# Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type

Luis Valdivieso\* Wim Schoutens<sup>†</sup> Francis Tuerlinckx<sup>‡</sup>

#### Abstract

In this article we propose a maximum likelihood methodology to estimate the parameters of a one-dimensional stationary process of Ornstein-Uhlenbeck type that is constructed via a self-decomposable distribution D. Our approach is based on the inversion of the characteristic function and the use of the classical or fractional discrete fast Fourier transform. The results are illustrated throughout an extensive simulation study. This includes the cases where D belongs to the gamma, tempered stable and normal inverse Gaussian family of distributions.

Keywords: Ornstein-Uhlenbeck processes; Likelihood inference; Fourier transform

 $<sup>^*</sup>$ Universitair Centrum voor Statistiek, K.U. Leuven, W. de Croylaan 54, 3001 Heverlee, Belgium. Email: l<br/>valdiv@pucp.edu.pe

<sup>&</sup>lt;sup>†</sup>Universitair Centrum voor Statistiek, K.U.Leuven, W. de Croylaan 54, 3001 Heverlee, Belgium. Email: Wim.Schoutens@wis.kuleuven.be

<sup>&</sup>lt;sup>‡</sup>Department of Phsychology, K.U. Leuven, Tiensestraat 102, 3000 Leuven, Belgium. Email: francis.tuerlinckx@psy.kuleuven.be

#### 1 Introduction

A stochastic process  $\mathbf{X} = \{X(t)\}$  is said to be an Ornstein-Uhlenbeck process if it satisfies the stochastic differential equation

$$\begin{cases} dX(t) = -\lambda X(t)dt + \sigma dW(t) \\ X(0) = X_0, \end{cases}$$

where  $\mathbf{Z} = \{Z(t)\}$  is a Gaussian process with mean 0 and variance  $t\sigma^2$ ,  $\lambda$  is a strictly positive intensity parameter and  $X_0$  is an independent random variable. One possible generalization of this process emerges from allowing  $\mathbf{Z}$  to be a Lévy process. Such a model is called a process of Ornstein-Uhlenbeck type and  $\mathbf{Z}$  is referred to as its background driving Lévy process (BDLP). In our discussion we will be mainly interested in stationary processes of this type, which can be generated by imposing certain conditions to the BDLP  $\mathbf{Z}$  or more interestingly, by designing a stationary self-decomposable law, say D, and finding then a BDLP  $\mathbf{Z}$  that exactly matches this distribution. This last process is called a D-OU process. Recently, Barndorff-Nielsen and Shepard [4] have proposed the use of these processes for modelling the volatility coefficient in a stock price process that follows a geometric Brownian motion. Cariboni and Schoutens [6] have also employed these processes for the dynamic of the default intensity in a intensity-based credit risk model.

The main contribution of this paper lies in the presentation of a new methodology to obtain maximum likelihood estimates for the parameters governing a general D-OU process. Our estimations are based on the data observed from the D-OU process. In this context, our work is related with the one of Barndorff-Nielsen [3] who proposes a simulation-based likelihood approximation method when D is an inverse Gaussian or a normal inverse Gaussian distribution. Jiang and Pedersen [8] and Prause [13] have also addressed this problem in the context of the stochastic volatility model above under a superposition of inverse Gaussian-OU processes for the volatility coefficient. They propose to estimate the intensity parameters by regression and in a second step the remaining parameters by treating the observed log stock price increments as coming from a normal inverse Gaussian random sample. From a non-parametric point of view, the estimation of  $\lambda$  and the Lévy measure of the BDLP  $\mathbf{Z}$  has been considered in [9].

In short, our methodology is based on the inversion of the characteristic function and the use of the classical or fractional discrete Fast Fourier transform. In order to asses the quality of our estimations we will perform an extensive simulation study covering the cases where the self-decomposable distribution D belongs to the gamma, tempered stable and normal inverse Gaussian family of distributions. The choice of these distributions is motivated not only by their potential applications, like in the context of stochastic volatility models, but also by their flexibility to cover different scenarios. While the BDLP of the gamma-OU and tempered stable-OU processes have paths of finite variation with finitely and infinitely, respectively, many positive jumps in any finite interval, the BDLP of the normal inverse Gaussian-OU process has paths of infinite variation with infinitely many positive and negative jumps in any finite interval.

The paper is organized as follows. In Section 2 we describe the inference problem for the classical Ornstein-Uhlenbeck process. Section 3 defines a D-OU process and introduces the gamma, tempered stable and normal inverse Gaussian candidates for D. In Section 4

we present the inference methodology for a general D-OU process together with explicit calculations in the case of the previously mentioned D-OU processes. In Section 5 we discuss the simulation of a D-OU processes. Section 6 presents the simulation study and Section 7 concludes.

# 2 The Ornstein-Uhlenbeck process

A stochastic process  $\mathbf{X} = \{X(t)\}$  is said to be an Ornstein-Uhlenbeck process if it satisfies the homogeneous linear stochastic differential equation:

$$\begin{cases} dX(t) = -\lambda X(t)dt + \sigma dW(t) \\ X(0) = X_0, \end{cases}$$
 (1)

where  $\lambda$  and  $\sigma$  are strictly positive parameters and  $X_0$  is a random variable independent of the standard Brownian motion  $\mathbf{W} = \{W(t)\}$ . As shown in [10], (1) has the unique strong solution:

$$X(t) = \exp(-\lambda t) \left( X_0 + \sigma \int_0^t \exp(\lambda s) dW(s) \right).$$

Moreover X is a continuous in probability Markov process satisfying:

- $E[X(t)] = \exp(-\lambda t)E[X_0]$
- $Var(X(t)) = \frac{\sigma^2}{2\lambda} + (Var(X_0) \frac{\sigma^2}{2\lambda}) \exp(-2\lambda t)$
- $Cov(X(s), X(s+t)) = \left(Var(X_0) + \frac{\sigma^2}{2\lambda}(\exp(2\lambda s) 1)\right) \exp(-\lambda(2s+t)).$

Observe that if  $X_0 \sim N(0, \frac{\sigma^2}{2\lambda})$ , **X** becomes a (strictly) stationary Gaussian process with covariance function:

$$c(t) = Cov(X(s), X(s+t)) = \frac{\sigma^2}{2\lambda} \exp(-\lambda t).$$

Note also that if we slightly modify (1) to:

$$\begin{cases} dX(t) = -\lambda X(t)dt + \sigma dZ(\lambda t) \\ X(0) = X_0, \end{cases}$$
 (2)

where  $X_0 \sim N(0, \frac{\sigma^2}{2})$  and  $\mathbf{Z} = \{Z(t)\}$  is an independent Gaussian process with  $Z(t) \sim N(0, t\sigma^2)$ , then the associated stationary distribution is normal with mean 0 and variance  $\frac{\sigma^2}{2}$ . In other words, (2) removes the dependence on  $\lambda$  in the stationary law.

Before exploring the possibility to extend equation (2), by allowing **Z** to be a Lévy process, it is illustrative to study the maximum likelihood estimation of the vector parameter  $\theta = (\sigma^2, \lambda)$  in (2). For this purpose, let  $X(t_0), X(t_1), X(t_2), \dots, X(t_n)$  be a sample of the stationary process defined by (2),  $x_0, x_1, x_2, \dots, x_n$  their observed values and  $\Delta_k = t_k - t_{k-1}$ .

Thanks to the Markovian and Gaussian property, the likelihood function of this sample is explicitly given by:

$$L(\theta) = f_{X(t_0)}(x_0) \prod_{k=1}^{n} f_{X(t_k)|X(t_{k-1})=x_{k-1}}(x_k)$$

$$= \frac{1}{\sigma\sqrt{\pi}} \exp\left(-\frac{x_0^2}{\sigma^2}\right) \prod_{k=1}^{n} \frac{\exp\left(\frac{-(x_k - \exp(-\lambda\Delta_k)x_{k-1}^2)}{\sigma^2(1 - \exp(-2\lambda\Delta_k))}\right)}{\sigma\sqrt{\pi(1 - \exp(-2\lambda\Delta_k))}}.$$

Hence, the observed maximum likelihood estimator can be easily found by numerically maximizing the log-likelihood function

$$K(\theta) = -\frac{(n+1)}{2}\log(\pi\sigma^2) - \frac{x_0^2}{\sigma^2} - \frac{1}{2}\sum_{k=1}^n (1 - \exp(-2\lambda\Delta_k))$$
$$-\sum_{k=1}^n \frac{(x_k - \exp(-\lambda\Delta_k)x_{k-1})^2}{\sigma^2(1 - \exp(-2\lambda\Delta_k))}.$$

# 3 Processes of Ornstein-Uhlenbeck type

Let  $\mathbf{Z} = \{Z(t)\}$  be an univariate Lévy process with generating triplet  $(\sigma_0, \gamma_0, \nu_0)$  and let  $\lambda > 0$ . In our notation  $\sigma_0 > 0$  represents the volatility parameter of the Brownian motion Lévy component,  $\gamma_0$  the drift and  $\nu_0$  the Lévy measure of the process.

A stochastic process  $\mathbf{X} = \{X(t)\}$  is said to be a process of Ornstein-Uhlenbeck type generated by  $(\sigma_0, \gamma_0, \nu_0, \lambda)$  if it is càdlàg and satisfies the stochastic differential equation

$$\begin{cases} dX(t) = -\lambda X(t)dt + dZ(\lambda t) \\ X(0) = X_0, \end{cases}$$

where  $X_0$  is a random variable independent of **Z**. Under this setting **Z** is termed the background driving Lévy process (BDLP).

Using the theory of integration of left-continuous predictable process with respect to semimartingales and in particular to Lévy processes (see [14]), we obtain for any  $t \ge 0$  and  $\Delta \ge 0$ :

$$X(t) = \exp(-\lambda t) \left( X_0 + \int_0^t \exp(\lambda s) dZ(\lambda s) \right),$$

or recursively

$$X(t + \Delta) = \exp(-\lambda \Delta) \left( X(t) + \exp(-\lambda t) \int_{t}^{t+\Delta} \exp(\lambda s) dZ(\lambda s) \right).$$
 (3)

As a result,  $\mathbf{X}$  is a Markov process. Furthermore, we have the next property related to the Lévy functional  $Z^*(\Delta) = \int_0^{\lambda\Delta} \exp(s) \mathrm{d}Z(s)$ .

**Proposition 3.1** For any  $t \ge 0$  and  $\Delta \ge 0$ 

$$Z^*(\Delta) \stackrel{d}{=} \int_0^{\Delta} \exp(\lambda s) dZ(\lambda s) \stackrel{d}{=} \exp(-\lambda t) \int_t^{t+\Delta} \exp(\lambda s) dZ(\lambda s),$$

where  $\stackrel{d}{=}$  means equality in distribution.

**Proof:** The first equality holds for simple predictable integrand processes and then, by density arguments, for the exponential function. The second equality follows from the application of the integration by parts rule for semi-martingales (see [14]). By this rule we can write  $\int_t^{t+\Delta} \exp(\lambda s) dZ(\lambda s)$  as

$$\exp(\lambda(t+\Delta))Z(\lambda(t+\Delta)) - \exp(\lambda t)Z(\lambda t) - \exp(\lambda t)\int_{\lambda t}^{\lambda(t+\Delta)} \exp(s)Z(s)ds.$$

On the other side,  $\int_{\lambda t}^{\lambda(t+\Delta)} \exp(s)Z(s)ds$  equals

$$\int_0^{\lambda \Delta} \exp(s) Z(s + \lambda t) ds \stackrel{d}{=} \int_0^{\lambda \Delta} \exp(s) Z(s) ds + Z(\lambda t) (\exp(\lambda \Delta) - 1).$$

Hence.

$$\int_{t}^{t+\Delta} \exp(\lambda s) dZ(\lambda s) \stackrel{d}{=} \exp(\lambda (t+\Delta)) Z(\lambda \Delta) - \exp(\lambda t) \int_{0}^{\lambda \Delta} \exp(s) Z(s) ds$$

and applying once more the integration by parts rule

$$\exp(-\lambda t) \int_{t}^{t+\Delta} \exp(\lambda s) dZ(\lambda s) \stackrel{d}{=} \int_{0}^{\Delta} \exp(\lambda s) dZ(\lambda s).$$

Throughout this paper we will understand any distribution  $\mu$  as a probability distribution on  $\mathbb{R}$ . The characteristic and cumulant function of  $\mu$  will be denoted, respectively, by  $\phi_{\mu}$  and  $C_{\mu} = \log(\phi_{\mu})$ .

**Definition 3.1** A distribution  $\mu$  is said to be self-decomposable if for any 0 < a < 1, there exists a distribution  $v_a$  such that

$$\phi_{\mu}(\vartheta) = \phi_{\mu}(a\vartheta)\phi_{\nu_{a}}(\vartheta).$$

Alternatively, a random variable X is said to have a self-decomposable distribution if for any 0 < a < 1, there exists a random variable  $Y_a$ , independent of X, such that

$$X \stackrel{d}{=} aX + Y_a$$
.

Sato [18] has shown that any non-degenerate self-decomposable distribution  $\mu$  is absolutely continuous and infinitely divisible. Then, the Lévy-Khintchine representation theorem can be invoked to guarantee the existence of a  $\mu$ -Lévy process with characteristic function  $\phi_{\mu}^{t}$  at the time index t>0. In addition, as proved in [21], this process has a Lévy measure with an unimodal density that we will denote hereafter by u.

In relation to self-decomposability, Sato [18] has also established the following fundamental result about the stationarity of a process of Ornstein-Uhlenbeck type.

**Proposition 3.2** If X is a process of Ornstein-Uhlenbeck type generated by  $(\sigma_0, \gamma_0, \nu_0, \lambda)$  such that

$$\int_{|x|>2} \log(|x|) \, dv_0(x) < \infty,\tag{4}$$

then X has a unique self-decomposable stationary distribution  $\mu$ .

Conversely, for any  $\lambda > 0$  and any self-decomposable distribution D, there exists a unique triplet  $(\sigma_0, \gamma_0, \nu_0)$  satisfying (4) and a process of Ornstein-Uhlenbeck type  $\mathbf{X}$  generated by  $(\sigma_0, \gamma_0, \nu_0, \lambda)$  such that D is the stationary distribution of  $\mathbf{X}$ .

The stationary process X in the converse result of Proposition 3.2 is called a D-OU process. From (3), the autocorrelation function of this process takes the form

$$\rho(\Delta) = \frac{Cov(X(t), X(t+\Delta))}{\sqrt{Var(X(t))Var(X(t+\Delta))}} = \exp(-\lambda|\Delta|), \tag{5}$$

for any  $t \geq 0$  and  $\Delta \in \mathbb{R}$ .

If  $(\sigma, \gamma, \nu)$  denotes the Lévy triplet of the stationary distribution D, then  $(\sigma_0, \gamma_0, \nu_0)$  satisfy

$$\sigma_0 = 2\sigma$$
 and  $\gamma_0 = \gamma - \int_{-\infty}^{-1} w(x) dx - \int_{1}^{\infty} w(x) dx$ ,

where w is the Lévy density associated to  $\nu_0$ . In [3] it is shown that if the Lévy density associated to  $\nu$ , u, is differentiable then w satisfies

$$\omega(x) = -u(x) - xu'(x).$$

Furthermore we have the following key relation among the cumulants of D,  $\mathbf{Z}$  and the Lévy functional  $Z^*(\Delta)$  (see [3] for the proof).

**Proposition 3.3** For any  $\Delta > 0$  and  $\vartheta \in \mathbb{R}$ :

$$C_{Z^*(\Delta)}(\vartheta) = \log(E[\exp(i\vartheta Z^*(\Delta))]) = \lambda \int_0^{\Delta} C_{Z(1)}(\vartheta \exp(\lambda s)) ds,$$

where  $C_{Z(1)}(\vartheta) = \vartheta \frac{dC_D(\vartheta)}{d\vartheta}$ .

We conclude this section by briefly introducing some important D-OU processes.

# 3.1 The gamma $\Gamma(a,b)$ -OU process

A random variable X has a gamma distribution  $\mu$  with parameters a > 0 and b > 0, for short a  $\Gamma(a, b)$  distribution, if its density function is given by:

$$f_X(x) = \frac{b^a x^{a-1}}{\Gamma(a)} \exp(-bx), \quad \forall x > 0,$$

where  $\Gamma$  denotes the gamma function.

The characteristic function of this distribution is known to be:

$$\phi_{\mu}(\vartheta) = \left(\frac{b}{b - \mathrm{i}\vartheta}\right)^{a}$$

As proved in [7], this is a self-decomposable distribution. Hence, by Proposition 3.1 the  $\Gamma(a,b)$ -OU process is well defined. Moreover, the BDLP **Z** of this process has, by virtue of Proposition 3.3, a cumulant characteristic function:

$$C_{Z(1)}(\vartheta) = \frac{a\vartheta i}{b - i\vartheta} = a\left(\frac{b}{b - i\vartheta} - 1\right).$$

Consequently, **Z** is a compound Poisson process with intensity parameter a and associated exponential distribution with parameter b.

### 3.2 The tempered stable $TS(\kappa, a, b)$ -OU process

A non-negative random variable has a tempered stable distribution  $\mu$  with parameters a > 0,  $b \ge 0$  and  $\kappa \in ]0,1[$ , for short a  $TS(\kappa,a,b)$  distribution, if its characteristic function is given by:

$$\phi_{\mu}(\vartheta) = \exp(a(b - (b^{\frac{1}{\kappa}} - 2i\vartheta)^{\kappa})).$$

As suggested by the name, this distribution is defined by exponentially tilting a positive and symmetric  $\kappa$ -stable distribution with characteristic function  $\varphi(\vartheta) = \exp(a2^{\kappa}\vartheta^{\kappa})$  and density function q. More precisely, the density function of this distribution is given by:

$$f(x) = \exp(ab - \frac{1}{2}b^{\frac{1}{\kappa}}x)g(x).$$

As it usually happens with a  $\kappa$ -stable distribution, f is known only in the form of series representation. A notable exception is  $\kappa = \frac{1}{2}$ . In this case  $\mu$  is known as an inverse Gaussian, or shortly, an IG(a,b) distribution and its density function is given by:

$$f(x) = \frac{a \exp(ab)}{\sqrt{2\pi x^3}} \exp(-\frac{1}{2}(\frac{a^2}{x} + b^2x)), \ \forall x > 0.$$

A further analysis of the tempered stable family of distributions and its extension to the case  $0 < \kappa < 2$  can be found in [16].

Since  $\mu$  is infinitely divisible, we can define a  $TS(\kappa, a, b)$ -Lévy process. This process has Lévy density

$$u(x) = \frac{a\kappa 2^{\kappa} x^{-(\kappa+1)}}{\Gamma(1-\kappa)} \exp(-\frac{b^{\frac{1}{\kappa}} x}{2})$$

and then it follows by [11], Theorem 5.11.2, that the tempered stable distribution is self-decomposable.

In [5] it is deduced that the BDLP  $\mathbf{Z} = \{Z(t)\}$  of a  $TS(\kappa, a, b)$ -OU process admits for any  $t \geq 0$  the decomposition:

$$Z(t) \stackrel{d}{=} I(t) + \sum_{j=1}^{N(t)} Y_j,$$

where  $\mathbf{I} = \{I(t)\}$  is a  $TS(\kappa, \kappa a, b)$ -Lévy process,  $\{N(t)\}$  is a Poisson process with intensity parameter  $ab\kappa$  and  $Y_1, Y_2, \ldots$  are independent and identically distributed  $\Gamma(1-\kappa, \frac{b^{\frac{1}{\kappa}}}{2})$  random variables.

#### 3.3 The normal inverse Gaussian $NIG(\alpha, \beta, \delta)$ -OU process

A random variable X has a normal inverse Gaussian distribution  $\mu$  with parameters  $\alpha$ ,  $\beta$  and  $\delta$ , for short a  $NIG(\alpha, \beta, \delta)$  distribution, if its density function is given by:

$$f(x) = \frac{\alpha}{\pi} \exp(\delta \sqrt{\alpha^2 - \beta^2} + \beta x) \frac{K_1(\alpha \delta \sqrt{1 + (\frac{x}{\delta})^2})}{\sqrt{1 + (\frac{x}{\delta})^2}}, \quad \forall x \in \mathbb{R},$$

where  $K_1$  is a modified Bessel function of the third kind of order 1 and the parameters  $\alpha, \beta, \delta$  satisfy the restrictions  $|\beta| < \alpha$  and  $\delta > 0$ . The characteristic function of this distribution is given by:

$$\phi_{\mu}(\vartheta) = \exp(\delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + i\vartheta)^2})),$$

showing that  $\mu$  is infinitely divisible and closed under convolutions if  $\alpha$  and  $\beta$  are kept fixed. One interesting property of a  $NIG(\alpha, \beta, \delta)$ -Lévy Process,  $\mathbf{Y} = \{Y(t)\}$ , is that it can be represented as an inverse Gaussian time-changed Brownian motion:

$$Y(t) = \beta \delta^2 IG(t) + \delta W(IG(t)),$$

where  $\mathbf{IG} = \{IG(t)\}\$  is an  $IG(1, \delta\sqrt{\alpha^2 - \beta^2})$ -Lévy process and  $\mathbf{W}$  is an standard Brownian motion. This process has a generating triplet  $(0, \gamma, v)$  with drift:

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

and Lévy measure:

$$v(B) = \int_B u(x) dx = \int_B \frac{\delta \alpha}{\pi |x|} K_1(\alpha |x|) \exp(\beta x) dx.$$

Halgreen [7] has shown that  $\mu$  is self-decomposable, so the  $NIG(\alpha, \beta, \delta)$ -OU process is well defined. This process possesses a pure-jump BDLP with density

$$w(x) = (1 - \beta x)u(x) + \frac{\delta \alpha^2}{\pi} K_0(\alpha |x|) \exp(\beta x).$$

Unlike with the so far analyzed up-jump D-OU processes having paths of bounded variation, the paths of the  $NIG(\alpha, \beta, \delta)$ -OU process are of unbounded variation with infinitely many up and down jumps in any finite interval.

# 4 Likelihood inference for a D-OU process

Let  $\mathbf{X} = \{X(t)\}$  be a D-OU process, where the stationary distribution D depends on an unknown parameter  $\theta \in \mathbb{R}^m$ . Suppose we are interested in estimating  $\theta$  based on a set of

n+1 observations  $x_0, x_1, \ldots, x_n$  coming from a sample  $X(0), X(t_1), X(t_2), \ldots, X(t_n)$  of **X**. By the Markov property, the likelihood function of this sample can be expressed as

$$L(\theta) = f_{X(0)}(x_0) \prod_{k=1}^{n} f_{X(t_k)|X(t_{k-1}) = x_{k-1}}(x_k),$$

where  $f_{X(t_k)|X(t_{k-1})=x_{k-1}}$  denotes the conditional density of the random variable  $X(t_k)$ , given  $X(t_{k-1}) = x_{k-1}$ .

From (3) and Proposition 3.1, we obtain the autoregressive scheme:

$$X(t_k) = \exp(-\lambda \Delta_k) \left( X(t_{k-1}) + \exp(-\lambda t_{k-1}) \int_{t_{k-1}}^{t_k} \exp(\lambda s) dZ(\lambda s) \right)$$

$$\stackrel{d}{=} \exp(-\lambda \Delta_k) \left( X(t_{k-1}) + Z^*(\Delta_k) \right),$$

where  $t_0 = 0$ ,  $\Delta_k = t_k - t_{k-1}$  and  $Z^*(\Delta_k) = \int_0^{\lambda \Delta_k} \exp(s) dZ(s)$ . Hence,

$$P(X(t_k) \le x_k | X(t_{k-1}) = x_{k-1}) = P(Z^*(\Delta_k) \le \exp(\lambda \Delta_k) x_k - x_{k-1})$$

and if  $Z^*(\Delta_k)$  is a continuous random variable

$$f_{X(t_k)|X(t_{k-1})=x_{k-1}}(x_k) = \exp(\lambda \Delta_k) f_{Z^*(\Delta_k)}(\exp(\lambda \Delta_k) x_k - x_{k-1}).$$

Note that if **Z** is a process of finite activity,  $Z^*(\Delta_k)$  turns out to be a mixed random variable. For instance, if **Z** is the BDLP of a  $\Gamma(a,b)$ -OU process  $Z^*(\Delta_k)$  takes the value 0 with probability  $p = \exp(-\lambda a \Delta_k) > 0$ . In such a situation we can slightly modify the conditional density above and write

$$f_{X(t_k)|X(t_{k-1})=x_{k-1}}(x_k)$$

$$= \begin{cases} p &, \text{ if } \xi_k = \exp(\lambda \Delta_k) x_k - x_{k-1} = 0\\ \exp(\lambda \Delta_k) f_{Z^*(\Delta_k)}^J(\xi_k) (1-p) &, \text{ if } \xi_k = \exp(\lambda \Delta_k) x_k - x_{k-1} > 0, \end{cases}$$

where  $f_{Z^*(\Delta_k)}^J$  denotes the density of  $Z^*(\Delta_k)$  conditioned to the presence of jumps.

Without loss of generality and in order to simplify our presentation, we will assume hereafter a regular grid  $\Delta_k = \Delta$  with k = 1, 2, ..., n and an absolutely continuous distribution for  $Z^*(\Delta_k)$ . With these assumptions the likelihood function becomes

$$L(\theta) = f_{X(0)}(x_0) \exp(n\lambda \Delta) \prod_{k=1}^{n} f_{Z^*(\Delta)}(\exp(\lambda \Delta) x_k - x_{k-1})$$
(6)

and the maximum likelihood inference can be performed by simply knowing the densities of D and  $Z^*(\Delta)$ .

In a few words, our methodology is based on the evaluation of the density  $f_{Z^*(\Delta)}$  via a discrete fast Fourier transform. To simplify matters, let us assume for the moment that the intensity parameter  $\lambda$  is known and we are interested in the vector parameter  $\theta$  indexing the stationary distribution D. Then we propose to consider the following steps

• Find an initial estimator  $\hat{\theta}_0$  of  $\theta$ . Using (3), we propose to consider the *n* independent and identically distributed random variables

$$Y_k = \int_{\lambda(k-1)\Delta}^{\lambda k\Delta} \exp(s) dZ(s) = \exp(\lambda \Delta) X(k\Delta) - X((k-1)\Delta)$$
 (7)

with k = 1, 2, ..., n, and obtain  $\hat{\theta}_0$  by matching for this sample the empirical and theoretical mean, variance, bias or kurtosis. The selection of these statistics or the consideration of higher order centered moments will depend on the dimension of  $\theta$ . In practice, this initial estimation can be improved by means of the generalized method of moments or another technique. In this work we have used for instance a modified differential evolution algorithm. This is a global optimization technique that has been successfully applied to different complex and multidimensional optimization problems. For details, we refer to [20].

- Use Proposition 3.3 to obtain the cumulant function of the BDLP  $\mathbf{Z}$  and the characteristic function of the Lévy functional  $Z^*(\Delta)$ . Note that this function can be used to evaluate all the required theoretical moments above.
- Use the preceding characteristic function and the classical or fractional discrete Fast Fourier transform to evaluate the density function  $f_{Z^*(\Delta)}$ . The procedure can be extended to  $f_{X(0)}$ , if D has not an explicit density function.
- Use a numerical method to optimize the likelihood function (6). Although a global optimization procedure can be used, like the differential evolution algorithm, it is better for speeding purposes to start with a good initial estimator  $\hat{\theta}_0$  and try then a local search algorithm around it.

If  $\lambda$  is not given, (5) suggests to estimate it by

$$\hat{\lambda}_0 = -\frac{\log(\hat{acf}(1))}{\Delta},$$

or by solving:

$$\hat{\lambda}_0 = \arg\min_{\lambda} \sum_{k=1}^m (\hat{acf}(k) - \exp(-\lambda k\Delta))^2,$$

where  $\hat{acf}(k)$  denotes the empirical autocorrelation function of lag k based on the data  $x_0, x_1, \ldots, x_n$ . As a criterion, we could truncate the sum above as soon as the empirical autocorrelation function reaches the level  $\frac{1.96}{\sqrt{n}}$ . This is the 95% upper confident limit interval for the autocorrelation function of a white noise process.

Once  $\hat{\lambda}_0$  is obtained, we could opt for two methods. The first is a plug in method consisting in following the scheme above after substituting  $\lambda$  with  $\hat{\lambda}_0$ . The second is a proper likelihood method that considers  $(\hat{\theta}_0(\hat{\lambda}_0), \hat{\lambda}_0)$  as an initial estimator in the search for  $(\theta, \lambda)$  that maximizes (6). All our results will be based on this last method.

We detail now our methodology with the  $D-{\rm OU}$  processes defined in Section 2. In what follows, we will make reference to the empirical mean  $\bar{Y} = \sum_{k=1}^n Y_k/n$ , variance  $S^2 = \sum_{k=1}^n (Y_k - \bar{Y})^2/n$ , bias  $B = \sum_{k=1}^n (Y_k - \bar{Y})^3/(nS^3)$  and kurtosis  $K = \sum_{i=1}^n (Y_i - \bar{Y})^4/(nS^4)$  of the random sample (7).

#### 4.1 Inference for a $\Gamma(a,b)$ -OU process

As explained, the initial estimators are found by matching

$$E[Y_1] = \frac{a(\exp(\lambda \Delta) - 1)}{b}$$
 and  $Var(Y_1) = \frac{a(\exp(2\lambda \Delta) - 1)}{b^2}$ 

with the empirical mean and variance. This yields  $\hat{\theta}_0 = (\hat{a}_0, \hat{b}_0)$ , where

$$\hat{a}_0 = \frac{\bar{Y}\hat{b}_0}{\exp(\lambda\Delta) - 1}$$
 and  $\hat{b}_0 = \frac{\bar{Y}(\exp(2\lambda\Delta) - 1)}{S^2(\exp(\lambda\Delta) - 1)}$ .

On the other side, by Proposition 3.3, the characteristic function of the Lévy functional  $Z^*(\Delta)$  equals:

$$\begin{split} \phi(\vartheta) &= \exp\left(\lambda \int_0^\Delta a (\frac{b}{b-\mathrm{i}\vartheta \exp(\lambda s)} - 1) \mathrm{d}s\right) \\ &= \exp\left(a\lambda b \left(\int_0^\Delta \frac{b}{b^2 + \vartheta^2 \exp(2\lambda s)} \mathrm{d}s + \mathrm{i} \int_0^\Delta \frac{\vartheta \exp(\lambda s)}{b^2 + \vartheta^2 \exp(2\lambda s)} \mathrm{d}s\right) - a\lambda \Delta\right) \\ &= \exp\left(a\log\left(\frac{b-\mathrm{i}\vartheta}{b-\mathrm{i}\vartheta \exp(\lambda\Delta)}\right)\right) \\ &= \left(\frac{b-\mathrm{i}\vartheta}{b-\mathrm{i}\vartheta \exp(\lambda\Delta)}\right)^a. \end{split}$$

This function is employed then to obtain

$$\phi^{J}(\vartheta) = \frac{\phi(\vartheta) - \exp(-\lambda a\Delta)}{1 - \exp(-\lambda a\Delta)},$$

which is the characteristic function associated to the conditional density  $f_{Z^*(\Delta)}^J$  introduced earlier in this section.

A peculiar property of this process emerges by observing that for any  $k \frac{y_k + x_{k-1}}{x_k} = \exp(\lambda \Delta)$ , being  $x_k$  and  $y_k$ , respectively, the observed values of the random variables  $X(k\Delta)$  and  $Y_k$  in (7). Hence, if the  $\Gamma(a,b)$ -OU process does not jump between the consecutive periods  $(k-1)\Delta$  and  $k\Delta$ , then  $\lambda$  can be exactly recovered by

$$\lambda = \frac{1}{\Delta} \log \left( \frac{x_{k-1}}{x_k} \right).$$

This happens finitely many times in the interval  $[0, n\Delta]$ .

# 4.2 Inference for a $TS(\kappa, a, b)$ -OU process

Since the stationary distribution of this process involves three parameters, we need to match now the theoretical mean  $E[Y_1] = 2a\kappa b^{\frac{\kappa-1}{\kappa}}(\exp(\lambda\Delta) - 1)$ , variance  $Var(Y_1) = 4a\kappa(1-\kappa)b^{\frac{\kappa-2}{\kappa}}(\exp(2\lambda\Delta) - 1)$  and bias  $\frac{E[(Y_1-E[Y_1])^3}{Var(Y_1)^{\frac{3}{2}}} = \frac{(2-\kappa)(\exp(3\lambda\Delta) - 1)}{\sqrt{ab\kappa(1-\kappa)}(\exp(2\lambda\Delta) - 1)^{\frac{3}{2}}}$  with the

corresponding empirical statistics. Solving these equations we obtain the moment-based initial estimator  $\hat{\theta}_0 = (\hat{\kappa}_0, \hat{a}_0, \hat{b}_0)$ , where

$$\hat{\kappa}_0 = \frac{B\bar{Y}(\exp(2\lambda\Delta) - 1)^2 - 2S(\exp(\lambda\Delta) - 1)(\exp(3\lambda\Delta) - 1)}{B\bar{Y}(\exp(2\lambda\Delta) - 1)^2 - S(\exp(\lambda\Delta) - 1)(\exp(3\lambda\Delta) - 1)},$$

$$\hat{b}_0 = (\frac{2\bar{Y}(1 - \hat{\kappa}_0)(\exp(2\lambda\Delta) - 1)}{S^2(\exp(\lambda\Delta) - 1)})^{\hat{\kappa}_0}$$

and

$$\hat{a}_0 = \frac{\bar{Y}\hat{b}_0^{\frac{1-\hat{\kappa}_0}{\hat{\kappa}_0}}}{2\hat{\kappa}_0(\exp(\lambda\Delta) - 1)}.$$

As it is easy to verify  $\hat{\kappa}_0 < 1$ , but no guarantee is given that  $\hat{\kappa}_0$  remains strictly positive. To circumvent the possibility that  $\hat{\kappa}_0 \leq 0$  we could opt for restricting the tempered stable distribution to an inverse Gaussian distribution ( $\kappa = \frac{1}{2}$ ) and using the estimators:

$$\hat{a}_0 = \frac{\hat{b}_0 \bar{Y}}{\exp(\lambda \Delta) - 1}$$

and

$$\hat{b}_0 = \frac{1}{S} \sqrt{\frac{\bar{Y}(\exp(2\lambda\Delta) - 1)}{\exp(\lambda\Delta) - 1}}.$$

Another alternative is to use the generalized method of moments.

On the other side, according to Proposition 3.3, the characteristic function of the Lévy functional  $Z^*(\Delta)$  is given by

$$\phi(\vartheta) = \exp\left(\lambda \int_{0}^{\Delta} C_{Z(1)}(\vartheta \exp(\lambda s)) ds\right)$$

$$= \exp\left(\lambda \int_{0}^{\Delta} 2i\vartheta \exp(\lambda s) a\kappa (b^{\frac{1}{\kappa}} - 2i\vartheta \exp(\lambda s))^{\kappa - 1} ds\right)$$

$$= \exp\left(-a\kappa \int_{b^{\frac{1}{\kappa}} - 2i\vartheta}^{b^{\frac{1}{\kappa}} - 2i\vartheta \exp(\lambda \Delta)} y^{\kappa - 1} dy\right)$$

$$= \exp\left(a\left((b^{\frac{1}{\kappa}} - 2i\vartheta)^{\kappa} - (b^{\frac{1}{\kappa}} - 2i\vartheta \exp(\lambda \Delta))^{\kappa}\right)\right).$$

# 4.3 Inference for a $NIG(\alpha, \beta, \delta)$ -OU process

Similarly to the tempered stable case, an initial estimator is obtained by matching  $E[Y_1] = \frac{\delta\beta\exp(\lambda\Delta)-1)}{\sqrt{\alpha^2-\beta^2}}, Var(Y_1) = \frac{\alpha^2\delta\exp(2\lambda\Delta)-1)}{\sqrt{(\alpha^2-\beta^2)^3}}$  and  $\frac{E[(Y_1-E[Y_1])^3}{Var(Y_1)^{\frac{3}{2}}} = \frac{3\beta\exp(3\lambda\Delta)-1)}{\alpha\sqrt{\delta(\alpha^2-\beta^2)^{\frac{1}{4}}}(\exp(2\lambda\Delta)-1)^{\frac{3}{2}}}$  with the empirical mean, variance and bias. The resulting moment-based initial estimator  $\hat{\theta}_0 = (\hat{\alpha}_0, \hat{\beta}_0, \hat{\delta}_0)$ , is given by:

$$\hat{\alpha}_0 = \frac{(\bar{Y}^2 + (\exp(\lambda \Delta) - 1)\hat{\delta}_0)^2)^{\frac{3}{2}}(\exp(2\lambda \Delta) - 1)}{\hat{\delta}_0^2 S^2(\exp(\lambda \Delta) - 1)^3},$$

$$\hat{\beta}_0 = \frac{\hat{\alpha}_0 \bar{Y}}{\sqrt{\bar{Y}^2 + ((\exp(\lambda \Delta) - 1)\hat{\delta}_0)^2}}$$

and

$$\hat{\delta}_0 = \sqrt{\frac{3S\bar{Y}(\exp(3\lambda\Delta) - 1)}{B(\exp(2\lambda\Delta) - 1)^2(\exp(\lambda\Delta) - 1)} - \left(\frac{\bar{Y}}{\exp(\lambda\Delta) - 1}\right)^2}.$$

If the term inside the last radical is negative, we propose to turn to the symmetric case  $(\beta = 0)$  and obtain the estimator by also matching the theoretical kurtosis

$$\frac{E[(Y_1 - E[Y_1])^4}{Var(Y_1)^2} = 3\left(1 + \frac{(\exp(\lambda\Delta) - 1)}{\alpha\delta(\exp(2\lambda\Delta) - 1)}\right)$$

with the empirical one. This leads us to the initial estimator

$$\hat{\alpha}_0 = \frac{\hat{\delta}_0(\exp(2\lambda\Delta) - 1)}{S^2}$$

and

$$\hat{\delta}_0 = \frac{S}{(\exp(2\lambda\Delta) - 1)} \sqrt{\frac{3(\exp(\lambda\Delta) + 1)}{K - 3}}.$$

Another alternative is to use the generalized method of moments.

On the other side, according to Proposition 3.3, the characteristic function of the Lévy functional  $Z^*(\Delta)$  is given by:

$$\phi(\vartheta) = \exp\left(\lambda \int_0^\Delta \frac{\delta i\vartheta \exp(\lambda s)(\beta + i\vartheta \exp(\lambda s))}{\sqrt{\alpha^2 - (\beta + i\vartheta \exp(\lambda s))^2}} ds\right)$$

$$= \exp\left(\delta \int_\vartheta^{\vartheta \exp(\lambda \Delta)} \frac{i(\beta + iy)}{\sqrt{\alpha^2 - (\beta + iy)^2}} dy\right)$$

$$= \exp\left(\delta(\sqrt{\alpha^2 - (\beta + i\vartheta)^2} - \sqrt{\alpha^2 - (\beta + i\vartheta \exp(\lambda \Delta))^2}\right).$$

# 5 On the simulation of a D-OU process

A general method to simulate a D-OU process is based on (3) and the Euler scheme:

$$X(k\Delta) = \exp(-\lambda \Delta) \left( X((k-1)\Delta) + \sum_{j=1}^{\left[\frac{\lambda \Delta}{\hbar}\right]} \exp(j\hbar) Z(\hbar) \right), \tag{8}$$

where k = 1, 2, ..., n,  $0 \le \hbar \le \lambda \Delta$  is sufficiently small, [x] denotes the maximum integer of x and X(0) is simulated from the stationary distribution D. This scheme is justified by the following result.

#### Proposition 5.1

$$\sum_{j=1}^{\left[\frac{\lambda\Delta}{\hbar}\right]} \exp(j\hbar) Z(\hbar) \xrightarrow{d} Z^*(\Delta) = \int_0^{\lambda\Delta} \exp(s) dZ(s) \quad as \quad \hbar \to 0,$$

where  $\stackrel{d}{\rightarrow}$  denotes convergence in distribution.

**Proof:** Let  $n = \left[\frac{\lambda \Delta}{\hbar}\right]$  be the maximum integer of  $\frac{\lambda \Delta}{\hbar}$  and consider the partition  $s_k = k\hbar$  with  $k = 0, 1, 2, \ldots, n + 1$ . Since the simple process

$$H_n(s) = \mathbf{1}_{\{0\}}(s) + \sum_{k=1}^{n+1} \exp((k-1)\hbar) \mathbf{1}_{](k-1)\hbar,k\hbar]}(s)$$

converges uniformly to the exponential function  $\exp(s)$  and, according to [14], **Z** is a semimartingale it follows that conforms  $\hbar \to 0$ ,  $\int_0^{\lambda\Delta} H_n(s) dZ(s)$  converges in probability to  $\int_0^{\lambda\Delta} \exp(s) dZ(s)$ . Proposition 5.1 follows then by the fact that

$$\int_0^{\lambda \Delta} H_n(s) dZ(s) = \sum_{k=1}^{n+1} \exp((k-1)\hbar) \left( Z(k\hbar) - Z((k-1)\hbar) \right)$$

$$= \sum_{j=1}^{n} \exp(j\hbar) \left( Z((j+1)\hbar) - Z(j\hbar) \right) \stackrel{d}{=} \sum_{j=1}^{n} \exp(j\hbar) Z(\hbar). \quad \Box$$

We discuss now in detail the simulation of the three particular D-OU processes introduced in Section 2.

# 5.1 Simulation of a $\Gamma(a,b)$ -OU process

As mentioned in Subsection 3.1, the BDLP  $\mathbf{Z} = \{Z(t)\}$  of this process is a compound Poisson process with intensity parameter a and associated exponential distribution with parameter b. Hence, it can be represented in law as:

$$Z(t) = \sum_{j=1}^{N(t)} Y_j,$$

where  $\mathbf{N}=\{N(t)\}$  is a Poisson process with intensity parameter a and  $Y_1,Y_2,...$  are independent and identically distributed exponential random variables with parameter b. Using this representation it is not difficult to see that the Lévy functional  $Z^*(\Delta)=\int_0^{\lambda\Delta}\exp(s)\mathrm{d}Z(s)$  is explicitly given by:

$$Z^*(\Delta) \stackrel{d}{=} \sum_{j=1}^{N(\lambda\Delta)} \exp(T_j) Y_j,$$

being  $T_1 < T_2 < T_3 \dots$  the arrival times of the Poisson process N. As a consequence, we obtain the following exact simulation scheme:

$$X(k\Delta) = \exp(-\lambda \Delta) \left( X((k-1)\Delta) + \sum_{j=1}^{N(\lambda \Delta)} \exp(T_j) Y_j \right),$$

where X(0) can be simulated from a  $\Gamma(a,b)$  distribution.

# 5.2 Simulation of a $T(\kappa, a, b)$ -OU process

As mentioned in Subsection 3.2, the BDLP  $\mathbf{Z} = \{Z(t)\}\$  of this process can be decomposed in law as:

$$Z(t) = I(t) + \sum_{j=1}^{N(t)} Y_j,$$

where  $\mathbf{I} = \{I(t)\}$  is a  $TS(\kappa, \kappa a, b)$ -Lévy process,  $\mathbf{N} = \{N(t)\}$  a Poisson process with intensity parameter  $ab\kappa$  and  $Y_1, Y_2, Y_3, \ldots$  are independent and identically distributed  $\Gamma(1 - \kappa, \frac{b^{\frac{1}{\kappa}}}{2})$  random variables. Hence, by (8) we could make use of the simulation scheme:

$$X(k\Delta) = \exp(-\lambda\Delta) \left( X((k-1)\Delta) + \sum_{j=1}^{\left[\frac{\lambda\Delta}{\hbar}\right]} \exp(j\hbar) I(\hbar) + \sum_{j=1}^{N(\lambda\Delta)} \exp(T_j) Y_j \right),$$

where  $T_1 < T_2 < T_3 < T_4 \ldots$  are the arrival times of the Poisson process  $\mathbb{N}$ ,  $\hbar$  is sufficiently small and X(0) is simulated from a  $TS(\kappa, a, b)$  distribution. This scheme results appropriate if  $\kappa = \frac{1}{2}$ , since here the inverse Gaussian random variable  $I(\hbar)$  can be easily simulated with the algorithm proposed by Michael et.al. [12]. Unfortunately, as already commented, the density function of a general tempered stable random variable  $I(\hbar)$  is not explicitly known and its series representation has no practical use. To circumvent this problem we invoke the method proposed by Rosiński [16, 17]. He shows that the process  $\{J_M(t)\}_{t\in[0,T]}$  with

$$J_M(t) = 2\sum_{j=1}^M \min\left\{ \left(\frac{a\lambda T\kappa}{b_j\Gamma(1-\kappa)}\right)^{\frac{1}{\kappa}}, \frac{e_j v_j^{\frac{1}{\kappa}}}{b^{\frac{1}{\kappa}}} \right\} \exp(\lambda T u_j) \mathbf{1}_{\{Tu_j \le t\}}$$

converges almost surely and uniformly as  $M \to \infty$  to the process  $\{J(t) = \int_0^{\lambda t} \exp(s) \mathrm{d}I(s)\}_{t \in [0,T]}$ . Here  $\{e_j\}$  is an independent and identically distributed sequence of exponential random variables with parameter 1,  $\{u_j\}, \{v_j\}$  are independent and identically distributed sequences of uniform random variables on the interval [0,1] and  $b_1 < b_2 < b_3, \ldots$  are the arrival times of a Poisson process with intensity parameter 1. All series are assumed to be independent of each other.

Note that the above method permits a simulation of the whole path of the process  $\{J(t)\}_{t\in[0,T]}$ . In addition to this, Rosiński [17] has also shown that the initial random variable X(0) can be approximated in law by

$$2\sum_{j=1}^{M} \min \left\{ \left( \frac{a}{b_j \Gamma(1-\kappa)} \right)^{\frac{1}{\kappa}}, \frac{e_j v_j^{\frac{1}{\kappa}}}{b^{\frac{1}{\kappa}}} \right\}. \tag{9}$$

As a result, we can rely on the simulation scheme

$$X(k\Delta) = \exp(-\lambda k\Delta) \left( X(0) + J_M(k\Delta) + \sum_{j=1}^{N(\lambda k\Delta)} \exp(T_j) Y_j \right),$$

where M is sufficiently large,  $T = n\Delta$  and X(0) can be simulated from (9).

### 5.3 Simulation of a $NIG(\alpha, \beta, \delta) - OU$ process

According to Barndorff-Nielsen [3] the BDLP of a  $NIG(\alpha, \beta, \delta) - OU$  process can be decomposed as the sum of a NIG- Lévy process, a compound Poisson process and a Lévy process for which only the characteristic function is known. This stops the possibility to use such decomposition as the base of a simulation scheme similar to the ones we have previously presented.

To simulate a  $NIG(\alpha, \beta, \delta) - OU$  process we will employ instead a lengthly but more general scheme that can be applied to any  $D-\mathrm{OU}$  process. This is based on (8) and a compound Poisson approximation of the BDLP  $\mathbf{Z}$  at  $\hbar$  (see [19] for details). Concretely, given  $0 < \epsilon < 1$  sufficiently small and the partition

$$a_0 < a_1 < a_2 < \ldots < a_k = -\epsilon < \epsilon = a_{k+1} < a_{k+2} < \ldots < a_{d+1}$$

 $Z(\hbar)$  can be approximated in law by

$$Z(\hbar) = \gamma_0 \hbar + \tilde{\sigma} W(\hbar) + \sum_{j=1}^{d} c_j \left( N_j(\hbar) - \hbar \tau_j \mathbf{1}_{|c_j| \le 1} \right), \tag{10}$$

where  $\tilde{\sigma} = \sqrt{\frac{2\alpha\delta\epsilon}{\pi}}$ ,  $\{W(t)\}$  is a standard Brownian motion,  $\{N_j(t)\}$  is a Poisson process with intensity parameter

$$\tau_j = \begin{cases} v_0([a_{j-1}, a_j]) &, \text{ for } 1 \le j \le k \\ v_0([a_j, a_{j+1}]) &, \text{ for } k+1 \le j \le d \end{cases}$$

and

$$c_{j} = \begin{cases} -\sqrt{\frac{1}{\tau_{j}} \int_{a_{j-1}}^{a_{j}} x^{2} dv_{0}(x)} &, \text{ for } 1 \leq j \leq k \\ \sqrt{\frac{1}{\tau_{j}} \int_{a_{j}}^{a_{j+1}} x^{2} dv_{0}(x)} &, \text{ for } k+1 \leq j \leq d. \end{cases}$$

The inclusion of the term  $\tilde{\sigma}$  was suggested by Asmussen and Rosiński [1] to improve the variability of the small jumps, whose sum is approximated above by the expected value.

As stated in Subsection 3.3 all the components in (10) can be calculated. For instance, the BDLP of the  $NIG(\alpha, \beta, \delta) - OU$  process has drift

$$\gamma_0 = \frac{2\delta\alpha}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x) dx - \int_{|x|>1} w(x) dx$$

and Lévy measure  $v_0(B) = \int_B w(x) dx$ , with

$$w(x) = (1 - \beta x)u(x) + \frac{\delta \alpha^2}{\pi} K_0(\alpha |x|) \exp(\beta x).$$

#### 6 Simulation results

This section presents the estimation results of the preceding methodology for a collection of data sets generated by the schemes in Section 5. For each distribution we simulated 100 sample paths on the interval [0, 10] using a regular partition of 1000 intervals. The procedures were programmed in Matlab using a 2.4 MHz PC. To cover processes with different degress of dependance, we fixed the intensity parameter  $\lambda$  in 0.5 and 10, respectively. For the  $\Gamma(a,b)$ -OU process we selected a=10 and b=5. For the  $TS(\kappa,a,b)$ -OU process, we set a=10, b=2 and two different values of  $\kappa$ ;  $\kappa=\frac{1}{2}$  and  $\kappa=\frac{2}{3}$ . Finally, for the  $NIG(\alpha,\beta,\delta)$ -OU process we set a=2, b=1 and  $\delta=5$ .

The optimization of the likelihood function was conducted with the Matlab code fmincon using the initial moment-based estimator proposed in Section 4 or an improvement as a starting point. In the tempered stable case, for instance, the moment-based estimator was improved by the application of a modified differential evolution algorithm having the likelihood as its objective function. The modification consisted on including the moment-based estimator in the initial population and on replacing unfeasible solutions by random parameters chosen from the admissible region. In all the cases the admissible region was set to pmin = [0.05, 0.05, 0.05, 0.05, 0.05] and pmax = [0.99, 100, 100, 100] and we set NP = 5, F = 0.8 and CR = 0.4 for the internal parameters in this procedure (see [20] for details). In the normal inverse Gaussian case, on the other hand, the moment-based estimator was improved by the use of the generalized method of moments. Only in cases of no convergence, we switched to the differential evolution algorithm.

As explained, the computation of the objective likelihood function was performed by means of the classical or the fractional discrete fast Fourier transform. To be precise, we applied the latter only in the normal inverse Gaussian process. The fractional technique, introduced by Bailey and Swarztrauber [2], turns out to be more convenient when, as it happens here, the associated  $Z^*(\Delta)$  characteristic function has a fast decay in the tails.

Throughout this section N will denote the number of terms in the discrete Fourier transform. Tables 1, 2, 3, 4 and 5 in appendix A report the mean, median, standard deviation (Std.Dev.) and root mean square error (RMSE) of the 100 moment-based initial estimations and final likelihood estimations for respectively the  $\Gamma(a,b)$ -OU,  $TS(\kappa,a,b)$ -OU and  $NIG(\alpha,\beta,\delta)$ -OU processes. Displayed on the last column appears also the mean, median and standard deviation of the total running time in seconds. In all the tables we include for reference the true parameter values.

As shown in Table 1, the moment-based initial estimators perform very well for both values of the intensity parameter  $\lambda$ . Recall that we are not including here results about  $\lambda$  since by the discussion in Subsection 4.1 this parameter can be exactly recovered from the data. The maximum likelihood estimations, on the other side, perform also very well despite a little bias in the parameter b for  $\lambda=0.5$ . However, as one can see from the standard deviation and specially from the root mean square error these estimations are more efficient than the initial ones. Table 1 reveals also the speed and stability of this procedure. With  $N=2^{15}$ , we did not find any problem in the convergence of the optimization algorithm.

Contrary to Table 1, the initial moment-based estimators in Table 2 perform very poorly. We found similar results with the generalized method of moments. On the other hand, the maximum likelihood estimations in Table 3 are remarkably good, especially for the intensity

parameter  $\lambda$ . The fact that the whole procedure takes in average between 1 and 2 minutes is partially due to the use of the intermediary differential evolution algorithm, the need to compute the tempered stable density when  $\kappa = \frac{2}{3}$  and the high concentration of the  $Z^*(\Delta)$  density in a neighborhood of 0 that demands at least a value  $N = 2^{15}$ . The latter problem aggravates if the parameter  $\kappa$  approaches to 0.

Although the  $NIG(\alpha, \beta, \delta)$ -OU and  $TS(\kappa, a, b)$ -OU processes share the same number of parameters, we see in Tables 4 and 5 that the estimation of the former is considerably faster than the one of the latter. This is due to the fast decay in the tails of the  $Z^*(\Delta)$  characteristic function and the use of the fractional discrete fast Fourier transform. Clearly the likelihood estimators outperform the moment-based estimators. This is more evident if  $\lambda = 0.5$ . The longer computation time in this last case is explained by the use of  $N = 2^{11}$  in contrast to the value  $N = 2^9$  for  $\lambda = 10$ .

#### 7 Conclusions

In this paper we have proposed a new methodology to obtain the maximum likelihood estimates of the parameters governing a one-dimensional stationary process of Ornstein-Uhlenbeck type, which is driven by a general Lévy process  $\mathbf{Z}$  and is constructed via a self-decomposable distribution D. By expressing the likelihood function of the sample of this process in terms of the density of the Lévy functional  $Z^*(\Delta) = \int_0^{\Delta} \exp(s) dZ(s)$ , our methodology is based on the inversion of the characteristic function of this variable and the use of the classical or fractional fast Fourier transform. Explicit details were given in the cases where D belongs to the family of gamma, tempered stable and normal inverse Gaussian distributions. Once we studied how to simulate these processes, we proceeded to apply our methodology to the simulated data. The results show that this is a fast, accurate and reliable methodology to estimate the parameters of the stationary distribution D and the intensity parameter  $\lambda$ . Out of the three studied processes, only the tempered stable distribution caused problems for a parameter  $\kappa$  approaching to 0. This is due to high concentration of the underlying density around 0 and the subsequent slow decay in the tails of the corresponding characteristic function.

This paper opens the possibility to extent our methodology to other similar processes and multivariate processes of Ornstein-Uhlenbeck type. Another interesting topic of further research is the study of the estimation problem when the underlying process is a convex linear combination of two or more independent D-OU processes (superposition). This represents a more challenging problem by virtue of the non-Markovian property of the resulting process.

#### References

- [1] S. Asmussen, J. Rosiński, Approximations of small jumps of Lévy processes with a view towards simulation, Journal of Applied Probability 38 (2001) 482-493.
- [2] D.H. Bailey, P. N. Swarztrauber, The Fractional Fourier Transform and applications, SIAM Review **33(3)** (1991) 389-404.

- [3] O.E. Barndorff-Nielsen, Processes of normal inverse Gaussian type, Finance and Stochastic 2 (1988) 41–68.
- [4] O.E. Barndorff-Nielsen, N. Shepard, Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics, Journal of the Royal Statistical Society B **63** (2001) 167–241.
- [5] O.E. Barndorff-Nielsen, N. Shepard, Normal modified stable processes, Theory of Probability and Mathematical Statistics 2 (2002) 1–20.
- [6] J. Cariboni, W. Schoutens, Jumps in Intensity models, Technical Report 1, University Centre for Statistics, K.U. Leuven, Leuven, Belgium, 2006.
- [7] C. Halgreen, Self-decomposability of the generalized inverse Gaussian and hyperbolic distributions, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **47** (1979) 13–18.
- [8] W. Jiang, J. Pedersen, Parameter Estimation for a Discretely Observed Stochastic Volatility Model with Jumps in the Volatility, Chinese Annals of Mathematics **24** (2003) 227–238.
- [9] G. Jongbloed, F.H. Van Der Meulen, A.W. Van Der Vaart, Nonparametric inference for Lévy-driven Ornstein-Uhlenbeck processes, Bernoulli 11 (2005) 759-791.
- [10] I. Karatzas, S.E. Shreve, Brownian Motion and Stochastic Calculus, second ed. Springer, New York, 1991.
- [11] E. Lukacs, Characteristic Functions, Charles Griffin and Co, London, 1970.
- [12] J.R. Michael, W.R. Schucany, R.W. Haas, Generating random variates using transformations with multiple roots, The American Statistician **30** (1976) 88–90.
- [13] K. Prause, The Generalized Hyperbolic model: estimation, financial derivatives, and risk measures, Phd thesis, Freiburg, 1999.
- [14] P. Protter, Stochastic Integration and Differential Equations, Springer-Verlag, Heidelberg, 1990.
- [15] S. Raible, Lévy processes in finance: theory, numerics, and empirical facts, Phd thesis, Freiburg, 2000.
- [16] J. Rosiński, Contribution to the discussion of a paper by Barndorff-Nielsen and Shephard, Journal of the Royal Statistical Society B **63** (2001) 230-231.
- [17] J. Rosiński, Tempered Stable Processes, in: O.E. Bandorff-Nielsen (Ed.), Miniproceddings of 2nd MaPhySto Conference on Lévy Processes: Theory and Applications, 2002, pp. 215-220.
- [18] K. Sato, Lévy Processes and Infinitely Divisible Distributions, Cambridge University Press, Cambridge, 1999.

- [19] W. Schoutens, Lévy Processes in Finance: Pricing Financial Derivatives, Wiley, Chichester, 2003.
- [20] R. Storn, K. Price, Differential Evolution a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, Journal of Global Optimization 11 (1997) 341 – 359.
- [21] M. Yamazato, Unimodality of infinitely divisible distribution functions of class L, Annals of Probability 6 (1978) 523–531.

# A Tables

Table 1: Parameter estimation for a  $\Gamma(a,b)$  – OU process.

	Initial estimations		Final estimations		Time
$\lambda = 0.5$	a = 10	b=2	a = 10	b=2	
Mean	10.3799	2.0688	10.2026	1.8946	5.7000
Median	10.2322	2.0495	10.3731	1.8630	4.7890
Std. Dev.	1.7720	0.4071	1.2864	0.2827	3.6370
RMSE	1.8036	0.4108	1.2959	0.3004	
$\lambda = 10$					
Mean	10.132	2.0234	10.005	1.9861	6.2006
Median	10.113	2.0156	10.028	1.977	5.57
Std. Dev.	0.5421	0.1100	0.3884	0.0750	2.3930
RMSE	0.5553	0.1119	0.3865	0.0759	

Table 2: Initial estimations for a  $TS(\kappa, a, b)-$  OU process.

			1	
	$\kappa = 0.5$	a = 10	b=2	$\lambda = 0.5$
Mean	0.6845	14.7926	2.4757	0.9581
Median	0.7512	4.5354	2.4083	0.8873
Std.Dev.	0.1882	70.9144	0.7504	0.4555
RMSE	0.2629	70.7215	0.8853	0.6445
	$\kappa = 0.5$	a = 10	b=2	$\lambda = 10$
Mean	0.4772	14.8061	1.9541	10.2701
Median	0.4960	10.1847	1.9838	10.2737
Std.Dev.	0.1183	18.3437	0.2283	1.2485
RMSE	0.1199	18.8739	0.2318	1.2713
	$\kappa = \frac{2}{3}$	a = 10	b=2	$\lambda = 0.5$
Mean	0.7684	11.1081	2.3976	1.0247
Median	0.7991	7.2036	2.2767	0.8740
Std.Dev.	0.1492	23.8775	0.6645	0.5562
RMSE	0.1800	23.7837	0.7715	0.7627
	$\kappa = \frac{2}{3}$	a = 10	b=2	$\lambda = 10$
Mean	0.6279	13.7993	1.9483	10.5372
Median	0.6533	10.1413	1.9432	10.6499
Std.Dev.	0.1022	27.5110	0.1766	1.3930
RMSE	0.1088	27.6355	0.1832	1.4865

Table 3: Final estimations for a  $TS(\kappa, a, b)-$  OU process.

	$\kappa = 0.5$	a = 10	b=2	$\lambda = 0.5$	Time
Mean	0.4954	10.4011	1.9632	0.4999	98.7351
Median	0.4954	10.2507	1.9847	0.4999	93.4300
Std.Dev.	0.0180	1.6749	0.4115	0.0004	30.2270
RMSE	0.0185	1.7141	0.4111	0.0004	
	$\kappa = 0.5$	a = 10	b=2	$\lambda = 10$	
Mean	0.4936	10.3267	1.9747	10.0204	68.7084
Median	0.4901	10.3179	1.9720	10.0002	64.7340
Std.Dev.	0.0288	1.4139	0.0698	0.1535	20.2560
RMSE	0.0293	1.4443	0.0739	0.1541	
	$\kappa = \frac{2}{3}$	a = 10	b=2	$\lambda = 0.5$	
Mean	0.6592	10.3704	1.9764	0.5070	106.6794
Median	0.6596	10.1474	1.9174	0.5070	100.8595
Std.Dev.	0.0238	2.1052	0.5222	0.0052	29.0779
RMSE	0.0248	2.1272	0.5201	0.0087	
	$\kappa = \frac{2}{3}$	a = 10	b=2	$\lambda = 10$	
Mean	0.6254	10.6873	1.9729	10.2271	78.7703
Median	0.6314	10.5098	1.9794	10.2220	72.2660
Std.Dev.	0.0337	1.3987	0.0963	0.4318	31.9978
RMSE	0.0532	1.5522	0.0996	0.4859	

Table 4: Initial estimations for a  $NIG(\alpha,\beta,\delta)-{\rm OU}$  process.

	$\alpha = 2$	$\beta = 1$	$\delta = 5$	$\lambda = 0.5$
Mean	4.5662	2.9456	4.3388	0.9490
Median	3.4839	1.7812	3.7967	0.8783
Std.Dev.	4.3160	4.1498	2.4137	0.5202
RMSE	5.0027	4.5644	2.4909	0.6852
	$\alpha = 2$	$\beta = 1$	$\delta = 5$	$\lambda = 10$
Mean	$\alpha = 2$ $1.9346$	$\beta = 1$ $0.9350$	$\delta = 5$ $4.8481$	$\lambda = 10$ $10.6849$
Mean Median		1-		
	1.9346	0.9350	4.8481	10.6849

Table 5: Final estimations for a  $NIG(\alpha,\beta,\delta)-{\rm OU}$  process.

	$\alpha = 2$	$\beta = 1$	$\delta = 5$	$\lambda = 0.5$	Time
Mean	2.0834	1.0198	5.1504	0.5016	4.7455
Median	2.0538	0.9739	5.1710	0.4995	3.3200
Std.Dev.	0.4516	0.3460	0.5298	0.0466	3.9205
RMSE	0.4571	0.3448	0.5482	0.0464	
	$\alpha = 2$	$\beta = 1$	$\delta = 5$	$\lambda = 10$	
Mean	1.9157	0.9521	4.6914	10.7067	1.0711
Median	1.9177	0.9315	4.6538	10.7491	1.0385
Std.Dev.	0.1722	0.1152	0.3918	0.7330	0.1784
RMSE	0.1909	0.1243	0.4972	1.0156	