerical theory. For instance in expressing the recursions as convolution integrals such that the Fast Fourier transform can be used. This is perhaps why it is less popular with practitioners than Monte Carlo simulations which mainly demands statistical skills and little numerical skill.

Generally, the choice of grid and its fineness is of great importance. Convergence was not achieced on the roughest grids. Meanwhile, the run times increased rapidly with finer grids. Reducing run times is at the essence of numerical theory. One cannot avoid the numerous transformations and inverse transformations, but one could improve the speed of the interpolations in each time step used in computing the joint density between grid points. Any improvements in the run time

of these computations will greatly affect the run time of the method because they are performed in each time step. For example, one could use interpolation once again to reduce the number of grid points at which computations are performed. One could perform computations at some subset of the grid and obtain values at the grid points outside the subset by interpolation.

Furthermore, one could argue that a grid expanding with time would be useful. Figure 15 shows that the probability mass is more centred around the mean at early time steps before becoming more smeared out at later time steps. An initially smaller grid would imply fewer grid points in total over all time steps. When attempting to improve the grid, it is important to remember that the Fast Fourier transformation requires equidistant grids. This is a major disadvantage of the PIFFT as for example intervals with inverse linear boundaries can be very useful - especially when working with distributions with excess kurtosis [60].

It should also be noted that the PIFFT method was shown to perform very well for a range of other options in earlier theses. For example Barrier and Lookback [50] and Spread options [65]. In fact, flexibility must be regarded as one of the advantages of the method. The requirements for the method is that the transition probability preserves the Markov property and that its characteristic function is known in closed form (the transition probability itself does not need to be known in closed form). This makes it possible to implement the method for Lévy processes which is a large class of market models. It also implies great freedom of choice of the transition probability and that the transition probability can be altered for each time step. Thus, the method is flexible regarding complex path-dependent structures and makes it possible to change the distribution parameters in each iteration. In finance, this is particularly useful because for example Δt is often non-constant (due to weekends and similar) and empirical studies show that also the volatility of an asset is non-constant (motivating the introduction of stochastic volatility [60]).

The swiftness of the method for Barrier and Spread options compared to Monte Carlo simulations was demonstrated in 2014[65]. While runtimes for Barrier and Spread options were approximately 50 and 40 times better than the equivalent Monte Carlo run times, Asian options prove to be more demanding. An important reason for this is that Asian options require the inverse Fast Fourier transform to be performed in each iteration in order to calculate the running mean. Barrier and Spread options can do all time steps in the frequency domain before applying the inverse Fast Fourier transform once at maturity. This suggests that when considering higher dimensional derivatives, Path Integration is more suitable for derivatives with dimensions that are independent (e.g. Spread Options on assets).

Furthermore, informally one can think of Path Integration and Monte Carlo simulations doing something very similar, only in a different order. Monte Carlo simulation considers one realisation over the entire time interval in each iteration

while Path Integration considers all possibilities at the current time step in each iteration. In this regard, Path Integration is conceptually very satisfying in the sense that one covers the entire sample space in one completion of the algorithm compared to Monte Carlo simulations which will not cover the entire sample space in finite time. Consequently, the price computed by Monte Carlo simulations will always have some uncertainty while the Path Integration method yields one sharp value. However, for option pricing this difference is of little importance. In risk analysis for example, where tail behaviour is essential, this difference can be important.

It is also important to note that the run time of the Monte Carlo simulations possibly could be reduced by a more thorough investigation into the minimal number of realisations required to obtain meaningful confidence intervals. The easiest, but also most time consuming approach would be to do this empirically. Alternatively, one could make a more theoretical approach. Monte Carlo simulations use the mean estimator for the option price. For the GBM model, one could derive the theoretical variance of this estimator by realising that the mean estimator consists of a scaled sum of a scaled sum of log-normally distributed random variables[18]. Then, one could perform simulations until the empirical variance of the mean estimator[19] becomes satisfyingly close to the theoretical variance. However, the properties of a sum of non-independent log-normally distributed random variables are quite complicated and have not been extensively investigated in the work on this thesis. I refer the reader to the work of Hcine and Bouallegue [45] for more on the properties of a sum of correlated log-normal random variables.

Comparing the PIFFT method to other methods in the Transform and Quadrature class, the CONV and COS method are the most comparable. The CONV method is very similar, but propages the payoff function forward in time instead of the density function. The immediate availability of the payoff function with the CONV method makes it well suited for pricing options with the possibility of early exercise. The CONV method also exploits the efficiency of the Fast Fourier transform. Additionally, the COS method also uses the Fast Fourier transform and is the most efficient method of the three. However, the COS method is also the most complicated and is therefore less likely to become popular with practitioners.

[18] The stochastic part of the price of an Asian option is the average at maturity. The average is a scaled sum the log-normally distributed asset prices. The mean estimator used in the Monte Carlo simulations is again a scaled sum of these prices.

[19] $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

Bibliography

[1] P. W. Duck D. P. Newton A. D. Andricopoulos, M. Widdicks. *Universal option valuation using quadrature methods.* In *Journal of financial Studies*, pages 447–471. Taiwan Finance Association, 2003.

[2] Stegun Abramowitz. *Handbook of Mathematical Functions.* United States Department of Commerce, 1968.

[3] J. M. Aldaz. *Sharp bounds for the difference between the arithmetic and geometric means.* `https://link.springer.com/content/pdf/10.1007%2Fs00013-012-0434-7.pdf`. [Downloaded: 25.09.2017].

[4] AoPS. *Proofs of AM-GM.* `http://artofproblemsolving.com/wiki/index.php?title=Proofs_of_AM-GM`. [Downloaded: 25.09.2017].

[5] O. E. Barndorff-Nielsen. *Exponentially Decreasing Distributions for the Logarithm of Particle Size.* In *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, Vol. 353*, pages 401–419. Royal Society, 1977.

[6] O. E. Barndorff-Nielsen. *Normal Inverse Guassian Processes and the Modelling of Stock Returns*, 1994.

[7] Ole E. Barndorff-Nielsen. *Normal Inverse Gaussian Distribution and Stochastic Volatility Modelling.* In *Scandinavian Journal of Statistics, Vol 24, No. 1*, pages 1–13. Wiley, 1997.

[8] T. Björk. *Arbitrage theory in continuous time.* Oxford University Press, 2004.

[9] Boyle. *Option valuation using a three-jump process.* In *International Options Journal, 3(7-12)*, pages 7–12. Scientific Research, 1986.

[10] P. P. Boyle. *Options: A Monte Carlo Approach.* In *Journal of Financial Economics*, pages 323–338. North-Holland Publishing Company, 1976.

[11] A. Næ ss C. Skaug. *Pricing of Asian Options by Numerical Path Integration*, 2005.

[12] A. Næ ss C. Skaug. *Fast and Accurate Pricing of Discretely Monitored Barrier Options by Numerical Path Integration.* Springer, 2007.

[13] E. C. Chang D. B. Madan, P. Carr. *The Variance Gamma Process and Option Pricing.* In *European Finance Review 2*, pages 79–105. Oxford University Press, 1998.

[14] D. P. Newton D. Chen, H. J. Hãrkônen. *Advancing the universality of quadrature method to any underlying process for option pricing.* In *Journal of Financial Economics.* North-Holland Publishing Company, 2014.

[15] W. Breymann D. Luethi. *ghyp: A package on generalized hyperbolic distributions*. https://cran.r-project.org/web/packages/ghyp/vignettes/Generalized_Hyperbolic_Distribution.pdf. [Downloaded 01.05.2018].

[16] W. Breymann D. Luethi. *Package 'ghyp'*. https://cran.r-project.org/web/packages/ghyp/ghyp.pdf. [Downloaded 26.04.2018].

[17] J. Dash. *Path Integrals and Options-I Parts I, II*. Centre National de la Recherche Scientifique, 1989.

[18] P.A.M. Dirac. *The Lagrangian in Quantum Mechanics*. http://www.hep.anl.gov/czachos/soysoy/Dirac33.pdf. [Downloaded: 24.10.2017].

[19] J. C. Stein E. M. Stein. *Stock price distributions with stochastic volatility: An analytical approach*. In *Review of financial studies 4(4)*, pages 727–752. Oxford Academic, 1991.

[20] Trading Economics. *Indicators - Countries*. https://tradingeconomics.com/countries. [Downloaded 04.06.2018].

[21] M. Scholes F. Black. *The pricing of options and corporate liabilities*. In *The journal of political economy*, pages 637–654. The University of Chicago Press, 1973.

[22] C. W. Oosterlee F. Fang. *A novel pricing method for European options based on Fourier-cosine series expansions*. In *Journal on Scientific Computing*, pages 826–848. SIAM, 2008.

[23] Jr. F. J. Massey. *The Kolmogorov-Smirnov Test for Goodness of Fit*. In *Journal of the American Statistical Association, Vol. 46, No. 253*, pages 68–78. American Statistical Association, 1951.

[24] W. Feller. *On the Kolmogorov-Smirnov limit theorems for empirical distributions*. In *The Annals of Mathematical Statistics, Vol. 19*, pages 177–189. Institute of Mathematical Statistics, 1948.

[25] R. P. Feynman. *The Principle of Least Action in Quantum Mechanics - Ph.D Thesis Princeton University*. PhD thesis, Princeton University, New Jersey, 1942.

[26] Fincad. *Asian Options*. http://www.fincad.com/resources/resource-library/wiki/asian-options. [Downloaded: 25.09.2017].

[27] I. V. Girsanov. *On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures*. In *Theory of Probability and Its Applications Vol. 5 (3)*, page 285–301. The Society for Industrial and Applied Mathematics, 1960.

[28] Engineering Statistics Handbook. *Anderson-Darling Test*. https://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htm. [Downloaded 22.05.2018].

[29] D. B. Hertz. *Risk Analysis in Capital Investment.* In *Harvard Business Review, 42(1)*, pages 95–106. Harvard Business Publishing, 1964.

[30] S. L. Heston. *A closed-form solution for options with stochastic volatility with applications to bond and currency options.* In *Review of financial studies, 6(2)*, pages 327–343. Oxford University Press, 1993.

[31] Investing.com. *DAX (GDAXI).* `https://www.investing.com/indices/germany-30-historical-data`. [Downloaded 26.04.2018].

[32] Investing.com. *Nasdaq 100 Index.* `https://www.investing.com/indices/nq-100-historical-data`. [Downloaded 26.04.2018].

[33] Investing.com. *OMXC20 Index.* `https://www.investing.com/indices/omx-copenhagen-20-historical-data`. [Downloaded 26.04.2018].

[34] Investing.com. *OMXS30 Index.* `https://www.investing.com/indices/omx-stockholm-30-historical-data`. [Downloaded 26.04.2018].

[35] Investing.com. *OSE Benchmark.* `https://www.investing.com/indices/ose-benchamrk-historical-data`. [Downloaded 26.04.2018].

[36] Investing.com. *S&P 500 Index.* `https://www.investing.com/indices/us-spx-500-historical-data`. [Downloaded 26.04.2018].

[37] Investopedia. *Filtering and Smoothing Data.* `https://se.mathworks.com/help/curvefit/smoothing-data.html`. [Downloaded: 07.07.2018].

[38] Investopedia. *How To Circumvent Limitations Of Black-Scholes Model.* `http://www.investopedia.com/articles/active-trading/041015/how-circumvent-limitations-blackscholes-model.asp?ad=dirN&qo=investopediaSiteSearch&qsrc=0&o=40186`. [Downloaded: 09.10.2017].

[39] Investopedia. *What kinds of derivatives are types of contingent claims?* `http://www.investopedia.com/ask/answers/052715/what-kinds-derivatives-are-types-contingent-claims.asp`. [Downloaded: 11.10.2017].

[40] S. A. Ross J. C. Cox. *The valuation of options for alternative stochastic processes.* In *The journal of financial economics, 3(1)*, pages 145–166. North-Holland Publishing Company, 1976.

[41] T. S. Kleppe. *Numerical Path Integration for Lévy Driven Stochastic Differential Equations.* Master's thesis, Norwegian University of Science and Technology, 2006.

[42] A. Kolmogoroff. *Sulla determinazione empirica di una legge di distribuzione.* In *Giornale dell'Istituto Italiano degli Attuari, Vol. 4*, pages 1–11. Istituto italiano degli attuari, 1933.

[43] E. Kreyszig. *Advanced Engineering Mathematics*. John Wiley , Sons, INC., 10. edition, 2011.

[44] V. Linetsky. *The Path Integration Approach to Financial Modelling and Options PricingPath Integrals and Options-I Parts I, II*. In *Computational Economics, 11(1-2)*, pages 129–163. Kluwer Academic Publishers, 1998.

[45] R. Bouallegue M. B. Hcine. *Highly Accurate Log Skew Normal Approximation to the Sum of Correlated Lognormals*. `https://arxiv.org/ftp/arxiv/papers/1501/1501.02347.pdf`, 2014. [Downloaded: 24.10.2017].

[46] M. Rutkowski M. Musiela. *Martingale Methods in Financial Modelling 2. ed.* Springer, 2005.

[47] Math3Ma. *Dominated Convergence Theorem*. `http://www.math3ma.com/mathema/2015/10/11/dominated-convergence-theorem`. [Downloaded: 12.10.2017].

[48] MathWorks. *2-D fast Fourier transform*. `https://se.mathworks.com/help/matlab/ref/fft2.html?requestedDomain=www.mathworks.com#bvhcnas`. [Downloaded: 27.09.2017].

[49] Weiss Montroll. *Random walks on lattices II*. In *Journal of Mathematical Physics 6(2)*, pages 167–181. American Institute of Physics, 1965.

[50] J. H. Østreim. *Pricing of Lookback Options with NIG and VG Dynamics, Using the Numerical Path Integration Method*. Master's thesis, Norwegian University of Science and Technology, 2009.

[51] C. O'Sullivan. *Path dependent option pricing under Lévy processes*. In *EFA 2005 Moscow Meetings Paper*. University College Dublin, 2005.

[52] D. Madan P. Carr. *Option valuation using the fast Fourier transform*. In *Journal of computational finance*, pages 61–73. Infopro Digital, 1999.

[53] J. Dewynne P. Wilmott, S. Howison. *The Mathematics of Financial Derivatives*. Cambridge University Press, 1. edition, 2009.

[54] P. Tankov R. Cont. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC: London, 2004.

[55] F. Bervoets C. W. Oosterlee R. Lord, F. Fang. *A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes*. In *Journal on Scientific Computing 30(4)*, pages 1678–1705. SIAM, 2008.

[56] E. Reiner. *Convolution methods for path-dependent options*. UBS Warburg Dillon Read, 2000.

[57] S. M. Ross. *Introduction to Probability Models*. Academic Press, 11 edition, 2014.

[58] T. H. Rydberg. *The normal inverse gaussian lévy process: simulation and approximation.* In *Communications in Statistics. Stochastic Models, 13:4*, pages 887–910. Taylor & Francis Group, 1997.

[59] K. Sato. *Lévy Processes and Infinitely Divisable Distributions.* Cambridge University Press, 1999.

[60] W. Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives.* Wiley Online Library, 2003.

[61] M. A. Sullivan. *Valuing American put options using Gaussian quadrature.* In *Review of financial Studies 13(1)*, pages 75–94. Oxford University Press, 2000.

[62] D. A. Darling T.W. Anderson. *Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes.* In *Annals of Mathematical Statistics. Vol. 23*, pages 193–212. Institute of Mathematical Statistics, 1952.

[63] S. Mulinacci P. Rossi U. Cherubini, G. Della Lunga. *Fourier Transform Methods in Finance.* Wiley, 2010.

[64] M. B. Urfjell. *Fast and Accurate Pricing of Asian Options Using Numerical Path Integration - With the Fast Fourier Transform*, 2017. Unpublished Project Thesis at NTNU. Arvid Næss as supervisor.

[65] A. R. Aadland V. L. Devold, E. T. Sæ bœ. *Pricing Exotic Options Using Numerical Path Integration - with Applications to Counterparty Credit Risk*, 2014. Unpublished Project Thesis at NTNU. Arvid Næss as supervisor.

[66] S. Wang. *A novel fitted finite volume method for the Black-Scholes equation governing option pricing.* In *IMA Journal of Numerical Analysis 24(4)*, pages 699–720. School of Mathematics & Statistics, The University of Western Australia, 2004.

[67] N. Wijst. *Finance - A Quantitative Introduction.* Cambridge University Press, 1. edition, 2013.

[68] Wikipedia. *Dominated Convergence Theorem.* `https://en.wikipedia.org/wiki/Dominated_convergence_theorem`. [Downloaded: 12.10.2017].

[69] Wikipedia. *Fast Fourier transform.* `https://en.wikipedia.org/wiki/Fast_Fourier_transform`. [Downloaded: 27.09.2017].

[70] Wikipedia. *Financial Crisis of 2007-2008.* `https://en.wikipedia.org/wiki/Financial_crisis_of_2007%E2%80%932008#Incorrect_pricing_of_risk`. [Downloaded: 09.10.2017].

[71] D. Williams. *Probability with Martingales.* Cambridge University Press, 1991.

[72] O. Pironneau Y. Achdou. *Computational Methods for Option Pricing.* Siam, vol. 30 edition, 2005.

[73] T. Zhang. *Martingale Theory for Finance*. http://www.maths.manchester.ac.uk/~tzhang/publications/Math67201-47201notes-4.pdf. [Downloaded: 11.10.2017].

Appendices

A Fourier Analysis

This section gives a brief introduction to the Fourier transform and the Fast Fourier transform which will be central in estimating the characteristic function of the transition probability for stochastic differential equations driven by Levy processes.

The Fourier transformation transforms a function from its original domain, e.g. x, to the frequency domain, u, by representing the function as an infinite sum of sine- and cosine terms with distinct frequencies. The Fourier transform, $\mathscr{F}(\cdot)$, of a function $f(x)$ and its inverse transform is defined as follows [43].

Definition 18: Fourier Transform and its Inverse

$$\begin{aligned} \mathscr{F}(f) = \hat{f}(u) &= \int_{-\infty}^{\infty} f(x)e^{-iux}dx \\ \mathscr{F}^{-1}(\hat{f}) = f(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(u)e^{iux}du \end{aligned} \tag{A.1}$$

For the Fourier transform of a function $f(x)$ to exist, the following conditions must be fulfilled

Theorem 7: Existence of the Fourier Transform

Let the function $f(x)$ be piecewise continuous on every finite interval and absolutely integrable on the x-axis, i.e.

$$\int_{-\infty}^{\infty} |f(x)|dx < \infty. \tag{A.2}$$

Then, the Fourier transform $\hat{f}(u) = \mathscr{F}(f(x))$ exists.

Note that a probability density function, $P(x)$ have that $P(x) \geq 0$ for all x and $\int_{-\infty}^{\infty} P(x)dx = 1$. Thus, the Fourier transform exists for all piecewise continuous probability density functions.

The Fourier transform can be extended into higher dimensions. Because Asian options are two dimensional, with the price of the underlying and its average being the two dimensions, it is sufficient to define the 2 dimensional Fourier transformation.

Definition 19: 2D Fourier Transform and its Inverse

$$\begin{aligned} \hat{f}(u_1, u_2) &= \int_{\infty}^{\infty} f(x_1, x_2) e^{-i(u_1 x_1 + u_2 x_2)} dx_1 dx_2 \\ f(x_1, x_2) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(u_1, u_2) e^{i(u_1 x_1 + u_2 x_2)} du_1 du_2 \end{aligned} \tag{A.3}$$

An advantage of using Fourier transformation is its effect on convolutions. A convolution is defined as

Definition 20: Convolution

$$h(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(p) g(x - p) dp = \int_{-\infty}^{\infty} f(x - p) g(p) dp.$$

On such convolutions, one has the following relation.

Theorem 8: Convolution Theorem

Let $f(x)$ and $g(x)$ be piecewise continuous, bounded and absolutely integrable functions on the x-axis. Then

$$\mathscr{F}(f * g) = \mathscr{F}(f)\mathscr{F}(g). \tag{A.4}$$

Informally, Theorem 8 means that the Fourier transform simplifies a convolution integral in the original domain to a product in the frequency domain. This result will be central for the numerical solution algorithm presented in section 5.2.

A.1 Discrete and Fast Fourier Transform

A common scenario in applications is to have a finite number of measurements of some function of interest. This is the case with financial derivatives where the price of the underlying is measured at discrete time points. The discrete Fourier transform was developed for such scenarios.
Assume N measurements are taken of a function $f(x)$ over an interval $x \in [0, T]$ at equidistant points, $\boldsymbol{x} = [x_0, x_1, ..., x_{N-1}]$, such that

$$x_k = \frac{Tk}{N}, \qquad k = 0, 1, ..., N - 1. \tag{A.5}$$

Then, the discrete Fourier transform is defined as follows.

Definition 21: Discrete Fourier Transform

Let f_k denote $f(x_k)$ of some function $f(x)$ with x_k from (A.5). Then, the discrete Fourier transform of a signal $\boldsymbol{f} = [f_0, ..., f_{N-1}]^T$, is the signal $\hat{\boldsymbol{f}} = [\hat{f}_0, ..., \hat{f}_{N-1}]$ with components

$$\hat{f}_n = \sum_{k=0}^{N-1} f_k e^{-inx_k}, \qquad n = 0, 1, ..., N-1, \tag{A.6}$$

Note, the superscript T denotes the transpose in Definition 21. $\hat{f}_n$ is the frequency spectrum of the signal. In vector notation, Definition 21 becomes

$$\hat{\boldsymbol{f}} = \boldsymbol{F_N} \boldsymbol{f}, \tag{A.7}$$

where the $N \times N$ Fourier matrix $\boldsymbol{F_N} = [e_{nk}]$ has the entries (given from (A.6))

$$e_{nk} = e^{-inx_k} = e^{-2\pi ink/N} = w^{nk}, \qquad w = w_N = e^{-2\pi i/N},$$

where $n, k = 0, ..., N-1$.

Solving the system in (A.7) with standard solution techniques requires $\mathcal{O}(N^2)$ operations. The fast Fourier transform is an improved computational method for solving (A.7) which requires $\mathcal{O}(Nlog_2N)$ operations. There are several variations of the fast Fourier transform. I will not give a detailed description of these in this paper, but refer to [69] and [48] for a thorough review.

B Fitted Parameters

This appendix elaborates on the relations in Table 3.
I have used built-in functions from the "ghyp" package in R to perform the fitting of NIG and VG distributions to several markets (See Section 4).

The functionality in the package considers initially a generalised hyperbolic distribution of which the NIG and VG distribution are special cases. In the following, let X be a d-dimensional, generalised hyperbolic distributed random variable. Under the "standard" parameterisation used in *ghyp*, the generalised hyperbolic distribution of X is defined as the distribution with the probability density function

$$f_X(x;\lambda,\chi,\psi,\mu,\Sigma,\gamma) = \frac{(\sqrt{\psi/\chi})^\lambda(\psi+\gamma'\Sigma\gamma)^{d/2-\lambda}}{(2\pi)^{d/2}|\Sigma|^{1/2}K_\lambda(\sqrt{\chi\psi})}\times \\ \frac{K_{\lambda-d/2}(\sqrt{(\chi+Q(x))(\psi+\gamma'\Sigma^{-1}\gamma)})}{(\sqrt{(\chi+Q(x))(\psi+\gamma'\Sigma^{-1}\gamma)})^{d/2-\lambda}}\exp((x-\mu)'\Sigma^{-1}\gamma), \quad \text{(B.1)}$$

where $Q(x) = (x-\mu)'\Sigma(x-\mu)$.
For computational purposes (details in [15]), the fitting is performed under the $(\lambda,\bar{\alpha},\mu,\Sigma,\gamma)$-parameterisation. This parameterisation has a given relation to the $(\lambda,\alpha,\mu,\Delta,\delta,\beta)$-parameterisation. This parameterisation is similar to the ones used in (3.9) and (3.20).

We are interested in the univariate case and set $d = 1$. We also need the observation discussed in [15] that the Bessel function is symmetric with respect to the index λ.

The switch from $(\lambda,\bar{\alpha},\mu,\Sigma,\gamma)$ to $(\lambda,\chi,\psi,\mu,\Sigma,\gamma)$ are given by the relations (B.2), (B.3) and that Σ and γ are unchanged.

$$\psi = \bar{\alpha}\frac{K_{\lambda+1}(\bar{\alpha})}{K_\lambda(\bar{\alpha})} \quad \text{(B.2)}$$

$$\chi = \frac{\bar{\alpha}^2}{\psi} = \bar{\alpha}\frac{K_\lambda(\bar{\alpha})}{K_{\lambda+1}(\bar{\alpha})} \quad \text{(B.3)}$$

The switch from $(\lambda,\chi,\psi,\mu,\Sigma,\gamma)$ to $(\lambda,\alpha,\mu,\Delta,\delta,\beta)$ is given by:

$$\Delta = |\Sigma|^{-1}\Sigma$$
$$\beta = \Sigma^{-1}\gamma$$
$$\delta = \sqrt{\chi^{|\Sigma|}}$$
$$\alpha = \sqrt{|\Sigma|^{-1}(\psi+\gamma'\Sigma^{-1}\gamma)}$$

The opposite switch is made by

$$\Sigma = \Delta, \quad \gamma = \Delta\beta, \quad \chi = \delta^2, \quad \psi = \alpha^2 - \beta'\Delta\beta$$

Δ in the $(\lambda, \alpha, \mu, \Delta, \delta, \beta)$-parameterisation is a matrix with determinant 1. In the univariate case, this simplifies to $\Delta = 1$ [15]. Therefore, in this parameterisation, in the univariate case, some of the relations simplifies. The results are

$$\Delta = \Sigma = 1 \tag{B.4}$$
$$\beta = \gamma \tag{B.5}$$
$$\delta = \sqrt{\chi} \tag{B.6}$$
$$\psi = \alpha^2 - \beta^2 \tag{B.7}$$

NIG Distribution

The NIG distribution is the special case of the generalised hyperbolic distribution where $\lambda = -1/2$, $\chi > 0$ and $\psi > 0$. Note that this simplifies (B.2) and (B.3) to

$$\psi = \bar{\alpha} \tag{B.8}$$
$$\chi = \bar{\alpha} \tag{B.9}$$

Taking one step at a time, $d = 1$ and $\lambda = -1/2$ simplifies (B.1) to

$$f_X(x; \chi, \psi, \mu, \Sigma, \gamma) = \frac{(\psi/\chi)^{-\frac{1}{4}}(\psi + \gamma^2)}{(2\pi)^{1/2} K_{-\frac{1}{2}}(\sqrt{\chi\psi})} \times \frac{K_{-1}(\sqrt{(\chi + Q(x))(\psi + \gamma^2)})}{\sqrt{(\chi + Q(x))(\psi + \gamma^2)}} \exp((x - \mu)\gamma),$$

where $Q(x) = (x - \mu)^2$. Furthermore, the Modified Bessel function of the third order can be stated explicitly for $\lambda = -1/2$:

$$K_{-1/2}(x) \overset{\text{sym.}}{=} K_{1/2} = \sqrt{\frac{\pi}{2x}} e^x, \quad x > 0. \tag{B.10}$$

Writing $K_{-\frac{1}{2}}(\sqrt{\chi\psi})$ explicitly and rearranging yields

$$f_X(x; \chi, \psi, \mu, \Sigma, \gamma) = \frac{\psi^{1/2}(\psi + \gamma^2)}{\pi} e^{\sqrt{\chi\psi}} \frac{K_1(\sqrt{(\chi + (x - \mu)^2)(\psi + \gamma^2)})}{\sqrt{(\chi + (x - \mu)^2)(\psi + \gamma^2)}} e^{(x-\mu)\gamma}$$

Using the relations (B.4),(B.5),(B.6) and (B.7), rearranging and merging the exponential functions yields

$$f_X(x; \alpha, \mu, \Delta = 1, \delta, \beta) = \frac{\alpha(\alpha^2 - \beta^2)^{1/2}}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + (x - \mu)\beta\right) \frac{K_1\left(\alpha\sqrt{\delta^2 + (x - \mu)^2}\right)}{\sqrt{\delta^2 + (x - \mu)^2}}$$

Combining (B.6) with (B.9), we have that $\delta^2 = \chi = \bar{\alpha}$. Furthermore, from (B.8), $\bar{\alpha}$ is also equal to ψ. Finally, combined with (B.7), this yields that $(\alpha - \beta)^{1/2} =$

δ. Conclusively, when excluding the parameter Δ which was equal to 1 in the univariate case, one gets

$$f_X(x;\alpha,\mu,\delta,\beta) = \frac{\alpha\delta}{\pi}\exp\left(\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)\right)\frac{K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\sqrt{\delta^2+(x-\mu)^2}},$$

which we recognise as (3.9).

This proves that the relation between the fitted parameters and the parameters initially used by Barndorff-Nielsen (and me in this thesis), is indeed given by 3.

VG Distribution

The Variance Gamma distribution is the special case of the generalised hyperbolic distribution where $\chi \to 0$ and $\lambda > 0$.
The Variance Gamma distribution as presented in Section (3.21) is already in the standard parameterisation of the ghyp-package.

To avoid unnecessary confusion, it is however worthy of mentioning that the code returns a value for a parameter ψ which is excluded in (3.21). The parameter ψ is the same ψ as found in (B.1) and satisfies the relation in (B.2). However, for the Variance Gamma distribution, (B.2) simplifies to

$$\psi = \bar{\alpha}\frac{K_{\lambda+1}(\bar{\alpha})}{K_\lambda(\bar{\alpha})} = 2\lambda.$$

I therefore limit myself to only using the parameters $(\lambda,\mu,\sigma,\gamma)$.

C Types of Averaging

In finance there are two frequently used methods for finding the average of a set of numbers [53, p. 223-225].

C.1 Arithmetic Mean

The arithmetic mean, μ_A, of a series, S, is given by

$$\text{Discrete: } \mu_A = \frac{1}{N}\sum_{i=1}^{N} S(t_i)$$
$$\text{Continuous: } \mu_A = \frac{1}{T}\int_0^T S(\tau)d\tau$$

C.2 Geometric Mean

The geometric mean, μ_G, of a series, S is given by

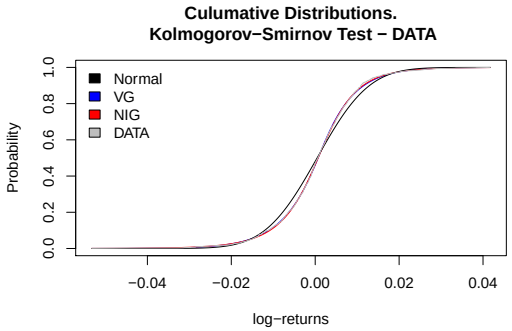
$$\text{Discrete: } \mu_G = \left(\prod_{i=1}^{N} S(t_i)\right)^{\frac{1}{N}}$$
$$\text{Continuous: } \mu_G = \exp\left(\frac{1}{T}\int_0^T \log S(\tau)d\tau\right)$$

The geometric mean is more complicated to compute than the arithmetic mean, but it has the advantage of taking compounding into account when dealing with non-independent values such as return rates.
When considering independent values, the arithmetic mean is more accurate and when considering non-independent values, the geometric mean is more accurate. Furthermore, $\mu_G \leq \mu_A$ [4], with the difference depending on the variance of the underlying in the following manner

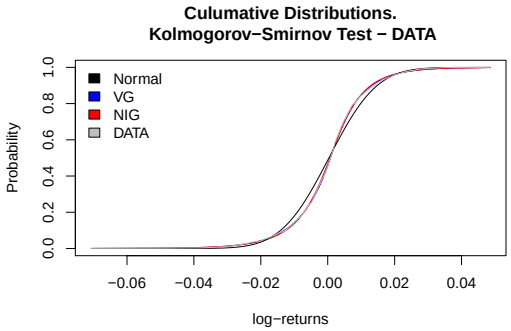
$$\mu_G \approx \mu_A - \frac{1}{2}\text{Var}[\Omega]$$

[3].
I will consider Asian options with a stock price as the underlying which does not require the same compounding as for example return rates. Thus, I will consider arithmetic Asian options which are the most common type of Asian option [26].

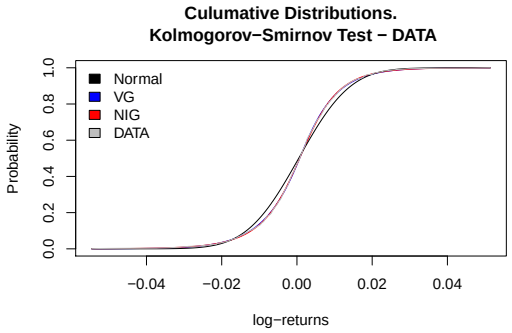
D Plots Associated with the Kolmogorov-Smirnov Test



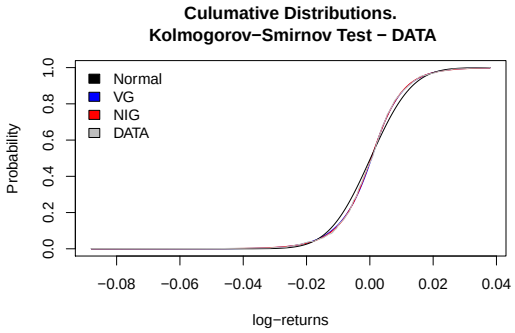
(a) OSEBX



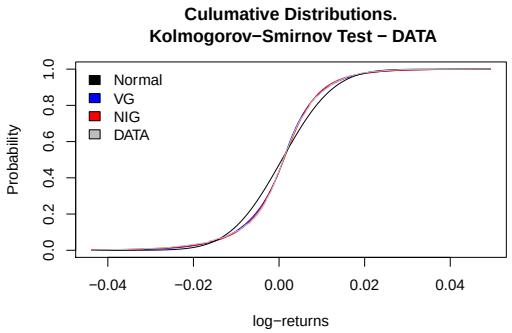
(b) DAX



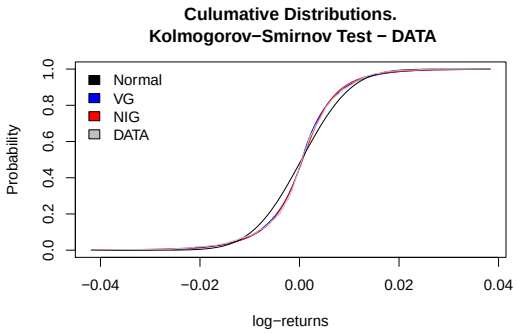
(c) OMXC20



(d) OMXS30



(e) NDX



(f) SPX

Figure 19: Cumulative distributions associated with the Kolmogorov-Smirnov Test discussed in Section 4.2.

E Code

E.1 MatLab Codes

E.1.1 Main

```
clear all;
clc;
initialiseAndGetParams;
overall_runtime = tic;

gridsizes = 2.^[8 9 10 11 12];
strikes = [0.99*strike, strike, 1.01*strike];

%% Pricing
gbmpiprices = zeros(length(strikes)+2, length(gridsizes)+1)
    ; % Gridsizes along columns, strikes along rows
[rows,cols] = size(gbmpiprices);
gbmpiprices(:,1) = [0, strikes, 0];
gbmpiprices(1,:) = [0, gridsizes];
for c = 2:(cols-1)
    pi_time = tic;
    gbmpiprices(2:(length(strikes)+1),c) = pi_fft_asian_gbm
        (s0,strikes,sd,r,q,T,dt,gridsizes(c-1)); % y-limits
        set for DAX
    gbmpiprices(length(strikes)+2,c) = toc(pi_time);
    fprintf("PI - GBM:\n Gridsize: %i, strike: %i, result:
        %d, runtime: %d \n",gridsizes(c-1),strike,
        gbmpiprices(round(rows/2),c),gbmpiprices(rows,c));
end

nigpiprices = zeros(length(strikes)+2, length(gridsizes)+1)
    ; % Gridsizes along columns, strikes along rows
[rows,cols] = size(nigpiprices);
nigpiprices(:,1) = [0, strikes, 0];
nigpiprices(1,:) = [0, gridsizes];
for c = 2:cols
    pi_time = tic;
    nigpiprices(2:(length(strikes)+1),c) = pi_fft_asian_nig
        (s0,strikes,alpha,delta,beta,munig,r,q,T,dt,
        gridsizes(c-1));
    nigpiprices(length(strikes)+2,c) = toc(pi_time);
    fprintf("PI - NIG:\n Gridsize: %i, strike: %i, result:
        %d, runtime: %d \n",gridsizes(c-1),strike,
        nigpiprices(round(rows/2),c),nigpiprices(rows,c));
end
```

```
31
32 vgpiprices = zeros(length(strikes)+2, length(gridsizes)+1);
       % Gridsizes along columns, strikes along rows
33 [rows,cols] = size(vgpiprices);
34 vgpiprices(:,1) = [0, strikes, 0];
35 vgpiprices(1,:) = [0, gridsizes];
36 for c = 2:cols
37     pi_time = tic;
38     vgpiprices(2:(length(strikes)+1),c) = pi_fft_asian_vg(
           s0,strikes,sigma,lambda,vggamma,muvg,r,q,T,dt,
           gridsizes(c-1));
39     vgpiprices(length(strikes)+2,c) = toc(pi_time);
40     fprintf("PI - VG:\n Gridsize: %i, strike: %i, result: %
           d, runtime: %d \n",gridsizes(c-1),strike,vgpiprices(
           round(rows/2),c),vgpiprices(rows,c));
41 end
42
43 %% Saving Results
44 overall_runtime = toc(overall_runtime);
45 xlswrite('gbmpiprices.xlsx',gbmpiprices);
46 xlswrite('nigpiprices.xlsx',nigpiprices);
47 xlswrite('nigpiprices_corrected.xlsx',nigpiprices_corrected
       );
48 xlswrite('vgpiprices.xlsx',vgpiprices);
49 xlswrite('vgpiprices_unsmoothed.xlsx',vgpiprices_unsmoothed
       );
```

E.1.2 PIFFT - GBM

```
function [price] = pi_fft_asian_gbm(s0,strikes,sigma,r,q,T,
    dt,N)

%% Parameters
mu = r-q; % Arbitrage-free drift of S_t
m = round(T/dt); % Number of time steps
T = m*dt; % Avoiding rounding error
% Mean and volatility of z_1 and transitions
mu_z1 = (mu-0.5*sigma^2)*dt;
sigma_z1 = sigma*sqrt(dt);

%% Initialising grid
% Defining the z range based on chebyshev's inequality
k = 10; % 1-1/k^2 of the probability included. 99 %
zmin = T*(mu-0.5*sigma^2) - k*sqrt(T*sigma^2);
zmax = T*(mu-0.5*sigma^2) + k*sqrt(T*sigma^2);
% Initialising grid
dz = (zmax-zmin)/(N-1);
z = (zmin:dz:zmax);
ymax = 1.2651e+04 + 4000;
ymin = 1.2651e+04 - 4000;
dy = (ymax-ymin)/(N-1);
y = ymin:dy:ymax;
% Defining the y,z grid
[Y,Z] = meshgrid(y,z);

%% Defining the transformation variable, z_1 = u(y_2|z_2),
    for f_{A_2|Z_2}
u = 3*Y/s0 - 1 - exp(Z); % u(y_2|z_2) = ln(tranz1)
correction = (u <= 0); % because e^z_1 > 0
u(correction) = 10^(-60);
u = log(u); % transz1 = u(y_2|z_2)
ind = ones(N);
ind(correction) = 0;
uder = 3./(3*Y-s0-s0*exp(Z));

%% Calculating the inital desity f_{A_2,Z_2}(y_2,z_2):
% f_{Z1}(z2-z1):
f_trans = (1/(sqrt(2*pi)*sigma_z1)) * exp(-0.5*(((Z-u)-
    mu_z1).^2)/...
      (sigma_z1^2)).*ind;
% f_{Z_1}((u(y_2|z_2)):
f = (1/(sqrt(2*pi)*sigma_z1)) * exp(-0.5*((u-mu_z1).^2)/(
    sigma_z1^2)).*ind;
```

```
41 % f_{A_2,Z_2}(y_2,z_2) (overwrites f_{Z_1} which is no
       longer needed)
42 f = f_trans.*f.*abs(uder);
43 neg = (f < 0);
44 f(neg) = 0;
45 % Scaling - Securing that the joint pdf integrates to 1
46 w = trapz(z,f,1);
47 w = trapz(y,w);
48 f = f/w;
49
50 %% Initialising the characteristic function of the
       transition prob.
51 freq = mod(1/2+(0:(N-1))/(N),1)-1/2;
52 v = freq * (2*pi/dz);
53 v = v';
54 phi = exp(-0.5*sigma_z1^2*v.^2).*exp(-1i*v*mu_z1);
55 phi_m = repmat(phi,1,length(phi));
56
57 %% Recursions
58 for j = 3:m
59     % Interpolating, setting f_i-1
60     YI = (j/(j-1))*Y - s0/(j-1)*exp(Z);
61     f = interp2(Y,Z,f,YI,Z,'spline');
62     corr = isnan(f);
63     f(corr) = 0;
64
65     % Calculating the convolution
66     f = fft(f);
67     f = (phi_m.*f);
68     f = (real(ifft(f)));
69     neg = (f<=0);
70     f(neg) = 0;
71
72 % Making sure the final joint density integrates to 1
73     w = trapz(y,f,1);
74     w = trapz(z,w);
75     f = f/w;
76
77     % Fixing numerical problems:
78     p1 = (abs(Z) > 0.5);
79     f(p1) = 0; % Very dangerous correction. Use with
           caution
80
81 %     if (mod(j,10) == 0)
82 %     figure;
```

```
83 %       surf(y,z,f)
84 %       end
85 end
86
87 %% Calculating the risk neutral expectation, using the
       trapesoidal rule
88 V = trapz(z,f);
89 price = zeros(1,length(strikes));
90 for i = 1:length(strikes)
91       func = max(y - strikes(i),0).*V;
92       price(1,i) = exp(-r*T)*trapz(y,func);
93 end
94
95 end
```

E.1.3 PIFFT - NIG

```
function [price] = pi_fft_asian_nig(s0,strikes,alpha,delta,
    beta,mu,r,q,T,dt,N)

%% Parameters
omega = delta*( sqrt(alpha^2-(beta+1)^2) - sqrt(alpha^2-
    beta^2) );
mu = (r - q + omega); % Mean correction
m = round(T/dt); % Number of time steps
T = m*dt; % To avoid rounding error

%% Initialising grid
% Defining the z range based on chebyshev's inequality (
    approx for NIG,
% based on GBM)
k = 10; % 1-1/k^2 of the probability included. 99 %
zmin = T*(mu + delta*beta/(sqrt(alpha^2-beta^2))) - k*...
     alpha*sqrt(delta*T*(alpha^2-beta^2)^(-3/2) );
zmax = T*(mu + delta*beta/(sqrt(alpha^2-beta^2))) + k*...
     alpha*sqrt(delta*T*(alpha^2-beta^2)^(-3/2) );
% Initialising grid
dz = (zmax-zmin)/(N-1);
z = (zmin:dz:zmax);
ymax = 16651; % Approx. mean +/- 4000 (Set from
    observations of MC sims)
ymin = 8651;
dy = (ymax-ymin)/(N-1);
y = ymin:dy:ymax;
% Defining the y,z grid
[Y,Z] = meshgrid(y,z);

%% Defining the transformation variable, z_1 = u(y_2|z_2),
    for f_{A_2|Z_2}
u = 3*Y/s0 - 1 - exp(Z); % u(y_2|z_2) = ln(tranz1)
correction = (u <= 0); % because e^z_1 > 0
u(correction) = 10^(-60);
u = log(u); % transz1 = u(y_2|z_2)
ind = ones(N);
ind(correction) = 0;
uder = 3./(3*Y-s0-s0*exp(Z));

%% Calculating the inital desity f_{A_2,Z_2}(y_2,z_2):

% f_{Z1}(z2-z1):
f_trans = alpha*dt*delta/pi * exp( dt*delta*sqrt(alpha^2-
```

```
    beta^2) +...
40      beta*(Z-u-dt*mu) ) .* besselk(1,alpha*sqrt(dt^2*delta^2
            +...
41      (Z-u-dt*mu).^2) ) ./ (sqrt(dt^2*delta^2 + (Z-u-dt*mu)
            .^2)).*ind;
42  % f_{Z_1}((u(y_2|z_2)):
43  f = alpha*dt*delta/pi * exp( dt*delta*sqrt(alpha^2-beta^2)
    +...
44      beta*(u-dt*mu) ) .* besselk(1,alpha*sqrt(dt^2*delta^2
            +...
45      (u-dt*mu).^2) ) ./ (sqrt(dt^2*delta^2 + (u-dt*mu).^2))
            .*ind;
46  % f_{A_2,Z_2}(y_2,z_2) (overwrites f_{Z_1} which is no
    longer needed)
47  f = f_trans.*f.*abs(uder);
48  neg = (f < 0); % Indecies to negative values
49  f(neg) = 0;
50  nan = isnan(f);
51  f(nan) = 0;
52  % Scaling - Securing that the joint pdf integrates to 1
53  w = trapz(z,f,1);
54  w = trapz(y,w);
55  f = f/w;
56
57  %% Initialising the characteristic function of the
    transition prob.
58  freq = mod(1/2+(0:(N-1))/(N),1) - 1/2;
59  v = freq * (2*pi/dz);
60  v = v';
61  phi = conj(exp( -delta*dt*( sqrt(alpha^2-(beta+1i*v).^2 )
    -...
62      sqrt(alpha^2-beta^2) ) + 1i*v*mu*dt ));
63  phi_m = repmat(phi,1,length(phi));
64
65  %% Recursions
66  for j = 3:m
67      % Interpolating, setting f_i-1
68      YI = (j/(j-1))*Y - s0/(j-1)*exp(Z);
69      f = interp2(Y,Z,f,YI,Z,'spline');
70      corr = isnan(f);
71      f(corr) = 0;
72
73      % Calculating the convolution
74      f = fft(f);
75      f = (phi_m.*f);
```

```
76      f = (real(ifft(f)));
77      neg = (f<=0);
78      f(neg) = 0;
79
80  % Making sure the final joint density integrates to 1
81      w = trapz(y,f,1);
82      w = trapz(z,w);
83      f = f/w;
84
85      % Fixing numerical problems:
86      p1 = (abs(Z) > 0.5);
87      f(p1) = 0; % Very dangerous correction. Use with
            caution.
88
89  %       if (mod(j,5) == 0)
90  %       figure;
91  %       surf(y,z,f)
92  %       end
93  end
94
95  %% Calculating the risk neutral expectation, using the
        trapesoidal rule
96  V = trapz(z,f);
97  price = zeros(1,length(strikes));
98  for i = 1:length(strikes)
99      func = max(y - strikes(i),0).*V;
100     price(1,i) = exp(-r*T)*trapz(y,func);
101 end
102
103 end
```

E.1.4 PIFFT - VG

```
function [price] = pi_fft_asian_vg(s0,strikes,sigma,lambda,
    vggamma,mu,r,q,T,dt,N)
% Note: Calling gamma for vggamma to avoid problems using
    the gamma
% function
%% Parameters
omega = lambda*log(1 - 1/2*sigma^2/lambda - vggamma/lambda)
    ;
mu = r - q + omega; % Mean correction.
m = round(T/dt); % Number of time steps
T = m*dt; % To avoid rounding error

%% Initialising grid
zmin = -0.8127; % DAX limits from NIG-code
zmax = 0.8077;
% Initialising grid
dz = (zmax-zmin)/(N-1);
z = (zmin:dz:zmax);
ymin = 8651;
ymax = 16651;
dy = (ymax-ymin)/(N-1);
y = ymin:dy:ymax;
% Defining the y,z grid
[Y,Z] = meshgrid(y,z);

%% Defining the transformation variable, z_1 = u(y_2|z_2),
    for f_{A_2|Z_2}
u = 3*Y/s0 - 1 - exp(Z); % u(y_2|z_2) = ln(tranz1)
correction = (u <= 0); % because e^z_1 > 0
u(correction) = 10^(-60);
u = log(u); % transz1 = u(y_2|z_2)
ind = ones(N);
ind(correction) = 0;
uder = 3./(3*Y-s0-s0*exp(Z));

%% Calculating the inital desity f_{A_2,Z_2}(y_2,z_2):
% f_{Z1}(z2-z1):
f_trans = 2*lambda^(lambda*dt)*exp((Z-u-mu*dt)*vggamma/
    sigma^2)/...
      (sqrt(2*pi)*sigma*gamma(lambda*dt))...
      .* besselk( lambda*dt-1/2 , 1/sigma^2*sqrt((Z-u-dt*mu)
          .^2*...
      (2*lambda*sigma^2+vggamma^2)) ) ./ ( (Z-u-mu*dt).^2/...
      (2*lambda*sigma^2+vggamma^2)).^(1/4-1/2*lambda*dt).*ind
```

```
        ;
% f_{Z_1}((u(y_2|z_2)):
f = 2*lambda^(lambda*dt)*exp((u-mu*dt)*vggamma/sigma^2)/(sqrt(2*pi)*...
    sigma*gamma(lambda*dt)).*...
    besselk( lambda*dt-1/2 , 1/sigma^2*sqrt((u-dt*mu).^2*...
    (2*lambda*sigma^2+vggamma^2)) ) ./...
    ( (u-mu*dt).^2/(2*lambda*sigma^2+vggamma^2)).^(1/4-1/2*lambda*dt).*ind;
% f_{A_2,Z_2}(y_2,z_2) (overwrites f_{Z_1} which is no longer needed)
f = f_trans.*f.*abs(uder);
neg = (f < 0); % Indecies to negative values
f(neg) = 0;
nan = isnan(f);
f(nan) = 1e-60;
% Scaling - Securing that the joint pdf integrates to 1
w = trapz(z,f,1);
w = trapz(y,w);
f = f/w;

%% Initialising the characteristic function of the transition prob.
freq = mod(1/2+(0:(N-1))/(N),1) - 1/2;
v = freq * (2*pi/dz);
v = v';
phi = conj( (1 - 1i*vggamma/lambda*v + sigma^2/...
    (2*lambda)*v.^2).^(-dt*lambda) .* exp(1i*v*mu*dt ));

%% Smoothing and Fixing numerical issues:
% realpart = real(phi);
% imagpart = imag(phi);
% px = (round(N*0.5*0.7)):(round(N*0.5*1.3)); % Problem area indeces
% py = imagpart(px);
% py = smooth(py,length(px),'rloess');
% imagpart(px) = py;
% neg = (real(realpart)<=0);
% realpart(neg) = 0;
% phi = realpart + 1i*imagpart;

phi_m = repmat(phi,1,length(phi));

%% Recursions
```

```
77  for j = 3:m
78      % Interpolating, setting f_i-1
79      YI = (j/(j-1))*Y - s0/(j-1)*exp(Z);
80      f = interp2(Y,Z,f,YI,Z,'spline');
81      corr = isnan(f);
82      f(corr) = 0;
83
84      % Calculating the convolution
85      f = fft(f);
86      f = (phi_m.*f);
87      f = (real(ifft(f)));
88      neg = (f<=0);
89      f(neg) = 0;
90
91  % Making sure the final joint density integrates to 1
92      q = trapz(y,f,1);
93      w = trapz(z,q);
94      f = f/w;
95
96      % Fixing numerical problems:
97      p1 = (abs(Z) > 0.5);
98      f(p1) = 0; % Very dangerous correction. Use with
           caution
99
100 %       if (mod(j,25) == 0)
101 %       figure;
102 %       surf(y,z,f)
103 %       end
104 end
105 q = trapz(y,f,1);
106 w = trapz(z,q);
107 f = f/w;
108
109
110 %% Calculating the risk neutral expectation, using the
       trapesoidal rule
111 V = trapz(z,f);
112 price = zeros(1,length(strikes));
113 for i = 1:length(strikes)
114     func = max(y - strikes(i),0).*V;
115     price(1,i) = exp(-r*T)*trapz(y,func);
116 end
117
118 end
```

E.2 R Codes

E.2.1 Monte Carlo Simulation

```
asianMC_strikes = function(params,strikes,q=0,T=1,dt=250,no
    =1000000,ifPlot=FALSE,n_plotpaths=10){
  # See "main" for descirptions of variables
  tictoc::tic(0)

  r = params$GLOBAL$riskfreeRate
  s0 = params$GLOBAL$s0
  days = round(T/dt)
  T = days*dt

  simprice = data.frame(matrix(0, ncol = 3, nrow = length(
      strikes)))
  colnames(simprice) = c("NIG_price","VG_price","GBM_price
      ")
  rownames(simprice) = strikes

  for(i in 1:3){ # 1 = NIG, 2 = VG, 3 = GBM
    tictoc::tic(i)

    if(i == 1){ # NIG
      alpha = params$NIG$alpha
      beta = params$NIG$beta
      delta = params$NIG$delta
      omega = delta*( sqrt(alpha^2-(beta+1)^2) - sqrt(alpha
          ^2-beta^2) )
      mu = r - q + omega # Mean corrected mu

      dx = matrix( rghyp(n=days*no,object=NIG.ad(alpha =
          alpha, delta = (delta)*(dt), beta = beta, mu = mu*
          dt)), nrow = no,ncol = days )
      s = matrix(s0, nrow = no, ncol = days)

      price = matrix(NA, nrow = length(strikes), ncol = no)

      for (i in 1:no) {
        for (t in 2:days) {
          s[i, t] = s[i, t - 1] * exp(dx[i,t])
          if(t==days & i%%10000 == 0){print(i)}
        }
        for (i_strike in 1:length(strikes)){
          price[i_strike,i] = exp(-r * T) * max(c((mean(s[i
              , ]) - strikes[i_strike]), 0))
```

```
36          }
37        }
38
39        for (i_strike in 1:length(strikes)){
40        simprice$NIG_price[i_strike] = mean(price[i_strike,])
41        }
42        nigtime = tictoc::toc(i, quiet = TRUE)
43        nigtime = as.numeric(nigtime$toc - nigtime$tic)
44
45        if (ifPlot) {
46          title = "NIG"
47          MC_plot(s,no,days,n_plotpaths,title)
48        }
49
50      }else if(i == 2){ # VG
51        lambda = params$VG$lambda
52        mu = params$VG$mu
53        sigma = params$VG$sigma
54        gamma = params$VG$gamma
55        omega = lambda*log( 1 - sigma^2/(2*lambda) - gamma/
              lambda )
56        mu = r - q + omega # Mean corrected mu
57
58        VGobject_t = VG(lambda = lambda*dt, mu = dt*mu, sigma
              = sigma*sqrt(dt), gamma = dt*gamma)
59        dx = matrix( rghyp(n=days*no,object=VGobject_t), nrow
              = no,ncol = days )
60
61        s = matrix(s0, nrow = no, ncol = days)
62
63        price = matrix(NA, nrow = length(strikes), ncol = no)
64
65        for (i in 1:no) {
66          for (t in 2:days) {
67            s[i, t] = s[i, t - 1] * exp(dx[i,t])
68            if(t==days & i%%10000 == 0){print(i)}
69          }
70          for (i_strike in 1:length(strikes)){
71            price[i_strike,i] = exp(-r * T) * max(c((mean(s[i
                  , ]) - strikes[i_strike]), 0))
72          }
73        }
74
75        for (i_strike in 1:length(strikes)){
76          simprice$VG_price[i_strike] = mean(price[i_strike
```

```
            ,])
77        }
78        vgtime = tictoc::toc(i, quiet = TRUE)
79        vgtime = as.numeric(vgtime$toc - vgtime$tic)
80
81        if (ifPlot) {
82          title = "VG"
83          MC_plot(s,no,days,n_plotpaths,title)
84        }
85
86      }else{#i = 3 # GBM
87        mu = params$GBM$mean
88        sd = params$GBM$sd
89        mu = r - q - 1/2*sigma^2 # Mean corrected mu
90
91        dx = matrix( rnorm(n=days*no,mean=mu*dt,sd=sqrt(dt)*
              sd), nrow = no,ncol = days )
92        s = matrix(s0, nrow = no, ncol = days)
93
94        price = matrix(NA, nrow = length(strikes), ncol = no)
95
96        for (i in 1:no) {
97          for (t in 2:days) {
98            s[i, t] = s[i, t - 1] * exp(dx[i,t])
99            if(t==days & i%%10000 == 0){print(i)}
100         }
101         for (i_strike in 1:length(strikes)){
102           price[i_strike,i] = exp(-r * T) * max(c((mean(s[i
                , ]) - strikes[i_strike]), 0))
103         }
104       }
105
106       for (i_strike in 1:length(strikes)){
107         simprice$GBM_price[i_strike] = mean(price[i_strike
              ,])
108       }
109       gbmtime = tictoc::toc(i, quiet = TRUE)
110       gbmtime = as.numeric(gbmtime$toc - gbmtime$tic)
111
112       if (ifPlot) {
113         title = "GBM"
114         MC_plot(s,no,days,n_plotpaths,title)
115       }
116
117     } #else
```

```
118
119     } # for(i in 1:3)
120
121     totaltime = tictoc::toc(0, quiet = TRUE)
122     times = data.frame(t(c(nigtime=nigtime,vgtime=vgtime,
            gbmtime=gbmtime,totaltime=totaltime)))
123
124     return(list(prices=simprice, times=times))
125 }
```

E.2.2 Parameter Fitting

```
fitParams = function(data, ifPlot=FALSE, ifTest=FALSE,
    ifPlotKS=FALSE){
# data: Data Frame containing: Time | logreturns
# ifPlot: Logical. True => Plot fitted models and display
    evaluating statistics
# ifTest: Logical. True => Perform a Goodness-of-fit test (
    Kologorov Smirnov)


  # Built in functionality to ID best fitted model based on
      AIC alone
  bestmodel = stepAIC.ghyp(data[,2])$best.model
  bestmodel_name = getElement(bestmodel, "model")[3]


  # Fitting ####
  fittedVG = ghyp::fit.VGuv(data[,2])
  fittedNIG = ghyp::fit.NIGuv(data[,2])
  fittednormal = MASS::fitdistr(data[,2], "normal")


  # Goodness-of-fit test (Kolmogorov-Smirnov test) ####
  ks = NULL # avoids error when ifTest = FALSE
  if(ifTest){

    filename = file.path("C:","Users/marku/Dropbox/Master/
        Code/fig/fit",paste(deparse(substitute(data)), "_ks.
        pdf", sep = ""))
    ks = kolmogorovSmirnovTest_v2(data,fittedNIG,fittedVG,
        fittednormal,ifPlotKS,filename)
    dev.set(2)

  }

  # Extract Parameters ####
  nigparams = data.frame(cbind(as.data.frame(coef(fittedNIG
      , type = "alpha.delta")),data.frame(AIC(fittedNIG),
      logLik(fittedNIG))))
  vgparams = data.frame(cbind(as.data.frame(coef(fittedVG,
      type = "chi.psi")),data.frame(AIC(fittedVG),logLik(
      fittedVG))))
  normalparams = data.frame(cbind(as.data.frame(t(
      fittednormal$estimate)),data.frame(AIC(fittednormal),
      fittednormal$loglik)))
```

```
32
33  # Store results
34  parameters = list(NIG=nigparams,VG=vgparams,GBM=
        normalparams,KS=ks,BESTMODEL=bestmodel_name,GLOBAL=
        NULL)
35
36  # Plotting ####
37  if(ifPlot){
38
39    print(noquote(paste(c('Best model (based on AIC alone)
          for',paste(deparse(substitute(data)),'is',
          bestmodel_name)))))
40
41    # Densities:
42    # filename = file.path("C:","Users/marku/Dropbox/Master
          /Code/fig/fit",paste(deparse(substitute(data)), "
          _density.pdf", sep = ""))
43    # pdf(file = filename, width=6, height=4)
44
45    main = paste("Fitted densities for ",toupper(deparse(
          substitute(data))), sep = "")
46    xlim = c(min(data[,2]),max(data[,2]))
47    x = seq(xlim[1],xlim[2],(diff(xlim))/99)
48    hist(data[,2], breaks = x, freq = FALSE, border = "gray
          ", main = main, xlab = "Log returns", ylab = "
          Density")
49    lines(fittedVG, col = "blue")
50    lines(fittedNIG, col = "red")
51    lines(x,dnorm(x, mean = normalparams$mean, sd =
          normalparams$sd), type = 'l', col = "black")
52    legend(x = "topleft", c("Normal","VG","NIG",toupper(
          deparse(substitute(data)))), fill = c("black","blue
          ","red","gray"),bty = "n")
53    # dev.off()
54
55    # QQ-Plot
56    # filename = file.path("C:","Users/marku/Dropbox/Master
          /Code/fig/fit",paste(deparse(substitute(data)), "_qq
          .pdf", sep = ""))
57    # pdf(file = filename, width=6, height=4)
58
59    # main = paste("QQ-Plot for ",toupper(deparse(
          substitute(data))))
60    # ghyp::qqghyp(fittedNIG, data = data[,2], ghyp.pch =
          1, gauss.pch = 1, ghyp.col = "red", gauss.lty = 1,
```

```
        ghyp.lty = 0, gaussian = TRUE, main = main, gauss.
        col = "black", plot.legend = FALSE)
61    # ghyp::qqghyp(fittedVG, data = data[,2], ghyp.col = "
        blue", gauss.pch = 1, gauss.lty = 0, main = main,
        add = TRUE, plot.legend = FALSE)
62    # legend(x = "topleft", c("Normal","VG","NIG"), fill =
        c("black","blue","red"),bty = "n")
63    # dev.off()
64    # dev.set(2)
65  }
66
67  return(parameters)
68
69 }
```

E.2.3 Kolmogorov's Goodness-of-fit-test

```
kolmogorovSmirnovTest_v2 = function(data, fittedNIG, fittedVG
    , fittednormal, ifPlotKS=FALSE, filename){
  # data: Data Frame containing: Time | logreturns
  # fitted...: Distributions to test the fit of (Code
     adapted to the specific object types of relevance)
  # ifPlotKS: if TRUE, plot cumulative distributions
  # filename: Filename of ks-plot.

  x = sort(data[,2])
  n = length(x)

  # Empirical Cumulative Distribution Function Fn(x): ####
  Fn = numeric(n)
  for(i in 1:n){
    Fn[i] = i*(1/n)
  }

  # Reference Cumulative Distribution F0(x): ####
  # NIG:
  Fnig = pghyp(q = x, object = fittedNIG)
  # VG:
  Fvg = pghyp(q = x, object = fittedVG)
  # Normal:
  params = coef(fittednormal)
  Fnormal = pnorm(q = x, mean = params["mean"], sd = params
     ["sd"])

  Dnnig = max(abs(Fn-Fnig))
  Dnvg = max(abs(Fn-Fvg))
  Dnnormal = max(abs(Fn-Fnormal))

  ks = round(c(NIG = sqrt(n)*Dnnig, VG = sqrt(n)*Dnvg,
     Normal = sqrt(n)*Dnnormal), digits = 4)

  if(ifPlotKS){

    # pdf(file = filename, width=6, height=4)

    main = paste("Culumative Distributions.\n Kolmogorov-
       Smirnov Test - ",toupper(deparse(substitute(data))),
        sep = "")
    plot(x,Fvg,type='l',col="blue",main=main,xlab="log-
       returns",ylab="Probability")
    lines(x,Fnig, col = "red")
```

```
38      lines(x,Fnormal, col = "black")
39      lines(x,Fn, col = "gray")
40      legend("topleft", c("Normal","VG","NIG",toupper(deparse
            (substitute(data)))), fill = c("black","blue","red
            ","gray"),bty = "n")
41      # dev.off()
42    }
43
44    return(ks)
45  }
```