DISA

Dipartimento di Informatica
e Studi Aziendali

2008/5

Characteristic function estimation of non-Gaussian Ornstein-Uhlenbeck processes

Emanuele Taufer

UNIVERSITÀ DEGLI STUDI
DI TRENTO

DISA Working Paper

DISA

Dipartimento di Informatica
e Studi Aziendali

2008/5

Characteristic function estimation of non-Gaussian Ornstein-Uhlenbeck processes

Emanuele Taufer

UNIVERSITÀ DEGLI STUDI
DI TRENTO

DISA WORKING PAPER

Ricardo Alberto MARQUES PEREIRA
Dipartimento di Informatica e Studi Aziendali
Università degli Studi di Trento
Via Inama 5, TN 38100 Trento ITALIA
Tel +39-0461-882147 Fax +39-0461-882124
E-mail: ricalb.marper@unitn.it

Characteristic function estimation of non-Gaussian Ornstein-Uhlenbeck processes

Emanuele Taufer*

Abstract

Continuous non-Gaussian stationary processes of the OU-type are becoming increasingly popular given their flexibility in modelling stylized features of financial series such as asymmetry, heavy tails and jumps. The use of non-Gaussian marginal distributions makes likelihood analysis of these processes unfeasible for virtually all cases of interest. This paper exploits the self-decomposability of the marginal laws of OU processes to provide explicit expressions of the characteristic function which can be applied to several models as well as to develop efficient estimation techniques based on the empirical characteristic function. Extensions to OU-based stochastic volatility models are provided.

Keywords: Ornstein-Uhlenbeck process, Lévy process, self-decomposable distribution, characteristic function, estimation.

MSC : 62F10, 62F12, 62M05.

1 Introduction

A continuous stationary process $\{X(t), t \geq 0\}$ is defined to be of the Ornstein-Uhlenbeck type (OU for short) if it is the solution of the stochastic differential equation

$$dX(t) = -\lambda X(t)dt + d\grave{Z}(t); \tag{1.1}$$

here $\lambda > 0$ and $\grave{Z}(t)$ is a homogeneous Lévy process, commonly referred to as the background driving Lévy process (BDLP), for which $\mathrm{E}[\log(1 + |\grave{Z}(1)|)] < \infty$. The modelling via the use of general Lévy processes, other than Brownian motion, allows to introduce specific non-Gaussian distributions for the marginal law of $X(t)$ and has received considerable attention in recent literature in an attempt to accomodate features such as jumps, semi-heavy tails and asymmetry which are well evident in real phenomena and are a point of remarkable interest in fields of application such as finance and econometrics.

Among recent contributions we find a completely new class of models, termed non-Gaussian Ornstein-Uhlenbeck models by which, stochastic processes with given correlation structure and (possibly non-Gaussian) marginal distribution are constructed; see Barndorff-Nielsen (1998, 2001), Barndorff-Nielsen et Al. (1998), Barndorff-Nielsen and Shephard (2001, 2003), Barndorff-Nielsen

*Department of Computer and Management Sciences, University of Trento, Via Inama 5, 38100 Trento - Italy, email:emanuele.taufer@unitn.it

and Leonenko (2005). Most notable examples include OU processes with marginal distributions such as the Normal Inverse Gaussian and the Inverse Gaussian (Barndorff-Nielsen, 1998), the Variance Gamma (Seneta, 2004), the Meixner (Schoutens and Teugel, 1998), the Normal and the Gamma. OU processes with positive jumps (subordinators) with marginal distributions such as the Inverse Gaussian are as well used as building blocks of stochastic volatility models (Barndorff-Nielsen and Shephard, 2001).

One of the key concepts related with these processes is that of self-decomposability. Recall that a random variable X is self-decomposable if, for all $c \in (0,1)$, there exists a characteristic function (ch.f.) $\phi_c(\zeta)$ such that the ch.f. of X, $\phi(\zeta)$, can be decomposed as $\phi(\zeta) = \phi(c\zeta)\phi_c(\zeta)$. Self-decomposability is closely related to stationary linear autoregressive time series of order 1, i.e. an AR(1) process. Essentially the only possible AR(1) processes are those for which the one-dimensional marginal law is self-decomposable and similarly for the OU process, i.e. an "AR(1)" in continuous time. The peculiar form of the ch.f. of $X(t)$ and its relation with the ch.f. of the underlying Lévy process has been exploited in Taufer and Leonenko (2008) to provide fast and reliable simulation procedures for OU processes. In this paper it will form the basis for providing explicit estimation procedures. With some more detail, it will be assumed that the marginal law of $X(t)$ belongs to a parametric family indexed by a parameter vector, and it will be discussed how to implement estimation of the parameters appearing in (1.1), including λ, by exploiting the empirical ch.f..

Efficient estimation of OU processes in the non-Gaussian case maybe quite cumbersome to implement. Usually one has availability of observations at discrete time points because continuous data are hardly available in practice or not economical to observe. Although in theory, a likelihood could be constructed by exploiting the markov property or independence of increments, one in practice has rarely availability of explicit or tractable expressions for the relevant densities. One way out is to resort to simulation based techniques, as it has been discussed in Barndorff-Nielsen (1998) or, in the context of OU stochastic volatility processes, by Roberts et al. (2004), Griffin and Steel (2006), Gander and Stephens (2007a, 2007b). The problem with these methods is that simulation of OU, and more generally, Lévy processes, is difficult owing to their jump character and one usually has to resort to approximations. Also, simulation based techniques are usually implemented more easily for subordinators rather than for general OU processes; for further details and references on simulation of Lévy processes see Todorov and Tauchen (2006). On the other hand, Jongbloed et al. (2005) and Jongbloed and Van der Meulen (2006) propose, respectively, non-parametric and parametric estimators of the Lévy density for subordinators; their estimators are based on the empirical cumulant function and in this sense they are close to those defined here. The approach of this paper is more general as it is not restricted to processe with positive jumps: by providing rules for easily constructing the estimators starting from the univariate ch.f. of $X(t)$, which is available in explicit form for several models of interest, a unified approach to ch.f. estimation of OU processes is introduced; previous literature concentrates mostly on processes with positive jumps, while the general case may turn quite complicated. Exploiting the peculiarity of self-decomposable random variables will allow to completely circumvent problems encountered with other techniques. Other works of interest here are those of Neumanna and Reiss (2008) which consider non-parametric estimation of the Lévy characteristic of general Lévy processes, Woerner (2004) which considers estimation of skewness parameter for Lévy processes and, for the Gaussian case, Baran, Pap, and van Zuijlen (2003), Gloter (2001), Florens-Landais and Pham (1999), Pap and van Zuijlen (1996) and the references therein.

Estimation based on the empirical ch.f. is well established and a good starting point is the paper of Feuerverger and McDunnough (1981) which show that arbitrarily high levels of efficiency can be obtained by such methods in the i.i.d. case; furthermore, they discuss the extension to dependent observations. Other references that are relevant to the subsequent development are those of Madan and Seneta (1987), Feuerverger (1990), Knight and Satchell (1997) and Knight and Yu (2002), Jiang and Knight (2002) which discuss in more depth empirical ch.f. estimation in a non i.i.d. setting. The contribution of this paper builds on previous literature in more ways: by implementing the procedure for a large variety of OU processes; by presenting a method that requires neither discretization nor simulation; by providing exact expressions for moments and ch.f. of OU processes; by comparing two approaches: one based on the bivariate ch.f of (1.1), the other based on the ch.f. of the 'error' term of the discretely observed process; this last approach may be seen as an attempt to substitue likelihood analisys of (1.1) based on i.i.d. observations.

This paper is organized as follows. The next section will provide some background information and some results on ch.f. of OU processes needed for estimation purposes. In section 3 the estimators are defined and their asymptotic properties studied. Section 4 is devoted to applications with simulated and real data. Section 5 discusses an extension of the ch.f. technique to OU based stochastic volatility models. The appendix provides the proofs of the results presented in the previous sections.

2 Characteristic functions and OU processes

The present section reviews some known results; for further details and generalizations the reader is referred to Wolfe (1982), Barndorff-Nielsen et Al. (1998), Barndorff-Nielsen (2001), Sato (1999). Equation (1.1) has a strong solution of the form

$$X(t) = e^{-\lambda t} X(0) + \varepsilon^{(\lambda)}(t). \tag{2.2}$$

where $\varepsilon^{(\lambda)}(t)$ is an error term, independent of $X(0)$, which is given by

$$\varepsilon^{(\lambda)}(t) = \int_0^t e^{-\lambda(t-s)} d\grave{Z}(s). \tag{2.3}$$

Define the ch.f. of $X(t)$ and $\varepsilon^{(\lambda)}(t)$ as

$$\phi(\zeta) = \mathrm{E}(e^{i\zeta X(t)}), \qquad \varphi_t(\zeta) = \mathrm{E}(e^{i\zeta\varepsilon^{(\lambda)}(t)}). \tag{2.4}$$

Also, let the cumulant functions of $X(t)$ and $\grave{Z}(1)$ be denoted by

$$\kappa(\zeta) = \log\phi(\zeta), \qquad \grave{\kappa}(\zeta) = \log \mathrm{E}(e^{i\zeta\grave{Z}(1)}). \tag{2.5}$$

If $X(t)$ is to be stationary, its ch.f. must have the form $\phi(\zeta) = \phi(e^{-\lambda t}\zeta)\varphi_t(\zeta)$ for all $t \geq 0$, hence $X(t)$ must be self-decomposable. The following result, due to Barndorff-Nielsen (1998) gives the exact relation between $\kappa(\cdot)$ and $\grave{\kappa}(\cdot)$ in order to have a specified marginal distribution for $X(t)$. It will be needed for estimation purposes.

Lemma 1. *(Barndorff-Nielsen, 1998, Theorem 2.3). Suppose that $\zeta\kappa'(\zeta)$ is continuous at 0, then the choice $\grave{\kappa}(\zeta) = \zeta\kappa'(\zeta)$ implies that for any $\lambda > 0$: i) $\exp\{\lambda\grave{\kappa}(\zeta)\}$ is an infinitely divisible ch.f.; ii) there is a process $\grave{Z}(t)$, say $\grave{Z}^{(\lambda)}(t)$, for which $\grave{Z}^{(\lambda)}(1)$ has ch.f. $\exp\{\lambda\grave{\kappa}(\zeta)\}$ such that a stationary version of $X(t)$ exists and has marginal distribution with ch.f. $\phi(\zeta)$.*

Recall that for the Lévy process $\grave{Z}(t)$ it holds that $\log \mathrm{E}(e^{i\zeta \grave{Z}(t)}) = t\grave{\kappa}(\zeta)$, from the discussion above we also note that $\grave{Z}^{(\lambda)}(t)$ is equal in distribution to $\grave{Z}^{(1)}(\lambda t)$ hence, equation (1.1) is often rewritten by substituting $d\grave{Z}(t)$ with $d\grave{Z}(\lambda t)$ so that it will turn out that whatever the value of λ, the marginal distribution of $X(t)$ is unchanged.

The stationary process $\{X(t), t \geq 0\}$ can be extended to a stationary process on the whole real line by introducing an independent càdlàg version of $\grave{Z}$, say $\grave{Z}^*$ and, for $t < 0$, define $\grave{Z}(t) = \grave{Z}^*(t)$ and let

$$X(t) = e^{-\lambda t} \int_{-\infty}^{t} e^{\lambda s} d\grave{Z}(\lambda s), \qquad t \in \mathbb{R}. \tag{2.6}$$

The following formula is of fundamental importance for working with ch.f. of OU processes and their functionals, for a proof see Barndorff-Nielsen and Shephard (2001).

Lemma 2. *Barndorff-Nielsen and Shephard (2001). For an integrable function with respect to $\grave{Z}$ we have*

$$\log \mathrm{E}[\exp\{i\zeta \int_0^\infty f(t)\grave{Z}(dt)\}] = \int_0^\infty \log \mathrm{E}[\exp\{i\zeta f(t)\grave{Z}(1)\}]dt. \tag{2.7}$$

Lemma 1 and Lemma 2 can be exploited to easily obtain explicit expressions for the ch.f. of (1.1) and their mutual relations. First, a result of Jurek and Vervaat (1983) states that a random variable X is self-decomposable if and only if it has a representation of the form $X = \int_0^\infty e^{-t} d\grave{Z}(t)$; application of (2.7) yields at once that

$$\kappa(\zeta) = \int_0^\infty \grave{\kappa}(e^{-s}\zeta)ds. \tag{2.8}$$

The above relation and formula (2.7) can be exploited to obtain a relation between $\kappa(\zeta)$ and $\log \varphi_t(\zeta)$. The next two results will be needed in our estimation procedure, we present them in the form of propositions.

Proposition 1. *The ch.f. $\varphi_t(\zeta)$ of $\varepsilon^{(\lambda)}(t)$ is*

$$\exp\{[\kappa(\zeta) - \kappa(\zeta e^{-\lambda t})]\}. \tag{2.9}$$

Proof. See the Appendix □

Note that from Proposition 1, by continuity, $\lim_{t\to\infty} \varphi_t(\zeta) = \phi(\zeta)$. There are a variety of relations of this flavour that can be obtained by exploiting formula (2.7) and the properties of OU processes. One that we will need, regards the joint ch.f. of $X(t_1), \ldots, X(t_m)$, $t_1, \ldots t_m \in \mathbb{R}$; from Barndorff-Nielsen and Leonenko (2005) it has the form

$$\mathrm{E}[\exp\{i[\zeta_1 X(t_1) + \cdots + \zeta_m X(t_m)]\}] = \exp\left\{\lambda \int_{\mathbb{R}} \grave{\kappa}\left(\sum_{j=1}^m \zeta_j \mathbb{I}_{(t_j > s)}(s) e^{-\lambda(t_j - s)}\right) ds\right\}. \tag{2.10}$$

As it is, this result is of little practical use, in order to proceed with ch.f.-based estimation we need to give it explicit form. We specialize it to the case $m = 2$ and $t_2 > t_1$, it will suffice for our purposes. We have

Proposition 2. *Let* $\psi(\zeta_1, \zeta_2) = \mathrm{E}[\exp\{i[\zeta_1 X(t_1) + \zeta_2 X(t_2)]\}$ *denote the joint ch.f. of* $X(t_1)$ *and* $X(t_2)$, $t_2 > t_1$. *Then*

$$\psi(\zeta_1, \zeta_2) = \exp\{\kappa(\zeta_1 - \zeta_2 e^{-\lambda(t_2-t_1)}) + \kappa(\zeta_2) - \kappa(\zeta_2 e^{-\lambda(t_2-t_1)})\}, \tag{2.11}$$

where $\kappa(\zeta)$ *is the cumulant function of* $X(t)$.

Proof. See the Appendix □

Remark 1. *A general form of* $\psi(\zeta_1, \zeta_2)$, *for* $t_1, t_2 \in \mathbb{R}$ *can be obtained by introducing* $t_* = \min(t_1, t_2)$, $t^* = \max(t_1, t_2)$ *and defining appropriate cases.*

Proposition 1 and Proposition 2 allow to easily obtain explicit forms for, respectively, $\varphi_t(\zeta)$ and $\psi(\zeta_1, \zeta_2)$ if an explicit expression for $\phi(\zeta)$ is available. This operational rule can be actually applied in several cases of interest and for models of wide applicability. Table (1) lists the cumulant function together with the parameter space and domain of some important examples by which OU processes with given marginal distribution can be constructed. Together with common distributions such as the Normal the t and the Gamma, we find some less known examples which find important applications in financial and econometric data: the Normal Inverse Gaussian and the Inverse Gaussian, see Barndorff-Nielsen (1997, 1998), Rydberg (1997), Barndorff-Nielsen and Shephard (2001); the Tempered Stable, see Tweedie (1984) and Hougaard (1986), of which the Inverse Gaussian is a special case when $\kappa = 1/2$; the Variance Gamma, discussed in Madan and Seneta (1990) and Seneta (2004); the Symmetric Gamma, for which the reader is referred to Dufresne (1997), Kotz et Al. (2001, p. 179) and Steutel and van Harn (2004, p. 504); the Euler's Gamma, see Grigelionis (2003); the Meixner, for which one can consult Schoutens and Teugel (1998) and Grigelionis (1999) and the Generalized z, (Grigelionis, 2001) of which the Meixner can be seen as a special case. The elegant expression for the ch.f. of the t-distribution has been provided by Heyde and Leonenko (2005). A general reference text on self-decomposable distributions and financial applications is Schoutens (2003).

3 Estimation

To enter the estimation problem, suppose that the marginal law of $X(t)$ is a member of a family of distributions indexed by a vector of parameters $\gamma \in \Gamma \subset \mathbb{R}^{p-1}$ which is our estimation goal. More generally, we may consider estimating the parameter vector $\theta \in \Theta \subset \mathbb{R}^p$ with $\theta = (\gamma', \lambda)'$ where $\lambda > 0$ is the auto-regression parameter of the OU process.

Suppose we observe the process at equispaced time points $0 < t_1 < \cdots < \dots t_n$ with $\Delta = t_j - t_{j-1}$, $j = 1, \dots n$, $t_0 = 0$. In order to slightly simplify notation, denote the observation at time t_j, $X(t_j)$, by X_j. It follows from the discussion in Wolfe (1982) that, for self-decomposable distributions, a discrete AR(1) process can be embedded into a continuous OU process. In our case, this amounts to say that the discretely observed OU process can be written as

$$X_j = e^{-\lambda\Delta} X_{j-1} + \varepsilon_j^{(\lambda)}, \tag{3.12}$$

where the $\varepsilon_j^{(\lambda)}$'s are *i.i.d.* random variables which are equal, in distribution, to the random variable $\varepsilon^{(\lambda)}(1)$, hence their ch.f. has the form $\exp\{\kappa(\zeta) - \kappa(\zeta e^{-\lambda\Delta})\}$. Note that we will not be able to distinguish between Δ and λ so that we will actually consider estimation of, say, $\lambda' = \lambda\Delta$ hence, from now on it will be assumed that $\Delta = 1$.

Name and Symbol	$\log \mathbf{E}(\exp\{\mathbf{i}\zeta\mathbf{X}\})$	**Constraints**
Normal $N(\mu,\delta)$	$\mu i\zeta - \delta\zeta^2/2$	$X \in I\!R$ $\mu \in I\!R$, $\delta > 0$
Normal Inv. Gaussian $NIG(\alpha,\beta,\mu,\delta)$	$i\mu\zeta - \delta(\sqrt{\alpha^2-(\beta+i\zeta)^2} - \sqrt{\alpha^2-\beta^2})$	$x \in I\!R$, $\mu \in I\!R$ $0 \leq \lvert\beta\rvert \leq \alpha$, $\delta > 0$
Inverse Gaussian $IG(\delta,\gamma)$	$-\delta(\sqrt{-2i\zeta+\gamma^2} - \gamma)$	$X > 0$, $\mu \in I\!R$ $\delta > 0$, $\beta > 0$
Tempered Stable $TS(\nu,\delta,\gamma)$	$\delta\gamma - \delta(\gamma^{1/\nu} - 2i\zeta)^\nu$	$X > 0$, $0 < \nu < 1$ $\delta > 0$, $\gamma \geq 0$
Gamma $\Gamma(\alpha,\beta)$	$\log\left(1 - \frac{i\zeta}{\beta}\right)^{-\alpha}$	$X > 0$ $\alpha > 0$, $\beta > 0$
Variance Gamma $V\Gamma(\alpha,\beta,\mu,\delta)$	$\mu i\zeta + 2\delta\log\left(\frac{\sqrt{\alpha^2-\beta^2}}{\sqrt{\alpha^2-(\beta+i\zeta)^2}}\right)$	$x \in I\!R$, $\mu \in I\!R$ $\delta > 0$, $\alpha > \lvert\beta\rvert > 0$
Symmetric Gamma $S\Gamma(\alpha,\beta)$	$\log\left(1 + \frac{\zeta^2}{\beta^2}\right)^{-\alpha}$	$X \in I\!R$ $\alpha > 0$, $\beta > 0$
Euler Gamma $E\Gamma(\alpha,\beta,\gamma,\delta)$	$\delta\left[\log\left(\Gamma(\gamma+i\alpha\zeta)/\Gamma(\gamma)\right) - i\alpha\zeta\log\beta\right]$	$X \in I\!R$, $\alpha \neq 0$ $\beta > 0$, $\gamma > 0$, $\delta > 0$
Meixner $M(\alpha,\beta,\mu,\delta)$	$i\mu\zeta + 2\delta\log\left(\frac{\cos(\beta/2)}{\cosh((\alpha\zeta-i\beta)/2)}\right)$	$x \in I\!R$, $\mu \in I\!R$, $\alpha > 0$ $-\pi < \beta < \pi$, $\delta > 0$
Generalized z $GZD(\alpha,\beta_1,\beta_2,\mu,\delta)$	$2\delta\left[\log\left(B(\beta_1 + \frac{i\alpha\zeta}{2\pi}, \beta_2 - \frac{i\alpha\zeta}{2\pi})/B(\beta_1,\beta_2)\right)\right] + i\mu\zeta$	$X \in I\!R$, $\alpha > 0$, $\mu \in I\!R$ $\beta_1 > 0$, $\beta_2 > 0$, $\delta > 0$
Symmetric scaled t $T(\nu,\delta,\mu)$	$i\mu\zeta + \log\left(\frac{K_{\nu/2}(\delta\lvert\zeta\rvert)}{\Gamma(\nu/2)}(\delta\lvert\zeta\rvert)^{\nu/2}2^{1-\nu/2}\right)$	$x \in I\!R$, $\mu \in I\!R$, $\delta > 0$ $\nu > 2$

Table 1: Self-decomposable distributions and their Ch.f.. $\Gamma(\cdot)$ and $B(\cdot,\cdot)$ denote, respectively, Euler's Gamma and Beta function while $K_\nu(\cdot)$ is the modified Bessel function of the third kind.

In this section we will slightly burden notation by introducing dependence on the parameters. So we will write $\phi_\gamma(\zeta)$ for $\mathrm{E}(e^{i\zeta X})$ and, suppressing dependence on t, $\varphi_\theta(\zeta)$ or $\varphi_{(\gamma,\lambda)}(\zeta)$ for $\mathrm{E}(e^{i\zeta\varepsilon^{(\lambda)}(1)})$ which, by Proposition 1, is $\phi_\gamma(\zeta)/\phi_\gamma(e^{-\lambda}\zeta)$; finally we write $\psi_\theta(\zeta_1,\zeta_2)$ for $\mathrm{E}[\exp\{i[\zeta_1 X_1 + \zeta_2 X_2]\}$.

The basic idea of the ch.f. estimation method is to minimize some (possibly weighted) distance function between an empirical estimator of the ch.f. and a model-based one. To enter in detail into the estimation problem, let us define first the empirical estimators of the ch.f. as

$$\psi_n(\zeta_1,\zeta_2) = \frac{1}{n-1}\sum_{j=1}^{n-1}\exp\{i[\zeta_1 X_j + \zeta_2 X_{j+1}]\}, \tag{3.13}$$

$$\varphi_{n,\lambda}(\zeta) = \frac{1}{n-1}\sum_{j=1}^{n-1}\exp\{i\zeta[X_{j+1} - e^{-\lambda}X_j]\}. \tag{3.14}$$

Note that $\varphi_{n,\lambda}(\zeta) = \psi_n(-e^{-\lambda}\zeta, \zeta)$. Next define the quantities.

$$Q_n^1(\theta) = \int\int_S |\psi_n(\zeta_1, \zeta_2) - \psi_\theta(\zeta_1, \zeta_2)|^2 dW(\zeta_1, \zeta_2), \tag{3.15}$$

$$Q_n^2(\theta) = \int_S |\varphi_{n,\lambda}(\zeta) - \varphi_\theta(\zeta)|^2 dW(\zeta). \tag{3.16}$$

The estimators are defined as the quantities minimizing the quadratic functions $Q_n^1(\theta)$ or $Q_n^2(\theta)$, i.e.

$$\hat{\theta}_1 = \arg\min_{\theta\in\Theta} Q_n^1(\theta), \qquad \text{and} \qquad \hat{\theta}_2 = \arg\min_{\theta\in\Theta} Q_n^2(\theta). \tag{3.17}$$

S denotes the region of integration and for convenience, according to the case at hand, may indicate either a subset of $\mathbb{R}$ or $\mathbb{R}^2$. In the case of $Q_n^2(\theta)$ which is one-dimensional, we may restrict attention the the case $S \subset \mathbb{R}^+$.

$W(\zeta_1, \zeta_2)$ and $W(\zeta)$ are to be considered weighting functions which may serve different purposes and characterize the estimation procedure. If they are chosen to be step functions we turn into a discrete set up which is much easier to implement from the computational point of view. Usually one chooses a grid of points $\zeta_1 \ldots \zeta_m$ at which the ch.f. are evaluated and then minimizes a sum instead of an integral. It is known that the choice of the grid of points has effect on the efficiency of the estimation procedure; Feuerverger and McDunnough (1981) show that, either in the i.i.d. or dependent case, using a weight given by the Fourier transform of the score and a grid sufficiently fine and extended the ch.f.-based estimation procedure is asymptotically equivalent to maximum likelihood estimation. In general the choice of an appropriate weight and grid may be impractical given the required quantities are seldom available explicitly and so one resorts to second best choices.

If we consider avoiding the choice of a grid and use a continuous version of $Q_n^1(\theta)$ and $Q_n^2(\theta)$, the choice of a proper weighting function is important in order to have a finite integral and for computational reasons since the procedure becomes cumbersome to implement. Knight and Yu (2002) discuss the use of an exponential weight (consider here $Q_n^2(\theta)$ for simplicity), i.e. $dW(\zeta) = \exp\{-\zeta^2\}d\zeta$, Epps (2005), in the context of goodness of fit testing, suggests using $dW(\zeta) = (|\varphi_\theta(\zeta)|^2/\int |\varphi_\theta(v)|^2 dv)d\zeta$. Both choices have the effect of damping out the persistent oscillations of $|\varphi_{n,\lambda}(\zeta) - \varphi_\theta(\zeta)|$, as $\zeta \to \infty$ assuring finiteness of the integral but may not be optimal for efficiency considerations. Note that this last weighting scheme depends on θ, while our proposed weights do not. Under suitably regularity conditions estimation based on $Q_n^1(\theta)$ and $Q_n^2(\theta)$ is equivalent to an estimating equation approach with weights depending on θ, for further details on equivalent approaches, see Knight and Yu (2002) and Jiang and Knight (2002).

Before discussing some of the properties of the above estimators let us note some of their features. $Q_n^1(\theta)$ is the estimator based on the bivariate ch.f. discussed in Feuerverger and McDunnough (1981) and Knight and Yu (2002), here we specialize it to the case of OU processes and give explicit solution for several cases of interest by exploiting Proposition 2.

$Q_n^2(\theta)$ is an estimator based on the ch.f. of the unobserved error term $\varepsilon^{(\lambda)}$ instead of that of the observed X_j. It exploits the peculiar relation between the ch.f. of X and that of $\varepsilon^{(\lambda)}$ and can be given explicit form by using Proposition 1, so it is applicable easily for all the models indicated in Table 1. Note that the auto-regression parameter λ is inserted into the empirical ch.f. estimator. Given the uniqueness of the Fourier-Stieltjes transformation, the amount of information contained in the ch.f. and the distribution function is the same. So this approach can be seen as an attempt to substitute likelihood analysis on the sequence of i.i.d. errors $\varepsilon^{(\lambda)}$ which otherwise would be possible essentially only in the Normal case.

This procedure however may have some drawbacks. First of all it may have identifiability problems. Consider for example the case where the marginal distribution of X is $N(0,\sigma^2)$, then $\phi(\zeta) = \exp\{-\sigma^2\zeta^2/2\}$ and consequently, $\varphi(\zeta) = \exp\{-\sigma^2(1-e^{-2\lambda})\zeta^2/2\}$ which is not identifiable for all possible values of σ^2 and λ. Next, the region of the λ values where Q_n^2 is minimized may be quite large, note in fact that, for $\lambda \to \infty$, we have that $\varphi_\theta(\zeta) = \psi_\theta(-e^{-\lambda}\zeta, \zeta) \to \phi_\gamma(\zeta)$ and $\varphi_{n,\lambda} \to (n-1)^{-1}\sum_j \exp i\zeta X_j$ and hence one can achieve very small values of Q_n^2 even with possibly wrong values of λ, this may be a problem for numerical minimization routines, indeed some simulations we run indicated that in certain instances, the algorithm converges to large values of λ. Third, for large data set the minimization procedure is numerically cumbersome as it requires to evaluate $\varphi_{n,\lambda}(\zeta)$ for each λ value in combination with each ζ value.

These features suggests that estimation of λ be undertaken exogenally; this makes the estimator considerably simpler to implement and computationally fast. We will then switch a little the problem, considering estimation of the marginal parameters of X by minimizing $Q_n^2(\gamma,\hat{\lambda})$ wrt γ, where $\hat{\lambda}$ is a square root-consistent estimator of λ. We define then the estimators

$$\hat{\gamma} = \arg\min_{\gamma\in\Gamma} Q_n^2(\gamma,\hat{\lambda}). \tag{3.18}$$

Note that this procedure exploits information on the structure of the process in order to estimate the marginal parameters and can lead to improved estimation procedures, as shown in Ghosh and Beran (2006) in the context of linear processes with long-range dependence.

The asymptotic theory for the estimators obtained from $Q_n^1(\theta)$ can be straightforwardly obtained from Knight and Yu (2002), under identifiability and regularity conditions stated there. Let $\theta_0 = (\gamma_0, \lambda_0)$ denote the true unknown parameter values.

Proposition 3. *Under regularity conditions (see Knight and Yu (2002)), $\hat{\theta}_1$ is consistent for θ_0 and*

$$\sqrt{n}\left(\hat{\theta}-\theta_0\right) \xrightarrow{D} \mathcal{N}\left(0, B(\theta_0)^{-1}A(\theta_0)B(\theta_0)^{-1}\right) \tag{3.19}$$

where, for $A_n(\theta) = \frac{1}{2}\frac{\partial Q_n^1(\theta)}{\partial\theta}$, $A(\theta) = \lim_{n\to\infty} n^{-1}\mathrm{E}\left[A_n(\theta)A_n(\theta)'\right]$ and

$$B(\theta) = \int\int_S \frac{\partial\psi_\theta(\zeta_1,\zeta_2)}{\partial\theta}\frac{\partial\psi_\theta(\zeta_1,\zeta_2)}{\partial\theta'}dW(\zeta_1,\zeta_2).$$

We need some more work to prove asymptotic results for $\hat{\gamma}$, in particular we need to add the condition that $\mathrm{E}X^{1+\delta} < \infty$ for some $\delta > 0$ in order to have convergence under exogenous estimation of λ.

Theorem 1. *Under conditions A1 - A6 (see the Appendix), $\hat{\gamma}$ is consistent for γ_0 and*

$$\sqrt{n}\left(\hat{\gamma}-\gamma_0\right) \xrightarrow{D} \mathcal{N}\left(0, B'(\theta_0)^{-1}A'(\theta_0)B'(\theta_0)^{-1}\right). \tag{3.20}$$

where, for $G_n(\theta) = \frac{1}{2}\frac{\partial Q_n^2(\gamma,\lambda)}{\partial\gamma}$, $A'(\theta) = \mathrm{E}\left[G_n(\theta)G_n(\theta)'\right]$ and

$$B'(\theta) = \int_S \frac{\partial\,\varphi_{\gamma,\lambda}(\zeta)}{\partial\gamma}\frac{\partial\,\varphi_{\gamma,\lambda}(\zeta)}{\partial\gamma'}dW(\zeta).$$

Remark 2. *As we will see in the Appendix, computation of $A'(\theta)$ does not require to consider covariance terms and hence turns out to be considerably simpler wrt $A(\theta)$.*

4 Applications

4.1 The Normal Inverse Gaussian OU process

As a first example of application consider the NIG-OU process introduced by Barndorff-Nielsen (1998). From Table 1 we see that the marginal law of the process is characterized by 4 parameters which makes it quite a flexible distribution. Here, δ is a scaling and μ a location parameter, whereas β is an asymmetry parameter and the quantities $\alpha \pm \beta$ determine the heavyness of the tails. Here we consider the case with $\mu = 0$. The law governing the BDLP of the NIG-OU process has been analitically derived by Barndorff-Nielsen (1998), in particular the BDLP is composed by three independent Lévy processes. Using simulation-based techniques for this case may turn an unwieldly task since one needs to (approximately) simulate three separate Lévy processes with jumps. Instead, it turns out to be quite simple to apply the procedure discussed here. In fact, from the univariate ch.f. of Table 1, for $\mu = 0$, Proposition 1 and Proposition 2 obtain that

$$\varphi_\theta(\zeta) = \exp\left\{ -\delta(\sqrt{\alpha^2 - (\beta + i\zeta)^2} - \sqrt{\alpha^2 - (\beta + ie^{-\lambda}\zeta)^2} \right\} \tag{4.21}$$

and

$$\psi_\theta(\zeta_1, \zeta_2) = \exp\{-\delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + i\zeta_2)^2} + \sqrt{\alpha^2 - (\beta + ie^{-\lambda}\zeta_2)^2} \\ - \sqrt{\alpha^2 - (\beta + i(\zeta_1 + e^{-\lambda}\zeta_2))^2}\}. \tag{4.22}$$

This example will be taken as an occasion to perform a small simulation study to compare the estimation procedures based on $Q_n^1(\theta)$, $Q_n^2(\gamma, \hat{\lambda})$ and moment-based estimators; the computing equations for these in the NIG case can be found, for example in Karlis (2002). We will consider an OU process with $\lambda = 1$ and a $NIG(\alpha = 2, \beta = 1, \delta = 1)$ marginal distribution. Several sample paths from such a process will be generated as suggested in Taufer and Leonenko (2008); it will be assumed that the sampling interval of the process is unity, i.e. $\Delta = 1$ and the case where $dW(\zeta)$ is a step function will be considered, turning in this way into a discrete estimation framework. The sampling intervals of the ch.f. will also be constant and set to $\zeta_j - \zeta_{j-1} = 0.05$, $j = 1, \dots m$. The choice of the step size maybe quite important in order to minimize effectively the distance of the empirical ch.f. and the model one. Feuerverger and McDunnough (1981) discuss optimizing procedures in order to obtain, for fixed sampling size of the ch.f., either an optimal fixed step or optimal sampling points $\zeta_1, \dots, \zeta_m$. As far as the number m of the sampling points, it was found to be very effective the heuristic approach of plotting $|\phi(\zeta)|^2$ against ζ and choose m, given the step, which samples the empirical ch.f. until it is sensibly different from 0. According to different cases, some calibration would be desirable, however a large simulation study is out of purpose here. The results of the simulations are in line with those of Feurverger and McDunnough (1981).

Table (2) gives the sample mean and the mean square error (MSE) of the moment-based estimators, indicated as ME, those based on $Q_n^1(\theta)$ and those based on $Q_n^2(\gamma, \hat{\lambda})$. The preliminary estimator of λ in the case of $Q_n^2(\gamma, \hat{\lambda})$ is the standard auto-correlation estimator between X_j and X_{j+1} which is connected with the maximum likelihood estimator of λ in the Normal case, this estimator has been used also in conjunction with ME. The results are based on 1000 replications of samples of sizes ranging from 500 to 10000. For each sample size, we reported in the line indicated with OV (overall variability), the algebric sum of the elements of the variance-covariance matrix of the estimators

		ME		$Q_n^2(\eta, \hat{\lambda})$		$Q_n^1(\theta)$	
n		Mean	MSE	Mean	MSE	Mean	MSE
500	$\alpha = 2$	6.127	567.969	2.104	0.14983	2.095	0.12538
	$\beta = 1$	4.878	562.614	1.024	0.03635	1.025	0.03070
	$\delta = 1$	1.122	0.10231	1.025	0.02358	1.028	0.02599
	$\lambda = 1$	1.020	0.01285	1.020	0.01285	1.070	0.01861
	OV		2257.16		0.34945		0.31289
1000	$\alpha = 2$	3.449	75.5322	2.051	0.06179	2.046	0.05353
	$\beta = 1$	2.303	74.0240	1.017	0.01868	1.016	0.01567
	$\delta = 1$	1.093	0.00549	1.009	0.00859	1.010	0.01059
	$\lambda = 1$	1.015	0.00702	1.015	0.00702	1.010	0.00928
	OV		298.813		0.14862		0.13735
5000	$\alpha = 2$	2.227	0.26587	2.014	0.01161	2.009	0.00996
	$\beta = 1$	1.180	0.15455	1.003	0.00344	1.003	0.00298
	$\delta = 1$	1.039	0.01485	1.002	0.00183	1.001	0.00214
	$\lambda = 1$	1.003	0.00116	1.003	0.00116	1.002	0.00166
	OV		1.03025		0.02775		0.02551
10000	$\alpha = 2$	2.130	0.12189	2.007	0.00556	2.004	0.00488
	$\beta = 1$	1.103	0.06852	1.003	0.00169	1.002	0.00143
	$\delta = 1$	1.023	0.00759	1.001	0.00089	1.000	0.00111
	$\lambda = 1$	1.002	0.00064	1.002	0.00064	1.001	0.00080
	OV		0.47849		0.01299		0.01229

Table 2: Sample Mean and MSE of estimates based on simulated data of NIG(2,1,1)-OU process with $\lambda = 1$, 1000 iterations.

corrected by the bias terms; it is an attempt to give a measure of the overall performance of each method in terms of MSE.

Placing attention on the results, we see that ME are highly unreliable for α and β for smaller sample sizes and the bias remains quite high even in large sample sizes. Estimation of δ and λ is much more reliable. Turning to the estimators based on the ch.f. we see both perform much better for all sample sizes, with very small bias even at smaller sample sizes. Overall it seems that the procedure based on $Q_n^1(\theta)$ is slightly superior even if differences are very small. At the level of single parameters $Q_n^1(\theta)$ is generally better for α and β whicle $Q_n^2(\gamma, \hat{\lambda})$ is generally better for δ and λ. Evidence seems to indicate that estimation of λ may be better if taken exogenally.

4.2 GDP growth rates

As an application to real data we consider estimation of the sequence of quarterly US GDP growth rates. It is defined as the logarithm difference of the real (in 2000 constant dollars), seasonally adjusted US GDP sequence from the first quarter of 1947 to the second quarter of 2007 for a total of 241 observations. The same sequence, but up to 2000, has been studied by Chan et al. (2004), together with other economic series with non-Normal errors. Chan et al. (2004) find that an $AR(1)$ model fits well to the data although they decidedly show non-Normality, they provide bootstrap procedures for quarterly intervals forecasts. Here, without the pretention to fully analyze the data set, and building on Chan et al. (2004), we proceed by adapting a discretely observed OU process with NIG marginal distribution which seems a good candidate given its flexibility in modelling data with asymmetry and/or heavier tails than the Normal. A quick moment estimates of the data reveals a slight asymmetry with a coefficient of 0.14 and an excess kurtosis of 1.37. Figure 1a shows a Q-Q plot of the data against a NIG model with $\alpha = 0.3468$, $\beta = 0.0500$, $\mu = 2.8950$, $\delta = 5.5160$, $\lambda = 1.1110$, the fit is remarkably good for experimental data, the parameter values have been determined by the two step Q_n^2 procedure. Figure 1b shows the plot of $|\varphi_{n,\hat{\lambda}}(\zeta)|^2$ and $|\varphi_{\hat{\theta}}(\zeta)|^2$ against ζ. One can note that adherence of the ch.f. of is remarkably good in all cases.

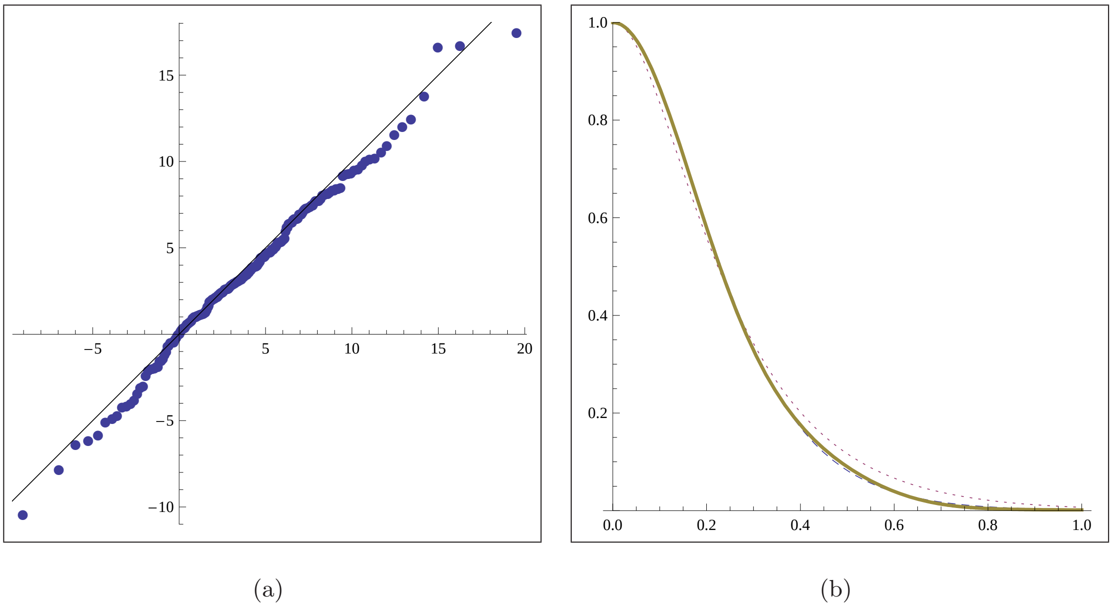


Figure 1: a) Q-Q plot of the data against a $NIG(0.3468, 0.0500, 2.8950, 5.5160)$ distribution; b) Plot of $|\varphi_{n,\hat{\lambda}}(\zeta)|^2$ (solid), $|\varphi_{\hat{\theta}}(\zeta)|^2$, $\hat{\theta}$ from Q_n^1, (dotted) and $|\varphi_{\hat{\gamma},\hat{\lambda}}(\zeta)|^2$, $\hat{\gamma}$ from Q_n^2, (dashed) against ζ, $\hat{\lambda}$ auto-correlation estimator.

5 An extension to Stochastic volatility models

Consider the following asset return process with stochastic conditional volatility of the log-asset price $S(t)$

$$dS(t) = \{\mu + \beta X(t)\}dt + \sqrt{X(t)}dN(t) \tag{5.23}$$

where $X(t)$ is a non-negative OU process as defined in (1.1) and $N(t)$ is a standard Brownian motion independent of $X(t)$, μ is a drift and β is a risk premium parameter. This model was introduced in Barndorff-Nielsen and Shephard (2001) with the aim of modelling stylized features of financial markets while maintining analytical tractability. The positive OU process $X(t)$, with no Gaussian component, moves entirely by jumps and decays exponentially between two jumps while $S(t)$ remains continuous; to introduce discontinuities one can introduce a Lévy component in equation (5.23). Suppose we observe the process $S(t)$ at fixed time instants $t_0 < t_1 < \cdots < t_m$ and define $\Delta S(t_j) = S(t_j) - S(t_{j-1})$, then $\Delta S(t_j)|\Delta x^*(t_j) \sim N(\mu(t_j - t_{j-1}) + \beta x^*(t_j), x^*(t_j))$ where $\Delta X^*(t_j) = X^*(t_j) - X^*(t_{j-1})$ and $X^*(t_j) = \int_0^{t_j} X(s)ds$ is the integrated volatility; $\Delta x^*(t_j)$ is termed actual volatility. The terms $\Delta S(t_j)|\Delta x^*(t_j)$ $j = 1 \dots n$ are conditionally independent and we note that the distribution of $\Delta S(t_j)$ will be a location scale mixture of normals. By an appropriate design of the stochastic process for $X(t)$ one can allow aggregate returns to be heavy tailed, skewed and exhibit volatility clustering. Following the discussion of Section 2, typical choices for the marginal distribution of $X(t)$ are the Inverse Gaussian and Gamma distributions, alternatively one may model directly the BDLP of $X(t)$ obtaining a variety of models which can adapt very well to practical situations. For further details on these models one can refer to Barndorff-Nielsen and Shephard (2001, 2003).

Estimating the parameters of a continuous stochastic volatility model is difficult owing to the inability to compute the appropriate likelihood function. Model-based estimation approaches are based on MCMC methods and references to recent related works are those of Barndorff-Nielsen and Shephard (2001), Gander and Stephens (2006, 2007) Griffin and Steel (2006). Alternatively one might consider non-model-based estimation approaches which exploits realized volatility, i.e. use the existence of high-frequency intraday data to directly estimate moments of integrated volatility, which was introduced by Barndorff-Nielsen and Shephard (2002), for extensions on this approach and further references to recent works one can consult Woerner (2007).

Here we sketch shortly how we could implement ch.f. based estimation. This can be done quite easily as explicit expressions may be derived with little work. Asimptotic properties of these estimators can be obtained by the method discussed in Knight et al. (2002) which has studied ch.f. based estimation of a stochastic volatility model based on a Gaussian volatility model.

Theorem 2 obtains the joint characteristic function of the increments of the stochastic volatility process in term of the cumulant generating function of the BDLP $\grave{Z}(1)$ of $X(t)$, this exists in closed form for key cases such as the Inverse Gaussian and the Gamma and allows to obtain results in closed form solution.

Theorem 2. *Let $k(\zeta) = \log \mathrm{E} e^{-\zeta \grave{Z}(1)}$, the joint characteristic function of $\Delta S(t_1), \dots \Delta S(t_m)$ is*

$$\exp\{i\mu \sum_{j=1}^{m} \zeta_j (t_j - t_{j-1})\} \exp\{\lambda \int_0^{\infty} k(J e^{-\lambda s}) ds\} \exp\{\lambda \sum_{l=0}^{m} \int_{t_{l-1}}^{t_l} k(H_l(s)) ds\}, \tag{5.24}$$

where

$$J = \sum_{j=1}^{m} \left(\frac{1}{2}\zeta_j^2 - i\beta\zeta_j \right) \varepsilon_\lambda(t_{j-1}, t_j), \tag{5.25}$$

$$H_l(s) = \sum_{j=l+1}^{m} \left(\frac{1}{2}\zeta_j^2 - i\beta\zeta_j \right) \varepsilon_\lambda(t_{j-1}, t_j) e^{\lambda s} + \mathbb{I}_{(l>0)} \left(\frac{1}{2}\zeta_l^2 - i\beta\zeta_l \right) \varepsilon_\lambda(0, t_l - s). \tag{5.26}$$

With $\varepsilon_\lambda(u, v) = \frac{1}{\lambda}(e^{-\lambda u} - e^{-\lambda v})$ and the convention that $t_{-1} = 0$ and $\sum_{j=m+1}^{m} f(j) = 0$.

5.1 Marginal models for $X(t)$

To get more specific in our modelling framework we present two important models for the marginal distribution of the volatility process which allow to obtain a variety of behaviours for the process $S(t)$.

The expressions for the ch.f. we have provided are in terms of $k(\zeta) = \log \mathrm{E}\{\exp(-\zeta \grave{Z}(1))\}$ hence one can model directly from the BDLP if has a direct expression for this. Alternatively if one prefers modelling by first specifying a marginal distribution for the process $X(t)$ and exploiting the relation

$$\grave{\kappa}(\zeta) = \zeta \frac{\partial \kappa(\zeta)}{\partial \zeta} \tag{5.27}$$

and the fact that $k(\zeta) = \kappa(-i\zeta)$.

A first important model for X is the Tempered Stable, see Table 1, i.e. $X \sim TS(\nu, \delta, \gamma)$. When $\nu = 1/2$ we have the important sub-case of the Inverse Gaussian distribution. For the TS case, exploiting (5.27) we obtain

$$k(\zeta) = -2\zeta\nu\delta^{2\nu}(\gamma^2 + 2\zeta)^{\nu-1}. \tag{5.28}$$

For this case, the ch.f. of $\Delta S(t_1), \ldots, \Delta S(t_m)$ can be given explicit form in terms of hypergeometric series. Here, to avoid large expressions we report the IG case for $m = 1$, let AT denote the ArcTanh function and $t_j - t_{j-1} = \Delta$, we have

$$C\{\zeta \ddagger \Delta S(t_j)\} = \frac{2\delta\left(\frac{\zeta^2}{2} - i\zeta\beta\right)\left[AT\left((B^{-1})\sqrt{1 + \frac{2(1-e^{-\lambda\Delta})(\zeta^2/2 - i\zeta\beta)}{\gamma^2\lambda}}\right) - AT(B^{-1})\right]}{\lambda\gamma B} \tag{5.29}$$

where $B = \sqrt{1 + \frac{2(\zeta^2/2 - i\zeta\beta)}{\gamma^2\lambda}}$. As we see, although cumbersome, the above function is straightforwardly applicable in numerical procedures.

Another important case is the Generalized Inverse Gaussian model, which we denote by $X \sim GIG(\nu, \delta, \gamma)$, with $\delta \geq 0, \gamma > 0$ if $\nu > 0$, $\delta > 0, \gamma > 0$ if $\nu = 0$ and $\delta > 0, \gamma \geq 0$ if $\nu < 0$. For details see Eberlein (2001) and Barndorff-Nielsen and Shephard (2001b); this model has ch.f.

$$\phi(\zeta)_{GIG(\nu,\delta,\gamma)} = \left(\frac{\gamma^2}{\gamma^2 - 2\zeta}\right)^{\nu/2} \frac{K_\nu(\delta\sqrt{\gamma^2 - 2\zeta})}{K_\nu(\delta\gamma)}. \tag{5.30}$$

where $K_\nu(x)$ denote the modified Bessel function of the third kind. The appropriate form of $k(\zeta)$ for the GIG model, not shown here, can be recovered by (5.27). Note that

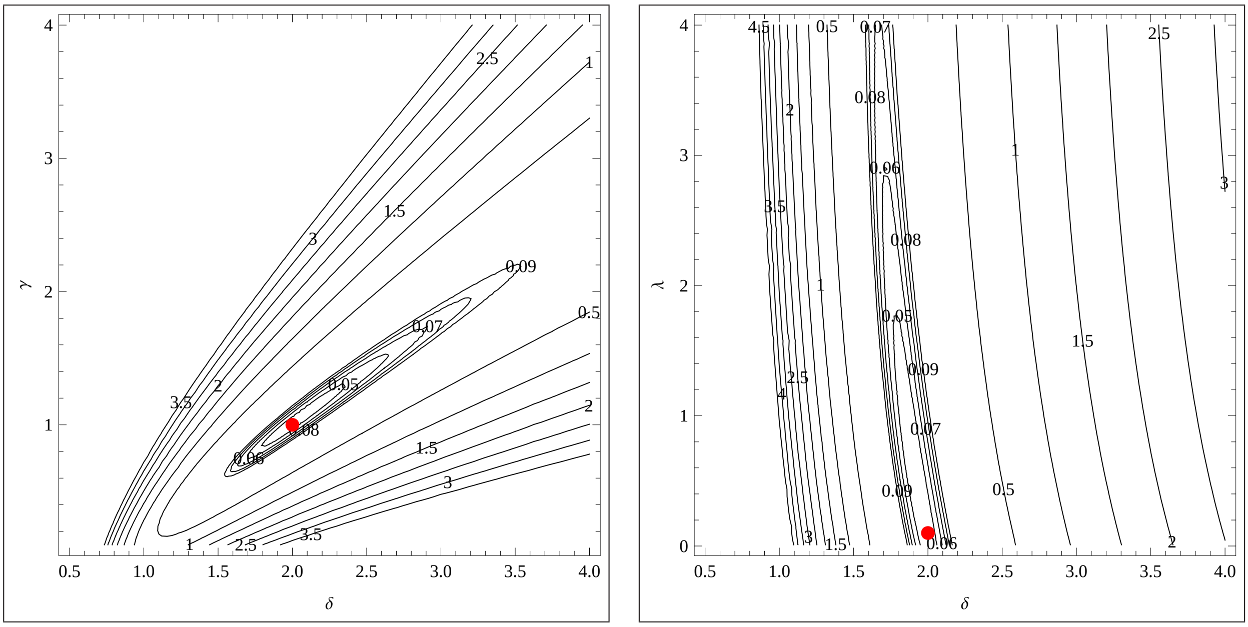
$$\frac{\partial K_\nu(x)}{\partial x} = -\frac{1}{2}(K_{\nu-1}(x) + K_{\nu+1}(x)). \tag{5.31}$$

In the limiting case $\nu > 0$, $\delta = 0$ it reduces to the density of a Gamma distribution $\Gamma(\gamma^2, \nu/2)$; for $\nu < 0$, $\gamma = 0$ one gets those of a reciprocal Gamma distribution. For the Gamma case we have the simple form

$$k(\zeta) = \frac{\nu\zeta}{2(\gamma^2 + \zeta)}. \tag{5.32}$$

Notably the IG distribution is common either to the TS and to the GIG distribution, being $IG(\delta,\gamma) \sim TS(1/2.\delta,\gamma) \sim GIG(-1/2,\delta,\gamma)$.

In the cases where an explicit form of the ch.f. of the process (5.23) cannot be obtained, it appears that (5.24) is well suited for numerical integration. Also, by using the same techniques discussed in Theorem 2 one can provide expressions for the ch.f. of stochastic volatility models with leverage and where superpositions of OU processes are used to model volatility. It appears that a thorough investigation of these models would require much additional work which seems suitable for a separate paper.



(a) (b)

Figure 2: Contour plots of the objective function for model (5.23) with $X(t)$ following and IG-OU process with $\delta = 2$, $\gamma = 1$, $\lambda = 0.01$. Lenght of series n=2500. a) (δ,γ); b) (δ,λ).

Here we give a brief report on some results with simulated data by estimating the paramters for a series of lenght 2500 from model (5.23) with $\mu = \beta = 0$ and where $X(t)$ follows and $IG(2,1)$ OU process with $\lambda = 0.01$. Figure 2 provides the contour plots of the objective function based on $Q_n^1(\theta)$ obtained by using (5.24) with $m = 2$ and the empirical estimator $\psi_n(\zeta_1,\zeta_2)$ with 400 pairs (ζ_1,ζ_2) generated from two independent $N(0,3)$ distributions.

Although the plots represent a single simulation, it appears that the objective function has quite a regular behaviour with very low values in the neighborhood of the true values of the parameters.

6 Appendix: Proofs

Proof of Proposition 1. Let $I\!I$ denote the indicator function; let $f(t) = e^{-\lambda(t-s)} I\!I_{(0,t]}(s)$ and apply formula (2.7) to obtain

$$
\begin{aligned}
\log \varphi_t(\zeta) &= \lambda \int_0^\infty \grave{\kappa}\left(e^{-\lambda(t-s)} I\!I_{(0,t]}(s)\zeta\right) ds \\
&= \lambda \int_0^t e^{-\lambda(t-s)} \zeta\, \kappa'\left(e^{-\lambda(t-s)}\zeta\right) ds \\
&= \kappa(\zeta) - \kappa(\zeta e^{-\lambda t})
\end{aligned}
$$

from which the result follows. □

Proof of Proposition 2. From (2.10), for $t_2 > t_1$, write

$$
\begin{aligned}
\log \psi(\zeta_1, \zeta_2) &= \lambda \int_{I\!R} \grave{\kappa}\left(\sum_{j=1}^{2} \zeta_j I\!I_{(t_j > s)}(s) e^{-\lambda(t_j - s)}\right) ds \\
&= \lambda \int_{-\infty}^{t_1} \grave{\kappa}\left(\sum_{j=1}^{2} \zeta_j e^{-\lambda(t_j - s)}\right) ds + \lambda \int_{t_1}^{t_2} \grave{\kappa}\left(\zeta_2 e^{-\lambda(t_2 - s)}\right) ds
\end{aligned}
$$

next use, as in Proposition 1, the relation $\grave{\kappa}(\zeta) = \zeta\kappa'(\zeta)$ to formally integrate and obtaining the result. □

The following technical result will be needed in the proof of Theorem 1.

Lemma 3. *Let $\lambda^* \xrightarrow{p} \lambda$ as $n \to \infty$ and assume that $\mathrm{E}|X|^{1+\delta} < \infty$ for some $\delta > 0$. Then*

$$
i)\quad \varphi_{n,\lambda^*}(\zeta) \quad \xrightarrow{p} \quad \varphi_\theta(\zeta), \qquad and \qquad ii)\quad \frac{\partial}{\partial\lambda}\varphi_{n,\lambda}(\zeta)\big|_{\lambda=\lambda^*} \quad \xrightarrow{p} \quad \frac{\partial}{\partial\lambda}\varphi_\theta(\zeta), \tag{6.33}
$$

where $\frac{\partial}{\partial\lambda}\varphi_\theta(\zeta) = i\zeta e^{-\lambda}\varphi_\theta(\zeta)\mathrm{E}(X)$.

Proof of Lemma 3. Denote by $o_p(1)$ a quantity converging in probability to 0 as $n \to \infty$ and recall that $|e^{iu}| = 1$. To prove *i)* we first prove that we can simply consider the empirical estimator of the ch.f. evaluated at λ instead of λ^*. Note that

$$
\begin{aligned}
\left|\varphi_{n,\lambda^*}(\zeta) - \varphi_{n,\lambda}(\zeta)\right| &= \left|\frac{1}{n-1}\sum_{j=1}^{n-1}\left(e^{i\zeta(X_{j+1} - e^{-\lambda^*}X_j)} - e^{i\zeta(X_{j+1} - e^{-\lambda}X_j)}\right)\right| \\
&\le \frac{1}{n-1}\sum_{j=1}^{n-1} |e^{i\zeta X_{j+1}}||e^{-i\zeta\lambda X_j}||(e^{i\zeta X_j(e^{-\lambda} - e^{-\lambda^*})} - 1| \\
&\le 2^{2-\delta}|e^{-\lambda} - e^{-\lambda^*}|^\delta \frac{|\zeta|^\delta}{n-1}\sum_{j=1}^{n-1}|X_j|^\delta \\
&= o_p(1).
\end{aligned}
$$

Here the last inequality is obtained from $|e^{iu}-1| \le 2^{2-\delta}|u|^\delta$, see, for example, Sen and Singer (1993, formula 1.4.59). Note that for this result to hold we only require that $\mathrm{E}|X|^\delta < \infty$.

Part *ii*) can be proved by the same devices. Note first that

$$\left|\frac{\partial}{\partial\lambda}\varphi_{n,\lambda}(\zeta)\big|_{\lambda=\lambda^*} - \frac{\partial}{\partial\lambda}\varphi_{n,\lambda}(\zeta)\right|$$

$$= \left|\frac{1}{n-1}\sum_{j=1}^{n-1}\left(i\zeta e^{-\lambda^*}X_j e^{i\zeta(X_{j+1}-e^{-\lambda^*}X_j)} - i\zeta e^{-\lambda}X_j e^{i\zeta(X_{j+1}-e^{-\lambda}X_j)}\right)\right|$$

$$\le |e^{-\lambda^*}-e^{-\lambda}|\frac{|\zeta|}{n-1}\sum_{j=1}^{n-1}|X_j| + 2^{2-\delta}|e^{-\lambda}-e^{-\lambda^*}|^\delta \frac{|\zeta|^\delta}{n-1}\sum_{j=1}^{n-1}|X_j|^{1+\delta}$$

$$= o_p(1).$$

Next, to obtain the result we note that from Sørensen (1999, Theorem 3.2)

$$i\zeta e^{-\lambda}\frac{1}{n-1}\sum_{j=1}^{n-1} e^{i\zeta(X_{j+1}-e^{-\lambda}X_j)}X_j \quad \xrightarrow{a.s.} \quad i\zeta e^{-\lambda}\mathrm{E}(e^{i\zeta(X_{j+1}-e^{-\lambda}X_j)}X_j) \tag{6.34}$$

and, by independence of $X_{j+1}-e^{-\lambda}X_j$ and X_j, $\mathrm{E}(e^{i\zeta(X_{j+1}-e^{-\lambda}X_j)}X_j) = \varphi_\theta(\zeta)\mathrm{E}(X)$. □

Proof of Theorem 1. We need the following regularity conditions and assumptions:

A1 The parametrization used brings to an identifiable problem, i.e. $\varphi_{\theta'}(\zeta) \ne \varphi_\theta(\zeta)$ if $\theta' \ne \theta$; Θ is a compact set and $\theta_0 \in \mathrm{Int}(\Theta)$.

A2. $\varphi_\theta(\zeta)$ is twice continuously differentiable in θ.

A3 $G_n(\gamma,\lambda) = \frac{1}{2}\frac{\partial Q_n^2(\gamma,\lambda)}{\partial\gamma}$, $\frac{\partial}{\partial\lambda}G_n(\gamma,\lambda)$ and $\frac{\partial}{\partial\gamma}G_n(\gamma,\lambda)$ are W-integrable functions over Θ.

A4. The matrix $B'(\theta)$ defined in Theorem 1 is a W-integrable function over Θ and is of full rank at θ_0.

A5. For a $\delta > 0$, $\mathrm{E}|X|^{1+\delta} < \infty$.

A6. $\hat\lambda$ is a $\sqrt{n}$-consistent estimator of λ.

For the proof of consistency, under $\mathrm{E}(X^\delta) < \infty$ and $\lambda^* \xrightarrow{p} \lambda$ we have from Lemma 3.i), that $|\varphi_{n,\lambda^*}(\zeta) - \varphi_{n,\lambda}(\zeta)| = o_p(1)$ and noting that $\varphi_{n,\lambda}(\zeta)$ is an empirical ch.f. based on *i.i.d.* observations, consistency follows by standard arguments by A1, A2 and A3.

Next, to prove asymptotic normality, by expanding $G_n(\hat\gamma,\hat\lambda)$ around γ we have, for $|\gamma^*-\gamma| \le |\hat\gamma-\gamma|$,

$$\sqrt{n}G_n(\hat\gamma,\hat\lambda) = \sqrt{n}G_n(\gamma,\hat\lambda) + \sqrt{n}(\hat\gamma-\gamma)\frac{\partial}{\partial\gamma}G_n(\gamma,\hat\lambda)\big|_{\gamma=\gamma^*}. \tag{6.35}$$

Asymptotic normality of $\sqrt{n}(\hat\gamma-\gamma)$ will follow if we prove that: a) $\sqrt{n}G_n(\gamma,\hat\lambda)$, with λ playing the role of nuisance parameter, is asymptotically Normal and b) $\frac{\partial}{\partial\gamma}G_n(\gamma,\hat\lambda)\big|_{\gamma=\gamma^*}\frac{\partial}{\partial\gamma'}G_n(\gamma,\hat\lambda)\big|_{\gamma=\gamma^*} \to B'(\theta)$ where $B'(\theta)$ is defined in A4.

To prove a), note that, by a further Taylor expansion around λ, with $|\lambda^*-\lambda| \le |\hat\lambda-\lambda|$ we have

$$\sqrt{n}G_n(\gamma,\hat\lambda) = \sqrt{n}G_n(\gamma,\lambda) + \sqrt{n}(\hat\lambda-\lambda)\frac{\partial}{\partial\lambda}G_n(\gamma,\lambda)\big|_{\lambda=\lambda^*}, \tag{6.36}$$

To ease a little notation, let $\frac{\partial}{\partial\lambda}\varphi_{n,\lambda}(\zeta)\big|_{\lambda=\lambda^*} = \dot\varphi_{n,\lambda^*}(\zeta)$, then we have

$$\frac{\partial}{\partial\lambda}G_n(\gamma,\lambda)\big|_{\lambda=\lambda^*} = \int_S (Re\,\dot\varphi_{n,\lambda^*}(\zeta) - Re\,\dot\varphi_{\gamma,\lambda^*}(\zeta))\frac{\partial}{\partial\gamma}\varphi_{\gamma,\lambda^*}(\zeta)dW(\zeta)+ \\ + \int_S (Re\,\varphi_{n,\lambda^*}(\zeta) - Re\,\varphi_{\gamma,\lambda^*}(\zeta))\frac{\partial}{\partial\gamma}\dot\varphi_{\gamma,\lambda^*}(\zeta)dW(\zeta) + Im \quad (6.37)$$

where Im denotes the same expression of the r.h.s. of the above formula but with Re replaced by Im. By dominated convergence theorem we have $\dot\varphi_{\gamma,\lambda^*}(\zeta) = i\zeta e^{-\lambda^*}\mathrm{E}[X_j \exp\{i\zeta(X_{j+1} - e^{-\lambda^*}X_j\}]$ which, using uniform continuity of the ch.f. and the same devices used in Lemma 1, converges in probability to $i\zeta e^{-\lambda}\varphi_{\gamma,\lambda}(\zeta)\mathrm{E}(X)$. Using this result, *Lemma 3.ii)* and A3 we can prove that $\frac{\partial}{\partial\lambda}G_n(\gamma,\lambda)\big|_{\lambda=\lambda^*}$ converges in probability to 0 and since, by assumption, $\sqrt{n}(\hat\lambda - \lambda) = O_p(1)$, we have that the asymptotic distributions of $\sqrt{n}G_n(\gamma,\hat\lambda)$ and $\sqrt{n}G_n(\gamma,\lambda)$ are the same. Namely, since $G_n(\gamma,\lambda)$ is a sum of *i.i.d.* random variables, the standard central limit theorem obtains

$$\sqrt{n}G_n(\gamma,\lambda) \quad \xrightarrow{D} \quad N(0, A'(\theta_0)) \quad (6.38)$$

where $A'(\theta_0) = \mathrm{E}G_n(\gamma,\lambda)G_n(\gamma,\lambda)'$. As far as part b) is concerned, note that we have

$$\frac{\partial}{\partial\gamma}G_n(\gamma,\lambda) = \int_S (Re\,\varphi_{n,\lambda}(\zeta) - Re\,\varphi_{\gamma,\lambda}(\zeta))\frac{\partial}{\partial\gamma\partial\gamma'}\varphi_{\gamma,\lambda}(\zeta)dW(\zeta)+ \\ + \int_S \frac{\partial}{\partial\gamma}Re\,\varphi_{\gamma,\lambda}(\zeta)\frac{\partial}{\partial\gamma}Re\,\varphi_{\gamma,\lambda}(\zeta)dW(\zeta) + Im. \quad (6.39)$$

Again, we can use the results of Lemma 1 and dominated convergence to get the desird result. The proof can be completed by standard arguments as in, for example, Feuerverger and McDunnough (1981). □

Proof of Theorem 2. The characteristic function of $\mathrm{E}\exp\{i\zeta\int_0^\infty f(t)dS(t)\}$ for a general positive function $f(u)$ can be obtained straightforwardly from Barndorff-Nielsen and Shephard (2001, formula 36); there $H(s) = \int_0^\infty\{\frac{1}{2}\zeta^2 f^2(u+s) - i\beta\zeta f(u+s)\}e^{-\lambda u}du$ and $J = H(0)$. To obtain our result, we need to set $\zeta = 1$ and plug in $f(s) = \sum_{j=1}^m \zeta_j \mathbb{I}_{(t_{j-1},t_j]}(s)$. Note first of all that, since $(t_{j-1},t_j]$ and $(t_{j'-1},t_{j'}]$ do not overlap, for $j \neq j'$, then $f^2(s) = \sum_{j=1}^m \zeta_j^2 \mathbb{I}_{(t_{j-1},t_j]}(s)$, this allows to obtain straightforwardly expression (5.25). To obtain the expression for $H(s)$ we write

$$\mathbb{I}_{(t_{j-1},t_j]}(u+s) = \mathbb{I}_{(t_{j-1}-s,t_j-s]}(u)\mathbb{I}_{(0,t_{j-1}]}(s) + \mathbb{I}_{(0,t_j-s]}(u)\mathbb{I}_{(t_{j-1},t_j]}(s) \quad (6.40)$$

to obtain

$$H(s) = \sum_{j=1}^m \left(\frac{1}{2}\zeta_j^2 - i\beta\zeta_j\right)\left[\varepsilon_\lambda(t_{j-1}-s, t_j-s)\mathbb{I}_{(0,t_{j-1}]}(s) + \varepsilon_\lambda(0, t_j-s)\mathbb{I}_{(t_{j-1},t_j]}(s)\right]. \quad (6.41)$$

In order to compute $\int k(H(s))ds$ it is convenient to separate the components of $H(s)$ over the disjoint intervals $(t_{j-1},t_j]$ and proceed to integrate over separate regions. To this end, write $\mathbb{I}_{(0,t_{j-1}]}(s) = \sum_{l=0}^{j-1}\mathbb{I}_{(t_{l-1},t_l]}(s)$ with the convention that $t_{-1} = 0$ and rearrange terms to obtain (5.26). □

7 Bibliography

Baran, S., Pap, G., van Zuijlen, M. C. A. (2003). Estimation of the mean of stationary and nonstationary Ornstein-Uhlenbeck processes and sheets. *Comput. Math. Appl.* 45, 563-579.

Barndorff-Nielsen, O.E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Statist.* 24, no. 1, 1-13.

Barndorff-Nielsen, O.E. (1998). Processes of Normal Inverse Gaussian type. *Finance and Stochastics* 2, 41-68.

Barndorff-Nielsen, O. E. (2001). Superposition of Ornstein-Uhlenbeck type processes. *Theory Probab. Appl.* 45, no. 2, 175-194. Translated from *Teor. Veroyatnost. i Primenen.* 45 (2000), no. 2, 289-311;

Barndorff-Nielsen, O.E, Jensen, J.L, Sørensen, M. (1998). Some stationary processes in discrete and continuous time. *Adv. Appl. Prob.* 30 (4), 989-1007.

Barndorff-Nielsen, O.E., Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol. 63* , no. 2, 167-241.

Barndorff-Nielsen, O.E., Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. R. Stat. Soc. Ser. B Stat. Methodol. 64* , no. 2, 253-280.

Barndorff-Nielsen, O.E., Shephard, N. (2003). Integrate OU processes and non-Gaussian OU-based stochastic volatility models. *Scand. J. Stat.*, 30, 277-295.

Barndorff-Nielsen, O. E., Leonenko, N. N. (2005). Spectral properties of superpositions of Ornstein-Uhlenbeck type processes. *Methodol. Comput. Appl. Probab.* 7, no. 3, 335-352.

Dufresne, D. (1997). Algebraic properties of Beta and Gamma distributions, and applications. *Adv. Appl. Math.* 20, 285-299.

Feuerverger, A. (1990). An efficiency result for the empirical characteristic function in stationary time-series models. *Canad. J. Statist.* 18, no. 2, 155-161.

Feuerverger, A. and McDunnough, P. (1981). On some fourier methods for inference. *J. Amer. Stat. Assoc.* 76, 379-387.

Florens-Landais, D., Pham, H. (1999). Large deviations in estimation of an Ornstein-Uhlenbeck model. *J. Appl. Probab.* 36, 60-77.

Gander, M.P.S., Stephens, D.A. (2007a). Stochastic volatility modelling in continuous time with general marginal distributions: inference, prediction and model selection. *J. Stat. Plann. and Infer.* 137, 3068-3081.

Gander, M.P.S., Stephens, D.A. (2007b). Simulation and inference for stochastic volatility models driven by Lé vy processes. *Biometrika* 94, no.3, 627-646.

Ghosh, S. and Beran, J. (2006). On estimating the cumulant generating function of linear processes. *Ann. Inst. Statist. Math.* 58, no. 1, 53-71.

Gloter, A. (2001). Parameter estimation for a discrete sampling of an integrated Ornstein-Uhlenbeck process. *Statistics* 35, 225–243.

Griffin, J.E., Steel, M.F.J. (2006). Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility. *J. Econometics* 134, 605-644.

Grigelionis, B. (1999). Processes of Meixner type. *Liet. Mat. Rink.* 39, no. 1, 40-51; translation in *Lithuanian Math. J.* 39 (1999), no. 1, 33-41.

Grigelionis, B. (2001). Generalized z-distributions and related stochastic processes. *Liet. Mat. Rink.* 41, no. 3, 303-319; translation in *Lithuanian Math. J.* 41 (2001), no. 3, 239-251.

Grigelionis, B. (2003). On the self-decomposability of Euler's Gamma function. *Liet. Mat. Rink.* 43, no. 3, 359-370; translation in *Lithuanian Math. J.* 43 (2003), no. 3, 295-305.

Heyde, C.C.; Leonenko, N. (2005). Student processes. *Adv. Appl. Prob.* 37, 342-365.

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 387-396.

Jiang, G.J.; Knight, J.L. (2002). Estimation of continuous-time processes via the empirical characteristic function. J. Bus. Econom. Statist. 20, no. 2, 198-212.

Jongbloed, G.; van der Meulen, F.H. (2006). Parametric estimation for subordinators and induced OU processes. *Scand. J. Statist.* 33, no. 4, 825-847.

Jongbloed, G.; van der Meulen, F. H.; van der Vaart, A. W. (2005). Nonparametric inference for Lvy-driven Ornstein-Uhlenbeck processes. *Bernoulli* 11, no. 5, 759-791.

Jurek, Z.J. and Vervaat, W. (1983). An integral representation for self-decomposable Banach space random variables. *Zeit. Wahrsch. ver. Gebiete* 62, 247-262.

Knight, J.L. and Satchell, S.E. (1997). The cumulant generating function estimation method. *Econometric Theory* 13, 170-184.

Knight, J.L. and Yu, J. (2002). Empirical characteristic function in time series estimation. *Econometric Theory* 18, 691-721.

Knight, J.L.; Satchell, S.E.; Yu, J. (2002). Estimation of the stochastic volatility model by the empirical characteristic function method. *Aust. N. Z. J. Stat.* 44, no. 3, 319-335.

Kotz, S., Kozubowski, T. J., Podgórski, K. (2001). *The Laplace Distributions and Generalizations.* Birkhäuser, Boston.

Madan, D.B. and Seneta, E. (1987). Simulation of estimates using the empirical characteristic function *International Statistical Review* 55, 153-161.

Madan, D.B., Seneta, E. (1990). The VG model for share market returns. *J. Bus.* 63, 511-524.

Neumann, M.J. and Reiss, M. (2008). Nonparametric estimation for Lévy processes from low-frequency observations. *Preprint.*

Pap, G., van Zuijlen, M. C. A. (1996). Parameter estimation with exact distribution for multidimensional Ornstein-Uhlenbeck processes. *J. Multiv. Anal.* 59, 153-165.

Roberts, G.O.; Papaspiliopoulos, O.; Dellaportas, P. (2004). Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 66, no. 2, 369-393.

Rydberg, H (1997). The Normal Inverse Gaussian Lévy process: simulation and approximation. *Communications in Statistics - Stochastic Models* 13 (4), 887-910.

Sato, K.I.(1999). *Levy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

Schoutens, W., Teugels, J.L. (1998). Lévy processes, polynomials and martingales. Special issue in honor of Marcel F. Neuts. *Comm. Statist. Stochastic Models* 14, no. 1-2, 335-349.

Schoutens, W. (2003). *Lévy Processes in Finance.* Wiley, Chicester.

Sen, P.K. and Singer, J.M. (1993). *Large Sample Methods in Statistics.* Chapman & Hall, London.

Seneta, E. (2004) Fitting of Variance-Gamma model to financial data. *J. Appl. Probab.* 41 A, 177-187.

Sørensen, M. (1999). On asymptotics of estimating functions. *Braz. J. Probab. Stat.* 13, no. 2, 111-136.

Steutel, F. W., van Harn, K. (2004). *Infinite Divisibility of Probability Distributions on the Real Line.* Marcel Dekker, New York.

Taufer, E., Leonenko, N. (2008). Simulation of Levy-driven Ornstein-Uhlenbeck processes with given marginal distribution. *Computational statistics & data analysis*, Articles in press.

Todorov, V., Tauchen, G. (2006). Simulation methods for Lvy-driven continuous-time autoregressive moving average (CARMA) stochastic volatility models. *J. Bus. Econom. Statist.* 24, no. 4, 455-469.

Tweedie, M.C.K. (1984). An index which distinguishes between some important exponential families. In *Statistics: Applications and New Directions: Proc. Indian Statistical Institute Golden Jubilee International Conference* (ed. J. Ghosh and J. Roy), pp. 579-604.

Woerner, J.H.C. (2004) Estimating the skewness in discretely observed Lévy processes. *Econometric Theory* 20, 927-942.

Woerner, J.H.C. (2007)Inference in Lévy-type stochastic volatility models. *Adv. Appl. Prob.* 39, 531-549.

Wolfe, J. (1982). On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$. *Stochastic Process. Appl.* 12, no. 3, 301-312.