

SUPERPOSITIONS OF ORNSTEIN-UHLENBECK TYPE PROCESSES:
INTERMITTENCY AND APPLICATIONS TO FINANCE

By

Irena Tesnjak

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Statistics - Doctor of Philosophy

2017

ABSTRACT

SUPERPOSITIONS OF ORNSTEIN-UHLENBECK TYPE PROCESSES: INTERMITTENCY AND APPLICATIONS TO FINANCE

By

Irena Tesnjak

Ornstein-Uhlenbeck (OU) type processes driven by Levy noise are useful in modeling the activity time in the fractal activity time geometric Brownian motion (FATGBM) model for a risky asset. Discrete superpositions of these processes can be constructed to incorporate non-Gaussian marginal distributions and long or short range dependence. While the partial sums of finite superpositions of OU type processes obey the central limit theorem, we show that the partial sums of a large class of infinite long range dependent superpositions are intermittent. We discuss the property of intermittency and behavior of the cumulants for the long-range dependent superpositions of OU type processes. In addition we show an application of finite superpositions in modeling financial time series and superiority of the model at hand compared to the Black-Scholes model when modeling log-returns.

TABLE OF CONTENTS

| | |
|--|-----------|
| LIST OF TABLES | v |
| LIST OF FIGURES | vi |
| Introduction | 1 |
| Chapter 1 Overview of the Ornstein-Uhlenbeck (OU) type processes and their superpositions | 4 |
| 1.1 Preliminaries | 4 |
| 1.2 Levy processes | 9 |
| 1.3 Infinitely divisible distributions | 11 |
| 1.3.1 Infinite divisibility and the Levy-Khintchine representation | 11 |
| 1.3.2 The correspondence between infinitely divisible distributions and Levy processes in law | 16 |
| 1.3.3 Self-decomposable distributions | 22 |
| 1.4 Ornstein-Uhlenbeck type processes | 26 |
| 1.4.1 Stochastic integral with respect to the Levy process | 28 |
| 1.4.2 Existence and uniqueness of the OU type process | 34 |
| 1.4.3 Properties of the the OU type processes | 37 |
| 1.4.3.1 The stationary OU type process | 37 |
| 1.4.3.2 The stationary OU type processes and self-decomposable distributions | 41 |
| 1.4.3.3 Covariance of the stationary OU type process | 44 |
| 1.5 Superpositions of the OU type processes | 45 |
| 1.5.1 Existence of supOU processes | 45 |
| 1.5.2 The covariance structure of supOU processes | 47 |
| 1.5.2.1 Finite supOU processes | 47 |
| 1.5.2.2 Infinite supOU processes | 47 |
| 1.5.3 Examples of OU type superpositions | 50 |
| 1.5.3.1 Gamma supOU process | 50 |
| 1.5.3.2 Inverse Gaussian supOU process | 51 |
| 1.5.3.3 Variance gamma supOU process | 52 |
| 1.5.3.4 Normal inverse Gaussian supOU process | 54 |
| 1.5.3.5 Tempered stable supOU process | 55 |
| Chapter 2 Intermittency of supOU processes | 57 |
| 2.1 Intermittency | 59 |
| 2.2 Limiting properties of partial sums of finite supOU processes | 62 |
| 2.3 Limiting properties of partial sums of infinite supOU processes | 64 |

| | | |
|---------------------|---|------------|
| Chapter 3 | Applications in finance | 79 |
| 3.1 | The Black-Scholes Model: | |
| | Geometric Brownian Motion (GBM) | 81 |
| 3.2 | Empirical evidence against the Black-Scholes model | 83 |
| 3.2.1 | The non-Gaussian character of the returns | 84 |
| | 3.2.1.1 Skewness (asymmetry) and excess kurtosis (fat tails) | 86 |
| 3.2.2 | The dependence structure of the log-returns: | |
| | regular, absolute and squared returns | 89 |
| 3.2.3 | Stochastic Volatility and clustering | 94 |
| 3.3 | Fractal Activity time geometric Brownian motion model with supOU as activity time | 98 |
| 3.3.1 | Fractal Activity time geometric Brownian motion process | 99 |
| 3.3.2 | Properties of log-returns in the FATGBM model | 101 |
| | 3.3.2.1 Moments, skewness and kurtosis | 102 |
| | 3.3.2.2 The dependence structure in the FATGBM model | 103 |
| | 3.3.2.3 Stochastic volatility - conditional heteroscedasticity | 105 |
| 3.3.3 | Distribution theory | 106 |
| | 3.3.3.1 Gamma distribution of the unit increments of the activity time process | 107 |
| 3.4 | Data calibration | 110 |
| | 3.4.1 VG parameter estimation using method of moments | 112 |
| | 3.4.2 VG parameter fitting - results | 115 |
| | 3.4.3 Conclusion | 123 |
| APPENDICES | | 124 |
| | Appendix A Mixing Coefficients and some important results | 125 |
| | Appendix B Financial prerequisites | 126 |
| BIBLIOGRAPHY | | 133 |

LIST OF TABLES

| | | |
|------------|---|----|
| Table 3.1: | Values for sample skewness, kurtosis and the resulting decisions of the normality test. | 85 |
| Table 3.2: | Values for sample skewness, kurtosis mean and standard deviation . | 89 |

LIST OF FIGURES

| | | |
|--------------|--|-----|
| Figure 3.1: | Histogram of the S&P500 returns against the Normal curve fit. . . . | 84 |
| Figure 3.2: | Empirical autocorrelation function of regular, squared and absolute log-returns for the S&P500 index. | 91 |
| Figure 3.3: | Autocorrelation function of regular, squared and absolute log-returns for the simulated Geometric Brownian motion, calibrated on the S&P500 index. | 93 |
| Figure 3.4: | S & P 500 real returns against the simulated GBM returns. | 95 |
| Figure 3.5: | The historical volatility for the S&P 500 index returns (daily readings). | 97 |
| Figure 3.6: | Fitting of VG distribution parameter to the S&P 500 Index returns. The plot shows the histogram of the empirical returns along with three differently calibrated VG densities. The methods presented are VG MOM method (example 2), VG MLE method (example3) and the Normal curve. | 120 |
| Figure 3.7: | Fitting VG, NIG and Normal distribution parameters to the S&P 500 Index returns. The plot shows the histogram of the empirical returns along with three different density. | 120 |
| Figure 3.8: | Tail probabilities of the Fitted VG (MOM, MLE), NIG and Normal calibrated on the S&P 500 returns. The plot shows how three distributions fit the empirical distribution. | 121 |
| Figure 3.9: | Fitting of VG distribution parameters to the S&P 500 Index returns. The plot shows three different examples explaining different methods of estimation. | 121 |
| Figure 3.10: | Fitting the VG distribution to the S&P 500 Index returns. The plot shows outputs of the MLE approach. | 122 |

Introduction

Fisher Black, Myron Scholes and Robert C. Merton have developed the Geometric Brownian motion (GBM) to model risky asset evolution and by provided the finance community with an explicit method for pricing a European option, see [12, 56].

Although, the classical Black-Scholes GBM model is undoubtedly, the most renowned and widely accepted model within both academic and practitioners community worldwide, the fit of the model has been questioned continuously over the last decade. Recently, a large body of literature has shown the departure of the model from the real financial data, see for instance Heyde and Liu (2001) [39] and Heyde (2009) [38], also Cont (2001) and Granger (2005) [16, 35] who thoroughly explored the characteristics of the empirical log-returns for the financial time-series. The 'stylized features' of the log-returns included heavier-tailed and higher peaked distributions for the log-returns, uncorrelated and stationary returns and the dependence of squared and absolute returns (both short and long range dependence), which contradicted the GBM model.

A number of models that incorporate non-Gaussian distributions and/or dependence in log returns have been proposed (see, [52], [51], [37], [36], [10], and many others) including the class of stochastic volatility models of Barndorff-Nielsen and Shephard, see [8]. The focus is motivated by fractal activity time GBM models (FATGBM) that, along with other

stochastic volatility models, address the stylized features. FATGBMs were first introduced in Heyde (1999) [37] and further developed for the Variance-Gamma, Normal-Inverse Gaussian and student distributions in [24–26, 36, 48, 49]. Heyde’s FATGBM model belongs to the class of subordinator type models, where a random time change (*activity time process*) with preferable characteristics, is used instead of the regular deterministic time to account for the stylized features of the log-returns.

Several versions of Heyde’s model have the capacity to capture the stylized features of the financial data if a suitable activity time is being chosen. As suggested by Heyde himself [37], if the activity time process T_t is chosen such that its increments exhibit dependence and has heavier tails than Gaussian, then these features are going to be inherited by the resulting model. Furthermore, there is a strong empirical evidence (Heyde and Liu (2001) [39]) that the activity time T_t should be asymptotically self-similar.

In this thesis we focus on FATGBM models with a specific time change, namely, the models that use superpositions of Ornstein-Uhlenbeck type processes (supOU) to construct the increments of the activity time, elaborated in [48, 49]. The finite or infinite supOU processes were first introduced by Bardnorff-Nielsen and other collaborators, see [3, 4, 8]. The motivation behind this approach lies in the fact that both finite and infinite supOU exhibit the properties that an ‘ideal’ activity time should have, as mentioned in Heyde (1999) [37] and established further in the literature. In addition to allowing for non-Gaussian marginals, the supOU process have very flexible correlation structure (summation of exponential terms) that allows for easy modeling of short and long range dependence. In addition, these models are analytically very tractable (their marginal distributions are known), which is a very useful property because implementation of these models is straightforward.

It can be shown that the partial sums of finite supOU processes have Gaussian limiting

distribution, hence they are asymptotically self-similar. In this thesis, the limiting properties of the infinite supOU process are investigated. We found that the partial sums of the infinite superpositions are intermittent, which in turn implies unusual cumulant behavior which may preclude central limit theorem (CLT) type of results. Therefore, any use of the infinite superposition processes would be very difficult from the application standpoint, since testing efficiency of estimators, constructing confidence intervals or performing hypothesis tests would be challenging. Another aspect from the finance standpoint is that, as a result of our findings, use of the infinite LRD supOU process as an activity time would not address the challenge of building a FATGBM model that incorporates long-range dependence of the log returns and the asymptotic self-similarity of the activity time.

The thesis is organized as follows. The first chapter includes the preliminary material introducing the superpositions of Ornstein-Uhlenbeck type processes. The second chapter presents new theoretical results. The third chapter focuses on applications of finite superpositions of the OU type processes for building tractable FATGBMs.

Chapter 1

Overview of the Ornstein-Uhlenbeck (OU) type processes and their superpositions

1.1 Preliminaries

This section contains standard terminology, definitions, and theorems that will be used throughout. Let (Ω, \mathcal{F}, P) be a probability space and let $X : \Omega \rightarrow \mathbb{R}$ be a $\mathcal{F}/\mathcal{B}(\mathbb{R})$ -measurable real-valued mapping or simply a random variable. Here $\mathcal{B}(\mathbb{R})$ is the sigma-algebra of Borel sets on \mathbb{R} . A probability measure induced by X , denoted as $F_X(B) = P(X \in B) = P(\{\omega : X(\omega) \in B\})$, for all $B \in \mathcal{B}(\mathbb{R})$, will be referred to as distribution of X , or sometimes a law of X . In addition, whenever two random variables, say X, Y on \mathbb{R} , not necessarily defined on the same probability space, have identical distributions, i.e. $F_X = F_Y$, we will use the following notation $X \stackrel{\mathcal{D}}{=} Y$.

The convolution μ of two distributions μ_1 and μ_2 on \mathbb{R} , denoted by $\mu = \mu_1 * \mu_2$ is a distribution defined by $\mu(B) = \int \int_{\mathbb{R} \times \mathbb{R}} \mathbf{1}_B(x+y) \mu_1(dx) \mu_2(dy)$, $B \in \mathcal{B}(\mathbb{R})$. Note, in case $X_1 \sim \mu_1$ and $X_2 \sim \mu_2$ are two independent random variables on \mathbb{R} , then their sum $X_1 + X_2$ has distribution $\mu_1 * \mu_2$.

A characteristic function of a random variable X with a distribution μ , usually denoted by $\varphi_X(z)$, or sometimes $\hat{\mu}(z)$, is given by the following expression:

$$\varphi_X(z) = \hat{\mu}(z) = \mathbb{E}e^{izX}, \quad z \in \mathbb{R}.$$

A cumulant (generating) function of a random variable X will be denoted by:

$$\kappa_X(z) = \log \mathbb{E}e^{izX}, \quad z \in \mathbb{R}.$$

Assuming it exists, one can define the m -th cumulant of X , denoted as $\kappa_X^{(m)}$, for $m \in \mathbb{N}$, using the m -th derivative of the cumulant function κ_X , evaluated at zero, written correctly below,

$$\kappa_X^{(m)} = (-i)^m \left. \frac{d^m}{dz^m} \kappa_X(z) \right|_{z=0}$$

In this context, we say that a characteristic function $\kappa_X(\cdot)$ is analytic around the origin, if and only if, at any point in the neighborhood of the origin, its Taylor series converges to the value of the function, i.e. if for any z around the origin, the following is true,

$$\kappa_X(z) = \sum_{m=1}^{\infty} \frac{(iz)^m}{m!} \kappa_X^{(m)}.$$

In this paragraph, we will briefly explain regularly and slowly varying functions, which is an essential tool when dealing with heavy tails, long range dependence, and domains of attraction. Intuitively, regularly varying functions, in the long run, behave like power functions, which in case of regularly varying distributions means that they have tails that are approximately power-law, like Pareto distribution for example.

Definition 1.1.1. *It is said that a measurable function $U : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is regularly varying at ∞ with an index $\rho \in \mathbb{R}$, if for any scale $x > 0$ the following holds,*

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\rho,$$

where the index ρ is called the exponent of variation.

Note that regularly varying functions are asymptotically scale invariant functions.

Definition 1.1.2. *A measurable function $L : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is slowly varying at ∞ if for any scale $x > 0$ the following holds,*

$$\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1,$$

i.e. it is regularly varying function with $\rho = 0$.

Notice that if $U \in RV_\rho$ then $U(x)/x^\rho \in RV_0$, which in turn means that any ρ -regularly varying function can be represented as $U(x) = x^\rho L(x)$. Simple examples of slowly varying function is $\log x, \log(\log x)$ and for the regularly varying are power functions. On the other hand, an exponential function e^x is an example of not regularly varying function.

In the next paragraph we are going to define notions related to stochastic processes and families of random variables. From here throughout, a probability space (Ω, \mathcal{F}, P) is fixed and random variables are defined on it. We will start with defining the family of finite-dimensional distributions and continue with stochastic processes, see below.

Definition 1.1.3. *For any fixed integer n , and any choice of $0 \leq t_1 < t_2 < \dots < t_n$, the joint distribution $P[X(t_1) \in B_1, X(t_2) \in B_2, \dots, X(t_n) \in B_n]$ determines a probability measure on $\mathcal{B}((\mathbb{R})^n)$. The family of probability measures over all possible choices of n and $\{t_i\}_{i=1}^n$'s, is called the system of finite-dimensional distributions of a process $\{X_t\}$.*

Definition 1.1.4. Let X_j be a random variable with values in \mathbb{R}^{d_j} , where $j = 1, \dots, n$. One says that the finite family of random variables $\{X_{t_1}, \dots, X_{t_n}\}$ is independent if for every $B_j \in \mathcal{B}(\mathbb{R}^{d_j})$, $j = 1, \dots, n$, the joint probability can be decomposed as following:

$$P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^n P(X_i \in B_i).$$

Note that an infinite family of random variables is independent, if every finite subfamily of it is independent.

Definition 1.1.5 (Modification). A stochastic process $\{Y_t\}$ is said to be a modification of a stochastic process $\{X_t\}$ if they are almost surely equal at each time point t , i.e. if for all $t \in [0, \infty)$, $P(X_t = Y_t) = 1$.

Definition 1.1.6. Two stochastic process $\{X_t\}$ and $\{Y_t\}$, not necessarily defined on the same probability space, are said to be equal in law (in distribution), denoted as $\{X_t\} \stackrel{\mathcal{D}}{=} \{Y_t\}$, if the systems of their finite-dimensional distributions are equal. Note that, from here throughout whenever we talk about equality in distribution of two stochastic processes we mean equality of their finite dimensional distributions.

Definition 1.1.7 (Stationarity).

- (i) The stochastic process $\{X_t\}$ is said to be stationary in a narrow sense if the system of its finite-dimensional distributions is shift invariant, i.e. if $\forall h \geq 0$, $\{X_t\}_{t \geq 0} \stackrel{\mathcal{D}}{=} \{X_{t+h}\}_{t \geq 0}$.
- (ii) The stochastic process $\{X_t\}$ is said to be stationary in a wide sense (weakly), also called second order stationary, if, one- and two-dimensional distributions of the process $\{X_t\}$ are shift invariant.

An important concept involving stochastic processes, either continuous or discrete, is

memory of the process, as $t \rightarrow \infty$. Intuitively, depending on the behavior of the process, its short or long term memory property will allow one to extrapolate across time scales and deduce long time behavior from short time behavior.

The most obvious way to measure the length of memory in a stochastic process that has second-order moments is by looking at the rate at which its correlations decay with time-lag.

Hence, consider a weakly or second-order stationary time series $\{X_n\}_{n \in \mathbb{Z}}$, assuming in addition, the existence of the second moment. Let's first define measures of dependence in time series, that is let $R_X(k)$ and $\rho_X(k)$ be autocovariance function (ACVF) and autocorrelation function (ACF) for the process $\{X_n\}$. For any time points $t, k \in \mathbb{Z}$, we define the following,

$$R_X(k) = Cov(X_0, X_k) = Cov(X_t, X_{t+k}), \quad (1.1)$$

and also,

$$\rho_X(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{Cov(X_0, X_k)}{Var(X_0)}. \quad (1.2)$$

Now, we can give definition of the notions of short and long range dependence as follows,

Definition 1.1.8. *A second-order stationary time series $\{X_n\}_{n \in \mathbb{Z}}$ with finite second moment is called short range dependent (SRD, in short) if,*

$$\sum_{k=-\infty}^{+\infty} |R_X(k)| < \infty. \quad (1.3)$$

and long range dependent (LRD in short) if,

$$\sum_{k=-\infty}^{+\infty} |R_X(k)| = \infty. \quad (1.4)$$

Note, that short range dependence is also referred to as short memory or weakly dependent and similarly, long range dependence is often referred to as long memory or strongly dependent series. Also, in case of the continuous time stochastic process the above series will be replaced by integrals over the time axis, $t \in \mathbb{R}$.

1.2 Levy processes

The Levy processes are named after the famous French mathematician Paul Lévy, whose contribution, although not alone in it, played an important role in understanding and characterizing the processes with stationary and independent increments.

Definition 1.2.1 (Lévy process). *Let $\{Z_t\}_{t \geq 0}$ be a stochastic process defined on a probability space (Ω, \mathcal{F}, P) . We say that a real-valued stochastic process Z_t is a Lévy process in law if the following conditions are satisfied:*

- (i) *It starts at zero almost surely, i.e. $Z_0 = 0$ a.s.*
- (ii) *Z_t has independent increments, i.e. for any $n \geq 1$ and any choice of $0 \leq t_0 < \dots < t_n, \leq$ the family of random variables $\{Z(t_k) - Z(t_{k-1})\}_{k=1}^n$ is independent.*
- (iii) *Z_t has stationary increments (temporal homogeneity), i.e. for any $s \in [0, \infty)$, the two processes are equal in law, $\{Z_t - Z_s\}_{t \geq 0} \stackrel{D}{=} \{Z_{t-s}\}_{t \geq 0}$. In other words, the distribution of $Z_{t+h} - Z_t$ doesn't depend on specific time point t , only on the time increment h .*
- (iv) *Z_t is stochastically continuous, or in other words if the following is true, for any $t \in [0, \infty)$ and $\epsilon > 0$, $\lim_{s \rightarrow t} P(|Z_s - Z_t| > \epsilon) = 0$.*

The above statements from (i)-(iv) define the Levy process in law (in distribution), and the Kolmogorov extension theorem guarantees the existence of a stochastic process with a

given system of finite-dimensional distributions.

Regarding the sample path properties of the Levy process, it can be shown that (i)-(iv) imply the existence of a modification that is right-continuous with left limits, or continue á droite, limite á gauche, in French, denoted as cádlág, see Sato (1999) [63]. Hence, for every Levy process in law, there exist $\Omega_0 \in \mathcal{F}$, $P(\Omega_0) = 1$, such that $\forall \omega \in \Omega_0$, $X_t(\omega)$ has cadlag sample paths. We will refer to this cádlág modification as a Levy process throughout. In addition to the right-continuity of their sample paths one can also show that they have at most countable number of random jump discontinuities occurring at random times, on each finite time interval.

Even though, above we defined a Levy process only on the positive real axis $\{Z_t\}_{t \geq 0}$, it can be extended it to the whole set of real numbers, i.e. $\{Z_t\}_{t \in \mathbb{R}}$, provided that on the negative time axes $t < 0$, process is defined as follows,

Definition 1.2.2. *Let $\{Z_t\}_{t \geq 0}$ be a Levy processes with $E[e^{izZ_t}] = e^{t\kappa_0(z)}$. Denote $\{\tilde{Z}_t\}_{t \geq 0}$ as an independent copy of $\{Z_t\}_{t \geq 0}$. Then the Levy process on the negative time axis $t < 0$, is defined as follows, for any $t < 0$, $Z_t \stackrel{\mathcal{D}}{=} -\tilde{Z}_{(-t)-}$.*

Below we listed some important properties of the Levy process.

- (i) The stationary and independent increments property imply that a Levy process is a Markov process, more precisely a temporally homogeneous Markov process. Combined with the almost sure right continuity of sample paths, one may show that Levy processes are strong Markov processes as well.
- (ii) They form a special subclass of semimartingales.
- (iii) They are 'analogs' of random walks in continuous time.

The next section addresses the specific type of marginal distributions that Levy processes can have by introducing a class of infinitely divisible distributions.

1.3 Infinitely divisible distributions

1.3.1 Infinite divisibility and the Levy-Khintchine representation

In this section we will define infinitely divisible distributions, give their representation and main properties. We will also briefly explain their correspondence to the Lévy processes in distribution. As standard notation we will write: $\mu^n = \mu * \dots * \mu$ as the n -fold convolution of a probability measure μ with itself.

Definition 1.3.1. *A probability measure μ on \mathbb{R} is said to be infinitely divisible (ID) if, for any positive integer n , there exist a probability measure μ_n on \mathbb{R} such that: $\mu = \mu_n^n$.*

Since the convolution of measures corresponds to the product of characteristic functions or sum of independent random variables, we can therefore say that probability measure μ is *infinitely divisible (ID)* if and only if, for each n , the n -th root of the characteristic function $\hat{\mu}$ is also a characteristic function of some probability measure: $\hat{\mu}(z) = (\hat{\mu}_n)^n$. Equivalently, one can also say that a random variable X is infinitely divisible (ID) iff for any $n \in \mathbb{N}$ there exists an i.i.d sequence of random variables Y_n such that: $X \stackrel{D}{=} Y_1^{(n)} + \dots + Y_n^{(n)}$.

As it turns out, all infinitely divisible distributions have a specific form of their characteristic function which is very well known representation, named the Lévy-Khintchine representation or the Lévy-Khintchine formula, see below.

Theorem 1.3.1 (Lévy-Khintchine). *If μ is an infinitely divisible distribution on \mathbb{R} , then*

there exists a unique triplet (A, ν, γ) (referred to as the generating triplet) such that:

$$\hat{\mu}(z) = e^{\kappa(z)} = \exp\left\{-\frac{1}{2}Az^2 + i\gamma z + \int_{\mathbb{R}} \left(e^{izx} - 1 - izx\mathbf{1}_{[-1,1]}(x)\right) \nu(dx)\right\}, \quad z \in \mathbb{R}. \quad (1.5)$$

where $\gamma \in \mathbb{R}$, $A \geq 0$ and ν is the Lévy measure, that is, a σ -finite Borel measure on \mathbb{R} such that

$$\nu(\{0\}) = 0, \quad \int_{\mathbb{R}} (1 \wedge x^2) \nu(dx) < \infty. \quad (1.6)$$

Conversely, given a generating triplet (A, ν, γ) , there exists an infinitely divisible distribution μ on \mathbb{R} whose characteristic function is given by (1.5). The exponent $\kappa(z) = \log(\hat{\mu}(z))$, in the above formula, is often referred to as the Levy exponent.

From the above theorem, one can conclude that the generating triplet (A, ν, γ) uniquely determines the infinitely divisible random variable X with law μ . The parameters A, ν, γ are called, respectively, the Gaussian term, the Lévy measure of μ , and the drift.

Bellow we have provided some examples of well known distributions that are elements of the class of infinitely divisible distributions and some that are not. Note, in order to show that a specific distribution belongs to the class of infinitely divisible distributions, one can either show that its characteristic function is of the Levy-Khinchine form (1.5) or directly use the definition 1.3.1 and suggest how to re-parameterize the form of its characteristic function in order to get the n -th root.

Example 1.3.1.

(i) Let X be a Gaussian random variable: $X \sim \mathcal{N}(\gamma, A)$. It is infinitely divisible since its

characteristic function is of the form,

$$\varphi_X(z) = e^{n(iz\frac{\gamma}{n} - \frac{1}{2}\frac{A^2}{n}z^2)} = (\varphi_{X_n}(z))^n$$

, where the n -th root is $X_n \sim \mathcal{N}(\frac{\gamma}{n}, \frac{A^2}{n})$. On the other hand, one can also see that $\varphi_X(z)$ is of the Levy-Khinchine form (1.5), with generating triplet $(A, 0, \gamma)$, where the variance of a normal variable is related to the Gaussian term A , a mean is related to the drift term γ , while the Levy measure is zero, $\nu = 0$. Note here, the Levy measure is zero, when and only when the distribution is Gaussian.

(ii) Let a Gamma distribution be defined by the following density,

$$\Gamma(\alpha, \beta) \sim \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x} \mathbf{1}_{[0, \infty)}(x), \alpha > 0, \beta > 0.$$

Similarly, it belongs to the class of infinitely divisible distributions. While the infinite divisibility can be easily seen by suggesting the n -th root of its characteristic function, i.e.

$$\varphi_{\Gamma(\alpha, \beta)}(z) = \left(1 - \frac{iz}{\alpha}\right)^{-\beta} = \left(\left(1 - \frac{iz}{\alpha}\right)^{-\beta/n}\right)^n = \left(\varphi_{\Gamma(\alpha, \beta/n)}(z)\right)^n,$$

the Levy-Khinchine form of its characteristic function is not as straightforward to find. It can be shown that it is given by the following expression,

$$\varphi_{\Gamma(\alpha, \beta)}(z) = e^{\beta \int_0^\infty (e^{izx} - 1)(e^{-\alpha x}/x) dx}.$$

From the above representation one can see that the generating triplet is of the form, $A = 0$, $\gamma = \int_0^1 x\nu(dx)$, and $\nu(dx) = (e^{\alpha x}/x)dx$ and that Levy measure has density with respect

to Lebesgue measure. In addition, since an exponential is a special case of the Gamma distribution, it is also infinitely divisible.

(iii) Following the similar procedure as above, one can see that the Poisson distribution is infinitely divisible as well, i.e.

$$\varphi_{P(\lambda)}(z) = e^{-\lambda(1-e^{iz})} = \left(e^{-\lambda/n(1-e^{iz})} \right)^n = \left(\varphi_{P(\lambda/n)}(z) \right)^n.$$

After rewriting the above characteristic function, one can see that the Levy-Khinchine exponent is of the form:

$$\kappa_{P(\lambda)}(z) = \lambda \int_{\mathbb{R}} \{ e^{izx} - 1 - izx \mathbf{1}_{[-1,1]}(x) \} \delta_1(dx),$$

and hence the generating triplet for Poisson distribution is given by, $(0, \lambda \delta_1, \int_{-1}^1 x \delta_1(dx))$, where δ_1 denotes the Dirac measure supported on $\{1\}$.

(iv) The compound Poisson distribution is given as: $\sum_{i=1}^N \xi_i \sim \mu$, where N is a Poisson random variable with parameter λ , $\{\xi_i\}_{i \geq 1}$ an i.i.d sequence of random variables with common law F , and it is assumed to be independent of N . Its characteristic function is of the form

$$\hat{\mu}(z) = e^{\lambda \int_{\mathbb{R}} (e^{izx} - 1) F(dx)} = e^{\lambda(\hat{F}(z) - 1)} = \left(e^{\lambda/n(\hat{F}(z) - 1)} \right)^n.$$

Hence, it belongs to the class of infinitely divisible distributions with the generating triplet, $(0, \lambda F(dx), \lambda \int_{-1}^1 x F(dx))$.

(v) Distribution of a Levy process $\{Z_t\}_{t \geq 0}$, at any time t , is an infinitely divisible distribution,

see below. Trivially we have that for any $n \geq 1$ and $t \in [0, \infty)$,

$$Z_t = (Z_{t_n} - Z_{t_{n-1}}) + \dots + (Z_{t_1} - Z_{t_0}),$$

where $t_k = kt/n$, $k = 0, 1, \dots, n$ and $t_n \equiv t$. Stationarity and independence of the Levy increments yields that, $\{Z_{t_k} - Z_{t_{k-1}}\}_{k=0}^n$ is a sequence of i.i.d random variables, hence, for a fixed time t , the random variable Z_t is infinitely divisible.

(vi) In general it turns out that, any distribution with bounded support (except δ -distributions) is not infinitely divisible. Hence, uniform and Binomial distribution are the examples of distributions that are not infinitely divisible.

Remark 1.3.1. Some other continuous distributions that are infinitely divisible are: Cauchy, one-sided strictly stable distributions of index $1/2$, while other discrete infinitely divisible distributions are geometric, negative binomial, etc... In the examples given above one can easily find the n -th root of the characteristic function, but there are many other distributions whose infinite divisibility is much harder to prove, like for example: Student's t -distribution, Pareto, F -distribution, Gumbel, Weibull, log-normal, logistic distribution, and some others, see [63]

For the sake of completeness, some interesting properties of infinitely divisible distributions are given in the form of lemmas below, see Sato (1999) [63][chapter 2] for more details.

Lemma 1.3.2.

(i) If μ is infinitely divisible distribution then its characteristic function $\hat{\mu}(z)$ has no zeros, i.e. $\hat{\mu}(z) \neq 0$ for any $z \in \mathbb{R}$.

- (ii) If μ_1 and μ_2 are infinitely divisible then their convolution, $\mu_1 * \mu_2$ is infinitely divisible.
- (iii) If $\{\mu_k\}_{k \geq 0}$ is a sequence of infinitely divisible distributions and $\{\mu_k\} \rightarrow \mu$, as $k \rightarrow \infty$, then the limiting distribution μ is infinitely divisible.
- (iv) If μ is infinitely divisible then, for any $t \in [0, \infty)$, μ^t is both, well-defined and infinitely divisible. Note $\mu^0 = \delta_0$.
- (v) Every infinitely divisible distribution is the weak limit of a sequence of compound Poisson distributions.

1.3.2 The correspondence between infinitely divisible distributions and Levy processes in law

Let us now present the details of a well known fact that there is one-to-one correspondence between the class of infinitely divisible distributions and Levy processes in law. We will explain this relationship in a very simple manner and then we'll state the theorem.

From the example 1.3.1 we already know that the Levy process Z_t is infinitely divisible for any fixed $t \geq 0$, which means that for any integer n , we can decompose Z_t as a sum of n i.i.d reparametrized distributions. Hence, the more interesting question is if we can decompose it as sum of t , i.i.d. Z_1 -distributions, where t is a real number, i.e. if and when the following is true, for any Levy process $\{Z_t\}$, and for any time $t \in \mathbb{R}^+$, $t \geq 0$,

$$\varphi_{Z_t}(z) = E[e^{izZ_t}] = e^{t\kappa(z)} = (\varphi_{Z_1})^t, \quad \kappa(z) = \log(E[e^{izZ_1}]).$$

From the following relationship, $Z_t = (Z_{t_n} - Z_{t_{n-1}}) + \dots + (Z_{t_1} - Z_{t_0})$, by taking specific

integer valued times, $t = m, n = m$ and $t = m, n = n$, one can easily get,

$$(\varphi_{Z_1})^m = \varphi_{Z_m} = (\varphi_{Z_{m/n}}(z))^n.$$

For any $n, m \in \mathbb{N}$ and hence for any rational number, the following equation holds,

$$\varphi_{Z_{m/n}} = (\varphi_{Z_1})^{\frac{m}{n}}.$$

If we simply choose the sequence of rationals $\{r_n\}$ such that $r_n \rightarrow t$, together with the fact that the Levy process is stochastically continuous we get that, $Z_{r_n} \rightarrow Z_t$ in probability, which implies convergence in distribution. This in turn implies, $Z_t \stackrel{\mathcal{D}}{=} (Z_1)^t$ and $\varphi_{Z_t} = e^{t\kappa(z)}$, for any real time $t \geq 0$ and not only for integer or rational.

From the above discussion it is clear that each Levy process can be associated with an infinitely divisible distribution, but what is not straightforward is, whether given an infinitely divisible distribution μ , one can construct a Levy process Z_t , such that $Z_1 \stackrel{\mathcal{D}}{=} \mu$. This question is answered in the theorem below, see Sato 1999, see [63] for more details. Intuitively, starting with an infinitely divisible distribution one needs to construct a corresponding Levy process in law, i.e. one needs to suggest an appropriate system of finite dimensional distributions such that it is well defined (satisfy Kolmogorov consistency theorem), $Z_1 \stackrel{\mathcal{D}}{=} \mu$, and have properties of the Levy process. If one shows that the proposed system of f.d.d. is well defined, then one only needs to show that the system is a Levy process in law. See the theorem below,

Theorem 1.3.3. *Let $\{Z_t\}_{t \geq 0}$ be a real-valued Levy process in law and denote $Z_1 \stackrel{\mathcal{D}}{=} \mu$.*

(i) *Then for every $t \in [0, \infty)$, Z_t is infinitely divisible and $Z_t \stackrel{\mathcal{D}}{=} (Z_1)^t \stackrel{\mathcal{D}}{=} \mu^t$.*

(ii) Conversely, if μ is an infinitely divisible distribution on \mathbb{R} , then there is a real-valued Levy process in law, denoted with $\{Z_t\}_{t \geq 0}$, such that $Z_1 \stackrel{\mathcal{D}}{=} \mu$.

(iii) If $\{Z_t\}$ and $\{Z'_t\}$ are two real-valued Levy process in law such that $Z_1 \stackrel{\mathcal{D}}{=} Z'_1$ then the processes are identical in law, i.e. $\{Z_t\}_{t \geq 0} \stackrel{\mathcal{D}}{=} \{Z'_t\}_{t \geq 0}$.

Note, μ is said to be an infinitely divisible distribution corresponding to the Levy process in law $\{Z_t\}$, where $Z_1 \stackrel{\mathcal{D}}{=} \mu$; conversely, the process $\{Z_t\}$ is said to be the Levy process in law corresponding to the infinitely divisible distribution μ . A generating triplet for the infinitely divisible distribution μ , denoted as (A, γ, ν) , is also called the Levy generating triplet.

Remark 1.3.2. *Even though the stationary and independent increments property of the Levy process $\{Z_t\}$ is enough to claim that its marginal distributions are infinitely divisible, it is not enough in order to prove the property (i). One still needs the stochastic continuity.*

Below we give couple of well known examples of the Levy processes. We are going to assume that (Ω, \mathcal{F}, P) is a probability space and that a given process is defined on it.

Example 1.3.2.

(i) A standard Brownian motion is a classical example of the Levy process with normal marginal distributions, i.e. a real-valued process with stationary and independent increments, starts at zero a.s., has a modification with continuous sample paths, and at each time $t > 0$, process B_t has normal distribution with variance t , or $B_t \sim \mathcal{N}(0, t)$. Note also that a linear Brownian motion or a BM with drift, i.e. $\tilde{B}_t \equiv B_t + \gamma t$ is a Levy process too.

(ii) A continuous time and a non-negative integer-valued process $\{N_t\}_{t \geq 0}$ is said to be a Poisson process with intensity $\lambda > 0$, if it is a Levy process and, in addition, at each time $t > 0$, N_t has a Poisson distribution with parameter λt . Since it is a Levy process, the characteristic function is of the Levy-Khintchine form, $E e^{izN_t} = e^{-\lambda t(1-e^{iz})}$.

Remark 1.3.3. *Note, another way of defining the Poisson is as a counting process. In fact, it is the only process that is both, Levy and counting at the same time. The Poisson counts the arrivals of randomly occurring events or events with exponential inter-arrival times in the time interval $[0, t]$, i.e. $N_t = \sum_{n \geq 1} \mathbf{1}_{t \geq T_n}$, where T_n is a random walk with exponentially distributed waiting times, $T_n = \sum_{i=1}^n \tau_i$, $\tau_i \sim \exp(\lambda)$. Since it is a Levy process, the sample paths of Poisson are a.s. right continuous by the definition, but some other path properties come more intuitively when using the second definition. Hence, from here we may also conclude that the trajectories of a Poisson are always increasing and piece-wise constant functions, with jumps of size 1 only, where jumps occur whenever the random events occur.*

Note, also that a Poisson process itself cannot be used to model asset prices because the jump size is always equal to one, which is way too restrictive, but it can be used as building block to construct more advanced financial models that exhibit jumps.

(iii) Let $\lambda > 0$ and let F be a distribution on \mathbb{R} such that $F\{0\} = 0$. A real-valued and continuous time process is a compound Poisson process $\{Z_t\}_{t \geq 0}$ associated with λ and F , if it is a Levy process and at each time $t > 0$ it has a compound Poisson distribution, $E e^{izZ_t} = e^{-\lambda t(1-\hat{F}(z))}$, $z \in \mathbb{R}$.

Another way of defining a compound Poisson process with a jump intensity λ and a jump size distribution F is as a random walk, $Z_t = \sum_{i=1}^{N_t} \xi_i$, where $\{\xi_i\}_{i \geq 1}$ is a sequence of independent random variables with law F and N_t is a Poisson process with with intensity λt and independent of the given sequence $\{\xi_i\}_{i \geq 1}$. One can see that, the jumps arrive randomly according to a Poisson process and the size of the jumps is also random, with a specified probability distribution.

Remark 1.3.4. *Similarly as a Poisson process, a compound Poisson process is a piece-wise*

constant process which jumps at jump times of a standard Poisson process, but as opposed to a Poisson which has fixed jump length of one, the jump sizes of a compound Poisson are i.i.d. random variables with a given law F .

(iv) As already explained, gamma distribution is infinitely divisible, hence by the theorem 1.3.3 there exist a Levy process Z_t with gamma marginal distributions, such that $Z_1 \stackrel{\mathcal{D}}{=} \Gamma(\alpha, \beta)$ and $Z_t \stackrel{\mathcal{D}}{=} (\Gamma(\alpha, \beta))^t$. Since Gamma distribution is closed under convolution in the first parameter (skewness parameter) we get, $Z_t \stackrel{\mathcal{D}}{=} \Gamma(t\alpha, \beta)$, i.e. the characteristic function of the gamma process is of the form, $E[e^{izZ_t}] = (1 - iz/\beta)^{-t\alpha}$.

Remark 1.3.5. *Stationarity, independent increments property and the fact increments $(Z(1))$ are gamma distributed imply that the process is a.s. strictly increasing. Even though it seems similar to a compound Poisson where the jump distribution F is concentrated on the positive real line $(0, \infty)$, there are two main differences. First, Levy measure of a Gamma process has infinite total mass unlike the Levy measure of a compound Poisson, which is necessarily finite and equal to the arrival rate of jumps¹. Second, while a compound Poisson with positive jumps does have an a.s. non-decreasing paths, it does not have paths that are almost surely strictly increasing².*

Remark 1.3.6. *Note that the Poisson and the compound Poisson process with strictly positive increments are the most well-known non-negative Levy processes. As mentioned their paths are non-decreasing piece-wise constant functions where jumps happen rather rarely. Consequently, their increments are often exactly equal to zero, even when measured over large time intervals. This feature is fundamentally different from the Gamma process since its increments are strictly positive regardless of how much time has elapsed. This means that*

¹Frequently referred to as jump intensity.

²Levy processes whose paths are almost surely non-decreasing are called subordinators.

at every time point this process jumps, hence it will have infinity number of small jumps in any finite time interval. The Levy processes that exhibit this feature are said to be of infinite activity (jump intensity). The path of these processes have rough upward trend with occasional large jumps and intuitively one can conclude that the paths are not continuous anywhere.

From the Definition 1.2.1 alone, it is difficult to see, just how rich the class of Levy processes really is. In general, there are many processes that satisfy the definition of the Levy process and on the first encounter they can seem considerably different from one another. For example, firstly, the Brownian motion has an a.s. continuous sample paths whereas the Poisson process is a jump process only. Secondly, the Brownian motion is a.s. nowhere differentiable, i.e. its paths are of unbounded variation over finite time horizons, while the Poisson process is a non-decreasing process and thus has paths of bounded variation over finite time horizons. Still, they both belong to a one large class of processes which makes modeling with Levy processes very flexible and suitable for various situations.

Beside their flexibility with the sample path properties which allows for both continuous or jump sample paths, they have also larger class of possible marginal distributions, as mentioned previously, class of infinitely divisible distributions. Therefore, the probability models driven by Levy processes are applicable to many different situations. One of the drawbacks when modeling directly with Levy is their independent increments property, which is not true in many real life situations. Hence, if one wants to model phenomena that has correlation structure present when modeling squared or absolute increments, Levy processes would not be suitable. So, from the modeling and application perspective, because of their correlation structure, the OU and OU type processes, see the section 1.4, allow for more flexibility when fitting the data. These can be also used to define more advanced models,

referred to in literature as superpositions of OU type processes, see the section 1.5. They were first introduced by Barndorff-Nielsen and other collaborators, see [3, 4, 8] and the class of superpositions is richer than standard OU type processes.

1.3.3 Self-decomposable distributions

In this section we will define a class of *self-decomposable* (*SD*) distributions, which, due to the specific form of their Levy measure, is a subclass of infinitely divisible distributions.

We will digress here and try to explain some applicational motivation behind the processes with self-decomposable distributions. The most standard models for option pricing (Black and Scholes, 1973 [12]) usually assume that continuously compounded returns are normally distributed. The primary motivation behind this assumption is the central limit theorem which says that if the price returns are realized as sum of large number of independent influences, properly scaled, then returns should be normally distributed. Now, the fact that distribution of the price returns for time-series and option data is not normal has lead many scientist to research into jump-diffusion models, stochastic volatility models, pure-jump Levy processes and some other. So, the question is, are there some alternatives to the Gaussian distribution as a limiting law, i.e. can we construct models that are limits of properly scaled partial sums, where the scaling function is not necessarily \sqrt{n} ?³ The answer is yes there is a class of distributions with these properties, so-called laws of class *L*, which were first defined by Khintchine (1938) and Levy (1937) as limiting distributions for sums of n independent variables that are centered and scaled by functions of n , and this class is identical to the class of self-decomposable distributions. Examples would be Normal and stable distributions. Regardless, the self-decomposable laws represent an important generalization of Gaussian

³There is no economic motivation for the scaling function to be \sqrt{n} over any other function of n .

and stable laws as they describe limit laws with more general scaling constants than only $\frac{1}{\sqrt{n}}$. This flexibility is important in applications when the independent sources of randomness being summed are of different orders of magnitude. Now that we explained why these models are interesting from finance/modeling perspective let's proceed with the definition.

Definition 1.3.2. *A random variable X is self-decomposable if, for all $c \in (0, 1)$, its characteristic function can be decomposed/factorized as, $\varphi_X(z) = \varphi_{cX}(z) \times \varphi_c(z)$, for some characteristic function $\varphi_c(z)$, $z \in \mathbb{R}$.*

Intuitively, a random variable is *self-decomposable* if it can be decomposed into a sum of scaled down version of itself and an independent residual term, i.e. if for every $c \in (0, 1)$, there exists an independent random variable X_c , such that: $X \stackrel{D}{=} cX + X_c$.

Remark 1.3.7. *To summarize, the self-decomposable laws arise naturally as limiting distributions for certain generalizations of the central limit problem. More generally, the class of all infinitely divisible distributions coincides with those distributions that arise as limits of row sums of uniformly asymptotically negligible triangular arrays of random variables.*

There is another important characterization of self-decomposable distributions and that is through a special form of their Lévy measure. In particular, since self-decomposable distributions are a subclass of infinitely divisible distributions, their characteristic function is already of the Lévy-Khintchine form, but it can be shown that their Lévy measure has a density with respect to the Lebesgue measure, see Sato (1999), [63][corolary 15.11], i.e. their Lévy measure is of the following form:

$$\nu(dx) = \frac{k(x)}{|x|} dx, \tag{1.7}$$

for some non-negative function $k(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$, that is increasing on $(-\infty, 0)$ and decreasing on $(0, \infty)$ and also satisfies the integrability condition: $\int_{-\infty}^{\infty} (1 \wedge x^2) \frac{k(x)}{|x|} dx < \infty$.

Another interesting property of self-decomposable distributions is that their densities are unimodal, see Yamazato 1978 [68] or Sato 1999 [63]. The proof that every density in the class of L or self-decomposable class is unimodal is non-trivial and an interested reader may refer to the references suggested above. This property can be helpful in determining if some distribution is not self-decomposable. Below is the example where a distribution is infinitely divisible but not self-decomposable.

Example 1.3.3.

(i) Any stable distribution on \mathbb{R} is self-decomposable. For example, let's look at the one-sided α -stable law, for $0 < \alpha < 2$. Levy generating triplet is of the form $(0, \nu, \gamma)$, where $\gamma \in \mathbb{R}$ and Levy measure is of the form,

$$\nu(dy) = \begin{cases} C\alpha y^{-\alpha-1} dy, & y > 0 \\ 0, & y < 0. \end{cases}$$

Directly from the above definition, one can see that the given Levy measure has density with respect to Lebesgue, hence it is self-decomposable. In addition, the characteristic function of a stable distribution is given by the following formula,

$$\hat{\nu}(z) = \begin{cases} e^{-C\Gamma(1-\alpha)(-iz)^\alpha}, & \text{for } 0 < \alpha < 1, \\ e^{C\frac{\Gamma(2-\alpha)}{(\alpha-1)}(-iz)^\alpha}, & \text{for } 1 < \alpha < 2. \end{cases}$$

(ii) As mentioned previously, the Levy-Khinchine exponent of the Gamma distribution is given

by the following expression,

$$\varphi_{\Gamma(\alpha,\beta)}(z) = e^{-\beta \left(\int_0^\infty (e^{izx} - 1)(e^{-\alpha x}/x) dx \right)}.$$

From the above representation one can see that the Levy measure has density with respect to the Lebesgue measure, $\nu(dx) = (e^{-\alpha x}/x)dx$, hence it is self-decomposable. In addition, since exponential is a special case of Gamma distribution it is self-decomposable as well.

(iii) Other well known and self-decomposable distributions are double-exponential (Laplacian), log-normal, Pareto, t -distribution, F -distribution, etc...

(iv) An interesting example of a distribution that is infinitely divisible but not self-decomposable can be constructed in the following way. Take sum of independent normal and poisson distribution. The convolution is going to be infinitely divisible but not self-decomposable since for different parameters λ the resulting distribution can be multi-modal, while the self-decomposable distributions are unimodal. Another way to construct a continuous distribution that is infinitely divisible but not self-decomposable is via Levy-Khintchine representation while making sure that the Levy measure is not absolutely continuous with respect to the Lebesgue.

In the next section one will see a different view of self-decomposable distributions and that is as a stationary distribution of a class of Markov processes called Ornstein-Uhlenbeck type. This interpretation of a self-decomposable distribution gives another representation of its Levy measure as opposed to the equation (1.7), given above.

1.4 Ornstein-Uhlenbeck type processes

Named after Leonard Ornstein and George Eugene Uhlenbeck, real-valued Ornstein-Uhlenbeck (OU) process $\{Y_t\}_{t \geq 0}$, induced by a Brownian motion B_t , is an a.s. and unique solution to the following stochastic differential equation (SDE), commonly referred to as the Langevin SDE,

$$dY_t = -\lambda Y_t dt + dB_t.$$

A standard Ornstein-Uhlenbeck process has many useful modeling properties, for example a fact that it is a continuous time analogue of discrete AR(1) process, relatively straightforward to simulate, it has a mean reversion property (converges towards its long-term mean, as $t \rightarrow \infty$). However, a potential drawback of modeling via standard OU process is that it can fit only data that exhibit normality. Their exponentially decaying dependence structure is not very flexible either, so there is an obvious need for a model that improves on these features.

In this thesis our focus will be on the processes of Ornstein-Uhlenbeck type (OU type for short) which are analogues of the above mentioned ordinary OU processes where a Brownian motion part is replaced by a general Levy process $\{Z_t\}_{t \geq 0}$,

$$dY_t = -\lambda Y_t dt + dZ_t.$$

As we will see later, the class of possible marginal distributions for the OU type processes coincides with the class of self-decomposable distributions, which includes Gamma, Inverse Gaussian, Normal Inverse Gaussian, and many other distributions useful for applications in finance. Self-decomposable distributions are specifically useful when modeling data that

exhibits heavier tails and higher peaks than normal (frequently referred to as *Leptokurtic distributions*). Hence, the OU type models are more flexible when fitting empirical data.

Here we summarize some of the available definitions and results from the Sato's and Rocha-Arteaga's book [60]. In this chapter, we give a layout for the construction of an Ornstein-Uhlenbeck type process, which is a necessary building block for the upcoming chapters. As suggested in [60][Chapter 2], in order to construct an Ornstein-Uhlenbeck type process, first a stochastic integral of a deterministic integrand with respect to a Levy process $\{Z_t\}$ needs to be defined, which is done in the subsection 1.4.1 by following the reference provided. Afterwards, a stochastic integral of a step function on a bounded time interval is defined. Due to the fact that any measurable function can be approximated with a sequence of simple functions, the Levy integral can then be defined as a limit in probability of the previously constructed simple integrals. The Ornstein-Uhlenbeck type process is afterwards introduced as a stochastic integral with respect to a Levy process, $\{Z_t\}$. Now as suggested in [60,63], since the Levy integral is well defined, the existence and uniqueness of an Ornstein-Uhlenbeck type process is established as an almost sure, strong and unique solution to the Langevin integral equation, see the section 1.4.2.

The section 1.4.3 has important results on the stationarity of the Ornstein-Uhlenbeck type process and its Markovian property, see Sato's book [63] and [1,60] for more details. It is well known that the ordinary OU process driven by the Brownian motion has a limiting stationary distribution as $t \rightarrow \infty$, which is Gaussian. On the other hand, the OU type process doesn't necessarily have a limiting stationary distribution, since it is driven by a general Levy noise that tends to have more erratic behavior due to jumps. In this section we will give the conditions on the Levy measure related to the driving Levy process, under which processes of OU type have stationary limiting distributions. After establishing the limiting

stationary distribution of an OU type process, we will briefly explain a correspondence between self-decomposable distributions and stationary processes of OU type.

For more properties of OU type processes and their generalizations see Mandrekar & Rudiger (2007) [54], Barndorff-Nielsen (2001) [4], Barndorff-Nielsen & Stelzer (2011) [9].

1.4.1 Stochastic integral with respect to the Levy process

In this section we summarize the approach to the definition of the stochastic integral of a bounded measurable function defined on a bounded closed interval on \mathbb{R} with respect to a given Levy process, see the details in [60]. Note that this is a special case of integration with respect to semi-martingale. In addition, the expression for the characteristic function of the stochastic integral with respect to a given Levy process, is provided, refer to the same [1, 60, 63]. It turns out that its characteristic function depends on the characteristic function of the driving Levy process.

We will start by defining a step function first.

Definition 1.4.1. *Let $0 \leq t_0 < t_1 < \infty$. We say that a function f , defined on the interval $[t_0, t_1]$, is a step function if, for the finite number of points $t_0 = s_0 < s_1 < \dots < s_n = t_1$ and for some $a_j \in \mathbb{R}$, $j = 1, \dots, n$, it can be written in the following form,*

$$f(s) = \sum_{j=1}^n a_j \mathbf{1}_{[s_{j-1}, s_j]}(s). \quad (1.8)$$

As we explained, first we will define the stochastic integral for a given step function f , defined on a bounded closed interval $[t_0, t_1]$, with respect to a given Levy process:

Definition 1.4.2. *Let $\{Z_t\}_{t \geq 0}$ be a Levy process on \mathbb{R} , with $Z_1 \stackrel{\mathcal{D}}{=} \mu_0$ uniquely determined*

by a generating triplet of $\mu_0 \sim (A_0, \nu_0, \gamma_0)$ and the characteristic function of the form,

$$\varphi_{Z_t}(z) = E[e^{izZ_t}] = e^{t\kappa_0(z)}, \quad (1.9)$$

where $\kappa_0(z)$ is the Levy exponent of $\hat{\mu}_0$ and it is given by,

$$\kappa_0(z) = -\frac{1}{2}A_0z^2 + i\gamma_0z + \int_{\mathbb{R}} \left(e^{izx} - 1 - izx\mathbf{1}_{[-1,1]}(x) \right) \nu_0(dx). \quad (1.10)$$

Then the Levy integral for a given step function f is defined as,

$$\int_{t_0}^{t_1} f(s)dZ_s = \sum_{j=1}^n a_j (Zs_j - Zs_{j-1}). \quad (1.11)$$

Following the steps in [60], we present with properties of the above defined Levy integral of a simple function.

Proposition 1.4.1. *The distribution of the random variable $Y \equiv \int_{t_0}^{t_1} f(s)dZ_s$ is infinitely divisible for any step function f , and its characteristic function is given in the form below:*

$$\varphi_Y(z) = E\left\{ \exp \left(iz \int_{t_0}^{t_1} f(s) dZ_s \right) \right\} = \exp \left\{ \int_{t_0}^{t_1} \kappa_0(zf(s)) ds \right\} \quad (1.12)$$

Proof. By the independent and stationary increment property of the Levy process combined

with the expression (1.9) we get,

$$\begin{aligned}
E\left\{\exp\left(iz\int_{t_0}^{t_1}f(s)dZ_s\right)\right\} &= \prod_{j=1}^n E\left\{e^{iz a_j(Zs_j-Zs_{j-1})}\right\} \\
&= \prod_{j=1}^n e^{(s_j-s_{j-1})\kappa_0(a_j z)} \\
&= \exp\left(\sum_{j=1}^n(s_j-s_{j-1})\kappa_0(a_j z)\right) \\
&= \exp\int_{t_0}^{t_1}\kappa_0(zf(s))ds
\end{aligned}$$

The infinite divisibility of the Levy integral in (1.11) is a direct consequence of its definition. □

In the subsequent steps we still follow the reference [60], hence we define the Levy integral for a larger class of functions, real-valued bounded measurable functions on $[t_0, t_1]$

Proposition 1.4.2. *Let $f(s)$ be a real-valued bounded measurable function on $[t_0, t_1]$, such that there are uniformly bounded step functions $\{f_n(s)\}_{n \geq 1}$ on $[t_0, t_1]$ satisfying $f_n \rightarrow f$. Then, the Levy integral $\int \mathbf{1}_{[t_0, t_1]} f_n dZ_s$ converges to a real-valued random variable Y in probability. The law of Y is infinitely divisible and its characteristic function is given by:*

$$Ee^{izY} = \exp\left(\int_{t_0}^{t_1}\kappa_0(zf(s))ds\right). \tag{1.13}$$

Note, the limit Y does not depend on the choice of f_n up to a probability zero.

Proof. As suggested in the reference [60], we start with lemma from the Sato's book [63]. Even though the characteristic function is uniformly continuous, it is not straightforward to show that the Levy characteristic exponent is continuous. The Lemma 7.6 (pg 33), in Sato

1999, says that any continuous function $\varphi(z) : \mathbb{R} \rightarrow \mathbb{C}$ that is $\varphi(0) = 1$ and $\varphi(z) \neq 0$ for any z has a unique and continuous Levy exponent $\kappa(z) : \mathbb{R} \rightarrow \mathbb{C}$ such that $\kappa(0) = 0$ and $e^{\kappa(z)} = \varphi(z)$. Now, due to the continuity of the Levy exponent κ_0 in (1.10), one gets the convergence of $\kappa_0((f_n(s) - f_m(s))z) \rightarrow 0$ for almost every s , as $n, m \rightarrow \infty$. Since the step functions are assumed to be uniformly bounded and the Levy exponent κ_0 is continuous, the Dominated Convergence theorem implies,

$$\int_{t_0}^{t_1} \kappa_0((f_n(s) - f_m(s))z) ds \rightarrow 0, \quad \text{as } n, m \rightarrow \infty.$$

By (1.12),

$$E e^{iz \int_{t_0}^{t_1} (f_n(s) - f_m(s)) dZ_s} = e^{\int_{t_0}^{t_1} \kappa_0(z(f_n(s) - f_m(s))) ds} \rightarrow 1, \quad (1.14)$$

which implies convergence to zero in probability of the above sequence, i.e. $\int_{t_0}^{t_1} (f_n(s) - f_m(s)) dZ_s \xrightarrow{P} 0$. This means that the sequence of integrals is Cauchy in probability, and since the metric of convergence in probability is complete one may conclude that, there exists a random variable Y which is limit in probability of the sequence of random variables defining the Levy integral $\int_{t_0}^{t_1} f_n(s) dZ_s$.

The law of Y is infinitely divisible because the lemma 1.3.2 implies that the weak limit of infinitely divisible sequence of distributions is infinitely divisible. Since the sequence of simple Levy integrals $\int_{t_0}^{t_1} f_n(s) dZ_s$ is infinitely divisible, see the proposition 1.4.1 and the rest follows.

It is straightforward to get the convergence in distribution for the sequence of Levy integrals. The continuity of the Levy exponent κ_0 implies, $\kappa_0(f_n(s)z) \xrightarrow{n \rightarrow \infty} \kappa_0(f(s)z)$, and the DCT, $\int_{t_0}^{t_1} \kappa_0(f_n(s)z) ds = \sum_{j=1}^n \kappa_0(a_j z)(s_j - s_{j-1}) \rightarrow \int_{t_0}^{t_1} \kappa_0(f(s)z) ds$. Finally the

proposition 1.4.1 implies, as $n \rightarrow \infty$,

$$E e^{iz \int_{t_0}^{t_1} f_n(s) dZ_s} \rightarrow e^{\int_{t_0}^{t_1} \kappa_0(f(s)z) ds}$$

The existence of the limit in probability, the above convergence of the characteristic functions combined with the continuity theorem give the distributional form of the limit, i.e. the above equation (1.13).

Uniqueness follows from the equation (1.14), with two approximating sequences given as $f_n(s), g_n(s)$ where both are uniformly bounded and converging a.s. to f . So, the following holds, $\int_{t_0}^{t_1} f_n(s) dZ_s - \int_{t_0}^{t_1} g_n(s) dZ_s \rightarrow 0$ in probability. \square

Remark 1.4.1. *It is explained in [60] that by using Lusin's theorem it is possible to construct a sequence of uniformly bounded step functions such that if f is a real-valued bounded measurable function on $[t_0, t_1]$, then $Y = \int_{t_0}^{t_1} f(s) dZ_s$ is definable and (1.4.2) holds.*

Definition 1.4.3. *The real-valued random variable Y in the proposition above is the stochastic integral of f with respect to the Levy process $\{Z_t\}$ and it is denoted by,*

$$Y \equiv \int_{t_0}^{t_1} f(s) dZ_s$$

Proposition 1.4.3. *If f is a real-valued bounded measurable function on $[t_0, t_1]$, then $Y = \int_{t_0}^{t_1} f(s) dZ_s$ is definable and (1.4.2) holds.*

Note, the above proposition shows the existence of the Levy integral on the bounded time interval $[t_0, t_1]$. Provided that the assumptions in the proposition 1.4.1 hold, couple of important properties of the Levy integrals will hold true, see the references for more details [1, 60]. See below.

Proposition 1.4.4 (Some properties of the Levy integrals).

(i) For any Levy integral of the type, defined in 1.4.3, the following holds,

$$\int_{t_0}^{t_1} f(s)dZ_{s+h} \stackrel{\mathcal{D}}{=} \int_{t_0}^{t_1} f(s)dZ_s.$$

The above statement is true due to the stationarity of Levy increments and due to the fact that convergence in probability implies convergence in distribution. In addition, if the distribution of $\int_0^\infty f(s)dZ_s$ exists, the following property holds as well, $\int_0^\infty f(s)dZ_{(s+h)} \stackrel{\mathcal{D}}{=} \int_0^\infty f(s)dZ_s$.

(ii) Provided that a Levy process on the negative time-axis is defined as, $\{Z'_t : t < 0\} \stackrel{\mathcal{D}}{=} \{-Z_t : t \geq 0\}$, where Z'_t is independent of Z_t , the following property is true,

$$\int_0^t e^{-\lambda(t-s)}dZ_s \stackrel{\mathcal{D}}{=} \int_0^t e^{-\lambda s}dZ_s.$$

Again, it holds because of the stationary increment of the Levy process and the fact that, $\int_0^t e^{\lambda s}dZ_{-s} \stackrel{\mathcal{D}}{=} -\int_t^0 e^{\lambda s}dZ_s$ as a consequence of stationarity of the Levy increments.

Remark 1.4.2. As soon as we start working with the integrals on the unbounded intervals, eg. $\int_{t_0}^\infty f(s)dZ_s$ we will need to strengthen the above assumptions in order to get a limiting distribution of $\int_{t_0}^{t_1} f(s)dZ_s$, as $t_1 \rightarrow \infty$. In the section 1.4.3.1, we will see that existence of a stationary OU type process is related to the existence of the limiting distribution of an unbounded Levy integral, hence a stronger set of assumptions will be established.

Note that, an additive process is a stochastically continuous process with independent increments that starts at zero a.s. It will be used in the next two proposition. Below we included two propositions, from [60] for the sake of completeness. Even though, we are not going to give a detailed proof of the existence of an OU type process, in order to define it,

the proposition of Fubini type involving the stochastic integrals is a necessary step so we included it here.

Proposition 1.4.5. *Let $f(s)$ be a locally bounded and measurable function on the interval $s \in [0, \infty)$. Then there exist an additive real-valued process $\{Y_t\}$ such that, for every $t > 0$ the following is true,*

$$P \left(Y_t = \int_0^t f(s) dZ_s \right) = 1.$$

Note, since it is understood that $\int_0^t f(s) dZ_s$ is a modification of Y_t , then the integral of the following form, $\int_{t_0}^{t_1} f(s) dZ_s$ is a modification of $Y_{t_1} - Y_{t_0}$.

Similarly, for the sake of completeness, we included the proposition that shows that Levy integral has a modification that is an additive process.

Proposition 1.4.6. *Let $f(s)$ and $g(s)$ be bounded measurable functions on $[t_0, t_1]$. Then*

$$\int_{t_0}^{t_1} g(s) ds \int_{t_0}^s f(u) dZ_u = \int_{t_0}^{t_1} f(u) dZ_u \int_u^{t_1} g(s) ds, \text{ a.s.} \quad (1.15)$$

1.4.2 Existence and uniqueness of the OU type process

Given a real-valued Levy process $\{Z_t\}_{t \geq 0}$ with a generating triplet (A_0, ν_0, γ_0) , satisfying the Levy-Khintchine formula (1.5), a real constant $\lambda > 0$ and a starting point $x \in \mathbb{R}$, we define a new process $\{Y_t\}_{t \geq 0}$ such that it is a.s. right continuous with left limits and it a.s. satisfies the integral equation given below:

$$Y_t(\omega) = x + Z_t(\omega) - \lambda \int_0^t Y_s(\omega) ds, \quad \text{for } t \geq 0 \quad (1.16)$$

Theorem 1.4.7. *Considering the equation (1.16) as an equation for $Y_t(\omega)$, there exist an*

almost sure and unique solution $Y_t(\omega)$, and it is given by:

$$Y_t(\omega) = xe^{-\lambda t} + e^{-\lambda t} \int_0^t e^{\lambda s} dZ_s(\omega). \quad (1.17)$$

In general, the solution proposed in (1.17) is well defined if the integral with respect to measure induced by the Levy process is well defined, see the section 1.4.1. Hence, the integral equation (1.16) has an a.s. strong and unique solution regardless of the type of the Levy process being used as a driving noise.

Remark 1.4.3. *In many applications, the driving Levy process is an increasing, positive Levy processes, frequently referred to as a subordinator. It can be shown that the sample functions of these processes are almost surely, of bounded variation on any finite time interval⁴, hence the above integral is well defined. So, in case of Levy being increasing (or non-decreasing) we can re-write the above integral equation (1.16) as a path-wise differential equation, which if viewed as a function of t and ω is considered stochastic.*

The OU type process can be viewed as an a.s. and unique solution to a stochastic differential equation (SDE) given below:

$$dY_t(\omega) = -\lambda Y_t(\omega)dt + dZ_t(\omega), \quad t \geq 0. \quad (1.18)$$

Since all the necessary elements were previously defined one can proceed with the definition of an Ornstein-Uhlenbeck type (or OU type) process.

Definition 1.4.4 (Ornstein-Uhlenbeck type process). *Let $\lambda > 0$, M be a real valued random variable, and Z_t being a Levy process on \mathbb{R} generated by the triple (A_0, ν_0, γ_0) , such that*

⁴This means that the last integral in (1.17) is a path-wise Lebesgue-Stieltjes integral in s .

M and $\{Z_t\}_{t \geq 0}$ are independent. The process that is an a.s. and unique solution to the equation,

$$Y_t = M + Z_t - \lambda \int_0^t Y_s ds, \quad \text{for } t \geq 0 \quad (1.19)$$

is called the Ornstein-Uhlenbeck type process generated by Z_t and λ , starting from $Y_0 = M$. One can also refer to it as the Ornstein-Uhlenbeck type process generated by the quadruple $(A_0, \nu_0, \gamma_0, \lambda)$. Equivalently, the appropriate stochastic differential equation is of the form,

$$dY_t = -\lambda Y_t dt + dZ_t, \quad Y_0 = M. \quad (1.20)$$

Note that, the above process Z_t is often referred to as *the background driving Levy process* (or *BDLP for short*) and that the parameter λ is usually referred to as the *mean reversion parameter*. Intuitively, the parameter λ controls how fast the process Y_t converges to its long term mean. Moreover, it controls the memory of the process Y_t through the autocorrelation function, see the section (2.3.4).

Remark 1.4.4. *Frequently in the literature, the BDLP of the given OU type process is rescaled from Z_t to $Z_{\lambda t}$, hence the above stochastic differential equation (1.20) becomes: $dY_t = -\lambda Y_t dt + dZ_{\lambda t}$. The given scaling is such that the marginal distributions of OU type process are free of the parameter λ . This way when fitting a distribution to the data at hand there will be one parameter less to estimate. The parameter λ is usually estimated from the empirical correlation function.*

1.4.3 Properties of the the OU type processes

1.4.3.1 The stationary OU type process

Recall that a process is stationary if and only if its family of f.d.d. is shift invariant, or in other words if for any $u \geq 0$, the following is true: $\{Y_t\}_{t \geq 0} \stackrel{D}{=} \{Y_{t+u}\}_{t \geq 0}$.

The Levy processes usually do not have limiting distributions as $t \rightarrow \infty$, except for a case of a zero process. On the other hand, since the OU type processes have drift toward the origin with the magnitude that is proportional to the distance from the origin, they may have a limit if certain assumptions are satisfied, in which case stationarity will hold.

One way to show the stationarity is to directly write down its characteristic function and see what are the conditions under which the limiting distribution exists, i.e. the distribution as $t \rightarrow \infty$. See below for more details, Barndorff (1998), see [5]. Another approach, proposed by Sato (1999) [63] is to explore Markovian properties of an OU type process and to find the necessary conditions under which the limiting distribution of a Markov process exists and is stationary.

To simplify, we will only explain the first approach and explore the stationarity of a marginal distribution of the OU type process, similarly as in Barndorff (1998), see [5]. The proposition below reveals simple, yet important property of the OU type process, which will be used throughout, for more details refer to [60].

Proposition 1.4.8. *Let $\{Y_t\}_{t \geq 0}$ be an Ornstein-Uhlenbeck type process given by the formula,*

$Y_t = e^{-\lambda t} Y_0 + \int_0^t e^{-\lambda(t-s)} dZ_s$, $\lambda > 0$, then the following is true:

$$Y_{t+u} = e^{-\lambda u} Y_t + \int_0^u e^{-\lambda(u-s)} dZ_s, \quad u \geq 0. \quad (1.21)$$

Proof.

$$\begin{aligned}
Y_{t+u} &= e^{-\lambda(t+u)}Y_0 + \int_0^{t+u} e^{-\lambda(t+u-s)}dZ_s \\
&= e^{-\lambda(t+u)}Y_0 + e^{-\lambda(t+u)} \left(\int_0^t e^{\lambda s}dZ_s + \int_t^{t+u} e^{\lambda s}dZ_s \right) \\
&= e^{-\lambda u} \left\{ e^{-\lambda t}Y_0 + \int_0^t e^{-\lambda(t-s)}dZ_s \right\} + e^{-\lambda(t+u)} \int_0^u e^{\lambda(s'+t)}dZ_{(s'+t)} \\
&\stackrel{D}{=} e^{-\lambda u}Y_t + e^{-\lambda u} \int_0^u e^{\lambda s'}dZ_{s'}.
\end{aligned}$$

The last step is true due to the fact that Levy process Z_t has stationary increments, i.e.

$$dZ_{(s'+t)} \stackrel{D}{=} dZ_{s'} \text{ which implies } \int_0^u e^{\lambda s'}dZ_{(s'+t)} \stackrel{D}{=} \int_0^u e^{\lambda s'}dZ_{s'}. \quad \square$$

Remark 1.4.5. *In the proof above, Y_0 and both integrals defined on disjoint intervals \int_0^t and \int_t^{t+u} are independent due to the independent increment property of the Levy process Z_t . Hence, we can conclude that Y_t and $\int_0^u e^{\lambda s'}dZ_{s'}$ are independent.*

Let's continue with the stationarity of an OU type process Y_t . The necessary and sufficient condition for the stationarity of a marginal distribution is that, for any $u > 0$, $\varphi_{Y_t}(z) = \varphi_{Y_{t+u}}(z)$. If we denote $\varphi_{Y_t^{(c)}}(z)$ as the characteristic function of a new random variable $Y_t^{(c)} = E[e^{iz \int_0^u e^{-\lambda(u-s)}dZ_s}]$, which represents the integral term in the equation (1.21), along with the fact that $Y_t^{(c)}$ and Y_t are independent (remark 1.4.5), we get the following formula,

$$\varphi_{Y_t}(z) = \varphi_{Y_{t+u}}(z) = \varphi_{Y_t}(e^{-\lambda u}z) \varphi_{Y_t^{(c)}}(z). \quad (1.22)$$

The proposition 1.21 implies,

$$\varphi_{Y_t^{(c)}}(z) = E[e^{iz \int_0^u e^{-\lambda(u-s)} dZ_s}] = \exp\left(\int_0^u \kappa_{Z_1}\left(z e^{-\lambda(u-s)}\right) ds\right), \quad (1.23)$$

where $\kappa_{Z_1}(z)$ as the Levy exponent of the BDLP Z_t , given by $\varphi_{Z_1}(z) = e^{\kappa(z)}$.

Remark 1.4.6. *Since the above equation (1.22) holds for all $u > 0$, it follows it will hold for all $c = e^{-\lambda u}$, $c \in (0, 1)$. Hence, one may conclude that if the stationarity of the marginal distribution for Y_t holds, as in (1.22), then its marginal distribution must be self-decomposable.*

We are still interested in finding the condition under which the OU type process is stationary. So, if we denote $\omega = z e^{-\lambda(u-s)}$, then it follows from (1.22) and (1.23),

$$\varphi_{Y_t}(z) = \varphi_{Y_t}(e^{-\lambda u} z) \exp\left(\lambda^{-1} \int_{z e^{-\lambda u}}^z \kappa_{Z_1}(\omega) \omega^{-1} d\omega\right).$$

Since any characteristic function is continuous at zero and so is $\varphi_{Y_t}(z)$, if we let $u \rightarrow \infty$ it follows that,

$$\varphi_{Y_t}(z) = \lim_{u \rightarrow \infty} \exp\left(\lambda^{-1} \int_{z e^{-\lambda u}}^z \kappa_{Z_1}(\omega) \omega^{-1} d\omega\right) = \exp\left(\lambda^{-1} \int_0^z \kappa_{Z_1}(\omega) \omega^{-1} d\omega\right). \quad (1.24)$$

Remark 1.4.7. *It is important to notice that the equation (1.24), provided above, tells us how to choose the distribution of Z_1 , or its Levy exponent κ_{Z_1} , in order to get a particular marginal distribution of the OU type process Y_t , with characteristic function φ_{Y_t} .*

Let's continue with the stationarity of an OU type process. In order that it holds, the right-hand side of the equation (1.24) needs to hold, i.e. the Levy exponent κ_{Z_1} must

necessarily satisfy the following condition:

$$\int_0^z |\kappa_{Z_1}(\omega)| \omega^{-1} d\omega < \infty, \quad \text{for } z > 0. \quad (1.25)$$

According to Wolfe (1982) the above condition is equivalent to the condition provided below,

$$E(\log^+ |Z_1|) = E(0 \vee \log |Z_1|) < \infty, \quad (1.26)$$

Jurek and Mason (1993), stated in the theorem 3.6.6. that the above condition is equivalent to the condition given in Sato (1999), namely,

$$\int_{|x|>2} \log |x| \nu(dx) < \infty, \quad (1.27)$$

where ν is the Levy measure of the background driving Levy process (BDLP) Z_t .

We have discussed the condition that allows an OU type process to be stationary. A Markov process will be stationary if its transition function has a stationary limiting distribution, as $t \rightarrow \infty$. Since an OU type process is a Markov process, we can explore the self-decomposability of its marginals and its stationarity through exploring if and when its limiting distribution exists, as $t \rightarrow \infty$. According to definition of the OU type process and proposition 1.4.4 the following holds,

$$\begin{aligned} Y_t &= e^{-\lambda t} Y_0 + \int_0^t e^{-\lambda(t-s)} dZ_s \\ &\stackrel{\mathcal{D}}{=} e^{-\lambda t} Y_0 + \int_0^t e^{-\lambda s} dZ_s. \end{aligned}$$

As $t \rightarrow \infty$, if the limiting distribution exists, it is given by,

$$Y_\infty \stackrel{\mathcal{D}}{=} \int_0^\infty e^{-\lambda s} dZ_s.$$

The existence of the the above limiting distribution will be closely related to the existence of a stationary (or invariant) distribution of an OU type process. Provided that a measure ν is a Levy measure of BDLP Z_1 , it can be shown, see [1], that the necessary condition for the existence of the distribution of Y_∞ is,

$$\int_{|x|>1} \log(1 + |x|) \nu(dx) < \infty.$$

Remark 1.4.8. *Provided that the above limiting distribution Y_∞ exists and the OU type process Y_t is stationary, the following are alternative representations of an Ornstein-Uhlenbeck process,*

$$Y_t \stackrel{\mathcal{D}}{=} \int_0^\infty e^{-\lambda s} dZ_s \stackrel{\mathcal{D}}{=} \int_{-\infty}^0 e^{\lambda s'} dZ_{s'} \stackrel{\mathcal{D}}{=} \int_{-\infty}^t e^{-\lambda(t-s)} dZ_s \quad (1.28)$$

Note, provided that the OU type process is stationary, all three representations above can be proved by using the above approaches and appropriate change of variable.

Remark 1.4.9. *From the third equality from the representation (1.28), one can see that the OU type process is a special case of a moving average process $Z_t = \int_{-\infty}^{+\infty} f(s-t) dZ_s$ where $f(s) = \mathbf{1}_{(\infty,0]}(s) e^{\lambda s}$, see [1].*

1.4.3.2 The stationary OU type processes and self-decomposable distributions

There is a one-to one correspondence between the class of self-decomposable distributions and the stationary OU type processes. As it was briefly mentioned in the remark 1.4.6, it is

not difficult to see that if an OU type process is stationary then it is also self-decomposable, see the proposition 1.4.9 below.

Proposition 1.4.9. *Let Y_t be a stationary OU type process on \mathbb{R} , given by the formula, $Y_t = e^{-\lambda t} Y_0 + \int_0^t e^{-\lambda(t-s)} dZ_s$, $\lambda > 0$. Then the distribution Y_0 is self-decomposable, with given parameters $c = e^{-\lambda t}$ and $Y(t)^{(c)} = \int_0^t e^{-\lambda s} dZ_s$.*

Remark 1.4.10. *Note, it can be proved that any self-decomposable measure $\mu = \mu^c * \mu_t^{(c)}$ is an invariant measure, see [1] and Applebaum's lectures [2]. This means that any self-decomposable distribution μ is a potential candidate for marginal distribution of some stationary Markov process Y_t , such that $Y_0 \stackrel{\mathcal{D}}{=} \mu$.*

In fact, it is possible to show the converse result is true, that any self-decomposable distribution μ is a weak limit of a stationary Ornstein-Uhlenbeck type process Y_t such that $Y_0 \stackrel{\mathcal{D}}{=} \mu$. We closely follow Applebaum's book on the Levy processes and stochastic calculus [1] and present the proposition below.

Proposition 1.4.10. *Given any self-decomposable random variable X with distribution μ , there exists a stationary Ornstein-Uhlenbeck type process Y_t , such that,*

$$X = Y_0 = \int_0^\infty e^{-\lambda s} dZ_s, \quad \text{and} \quad \mathcal{L}(Y_0) \stackrel{\mathcal{D}}{=} \mu.$$

Proof. Let us give a sketch of the proof. Intuitively, since any self-decomposable measure can be decomposed for any $c \in (0, 1)$, i.e. $\mu = \mu^c * \mu_t^{(c)}$, take specific $c = e^{-\lambda t}$, to get, for any $t > 0$, $\mu = \mu^{e^{-\lambda t}} * \mathcal{L}(Y(t)^{(c)})$, where $Y(t)^{(c)} = \int_0^t e^{-\lambda s} dZ_s$. From here, one can see that it is possible to construct a stationary OU type process Y_t and its BDLP levy process

⁵Remember that a Markov process X_t is stationary iff $\mu \stackrel{\Delta}{=} \mathcal{L}(X_0)$ is an invariant measure.

Z_t that would satisfy the above equality, but it is not straightforward, see the reference for details (Jurek and Vervaat (1983), [43]). \square

Note, the result in the above proposition can be interpreted as a complete characterization of a self-decomposable distribution, see below.

Proposition 1.4.11. *Let X be a self-decomposable random variable with Levy measure \mathcal{Q} of the form, $\mathcal{Q}(dx) = q(x)dx$, where $q(x) = k(x)/|x|$, for some non-negative real function $k(\cdot)$, increasing on $(-\infty, 0)$, decreasing on $(0, \infty)$, and satisfies, $\int_{-\infty}^{\infty} (1 \wedge x^2) \frac{k(x)}{|x|} dx < \infty$. It follows that, if a random variable X is self-decomposable and if $\int_{|x|>1} \log(1+|x|) \mathcal{Q}(dx) < \infty$ then X can be represented as,*

$$X = \int_0^{\infty} e^{-s} dZ_s, \quad (1.29)$$

where Z_t is a Levy process whose law is uniquely determined by that of X . In fact it is a BDLP corresponding to X , where Z_1 is uniquely determined by its Levy measure $W(dx)$. Furthermore, the Levy measure $W(dx)$ of Z_1 is related to the Levy density $q(x)$ of X by,

$$W^+(x) = xq(x), \text{ for } x > 0, \quad \text{and} \quad W^-(x) = |x|q(x), \text{ for } x < 0, \quad (1.30)$$

In addition, if the Levy density $q(x)$ of X is differentiable then the Levy measure $W(dx)$ of Z_1 has the density w with respect to Lebesgue measure and the densities $w(x)$ and $q(x)$ are related by,

$$w(x) = -q(x) - xq'(x) \quad (1.31)$$

1.4.3.3 Covariance of the stationary OU type process

The nature of the covariance structure of any processes is important from many different aspects. For example in modeling it is essential to know if the data exhibit long or short range dependence and how it evolves with time; statistical estimation and inference techniques are highly reliant on the dependence structure of the underlying process and if the given process satisfy mixing conditions.

The covariance of the OU type processes can be derived from the elementary calculations, see below:

$$Cov(Y_t, Y_{t+u}) = Cov\left(Y_t, e^{-\lambda u} Y_t + \int_0^u e^{-\lambda(u-s)} dZ_s\right) \quad (1.32)$$

$$= e^{-\lambda u} Var(Y_t) + Cov\left(Y_t, \int_0^u e^{-\lambda(u-s)} dZ_s\right) \quad (1.33)$$

$$= e^{-\lambda u} Var(Y_t) \quad (1.34)$$

where Y_t and the integral $\int_0^u e^{-\lambda(u-s)} dZ_s$ are independent, see the proposition (1.4.8). From the above formula for the covariance, one can easily see that autocorrelation function of an OU type process is given by:

$$\rho(u) = e^{-\lambda|u|}. \quad (1.35)$$

From the above formula we can see that parameter λ controls the memory of the OU type process, where autocorrelation is decaying slower as λ decreases.

Remark 1.4.11. *Note, because its auto-correlation function is decreasing exponentially as number of lags is increasing, it follows that the Ornstein-Uhlenbeck process is short range*

dependent (SRD).

1.5 Superpositions of the OU type processes

1.5.1 Existence of supOU processes

Superpositions of the OU type processes, or supOU processes for short, were introduced in the papers of Barndorff-Nielsen [4], Barndorff-Nielsen and Shephard [8] and Barndorff-Nielsen and Leonenko [7], among others. In this section we will give a brief overview of the superposition type models, i.e. their existence, uniqueness, and the main properties including their covariance structure. The focus of this thesis will be mainly on their limiting behavior, see chapter 2. We will start with the definition of finite and infinite superpositions and the assumptions necessary so that they are well defined, see below.

Definition 1.5.1. *Let's assume the following,*

(A) *Let $\{Y^{(k)}(t)\}_{k \geq 1}$ be the sequence of independent stationary OU type processes, i.e. independent processes such that, each $Y^{(k)}(t)$ is a stationary solution of the equation:*

$$dY^{(k)}(t) = -\lambda_k Y^{(k)}(t)dt + dZ^{(k)}(\lambda_k t), \quad t \geq 0, \quad (1.36)$$

in which the Lévy processes $\{Z^{(k)}\}_{k \geq 1}$ are independent, and $\lambda_k > 0$ for all $k \geq 1$.

In addition, we assume that the self-decomposable distribution of $Y^{(k)}$ has finite moments of up to q , $q \geq 2$, it is closed under convolution with respect to at least one distributional parameter δ_k , and that cumulants of order $q \geq 2$ of $Y^{(k)}$ are proportional to parameter δ_k .

(B) Another set of assumption is related to the growth of moments of Y_t , see below,

$$\sum_{k=1}^{\infty} EY^{(k)}(t) < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \text{Var}Y^{(k)}(t) < \infty. \quad (1.37)$$

Let's define either finite for an integer m , or infinite superpositions of the OU type processes as the following,

(i) If the assumption (A) with $q = 2$ holds finite superposition for an integer m is defined as,

$$Y_m(t) = \sum_{k=1}^m Y^{(k)}(t), \quad t \in \mathbb{R} \quad (1.38)$$

(ii) Provided that both assumptions, (A) with $q = 2$ and (B) hold, infinite superposition is,

$$Y_{\infty}(t) = \sum_{k=1}^{\infty} Y^{(k)}(t), \quad t \in \mathbb{R}, \quad (1.39)$$

Note, the construction with infinite superposition is well-defined in the sense of mean-square or almost-sure convergence provided that the assumption (B) holds.

Although assumption (A) may seem restrictive, it is satisfied for many examples with tractable distributions of superpositions. The section 1.5.3 provides a number of examples where both assumptions (A) and (B) are satisfied. These examples include Gamma, inverse Gaussian and other well known distributions. Their superpositions have the marginal distributions that belong to the same class as the marginal distributions of the components of superposition.

1.5.2 The covariance structure of supOU processes

1.5.2.1 Finite supOU processes

In the case of finite superposition, the covariance function of the resulting process is given by the following formula,

$$R_{Y_m}(t) = \text{cov}(Y_m(0), Y_m(t)) = \sum_{k=1}^m \text{Var}(Y^{(k)}(t))e^{-\lambda_k t},$$

Recall the section 1.1, a stochastic process is short-range dependent if its covariance function is integrable, i.e. $\int_{t=-\infty}^{+\infty} |R_{Y_m}(t)| dt < \infty$.

Since we have,

$$\int_{t=-\infty}^{+\infty} |R_{Y_m}(t)| dt = \int_{t=-\infty}^{+\infty} \sum_{k=1}^m \text{Var}(Y^{(k)}(t))e^{-\lambda_k t} dt < \infty,$$

we may conclude that the finite superposition is a short-range dependent process.

1.5.2.2 Infinite supOU processes

In the case of infinite superposition, the covariance function is of the form,

$$R_{Y_\infty}(t) = \text{cov}(Y_\infty(0), Y_\infty(t)) = \sum_{k=1}^{\infty} \text{Var}(Y^{(k)}(t))e^{-\lambda_k t},$$

and under the condition (A) the variance of $Y^{(k)}(t)$ is proportional to δ_k , that is

$$\text{Var}(Y^{(k)}(t)) = \delta_k C_2,$$

where constant C_2 does not depend on k and reflects parameters of the marginal distribution of $Y^{(k)}$. If one chooses specific δ_k ,

$$\delta_k = k^{-(1+2(1-H))}, \quad \frac{1}{2} < H < 1, \quad \lambda_k = \lambda/k$$

for some $\lambda > 0$, then we get that the covariance function of an infinite superposition is,

$$R_{Y_\infty}(t) = C_2 \sum_{k=1}^{\infty} \frac{1}{k^{1+2(1-H)}} e^{-\lambda t/k}. \quad (1.40)$$

The lemma below shows that the correlation function (1.40) is not integrable for the chosen parameters δ_k and λ_k , thus the process obtained via infinite superposition exhibits long-range dependence.

Lemma 1.5.1. *For the infinite superposition of OU type processes that satisfy condition (A) with $q = 2$ and condition (B), the covariance function of $Y_\infty(t)$ given by with $\lambda^{(k)} = \lambda/k$ and $\delta_k = k^{-(1+2(1-H))}$, $\frac{1}{2} < H < 1$, can be written as*

$$R_{Y_\infty}(t) = \frac{L(t)}{t^{2(1-H)}}, \quad t > 0$$

where L is a slowly varying at infinity function.

Proof. The proof of this lemma is essentially the same as the proofs presented for particular cases of superpositions of OU processes in [47], [49]. We provide it here for completeness and for the remark that follows. The remark will be used for proofs later in the paper. Let

$$L(t) = C_2 t^{2(1-H)} \sum_{k=1}^{\infty} \frac{1}{k^{1+2(1-H)}} e^{-\lambda t/k}.$$

Estimate the sum appearing in the expression for L as follows:

$$\int_1^\infty \frac{e^{-\lambda t/u}}{u^{1+2(1-H)}} du \leq \sum_{k=1}^\infty \frac{1}{k^{1+2(1-H)}} e^{-\lambda t/k} \leq \int_1^\infty \frac{e^{-\lambda t/u}}{u^{1+2(1-H)}} du + e^{-\lambda t}.$$

Transform the variables $\lambda t/u = s$ to get

$$\frac{C_2}{\lambda^{2(1-H)}} \int_0^{\lambda t} e^{-s} s^{2(1-H)-1} ds \leq L(t) \leq \frac{C_2}{\lambda^{2(1-H)}} \int_0^{\lambda t} e^{-s} s^{2(1-H)-1} ds + C_2 e^{-\lambda t} t^{2(1-H)}.$$

Since

$$\int_0^{\lambda t} e^{-s} s^{2(1-H)-1} ds \rightarrow \Gamma(2(1-H))$$

as $t \rightarrow \infty$, it follows that $\lim_{t \rightarrow \infty} L(tv)/L(t) = 1$ for any fixed $v > 0$. □

Remark 1.5.1. *From proof of Lemma 1.5.1*

$$\begin{aligned} L([Nt]) &\leq \frac{C_2}{\lambda^{2(1-H)}} \int_0^{\lambda[Nt]} e^{-s} s^{2(1-H)-1} ds + C_2 e^{-\lambda[Nt]} [Nt]^{2(1-H)} \\ &\leq \frac{C_2}{\lambda^{2(1-H)}} \Gamma(2(1-H)) + C_2 e^{-2(1-H)} \left(\frac{2(1-H)}{\lambda} \right)^{2(1-H)} \end{aligned}$$

for all $N \geq 1$ and $t \in [0, 1]$ since the function $x^{2(1-H)} e^{-x}$ is bounded (attains its maximum at $x = 2(1-H)$). Also from the proof of Lemma 1.5.1

$$L(N) \geq \frac{C_2}{\lambda^{2(1-H)}} \int_0^{\lambda N} e^{-s} s^{2(1-H)-1} ds \geq \frac{C_2}{\lambda^{2(1-H)}} \int_0^\lambda e^{-s} s^{2(1-H)-1} ds$$

for all $N \geq 1$. Also note that $L(0) = 0$. Therefore the ratio $L([Nt])/L(N)$ is bounded uniformly in $N \geq 1$ and $t \geq 0$.

1.5.3 Examples of OU type superpositions

The examples in this section have been discussed in [4, 47]. We briefly present them to illustrate that conditions (A) and (B) are satisfied for a number of OU type processes.

1.5.3.1 Gamma supOU process

Gamma distribution $\Gamma(\alpha, \beta)$, $\alpha > 0, \beta > 0$ is defined by the following probability density function (pdf),

$$\pi(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbf{1}_{[0, \infty)}(x).$$

It is self-decomposable and hence there is a stationary OU type process $\{Y(t), t \geq 0\}$ with gamma marginal distribution, often referred to in literature as stationary Gamma OU type process. Its covariance function is of the form,

$$R_Y(t) = \frac{\alpha}{\beta^2} \exp\{-\lambda|t|\}.$$

The cumulant generating function is given by,

$$\kappa_Y(z) = \log E e^{izY(t)} = -\alpha \log \left(1 - \frac{iz}{\beta} \right) = \sum_{m=1}^{\infty} \frac{\alpha (iz)^m}{m \beta^m}, \quad z < \beta, \quad (1.41)$$

so that

$$EY(t) = \frac{\alpha}{\beta}, \quad \text{Var}Y(t) = \frac{\alpha}{\beta^2}.$$

Let's look at the superposition of the stationary Gamma OU type processes. Suppose $\{Y^{(k)}(t), t \geq 0\}$, $k \geq 1$ are independent stationary OU type Gamma processes with marginals

$\Gamma(\alpha_k, \beta)$, $k \geq 1$ where

$$\alpha_k = \alpha k^{-(1+2(1-H))}, \quad \frac{1}{2} < H < 1.$$

It follows from our setup and (1.41) that the processes $\{Y^{(k)}(t), t \geq 0\}$, $k \geq 1$ satisfy condition (A) with the choice of $\delta_k = \alpha_k$. In addition, if $\sum_{k=1}^{\infty} \alpha_k < \infty$, the condition (B) is satisfied as well, hence infinite superposition of the stationary Gamma OU type processes is well defined. Thus, the supOU process given by,

$$Y_{\infty}(t) = \sum_{k=1}^{\infty} Y^{(k)}(t), \quad t \geq 0,$$

has a marginal $\Gamma(\sum_{k=1}^{\infty} \alpha_k, \beta)$ distribution and the covariance function is of the following form,

$$R_{Y_{\infty}}(t) = \frac{1}{\beta^2} \sum_{k=1}^{\infty} \alpha_k e^{-\lambda^{(k)} t}.$$

Remark 1.5.2. *Note, with above choice of δ_k and a particular choice of $\lambda^{(k)}$ as λ/k we can get a long range dependent stationary Gamma OU type process.*

1.5.3.2 Inverse Gaussian supOU process

Inverse Gaussian distribution $IG(\delta, \gamma)$, $\delta > 0$, $\gamma > 0$ given with pdf

$$\pi(x) = \frac{1}{\sqrt{2\pi}} \frac{\delta e^{\delta\gamma}}{x^{3/2}} \exp\left\{-\frac{1}{2}\left(\frac{\delta^2}{x} + \gamma^2 x\right)\right\} \mathbf{1}_{[0,\infty)}(x), \quad (1.42)$$

is self-decomposable, so there exists a stationary OU type process $\{Y(t), t \geq 0\}$ with $IG(\delta, \gamma)$ marginal distribution and correlation function $\exp\{-\lambda|t|\}$. The cumulant generating func-

tion is

$$\kappa_Y(z) = \log Ee^{izY(t)} = \delta \left(\gamma - \sqrt{\gamma^2 - 2z} \right) = \sum_{m=1}^{\infty} \frac{\delta(2m)!(iz)^m}{(2m-1)(m!)^2 2^m \gamma^{2m-1}},$$

and

$$EY(t) = \frac{\delta}{\gamma}, \quad \text{Var}Y(t) = \frac{\delta}{\gamma^3}.$$

It follows that independent stationary OU type processes $\{Y^{(k)}(t), t \geq 0\}$, $k \geq 1$ with marginals $IG(\delta_k, \gamma)$, $k \geq 1$ where

$$\delta_k = \delta k^{-(1+2(1-H))}, \quad \frac{1}{2} < H < 1,$$

satisfy conditions (A) and (B), and we obtain the inverse Gaussian supOU process as,

$$Y_{\infty}(t) = \sum_{k=1}^{\infty} Y^{(k)}(t), \quad t \geq 0,$$

with marginal $IG(\sum_{k=1}^{\infty} \delta_k, \gamma)$ distribution. The covariance function is

$$R_{Y_{\infty}}(t) = \frac{1}{\gamma^3} \sum_{k=1}^m \delta_k \exp\{-\lambda_k |t|\}.$$

1.5.3.3 Variance gamma supOU process

The Variance Gamma distribution $VG(\kappa, \alpha, \beta, \mu)$ with the parameters

$$\kappa > 0, \quad \alpha > |\beta| > 0, \quad \mu \in \mathbb{R}, \quad \gamma^2 = \alpha^2 - \beta^2,$$

is defined by the following pdf

$$\pi(x) = \frac{\gamma^{2\kappa}}{\sqrt{\pi}\Gamma(\kappa)(2\alpha)^{\kappa-1/2}} |x - \mu|^{\kappa-1/2} K_{\kappa-1/2}(\alpha|x - \mu|) e^{\beta(x-\mu)}, \quad x \in \mathbb{R},$$

where $K_\lambda(z)$ denotes modified Bessel function of the second kind.

Since the Variance Gamma distribution is self-decomposable, the corresponding stationary OU type process, denoted as $\{Y(t), t \geq 0\}$, exists and its cumulant generating function is given by the following formula,

$$\kappa_Y(z) = \log Ee^{izY(t)} = i\mu z + 2\kappa \log \left(\frac{\gamma}{\alpha^2 - (\beta + iz)^2} \right), \quad |\beta + z| < \alpha.$$

It follows that $VG(\kappa, \alpha, \beta, \mu)$ distribution is closed under convolution with respect to parameters κ and μ . Independent stationary OU type processes $\{Y^{(k)}(t), t \geq 0\}$, $k \geq 1$ with marginals $VG(\kappa_k, \alpha, \beta, \mu_k)$, $k \geq 1$ where $\sum_{k=1}^{\infty} \mu_k$ converges and a specific choice for the parameter κ_k , i.e.

$$\kappa_k = \kappa k^{-(1+2(1-H))}, \quad \frac{1}{2} < H < 1,$$

satisfy conditions (A) and (B), and we obtain the Variance Gamma supOU process as,

$$Y_\infty(t) = \sum_{k=1}^{\infty} Y^{(k)}(t), \quad t \geq 0,$$

with marginal $VG(\sum_{k=1}^{\infty} \kappa_k, \alpha, \beta, \sum_{k=1}^{\infty} \mu_k)$ distribution. The covariance function is of the form,

$$R_{Y_\infty}(t) = \frac{2}{\gamma^2} \left(1 + 2 \left(\frac{\beta}{\gamma} \right)^2 \right) \sum_{k=1}^{\infty} \kappa_k \exp\{-\lambda_k |t|\}.$$

1.5.3.4 Normal inverse Gaussian supOU process

The normal inverse Gaussian distribution $NIG(\alpha, \beta, \delta, \mu)$ with following parameters,

$$\alpha \geq |\beta| \geq 0, \delta > 0, \mu \in \mathbb{R},$$

is defined by the pdf

$$\pi(x) = \frac{\alpha\delta}{\pi} e^{\delta\gamma} \frac{K_1\left(\alpha\sqrt{\delta^2 + (x - \mu)^2}\right)}{\sqrt{\delta^2 + (x - \mu)^2}} e^{\beta(x - \mu)}, \quad x \in \mathbb{R},$$

where $K_\lambda(z)$ denotes modified Bessel function of the second kind. Since it is self-decomposable, the corresponding stationary OU type process $\{Y(t), t \geq 0\}$ exists and its cumulant generating function is given by,

$$\kappa_Y(z) = \log Ee^{izY(t)} = i\mu z + \delta \left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + iz)^2} \right), \quad |\beta + z| < \alpha.$$

It follows that $NIG(\alpha, \beta, \delta, \mu)$ distribution is closed under convolution with respect to parameters δ and μ . Independent stationary OU type processes $\{Y^{(k)}(t), t \geq 0\}$, $k \geq 1$ with marginals $NIG(\alpha, \beta, \delta_k, \mu_k)$, $k \geq 1$ with convergent $\sum_{k=1}^{\infty} \mu_k$, and δ_k given as,

$$\delta_k = \delta k^{-(1+2(1-H))}, \quad \frac{1}{2} < H < 1,$$

satisfy conditions (A) and (B), and we obtain the normal inverse Gaussian supOU process as,

$$Y_\infty(t) = \sum_{k=1}^{\infty} Y^{(k)}(t), \quad t \geq 0,$$

with marginal $NIG(\alpha, \beta, \sum_{k=1}^{\infty} \delta_k, \sum_{k=1}^{\infty} \mu_k)$ distribution. The covariance function is given by,

$$R_{Y_{\infty}}(t) = \frac{\alpha^2}{(\alpha^2 - \beta^2)^{3/2}} \sum_{k=1}^{\infty} \delta_k \exp\{-\lambda_k |t|\}.$$

1.5.3.5 Tempered stable supOU process

The positive tempered stable distribution $TS(\kappa, \delta, \gamma)$, $\kappa \in (0, 1)$, $\delta > 0$, $\gamma > 0$ is self-decomposable, hence there exists a stationary OU type process $\{Y(t), t \geq 0\}$ with $TS(\kappa, \delta, \gamma)$ as its marginal distribution. The distribution of $Y(t)$ can be specified through its cumulant generating function given by,

$$\kappa_Y(z) = \log E^{izY(t)} = \delta\gamma - \delta \left(\gamma^{1/\kappa} - 2iz \right)^{\kappa}, \quad 0 < z < \frac{\gamma^{1/\kappa}}{2}.$$

Thus, the $TS(\kappa, \delta, \gamma)$ distribution is closed under convolution with respect to parameter δ . Independent stationary OU type processes $\{Y^{(k)}(t), t \geq 0\}$, $k \geq 1$ with marginals $TS(\kappa, \delta_k, \gamma)$, $k \geq 1$ and where δ_k is chosen as,

$$\delta_k = \delta k^{-(1+2(1-H))}, \quad \frac{1}{2} < H < 1,$$

satisfy conditions (A) and (B), and we obtain the tempered stable supOU process as,

$$Y_{\infty}(t) = \sum_{k=1}^{\infty} Y^{(k)}(t), \quad t \geq 0,$$

with marginal $TS(\kappa, \sum_{k=1}^{\infty} \delta_k, \gamma)$ distribution. The covariance function is given by,

$$R_{Y_{\infty}}(t) = 4\kappa(1 - \kappa) \gamma^{\frac{\kappa-2}{\kappa}} \sum_{k=1}^{\infty} \delta_k \exp\{-\lambda_k |t|\}.$$

Remark 1.5.3. *More examples of supOU type processes satisfying condition (A) can be derived from other distributions, like for example: normal tempered stable, Euler's gamma distribution and z -distribution.*

Chapter 2

Intermittency of supOU processes

The phenomenon of intermittency has been widely discussed in physics literature (see for example [11, 28, 57, 67, 69] and [27, Chapter 8]). The term is used to describe models exhibiting high degree of variability and enormous fluctuations which escape from the scope of the usual limit theory. Terms multifractality, separation of scales, dynamo effect are often used interchangeably with intermittency. For a formal definition of intermittency appearing in the theory of stochastic partial differential equations (SPDE) we follow [14] and [44, Chapter 7]. There, a nonnegative random field $\{\psi_t(x), t \geq 0, x \in \mathbb{R}\}$ stationary in parameter x is said to be intermittent if the function $k \mapsto \gamma(k)/k$ is strictly increasing on $[2, \infty)$ where $\gamma(k)$ is the k -th moment Lyapunov exponent of ψ defined by

$$\gamma(k) = \lim_{t \rightarrow \infty} \frac{\log E (\psi_t(x))^k}{t}, \quad (2.1)$$

assuming the limit exists and is finite. This approach to intermittency is tailored for the analysis of SPDEs and characterizes fields with progressive growth of moments.

To compare intermittency to a slower growth of moments, consider the sum $\phi_n = \sum_{i=1}^n \xi_i$, where ξ_i are positive independent identically distributed (iid) random variables with finite moments. The k -th moment of ϕ_n grows as $n^k (E\xi_1)^k$, therefore

$$\gamma(k) = \lim_{n \rightarrow \infty} \frac{k \log n + k \log E\xi_1}{n} = 0 \quad (2.2)$$

for all $k \geq 1$. With the appropriate centering and norming, the classical central limit theorem holds.

In contrast, for a sequence of products of positive random variables $\psi_n = \prod_{i=1}^n \xi_i$

$$\gamma(k) = \lim_{n \rightarrow \infty} \frac{\log E\psi_n^k}{n} = \log E\xi_1^k.$$

If ξ_i are not constant a.s., then from Jensen's inequality it follows that for $l > k$

$$E\xi_1^k < \left(E\xi_1^l\right)^{\frac{k}{l}},$$

showing that $\gamma(k)/k$ is strictly increasing. The wild growth of moments of ψ_n provides the main heuristic argument that intermittency implies unusual limiting behavior. A formal argument showing that under some assumptions intermittency implies large peaks in the space coordinate of the random field can be found in [44], some ideas of which will be used later in this work.

By far the most investigated model exhibiting intermittent behavior is the parabolic Anderson model (see [29–32]). In this work we consider models provided by the partial sums of discrete superpositions of Lévy driven Ornstein-Uhlenbeck (OU) type processes. While models based on Lévy flights describe the position of particle, models given by OU dynamics describe the velocity of particle trapped in a field generated by quadratic potential ([20]). Applications of Lévy-driven OU type processes include financial econometrics [8, 48, 50], fluid dynamics [66], plasma physics [15] and biology [59]. The stochastic model discussed in this thesis provides another example of intermittency model based on the velocity (see [27, Section 8.5]). First, we modify the preceding definition of intermittency to tailor it to the analysis

of sequences of partial sum processes. In the case of finite superpositions we show that the central limit theorem holds. In the case of infinite long range dependent superpositions, we show that the growth of cumulants is such that the partial sum process is intermittent. The section 1.5.3 contains examples that fit our assumptions which cover, to our knowledge, all the examples with tractable distributions of superpositions.

2.1 Intermittency

For a process $\{Y(t), t \geq 0\}$, denote

$$\bar{q} = \sup\{q > 0 : E|Y(t)|^q < \infty \forall t\}.$$

Our definition of intermittency is based on the version of Lyapunov exponent that replaces t in the denominator of (2.1) with $\log t$. For a stochastic process $\{Y(t), t \geq 0\}$, define the *scaling function* at point $q \in [0, \bar{q})$ as

$$\tau(q) = \lim_{t \rightarrow \infty} \frac{\log E|Y(t)|^q}{\log t}, \quad (2.3)$$

assuming the limit exists and is finite for every $q \in [0, \bar{q})$. Objects similar to the scaling function (2.3) appear in the theory of multifractal processes (see e.g. [33]), however, there are some important differences [45]. The following proposition gives some properties of τ .

Proposition 2.1.1. *The scaling function τ defined by (2.3) has the following properties:*

(i) τ is non-decreasing and so is $q \mapsto \tau(q)/q$;

(ii) τ is convex;

(iii) if for some $0 < p < r < \bar{q}$, $\tau(p)/p < \tau(r)/r$, then there is a $q \in (p, r)$ such that

$$\tau(p)/p < \tau(q)/q < \tau(r)/r.$$

Proof. (i) For $0 \leq q_1 < q_2 < \bar{q}$ Jensen's inequality implies

$$E|Y(t)|^{q_1} = E(|Y(t)|^{q_2})^{\frac{q_1}{q_2}} \leq (E|Y(t)|^{q_2})^{\frac{q_1}{q_2}}$$

and thus

$$\tau(q_1) \leq \frac{q_1}{q_2} \tau(q_2)$$

proving part (i).

(ii) Take $0 \leq q_1 < q_2 < \bar{q}$ and $w_1, w_2 \geq 0$ such that $w_1 + w_2 = 1$. It follows from Hölder's inequality that

$$E|Y(t)|^{w_1 q_1 + w_2 q_2} \leq (E|Y(t)|^{q_1})^{w_1} (E|Y(t)|^{q_2})^{w_2}.$$

Taking logarithms, dividing by $\log t$ for $t > 1$ and letting $t \rightarrow \infty$ we have

$$\tau(w_1 q_1 + w_2 q_2) \leq w_1 \tau(q_1) + w_2 \tau(q_2).$$

(iii) This is clear since $q \mapsto \tau(q)/q$ is continuous by (ii). □

We now define intermittency for a stochastic process and for a sequence of random variables by using the corresponding partial sum process.

Definition 2.1.1. A stochastic process $\{Y(t), t \geq 0\}$ is intermittent if there exist $p, r \in (0, \bar{q})$ such that

$$\frac{\tau(p)}{p} < \frac{\tau(r)}{r}.$$

Later in the thesis, we will investigate intermittency of a stationary sequence of random variables $\{Y_i, i \in \mathbb{N}\}$ with finite mean. In this sense, intermittency will be considered as intermittency of the centered partial sum process

$$S(t) = \sum_{i=1}^{\lfloor t \rfloor} Y_i - \sum_{i=1}^{\lfloor t \rfloor} EY_i, \quad t \geq 0.$$

Proposition 2.1.1(i) shows that the function $q \mapsto \tau(q)/q$ is always non-decreasing. What makes the process intermittent is the existence of points of strict increase. In the next two sections, we connect this property to the limiting behavior of cumulants of partial sums of superpositions of Ornstein-Uhlenbeck type processes. We show that while the partial sums of finite superpositions obey the central limit theorem, partial sums of infinite long-range dependent superpositions provide examples of intermittent processes.

Remark 2.1.1. *To conclude this section, we note that intermittent sequences exhibit erratic behavior which may preclude the usual central limit theorem type of results. For example, take the case of positive, i.i.d random variables, see (2.2), for which we already showed are not intermittent. It is well known that they do obey central limit theorem type of results if properly normalized. Now, on the contrary, for an intermittent sequence such normalization does not exist, which we will see in the next section for the case of partial sums of infinite supOU processes, i.e. we will see that the partial sums of finite superpositions of Ornstein-Uhlenbeck type processes obey the central limit theorem, while partial sums of infinite long-range dependent superpositions provide examples of intermittent processes and most likely do not obey CLT type of results.*

2.2 Limiting properties of partial sums of finite supOU processes

We begin with the limit distribution of the partial sum process for the finite superposition. Hence, for $t > 0$, we define a partial sum processes as following,

$$S_m(t) = \sum_{i=1}^{\lfloor t \rfloor} Y_m(i), \quad (2.4)$$

where the supOU process $Y_m(i)$, at time i , is a given by, $Y_m(i) = \sum_{k=1}^m Y^{(k)}(i)$, $i \in \mathbb{Z}$.

In the following paragraph we provide a brief proof of the asymptotic normality of the partial sum process, which is easy to prove using the strong mixing property (α -mixing) of the OU type processes. Strong mixing property was established in the work of Jongbloed, Van Der Meulen, and Van Der Vaart, see [42]. They showed that the integrability condition on the Levy measure of BDLP, $\int_{|x|>2} \log|x| \nu(dx) < \infty$, given in the (1.27), is enough to show the β -mixing which, in turn implies the strong mixing property of the stationary OU type process. The β -mixing of general, multi-dimensional OU type process is treated in Masuda, see [55]. He made a stronger assumption than (1.27), related to existence of the absolute moments of order α of the marginal distribution. More specific, Masuda assumed that the process is strictly stationary and that the following holds, $\int |x|^\alpha \pi(dx)^1 < \infty$, for some $\alpha > 0$ which in turn yields a stronger result than only the β -mixing itself. It follows that β -mixing coefficient converges to zero with an exponential rate, i.e. for some $a > 0$, $\beta_Y(t) = O(e^{-at})$, as $t \rightarrow \infty$. From this stronger assumption the asymptotic normality of the partial sum process follows directly, see below. Note that previously, asymptotic normality

¹Note that π is unique invariant probability measure.

of partial sums was reported for inverse Gaussian and gamma finite superpositions [48]. The result below is a straightforward generalization to a more general class of processes.

Theorem 2.2.1. *For a fixed integer $m > 0$, let Y_m be defined by (1.38), where OU type processes $\{Y^{(k)}, k = 1, \dots, m\}$ defined by (1.36) are independent and $E|Y^{(k)}|^{2+d} < \infty$ for some $d > 0$ and all $k = 1, \dots, m$. Then the partial sums process (2.4), centered and appropriately normed, converges to the Brownian motion*

$$\frac{1}{c_m N^{1/2}} \left(S_m([Nt]) - ES_m([Nt]) \right) \rightarrow B(t), \quad t \in [0, 1],$$

as $N \rightarrow \infty$ in the sense of weak convergence in Skorokhod space $D[0, 1]$. The norming constant c_m is given by

$$c_m = \left(\sum_{k=1}^m \text{Var} \left(Y^{(k)}(0) \right) \frac{1 - e^{-\lambda^{(k)}}}{1 + e^{-\lambda^{(k)}}} \right)^{1/2}.$$

Proof. Since each OU type process in the superposition has a finite second moment, β -mixing (absolute regularity, see the Appendix A) for each OU process holds with the exponential rate. Namely, there exists $a_k > 0$ such that the mixing coefficient $\beta_{Y^{(k)}}(t) = O(e^{-a_k t})$ [55, Theorem 4.3]. Denote by $\alpha^{(k)}(t)$ the strong mixing coefficient of the process $Y^{(k)}$, then since β -mixing implies α -mixing, [13], we get $2\alpha^{(k)}(t) \leq \beta^{(k)}(t) \leq D_k e^{-a_k t}$ for a constant D_k , for each $k = 1, \dots, m$. A finite sum of α -mixing processes with exponentially decaying mixing coefficients is also α -mixing, see again [13] with exponentially decaying mixing coefficient, therefore weak convergence of partial sums of the process Y_m in $D[0, 1]$ follows from Davydov, [18, Theorem 4.2], see the Appendix A). □

2.3 Limiting properties of partial sums of infinite supOU processes

Before we provide the results from this section, we will briefly present an important theorem related to the regular variation, previously defined in the Preliminaries section (1.1). There, we defined the notion of regularly varying (RV) function or a function that asymptotically behaves like a power function. The Karamata's theorem, stated in detail below, simply says that the ρ -regularly varying function integrates the way one would expect a common power function to integrate. In other words, for the purpose of integration, a ρ -varying function behaves roughly like x^ρ . A brief simple explanation is that any RV function $U(x) \in RV_\rho$ can be represented as, $U(x) = x^\rho L(x)$, where $L(x)$ is a slowly varying function. Since a slowly varying function asymptotically behaves like a constant, one would simply guess that it can be passed through the integral. Hence, one is left with integrating only a power function. See the integral expression below,

$$\int_0^x U(t) dt = \int_0^x t^\rho L(t) dt \simeq L(x) \int_0^x t^\rho dt = L(x) x^{\rho+1}/(\rho+1)$$

The Karamata's theorem is stated below.

Theorem 2.3.1 (Karamata's theorem).

(i) Suppose $\rho \geq -1$ and the $U \in RV_\rho$. Then $\int_0^x U(t)dt \in RV_{\rho+1}$ and

$$\lim_{x \rightarrow \infty} \frac{xU(x)}{\int_0^x U(t)dt} = \rho + 1$$

(ii) Suppose $\rho < -1$ and then $U \in RV_\rho$ implies that $\int_x^\infty U(t)dt$ is finite and $\int_x^\infty U(t)dt \in$

$RV_{\rho+1}$ and

$$\lim_{x \rightarrow \infty} \frac{xU(x)}{\int_x^\infty U(t)dt} = \rho - 1$$

Now, we can continue with the limiting properties of the infinite superposition process.

For $t > 0$, let's consider a partial sum process of the infinite supOU process as follows,

$$S_\infty(t) = \sum_{i=1}^{[t]} Y_\infty(i). \quad (2.5)$$

We now proceed with the limit distribution of the partial sum process for the infinite superposition (1.39). The variance of this process has been computed in [47, Equation (5.3)], however the result on the asymptotic normality of the partial sum process [47, Theorem 3] was not correct. Also incorrect was statement (30) of [6, Theorem 5]. Here we provide the derivation of the variance and correct the result on the limit distribution.

Lemma 2.3.2. *For the infinite superposition (1.39) of OU type processes that satisfy condition (A) with $q = 2$ and condition (B), set $\lambda^{(k)} = \lambda/k$ and $\delta_k = k^{-(1+2(1-H))}$, $\frac{1}{2} < H < 1$.*

Then

$$\text{Var}(S_\infty([Nt])) = \frac{L(N)[Nt]^{2H}}{H(2H-1)} (1 + o(1)), \text{ as } N \rightarrow \infty, \quad (2.6)$$

where L is a slowly varying at infinity function.

Proof. Using the expression for the covariance function of the infinite superposition from Lemma 1.5.1, write

$$\begin{aligned}
\text{Var}(S_\infty([Nt])) &= \sum_{m,n=1}^{[Nt]} \text{cov}(Y_\infty(m), Y_\infty(n)) \\
&= [Nt]\text{Var}(Y_\infty(n)) + 2 \sum_{m,n=1, m>n}^{[Nt]} \frac{L(m-n)}{(m-n)^{2(1-H)}} \\
&= C_2[Nt]\zeta(1+2(1-H)) + 2 \sum_{j=1}^{[Nt]-1} ([Nt]-j) \frac{L(j)}{j^{2(1-H)}},
\end{aligned}$$

where $\zeta(\cdot)$ is Riemann's zeta function. The sum appearing in the expression for the variance

$$\sum_{j=1}^{[Nt]-1} ([Nt]-j) \frac{L(j)}{j^{2(1-H)}}$$

is a Riemann sum for the following integral:

$$\int_0^1 ([Nt] - [Nt]u) \frac{L([Nt]u)}{([Nt]u)^{2(1-H)}} [Nt] du = [Nt]^{2H} \int_0^1 (1-u) u^{2H-2} L([Nt]u) du.$$

Consider the integral

$$\int_0^1 u^{2H-2} L([Nt]u) du = \frac{1}{[Nt]^{2H-1}} \int_0^{[Nt]} v^{2H-2} L(v) dv,$$

and apply Karamata's theorem [58, Theorem 2.1] to get

$$\int_0^{[Nt]} v^{2H-2} L(v) dv = \frac{L(N)[Nt]^{2H-1}}{2H-1} (1 + o(1))$$

as $N \rightarrow \infty$. Similarly,

$$\int_0^1 u^{2H-1} L([Nt]u) du = \frac{L(N)}{2H} (1 + o(1))$$

as $N \rightarrow \infty$, and therefore

$$\int_0^1 ([Nt] - [Nt]u) \frac{L([Nt]u)}{([Nt]u)^{2(1-H)}} [Nt] du = \frac{L(N)[Nt]^{2H}}{2H(2H-1)} (1 + o(1)).$$

Since $L([Nt]u)/L(N) \rightarrow 1$ as $N \rightarrow \infty$ uniformly on bounded intervals [64, Theorem 1.1],

the uniform convergence implies that

$$\int_0^1 (1-u) u^{2H-2} \frac{L([Nt]u)}{L(N)} du = \int_0^1 (1-u) u^{2H-2} du (1 + o(1)) = \frac{1}{2H(2H-1)} (1 + o(1)).$$

For $\frac{1}{2} < H < 1$, the second term in the expression for the variance of $S_\infty([Nt])$ dominates the first, and (2.6) follows. \square

In order to characterize the limit distribution of the partial sums of the infinite superpositions, we use the representation of the discretized stationary OU process as a first order autoregressive sequence

$$Y^{(k)}(i) = e^{-\lambda_k} Y^{(k)}(i-1) + W^{(k)}(i), \quad (2.7)$$

where $W^{(k)}(i)$ is independent of $Y^{(k)}(j)$ for all $j < i$. Denote by $\rho_k = e^{-\lambda_k}$. The following lemma provides a useful representation of the partial sums process for the infinite superposition.

Lemma 2.3.3. *The centered partial sums of the superposition of processes that satisfy condi-*

tion (A) with $q = 2$ and condition (B) with $\lambda^{(k)} = \lambda/k$ and $\delta_k = k^{-(1+2(1-H))}$, $\frac{1}{2} < H < 1$, can be written as

$$S_\infty([Nt]) - ES_\infty([Nt]) = \sum_{k=1}^{\infty} b_{[Nt]}^{(k)} \tau^{(k)}(0) + \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} a_{[Nt]-j}^{(k)} V^{(k)}(j), \quad (2.8)$$

where $\tau^{(k)}(0)$, $V^{(k)}(j)$ are independent for different k , for each k : $V^{(k)}(j)$ are independent for different j and also independent of $\tau^{(k)}(0)$. The series in (2.8) converges almost surely, and the coefficients are given by

$$b_{[Nt]}^{(k)} = \sum_{i=1}^{[Nt]} \rho_k^i = \frac{\rho_k(1 - \rho_k^{[Nt]})}{1 - \rho_k}, \quad (2.9)$$

and

$$a_{[Nt]-j}^{(k)} = \sum_{i=0}^{[Nt]-j} \rho_k^i = \frac{1 - \rho_k^{[Nt]-j+1}}{1 - \rho_k}. \quad (2.10)$$

Proof. Center the variables

$$\tau^{(k)}(i) = Y^{(k)}(i) - EY^{(k)}(i), \quad V^{(k)}(i) = W^{(k)}(i) - EW^{(k)}(i)$$

to arrive at centered version of (2.7)

$$\tau^{(k)}(i) = \rho_k \tau^{(k)}(i-1) + V^{(k)}(i). \quad (2.11)$$

Iterate (2.11) to obtain

$$\tau^{(k)}(i) = \rho_k^i \tau^{(k)}(0) + \sum_{j=1}^i \rho_k^{i-j} V^{(k)}(j).$$

Now the partial sum of $\tau^{(k)}$ can be written

$$\begin{aligned}
\sum_{i=1}^{[Nt]} \tau^{(k)}(i) &= \tau^{(k)}(0) \sum_{i=1}^{[Nt]} \rho_k^i + \sum_{i=1}^{[Nt]} \sum_{j=1}^i \rho_k^{i-j} V^{(k)}(j) \\
&= \tau^{(k)}(0) \sum_{i=1}^{[Nt]} \rho_k^i + \sum_{j=1}^{[Nt]} V^{(k)}(j) \sum_{i=j}^{[Nt]} \rho_k^{i-j} \\
&= \tau^{(k)}(0) \sum_{i=1}^{[Nt]} \rho_k^i + \sum_{j=1}^{[Nt]} V^{(k)}(j) \sum_{m=0}^{[Nt]-j} \rho_k^m \\
&= b_{[Nt]}^{(k)} \tau^{(k)}(0) + \sum_{j=1}^{[Nt]} a_{[Nt]-j}^{(k)} V^{(k)}(j),
\end{aligned}$$

where the coefficients are given by (2.9) and (2.10). Note that for different j , $V^{(k)}(j)$ are independent due to (2.7), and they are also independent of $\tau^{(k)}(0)$. For different k , independence follows from the independence of OU type processes $Y^{(k)}$. Summing with respect to k completes the derivation of (2.8), provided that the series in (2.8) converge almost surely. Series convergence holds because the terms have zero mean, and the series of second moments converge. Indeed, for the first series in (2.8)

$$\begin{aligned}
\sum_{k=1}^{\infty} (b_{[Nt]}^{(k)})^2 E(\tau^{(k)}(0))^2 &= C_2 \sum_{k=1}^{\infty} (b_{[Nt]}^{(k)})^2 \delta_k \\
&= C_2 \sum_{k=1}^{\infty} \sum_{j,i=1}^{[Nt]} \rho_k^{i+j} \delta_k = \sum_{j,i=1}^{[Nt]} \frac{L(i+j)}{(i+j)^{2(1-H)}}.
\end{aligned}$$

The sum can be viewed as a Riemann sum for the double integral:

$$\begin{aligned}
\frac{1}{[Nt]^2} \sum_{j,i=1}^{[Nt]} \frac{L(i+j)}{(i+j)^{2(1-H)}} &= \int_0^1 \int_0^1 \frac{L([Nt](x+y))}{([Nt](x+y))^{2(1-H)}} dx dy (1 + o(1)) \\
&= \frac{L(N)}{[Nt]^{2(1-H)}} \int_0^1 \int_0^1 \frac{dx dy}{(x+y)^{2(1-H)}} (1 + (o(1)))
\end{aligned}$$

as $N \rightarrow \infty$. The last equality is justified using Karamata's theorem as in Lemma 2.3.2, or by considering

$$L(N) \int_{x=\epsilon}^1 \int_{y=0}^1 \frac{L([Nt](x+y))}{L(N)} \frac{dx dy}{(x+y)^{2(1-H)}}$$

and using remark 1.5.1 and the dominated convergence theorem. More precisely, the fact that $L([Nt](x+y))/L(N)$ converges to 1 uniformly in $x+y \in [\epsilon, 1]$ by [64, Theorem 1.1], and letting $\epsilon \rightarrow 0$. Therefore the variance of the first series in (2.8) is of the order $L(N)N^{2H}$.

For the second term in (2.8), the series of the variances is

$$\sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \left(a_{[Nt]-j}^{(k)} \right)^2 E \left(V^{(k)}(j) \right)^2 = \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \left(\sum_{i=0}^{[Nt]-j} \rho_k^i \right)^2 (1 - \rho_k^2) C_2 \delta_k,$$

since $Var(V^{(k)}(j)) = (1 - \rho_k^2) Var(\tau^{(k)})$. The series of variances becomes

$$\begin{aligned} & \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \sum_{i_1, i_2=0}^{[Nt]-j} \rho_k^{i_1+i_2} (1 - \rho_k^2) C_2 \delta_k \\ &= \sum_{j=1}^{[Nt]} \sum_{i_1, i_2=0}^{[Nt]-j} \left(\frac{L(i_1+i_2)}{(i_1+i_2)^{2(1-H)}} - \frac{L(i_1+i_2+2)}{(i_1+i_2+2)^{2(1-H)}} \right) = \frac{[Nt]^3}{\lambda^{2(1-H)}} \times \\ & \int_{x=0}^1 \int_{y=0}^{1-x} \int_{z=0}^{1-x} \left(\frac{L([Nt](y+z))}{([Nt](y+z))^{2(1-H)}} - \frac{L([Nt](y+z)+2)}{([Nt](y+z)+2)^{2(1-H)}} \right) dx dy dz \\ & \times (1 + o(1)). \end{aligned}$$

Arguing in the same way as for the variance of first term in (2.8), we have

$$\begin{aligned}
& \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \sum_{i_1, i_2=0}^{[Nt]-j} \rho_k^{i_1+i_2} (1 - \rho_k^2) C_2 \delta_k = [Nt]^{2H+1} L(N) \\
& \times \int_{x=0}^1 \int_{y=0}^{1-x} \int_{z=0}^{1-x} \left(\frac{1}{(y+z)^{2(1-H)}} - \frac{1}{((y+z) + 2/[Nt])^{2(1-H)}} \right) dx dy dz \\
& = \frac{[Nt]^{2H} L(N)}{(2H-1)} \times \int_{x=0}^1 \int_{y=0}^{1-x} [Nt] \left((y + 2/[Nt])^{2H-1} - y^{2H-1} \right. \\
& \quad \left. - \left((y + 1 - x + 2/[Nt])^{2H-1} - (y + 1 - x)^{2H-1} \right) \right) dx dy (1 + o(1)).
\end{aligned}$$

It is not hard to see that as $[Nt] \rightarrow \infty$ the integrand converges to

$$2(2H-1) \left(y^{2H-2} - (y+1-x)^{2H-2} \right).$$

Also, by the mean value theorem

$$[Nt] \left((y + 2/[Nt])^{2H-1} - y^{2H-1} \right) = \frac{2((2H-1))}{\theta^{2(1-H)}} \leq \frac{2(2H-1)}{y^{2(1-H)}},$$

for some $\theta \in (y, y + 2/[Nt])$. Similar integrable bound holds for the second difference in the integrand above, and the dominated convergence theorem yields

$$\begin{aligned}
& \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \sum_{i_1, i_2=0}^{[Nt]-j} \rho_k^{i_1+i_2} (1 - \rho_k^2) C_2 \delta_k \\
& = 2[Nt]^{2H} L(N) \int_{x=0}^1 \int_{y=0}^{1-x} \left(y^{2H-2} - (y+1-x)^{2H-2} \right) dx dy (1 + o(1)),
\end{aligned}$$

which shows that the series in the second term converges almost surely, and that the variance of the second term has the same order as the variance of the first term, namely $L(N)[Nt]^{2H}$.

□

The next theorem gives the asymptotic behavior of the cumulants of the partial sums process.

Theorem 2.3.4. *Assume that the m -th cumulant of the centered partial sums of the superposition of processes satisfies condition (A) for all $q \geq 2$, condition (B), and has $\lambda^{(k)} = \lambda/k$ and $\delta_k = k^{-(1+2(1-H))}$, $\frac{1}{2} < H < 1$. In addition, for each k , the cumulant function of τ_k , given in the equation (2.11), is analytic around the origin. Then the m -th cumulant has the following asymptotic behavior:*

$$\kappa_{m,N} = D_m L(N) [Nt]^{m-2(1-H)} (1 + o(1))$$

as $N \rightarrow \infty$, where the $D_m = C_m K$ for some positive constant K .

Proof. Using (2.8), the logarithm of the characteristic function of the partial sums process can be written as

$$\begin{aligned} & \log E \exp \{iu(S_\infty([Nt]) - ES_\infty([Nt]))\} \\ &= \sum_{k=1}^{\infty} \log E \exp \left\{ ib_{[Nt]}^{(k)} u \tau^{(k)}(0) \right\} + \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \log E \exp \left\{ ia_{[Nt]-j}^{(k)} u V^{(k)}(j) \right\}. \end{aligned}$$

Under this theorem's assumption (basically, assumption (A) and that the cumulant function of τ_k is analytic), the logarithm of the characteristic function of $\tau^{(k)}(0)$ can be expanded

$$\log E \exp \left\{ iu \tau^{(k)}(0) \right\} = \sum_{m=2}^{\infty} \frac{(iu)^m}{m!} C_m \delta_k,$$

where the summation is from $m = 2$ due to centering. From (2.11), the logarithm of the

characteristic function of $V^{(k)}(j)$ can also be expanded as follows:

$$\begin{aligned}
\log E \exp \left\{ iuV^{(k)}(j) \right\} &= E \exp \left\{ iu\tau^{(k)}(i) \right\} - E \exp \left\{ iu\rho_k\tau^{(k)}(i-1) \right\} \\
&= \sum_{m=2}^{\infty} \frac{(iu)^m}{m!} C_m \delta_k - \sum_{m=2}^{\infty} \frac{(iu\rho_k)^m}{m!} C_m \delta_k \\
&= \sum_{m=2}^{\infty} \frac{(iu)^m}{m!} C_m (1 - \rho_k^m) \delta_k.
\end{aligned}$$

Therefore the m -th cumulant of the centered partial sums process is

$$\kappa_{m,N} = C_m \sum_{k=1}^{\infty} \left(b_{[Nt]}^{(k)} \right)^m \delta_k + C_m \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \left(a_{[Nt]-j}^{(k)} \right)^m (1 - \rho_k^m) \delta_k = I + II.$$

Consider the first term:

$$\begin{aligned}
I &= C_m \sum_{k=1}^{\infty} \left(b_{[Nt]}^{(k)} \right)^m \delta_k = C_m \sum_{k=1}^{\infty} \delta_k \left(\sum_{i=1}^{[Nt]} \rho_k^i \right)^m \\
&= C_m \sum_{i_1, \dots, i_m=1}^{[Nt]} \sum_{k=1}^{\infty} \delta_k \rho_k^{i_1 + \dots + i_m} = \frac{C_m}{C_2} \sum_{i_1, \dots, i_m=1}^{[Nt]} \frac{L(i_1 + \dots + i_m)}{(i_1 + \dots + i_m)^{2(1-H)}} \\
&= \frac{C_m [Nt]^m}{C_2} \int_0^1 \dots \int_0^1 \frac{L([Nt](x_1 + \dots + x_m))}{([Nt](x_1 + \dots + x_m))^{2(1-H)}} dx_1 \dots dx_m (1 + o(1)) \\
&= \frac{C_m [Nt]^m L(N)}{C_2 [Nt]^{2(1-H)}} \int_0^1 \dots \int_0^1 \frac{dx_1 \dots dx_m}{(x_1 + \dots + x_m)^{2(1-H)}} (1 + o(1)),
\end{aligned}$$

where we used Remark 1.5.1 and the dominated convergence argument for the slowly varying function. This shows that the first part of the expression for the m -th cumulant behaves like $L(N)[Nt]^{m-2(1-H)}$ multiplied by a constant

$$D_{m,I} = \frac{C_m}{C_2} \int_0^1 \dots \int_0^1 \frac{dx_1 \dots dx_m}{(x_1 + \dots + x_m)^{2(1-H)}}.$$

Now consider the second term

$$\begin{aligned}
II &= C_m \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \left(a_{[Nt]-j}^{(k)} \right)^m (1 - \rho_k^m) \delta_k \\
&= C_m \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \left(\sum_{i=0}^{[Nt]-j} \rho_k^i \right)^m (1 - \rho_k^m) \delta_k \\
&= C_m \sum_{j=1}^{[Nt]} \sum_{k=1}^{\infty} \sum_{i_1, \dots, i_m=0}^{[Nt]-j} \rho_k^{i_1 + \dots + i_m} (1 - \rho_k^m) \delta_k = \frac{C_m}{C_2} \\
&\times \sum_{j=1}^{[Nt]} \sum_{i_1, \dots, i_m=0}^{[Nt]-j} \left(\frac{L(i_1 + \dots + i_m)}{(i_1 + \dots + i_m)^{2(1-H)}} - \frac{L(i_1 + \dots + i_m + m)}{(i_1 + \dots + i_m + m)^{2(1-H)}} \right) = \frac{C_m [Nt]^{m+1}}{C_2 \lambda^{2(1-H)}} \\
&\times \int_{x=0}^1 \int_{y_1=0}^{1-x} \dots \int_{y_m=0}^{1-x} \left(\frac{L([Nt](y_1 + \dots + y_m))}{([Nt](y_1 + \dots + y_m))^{2(1-H)}} - \frac{L([Nt](y_1 + \dots + y_m) + m)}{([Nt](y_1 + \dots + y_m) + m)^{2(1-H)}} \right) \\
&\times dy_1 \dots dy_m dx (1 + o(1))
\end{aligned}$$

Remark 1.5.1 and the dominated convergence argument followed by integration with respect to y_m yield

$$\begin{aligned}
II &= \frac{C_m L(N) [Nt]^{m-2(1-H)+1}}{C_2} \\
&\times \int_{x=0}^1 \int_{y_1=0}^{1-x} \dots \int_{y_m=0}^{1-x} \left(\frac{1}{(y_1 + \dots + y_m)^{2(1-H)}} - \frac{1}{((y_1 + \dots + y_m) + m/[Nt])^{2(1-H)}} \right) \\
&\times dy_1 \dots dy_m dx (1 + o(1)) = \frac{C_m L(N) [Nt]^{m-2(1-H)}}{C_2 (2H-1)} \\
&\times \int_{x=0}^1 \int_{y_1=0}^{1-x} \dots \int_{y_{m-1}=0}^{1-x} [Nt] \left((y_1 + \dots + y_{m-1} + m/[Nt])^{2H-1} - (y_1 + \dots + y_{m-1})^{2H-1} \right. \\
&\left. - ((y_1 + \dots + y_{m-1} + 1 - x + m/[Nt])^{2H-1} - (y_1 + \dots + y_{m-1} + 1 - x)^{2H-1}) \right) \\
&\times dy_1 \dots dy_{m-1} (1 + o(1)).
\end{aligned}$$

Note as $[Nt] \rightarrow \infty$ the limit of the integrand is

$$m(2H - 1)((y_1 + \cdots + y_{m-1})^{2H-2} - (y_1 + \cdots + y_{m-1} + 1 - x)^{2H-2}).$$

The same argument as in the proof of Lemma 2.3.3 and the dominated convergence theorem yield

$$\begin{aligned} II &= \frac{mC_m L(N)[Nt]^{m-2(1-H)}}{C_2} \int_{x=0}^1 \int_{y_1=0}^{1-x} \cdots \int_{y_{m-1}=0}^{1-x} \left((y_1 + \cdots + y_{m-1})^{2H-2} \right. \\ &\quad \left. - (y_1 + \cdots + y_{m-1} + 1 - x)^{2H-2} \right) dy_1 \dots dy_{m-1} dx (1 + o(1)) \\ &= D_{m,II} L(N)[Nt]^{m-2(1-H)} (1 + o(1)) \end{aligned}$$

with

$$\begin{aligned} D_{m,II} &= \frac{mC_m}{C_2} \int_{x=0}^1 \int_{y_1=0}^{1-x} \cdots \int_{y_{m-1}=0}^{1-x} \\ &\quad \left((y_1 + \cdots + y_{m-1})^{2H-2} - (y_1 + \cdots + y_{m-1} + 1 - x)^{2H-2} \right) dy_1 \dots dy_{m-1} dx. \end{aligned}$$

Thus the asymptotic behavior of the second term is the same as of the first term, namely $L(N)[Nt]^{m-2(1-H)}$. \square

Remark 2.3.1. *The analyticity of the cumulant function κ_X in the above theorem ensures the existence of all the cumulants of the marginal distribution of the underlying supOU process X . It follows from Lemma 3.1, paper [34], that the cumulant function of $X(t)$ is analytic in the neighborhood of the origin if there exists $a > 0$ such that $Ee^{a|X(t)|} < 1$. This implies in particular that all the moments and cumulants of $X(t)$ exist. This condition is satisfied for many self-decomposable distributions, for example the IG, NIG, and others. Other examples of supOU processes satisfying the above conditions can be obtained by taking the marginal*

distribution to be gamma, variance gamma, tempered stable, Eulers gamma, or z-distribution.

For more details, see [34].

Corollary 2.3.5. *Under the assumptions of Theorem 2.3.4, the centered partial sum process $\{S_\infty(t) - ES_\infty(t), t \geq 0\}$ is intermittent.*

Proof. Let $Y(u) = S_\infty(u) - ES_\infty(u)$. We show intermittency at $p = 2$ and $r = 4$. From Theorem 2.3.4, the m -th cumulant of $Y([Nt])$ equals

$$D_m L(N)[Nt]^{m-2(1-H)}(1 + o(1))$$

as $N \rightarrow \infty$. Since $L([Nt])/L(N) \rightarrow 1$ as $N \rightarrow \infty$ for any $t > 0$, the m -th cumulant of $Y(u)$, denoted by $\tilde{\kappa}_{m,u}$, equals

$$D_m L(u)u^{m-2(1-H)}(1 + o(1))$$

as $u \rightarrow \infty$.

Using the relation between moments and cumulants it follows from Theorem 2.3.4 that

$$E|Y(u)|^2 = \tilde{\kappa}_{2,u} + \tilde{\kappa}_{1,u}^2 = D_2 L(u)u^{2H}(1 + o(1)),$$

$$E|Y(u)|^4 = \tilde{\kappa}_{4,u} + 3\tilde{\kappa}_{2,u}^2 = D_4 L(u)u^{2H+2}(1 + o(1)) + 3D_2^2 L(u)^2 u^{4H}(1 + o(1))$$

as $u \rightarrow \infty$. Since $H < 1$ implies $2H + 2 > 4H$, we have

$$\tau(2) = 2H,$$

$$\tau(4) = 2H + 2,$$

and thus $\tau(2)/2 < \tau(4)/4$. □

Note that the behavior of moments shown in the proof implies that

$$EY(u)^4/(EY(u)^2)^2$$

grows to infinity as $u \rightarrow \infty$, the behavior noted by Frisch ([27, Section 8.2] as a manifestation of intermittency. Other examples of unusual growth of moments are given in [62] in the context of fractional diffusion.

Similar behavior of cumulants was obtained in [4, Example 4.1] for a case of continuous (integrated) superpositions of OU type processes and also in [22] in the context of proving central limit theorem type results. The authors of [22] noted that the existence of the limit was unlikely given the behavior of cumulants. We concur with this statement, but showing that there is no weak limit under intermittency in the usual partial sum setting remains an open problem. The consideration of why the existence of a weak limit is unlikely is as follows.

If the limit of the partial sum process for the infinite superposition existed in the sense of convergence of all finite dimensional distribution, then by the Lamperti's theorem (see, for example, [21, Theorem 2.1.1]), the norming had to be a regularly varying function of N . That is, the weak convergence would hold for

$$N^a L_1(N) (S_\infty([Nt]) - ES_\infty([Nt]))$$

for some $a \in \mathbb{R}$ and a slowly varying at infinity function L_1 . However, no matter what $a \in \mathbb{R}$ is chosen, all cumulants of the centered and normed partial sum cannot converge. This is because the m -th cumulant of $N^a (S_\infty([Nt]) - ES_\infty([Nt]))$ behaves like $N^{m(a+1)-2(1-H)}$.

Also note that for even q , the scaling function defined in (2.3) in this case is $\tau(q) = q - 2(1 - H)$, and

$$\frac{\tau(q)}{q} = 1 - \frac{2(1 - H)}{q}$$

is strictly increasing in q . The term $-2(1 - H)$ in the exponent of the asymptotic behavior of the cumulants

$$\kappa_{q,N} = D_q L(N) [Nt]^{q-2(1-H)} (1 + o(1))$$

gives the reason for both the increasing behavior of $\tau(q)/q$ and for the lack of norming that would make cumulants converge. Of course, convergence of cumulants provides a sufficient means for proving the existence of the limit by showing the convergence of the characteristic function. The formal link between intermittency and lack of the limit theorems needs to be further developed for the partial sums and other sequences of stochastic processes.

Chapter 3

Applications in finance

For the past thirty years, an extensive research has been devoted to finding classes of stochastic models that can capture the essential statistical properties of stock price returns.

The Black-Scholes model, the most renowned and widely used model by far, also referred to as geometric Brownian motion model or GBM for short, was first introduced by Black, Scholes and Merton, see [12]. It models the stock price evolution as,

$$S_t = S_0 \exp \left\{ \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B(t) \right\}, t > 0. \quad (3.1)$$

Even though the Black-Scholes model is the most widely used model so far, as we'll see later on, it is not able to capture some important features of the log-returns and therefore some new alternatives were suggested in order to improve the overall model fit. The assumptions it makes about the nature of the stock price process, in turn impose the log-returns to be independent, stationary and normally distributed, the characteristics that differ from the empirical features of the returns listed above.

First improvement on the GBM in (3.1), that uses Levy processes instead of Brownian motion as a noise component was suggested by Merton in 1970's [56]. Since most of the properties of the Levy process, for example independent and stationary increments, having a version that is a.s. cadlag will hold true for the Merton's improved model as well, hence the new model allows for jumps in the stock price evolution. The class of possible marginal

distributions or infinitely divisible distributions is also larger than only Gaussian in geometric Brownian motion model.

The next refinement of the Black-Scholes model, proposed by Mandelbrot (1967) [53], had a completely different flavor. He used fractional Brownian motion $B_H(t)$ that is subordinated by positive stochastic process $\theta(t)$ that is independent of FBM, i.e. he modeled a stock price evolution with the process: $B_H(\theta(t))$. Mandelbrot's model belongs to a group of subordinator models, i.e. group of models where instead of the deterministic one is taking an increasing stochastic process. Note that financial interpretation of the time in the subordinator type models is completely different than in the regular stochastic models, see Howison and Lamper (2001) [41].

Another popular class of models that aimed to correct for the volatility parameter σ , in the equation (3.1), is called a class of *stochastic volatility models*. Assumptions of the GBM imply that volatility is constant, while in reality this parameter depends on both, time and the strike price. This phenomenon, in financial literature is referred to as the *volatility smile*. Widely recognized and used models from this class are Heston, Banrdorff-Nielsen and Shepard (BNS) and superposition of Ornstein-Uhlenbeck type (SupOU) model [4, 8].

Finally, a model that we are going to use for calibration purposes is a slight modification of the Heyde's Fractional Activity time GBM model, or FATGBM for short, a version that is in the literature frequently referred to as skewed FATGBM model. This model belongs to a class of subordinator models, i.e. models where Brownian motion process is subordinated to a specific activity time process, finite (or infinite) supOU process. In fact, it's been shown in [8] that the BNS stochastic volatility model is a special type of FATGBM model with a specific choice of the activity time. The equivalence between the two is a consequence of self-similarity of Brownian motion.

3.1 The Black-Scholes Model:

Geometric Brownian Motion (GBM)

In this section, we will briefly present the rationale behind stock price modeling and why Geometric Brownian motion seems to be a good representation of the reality.

In various option pricing models the time evolution of a stock price $\{S_t\}_{t \geq 0}$ is modeled by describing a price change in a small time interval: $\Delta S_t = S_{t+\Delta t} - S_t$. Now, the question is if it is better to use absolute, relative or log-returns. One can find in the finance literature that the most commonly used metrics are relative and log-returns, both having their strengths and weaknesses and one will be used over the other depending on the specific situation.

In order to be able to compare investments in different securities, instead of the absolute price change: ΔS_t , we will focus on modeling the relative price change (relative return)¹: $\Delta S_t/S_t$. It is reasonable to expect that in a small time interval Δt , return can be decomposed into a systematic and a random component:

$$\Delta S_t/S_t = \underbrace{\mu \Delta t}_{\text{systematic}} + \underbrace{\sigma \Delta B(t)}_{\text{random}}, \quad S_0 > 0 \quad (3.2)$$

where the systematic part is expected value of the return and the random part is noise.

Remark 3.1.1. *The equation 3.2 simply says that since the stock price fluctuates stochastically, it is reasonable to assume that it oscillates around some expected (deterministic) value and that oscillation is due to random noise.*

It is natural to assume that over a short period of time, both, the expected return (sys-

¹In finance, the relative price change is usually called a simple return. When evaluating portfolios that have more than one security simple returns are more convenient metric to use than absolute or log-returns. However, most authors in the financial literature prefer working with log-returns due to their additivity over time and the exponential nature of many financial models.

tematic term) and the variance of random component are proportional to the time increment Δt . Financial intuition here, is that in a short time span, the expected increase or decrease in price is proportional to the time length and also the longer the Δt the more volatile stock price may be.

Therefore, in the Black-Scholes setup, the deterministic part is of the form: $\mu\Delta t$, where the proportionality parameter μ , called the *drift*, represents the mean rate of return of the stock and it is assumed to be constant over time. At the same time, the random component in the Black-Scholes model is modeled by: $\sigma\Delta B(t)$ where $\Delta B(t)$ is the noise increment driven by Brownian motion and $\sigma > 0$ is assumed to be constant over time. So, the random term is driven by normally distributed error with variance $\sigma^2\Delta t$. Parameter σ governs how volatile² the stock price is and is commonly referred to as *volatility* of the stock.

Remark 3.1.2. *The above decomposition of the relative stock return into deterministic and random component can be applied in general, regardless of the model. Hence, how the deterministic term μ changes with time and if it should be constant function of time or random and how the error term is distributed is all tied to a specific model.*

Heuristically, as $\Delta t \rightarrow 0$, the equation 3.2 becomes SDE of the form:

$$dS_t/S_t = \mu dt + \sigma dB(t), \quad S_0 > 0 \tag{3.3}$$

where $\{B(t)\}_{t \geq 0}$ is a BM process. The stochastic differential equation 3.3 has an almost surely unique, strong solution, called *Geometric Brownian Motion (GBM)* and it is given by

² σ controls how much effect the noise have on the stock return, i.e. how much the stock price fluctuates.

the following expression:

$$S_t = S_0 \exp \left\{ \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B(t) \right\}, \quad t > 0. \quad (3.4)$$

Remark 3.1.3. Notice that GBM is an exponential of the Brownian motion process with drift. From 3.1, we can conclude that log returns of the stock price process are normally distributed: $\log \frac{S_t}{S_0} \sim \mathcal{N} \left(\left(\mu - \frac{1}{2} \sigma^2 \right) t, \sigma^2 t \right)$ and that stock price itself is log-normally distributed.

3.2 Empirical evidence against the Black-Scholes model

As we mentioned earlier, the empirical evidence suggests that the classical Black-Scholes model does not describe the properties of financial time series very well. We present here, the essential ways in which statistical features of real log-returns differ from those implied by the Black-Scholes, see the list below.

1. Empirical distribution of returns seems to be more sharply peaked with fatter tails.
2. Dependence structure of the returns:
 - (a) The empirical returns are uncorrelated (independent in the Gaussian case).
 - (b) The real absolute and squared returns exhibit strong and persistent dependence (as opposed to Black-Scholes returns).
3. The conditional variance of the returns $E[(X_t - \mu)^2 \mid \mathcal{F}_{t-1}]$, for the empirical returns behaves erratically³ as a function of time, i.e. seems to behave randomly and exhibits

³Note, recall the volatility parameter from GBM SDE (equations 3.2, 3.3), then this statement simply says that the volatility parameter changes stochastically over time.

possible signs of clustering and long-range dependence, while under the Black-Scholes assumption, it is constant: $E[(X_t - \mu)^2 | \mathcal{F}_{t-1}] = \sigma^2$, for all $t > 0$.

In subsequent sections, we provide more details about each stylized fact along with accompanying figures and tables. The data consists of daily adjusted closing prices of the Standard and Poor's 500 Index (S&P 500), from the 9th of September 2005 till the 9th of September 2016.

3.2.1 The non-Gaussian character of the returns

While the GBM assumption about the stock price evolution implies that log-returns are normally distributed (see the remark 3.1.3), it has been shown in practice that the returns have actually higher peaks and fatter tails (Leptokurtic distributions), see figure 3.1.

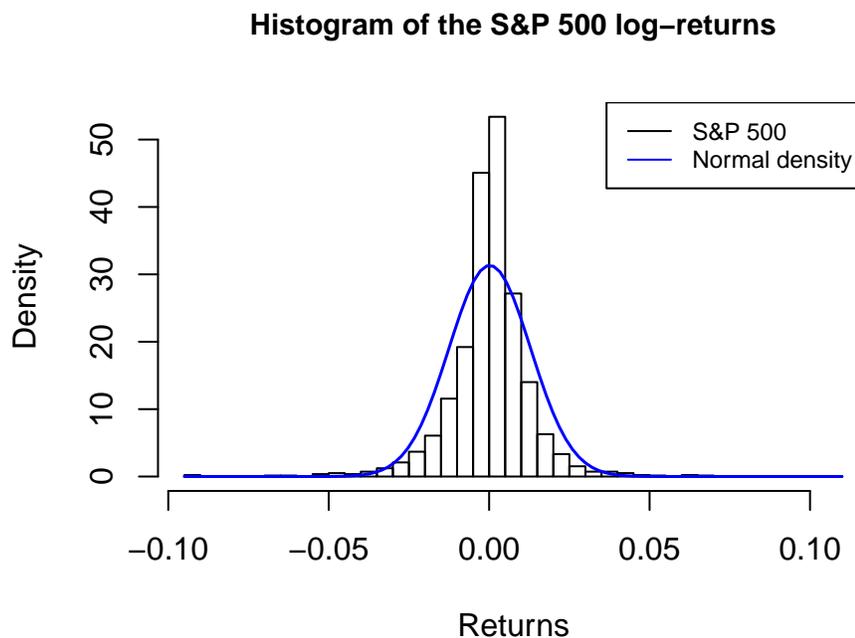


Figure 3.1: Histogram of the S&P500 returns against the Normal curve fit.

There are different ways one can showcase non-normality of the returns. The most popular among the practitioners is to test normality of the returns by using either Jarque-Bera, Kolmogorov-Smirnov, Shapiro-Wilk or Anderson-Darling test, where each has its own benefits and weaknesses. Our null hypothesis is that the log returns are i.i.d. normal under the GBM, and we look for the evidence against this null hypothesis. In the table 3.1 below, we present the results of the tests for the S&P 500 real and simulated returns alongside other stock returns, from the period of 2005 to 2016. One can clearly see evidence against the normality, similarly as in the figure above 3.1. As we continue we will see more and more evidence against the normality assumption.

| Asset returns | N | Sam. skewness | Sam. kurtosis | p-value | Decision (H0) |
|---------------|------|---------------|---------------|-------------|---------------|
| S&P500 (real) | 2770 | -0.3325 | 13.5867 | $\ll 0.001$ | reject |
| S&P500 (sim) | 2770 | -0.0033 | 3.0491 | > 0.5 | retain |
| DJI | 2770 | -0.0961 | 13.4402 | $\ll 0.001$ | reject |
| NDX | 2770 | -0.1746 | 10.9485 | $\ll 0.001$ | reject |
| GSK | 2770 | -0.3673 | 8.7517 | $\ll 0.001$ | reject |
| PFE | 2770 | -0.194 | 10.2959 | $\ll 0.001$ | reject |
| COF | 2770 | -0.3491 | 17.6885 | $\ll 0.001$ | reject |
| BAC | 2770 | -0.3049 | 24.1146 | $\ll 0.001$ | reject |
| MSFT | 2770 | 0.0729 | 13.2692 | $\ll 0.001$ | reject |
| IBM | 2770 | -0.1871 | 8.5777 | $\ll 0.001$ | reject |
| GOOG | 2770 | 0.5622 | 13.2928 | $\ll 0.001$ | reject |

Table 3.1: Values for sample skewness, kurtosis and the resulting decisions of the normality test.

Note that Jarque-Bera⁴ is a goodness of fit type test, which tests if sample skewness and kurtosis match the ones of normal distribution, i.e. the null hypothesis is a joint hypothesis of both, skewness and excess kurtosis being zero. Hence, it is similar to the tests we suggested. In the next section, as we shall see in the table 3.2, even only the sample skewness and kurtosis may indicate a degree up to a given distribution is asymmetric and has heavier tails than normal.

3.2.1.1 Skewness (asymmetry) and excess kurtosis (fat tails)

Let us denote $\mu = \mu_1 = E[X]$ as mean, $\mu_2 = E[(X - \mu)^2]$ as variance and $\mu_k = E[(X - \mu)^k]$ as the k-th central moment of a random variable X . Let us also assume that $\{X_i\}_{i=1}^N$, is a sample of size N , from the distribution X , and that its sample (central) moments are given by, $\hat{\mu}_k^{(N)} = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X}^{(N)})^k$, for $k = 1, 2, 3, 4$, and where $\bar{X}^{(N)} = \hat{\mu}_1^{(N)} = \frac{1}{N} \sum_{i=1}^N X_i$ is its sample average. In our case, we will assume that $\{X_i\}_{i=1}^N$, is a sample of returns that are uncorrelated (not necessarily independent) and identically distributed random variables.

Finally, skewness (γ_1) is defined as a quantity that measures the degree to which a given distribution X is asymmetric, while excess kurtosis (γ_2) measures how heavy its tails are comparing to normal. Below we give both formulas, for skewness and kurtosis respectively:

$$\gamma_1 = \frac{E[(X - \mu)^3]}{(Var[X])^{3/2}} = \frac{\mu_3}{(\mu_2)^{3/2}}, \quad \gamma_2 = \frac{E[(X - \mu)^4]}{(var[X])^2} = \frac{\mu_4}{(\mu_2)^2}.$$

Since we assumed that returns are uncorrelated, in order to estimate the skewness and kurtosis parameters of the marginal distribution of the returns, we need to use estimation techniques that are robust to the presence of correlation in our data. So, given the assump-

⁴This test is overly sensitive for smaller sample sizes and typical rule of thumb here is a sample size larger than 2000.

tions below are satisfied, we will employ the method of moments estimation, which allows us to exploit the fact that, for N large enough, theoretical moments of the distribution are relatively close to the sample moments. This is true provided that certain mixing properties and therefore ergodicity of the returns are satisfied. This implies consistency and asymptotic normality of the sample moments and hence of sample skewness and kurtosis. Note that in our case mixing property is satisfied, see the theorem 2.4. Finally, the formulas for the sample skewness and kurtosis are given below,

$$\hat{\gamma}_1^{(N)} = \frac{\hat{\mu}_3^{(N)}}{\left(\hat{\mu}_2^{(N)}\right)^{3/2}}, \quad \hat{\gamma}_2^{(N)} = \frac{\hat{\mu}_4^{(N)}}{\left(\hat{\mu}_2^{(N)}\right)^2}.$$

Remark 3.2.1. *Note that with respect to skewness, there are two types of distributions: a negatively skewed (has longer tail to the left than to the right) and a positively skewed distribution (reverse is true). Regarding the kurtosis, a mesokurtic distribution (behaves like a Normal with kurtosis value of three), a platykurtic distribution (has flatter top and kurtosis less than three), and finally a leptokurtic distribution (has higher peaks and kurtosis greater than three). In general, kurtosis greater than 3 indicates a distribution whose tails decrease slower than the Normal distribution and it is more sharply peaked.*

Both skewness and kurtosis are usually measured as a reference to a Normal distribution, that has skewness equal to zero and kurtosis equal to three (or excess kurtosis to zero). In addition to the normality tests performed above (table 3.1), below we provide a table with values of sample skewness and kurtosis, table 3.2, which seems to indicate that, for the majority of our data sets, sample skewness is slightly negative compared to Normal, while sample kurtosis is significantly higher. To be more specific, the sample skewness of our S&P 500 index is only slightly negative, around -0.3325 , while typically, the observed

daily log-returns of different indices may exhibit at least some significant negative skewness. Regarding the kurtosis, we clearly see that S&P 500 returns, as well as majority of other securities have kurtosis significantly higher than three.

From the results we conclude that for majority of our data, the empirical distribution is more Leptokurtic and has also slightly longer tail to the left. The fact that the distribution of the real returns are more leptokurtic than Normal has already been noted by Fama (1965), see [23]. The financial intuition behind a model with high kurtosis is that the large movements in a stock price occur more frequently than in the model with Normally distributed increments. This phenomena has been observed in the markets for quite some time, and one recent example is the 2007/2008 financial crisis or the mortgage crisis. This feature is one of the reasons for considering modeling a stock price processes via jumps. Regarding skewness, one possible explanation behind the negative skewness⁵ is that investors tend to react more strongly to negative information rather than the positive information, see Rydberg (2000) [61].

⁵In addition, another interesting perspective to positively/negatively skewed returns, is that investors are more willing to buy stocks that are positively skewed (meaning frequent small losses and a few extreme gains), which intuitively means that investors would rather accept smaller expected returns even slightly negative ones with an occasional extreme return (either gain or loss) than invest in the negatively skewed stock returns. A real life example of this is buying lottery tickets or "penny stocks", small-cap growth stocks, etc...

| Risky asset returns | N | Sam. Mean | Sam. St.dev | Sam. Skewness | Kurtosis |
|---------------------|------|-------------|-------------|---------------|----------|
| S&P500 (real) | 2770 | 0.00019458 | 0.0127 | -0.3325 | 13.5867 |
| S&P500 (sim GBM) | 2770 | -0.00010136 | 0.0125 | -0.0033 | 3.0491 |
| DJI | 2770 | 0.00019027 | 0.0117 | -0.0961 | 13.4402 |
| NDX | 2770 | 0.00038597 | 0.0137 | -0.1746 | 10.9485 |
| GSK | 2770 | 0.00013589 | 0.0137 | -0.3673 | 8.7517 |
| PFE | 2770 | 0.00025785 | 0.0144 | -0.194 | 10.2959 |
| COF | 2770 | 0.000016862 | 0.0312 | -0.3491 | 17.6885 |
| BAC | 2770 | -0.00027863 | 0.0353 | -0.3049 | 24.1146 |
| MSFT | 2770 | 0.00036112 | 0.0174 | 0.0729 | 13.2692 |
| IBM | 2770 | 0.00031453 | 0.0138 | -0.1871 | 8.5777 |
| GOOG | 2770 | 0.00058731 | 0.0192 | 0.5622 | 13.2928 |

Table 3.2: Values for sample skewness, kurtosis mean and standard deviation

3.2.2 The dependence structure of the log-returns: regular, absolute and squared returns

Another interesting phenomenon that is not being captured well by the Black-Scholes model is the dependence structure of the log-returns and their absolute and squared values. The dependence structure of the returns is very important because it tell us how predictable stock returns are and how much we can infer about the future by only using historical data.

A standard graphical approach for exploring dependence in returns is the autocorrelation function (ACF) which measures correlation between returns that are k-lags away. In order to investigate further, assuming that we are given a sample of N-log-returns, we compute

and plot a sample autocovariance and autocorrelation function, given by the formulas below,

$$\hat{R}^{(N)}(k) = \frac{1}{N} \sum_{t=1}^{N-k} (X_t - \bar{X})(X_{t+k} - \bar{X}),$$

and

$$\hat{\rho}^{(N)}(k) = \frac{\hat{R}^{(N)}(k)}{\hat{R}^{(N)}(0)},$$

where \bar{X} is the sample average and k represents number of lags.

If we assume that given sample of log-returns is coming from the population generated by white noise then, for relatively large N , the sample autocorrelation estimator has Normal distribution with variance $1/N$, where number of lags should be much smaller than total number of observations, $1 \leq k \leq m \ll N$, see Anderson (1942). This means that we can test if given returns are purely random (white noise) by checking if observed values of autocorrelation lie within the 95% confidence bounds based on the asymptotic values of the sample autocorrelation coefficient under the white noise assumption. These bounds correspond to the levels of $\pm 1.96/\sqrt{N}$. The interpretation of the test is following: we would reject the H_0 : population being generated by white noise, only if, for the fixed number of lags $1 \leq k \leq m \ll N$, we observe that $\hat{\rho}^{(N)}(k)$ is falling outside of the interval bounds more than 5% of the time, in which case we would call these correlations significant.

In our case, from the figure below 3.2, we see that the sample autocorrelations of real log-returns of various risky assets diminish rapidly and that there is little or no autocorrelation present in returns past couple of lags, i.e. we can say that the correlations are statistically insignificant (follow the white noise), see the green line. This is not enough to say that returns are independent, one can only conclude that they are uncorrelated. If we want to understand the correlation structure of the returns better, than it is necessary to study also

a serial correlation of the quadratic and absolute returns. This, in turn, may reveal some relevant information about how the variance changes with time and if the accumulation of large/small absolute returns persists as $t \rightarrow \infty$. For example, the strong dependence of the absolute returns, similarly as with squared returns, tells us if any accumulation of big movements of risky asset prices exist in our data see below.

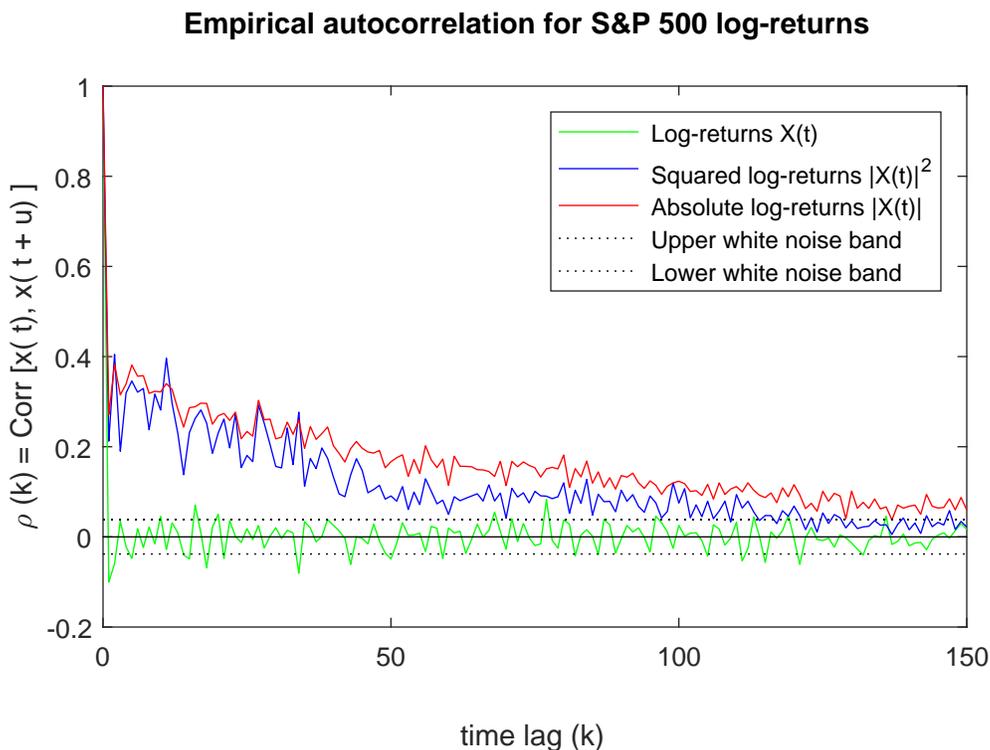


Figure 3.2: Empirical autocorrelation function of regular, squared and absolute log-returns for the S&P500 index.

In addition, in the figure 3.2, we also present sample autocorrelations for the absolute and squared log-returns, $|X_t|$ and X_t^2 respectively, $t = 0, \dots, N - k$, and where the black dotted lines on the plot represent the upper and lower bound of the 95% confidence interval for the sample autocorrelations under the H_0 (Gaussian white noise assumption). Interestingly, the figure does show some strong persistence in autocorrelations for both, squared and absolute

log-returns, and so there is a strong evidence against the GBM hypothesis. This empirical evidence matches the previous observations of Heyde and Liu (2001) [39], Cont (2001) [16], and also supports the "stylized facts" outlined in Granger (2005) [35]. Nowadays, most of the authors will consider the empirical features of risky asset returns as a foundation for building a reliable model that is a good representation of reality.

Let us compare our finding about the empirical log-returns, given in the figure above against the simulated GBM returns, calibrated on the same S&P500 data. We have decided to take the most simple approach for calibration, i.e. we used (historical) sample mean and variance in order to estimate the drift and volatility parameters. From this plot, see figure 3.3, one can clearly see that there is little autocorrelation present for all three sequences of log-returns, regular, squared and absolute. This is considered to be an additional evidence of the Black-Scholes deficiencies.

Recall, that a process is long-range dependent (LRD)⁶ if its autocorrelation function is not integrable⁷, which in continuous case turns out to be, $\int_{-\infty}^{+\infty} |\rho(t)| dt = \infty$, for $t \in \mathbb{R}$, and $\sum_{-\infty}^{+\infty} |\rho(k)| = \infty$, for the discrete case. In contrast, a series whose autocorrelation is integrable in \mathbb{R} may display short-range dependence.

Many authors consider long-range/short-range dependence as very important empirical feature of the returns. In general there are two schools of thought on this subject. One group of authors claims that the long-range dependence exists for risky assets and that it has been empirically proven, see Heyde and Yang (1997) [40] and references therein, however, this is not an universally accepted view. One way to distinguish between whether we have strong

⁶Some other acronyms for the LRD/SRD processes is either a long/strong memory process or strong dependence process, while for the short range dependence we have either short memory, weak memory or weak dependence process.

⁷Note, that there is also one more representation of the LRD and that is via product of power function and a slowly varying function.

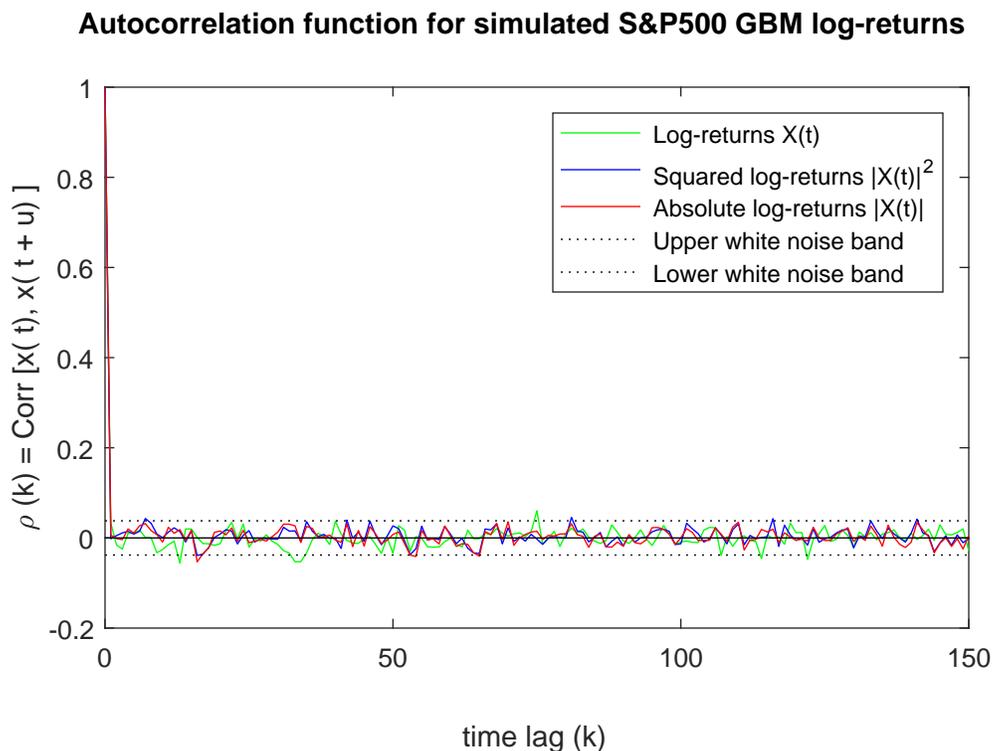


Figure 3.3: Autocorrelation function of regular, squared and absolute log-returns for the simulated Geometric Brownian motion, calibrated on the S&P500 index.

or weak dependence in our data is to estimate, the so called Hurst parameter, which, in a way, represent a direct link between the intensity of dependence of a given process and its self-similar scaling nature.

One of the reasons that the infinite superposition model was introduced was as a long range dependent solution that will keep the the most useful properties of the finite superposition model⁸ and introduce the long range dependence in addition. For this thesis, due to our findings we will refrain ourselves from using the long range dependence model and instead work only with the finite superposition model, that we know they is short range dependent.

⁸To mention the most important ones, tractability, non-normal class of marginal distributions and flexible correlation structure.

If these correlations are statistically significant we have an evidence of predictability. One can use Ljung-Box (LB) test in order to check joint significance of the auto correlation coefficients over more than one lag.

3.2.3 Stochastic Volatility and clustering

In a stochastic volatility model, price X_t is a solution of the following SDE, given by

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

which exists under certain conditions on the coefficients. The functions μ and σ are known as drift and the volatility of the underlying process X_t . The volatility function is also referred to an instantaneous absolute volatility. In the above scenario, assuming that drift is deterministic, the volatility function term $\sigma(t, X_t)$ will directly influence the variance of the Brownian motion. In particular, in the Black-Scholes model where GBM evolution is assumed, both drift and volatility functions are constants, $\mu(t, X_t) = \mu$ and $\sigma(t, X_t) = \sigma$, hence the volatility parameter is equal to the variance of the returns. Hence it simply describes how much effect does a noise term have on the price, see the figures 3.4 below.

In order to understand and model the volatility, one must start with the following question: How to use the Black-Scholes model in order to price an option today (now) by using the past information we have (historical price movements of the underlying stock)?

A first step in option pricing is to find/estimate the value of the volatility parameter σ , which is assumed to be constant under the GBM. There is rich literature on how to measure and estimate the volatility function $\sigma(t, X_t)$ in general, but since this is not the topic of the discussion, for the volatility calibration purposes we will focus only on estimating volatility

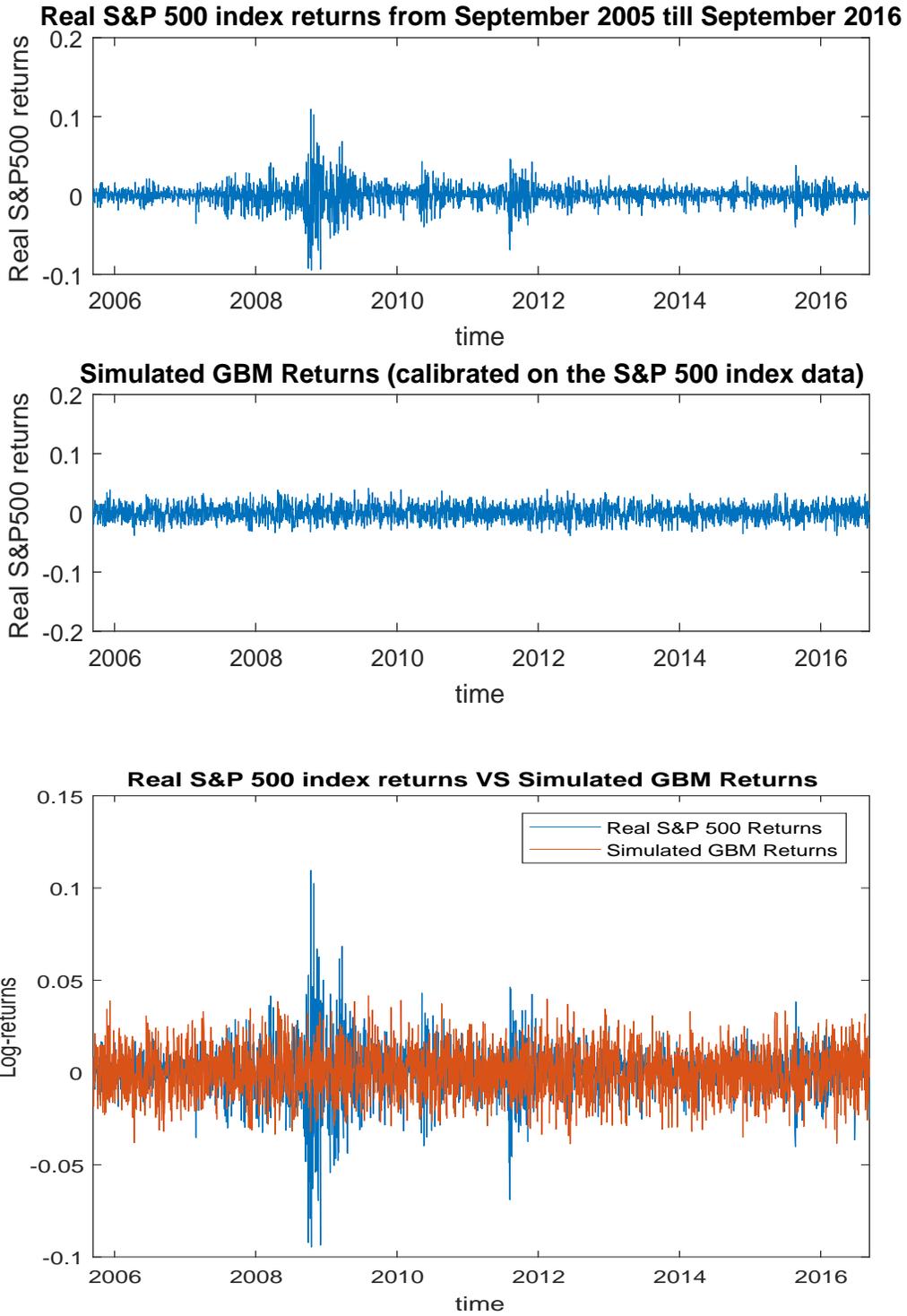


Figure 3.4: S & P 500 real returns against the simulated GBM returns.

under the GBM model. A key difficulty in measuring the market volatility is that it is not directly observable, it is a latent variable and therefore it must be inferred from the market price values.

Two most common techniques for calibrating the volatility parameter are historical and implied volatility. Here, we will focus only on the historical volatility, which is a retrospective measure that reflects how volatile the asset has been in the past. Different historical volatility type models can be used, from the most simple, like moving average (MA) and exponentially weighting moving average (EWMA) models to GARCH and stochastic volatility type models.

A simple moving average model for one-day volatility at time $t \geq 0$, is a sample standard deviation of all daily observed log-returns over a fixed time-window W preceding t . The past log-returns are denoted as $X_i, i = t - W \dots t - 1$, where time window W can be one year, month, day. MA model of volatility is given by the following formula,

$$\hat{\sigma}_t = \frac{1}{W-1} \sum_{i=1}^W (X_{t-i} - \bar{X})^2,$$

where X_i are observed log-returns and $\bar{X} = \frac{1}{W} \sum_{i=1}^W X_{t-i}$. Usually, the daily standard deviations are annualized by multiplying $\hat{\sigma}_t$ by square root of the number of trading days in one calendar year, which is typically 250 days a year. This annualized standard deviation is what is most commonly thought of the historical volatility.

Every day t , the above moving average formula for the spot volatility will be adjusted correspondingly, annualized and used in the Black-Scholes model to price an option on that particular day. Even though, under the GBM volatility should be constant, plots of the one-year and one-month MA annualized historical volatilities of the S&P 500 index, show that the estimated volatility parameter changes over time, see the figures 3.5a, 3.5b. This

effect can be observed in the previous plots of the regular S&P 500 returns, see figure 3.4. Moreover, one can even recognize a volatility clustering effect in both figures.

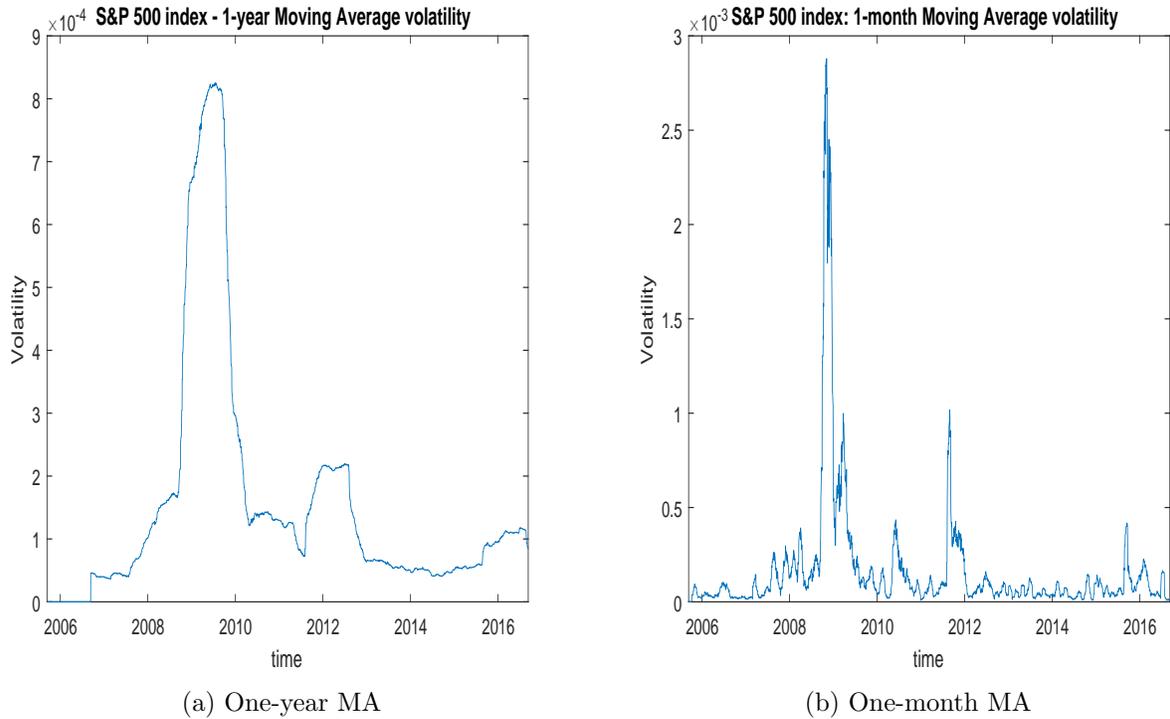


Figure 3.5: The historical volatility for the S&P 500 index returns (daily readings).

From the figures 3.5a, 3.5b the volatility seems to behave erratically over time which prompts one to consider non-deterministic models for this parameter. Moreover, one can also see there is a slight mean reversion effect, along with a periodic behavior of volatility, which can be better observed over the long term, see figure 3.5b. The phenomenon where periods with high variance of returns are followed by periods with low variance is called clustering effect and can be, intuitively, explained by the fact that large price variations are more likely to be followed large price variations, and vice versa. The clustering effect can be clearly observed in the figure 3.4. The fact that volatility exhibits clustering periodically in time makes it partially predictable, which may be of great help in the fields of risk management and option pricing.

3.3 Fractal Activity time geometric Brownian motion model with supOU as activity time

Since considerable investigation of the characteristics of the stock price returns show signs of non-Gaussianity and a very delicate dependence structure, new models were proposed that try to describe empirical log-returns more realistically. Most common problems that any of the more advanced models face, are complexity and intractability. These models become complicated either by introducing a large number of parameters to be estimated or because, simply their marginal densities are intractable, i.e. don't have a closed form solution.

Recall that in the previous chapters we mentioned that finite/infinite supOU processes have tractable distributions if suitable marginal distributions are chosen. Hence, to keep the tractability, for the calibration of the S&P 500 data, we suggested using Heyde's skewed FATGBM subordinator model where the activity time is constructed via finite superpositions of OU type processes. This model generalizes classical GBM model by introducing a time-changed version of the Brownian motion and most importantly subordinated process is chosen so that final model supports all of the desired features of the empirical log returns.

Remark 3.3.1. *It is important to notice that this approach proposed by Heyde [37] combines all of the desired properties: non-Gaussian distributional features, dependence structure of the log returns and tractable marginal distributions.*

There are various types of models that try to incorporate non-Gaussian distributions that have been proposed and investigated by many authors, see Prsetx, P.D, Sorensen, Rydberg. For example, approach that uses Levy processes instead of a Brownian motion in GBM model, proposed by Eberlein and Raible (1999) [19], allows larger class of possible marginal

distributions (*infinitely divisible distributions*) and it also introduces jumps in the stock price process. Then, already mentioned, Mandelbrot's Fractional Brownian Motion (FBM) subordinator model [53] and the Stochastic Volatility type models of Barndorff-Nielsen and Shephard [8].

Here, we focus more on models that try to describe volatility as a stochastic process. There are two approaches to building these type of models. One is modeling σ , the volatility parameter, directly as a stochastic process. The models of this type are commonly called *Stochastic Volatility type models*. The other way is through the subordination of a Brownian Motion (or a Levy process) and by treating time as a random activity time process. These are usually called subordinator or change of time models. Clearly, the model we suggested for the calibration is of the second type.

Moreover, different activity time type models have been proposed depending on how the activity time was constructed. In particular, use of Levy processes to model the activity time was suggested by Madan and Seneta (1990), Madan, Carr and Chang (1998), see [51, 52]. Authors, Bender and Marquardt (2009) [10] offered a construction of activity time via convolution between a Levy subordinator and a deterministic kernel. While all these constructions have many desirable features they do not specify the marginal distribution of the returns in a closed form which makes these models extremely hard to implement.

3.3.1 Fractal Activity time geometric Brownian motion process

This section introduces a *fractal activity time geometric Brownian motion* process, or FATGBM for short. FATGBM models, that belong in the group of subordinator models, were first introduced by Heyde (1999) [37], and see also Heyde and Liu (2001) [39]. The FATGBM model describes the stock price process $\{S_t, t \geq 0\}$, driven by standard Brownian

motion B_t and evaluated at a random activity time T_t , independent of B_t . The stock price process is given by the following expression:

$$\log S_t = \log S_0 + \mu t + \theta T_t + \sigma B(T_t), \quad (3.5)$$

where $\mu, \theta \in \mathbb{R}$, and $\sigma > 0$ are constants referred to as drift, asymmetry and volatility. The *activity time* process $\{T_t\}_{t \geq 0}$ is positive, nondecreasing with stationary (but not independent) increments, denoted as $\tau_t = T_t - T_{t-1}$, and it starts at zero a.s., $T_0 = 0$. In addition, processes B_t and T_t are \mathcal{F}_t -adapted.

Remark 3.3.2. *In the case when $T_t = t$ of the FATGBM model, equation (3.5) becomes classical GBM model, with corresponding log-returns being independent and identically distributed normal random variables, for any point in time t .*

Remark 3.3.3. *Note that the restricted case when $\theta = 0$ is frequently called a symmetric model. In which case the price process is of the form, $S_t = S_0 \exp\{\mu t + \sigma B(T_t)\}$.*

If certain conditions, see Kobayashi (2010) [46] are satisfied ⁹, the FATGBM process in equation (3.5) is a unique, strong solution to the time-change stochastic differential equation, given below:

$$\frac{dS_t}{S_t} = \mu dt + \left(\theta + \frac{1}{2}\sigma^2\right)dT_t + \sigma dB_{T_t} \quad (3.6)$$

The activity time process T_t has an attractive interpretation of time over which the market price evolves. In the finance literature, it is often associated with a trading volume or with the flow of new information. In other words, the more information is released into

⁹Since the SDE 3.6 is homogeneous linear SDE, beside the standard assumptions for the existence and uniqueness of the solution, one needs additional assumptions about the process T_t . In specific, the activity time is assumed to be positive, non-decreasing, with continuous sample paths and $T_0 = 0$

the market, or the more "frenzied" trading becomes, the faster the activity time flows, see Howison and Lamper [41].

3.3.2 Properties of log-returns in the FATGBM model

As with the Black-Scholes model, one is more interested in modeling log returns directly instead of the stock price process, hence the FATGBM log-returns are of the form,

$$X_t = \log S_t - \log S_{t-1} = \mu + \theta(T_t - T_{t-1}) + \sigma(B(T_t) - B(T_{t-1})) \quad (3.7)$$

$$\stackrel{\mathcal{D}}{=} \mu + \theta(\tau_t) + \sigma B(\tau_t) \quad (3.8)$$

$$\stackrel{\mathcal{D}}{=} \mu + \theta(\tau_t) + \sigma\sqrt{\tau_t}B(1), \quad (3.9)$$

provided that $\tau_t = T_t - T_{t-1}$.

The above log-returns can be simply expressed in terms of standard Normal errors,

$$X_t = \mu + \theta(\tau_t) + \sigma\sqrt{\tau_t}\xi(t), \quad t = 1, 2, \dots$$

where $\xi(t)$ is an iid sequence of standard Normal random variables, independent of the activity time τ_t , that is assumed by the model.

In the subsequent sections we will explain properties of the FATGBM log-returns and how they relate to the stylized facts of the empirical stock returns. We will see how moments, skewness, kurtosis and dependence structure of the returns behaves as a result of our assumptions about the parameters θ , μ , and $E\tau_t$. In addition, the fact that the variance of the FATGBM log-returns is heteroscedastic will also be explored.

3.3.2.1 Moments, skewness and kurtosis

Assuming the finiteness of at least fourth order moment of τ_t , the expressions for the moments of the FATGBM log-returns of up to the fourth order are given below,

$$EX_t = \mu + \theta E\tau_t, \quad (3.10)$$

$$E(X_t - EX_t)^2 = \sigma^2 E\tau_t + \theta^2 M_2, \quad (3.11)$$

$$E(X_t - EX_t)^3 = 3\theta\sigma^2 M_2 + \theta^3 M_3, \quad (3.12)$$

$$E(X_t - EX_t)^4 = 3\sigma^4(M_2 + (E\tau_t)^2) + 6\sigma^2\theta^2(E\tau_t M_2 + M_3) + \theta^4 M_4, \quad (3.13)$$

where $M_i = E(\tau_t - E\tau_t)^i$, $i = 2, 3, 4$. Note that Seneta (2004) [65] assumed that $E\tau_t = 1$ in his paper, and we will exploit this assumption when calibrating the Variance-Gamma parameters, see the section 3.4. In the above expressions one can see how different assumptions about θ , μ , and $E\tau_t$ can affect the properties of the returns.

The coefficients of skeweness (asymmetry) and of kurtosis (fatness of tails) are defined respectively by,

$$\gamma_1 = \frac{E[(X_t - \mu)^3]}{(Var[X])^{3/2}} = \frac{3\theta\sigma^2 M_2 + \theta^3 M_3}{(\sigma^2 E\tau_t + \theta^2 M_2)^{3/2}}. \quad (3.14)$$

$$\gamma_2 = \frac{E[(X - \mu)^4]}{(var[X])^2} = \frac{3\sigma^4(M_2 + (E\tau_t)^2) + 6\sigma^2\theta^2(E\tau_t M_2 + M_3) + \theta^4 M_4}{(\sigma^2 E\tau_t + \theta^2 M_2)^2}. \quad (3.15)$$

Notice that, symmetric log-returns¹⁰ ($\gamma_1 = 0$) correspond to the case when $\theta = 0$, while skewed returns correspond to the case when $\theta \neq 0$. In addition, in the case of the symmetric

¹⁰Symmetric about the mean μ .

log-returns: $\theta = 0$, the kurtosis is:

$$\gamma_2 = \frac{3(M_2 + (E[\tau_t])^2)}{(E[\tau_t])^2},$$

and it is still ≥ 3 . This allows for the returns to have heavier tails than normal distribution even in the case of symmetric returns.

From the above equations (3.10) one can see that even when $\theta = 0$, the variance can be heteroskedastic, i.e. one can pick $E(\tau_t)$ to be time-dependent.

3.3.2.2 The dependence structure in the FATGBM model

In this section we will show that the FATGBM subordinator model is capable of reproducing the dependence properties of the simple, squared and absolute empirical log-returns, as contrast to the Black-Scholes model.

First we'll explain some of the properties of covariance of the log-returns in terms of the activity time process $\{\tau_t\}$. If finiteness of the moments is assumed accordingly, then for $k \geq 1$ one gets:

$$\begin{aligned} cov(X_t, X_{t+k}) &= cov(\theta\tau_t + \sigma\sqrt{\tau_t}B_1(1), \theta\tau_{t+k} + \sigma\sqrt{\tau_{t+k}}B_2(1)) \\ &= \theta^2(E\tau_t\tau_{t+k} - E\tau_tE\tau_{t+k}) \\ &= \theta^2cov(\tau_t, \tau_{t+k}), \end{aligned}$$

where B_1 and B_2 are independent copies of Brownian motion.

Now, let us see how the dependence structure changes as we pick different parameters. For example when $\theta = 0$ then $cov(X_t, X_{t+k}) = 0$, i.e. the log-returns are uncorrelated or

independent in Gaussian case. In addition, from the above expression we can see that if the increments of the activity time τ_t in the FATGBM model are independent or when activity time is deterministic ($\tau_t = t$), the log-returns turn out to be uncorrelated as well.

Remark 3.3.4. *Note, when $\tau_t = t$, FATGBM behaves like a classical Geometric Brownian motion with i.i.d. Gaussian log-returns.*

Now, let us explore the dependence structure of the squared and log returns in a general case, when $\mu = \theta \neq 0$,

$$\begin{aligned} \text{cov}(X_t^2, X_{t+k}^2) &= \text{cov}((\mu + \theta\tau_t + \sigma\sqrt{\tau_t}B_1(1))^2, (\mu + \theta\tau_{t+k} + \sigma\sqrt{\tau_{t+k}}B_2(1))^2) \\ &= (\sigma^4 + 4\theta^2\mu^2 + 4\theta\mu\sigma^2)\text{cov}(\tau_t, \tau_{t+k}) + \theta^4\text{cov}(\tau_t^2, \tau_{t+k}^2) + \\ &\quad + (\theta^2\sigma^2 + 2\theta^3\mu)(\text{cov}(\tau_t^2, \tau_{t+k}) + \text{cov}(\tau_t, \tau_{t+k}^2)), \\ &\stackrel{\theta=0}{=} \sigma^4\text{cov}(\tau_t, \tau_{t+k}). \end{aligned}$$

From the above expression for the covariance function, one can conclude that, even in the symmetric case ($\theta = 0$) which implies uncorrelated log-returns ($\text{cov}(X_t, X_{t+k}) = 0$), the covariance of the squared returns is still persistent and depends only on the covariance of the subordinator, $\text{cov}(X_t^2, X_{t+k}^2) = \sigma^4\text{cov}(\tau_t, \tau_{t+k})$ for $\theta = 0$.

Remark 3.3.5. *One can conclude that dependence properties of the activity time process $\{\tau_t\}$ directly imply the properties of the squared log-returns $\{X_t^2\}$.*

In the next paragraph we will explore the covariance structure for the absolute log-returns

when $\mu = \theta = 0$, see the details below:

$$\begin{aligned}
\text{cov}(|X_t|, |X_{t+k}|) &= \text{cov}(\sigma\sqrt{\tau_t} |B_1(1)|, \sigma\sqrt{\tau_{t+k}} |B_2(1)|) \\
&= \sigma^2 E(|B_1(1)|) E(|B_2(1)|) \text{cov}(\sqrt{\tau_t}, \sqrt{\tau_{t+k}}) \\
&= \frac{2}{\pi} \sigma^2 \text{cov}(\sqrt{\tau_t}, \sqrt{\tau_{t+k}})
\end{aligned}$$

Similarly as for quadratic returns, we see that assuming $\theta = 0$ implies uncorrelated log-returns, while absolute log-returns are still directly dependent on the covariance structure of the activity time process.

3.3.2.3 Stochastic volatility - conditional heteroscedasticity

In this section we will see that, with the appropriate choice of the activity time τ_t , the FATGBM log-returns display the property of conditional heteroscedasticity, that is a time dependent conditional variance. To be more precise, if one chooses the activity time τ_t that has specific dependence structure, the heteroscedasticity simply follows.

Let $\mathcal{F}_t = \sigma(\{B(u), u \leq T_t\}, \{T_u, u \leq t\})$ be the sigma-algebra of information available up to time t , then the conditional variance is of the form:

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = E(X_t^2 | \mathcal{F}_{t-1}) - E(X_t | \mathcal{F}_{t-1})^2 \quad (3.16)$$

$$= \theta^2 E(\tau_t^2 | \mathcal{F}_{t-1}) + (\sigma^2 + 2\mu\theta) E(\tau_t | \mathcal{F}_{t-1}) - \quad (3.17)$$

$$- (2\mu\theta E(\tau_t | \mathcal{F}_{t-1}) + \theta^2 E(\tau_t | \mathcal{F}_{t-1})^2) \quad (3.18)$$

$$= \theta^2 \text{Var}(\tau_t | \mathcal{F}_{t-1}) + \sigma^2 E(\tau_t | \mathcal{F}_{t-1}). \quad (3.19)$$

In the case of restricted model, when $\theta = 0$, the above expression reduces to: $\text{Var}(X_t | \mathcal{F}_{t-1}) =$

$\sigma^2 E(\tau_t | \mathcal{F}_{t-1})$, see Heyde and Liu (2001) [39].

Remark 3.3.6. *What is interesting here, is when one does comparison of the time-dependent conditional variance of the FATGBM model to the constant conditional variance of the Black-Scholes model, one can conclude that the first model offers considerably more flexibility. That is, one can model conditional variance or the "volatility") as they are pleased: either as a random function of time or only a function of time. Note that, even in the reduced/symmetric FATGBM model, where $\theta = 0$, even though the returns are uncorrelated, one can still model the conditional variance.*

Remark 3.3.7. *Note, it is common in industry and finance to refer to the parameter σ , which in the Black-Scholes model simply represents a standard deviation of the returns, as the volatility parameter. Even though it may be confusing at times, the volatility would be referred to as random or stochastic, throughout the applied literature. Even though the context is understood, it is not correct mathematically, since expectation cannot be random. Hence here, we use the term volatility for conditional variance only - for which we know it can be random.*

3.3.3 Distribution theory

Since the FATGBM log returns are of the form:

$$X_t \stackrel{\mathcal{D}}{=} \mu + \theta\tau_t + \sigma\sqrt{\tau_t}B(1),$$

then the conditional distribution of the returns: X_t , given the variance $\tau_t = V$ is normal with mean μ and variance σ^2V when $\theta = 0$, while when $\theta \neq 0$, the mean is $\mu + \theta V$ and the

variance is the same. See below,

$$X_t | \tau_t \stackrel{\mathcal{D}}{=} \mathcal{N}(\mu + \theta\tau_t, \sigma^2\tau_t).$$

Since the distribution of the log-returns X_t is conditionally normal, they belong to the class of normal mixed distributions or so-called generalized hyperbolic distributions.

3.3.3.1 Gamma distribution of the unit increments of the activity time process

Suppose that the activity process τ_t follows a Gamma distribution, then the distribution of the log-returns X_t is going to be a Variance Gamma (VG) distribution. This is actually the reason why this distribution is called Variance-Gamma, because its variance is Gamma distribution: $VG(\mu, \theta, \sigma^2, \alpha, \beta) | V \sim \mathcal{N}(\mu, \sigma^2 V)$, where V has Gamma distribution.

Although VG distribution was explained in the section 1.5.3.3 in detail, here we will introduce a different parametrization of the VG distribution, the one that is later on going to be used for calibration of the returns.

As a reminder, if τ_t is distributed as $\Gamma(\alpha, \beta)$, where $\alpha, \beta > 0$, then its density and the characteristic function are given by the expressions:

$$f_{\Gamma}(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0, \quad \phi_{\Gamma}(u) = \left(1 - \frac{iu}{\beta}\right)^{-\alpha}, \quad u \in \mathbb{R}. \quad (3.20)$$

The moments of the τ_t -marginal Gamma distribution are

$$E\tau_t = \frac{\alpha}{\beta}, \quad M_2 = \frac{\alpha}{\beta^2}, \quad M_3 = \frac{2\alpha}{\beta^3}, \quad M_4 = \frac{3\alpha(\alpha+2)}{\beta^4}. \quad (3.21)$$

where $M_i = E(X_t - EX_t)^i$.

The specific choice of a Gamma distribution as a marginal distribution of the activity time τ_t coupled with the FATGBM model setup results in log-returns X_t having the marginal skew Variance Gamma distribution, see [51], with the density given by,

$$f_{VG}(x) = \sqrt{\frac{2}{\pi}} \frac{\beta^\alpha e^{-\frac{(x-\mu)\theta}{\sigma^2}}}{\sigma \Gamma(\alpha)} \left(\frac{|x-\mu|}{\sqrt{\theta^2 + 2\beta\sigma^2}} \right)^{\alpha-\frac{1}{2}} K_{\alpha-\frac{1}{2}} \left(\frac{|x-\mu|\sqrt{\theta^2 + 2\beta\sigma^2}}{\sigma^2} \right), \quad (3.22)$$

where $K_\eta(\omega)$ is a modified Bessel function of the third kind with index η ¹¹. For $\eta \in \mathbb{R}$ and $\omega > 0$ is given by,

$$K_\eta(\omega) = \frac{1}{2} \int_0^\infty z^{\eta-1} e^{-\omega/2(z+1/z)} dz.$$

Note that the behavior of the tail of the VG distribution is different from that of a Normal distribution. VG is considered to belong to a class of semi-heavy tails distributions that decay slower than the Normal and faster than the heavier tail distributions. The probability of its tail is a product of both a power and an exponential function. Namely, if X_t has VG distribution, then as $x \rightarrow \infty$, the tail probability is given by¹²:

$$P(|X_t| > x) \sim \text{const}(\alpha, \beta, \sigma) x^{\alpha-1} e^{-x\sqrt{2\beta/\sigma^2}}.$$

The characteristic function of the VG log-returns X_t is given by:

$$\phi_{VG}(u) = e^{i\mu u} \left(1 - \frac{i\theta u}{\beta} + \frac{1}{2\beta} \sigma^2 u^2 \right)^{-\alpha}, \quad u \in \mathbb{R}. \quad (3.23)$$

¹¹Note that K_η is referred to as a modified Bessel function of the second kind in some texts.

¹²The notation: $f(x) \sim g(x)$ means that $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$

In the symmetric case ($\theta = 0$) the expression for the characteristic function is of the form:

$$\phi_{VG}(u) = e^{i\mu u} \left(1 + \frac{1}{2\beta} \sigma^2 u^2\right)^{-\alpha}. \quad (3.24)$$

Remark 3.3.8. *The distribution defined in (3.22) and in (3.23) we will denote as $VG(\mu, \theta, \sigma^2, \alpha, \beta)$.*

Specializing the equations for the central moments of the FATGBM model, in 3.10, by taking that τ_t has Gamma marginal distribution, the central moments of the log-returns X_t are given by:

$$\begin{aligned} EX_t &= \mu + \theta\alpha/\beta, \\ E(X_t - EX_t)^2 &= \sigma^2 \frac{\alpha}{\beta} + \frac{\theta^2 \alpha}{\beta^2}, \\ E(X_t - EX_t)^3 &= \frac{3\theta\sigma^2 \alpha}{\beta^2} + \frac{2\theta^3 \alpha}{\beta^3}, \\ E(X_t - EX_t)^4 &= \frac{3\sigma^4 \alpha(\alpha + 1)}{\beta^2} + \frac{6\sigma^2 \theta^2 \alpha(\alpha + 2)}{\beta^3} + \frac{3\theta^4 \alpha(\alpha + 2)}{\beta^4}. \end{aligned}$$

Remark 3.3.9. *Note that, till now we defined only the univariate characteristics of the activity time process $\{\tau_t\}$, i.e. we fully defined the marginal distributions. We haven't yet introduced any time-dependence. This will be done through the particular choice of the covariance structure for the activity time process. This step is necessary in order to get the time-dependent conditional variance and the suitable dependence structure of the returns.*

Remark 3.3.10. *Notice here, that if the appropriate activity time process τ_t along with the convenient dependence structure is chosen, then all the empirical properties of the stock log-returns can be modeled accordingly.*

3.4 Data calibration

In this section, we would like to validate the approach of using a version of the FATGBM process to model the stock price evolution. The marginal distribution for the activity time is chosen to be a Gamma distribution. This, in turn implies that the marginal distribution of the log-returns is Variance-Gamma. We will showcase the strengths of this approach through series of plots and tables presenting the stylized facts of the log-returns discussed previously, while comparing it to the classical Geometric Brownian Motion (GBM) model.

The objective of this calibration exercise is to estimate the parameters of the Variance-Gamma (V-G) distribution whilst trying to keep the assumption of the dependent log-returns. While in the option pricing literature there are many different ways to approach model calibration problems, a frequently used method is to minimize squared residual errors between the underlying model and data at hand. Although, this approach is effective and it does not require many assumptions about the underlying distribution of the returns, it is fairly more complex to implement than the classical maximum likelihood (ML) or method of moments (MOM). Due to the fact that the returns in our model are dependent, so the regular MLE is not appropriate, we have decided to take a simpler route and to derive parameter estimates for the VG distribution using the method of moments technique. Importantly, this technique, if relying on the stationarity of increments and ergodicity property gives consistent estimators. As an exploration, we used the version of the MLE technique called pseudo-likelihood for the comparison purposes. It is explained below and in the results section in more details.

We will digress here for a moment to explain briefly and compare strengths and weakness of both, the MOM estimation technique and the Maximum Likelihood (ML) approach. It

is well known that, even though the MOM is relatively simple and reliable technique which, under very mild assumptions gives consistent estimators, these estimators are often biased. In addition, when working with small samples, the estimates derived by the MOM can fall outside of the parameter space. On the contrary, when estimating parameters from a known family of distributions, the maximum likelihood estimate is very efficient. Now, there are some situations when using the method-of-moments technique is more appropriate, for example, when likelihood equations are intractable or when one does not want to make any distributional assumptions. In our case, we want to keep the dependent observations assumption and use the estimation techniques that take this into account, hence the regular MLE would not be appropriate.

In addition, we will briefly dabble with pseudo-likelihood (PL) method, where the PL is said to be an approximation of the joint likelihood of a set of observed data. Pseudo-likelihood and Quasi-likelihood are very useful techniques to use in the situation where the full likelihood is difficult to work with or unknown. In our case, there is dependency in the data hence the likelihood is unknown. The pseudo-likelihood approach maximizes the product of marginal densities, as opposed to the full joint density. If observations are independent then these two coincide. In certain cases one can show that the pseudo-likelihood estimators are even consistent, see Cox and Reid (2004) [17]. We will attempt to use the MLE technique along with the unrealistic assumption of the independent increments, which is equivalent, in a way, to using the pseudo-likelihood method. That is, we are maximizing the product of the marginal distributions since we assumed "the observations are independent", which is nothing but the product approximation of the joint likelihood and the weaker the dependence in the data the two will be closer. This means that the MLE parameter estimates may not be that far from the true values of the parameters, especially if dependence is not as

pronounced. As we shall see in the results section 3.4.2, our calibration exercise showed that MLE estimators seemed to fit better than MOM which means that PL is relatively close to the joint likelihood of the data. In addition parameter θ seemed to be close to zero which intuitively implies that dependencies are not that strong.

3.4.1 VG parameter estimation using method of moments

As mentioned, we are considering FATGBM returns, see (3.5), (3.7), given by the expression,

$$X_t = \log S_t - \log S_{t-1} = \mu + \theta(\tau_t) + \sigma\sqrt{\tau_t}B(1), \quad (3.25)$$

where μ, θ , and $\sigma > 0$ are real constants and $\tau_t = T_t - T_{t-1}$ is an activity time process: positive, increasing, with stationary increments, and independent of Brownian motion.

We assume that increments of the activity time follow a marginal gamma distribution $\Gamma(\alpha, \beta)$. In addition, since $E\tau_t < \infty$, without loss of generality we can assume that

$$E\tau_t = 1, \quad (3.26)$$

which simply scales down the expected activity time change to be in units of one, over the unit calendar time. The scaling constant gets absorbed into θ and σ .

Since the $E\tau_t = 1$ the moments of the Gamma distribution, given in (3.27) are the following,

$$E\tau_t = 1, \quad M_2 = \frac{1}{\alpha}, \quad M_3 = \frac{2}{\alpha^2}, \quad M_4 = \frac{3(\alpha + 2)}{\alpha^3}. \quad (3.27)$$

Given the above assumptions, when increments of activity time follow a marginal gamma distribution $\Gamma(\alpha, \alpha)$, the returns follow a marginal (skew) VG distribution, denoted as

$VG(\mu, \theta, \sigma^2, \alpha)$ and given by,

$$f_{VG}(x) = \sqrt{\frac{2}{\pi}} \frac{\alpha^\alpha e^{-\frac{(x-\mu)\theta}{\sigma^2}}}{\sigma \Gamma(\alpha)} \left(\frac{|x-\mu|}{\sqrt{\theta^2 + \alpha\sigma^2}} \right)^{\alpha-\frac{1}{2}} K_{\alpha-\frac{1}{2}} \left(\frac{|x-\mu|\sqrt{\theta^2 + \alpha\sigma^2}}{\sigma^2} \right) \quad (3.28)$$

Considering the above $VG(\mu, \theta, \sigma^2, \alpha)$ distribution and given the equations (3.10), we obtain the following for its centralized moments,

$$EX_t = \mu + \theta, \quad (3.29)$$

$$E(X_t - EX_t)^2 = \sigma^2 + \theta^2 M_2, \quad (3.30)$$

$$E(X_t - EX_t)^3 = 3\theta\sigma^2 M_2 + \theta^3 M_3, \quad (3.31)$$

$$E(X_t - EX_t)^4 = 3\sigma^4(M_2 + 1) + 6\sigma^2\theta^2(M_2 + M_3) + \theta^4 M_4, \quad (3.32)$$

Consequently, after substituting the moment M_1, \dots, M_4 in the above equations, we obtain the following equations for the central moments,

$$EX_t = \mu + \theta, \quad (3.33)$$

$$E(X_t - EX_t)^2 = \sigma^2 + \frac{\theta^2}{\alpha}, \quad (3.34)$$

$$E(X_t - EX_t)^3 = \frac{3\theta\sigma^2}{\alpha} + \frac{2\theta^3}{\alpha^2}, \quad (3.35)$$

$$E(X_t - EX_t)^4 = 3\sigma^4 \left(1 + \frac{1}{\alpha} \right) + \frac{6\sigma^2\theta^2}{\alpha} \left(1 + \frac{2}{\alpha} \right) + \frac{3\theta^4}{\alpha^2} \left(1 + \frac{1}{\alpha} \right) \quad (3.36)$$

From the above equations, we can obtain expressions for the coefficients of skewness and of kurtosis. If we assume that $\theta \approx 0$, we can ignore all terms higher than $\theta^2, \theta^3, \theta^4 \dots$ in the

above equations (3.33), and so we get

$$\mu_1 = EX_t = \mu + \theta, \quad (3.37)$$

$$\mu_2 = E(X_t - EX_t)^2 \approx \sigma^2, \quad (3.38)$$

$$\mu_3 = E(X_t - EX_t)^3 \approx \frac{3\theta\sigma^2}{\alpha} \quad (3.39)$$

$$\mu_4 = E(X_t - EX_t)^4 \approx 3\sigma^4 \left(1 + \frac{1}{\alpha}\right) \quad (3.40)$$

If we denote the sample (central) moments as, $\hat{\mu}_k^{(N)} = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X}^{(N)})^k$, for $k = 1, 2, 3, 4$, and where $\bar{X}^{(N)} = \hat{\mu}_1^{(N)} = \frac{1}{N} \sum_{i=1}^N X_i$ is its sample average. Now by equating the theoretical moments to the sample moments, we get

$$\begin{aligned} \hat{\mu}_1 &= \mu + \hat{\theta}, \\ \hat{\mu}_2 &= \hat{\sigma}^2, \\ \hat{\mu}_3 &= \frac{3\hat{\theta}\hat{\sigma}^2}{\hat{\alpha}} \\ \hat{\mu}_4 &= 3\hat{\sigma}^4 \left(1 + \frac{1}{\hat{\alpha}}\right), \end{aligned}$$

From now on, in order to simplify the presentation of the results, instead of the α -parameter we'll be using $\nu = 1/\alpha$, in which case the MOM estimators are given by the following set of formulas,

$$\hat{\mu} = \hat{\mu}_1 - \hat{\theta}, \quad \hat{\sigma}^2 = \hat{\mu}_2, \quad \hat{\theta} = \frac{\hat{\alpha}\hat{\mu}_3}{3\hat{\sigma}^2} \left(= \frac{\hat{\mu}_3}{3\hat{\sigma}^2\hat{\nu}} \right), \quad \hat{\alpha} = \frac{1}{\left(\frac{\hat{\mu}_4}{3\hat{\sigma}^4} - 1\right)} \left(\hat{\nu} = \frac{\hat{\mu}_4}{3\hat{\sigma}^4} - 1 \right) \quad (3.41)$$

Remark 3.4.1. *Note, if the $\hat{\theta}$ in the true sample is small then the full set of equations (3.33) will be very closely satisfied by the MOM estimators $\hat{\mu}, \hat{\theta}, \hat{\sigma}$, and $\hat{\alpha} (= 1/\hat{\nu})$.*

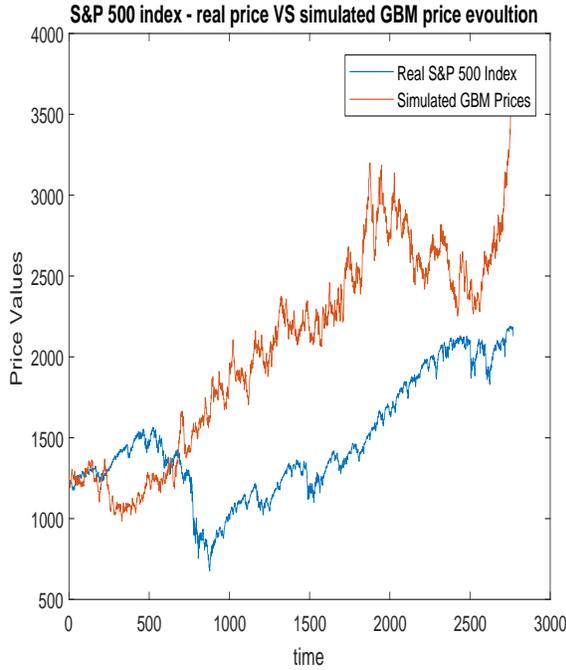
Remark 3.4.2. *Ideally we would have liked to estimate full set of parameters for VG distribution $VG(\mu, \theta, \sigma^2, \alpha, \beta)$, but due to the fact that solving a system of large number of nonlinear equations can be challenging at best, we are making some small simplifications in order to obtain meaningful results. Note that, there is no need for the independence of increments assumption in order to estimate the above parameters.*

Remark 3.4.3. *Notice here, that even though we have assumed that θ is small, we have avoided to take $\theta = 0$ from start, in which case we would be working with the symmetric VG distribution, and the model that would have the skewness parameter equal to zero and uncorrelated returns. To be more specific, setting the parameter $\theta = 0$ in (3.7) tells us that the returns X_t are distributed symmetrically about their mean μ with the variance σ^2 . Actually, in the financial context, the symmetric model assumption is not that far away from the truth, since symmetry is usually the case when working with the empirical data.*

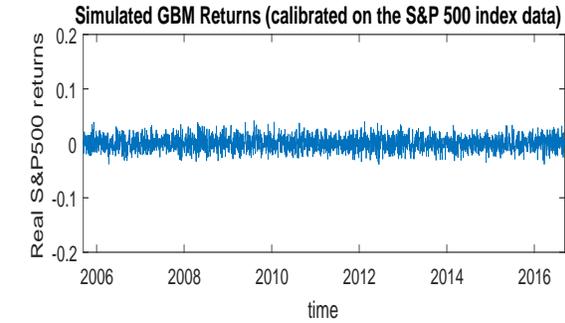
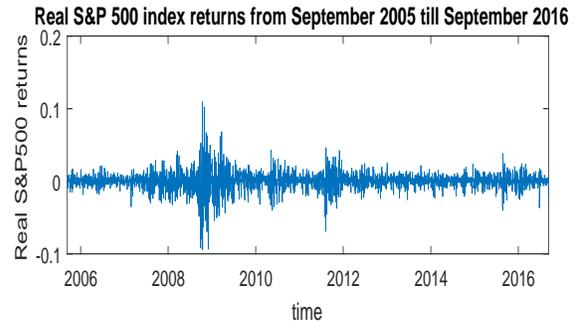
3.4.2 VG parameter fitting - results

Data used for this calibration exercise is the same as the one used in previous chapters. It consists of daily adjusted closing prices of the Standard and Poor's 500 Index (S&P 500), in the eleven year period, from the 9th of September 2005 till the 9th of September 2016. The total number of log daily prices was $N=2770$. The figure 3.6a shows the S&P500 price evolution in the time period from 2005-2016, and in the figure 3.6b the log-returns from the same period are presented.

We tried couple of different ways to calibrate the parameters, where two of the method of moments approaches are explained in first and second, while the MLE approach is given in the third example.



(a) Empirical price path for S&P500 index



(b) Empirical log-return sequence for S&P500 index

Example 3.4.1. *In this example we exactly followed the estimation approach suggested in the Seneta's (2004) paper [65] and what we explained previously, i.e. we assumed that $\theta \approx 0$, so the terms $\theta^2, \theta^3, \theta^4 \dots$ are ignored. We used the MOM estimates provided in (3.37), and results are the following:*

$$\hat{\mu} = 1.9458 \times 10^{-4},$$

$$\hat{\sigma} = 0.0127,$$

$$\hat{\theta} = -1.0477 \times 10^{-11},$$

$$\hat{\nu} = 3.5266, (\hat{\beta} = 0.2836)$$

First, notice that $\hat{\theta}$ is very small, hence $\theta^2, \theta^3, \theta^4 \dots$ are nearly negligible, which in turn implies the above estimates can be considered to be relatively close to the moment estimates. The

resulting VG density is shown in the figure 3.9a. To compare this method to the other two approaches in the examples below, please see the figure 3.6.

Example 3.4.2. In this example we kept the terms θ, θ^2 , while ignoring $\theta^3, \theta^4 \dots$ in the equations (3.33) and solved the system of non-linear equations in the MATLAB software.

The estimates we got as a solution are given below:

$$\begin{aligned}\hat{\mu} &= 0.0005978, \\ \hat{\sigma} &= 0.0127071, \\ \hat{\theta} &= -0.0004032197, \\ \hat{\nu} &= 3.50211.\end{aligned}$$

Here too $\hat{\theta}$ is relatively small, larger than in the previous example. Note that while σ and ν are unchanged, the estimates for μ and theta are considerably affected in this compared to the previous example, which casts doubt into the assumption that $\theta = 0$. The VG density with the estimated parameters is shown in the figure 3.9b. To compare this method to the other two approaches see the figure 3.6.

Example 3.4.3. In the final example we fitted the parameters of the VG distribution by maximizing the corresponding likelihood along with an additional assumption of independence of increments. For the initial values of the algorithm we used the MOM estimates which is a standard procedure. We interpreted the resulting estimates as quasi-likelihood type, or in our case better to say pseudo-likelihood type. Since pseudo-likelihood is simply an approximation of the joint likelihood, any differences between these and the MOM estimates we can explain with possibly persistent dependencies in our data or simply that the independent increment

assumption wasn't appropriate to start with. The PLE estimates are given below:

$$\hat{\mu} = 0.0008363,$$

$$\hat{\sigma} = 0.0119243,$$

$$\hat{\theta} = -0.0006523,$$

$$\hat{\nu} = 1.3369743.$$

In general, if we compare the pseudo-likelihood and MOM estimates from the previous example, they seem to be at least of the same order, where σ , θ seem close while ν and μ seem to differ. If we look at the plot comparing all three fitting examples against the empirical log-returns 3.6, we can see that MLE estimates provided a far better fit than the both MOM estimates. The possible explanation is that under the circumstances the approximated likelihood is very close to the real joint likelihood, see the conclusion section below for more detailed explanation. In the figure 3.9c we plotted the calibrated VG densities for all three methods. Observe that the MLE fitted plot seems the least peaked. Finally, since MLE performed the best, we included other plots related to the MLE estimation, see the figure 3.10a, 3.10b, 3.10c, 3.10d.

To illustrate the fit we compared the empirical returns, VG and inferior Gaussian distribution in the figure 3.7. One can clearly see that both VG is better fitting the data than normal distribution. So, in the figure 3.7, in addition to higher peaks than Gaussian distribution one can observe that typical returns data has heavier tails also.

In order to illustrate the tail fit, we plotted the tails for the VG distribution, fitted by MOM and MLE, alongside the NIG and Normal distribution, see the figures 3.8a, 3.8b. Conclusions are similar as previously stated, i.e. that MLE method seems to does metter

than MOM in general, and that both VG and NIG seem to be more appropriate for the data.

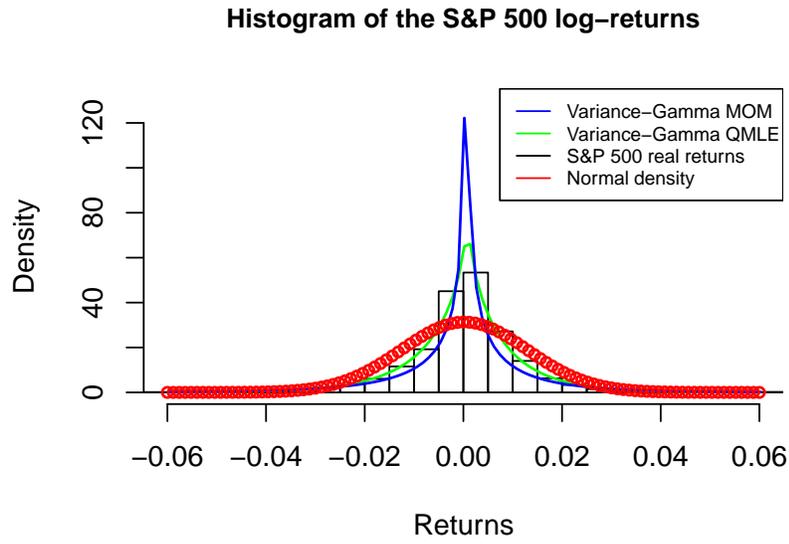


Figure 3.6: Fitting of VG distribution parameter to the S&P 500 Index returns. The plot shows the histogram of the empirical returns along with three differently calibrated VG densities. The methods presented are VG MOM method (example 2), VG MLE method (example3) and the Normal curve.

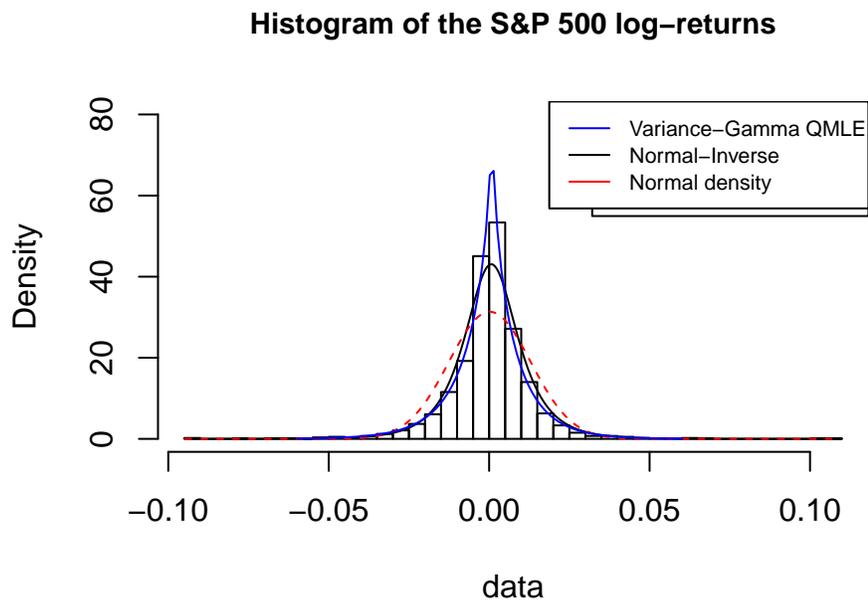


Figure 3.7: Fitting VG, NIG and Normal distribution parameters to the S&P 500 Index returns. The plot shows the histogram of the empirical returns along with three different density.

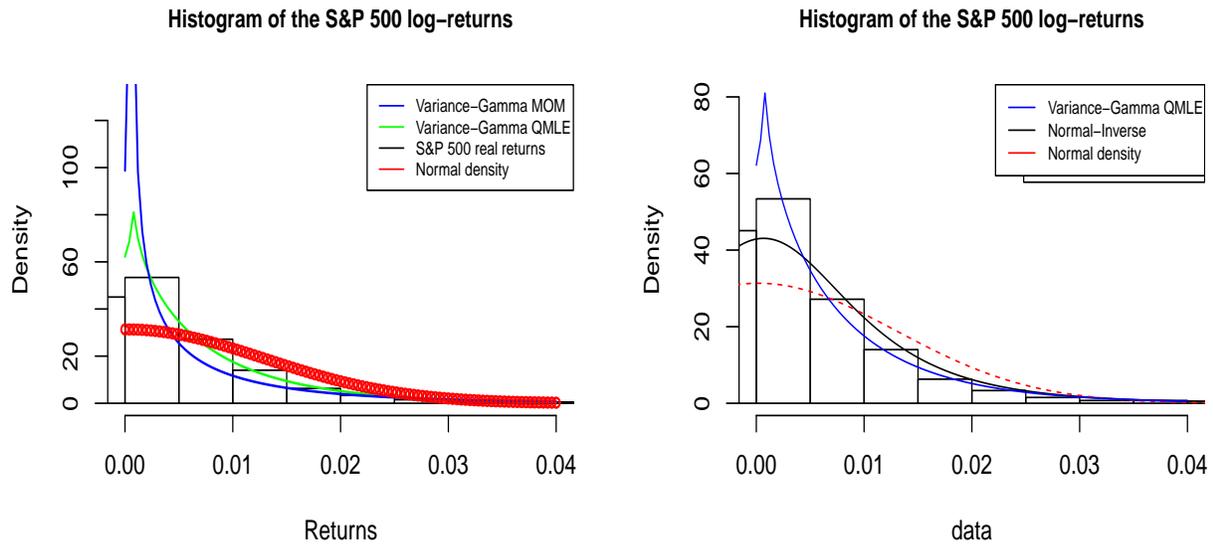


Figure 3.8: Tail probabilities of the Fitted VG (MOM, MLE), NIG and Normal calibrated on the S&P 500 returns. The plot shows how three distributions fit the empirical distribution.

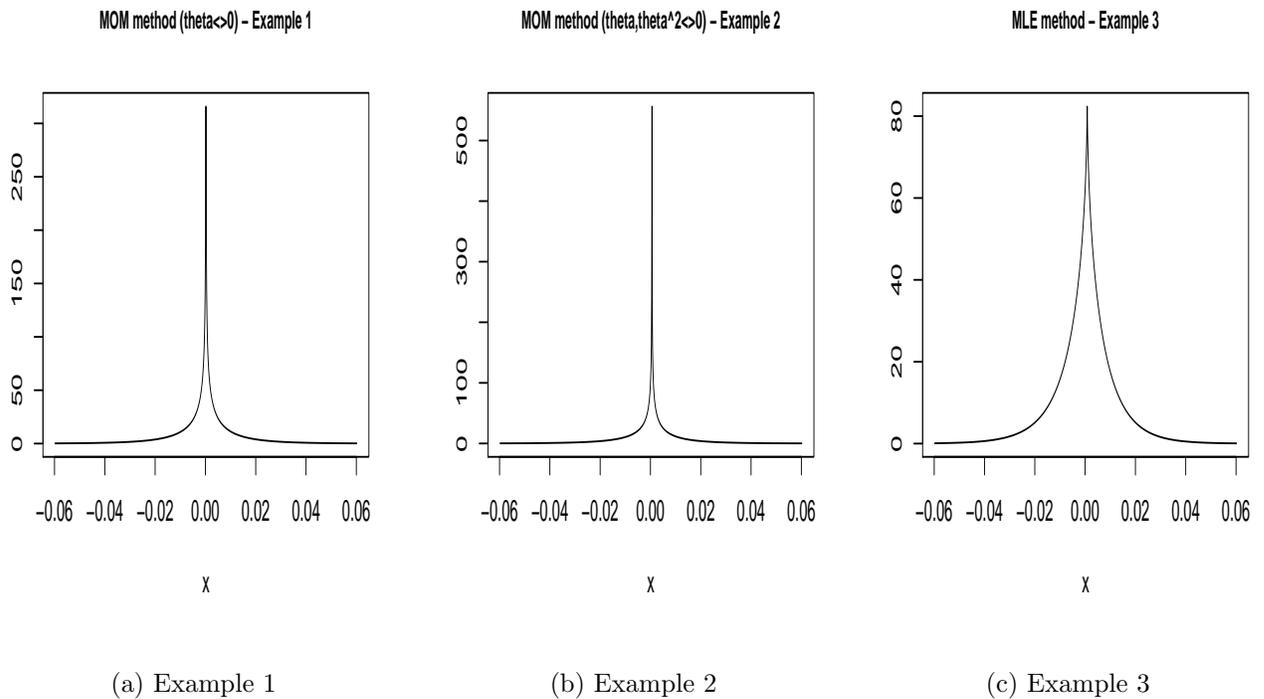
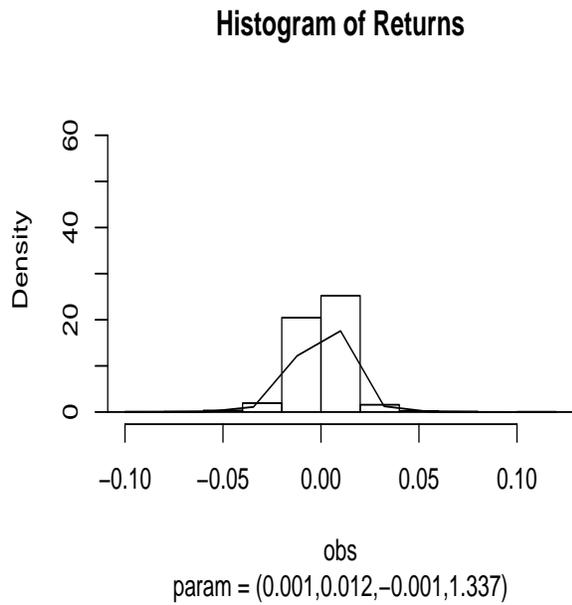
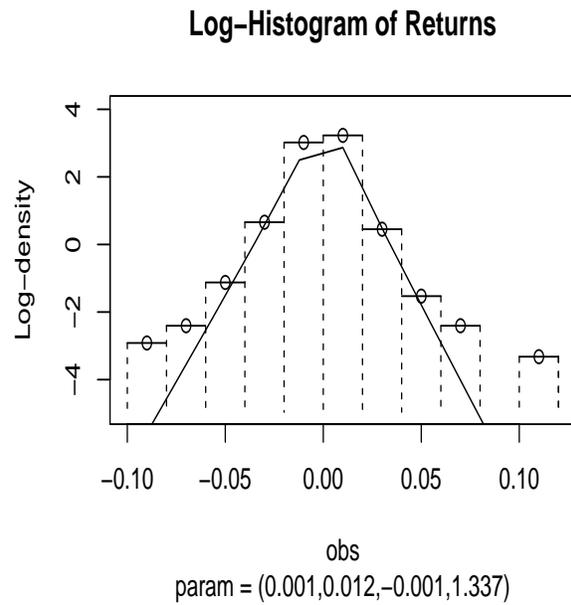


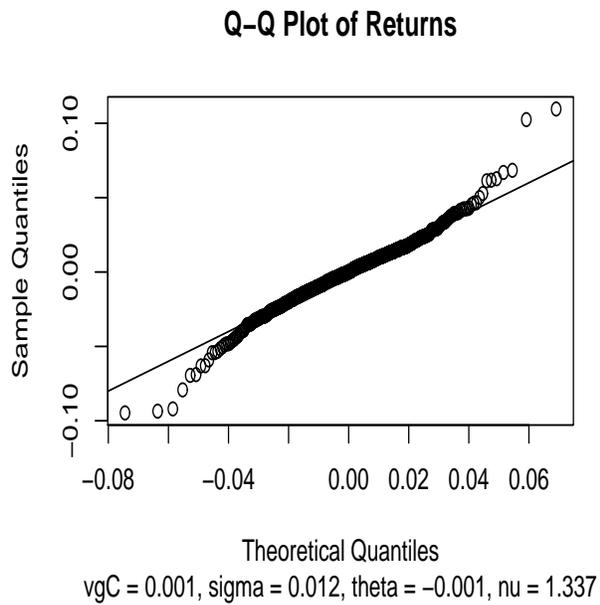
Figure 3.9: Fitting of VG distribution parameters to the S&P 500 Index returns. The plot shows three different examples explaining different methods of estimation.



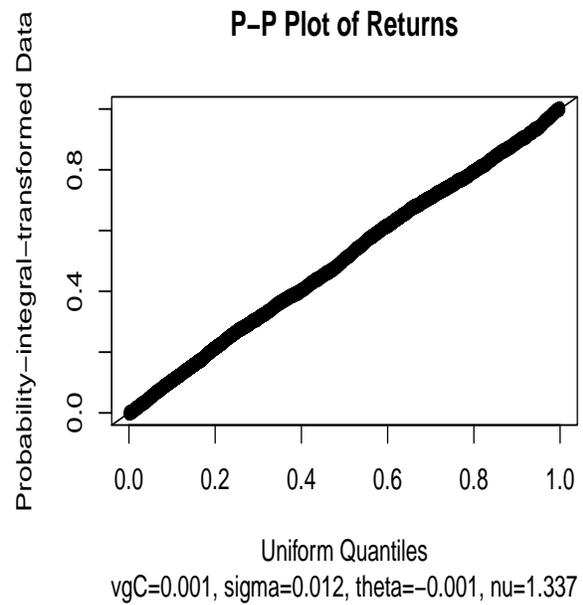
(a) VG density



(b) VG log-density



(c) VG QQ plot



(d) VG PP plot

Figure 3.10: Fitting the VG distribution to the S&P 500 Index returns. The plot shows outputs of the MLE approach.

3.4.3 Conclusion

Based on all provided results, we see that PL (pseudo-likelihood) gives the best fit with respect to the data at hand, which seems to be counter intuitive in this situation, see the figure 3.6. One of the main properties of this model, is the fact that $\theta \neq 0$ or existence of the dependence structure of the log-returns. If we just simply disregard this by assuming independence, one would expect to get far worse estimates, hence we would expect MOM to be superior to the MLE. One possible explanation is that the degree of autocorrelation for increments in the data is not as high as one would expect, i.e. that parameter θ is quite small.

This leads us to the conclusion that, unless the data at hand is extremely skewed and/or correlated, and in addition, interest lies in particularly estimating only the symmetry (dependence parameter) (θ), then pseudo-likelihood estimation is probably the best bet. Note, however that such highly skewed and dependent data one is not very likely to encounter in practice, the data we worked with are particularly realistic values of actual stock prices.

Even though the results are somewhat surprising, even more so provided that in the non-independence case pseudo-likelihood has theoretical drawbacks. Provided that the maximum likelihood is the superior estimation method for the independent data, it seems that, at least for the cases considered, the dependence structure did not impose the true empirical joint density to be sufficiently different from the product of marginal densities in order to cause the MOM method to perform better than the MLE.

APPENDICES

Appendix A

Mixing Coefficients and some important results

Let's denote $\alpha(t)$ a strong mixing coefficient and $\beta(t)$ a β -mixing coefficient or absolute regularity, see Bradley (2008) [13], as following:

$$\alpha(t) = \sup |P(A \cap B) - P(A)P(B)|, \quad A \in \mathcal{F}_s, B \in \mathcal{F}_{s+t}$$

and

$$\beta(t) = \sup |P(A/B) - P(A)P(B)|, \quad A \in \mathcal{F}_s, B \in \mathcal{F}_{s+t},$$

where $\mathcal{F}_s = \sigma(Y(s) : s \leq t)$, $\mathcal{F}_t = \sigma(Y(s) : s \geq t)$, for any $t \geq 0$.

Below we present a very important result on mixing, suggested in Davydov, [18].

Theorem A.0.1 (Davydov). *Let Y_t satisfy strong mixing conditions, i.e. $\alpha(t) \rightarrow 0$, as $t \rightarrow \infty$. Let also $E|Y_0|^{2+\sigma} = \tau^{2+\sigma} < \infty$ and $\sum_n \alpha(n)^{\sigma/2(2+\sigma)} = C < \infty$ for some $\sigma > 0$. Now, if $\sigma_n^2 = \text{Var} \left(\sum_{j=0}^{n-1} Y_n \right) \rightarrow \infty$, then we have $X_n \xrightarrow{Fdd} W(t)$, as $n \rightarrow \infty$ where $Y_n(t)$ is an interpolation process on $[0, 1]$ and $W(t)$ is a Brownian motion.*

Appendix B

Financial prerequisites

Financial instruments

In this section, we give a brief overview of the basic financial instruments. There are different ways in which companies can raise funds for working capital and capital expenditures. One is raising capital by equity financing. It refers to the sale of an ownership interest in a company by selling shares/stocks. The value of the company's shares typically reflects the value of real assets of the company along with its earning power. Another way of financing is called debt financing and it occurs when a firm raises money by selling bonds, bills or notes to individuals and institutional investors. The investors or creditors, in return for lending the money receive a promise that both, principal and interest on the debt will be repaid at some specified time in the future.

Below we give definitions of most commonly used instruments in the market.

Definition B.0.1. *A stock (also known as share or equity) is a type of security that provides partial ownership¹ in a company and holder of a stock (or shareholder) has a claim on part of companies assets and earnings.*

Definition B.0.2. *A stock index is a measure that follows a value of a portfolio of stocks that are representing the overall or only a portion of the market².*

¹The level of ownership is determined by the ratio of shares held relative to number of shares outstanding.

²Typically, depending on what sector of the stock market one is interested in, a smaller sample of stocks is selected as a good representative of the whole and used to track the performance of that sector.

Ideally, a change in the index price should be proportional to a change in the price of the stocks included. In the past, the price of the stock market index used to be calculated as an average of the stock prices included, while nowadays weighted average is used with either price-based³ weighing or most commonly used, market capitalization⁴ based weighing.

Keep in mind that the stock index is nothing more than a list of stocks and that anybody can create one. The most commonly known indices are Dow Jones Industrial Average (DJIA), Standard and Poor's 500 (S&P 500), Nasdaq Composite, etc. DJIA is one of the most recognized indices and it used to be referred to as "the market". It contains 30 of the most largest and influential companies in the U.S. On the other hand, the S&P 500 index contains 500 most widely held companies in the U.S., which makes it a better representative of the entire U.S. stock market and is commonly used as a market benchmark.

Modeling assumptions

Models that we used in the thesis have common underlying assumptions about the market, its participants and mathematical structure of the price processes. In this section we'll give a brief overview and rationale behind these. In every model, we assumed a simplistic market framework which states that only two assets are traded continuously in the market, a risky and a riskfree asset. Hence, a first group of assumptions is related to risky instruments and is frequently referred to as: set of "*frictionless market assumptions*" while other set of assumptions is associated with the riskfree asset itself.

³very susceptible to stock splits that in turn affect a company's weight in the index.

⁴Also referred to as "Market Cap". It is the total dollar market value of company's outstanding shares. It is calculated by multiplying the shares outstanding of a company with the current market price of one share.

Frictionless market assumptions:

1. No transaction costs:

By transaction costs we consider broker's commission, bid/ask spread, taxes, margin requirements, etc. Note that this is not as strong assumption as it seems, since most of the above requirements are satisfied for large market participants and financial institutions. This makes it a very reasonable approximation to reality.

2. Perfectly liquid markets:

This assumption implies that it is possible to trade any amount or fraction of a financial instrument available on the market, at any given time. More precisely, it is a market with large number of bid and ask offers, low spreads and also lower volatility. In these markets it is straightforward to execute a trade at any desirable price and quickly since there are numerous buyers and sellers. Also, relative changes in supply and demand have a small impact on the price of an instrument. The opposite of a liquid market is often called a *thin or illiquid market*. An example of a perfectly liquid market is a market for the stock of a Fortune 500 company or the forex market which is considered the largest and the most liquid market. The above assumption of a perfectly liquid market is well justified considering that, we are modeling stocks from the Fortune 500 companies for which demand and supply is very high.

3. No restrictions nor penalties or fees on short sales:

In February 2010, Securities and Exchange Commission (SEC) adopted a new short sale price restriction, which is commonly referred to as an *alternative uptick rule*⁵, designed

⁵The *uptick rule* or a *short sale rule* was put in place in the period from 1938 till July 2007 as a way to control a declining market, since at that time mechanisms and liquidity couldn't be guaranteed to prevent panic stock declines or any type of manipulation. The rule states that stocks could be shorted only on an uptick (last share price higher than the one before) or a zero-plus tick (last trade is the same as the previous one which was an uptick) and not on a downtick. However, the *naked shorting*, or selling the shares that

to preserve investor confidence and promote market efficiency, as short selling can potentially have both a beneficial and a harmful impact on the market. The alternative uptick rule is put in order to restrict short selling from further driving down the price of a stock that has dropped more than 10 percent in one day compared to the closing price on the previous day. Once the circuit breaker⁶ is triggered, it will give an advantage to long sellers over short sellers so they can sell their shares and get out of their positions before the price drops down even further.

4. No transaction delays:

This means that trading is continuous, i.e. that markets never close and there is no delay in execution of the trades. Note that with this assumption we are excluding high-frequency trading as a possibility.

5. No default risk:

The risk of failure of a company or bankruptcy along with other risks that are due to economic and political activity are ignored for simplicity. This makes an assumption of equal compounded interests rates for borrowing and lending a plausible assumption.

6. Competitive markets (market participants act as price takers)

Price takers are individuals and companies who don't have enough market share to influence market price on their own, so they have to accept the current available spot price. This means that market is not influenced by trading of individuals and even large amounts of trading of one participant cannot influence the prices in the market. In addition, we assume that brokers accept all trades.

7. Rational agents (market participants prefer more to less)

can't be verified is still forbidden.

⁶Also known as *collars* are control measures approved by SEC in order to prevent panic sellouts on the U.S. stock exchange and to manage excessive volatility in the market.

This is a relatively weak assumption, since it doesn't impose anything on the preferences and beliefs of the market participants except that they are rational in sense that they prefer more money for the same services.

8. No arbitrage

The rationale behind this assumption is: if there exists such a contract in the market that offers "something for nothing" (also known as a "free lunch") then the simple principle of supply and demand would force the price to converge to an equilibrium price.

Riskfree asset assumptions:

Second set of assumptions is concerned with a riskless asset or a bank account and it assumes the following list of conditions:

1. As mentioned previously, our market framework consists of two assets: a stock as a risky and a bank account as a riskless asset. Instead of a stock one can use an index and instead of a bank account one can use a riskfree bond. Also, it is assumed that in the market all investors are allowed to trade continuously up to a fixed finite time horizon T .
2. A bank account B_t is assumed to grow exponentially as it is described by the following differential equation: $dB_t/B_t = rdt, /, r \geq 0$. In other words, we assume that the continuously compounded rate of return for asset B_t is r , where r is deterministic and constant.
3. It is assumed that investors can deposit and borrow money at the same interest rate, i.e. borrowing and lending rates are equal to the riskfree rate of return r .

Probability setup

Even though in practice, stock prices are observed only when the exchange is open and the prices are restricted to discrete values (often multiples of 0.01 dollars), we use continuous time stochastic process with continuous realizations to model a stock price evolution. Hence, we will model the price process of our asset by a continuous-time stochastic process $\{S_t : 0 \leq t \leq T\}$, where S_t gives us the price of a stock at time $t \geq 0$.

Let (Ω, \mathcal{F}, P) be a probability space. We assume that Ω is set of all possible outcomes that we are interested in, eg. set of all possible outcomes for one stock. \mathcal{F} contains the set of all events we would like to evaluate, eg. find how probable is the event that the stock price will be above \$100. The probability measure P is a measure that reflects everything that happens in the market, including both, investors' personal belief about the price and their risk preference, i.e. measure P includes investors' individual biases. Example is given below.

Example B.0.1. *When one is deciding whether to go long or short with a certain stock, they would like to know what are the chances of a stock increasing or decreasing. So, they will first look at the historical data to see if there are any increasing/decreasing trends which will provide an initial belief about the distribution of up and down movements. Second, they will look at how volatile the stock was in the past and based on their risk preference they will adjust their initial probabilities. To summarize, based on historical data and their risk preference they will construct a subjective probability measure they will use to make decisions.*

Remark B.0.1. *Notice, the investor's risk preference, risk averse versus risk seeking is an important factor in both, when building an investment portfolio and when pricing an instrument with the same return value but different levels of risk. As such, the risk preference*

has a strong influence on investors' decision making process and therefore it is reflected in the overall market probability measure P .

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] D. Applebaum. *Levy Processes and Stochastic Calculus*. Cambridge University Press, Cambridge, 2009.
- [2] D. Applebaum. Lectures on Levy Processes and Stochastic Calculus (University of Sheffield). Lecture 5: The Ornstein-Uhlenbeck Process, July 2010.
- [3] Ole E Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance and stochastics*, 2(1):41–68, 1997.
- [4] Ole E Barndorff-Nielsen. Superposition of Ornstein–Uhlenbeck type processes. *Theory of Probability & Its Applications*, 45(2):175–194, 2001.
- [5] Ole E Barndorff-Nielsen, J.L Jensen, and M. Sorensen. Some stationary processes in discrete and continuous time. *Advances in Applied Probability*, 30:9891007, 1998.
- [6] Ole E Barndorff-Nielsen and N.N. Leonenko. Burgers’ turbulence problem with linear or quadratic external potential. *Journal of applied probability*, 42(2):550–565, 2005.
- [7] Ole E. Barndorff-Nielsen and N.N. Leonenko. Spectral properties of superpositions of Ornstein-Uhlenbeck type processes. *Methodology and computing in applied probability*, 7(3):335–352, 2005.
- [8] Ole E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):167–241, 2001.
- [9] Ole E Barndorff-Nielsen and R. Stelzer. Multivariate supOU processes. *The Annals of Applied Probability*, 21(1):140–182, 2011.
- [10] C. Bender and T. Marquardt. Integrating volatility clustering into exponential levy models. *Journal of applied probability*, 46(3):609–628, 2009.
- [11] L. Bertini and N. Cancrini. The stochastic heat equation: Feynman-kac formula and intermittence. *Journal of Statistical Physics*, 78(5-6):1377–1401, 1995.

- [12] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Politics and Economics*, 81:637–654, 1973.
- [13] R.C. Bradley. Basic properties of strong mixing conditions. a survey and some open questions. *Probability surveys*, 2(2):107–144, 2005.
- [14] R. Carmona and S.A. Molchanov. *Parabolic Anderson Problem and Intermittency*, volume 518. American Mathematical Soc., 1994.
- [15] AV Chechkin, V. Yu Gonchar, M. Szydl, et al. Fractional kinetics for relaxation and superdiffusion in a magnetic field. *Physics of Plasmas*, 9(1):78–88, 2002.
- [16] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.
- [17] D.R. Cox and N. Reid. A note on pseudolikelihood constructed from marginal densities. *Biometrika*, 91:729–737, 2004.
- [18] Yu A Davydov. Convergence of distributions generated by stationary stochastic processes. *Theory of Probability & Its Applications*, 13(4):691–696, 1968.
- [19] E. Eberlain and S. Raible. Term structure models driven by general levy processes. *Mathematical Finance*, 9:31–53, 1999.
- [20] I. Eliazar and J. Klafter. Lévy, Ornstein–Uhlenbeck, and subordination: Spectral vs. jump description. *Journal of statistical physics*, 119(1-2):165–196, 2005.
- [21] P. Embrechts and M. Maejima. *Selfsimilar Processes*. Princeton University Press, Princeton, NJ, 2002.
- [22] I. Endre and T. György. Superposition of diffusions with linear generator and its multifractal limit process. *ESAIM: Probability and Statistics*, 7:23–88, 2003.
- [23] E. Fama. The behavior of stock prices. *Journal of Business*, 38:34105, 1965.
- [24] R. Finlay and E. Seneta. Stationary-increment student and variance-gamma processes. *Journal of Applied Probability*, 43(4):441–453, 2006.
- [25] R. Finlay and E. Seneta. Option pricing with vg-like models. *International Journal of Theoretical and Applied Finance*, 11(8), 2007.

- [26] R. Finlay and E. Seneta. A gamma activity time process with noninteger parameter and self-similar limit. *Journal of Applied Probability*, 44(4):943–955, 2008.
- [27] Uriel Frisch. *Turbulence: the legacy of A.N. Kolmogorov*. Cambridge University Press, Cambridge, 1995.
- [28] Hirokazu Fujisaka. Theory of diffusion and intermittency in chaotic systems. *Progress of theoretical physics*, 71(3):513–523, 1984.
- [29] J. Gärtner, F Den H., and G. Maillard. Intermittency on catalysts: voter model. *The Annals of Probability*, 38(5):2066–2102, 2010.
- [30] J. Gärtner and F. den Hollander. Intermittency in a catalytic random medium. *The Annals of Probability*, 34(6):2219–2287, 2006.
- [31] J. Gärtner, W. König, and S. Molchanov. Geometric characterization of intermittency in the parabolic Anderson model. *The Annals of Probability*, 35(2):439–499, 2007.
- [32] J. Gärtner and S.A. Molchanov. Parabolic problems for the Anderson model. *Communications in Mathematical Physics*, 132(3):613–655, 1990.
- [33] D. Grahovac and N. N. Leonenko. Bounds on the support of the multifractal spectrum of stochastic processes. *arXiv preprint arXiv:1406.2920*, 2015.
- [34] D. Grahovac, N.N. Leonenko, A. Sikorski, and M. Taqqu. The unusual properties of long-range dependent superpositions of ornstein-uhlenbeck type processes (under review). 2016.
- [35] C.W.L. Granger. The past and future of empirical finance: Some personal comments. *Journal of Econometrics*, 129:35–40, 2005.
- [36] C. C. Heyde and N.N. Leonenko. Student processes. *Journal of Applied Probability*, 37:342–365, 2005.
- [37] C.C Heyde. A risky asset model with strong dependence through fractal activity time. *Journal of Applied Probability*, 36:1234–1239, 1999.
- [38] C.C. Heyde. Scaling issues for risky modeling. *Math. meth. Oper. Res.*, 69:593–603, 2009.

- [39] C.C. Heyde and S Liu. Empirical realities for a minimal description risky asset model. the need for fractal features. *J. Korean Math Soc.*, 38:1047–1059, 2001.
- [40] C.C. Heyde and Y. Yang. On defining long-range dependence. *Journal of Applied Probability*, 34:939–944, 1997.
- [41] S. Howison and D. Lamper. Trading volume in models of financial derivatives. *Applied Mathematical Finance*, 8:119–135, 2001.
- [42] G. Jongbloed, F.H. Van Der Meulen, and A.W. Van Der Vaart. Nonparametric inference for Lévy-driven Ornstein-Uhlenbeck processes. *Bernoulli*, 11(5):759–791, 2005.
- [43] Mz.J. Jurek and W. Vervaat. An integral representation for selfdecomposable banach space-valued random variables. *Z. Wahrsch. verw. Geb.*, 62:24762, 1983.
- [44] D. Khoshnevisan. *Analysis of stochastic partial differential equations*, volume 119. American Mathematical Soc., 2014.
- [45] D. Khoshnevisan, K. Kim, and Y. Xiao. Intermittency and multifractality: A case study via parabolic stochastic PDEs. *arXiv preprint arXiv:1503.06249*, 2015.
- [46] Kobayashi. Stochastic calculus for a time-changed semimartingale and the associated stochastic differential equations. *Journal of Theoretical Probability*, 24(3):789–820, 2011.
- [47] N. Leonenko and E. Taufer. Convergence of integrated superpositions of Ornstein-Uhlenbeck processes to fractional Brownian motion. *Stochastics: An International Journal of Probability and Stochastics Processes*, 77(6):477–499, 2005.
- [48] Nikolai N Leonenko, S Petherick, and A Sikorskii. Fractal activity time models for risky asset with dependence and generalized hyperbolic distributions. *Stochastic Analysis and Applications*, 30(3):476–492, 2012.
- [49] N.N. Leonenko, S Petherick, and A Sikorskii. A normal inverse Gaussian model for a risky asset with dependence. *Statistics & Probability Letters*, 82(1):109–115, 2012.
- [50] L. Li and V. Linetsky. Time-changed Ornstein–Uhlenbeck processes and their applications in commodity derivative models. *Mathematical Finance*, 24(2):289–330, 2014.
- [51] D.B. Madan, P.P. Carr, and E.C. Chang. The variance-gamma process and option pricing. *European Finance Review*, 2:79105, 1998.

- [52] D.B. Madan and E. Seneta. The variance-gamma (v.g.) model for share market returns. *The Journal of Business*, 63(4):511–524, 1990.
- [53] B. Mandelbrot and H.M. Taylor. On the distribution of stock price differences. *Operations Research*, 15(6):1057–1062, 1967.
- [54] V. Mandrekar and Rudiger B. *Stochastic integration in Banach spaces*. Springer Verlag, 2007.
- [55] H. Masuda. On multidimensional Ornstein-Uhlenbeck processes driven by a general Lévy process. *Bernoulli*, 10(1):97–120, 2004.
- [56] R.C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41:867–887, 1973.
- [57] S.A. Molchanov. Ideas in the theory of random media. *Acta Applicandae Mathematica*, 22(2-3):139–282, 1991.
- [58] Sidney I. Resnick. *Heavy-tail Phenomena: Probabilistic and Statistical Modeling*. Springer Science & Business Media, New York, 2007.
- [59] L.M. Ricciardi and L. Sacerdote. The Ornstein-Uhlenbeck process as a model for neuronal activity. *Biological cybernetics*, 35(1):1–9, 1979.
- [60] A. Rocha-Arteaga and K. Sato. *Topics in Infinitely Divisible Distributions and Levy Processes*. Sociedad Matematica Mexicana, Mexico, 2003.
- [61] T.H. Rydberg. Realistic statistical modelling of financial data. *Inter. Stat. Review*, 68:233–25, 2000.
- [62] T. Sandev, A.V. Chechkin, N. Korabel, H. Kantz, I.M. Sokolov, and R. Metzler. Distributed-order diffusion equations and multifractality: Models and solutions. *Physical Review E*, 92(4):042117, 2015.
- [63] Kenji-Iti Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999.
- [64] E. Seneta. *Regularly Varying Functions*. Springer Verlag, Berlin-Heidelberg-New York, 1976.

- [65] E. Seneta. Fitting the variance-gamma model to financial data. *Journal of Applied Probability*, 41(A):177–187, 2004.
- [66] T.H. Solomon, E.R. Weeks, and H.L. Swinney. Observation of anomalous diffusion and lévy flights in a two-dimensional rotating flow. *Physical Review Letters*, 71(24):3975, 1993.
- [67] Wojbor A. Woyczyński. *Burgers-KPZ turbulence(Göttingen lectures)*. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1998.
- [68] M. Yamazato. Unimodality of infinitely divisible distribution functions of class 1. *Ann. of Probability*, 6:523–531, 1978.
- [69] Ya B. Zel’dovich, S.A. Molchanov, A.A. Ruzmaïkin, and D.D. Sokolov. Intermittency in random media. *Soviet Physics Uspekhi*, 30(5):353, 1987.