Stat Infer Stoch Process (2009) 12:1–19
DOI 10.1007/s11203-008-9021-8

# Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type

**Luis Valdivieso · Wim Schoutens · Francis Tuerlinckx**

Received: 19 January 2007 / Accepted: 12 October 2007 / Published online: 29 April 2008
© Springer Science+Business Media B.V. 2008

**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

**JEL codes** 62M05 · 60G10 · 60H05 · 65T20

## 1 Introduction

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

$$\begin{cases} \mathrm{d}X(t) = -\lambda X(t)\mathrm{d}t + \mathrm{d}Z(t) \\ X(0) = X_0, \end{cases}$$

where $\mathbf{Z} = \{Z(t)\}$ is a Gaussian process such that $Z(t) \sim N(0, t\sigma^2)$, $\lambda$ 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 (see for instance (Kyprianou 2006; Sato 1999) or (Schoutens 2003)). 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

L. Valdivieso (✉) · W. Schoutens
Universitair Centrum voor Statistiek, K.U. Leuven, W. de Croylaan 54, 3001 Heverlee, Belgium
e-mail: lvaldiv@pucp.edu.pe

F. Tuerlinckx
Department of Phsychology, K.U. Leuven, Tiensestraat 102, 3000 Leuven, Belgium

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 (2001) 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 (2006) 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 (1988) who proposes a simulation-based likelihood approximation method when $D$ is an inverse Gaussian or a normal inverse Gaussian distribution. Jiang and Pedersen (2003) and Prause (1999) have also addressed this problem in the context of the Barndorff-Nielsen and Shepard model with 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 Jongbloed et al. (2005).

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 Sect. 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 Sect. 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 Sect. 5 we discuss the simulation of a D-OU process. Sect. 6 presents the simulation study and Sect. 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:

$$\begin{cases} \mathrm{d}X(t) = -\lambda X(t)\mathrm{d}t + \sigma \mathrm{d}W(t) \\ X(0) = X_0, \end{cases} \tag{1}$$

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 Karatzas and Shreve (1991), (1) has the unique strong solution:

$$X(t) = \exp(-\lambda t)\left(X_0 + \sigma \int_0^t \exp(\lambda s)\mathrm{d}W(s)\right).$$

Moreover, $\mathbf{X}$ is a continuous in probability Markov process with mean, variance and covariance functions:

$$E[X(t)] = \exp(-\lambda t)E[X_0],$$

$$Var(X(t)) = \frac{\sigma^2}{2\lambda} + \left(Var(X_0) - \frac{\sigma^2}{2\lambda}\right)\exp(-2\lambda t)$$

and

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

respectively. Observe that if $X_0 \sim N(0, \frac{\sigma^2}{2\lambda})$, $\mathbf{X}$ 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:

$$\begin{cases} \mathrm{d}X(t) = -\lambda X(t)\mathrm{d}t + \mathrm{d}Z(\lambda t) \\ X(0) = X_0, \end{cases} \tag{2}$$

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 Eq. 2, by allowing $\mathbf{Z}$ 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 $X(t_0), X(t_1), X(t_2), \ldots, X(t_n)$ be a sample of the stationary process defined by (2), $x_0, x_1, x_2, \ldots, x_n$ their observed values and $\Delta_k = t_k - t_{k-1}$. Thanks to the Markovian and Gaussian properties, the likelihood function of this sample is explicitly given by:

$$\begin{aligned} L(\theta) &= f_{X(t_0)}(x_0)\prod_{k=1}^{n} f_{X(t_k)|X(t_{k-1})=x_{k-1}}(x_k) \\ &= \frac{1}{\sigma\sqrt{\pi}}\exp\left(-\frac{x_0^2}{\sigma^2}\right)\prod_{k=1}^{n}\frac{\exp\left(\frac{-(x_k-\exp(-\lambda\Delta_k)x_{k-1})^2}{\sigma^2(1-\exp(-2\lambda\Delta_k))}\right)}{\sigma\sqrt{\pi(1-\exp(-2\lambda\Delta_k))}}. \end{aligned}$$

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

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

## 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 this 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

$$\begin{cases} \mathrm{d}X(t) = -\lambda X(t)\mathrm{d}t + \mathrm{d}Z(\lambda t) \\ X(0) = X_0, \end{cases}$$

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

Using the theory of integration of left-continuous predictable process with respect to semi-martingales and in particular to Lévy processes (see Protter 1990), we obtain for any $t \geq 0$ and $\Delta \geq 0$:

$$X(t) = \exp(-\lambda t)\left(X_0 + \int_0^t \exp(\lambda s)\mathrm{d}Z(\lambda s)\right),$$

or recursively

$$X(t+\Delta) = \exp(-\lambda\Delta)\left(X(t) + \exp(-\lambda t)\int_t^{t+\Delta} \exp(\lambda s)\mathrm{d}Z(\lambda s)\right). \tag{3}$$

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 $t \geq 0$ and $\Delta \geq 0$*

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

*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 first and the application of the integration by parts rule for semi-martingales (see Protter 1990). More precisely, we can write $\int_t^{t+\Delta} \exp(\lambda s)\mathrm{d}Z(\lambda s)$ as

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

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

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

Hence,

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

and applying once more the integration by parts rule

$$\exp(-\lambda t)\int_t^{t+\Delta}\exp(\lambda s)\mathrm{d}Z(\lambda s)\stackrel{d}{=}\int_0^{\Delta}\exp(\lambda s)\mathrm{d}Z(\lambda s). \qquad \square$$

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

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

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 (1999) has shown that any non-degenerate self-decomposable distribution $\mu$ is absolutely continuous and infinitely divisible. Then, the Lévy-Khintchine representation theorem guarantees the existence of a $\mu$-Lévy process with characteristic function $\phi_\mu^t$ at the time $t > 0$. In addition, as proved in Yamazato (1978), this process has a Lévy measure with a unimodal density that we will denote hereafter by $u$.

In relation to self-decomposability, Sato (1999) has also established the following result

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

$$\int_{|x|>2}\log(|x|)d\nu_0(x) < \infty, \tag{4}$$

*then* $\mathbf{X}$ *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 $\mathbf{X}$ 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

$$\sigma_0 = 2\sigma \ \text{ and } \ \gamma_0 = \gamma - \int_{-\infty}^{-1} w(x)\mathrm{d}x - \int_1^{\infty} w(x)\mathrm{d}x,$$

where $w$ is the Lévy density associated to $\nu_0$. In Barndorff-Nielsen (1988) it is shown that if the Lévy density associated to $\nu$, $u$, is differentiable then $w$ satisfies

$$\omega(x) = -u(x) - x\frac{\mathrm{d}u}{\mathrm{d}x}(x).$$

Furthermore we have the following key relation among the cumulants of $D$, $\mathbf{Z}$ and the Lévy functional $Z^*(\Delta)$ (see Barndorff-Nielsen (1988) 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* $C_{Z(1)}(\vartheta) = \vartheta \frac{dC_D}{d\vartheta}(\vartheta)$.

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:

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

As proved in Halgreen (1979), this is a self-decomposable distribution. Hence, by Proposition 3.1 the $\Gamma(a, b)$-OU process is well defined. Moreover, the BDLP $\mathbf{Z}$ 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, $\mathbf{Z} = \{Z(t)\}$ 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 $a > 0$, $b \geq 0$ 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 $g$. More precisely, the density function of this distribution is given by:

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

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:

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

A further analysis of the tempered stable family of distributions and its extension to the case $0 < \kappa < 2$ can be found in Rosiński (2007).

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

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

and then it follows by Lukacs (1970), Theorem 5.11.2, that the tempered stable distribution is self-decomposable.

In Barndorff-Nielsen and Shepard (2002) it is shown 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:

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

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:

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

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 $\mathbf{W}$ a standard Brownian motion. Moreover, $\mathbf{Y}$ has no Brownian motion part, drift $\gamma = \frac{2\delta\alpha}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x) \mathrm{d}x$ and Lévy measure $v(B) = \int_B u(x)\mathrm{d}x = \int_B \frac{\delta\alpha}{\pi|x|} K_1(\alpha|x|) \exp(\beta x)\mathrm{d}x$. On the other hand, Halgreen (1979) 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 $\mathbf{X}$. By the Markov property, the likelihood function of this sample is given by

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

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:

$$\begin{aligned} X(t_k) &= \exp(-\lambda\Delta_k)\left(X(t_{k-1}) + \exp(-\lambda t_{k-1})\int_{t_{k-1}}^{t_k} \exp(\lambda s)\mathrm{d}Z(\lambda s)\right) \\ &\stackrel{d}{=} \exp(-\lambda\Delta_k)\left(X(t_{k-1}) + Z^*(\Delta_k)\right), \end{aligned}$$

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

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 $\mathbf{Z}$ is a process of finite activity, $Z^*(\Delta_k)$ turns out to be a mixed random variable. For instance, if $\mathbf{Z}$ 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 $f_{X(t_k)|X(t_{k-1})=x_{k-1}}(x_k)$ as

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

where by $f^J_{Z^*(\Delta_k)}$ we denote the density of $Z^*(\Delta_k)$ conditioned to the presence of jumps.

Through this paper we will assume a regular grid $\Delta_k = \Delta$ with $k = 1, 2, \ldots, n$ and an absolutely continuous distribution for $Z^*(\Delta)$. With these assumptions the likelihood function becomes

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

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 things, 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

$$Y_k = \int_{\lambda(k-1)\Delta}^{\lambda k\Delta} \exp(s)\mathrm{d}Z(s) = \exp(\lambda\Delta)X(k\Delta) - X((k-1)\Delta) \tag{7}$$

  with $k = 1, 2, \ldots, n$, 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.
- 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 for example the differential evolution algorithm (see Storn and Price (1997) for details), it is recommended 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 $x_0, x_1, \ldots, x_n$ and the sum above is truncated at some integer value $m < n$.

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.

It is important to remark, that we are actually obtaining maximum likelihood estimates from a random sample of $Z^*(\Delta)$ and an independent sample from $D$ at time 0. As a consequence, the corresponding maximum likelihood estimator exhibits, under the usual regularity conditions, asymptotic unbiasedness, efficiency and normality.

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

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

As explained, the initial estimators are found by matching

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

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

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

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

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

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 $f^J_{Z^*(\Delta)}$ introduced earlier in this section.

A peculiar property of this process emerges by observing that for any $k = 1, 2, \ldots, n$, $\frac{y_k + x_{k-1}}{x_k} = \exp(\lambda\Delta)$, 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 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 $B(Y_1) = \frac{E[(Y_1 - E[Y_1])^3}{Var(Y_1)^{\frac{3}{2}}} = \frac{(2-\kappa)(\exp(3\lambda\Delta)-1)}{\sqrt{ab\kappa(1-\kappa)}(\exp(2\lambda\Delta)-1)^{\frac{3}{2}}}$ 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

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

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

Clearly $\hat{\kappa}_0 < 1$, 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:

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

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

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

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

Similarly to the tempered stable case, an initial estimator can be found 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 $B(Y_1) = \frac{3\beta\exp(3\lambda\Delta)-1)}{\alpha\sqrt{\delta(\alpha^2-\beta^2)^{\frac{1}{2}}}(\exp(2\lambda\Delta)-1)^{\frac{3}{2}}}$ 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 and empirical kurtosis

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

This leads us to the initial estimator

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

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:

$$\begin{aligned}\phi(\vartheta) &= \exp\left(\lambda\int_0^{\Delta}\frac{\delta \mathrm{i}\vartheta\exp(\lambda s)(\beta+\mathrm{i}\vartheta\exp(\lambda s))}{\sqrt{\alpha^2-(\beta+\mathrm{i}\vartheta\exp(\lambda s))^2}}\mathrm{d}s\right)\\ &= \exp\left(\delta\int_{\vartheta}^{\vartheta\exp(\lambda\Delta)}\frac{\mathrm{i}(\beta+iy)}{\sqrt{\alpha^2-(\beta+\mathrm{i}y)^2}}\mathrm{d}y\right)\\ &= \exp\left(\delta(\sqrt{\alpha^2-(\beta+\mathrm{i}\vartheta)^2}-\sqrt{\alpha^2-(\beta+\mathrm{i}\vartheta\exp(\lambda\Delta))^2}\right).\end{aligned}$$

## 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:

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

where $k=1,2,\ldots,n$, $0\le\hbar\le\lambda\Delta$ 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**

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

*where* $\xrightarrow{d}$ *denotes convergence in distribution.*

*Proof* Let $n=[\frac{\lambda\Delta}{\hbar}]$ 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 Protter (1990), $\mathbf{Z}$ is a semimartingale it follows that conforms $\hbar\to 0$, $\int_0^{\lambda\Delta}H_n(s)\mathrm{d}Z(s)$ converges in probability to $\int_0^{\lambda\Delta}\exp(s)\mathrm{d}Z(s)$. Proposition 5.1 follows then by the fact that

$$\begin{aligned}\int_0^{\lambda\Delta}H_n(s)\mathrm{d}Z(s)&=\sum_{k=1}^{n+1}\exp((k-1)\hbar)\,(Z(k\hbar)-Z((k-1)\hbar))\\ &=\sum_{j=1}^{n}\exp(j\hbar)\,(Z((j+1)\hbar)-Z(j\hbar))\overset{d}{=}\sum_{j=1}^{n}\exp(j\hbar)Z(\hbar).\end{aligned}$$

□

We discuss now in detail the simulation of the processes introduced in Sect. 2.

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

As mentioned in Subsect. 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 $Y_1, Y_2, \ldots$ are independent and identically distributed exponential random variables with parameter $b$. Using this representation it is not difficult to see that the Lévy functional $Z^*(\Delta) = \int_0^{\lambda\Delta} \exp(s)\mathrm{d}Z(s)$ can be explicitly written as:

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

being $T_1 < T_2 < T_3 \ldots$ the arrival times of the Poisson process $\mathbf{N}$. 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 Subsect. 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, $\mathbf{N} = \{N(t)\}$ is 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:

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

where $T_1 < T_2 < T_3 < T_4 \ldots$ are the arrival times of the Poisson process $\mathbf{N}$, $\hbar$ 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. (1976). 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 (2001, 2002, 2007). 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 a.s. 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 (2002) 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 (1988) 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-OU process. This is based on (8) and a compound Poisson approximation of the BDLP $\mathbf{Z}$ at $\hbar$ (see Schoutens 2003 for details). Concretely, given $0 < \epsilon < 1$ sufficiently small and the partition

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

$Z(\hbar)$ 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

$$c_j = \begin{cases} -\sqrt{\frac{1}{\tau_j}\int_{a_{j-1}}^{a_j} x^2 \mathrm{d}\nu_0(x)}, & \text{for } 1 \le j \le k \\ \sqrt{\frac{1}{\tau_j}\int_{a_j}^{a_{j+1}} x^2 \mathrm{d}\nu_0(x)}, & \text{for } k+1 \le j \le d. \end{cases}$$

The inclusion of $\tilde{\sigma}$ was suggested by Asmussen and Rosiński (2001) to improve the variability of the small jumps, whose sum in (10) is approximated by its expected value.

## 6 Simulation results

This section presents the estimation results of the preceding methodology for a collection of data sets generated by the schemes in Sect. 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 degrees 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 Sect. 4 or an improved estimator 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 with the likelihood as the 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 pmin =[0.05, 0.05, 0.05, 0.05] and pmax = [0.99, 100, 100, 100]. We set $NP = 5$, $F = 0.8$ and $CR = 0.4$ for the internal parameters of this procedure (see Storn and Price 1997 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 before, the computation of the objective likelihood function was conducted by means of the classical or the fractional discrete fast Fourier transform. To be precise, we applied the latter technique only for the normal inverse Gaussian process. The fractional technique, introduced by Bailey and Swarztrauber (1991), 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 1–5 report the mean, median, standard deviation (S.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 of Tables 1, 4 and 5 appears also the mean, median and standard deviation of the total running time in seconds.

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 Subsect. 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

**Table 1** Parameter estimation for a $\Gamma(a, b)$— OU process

| | Initial estimations | | Final estimations | | Time |
|---|---|---|---|---|---|
| | $a = 10$ | $b = 2$ | $a = 10$ | $b = 2$ | |
| $\lambda = 0.5$ | | | | | |
| 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 | 1.9770 | 5.5700 |
| 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

| | $\kappa = 0.5$ | $a = 10$ | $b = 2$ | $\lambda = 0.5$ |
|---|---|---|---|---|
| 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 |

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 optimization.

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

**Table 3** Initial estimations for a $NIG(\alpha, \beta, \delta)$-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 | 1.9346 | 0.9350 | 4.8481 | 10.6849 |
| Median | 1.8992 | 0.9091 | 4.7385 | 10.7420 |
| Std. Dev. | 0.2829 | 0.1927 | 0.6636 | 1.5041 |
| Rmse | 0.2890 | 0.2024 | 0.6775 | 1.6458 |

**Table 4** 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 | |

maximum likelihood estimations in Table 4 are remarkably good, especially for the intensity parameter $\lambda$. The fact that the whole procedure takes in average between 1 and 2 min is partially due to the use of the intermediary differential evolution algorithm, the need to compute the tempered stable density for $\kappa = \frac{2}{3}$ 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 when $\kappa \to 0$.

**Table 5** Final estimations for a $NIG(\alpha, \beta, \delta)$-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 | |

Although the $NIG(\alpha, \beta, \delta)$-OU and $TS(\kappa, a, b)$-OU processes share the same number of parameters, we see in Table 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 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 in terms of the density of the Lévy functional $Z^*(\Delta) = \int_0^{\lambda\Delta} \exp(s)\mathrm{d}Z(s)$, 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 causes 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, however, a more challenging problem by virtue of the non-Markovian property of the resulting process.

## References

Asmussen S, Rosiński J (2001) Approximations of small jumps of Lévy processes with a view towards simulation. J Appl Proba 38:482–493

Bailey DH, Swarztrauber PN (1991) The fractional fourier transform and applications. SIAM Rev 33(3): 389–404

Barndorff-Nielsen OE (1988) Processes of normal inverse Gaussian type. Finance Stochastic 2:41–68

Barndorff-Nielsen OE, Shepard N (2001) Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. J R Stat Soc B 63:167–241

Barndorff-Nielsen OE, Shepard N (2002) Normal modified stable processes. Theory Probab Math Stat 2:1–20

Cariboni J, Schoutens W (2006) Jumps in intensity models. Technical Report 1, University Centre for Statistics, K.U. Leuven, Belgium

Halgreen C (1979) Self-decomposability of the generalized inverse Gaussian and hyperbolic distributions. Z Wahrscheinlichkeit 47:13–18

Jiang W, Pedersen J (2003) Parameter estimation for a discretely observed stochastic volatility model with jumps in the volatility. Chi Ann Math 24:227–238

Jongbloed G, Van Der Meulen FH, Van Der Vaart AW (2005) Nonparametric inference for Lévy-driven Ornstein-Uhlenbeck processes. Bernoulli 11:759–791

Karatzas I, Shreve SE (1991) Brownian motion and stochastic calculus. Springer, New York

Kyprianou A (2006) Introductory lectures on fluctuations of Lévy processes with applications. Springer

Lukacs E (1970) Characteristic functions. Charles Griffin and Co, London

Michael JR, Schucany WR, Haas RW (1976) Generating random variates using transformations with multiple roots. Am Stat 30:88–90

Prause K (1999) The generalized hyperbolic model: estimation, financial derivatives, and risk measures. PhD thesis, Freiburg

Protter P (1990) Stochastic integration and differential equations. Springer-Verlag, Heidelberg

Raible S (2000) Lévy processes in finance: theory, numerics, and empirical facts. PhD thesis, Freiburg

Rosiński J (2001) Contribution to the discussion of a paper by Barndorff-Nielsen and Shephard. J R Stat Soc 63:230–231

Rosiński J (2002) Tempered stable processes. In: Bandorff-Nielsen OE (ed) Miniproceddings of 2nd MaPhySto Conference on Lévy processes: Theory and applications, pp 215–220

Rosiński J (2007) Tempering stable processes. Stoch Process Appl 117:677–707

Sato K (1999) Lévy processes and infinitely divisible distributions. Cambridge University Press

Schoutens W (2003) Lévy processes in finance: pricing financial derivatives. Wiley, Chichester

Storn R, Price K (1997) Differential evolution—a simple and efficient Heuristic for global optimization over continuous spaces. J Glob Optim 11:341–359

Yamazato M (1978) Unimodality of infinitely divisible distribution functions of class L. Ann Prob 6:523–531