Bernoulli **11**(2), 2005, 191–220

Diffusion-type models with given marginal distribution and autocorrelation function

BO MARTIN BIBBY[1]*, IB MICHAEL SKOVGAARD[1]** and MICHAEL SØRENSEN[2]

[1]*Department of Mathematics and Physics, The Royal Veterinary and Agricultural University, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark. E-mail: *bibby@dina.kvl.dk; **ims@kvl.dk*

[2]*Department of Applied Mathematics and Statistics, Institute of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 København Ø, Denmark. E-mail: michael@math.ku.dk*

Flexible stationary diffusion-type models are developed that can fit both the marginal distribution and the correlation structure found in many time series from, for example, finance and turbulence. Diffusion models with linear drift and a known and prespecified marginal distribution are studied, and the diffusion coefficients corresponding to a large number of common probability distributions are found explicitly. An approximation to the diffusion coefficient based on saddlepoint techniques is developed for use in cases where there is no explicit expression for the diffusion coefficient. It is demonstrated theoretically as well as in a study of turbulence data that sums of diffusions with linear drift can fit complex correlation structures. Any infinitely divisible distribution satisfying a weak regularity condition can be obtained as a marginal distribution.

Keywords: ergodic diffusions; generalized hyperbolic distributions; long-range dependence; saddlepoint approximation; stochastic differential equation; turbulence

1. Introduction

We consider the problem of choosing a continuous-time model based on discrete-time observations $X_{t_1}, \ldots, X_{t_n}$. Ideally the choice of a model should be based on an understanding of the processes governing the system from which the data are obtained. Often such a description of a system is made using a number of ordinary differential equations – in the case of a single ordinary differential equation,

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = b(X_t), \qquad t \geqslant 0.$$

A natural extension of this model is to add a white noise term,

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sqrt{v(X_t)}\mathrm{d}W_t, \qquad t \geqslant 0,$$

where W is a standard Wiener process. This introduces an uncertainty in the description of the system behind the data and results in dependence between the observations. See Pedersen (2000) for an example of this approach in the modelling of nitrous oxide emissions from the soil surface. In this paper we show how, with a given drift function b, any probability density satisfying weak regularity conditions can be obtained as a marginal distribution by choosing v suitably. This result is useful when choosing a parametrized class of diffusion coefficients v in the light of data. A linear specification of b is studied in detail.

In many cases the mechanisms driving the process of interest are not understood well enough or are too complicated to be described using a simple drift function, b, and a more data-driven approach must be taken. The main aim of this paper is to propose a method of choosing a model based on data also in such cases. Specifically, we show how to construct a model for X with a given marginal density f, $X_t \sim f$, and autocorrelation function $\rho(t) = \mathrm{corr}(X_s, X_{s+t})$, $s, t \geqslant 0$, where f is infinitely divisible and satisfies a weak regularity condition, and where $\rho(t)$ belongs to a large and very flexible class of autocorrelation functions. The model is usually not Markovian. Expressions for f and ρ are typically chosen so that they fit a histogram of the data and the empirical autocorrelation function. Aït-Sahalia (1996) took the same approach as we do in the case of an exponentially decreasing autocorrelation function, but instead of a parametric model for the marginal density, he estimated this density nonparametrically. In Bibby and Sørensen (1997; 2001) a similar approach based only on the marginal density f was used in connection with financial data. The construction in this paper, which involves sums of diffusion processes, is related to the sums of Ornstein–Uhlenbeck processes driven by Lévy processes introduced by Barndorff-Nielsen *et al.* (1998) and provides a continuous-path alternative to these jump processes. Therefore, the models introduced in this paper can be used to construct stochastic volatility models without jumps that are alternatives to the models of Barndorff-Nielsen and Shephard (2001); see Bibby and Sørensen (2004). Constructions different from ours of Markovian martingales with prescribed marginal distributions have recently been considered by Madan and Yor (2002).

In Section 2 we introduce the method in the situation where X is a diffusion process with a linear drift and hence has an exponentially decreasing autocorrelation function. For a large number of commonly used probability distributions, explicit diffusion models are given with linear drift and with these distributions as marginal distributions. Moreover, general expressions for exponential families and normal variance mixtures are established. Also nonlinear drift functions are considered. Section 3 contains a result on an approximation of the squared diffusion coefficient that enlarges the class of possible marginal densities for which a diffusion model can be handled in practice. The approximation is based on saddlepoint techniques, and the marginal density of the resulting model is approximately proportional to the saddlepoint approximation of the original marginal density. In Section 4 models for X with more realistic autocorrelation functions are constructed based on the results in Sections 2 and 3. These models are finite sums of diffusion processes and hence not Markovian. Here the marginal distribution must be infinitely divisible. Relations to long-range dependence are investigated. Infinite sums of diffusions are also briefly considered. In Section 5 multivariate models are introduced. Finally, in Section 6 we study an example involving turbulence data.

2. Construction of diffusions

In this section we describe the construction of diffusion process models with an exponential autocorrelation function and a specified marginal distribution. The diffusion will be constructed such that the marginal distribution is concentrated on a set (l, u) $(-\infty \leqslant l < u \leqslant \infty)$, and has a prespecified density f with respect to the Lebesgue measure on the state space (l, u). The construction of a diffusion from its marginal distribution and an exponential autocorrelation function first appeared in Aït-Sahalia (1996), who instead of using a parametric model estimated the marginal density nonparametrically. In this way he obtained a nonparametric estimator of the diffusion coefficient. In particular, Aït-Sahalia (1996) also derived the basic equations (2.2) (2.4) and (2.9) below. In the rest of this section, let f be a probability density satisfying the following condition.

Condition 2.1. *The probability density f is continuous, bounded, and strictly positive on (l, u), zero outside (l, u), and has finite variance.*

Consider the stochastic differential equation

$$dX_t = -\theta(X_t - \mu)dt + \sqrt{v(X_t)}dW_t, \qquad t \geqslant 0, \tag{2.1}$$

where $\theta > 0$, $\mu \in (l, u)$ and v is a non-negative function defined on the set (l, u). Theorem 2.1 below shows that if

$$v(x) = \frac{2\theta \int_l^x (\mu - y) f(y) dy}{f(x)} = \frac{2\theta\mu F(x) - 2\theta \int_l^x y f(y) dy}{f(x)}, \qquad l < x < u, \tag{2.2}$$

where F is the distribution function associated with the density f, then the solution X is ergodic with invariant density f. The process X is mean-reverting, and if it is stationary the autocorrelation function is $e^{-\theta t}$.

Theorem 2.1. *Suppose the probability density f has expectation μ and satisfies Condition 2.1. Then the following holds.*

(i) *The stochastic differential equation given by* (2.1) *and* (2.2) *has a unique Markovian weak solution. The diffusion coefficient is strictly positive for all $l < x < u$.*
(ii) *The diffusion process X that solves* (2.1) *and* (2.2) *is ergodic with invariant density f.*
(iii) *The function $f(x)v(x)$ satisfies*

$$\int_l^u v(x) f(x) dx < \infty, \tag{2.3}$$

and $E(X_{s+t} \mid X_s = x) = x e^{-\theta t} + \mu(1 - e^{-\theta t})$. If $X_0 \sim f$, then X is stationary, and the autocorrelation function of X is given by

$$\mathrm{corr}(X_{s+t}, X_s) = e^{-\theta t}, \qquad s, t \geqslant 0. \tag{2.4}$$

(iv) *If $-\infty < l$ or $u < \infty$, then the diffusion given by* (2.1) *and* (2.2) *is the only ergodic diffusion with drift $-\theta(x - \mu)$ and invariant density f. If the state space is*

$\mathbb{R}$, it is the only ergodic diffusion with drift $-\theta(x-\mu)$ and invariant density f for which the condition (2.3) *is satisfied.*

Remark. When f has infinite second moment but finite first moment, the stochastic differential equation given by (2.1) and (2.2) also has a unique Markovian weak solution with invariant density f. In this case (2.3) is not satisfied. A finite first moment is obviously needed for the construction (2.2).

Remark. If the state space is the real line, the stochastic differential equation given by (2.1) and (2.8) with $C > 0$ has a unique Markovian weak solution with invariant density f.

Remark. When one (or both) of l or u is finite, densities that are unbounded at the finite boundary can be of interest, for instance a gamma distribution with shape parameter smaller than one. The scale measure will in all cases of practical interest be finite at the boundary where the density is unbounded, so that with positive probability the diffusion will hit this boundary at a finite time. A stationary diffusion with density f is also obtained in this case with v given by (2.2) provided that the diffusion is made instantly reflecting at the boundary, because f will satisfy the appropriate differential equation. This is a nice theoretical solution to the problem that is, however, less easy to implement when simulating the process.

The following lemma is used in the proof of Theorem 2.1.

Lemma 2.2. *Suppose the expectation of f is smaller than or equal to μ, and that v is given by* (2.2). *Then the function*

$$g(x) = f(x)v(x) = 2\theta\int_l^x (\mu - y)f(y)\mathrm{d}y \tag{2.5}$$

is strictly positive for all $l < x < u$, and $\lim_{x\to l} g(x) = 0$. If f has expectation equal to μ, then $\lim_{x\to u} g(x) = 0$.

Proof. Since $g(x) = 2\theta\int_l^x(\mu - y)f(y)\mathrm{d}y$, we see that g is strictly increasing on (l, μ) and strictly decreasing on (μ, u), and that $\lim_{x\to l} g(x) = 0$ and $\lim_{x\to u} g(x) \geqslant 0$. Hence $g(x) > 0$ for all $l < x < u$. □

Proof of Theorem 2.1. That $v(x) > 0$ for all $l < x < u$ follows from Lemma 2.2 and the fact that f is continuous. For $l < x < u$, define the scale density

$$s(x) = \exp\left(2\theta\int_{x^*}^x \frac{y-\mu}{v(y)}\mathrm{d}y\right) = \frac{g(x^*)}{g(x)}, \tag{2.6}$$

for some interior point $l < x^* < u$, and the scale function

$$S(x) = \int_{x^*}^x s(y)\mathrm{d}y = g(x^*)\int_{x^*}^x \frac{1}{g(y)}\mathrm{d}y.$$

The function g is given by (2.5), and we have used the fact that $(\log g(y))' = -2\theta(y-\mu)/v(y)$. The function S is strictly increasing, twice continuously differentiable and maps (l, u) onto $\mathbb{R}$. If $(l, u) = \mathbb{R}$, this follows immediately from Lemma 2.2. If u is finite, it follows from Condition 2.1 that there exists a $K > 0$ such that

$$g(x) = 2\theta\int_x^u (y-\mu)f(y)\mathrm{d}y \leqslant K(u-x),$$

which implies that $\lim_{x\to u} S(x) = \infty$. If l is finite, a similar argument shows that $\lim_{x\to l} S(x) = -\infty$.

The stochastic differential equation

$$\mathrm{d}Y_t = s(S^{-1}(Y_t))\sqrt{v(S^{-1}(Y_t))}\mathrm{d}W_t \tag{2.7}$$

satisfies the conditions of Engelbert and Schmidt's (1985) Theorem 2.2 because the function $s(S^{-1}(x))\sqrt{v(S^{-1}(x))}$ is continuous on $\mathbb{R}$. Hence it has a unique Markovian weak solution with state space $\mathbb{R}$. By Itô's formula, the process $S^{-1}(Y_t)$ solves (2.1). This is the only solution because if X is a solution of (2.1), then $S(X_t)$ solves (2.7), again by Itô's formula. We have thus proved (i).

Regarding (ii), we need only check that the scale measure diverges at both end-points and that the speed measure has a density proportional to f (and hence is finite); see, for example, Skorokhod (1989). The invariant density is proportional to the density of the speed measure; see Karlin and Taylor (1981). We have already proved the first assertion, and the second follows easily because the speed measure has density

$$\frac{1}{v(x)s(x)} = \frac{f(x)}{g(x^*)},$$

where we have used (2.5) and (2.6).

Now to (iii). Note that if we can show that (2.3) holds, then it easily follows from (2.1) that $\mathrm{E}(X_{s+t}\,|\,X_s = x) = x\mathrm{e}^{-\theta t} + \mu(1-\mathrm{e}^{-\theta t})$, which again implies (2.4). If $-\infty < l$ and $u < \infty$, (2.3) follows from Lemma 2.2. Otherwise it must be checked that $v(x)f(x)$ goes sufficiently fast to zero at infinite boundaries. The condition that f has finite variance is exactly enough to ensure this. If $u = \infty$,

$$\begin{aligned}\int_\mu^\infty g(x)\mathrm{d}x &= 2\theta\int_\mu^\infty\int_x^\infty (y-\mu)f(y)\mathrm{d}y\,\mathrm{d}x = 2\theta\int_\mu^\infty\int_\mu^y \mathrm{d}x(y-\mu)f(y)\mathrm{d}y \\ &= 2\theta\int_\mu^\infty (y-\mu)^2 f(y)\mathrm{d}y < \infty,\end{aligned}$$

where we have used Tonelli's theorem. If $l = -\infty$, (2.3) is verified in a similar way.

Finally, to show (iv), note that for an ergodic diffusion of the form (2.1) with invariant density f, necessarily

$$f(x) = \frac{K}{v(x)}\exp\left(-2\theta\int_{x^*}^x \frac{y-\mu}{v(y)}\mathrm{d}y\right)$$

for some positive constant K. Here we have used the general expression for the speed measure. We see that the function $g = fv$ is differentiable and that $(\log g(x))' = -2\theta(x-\mu)/v(x)$ or $g'(x) = -2\theta(x-\mu)f(x)$. Thus

$$v(x) = \frac{2\theta\int_l^x (\mu - y)f(y)\mathrm{d}y + C}{f(x)}, \tag{2.8}$$

for some constant C. To ensure that $v(x) > 0$ for all $l < x < u$, it is necessary that $C \geqslant 0$, since by Lemma 2.2 the integral goes to zero at the boundaries. If one of the boundaries is finite, it is necessary that $C = 0$ for the scale measure $1/(fv)$ to diverge at that boundary, again because the integral in (2.8) goes to zero at the boundaries. If both boundaries are infinite, (2.8) defines an ergodic diffusion with invariant density f for all $C \geqslant 0$. However, (2.3) holds only when $C = 0$. □

By the arguments used to prove (2.3) for $u = \infty$ and $l = -\infty$, it follows that under the assumptions of Theorem 2.1,

$$\int_l^u v(x)f(x)\mathrm{d}x = 2\theta\int_l^u (y-\mu)^2 f(y)\mathrm{d}y = 2\theta\,\mathrm{var}(X_0).$$

The last equality holds, of course, only when $X_0 \sim f$.

The construction in Theorem 2.1 is a particular case of the following general result, the proof of which is analogous to the proof of Theorem 2.1.

Theorem 2.3. *Let b be a drift function with reversion defined on (l, u), that is, there exists a $\kappa \in (l, u)$ such that $b(x) > 0$ for $l < x < \kappa$ and $b(x) < 0$ for $\kappa < x < u$. Suppose f is a strictly positive, continuous probability density on (l, u) satisfying*

$$\int_l^u b(x)f(x)\mathrm{d}x = 0,$$

and that the function bf is continuous and bounded on (l, u) Then

$$v(x) = \frac{2\int_l^x b(y)f(y)\mathrm{d}y}{f(x)} > 0 \tag{2.9}$$

for all $l < x < u$, and the stochastic differential equation

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sqrt{v(X_t)}\mathrm{d}W_t, \qquad t \geqslant 0$$

has a unique Markovian weak solution which is ergodic with invariant density f.

The condition that b has reversion is only made for convenience. A sufficient condition is that the inequality (2.9) holds for all $l < x < u$.

In Bibby and Sørensen (2001) another method of constructing diffusion processes with a given marginal density was discussed. In that paper the squared diffusion coefficient was chosen proportional to the inverse of the marginal density raised to a power and an expression for the drift was then determined from the relationship between the drift, the diffusion coefficient and the invariant density. In Bibby and Sørensen (1997) a special case of this approach was considered, namely a diffusion process with no drift and diffusion coefficient proportional to $1/\sqrt{f}$.

When the invariant density belongs to an exponential family with a linear component in the canonical statistic the squared diffusion coefficient can be determined from the following theorem.

Theorem 2.4. *Consider an invariant density for a diffusion process which belongs to an exponential family of the form*

$$f(x;\, \xi) = a(\xi)b(x)\mathrm{e}^{\xi_1 x + \alpha(\xi)\cdot t(x)}, \tag{2.10}$$

where $\xi = (\xi_1, \ldots, \xi_p)$, *and where* α *and* t *may be vectors. Then the squared diffusion coefficient is given by*

$$v(x;\, \xi) = -\frac{2\theta}{f(x;\, \xi)} \frac{\partial}{\partial \xi_1} F(x;\, \xi), \qquad l < x < u. \tag{2.11}$$

Proof. Since the cumulant transform for f is given by

$$\kappa(t) = \log a(\xi) - \log a(\xi_1 + t, \xi_2, \ldots, \xi_p),$$

we obtain that

$$\mu = -\frac{1}{a(\xi)} \frac{\partial}{\partial \xi_1} a(\xi),$$

and hence that

$$\begin{aligned} \frac{\partial}{\partial \xi_1} F(x;\, \xi) &= \frac{F(x;\, \xi)}{a(\xi)} \frac{\partial}{\partial \xi_1} a(\xi) + \int_l^x y f(y;\, \xi)\mathrm{d}y \\ &= -\mu F(x;\, \xi) + \int_l^x y f(y;\, \xi)\mathrm{d}y, \end{aligned}$$

yielding (2.11). □

The result of simple linear transformations is given in the following lemma, from which it follows that we need only consider centred and standardized distributions.

Lemma 2.5. *Let* X *be a stationary diffusion process with linear drift and invariant density* f. *Consider the linear transformation given by*

$$Y_t = \alpha + \sigma X_t, \qquad \sigma > 0,\, \alpha \in \mathbb{R}.$$

Then

$$v_g(y) = \sigma^2 \cdot v_f\left(\frac{y - \alpha}{\sigma}\right),$$

where g denotes the invariant density of Y, and v_f and v_g denote the squared diffusion coefficients obtained by (2.2) *from f and g, respectively.*

We shall now give an extensive list of examples to show that the diffusion coefficient can be found explicitly for many much-used probability distributions, so that our modelling approach is often easy to use in practice. We begin by considering in detail examples of diffusions with an invariant density on the whole real line, that is $-l = u = \infty$, on the half-line, and with compact support.

Example 2.1 *Student's t distribution.* In this example we consider a diffusion process with invariant density equal to a *t(ν)* distribution, that is,

$$f(x) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)}\left(1 + \frac{1}{\nu}x^2\right)^{-(\nu+1)/2}, \qquad x \in \mathbb{R},\ \nu > 0.$$

Here we have taken $\mu = 0$. We only consider t distributions for which the variance exists, so we assume that $\nu > 2$. In this case

$$\int_{-\infty}^{x} y f(y)\mathrm{d}y = -\frac{\Gamma((\nu+1)/2)\nu^{\nu/2}}{(\nu-1)\sqrt{\nu}\Gamma(\nu/2)}(\nu + x^2)^{-(\nu-1)/2},$$

so we obtain that

$$v(x) = \frac{2\theta}{\nu - 1}(\nu + x^2), \qquad x \in \mathbb{R}.$$

The function v is also well defined for $\nu = 2$, and we saw above that it defines an ergodic diffusion with the *t(2)* distribution as invariant distribution.

In the following example we consider an invariant density on the half-axis (l, ∞), where $l > -\infty$. In this situation it may be more convenient to rewrite the expression in (2.2) as

$$v(x) = \frac{2\theta\left(\int_x^\infty (1 - F(y))\mathrm{d}y - (\mu - x)(1 - F(x))\right)}{f(x)}, \qquad x > l. \tag{2.12}$$

In the case of positive diffusions, where $l = 0$, the squared diffusion coefficient can be expressed in terms of the hazard function λ and the integrated hazard function Λ in the following way:

$$v(x) = \frac{2\theta\left(\mathrm{e}^{\Lambda(x)}\int_x^\infty \mathrm{e}^{-\Lambda(y)}\mathrm{d}y + x - \mu\right)}{\lambda(x)}, \qquad x > 0. \tag{2.13}$$

Example 2.2 *The gamma distribution.* Consider a diffusion process with an invariant density from the gamma distribution, that is,

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\lambda x} \qquad x > 0,\ \alpha > 0,\ \lambda > 0.$$

In order for the density to be bounded, we suppose that $\alpha \geqslant 1$. In this case the expectation is $\mu = \alpha/\lambda$. The distribution function is given by

$$F(x) = \frac{\Gamma(\lambda x;\, \alpha)}{\Gamma(\alpha)},$$

where

$$\Gamma(x;\, \alpha) = \int_0^x y^{\alpha-1} \mathrm{e}^{-y}\, \mathrm{d}y$$

is an incomplete gamma function. For the gamma invariant density we get that

$$\int_0^x yf(y)\mathrm{d}y = \frac{\alpha}{\lambda} F(x) - \frac{x}{\lambda} f(x),$$

and therefore

$$v(x) = \frac{2\theta x}{\lambda}.$$

This process is well known and was proposed by Cox *et al.* (1985) as a model for the short-term interest rate.

The following is a simple example of an invariant density with compact support.

Example 2.3 *The beta distribution.* Consider a diffusion process with an invariant density corresponding to the beta distribution, that is,

$$f(x) = B(\alpha,\, \beta)^{-1} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1,\, \alpha > 0,\, \beta > 0,$$

where $B(a,\, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. In this case the distribution function is given by

$$F(x) = I_x(\alpha,\, \beta) = \frac{\int_0^x y^{\alpha-1}(1-y)^{\beta-1}\mathrm{d}y}{B(\alpha,\, \beta)}, \qquad 0 < x < 1,$$

and the mean is $\mu = \alpha/(\alpha+\beta)$. Similarly, we obtain that

$$\int_0^x yf(y)\mathrm{d}y = \frac{\alpha}{\alpha+\beta} I_x(\alpha+1,\, \beta), \qquad 0 < x < 1.$$

Since we have that

$$I_x(\alpha,\, \beta) - I_x(\alpha+1,\, \beta) = \frac{\Gamma(\alpha+\beta+1)}{\Gamma(\alpha+1)\Gamma(\beta+1)} x^{\alpha}(1-x)^{\beta},$$

the squared diffusion coefficient takes the form

$$v(x) = \frac{2\theta}{\alpha+\beta} x(1-x), \qquad 0 < x < 1.$$

Table 1. The squared diffusion coefficient for the most common distributions. For some parameter values the Student, Pareto, inverse gamma, and the F distributions do not have finite variance

Distribution	Density function $f(x)$	State space (l, u)	Mean μ	Parameter space	Squared diffusion $v(x)$
Normal	$\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$	$(-\infty, \infty)$	0	–	2θ
Student	$\frac{\Gamma((\nu+1)/2)\nu^{\nu/2}}{\sqrt{\pi}\Gamma(\nu/2)}(\nu+x^2)^{-(\nu+1)/2}$	$(-\infty, \infty)$	0	$\nu > 1$	$\frac{2\theta}{\nu-1}(\nu+x^2)$
Laplace	$\frac{\alpha^2-\beta^2}{2\alpha}e^{\beta x-\alpha\vert x\vert}$	$(-\infty, \infty)$	$\frac{2\beta}{\alpha^2-\beta^2}$	$\alpha^2 > \beta^2$	$\frac{2\theta}{\alpha^2-\beta^2}(1+\alpha\vert x\vert+\beta x)$
Logistic	$\frac{e^x}{(1+e^x)^2}$	$(-\infty, \infty)$	0	–	$2\theta[(e^x+e^{-x}+2)\log(1+e^x)$ $-x(1+e^x)]$
Extreme value	$e^{-x-e^{-x}}$	$(-\infty, \infty)$	γ	–	$2\theta e^x\left(\gamma - x + e^{e^{-x}}Ei(-e^{-x})\right)$
Pareto	$\alpha(1+x)^{-\alpha-1}$	$(0, \infty)$	$\frac{1}{\alpha-1}$	$\alpha > 1$	$2\theta\mu x(1+x)$
Exponential	$\lambda e^{-\lambda x}$	$(0, \infty)$	$\frac{1}{\lambda}$	$\lambda > 0$	$\frac{2\theta}{\lambda}x$
Gamma	$\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}$	$(0, \infty)$	$\frac{\alpha}{\lambda}$	$\alpha \geqslant 1,\ \lambda > 0$	$\frac{2\theta}{\lambda}x$
χ^2	$\frac{1}{2^{\nu/2}\Gamma(\nu/2)}x^{(\nu/2)-1}e^{-x/2}$	$(0, \infty)$	ν	$\nu \geqslant 2$	$4\theta x$

Inverse gamma	$\frac{\delta^{\lambda}}{\Gamma(\lambda)} x^{-\lambda-1} e^{-\delta/x}$	$(0, \infty)$	$\frac{\delta}{\lambda-1}$	$\delta > 0,\ \lambda > 1$	$\frac{2\theta}{\lambda-1} x^2$
Inverse Gaussian	$\sqrt{\frac{\lambda}{2\pi x^3}} e^{-\lambda(x-\delta)^2/(2\sigma^2 x)}$	$(0, \infty)$	δ	$\lambda > 0,\ \delta > 0$	$\frac{4\theta\delta}{f(x)} e^{2\lambda/\delta} \Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\delta}+1\right)\right)$
F	$\frac{\alpha^{\alpha/2}\beta^{\beta/2}}{B(\alpha/2, \beta/2)} \frac{x^{(\alpha/2)-1}}{(\beta+\alpha x)^{(\alpha+\beta)/2}}$	$(0, \infty)$	$\frac{\beta}{\beta-2}$	$\alpha \geqslant 2,\ \beta > 2$	$\frac{4\theta}{\alpha(\beta-2)} x(\beta+\alpha x)$
log-normal	$\frac{1}{\sqrt{2\pi\sigma^2 x}} e^{-(\log x-\delta)^2/(2\sigma^2)}$	$(0, \infty)$	$e^{\delta+2\sigma^2}$	$\sigma^2 > 0$	$\frac{2\theta\mu}{f(x)}\left(\Phi\left(\frac{\log x-\delta}{\sigma}\right)-\Phi\left(\frac{\log x-\delta}{\sigma}-\sigma\right)\right)$
Weibull	$cx^{c-1}e^{-x^c}$	$(0, \infty)$	$\Gamma\left(\frac{1}{c}+1\right)$	$c > 0$	$\frac{2\theta}{f(x)}\left(\Gamma\left(\frac{1}{c}+1\right)(1-e^{-x^c})-\Gamma\left(x^c; \frac{1}{c}+1\right)\right)$
Uniform	1	$(0, 1)$	$\frac{1}{2}$	–	$\theta x(1-x)$
Beta	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$	$(0, 1)$	$\frac{\alpha}{\alpha+\beta}$	$\alpha > 0,\ \beta > 0$	$\frac{2\theta}{\alpha+\beta} x(1-x)$

This process has been used to model the variation of exchange rates in a target zone by De Jong *et al.* (2001) (for $\alpha = \beta$) and Larsen and Sørensen (2003).

In Table 1 the squared diffusion coefficient is given for a large number of common distributions. In the table, Φ denotes the standard normal distribution function, $\Gamma(x;\,\alpha)$ the incomplete gamma function given by (2.14), and Ei is the exponential integral function given by

$$Ei(x) = -\int_{-x}^{\infty} y^{-1}\mathrm{e}^{-y}\,\mathrm{d}y, \qquad x < 0.$$

Furthermore, γ denotes Euler's constant, $\gamma \approx 0.577\,22$.

Example 2.4 *Normal variance mixtures.* Consider the normal variance mixture

$$f(x) = \int_0^{\infty} \frac{1}{\sqrt{2\pi u}} \mathrm{e}^{-x^2/(2u)} h(u)\mathrm{d}u,$$

where the mixing distribution has density h on $\mathbb{R}_+$ and finite expectation μ_h. Let f^* be the normal variance mixture with mixing density $h^*(u) = uh(u)/\mu_h$. From the fact that a diffusion with marginals that are normally distributed with variance u is obtained when $v(x) = 2\theta u$ (the Ornstein–Uhlenbeck process), it follows easily that a diffusion with marginal density f emerges when

$$v(x) = \frac{2\theta\mu_h f^*(x)}{f(x)}.$$

If h belongs to a family of densities with a factor of the form x^κ ($\kappa > 0$), h^* belongs to the same class. An example is the class of generalized inverse Gaussian densities which contains, among many others, the inverse Gaussian densities, the gamma densities and the inverse gamma densities. When h is a generalized inverse Gaussian density, both f and f^* are explicitly known generalized hyperbolic densities. This result provides an alternative derivation in the case of the Student distribution, which is a normal variance mixture with an inverse gamma mixing distribution. As another example, a diffusion with a symmetric variance-gamma density, that is, (3.9) with $\beta = 0$, is obtained when

$$v(x) = |x - \mu| \frac{K_{\lambda+\frac{1}{2}}(\alpha|x-\mu|)}{K_{\lambda-\frac{1}{2}}(\alpha|x-\mu|)},$$

where K_λ is the modified Bessel function of the third kind with index λ. The variance-gamma distribution was introduced in the finance literature by Madam and Seneta (1990) and has often been connected with jump processes. For details of generalized inverse Gaussian and generalized hyperbolic distributions, see Bibby and Sørensen (2003).

3. Approximations

For some useful classes of distributions it is not possible to determine an explicit expression for the squared diffusion coefficient. However, for several such distributions the moment generating function exists and is known explicitly, so that the following approximation can be applied. Let M denote the moment generating function corresponding to the density f, that is,

$$M(t) = \int_l^u e^{tx} f(x) dx, \tag{3.1}$$

defined for t in the set

$$T = \left\{ t \in \mathbb{R} \middle| \int_l^u e^{tx} f(x) dx < \infty \right\}.$$

Similarly, we let κ denote the cumulant transform, $\kappa(t) = \log M(t)$, and note that it is twice differentiable for all $t \in \mathrm{int}(T)$. Consider the approximation to v given by

$$\tilde{v}(x) = \frac{2\theta(x-\mu)}{\hat{t}_x}, \tag{3.2}$$

where $\hat{t}_x$ is the saddlepoint given by

$$\kappa'(\hat{t}_x) = x. \tag{3.3}$$

Clearly $\tilde{v}(x)$ is positive for $l < x < u$ since $x - \mu = \kappa'(\hat{t}_x) - \kappa'(0)$ and κ is a convex function. Since κ is analytic the singularity of $\tilde{v}(x)$ at $x = \mu$ is removable; in fact, the limiting value of $\tilde{v}$ is $2\theta\kappa''(0)$ and $\tilde{v}$ has derivatives of all orders.

The function $\tilde{v}$ emerges in a natural way when making a substitution in the expression for v in (2.2). Define

$$r_x = \mathrm{sign}(\hat{t}_x)\sqrt{2(x\hat{t}_x - \kappa(\hat{t}_x))}. \tag{3.4}$$

Then the saddlepoint approximation to the density can be written

$$f(x) \approx (\kappa''(\hat{t}_x))^{-1/2} \varphi(r_x), \tag{3.5}$$

where φ is the standard normal density function. For the following computation, note that r_x is increasing in x, that $r_x \, dr_x = \hat{t}_x \, dx$ and that $(x-\mu)/r_x$ is a differentiable function when extended by continuity at $x = \mu$ where $r_x = 0$. Now define

$$I(x) = \int_x^u (y-\mu) f(y) \, dy = \int_{r_x}^{r_u} r_y \varphi(r_y) G(r_y) \, dr_y,$$

where

$$G(r_y) = \frac{(y-\mu)}{r_y} \frac{f(y)}{\varphi(r_y)} \frac{dy}{dr_y} = \frac{(y-\mu)}{\hat{t}_y} \frac{f(y)}{\varphi(r_y)}.$$

Integration by parts now yields

$$I(x) = \left[-\varphi(r_y)G(r_y)\right]_{r_x}^{r_u} + \int_{r_x}^{r_u} \varphi(r_y)G'(r_y)\,\mathrm{d}r_y$$
$$\approx \varphi(r_x)G(r_x) = f(x)\frac{x-\mu}{\hat{t}_x},$$

from which the approximation (3.2) is obtained using $v(x) = 2\theta I(x)/f(x)$. In this computation we have discarded two terms for the following reasons. First the upper limit r_u is usually infinity even if u is finite, but if r_u is finite, the term $-\varphi(r_u)G(r_u)$ is exponentially small in standard asymptotic analysis because of the factor $\varphi(r_u)$. Second, although there are no asymptotic considerations in the present setting, we may consider what happens when the density, f, corresponds to (a standardized version of) a convolution of n independent replications, thus approaching the normal. In particular, this fits in naturally with the infinitely divisible distributions. In that case the integral arising in the integration by parts above will be of low order compared to the leading term. More precisely, it is of order $O(n^{-1/2})$ relative to the leading term uniformly, improving to a relative error of order $O(n^{-1})$ for large deviations, that is, for arguments $x-\mu$ growing proportionally to $\sqrt{n}$ in the standardized scale. In view of Condition 4.1 in the next section, it may be noted that these asymptotic results are valid as $n \to \infty$ for a family of densities, f_n say, with characteristic functions

$$C_n(t) = \{C_0(t/\sqrt{n})\}^n \mathrm{e}^{\mathrm{i}t\mu}, \tag{3.6}$$

where C_0 is the characteristic function, some power of which must be integrable, of a centred distribution with finite Laplace transform in some neighbourhood of zero. For integer values of n this follows from asymptotic results for saddlepoint approximations. Using the method of contour integrals for the saddlepoint approximation (see Daniels 1954; 1987), the same techniques may be used to prove the validity for real (positive) values of n when C_0 corresponds to an infinitely divisible distribution. In summary, we may expect the approximation $\tilde{v}$ to work reasonably well near the mean and very well in the tails.

The approximation may be refined by inclusion of further terms according to the method outlined in Bleistein (1966). In asymptotic analysis as described above, the order of error would improve to $O(n^{-1})$ by the approximation

$$I(x) \approx \varphi(r_x)G(r_x) + G'(0)(1 - \Phi(r_x))$$

where Φ is the standard normal distribution function and

$$G'(0) = \frac{f(0)}{\varphi(0)}\left(-\frac{\kappa_3}{2\sqrt{\kappa_2}} + \frac{f'(0)}{f(0)}\kappa_2^{3/2}\right),$$

where κ_2 and κ_3 are the second and third cumulants of the distribution with density f.

Some properties of a diffusion process

$$\mathrm{d}X_t = \theta(\mu - X_t)\mathrm{d}t + \sqrt{\tilde{v}(X_t)}\mathrm{d}W_t, \qquad t \geqslant 0, \tag{3.7}$$

with $\tilde{v}$ as squared diffusion coefficient, are stated in the following theorem.

Theorem 3.1. *Let the density f have expectation μ and satisfy Condition 2.1. Assume that the function*

$$h(x) = x\hat{t}_x - \kappa(\hat{t}_x)$$

is such that $\int_\mu^x \exp\{h(y)\}\,\mathrm{d}y$ tends to ∞ as x tends to l and as x tends to u. Then, for a normalizing constant $c > 0$, the density

$$\tilde{f}(x) = \frac{c}{\tilde{v}(x)} \mathrm{e}^{-(x\hat{t}_x - \kappa(\hat{t}_x))}, \qquad x \in (l, u), \tag{3.8}$$

has mean μ, is the marginal density of a diffusion process given by the stochastic differential equation (3.7)*, and conclusions* (i)*,* (ii)*, and* (iii) *of Theorem 2.1 hold with v and f replaced by $\tilde{v}$ and $\tilde{f}$.*

Remark. Note that $\tilde{f}$ in (3.8) is approximately proportional to the saddlepoint approximation to f. This is seen by observing that both $\tilde{v}(x)$ and $\sqrt{\kappa''(\hat{t}_x)}$ are approximately proportional to $\kappa''(0) + \frac{1}{2}\kappa^{(3)}(0)\hat{t}_x$ near the mean of the distribution, while the exponential is identical to that of the saddlepoint approximation. Moreover, $\tilde{v}$ is continuous, and hence so is $\tilde{f}$.

Remark. The condition on the function h is satisfied at the upper end if $u = \infty$ and also if u is finite and either the limiting value or any of the derivatives of f at u is non-zero. Similarly at the lower end, l. The proof of this assertion is trivial in the case $u = \infty$ because $h(x)$ tends to infinity; the other part is derived from Tauberian theorems on the Laplace transform of the density. It seems a reasonable conjecture that the theorem holds without the condition on h, but we have not been able to prove this. Incidentally, the inverse Gaussian distribution provides an example of a density with a (lower) end-point of support at which the density and all its derivatives vanish; the conclusion of the theorem is, however, also valid for this distribution.

Proof of Theorem 3.1. Notice first that h is strictly convex with derivative $h'(x) = \hat{t}_x$ and with minimum $h(\mu) = 0$. For later use we now prove that $h(x)$ tends to infinity as x tends to u. This is trivial if $u = \infty$; otherwise, assume without loss of generality that $u = 0$. Then $\kappa(t)$ is decreasing in t with $\kappa(t) \to -\infty$ as $t \to \infty$. For x satisfying $l < x < 0$ we have

$$tx - \kappa(t) < h(x),$$

for any $t \neq \hat{t}_x$, because κ is strictly convex and the derivative of the left-hand side vanishes at $t = \hat{t}_x$. For arbitrary but fixed $t > 0$, we see that $h(x) \geqslant -\kappa(t)$ for x sufficiently close to zero, because h is increasing and hence has a limit. Since this holds for any $t > 0$ and $-\kappa(t)$ tends to infinity, so does $h(x)$ as x approaches zero. Thus, h tends to infinity at the upper end-point, u, and the same result holds for the lower end-point, l, by the same argument.

Next we prove that the squared diffusion coefficient corresponding to $\tilde{f}$ derived from (2.2) is $\tilde{v}$ from (3.2). For $x > \mu$, consider the integral

$$\int_\mu^x (y - \mu)\tilde{f}(y)\,\mathrm{d}y = \int_0^{h(x)} \frac{c}{2\theta}\,\mathrm{e}^{-h}\,\mathrm{d}h = \frac{c}{2\theta}\left(1 - \mathrm{e}^{-h(x)}\right),$$

where we have used $h'(x) = \hat{t}_x$ to substitute h for y in the integral. Similarly, for $x < \mu$, we have

$$\int_x^\mu (y-\mu)\tilde{f}(y)\,\mathrm{d}y = -\frac{c}{2\theta}\left(1 - \mathrm{e}^{-h(x)}\right).$$

Thus, since h tends to infinity at both ends, the mean of $\tilde{f}$ is μ. Furthermore, substitution in (2.2) shows that $\tilde{v}$ is indeed the squared diffusion coefficient calculated from this equation when the density is $\tilde{f}$.

The proof that the pair consisting of $\tilde{v}$ and $\tilde{f}$ admits the remaining conclusions in (i)–(iii) of Theorem 2.1 now copies the arguments of the proof of that theorem, except that instead of providing Condition 2.1 for $\tilde{f}$ we have directly assumed that the scale function is unbounded at the two end-points. Furthermore, notice that the integral in (2.3) in the present case is proportional to $\int \exp\{-h(x)\}\,\mathrm{d}x$, so that the convexity of h directly implies that the integral is finite. □

Let us consider some examples. For background material and detail about the variance-gamma distribution, the normal-inverse Gaussian distribution, and other generalized hyperbolic distributions, see Bibby and Sørensen (2003).

Example 3.1 *The variance-gamma distribution.* The variance-gamma distribution is a special case of the generalized hyperbolic distribution that has proved useful in the modelling of turbulence and financial data. The density function is given by

$$f(x) = \frac{(\alpha^2 - \beta^2)^\lambda}{\sqrt{\pi}\Gamma(\lambda)(2\alpha)^{\lambda-1/2}}|x-\delta|^{\lambda-1/2}K_{\lambda-1/2}(\alpha|x-\delta|)\mathrm{e}^{\beta(x-\delta)}, \qquad x \in \mathbb{R}, \tag{3.9}$$

where K_λ is the modified Bessel function of the third kind with index λ. The domain of the four parameters is $\lambda > 0$, $\alpha > |\beta|$ and $\delta \in \mathbb{R}$. The mean is of the form

$$\mu = \delta + \frac{2\beta\lambda}{\alpha^2 - \beta^2}. \tag{3.10}$$

Apart from the symmetric case ($\beta = 0$) treated in Example 2.4, it is not obvious how to determine an expression for the squared diffusion coefficient v. The moment generating function is, however, rather simple:

$$M(t) = \mathrm{e}^{\delta t}\left(\frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + t)^2}\right)^\lambda, \qquad |\beta + t| < \alpha. \tag{3.11}$$

The cumulant transform and its first derivative are given by

$$\kappa(t) = \delta t + \lambda(\log(\alpha^2 - \beta^2) - \log(\alpha^2 - (\beta + t)^2)),$$

$$\kappa'(t) = \delta + \frac{2\lambda(\beta + t)}{\alpha^2 - (\beta + t)^2},$$

and so the saddlepoint is

$$\hat{t}_x = \begin{cases} -\beta, & x = \delta, \\ \dfrac{\sqrt{\lambda^2 + \alpha^2(x-\delta)^2} - \lambda}{x - \delta} - \beta, & x \neq \delta. \end{cases}$$

The approximate squared diffusion coefficient thus takes the form

$$\tilde{v}(x) = \begin{cases} \dfrac{4\theta\lambda}{\alpha^2 - \beta^2}, & x = \delta, \\ \dfrac{2\theta(x-\delta)(x-\delta-2\beta\lambda/(\alpha^2-\beta^2))}{\sqrt{\lambda^2+\alpha^2(x-\delta)^2} - \lambda - \beta(x-\delta)}, & x \neq \delta. \end{cases} \tag{3.12}$$

Example 3.2 *The normal-inverse Gaussian distribution.* The normal-inverse Gaussian distribution is another member of the class of generalized hyperbolic distributions. According to Barndorff-Nielsen and Shephard (2001), it often fits the distribution of financial returns very well. Its density is given by

$$f(x) = \frac{\alpha\lambda K_1(\alpha\sqrt{\lambda^2 - (x+\delta)^2})}{\pi\sqrt{\lambda^2 + (x-\delta)^2}} \cdot \mathrm{e}^{\lambda\sqrt{\alpha^2-\beta^2}+\beta(x-\delta)}, \qquad x \in \mathbb{R}, \tag{3.13}$$

where we assume that $\lambda > 0$, $\alpha > |\beta|$, and $\delta \in \mathbb{R}$. The mean is

$$\mu = \delta + \frac{\beta\lambda}{\sqrt{\alpha^2 - \beta^2}}.$$

As in the previous example, the symmetric case ($\beta = 0$) can be handled by the result in Example 2.4, whereas it is hard to determine the squared diffusion coefficient explicitly in the general case. Again the approximation is readily obtained. The moment generating function of the normal-inverse Gaussian distribution is of the form

$$M(t) = \mathrm{e}^{\delta t + \lambda(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+t)^2})}, \qquad |\beta + t| < \alpha, \tag{3.14}$$

giving the following expression for the derivative of the cumulant transform,

$$\kappa'(t) = \delta + \frac{\lambda(\beta+t)}{\sqrt{\alpha^2 - (\beta+t)^2}}.$$

This means that the saddlepoint is given by

$$\hat{t}_x = \frac{\alpha(x-\delta)}{\sqrt{\lambda^2 + (x-\delta)^2}} - \beta,$$

and therefore the following approximate squared diffusion coefficient emerges:

$$\tilde{v}(x) = \frac{2\theta\sqrt{\lambda^2+(x-\delta)^2}(x-\delta-\beta\lambda/\sqrt{\alpha^2-\beta^2})}{\alpha(x-\delta) - \beta\sqrt{\lambda^2+(x-\delta)^2}}. \tag{3.15}$$

□

Example 3.3 *The inverse Gaussian distribution.* For the inverse Gaussian distribution we have that

$$v(x) = 4\theta\delta\sqrt{\frac{2\pi x^3}{\lambda}}\, \mathrm{e}^{2\lambda/\delta}\mathrm{e}^{\lambda(x-\delta)^2/2\delta^2 x}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\delta}+1\right)\right), \qquad x > 0; \tag{3.16}$$

see Table 1. In this case the moment generating function is given by

$$M(t) = \exp\left(\frac{\lambda}{\delta}\left(1-\sqrt{1-\frac{2\delta^2 t}{\lambda}}\right)\right) \qquad t \leqslant \frac{\lambda}{2\delta^2}.$$

The cumulant transform and its first derivative take the form

$$\kappa(t) = \frac{\lambda}{\delta}\left(1-\sqrt{1-\frac{2\delta^2 t}{\lambda}}\right),$$

$$\kappa'(t) = \frac{\delta\sqrt{\lambda}}{\sqrt{\lambda - 2\delta^2 t}}.$$

This means that the saddlepoint is given by

$$\hat{t}_x = \frac{\lambda(x+\delta)(x-\delta)}{2\delta^2 x^2},$$

and so

$$\tilde{v}(x) = \frac{4\theta\delta^2 x^2}{\lambda(x+\delta)}. \tag{3.17}$$

In Figure 1 the two versions of the squared diffusion coefficient, (3.16) and (3.17), corresponding to the parameter values $\theta = 1$, $\lambda = 5$ and 25, and $\delta = 5$ and 25 are drawn. Note that $M(t)$ is of the form $M_0(t)^\nu$ with $\nu = \lambda/\delta$ and $M_0(t) = \exp(1 - \sqrt{1 - 2\delta^2 t/\lambda})$, so from the remarks after (3.6) we expect the approximation to improve as λ/δ increases, which is in accordance with Figure 1.

Just as Theorem 2.1 could be generalized to diffusions with nonlinear drift function, as shown in Theorem 2.3, we may also generalize Theorem 3.1 to such cases. This may be viewed not only as an approximation but also as a result providing a class of diffusions with nonlinear drift and an (exact) analytic expression for the stationary density. The approximation is derived just as for the case with linear drift, and the proof follows that of Theorem 3.1.

Theorem 3.2. *Consider a probability density, f, satisfying the conditions of Theorem 3.1, and a drift function, $b(x)$, satisfying $b(x) > 0$ for $l < x < \mu$ and $b(x) < 0$ for $\mu < x < u$. Assume further that*

$$\int_l^u \frac{1}{\tilde{v}(x)}\, e^{-(x\hat{t}_x - \kappa(\hat{t}_x))}\, \mathrm{d}x < \infty,$$

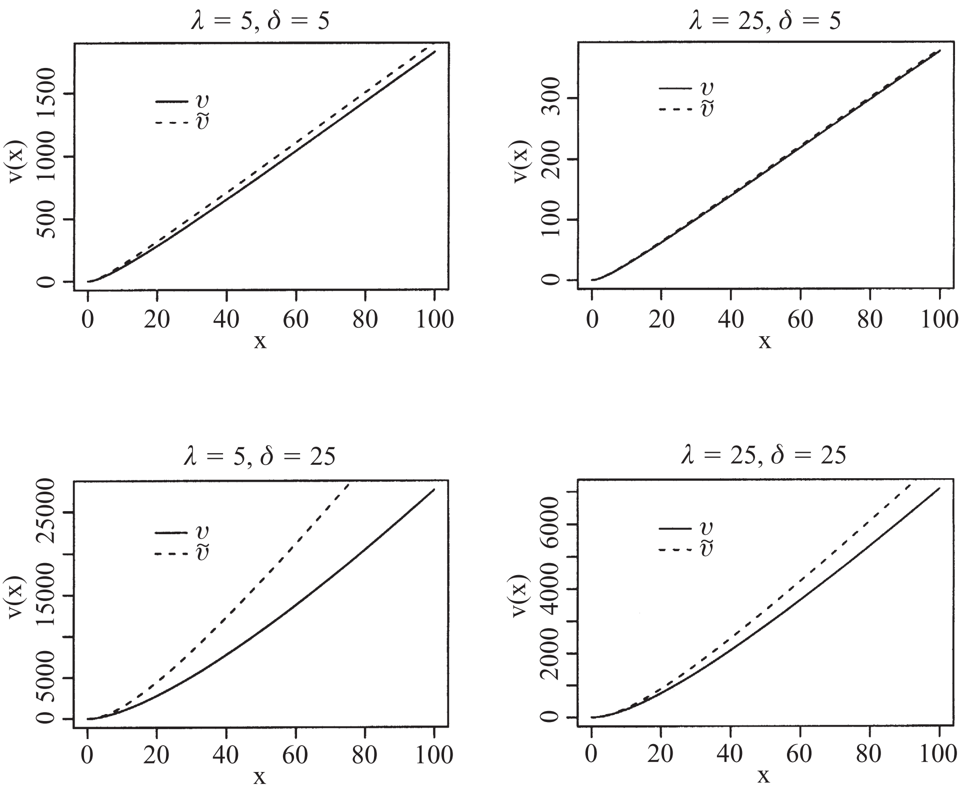

Figure 1. The two squared diffusion coefficients, (3.16) and (3.17), corresponding to the parameter values $\theta = 1$, $\lambda = 5$ and 25, and $\delta = 5$ and 25.

where

$$\tilde{v}(x) = \frac{-2b(x)}{\hat{t}_x}$$

(defined by continuity at x= *0) replaces* (3.2). *Then the differential equation*

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sqrt{\tilde{v}(X_t)}\,\mathrm{d}W_t, \qquad t \geqslant 0,$$

has a unique Markovian solution which is ergodic with invariant probability density $\tilde{f}$ given by (3.8).

Remark. Unlike the linear case, it is no longer true in general that μ is the mean of the distribution with density $\tilde{f}$. But the mean of the drift function, $b(X_t)$, is zero (provided X is stationary), so when $b(x)$ is anti-symmetric around μ, the mean is still μ. Similarly, an anti-symmetric drift guarantees that the same approximation results, relating $\tilde{f}$ to f, hold as in the case with linear drift, but in the general case $\tilde{f}$ may not comply with the saddlepoint approximation to f to the same degree of accuracy around $x = 0$; see the remark just below Theorem 3.1.

4. Sums of diffusions

Very often the correlation structure found in time series data is more complex than the exponentially decreasing autocorrelation of the models defined in Section 2. For diffusion models with a nonlinear drift the autocorrelation function is usually not known explicitly, but the autocorrelation function is bounded by a decreasing exponential function for all ρ-mixing diffusions. A stationary, ergodic diffusion is ρ-mixing under rather weak conditions; see Genon-Catalot *et al.* (2000). In order to obtain models with a more flexible correlation structure, we will therefore consider stochastic processes that are sums of processes of the type introduced in Section 2. Such processes have an explicit autocorrelation function of the form

$$\rho(t) = \phi_1 e^{-\theta_1 t} + \phi_2 e^{-\theta_2 t} + \ldots + \phi_m e^{-\theta_m t}, \qquad t \geqslant 0, \tag{4.1}$$

where $\phi_i > 0$, $i = 1, \ldots, m$ and $\phi_1 + \phi_2 + \ldots + \phi_m = 1$. This functional form is very flexible and can be fitted to a lot of empirical autocorrelation functions; see the discussion below. The construction considered in this section is closely related to the sums of Ornstein–Uhlenbeck processes driven by Lévy processes introduced in Barndorff-Nielsen *et al.* (1998) and used by Barndorff-Nielsen & Shephard (2001) to model financial volatility.

Our aim is to construct a stationary process X with a given marginal density f and with autocorrelation function given by (4.1) for some given integer m. We assume that f satisfies the following condition.

Condition 4.1. *The probability density f, with characteristic function C, is infinitely divisible, that is C^ϕ is a characteristic function for all positive ϕ. Assume, moreover, that there exists a $\phi_0 \geqslant 0$ such that for $\phi > \phi_0$ the probability distribution corresponding to C^ϕ has a density satisfying Condition 2.1.*

Note that Condition 4.1 excludes all distributions on a bounded interval since such distributions cannot be infinitely divisible. If f satisfies Condition 2.1, the only problem in the last part of Condition 4.1 is the boundedness of the density corresponding to C^ϕ because infinitely divisible densities are necessarily positive on (l, u). Properties of infinitely divisible distributions are reviewed in Steutel (1983). The reader is reminded that if C^ϕ is Lebesgue integrable on $\mathbb{R}$, then the corresponding distribution has a density with respect to the Lebesgue measure that is continuous and bounded as required in Condition 2.1. This is useful when the density cannot be found explicitly.

Let $f^{(i)}$ denote the density function corresponding to the characteristic function C^{ϕ_i} ($i = 1, \ldots, m$). Since f satisfies Condition 4.1, we can, according to Theorem 2.1, define a stationary process $X^{(i)}$ of the type introduced in Section 2 with marginal density $f^{(i)}$, provided that $\phi_i > \phi_0$, $i = 1, \ldots, m$. We will assume this to be the case. Specifically, let $W^{(1)}, W^{(2)}, \ldots, W^{(m)}$ be m mutually independent Wiener processes, define

$$v_i(x) = \frac{2\theta_i \int_l^x (\phi_i \mu - y) f^{(i)}(y) \mathrm{d}y}{f^{(i)}(x)}, \qquad i = 1, \ldots, m, \tag{4.2}$$

where μ is the expectation of f, and let $X^{(i)}$ be the solution of the stochastic differential equation

$$\mathrm{d}X_t^{(i)} = \theta_i(\phi_i \mu - X_t^{(i)})\mathrm{d}t + \sqrt{v_i(X_t^{(i)})}\mathrm{d}W_t^{(i)}, \qquad i = 1, \ldots, m. \tag{4.3}$$

Then the processes $X_t^{(1)}, \ldots, X_t^{(m)}$ are independent, and $X_t^{(i)} \sim f^{(i)}$, $i = 1, \ldots, m$, that is, the distribution of $X_t^{(i)}$ has characteristic function C^{ϕ_i}. Hence the process X constructed as the sum

$$X_t = X_t^{(1)} + X_t^{(2)} + \ldots + X_t^{(m)}, \tag{4.4}$$

has marginal density f, and since

$$\mathrm{corr}(X_{s+t}^{(i)}, X_s^{(i)}) = \mathrm{e}^{-\theta_i t}, \qquad i = 1, \ldots, m, \tag{4.5}$$

the autocorrelation function of X is given by (4.1). It is not difficult to see that

$$\phi_i = \frac{\mathrm{var}(X_t^{(i)})}{\mathrm{var}(X_t)}, \qquad i = 1, 2, \ldots, m. \tag{4.6}$$

The spectral density of the process X is given by

$$e(\omega) = \frac{2}{\pi}\left(\frac{\phi_1\theta_1}{\theta_1^2 + \omega^2} + \ldots + \frac{\phi_m\theta_m}{\theta_m^2 + \omega^2}\right), \tag{4.7}$$

which follows immediately from the fact that a process with autocorrelation function $\mathrm{e}^{-\theta t}$ has spectral density $2\theta/(\pi(\theta^2 + \omega^2))$.

The motivation for models of the type (4.4) is that the random variation quite frequently is a compound of processes with different time scales. An example is the velocity fluctuations in a turbulent wind that are caused by eddies with different time scales. The process $X^{(i)}$ represents random variation with a time scale θ_i^{-1}.

The construction of the process X is particularly simple if the marginal distribution of X belongs to a class of distributions which is closed under convolution. The following two examples illustrate this.

Example 4.1 *The gamma distribution.* Here we construct a stationary stochastic process X for which the marginal density is a gamma distribution, $X_t \sim \Gamma(\alpha, \lambda)$, and the autocorrelation function is of the form (4.1). This process can be obtained as the sum of m independent diffusion processes (4.4), where $X_t^{(i)}$ is the solution of

$$\mathrm{d}X_t^{(i)} = \theta_i\big(\phi_i\alpha\lambda^{-1} - X_t^{(i)}\big)\mathrm{d}t + \sqrt{2\theta_i\lambda^{-1}X_t^{(i)}}\mathrm{d}W_t^{(i)}, \qquad i = 1, \ldots, m. \tag{4.8}$$

According to Example 2.2, $X_t^{(i)} \sim \Gamma(\phi_i \alpha, \lambda)$, $i = 1, \ldots, m$, and $X_t^{(i)}$ satisfies (4.5). Here the ϕ_0 of Condition 4.1 equals α^{-1}, so the construction is only possible when $\phi_i \geqslant \alpha^{-1}$, $i = 1, \ldots, m$. If there exists a $\phi_i < \alpha^{-1}$, then the process $X^{(i)}$ is not ergodic and can hit the boundary zero in finite time with positive probability.

Example 4.2 *The variance-gamma distribution.* In this example we construct a stochastic process X whose marginal density is a variance-gamma distribution, $X_t \sim VG(\lambda, \alpha, \beta, \delta)$ – see Example 3.1 – and whose autocorrelation function is of the form (4.1). Let $X^{(1)}, \ldots, X^{(m)}$ be independent diffusions constructed according to (4.3) and (4.2) with μ given by (3.10). Then

$$X_t^{(i)} \sim VG(\phi_i \lambda, \alpha, \beta, \phi_i \delta), \qquad i = 1, \ldots, m,$$

and X given by the sum (4.4) has the right distribution and autocorrelation function. In practice v_i has to be replaced by the approximation $\tilde{v}_i$; see Example 3.1 and Section 6.

Finally, a more difficult example.

Example 4.3 *The hyperbolic distribution.* The moment generating function of the centred symmetric hyperbolic distribution is

$$M(t) = \frac{\alpha \cdot K_1(\delta(\alpha^2 - t^2))}{\sqrt{\alpha^2 - t^2} \cdot K_1(\delta\alpha)}, \qquad |t| < \alpha.$$

The hyperbolic distribution is infinitely divisible, so $M(t)^{\phi_i}$ is again a moment generating function, but there seems to be no way of inverting it to get an expression for $f^{(i)}$. If one wishes to simulate the process of type (4.4) with centred symmetric hyperbolic marginal distribution, it is therefore necessary to use the approximation introduced in Section 3. This can clearly only be done numerically.

The following theorem states exactly which autocorrelation functions can be approximated by an autocorrelation function of the form (4.1).

Theorem 4.1. *The class of functions obtained as limits, as $m \to \infty$, of pointwise convergent sequences $\rho_m(t)$ of autocorrelation functions given by* (4.1) *equals the class of all Laplace transforms for distributions on* $(0, \infty)$*, that is, the class of functions given by*

$$r(u) = \int_0^\infty e^{-uv} dP(v), \qquad u \geqslant 0,$$

for some probability measure P on $(0, \infty)$.

Proof. An autocorrelation function

$$\rho_m(t) = \phi_1^{(m)} e^{-\theta_1^{(m)} t} + \ldots + \phi_m^{(m)} e^{-\theta_m^{(m)} t}$$

is equal to the Laplace transform of the distribution concentrated in $\theta_1^{(m)}, \ldots, \theta_m^{(m)}$ with probabilities $\phi_1^{(m)}, .\phi_m^{(m)}$. If the sequence $\rho_m(t)$ is convergent, the sequence of distributions converges weakly to a probability distribution on $(0, \infty)$ and the limit function is the Laplace transform of this distribution. On the other hand, any probability distribution on $(0, \infty)$ can be obtained as the limit of probability distributions concentrated on a finite set. To see this consider a suitable sequence of discretizations of the distribution in question. □

We see, in particular, that we can only approximate autocorrelation functions that are decreasing and convex. Moreover, the logarithm of the autocorrelation function must also be convex. In fact, it is well known that the class of all Laplace transforms of distributions on $(0, \infty)$ equals the class of *completely monotone functions* r with $r(0) = 1$; see Feller 1971, p. 439). A function r on $[0, \infty)$ is called completely monotone if

$$(-1)^n r^{(n)}(u) \geqslant 0, \qquad u > 0$$

for all $n \in \mathbb{N}$, where $r^{(n)}$ is the nth derivative of r.

One motivation for using models with autocorrelations of the type (4.1) is to be able to fit a relatively simple model to data to which some might think it necessary to fit a model with *long-range dependence*. Let us therefore briefly discuss the fact that an autocorrelation function of the type (4.1) can be close to an autocorrelation function of a process with long memory. Suppose the series

$$r(u) = \sum_{j=1}^{\infty} \phi_j e^{-\theta_j u} \tag{4.9}$$

is convergent. If we, moreover, assume that

$$\sum_{j=1}^{\infty} \phi_j/\theta_j = \infty, \tag{4.10}$$

then

$$\int_0^{\infty} r(u) \mathrm{d}u = \infty,$$

so $r(u)$ is the autocorrelation function of a process with long memory that can be approximated as well as we wish by an autocorrelation function of the type (4.1). To give a specific example, we choose

$$\phi_j \sim j^{-1-2(1-H)}, \qquad \theta_j \sim j^{-1}$$

with $0 < H < 1$. Then

$$\phi_j/\theta_j \sim j^{-2(1-H)},$$

and when $\frac{1}{2} \leqslant H < 1$,

$$r(u) \sim L(u) u^{-2(1-H)},$$

where L is a slowly varying function (for a definition, see Feller 1971, p. 276), so a process with autocorrelation function r has long memory with Hurst exponent H.

The convergence of the sum (4.9) implies mean-square and hence almost sure convergence of the sum

$$X_t = \sum_{i=1}^{\infty} X_t^{(i)}, \tag{4.11}$$

where the random variables $X_t^{(i)}$ are given by (4.2) and (4.3) with $f^{(i)}$ denoting the density function corresponding to $C(t)^{\phi_i}$. It is again assumed that $f^{(i)}$, $i = 1, \ldots$, are continuous and bounded on their support. The limit process X is stationary with marginal density f and has autocorrelation function $r(t)$. It is thus possible to define an infinite version of the sum (4.4). Usually this is, however, an unnecessary complication because a good fit to data can be obtained for a small value of m in (4.4). When the long-memory condition (4.10) is satisfied, the limit process (4.11) is closely related to the long-range dependent processes constructed in Cox (1984), Barndorff-Nielsen *et al.* (1990), and Barndorff-Nielsen (1998).

5. Multivariate models

In this section we shall briefly show how to construct multivariate processes where each coordinate is a process of the type introduced in Section 4.

As in Section 4, we consider a probability density f with characteristic function C satisfying Condition 4.1. For given $\phi_i > \phi_0$ $(i = 1, \ldots, m)$ satisfying $\phi_1 + \ldots + \phi_m = 1$, define $v_i(x)$ by (4.2). Let the processes $X_t^{(k,i)}$, $i = 1, \ldots, m$, $k = 1, \ldots, d$, be given by

$$\mathrm{d}X_t^{(k,i)} = -\theta_i\left(X_t^{(k,i)} - \phi_i\mu\right)\mathrm{d}t + \sqrt{v_i\left(X_t^{(k,i)}\right)}\,\mathrm{d}W_t^{(k,i)},$$

where $W^{(1,1)}, \ldots, W^{(d,m)}$ are independent standard Wiener processes. Then we can define a d-dimensional process X by

$$X_t = (X_{1t}, \ldots, X_{\mathrm{d}t})$$

with

$$X_{jt} = X_t^{(\nu_{j1},1)} + \ldots + X_t^{(\nu_{jm},m)},$$

where $\nu_{ji} \in \{1, \ldots, d\}$, $i = 1, \ldots, m$. The point is that some of the ν_{ji} can be identical for different j so that the same process appears in different coordinates. As previously, we can interpret the process $X_t^{(k,i)}$ as random variation with time scale θ_i^{-1}. Dependence between two coordinates is thus caused by the random variation on certain time scales being the same for the two coordinates. Extra flexibility in the modelling of dependence can be obtained by taking some of the θ_i to be identical, so that it is possible for only a part of the random variation on a certain time scale to be the same in two coordinates.

The density of X_{jt} is f, and the autocorrelation function of X_j is given by (4.1). Moreover,

$$\operatorname{corr}(X_{j_1 t}, X_{j_2 t}) = \sum_{i \in M_{j_1 j_2}} \phi_i,$$

where $M_{j_1 j_2} = \{i \mid \nu_{j_1 i} = \nu_{j_2 i}\}$. The final process is obtained by making location–scale transformations of the marginals X_{jt}.

6. Case study

In this section we consider a data set consisting of 5415 measurements of the streamwise wind velocity component measured at Ferring on the Danish west coast in an experiment carried out in September 1985. The data were recorded using a sonic anemometer on a 30 metre mast erected on the shore around 60 m from the shoreline, with a 10 Hz frequency. The experiment is described in further detail in Mikkelsen (1988; 1989); the data were analysed in Barndorff-Nielsen *et al.* (1993) using a sum of independent autoregressions and in Bibby and Sørensen (2001) based on a hyperbolic diffusion model. The data are presented in Figure 2. See also Barndorff-Nielsen *et al.* (1990).

In Figures 3 and 4 a histogram and a log histogram of the wind velocity data are given, along with fitted curves corresponding to a variance-gamma density function; see Example 3.1. The fitted curves are determined by maximum likelihood based on a multinomial likelihood function where the groups are defined by the points (mid-points) in Figures 3 and 4.

In Figures 5 and 6 the empirical autocorrelation function and its logarithm are drawn for lag values up to 500. The following three exponential functions (and their logarithms) are included in the figures:

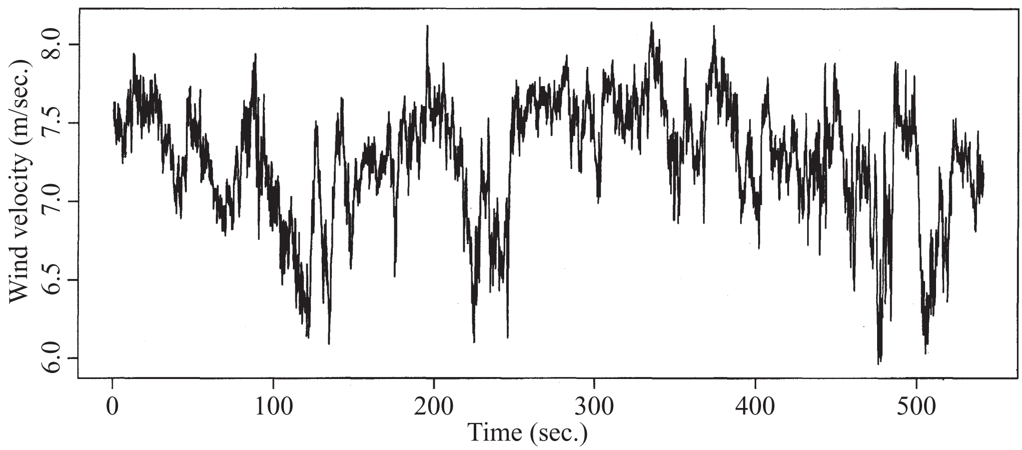

Figure 2. Streamwise wind velocity component in metres per second plotted against time in seconds.

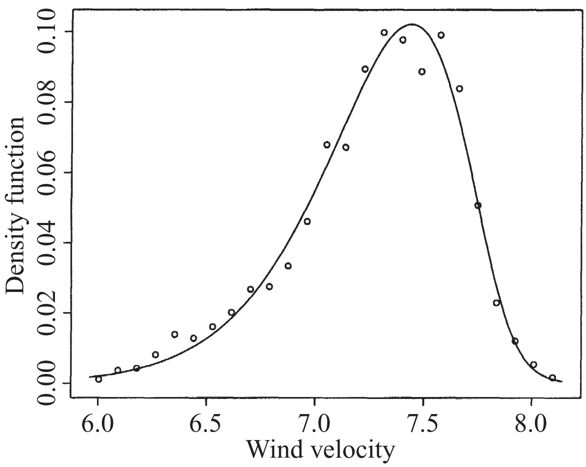

Figure 3. Histogram of the wind velocity data with a fitted variance-gamma density function

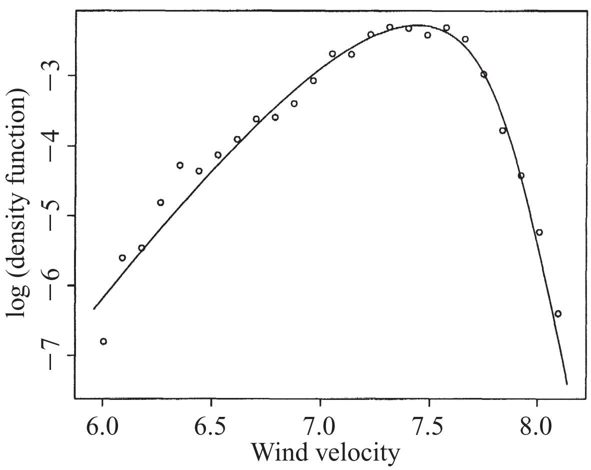

Figure 4. Log histogram of the wind velocity data with a fitted variance-gamma log-density function.

$$\rho_1(t) = \mathrm{e}^{-0.0093t},$$

$$\rho_2(t) = 0.83\mathrm{e}^{-0.0154t} + 0.17\mathrm{e}^{-0.0009t},$$

$$\rho_3(t) = 0.80\mathrm{e}^{-0.0125t} + 0.09\mathrm{e}^{-0.0986t} + 0.11\mathrm{e}^{-0.0001t}.$$

Based on Figures 3–6, we wish to consider a stochastic process with a variance-gamma marginal distribution,

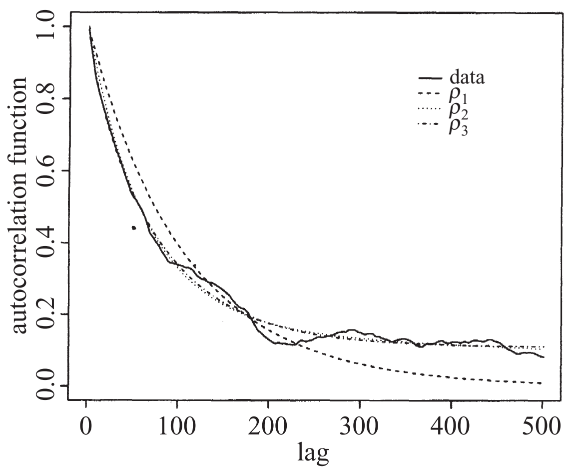

Figure 5. The empirical autocorrelation function of the wind velocity data with fitted curves corresponding to one, two, and three exponential functions.

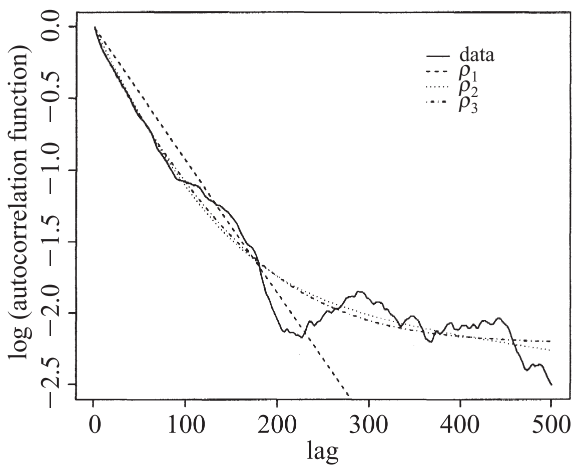

Figure 6. The log empirical autocorrelation function of the wind velocity data with fitted curves corresponding to the logarithm of one, two, and three exponential functions.

$$X_t \sim VG(\lambda, \alpha, \beta, \delta),$$

given as the sum of two diffusion processes,

$$X_t = X_t^{(1)} + X_t^{(2)},$$

where

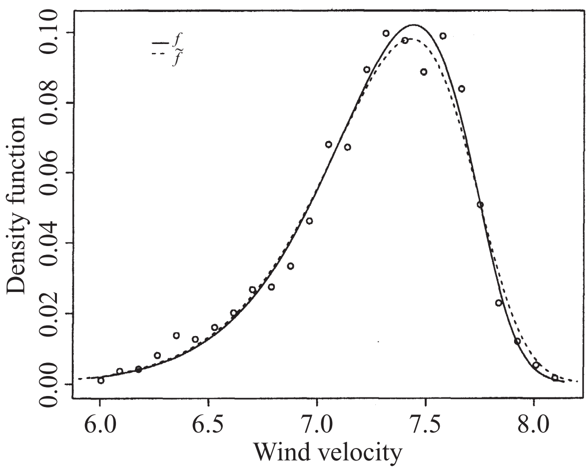

Figure 7. Histogram of the wind velocity data with fitted corresponding to a variance-gamma density (f) and an approximate variance-gamma density ($\tilde{f}$).

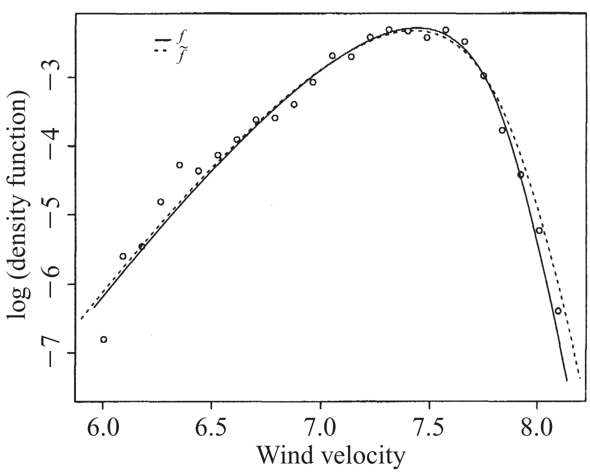

Figure 8. Log histogram of the wind velocity data with fitted curves corresponding to a variance-gamma log-density (f) and an approximate variance-gamme log-density ($\tilde{f}$).

$$\mathrm{corr}(X_{s+t}^{(i)}, X_s^{(i)}) = \mathrm{e}^{-\theta_i t}, \quad i = 1, 2.$$

This can be done using the construction given in Example 4.2.

From the fit of the histogram and the empirical autocorrelation function we obtain

$$\hat{\lambda} = 3.9134, \quad \hat{\alpha} = 13.1760, \quad \hat{\beta} = -7.8232, \quad \hat{\delta} = 7.8128,$$

$$\hat{\phi}_1 = 0.8346, \quad \hat{\phi}_2 = 0.1654, \quad \hat{\theta}_1 = 10.0154, \quad \hat{\theta}_2 = 0.0009.$$

A problem here is that v_1 and v_2 cannot be determined explicitly by (2.2). Instead we can consider the approximations given by (3.2). In Figures 7 and 8 the histogram and log histogram in Figures 3 and 4 are reproduced, now with the addition of the estimated $\tilde{f}$ given by the convolution of $\tilde{f}_1$ and $\tilde{f}_2$ from (3.8). The convolution had to be done numerically.

Acknowledgements

We are grateful to Harald Mikkelsen and Keld Rømer Rasmussen for putting the wind data at our disposal. We are equally grateful to Nakahiro Yoshida for his comments, and in particular for suggesting Theorem 4.1. The research of Michael Sørensen was supported by MaPhySto, The Danish National Research Foundation Network in Mathematical Physics and Stochastics, by the European Commission through the Research Training Network DYNSTOCH under the Human Potential Programme, and by the Centre for Analytical Finance funded by the Danish Social Science Research Council.

References

Aït-Sahalia, Y. (1996) Nonparametric pricing of interest rate derivative securities. *Econometrica*, **64**, 527–560.

Barndorff-Nielsen, O.E. (1998) Processes of normal inverse Gaussian type. *Finance Stochastics*, **2**, 41–68.

Barndorff-Nielsen, O.E. and Shephard, N. (2001) Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial econometrics (with discussion). *J. Roy. Statist. Soc. Ser. B.*, **63**, 167–241.

Barndorff-Nielsen, O.E., Jensen, J.L. and Sørensen, M. (1990) Parametric modelling of turbulence. *Philos. Trans. Roy. Soc. Lond. Ser. A.*, **332**, 439–455.

Barndorff-Nielsen, O.E., Jensen, J.L. and Sørensen, M. (1993) A statistical model for the streamwise component of a turbulent velocity field. *Ann. Geophys.*, **11**, 99–103.

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

Bibby, B.M. and Sørensen, M. (1997) A hyperbolic diffusion model for stock prices. *Finance Stochastics*, **1**, 25–41.

Bibby, B.M. and Sørensen, M. (2001) Simplified estimating functions for diffusion models with a high-dimensional parameter. *Scand. J. Statist.*, **28**, 99–112.

Bibby, B.M. and Sørensen, M. (2003) Hyperbolic processes in finance. In S. Rachev (ed.), *Handbook of Heavy Tailed Distributions in Finance*, pp. 211–248. Amsterdam: Elsevier Science.

Bibby, B.M. and Sørensen, M. (2004) Flexible stochastic volatility models of the diffusion type. In preparation

Bleistein, N. (1966) Uniform asymptotic expansions of integrals with stationary point near algebraic singularity. *Comm. Pure Appl. Math.*, **19**, 353–370.

Cox, D.R. (1984) Long-range dependence: A review. In H.A. David and H.T. David (eds), *Statistics: An Appraisal.* Ames: Iowa State University Press.

Cox, J.C., Ingersoll Jr., J.E. and Ross, S.A. (1985) A theory of the term structure of interest rates. *Econometrica*, **53**, 385–407.

Daniels, H.E. (1954) Saddlepoint approximations in statistics. *Ann. Math. Statist.*, **25**, 631–650.
Daniels, H.E. (1987) Tail probability approximations. *Internat. Statist. Rev.*, **55**, 37–48.
De Jong, F., Drost, F.C. and Werker, B.J.M. (2001) A jump-diffusion model for exchange rates in a target zone. *Statist. Neerlandica*, **55**, 270–300.
Engelbert, H.J. and Schmidt, W. (1985) On solutions of one-dimensional stochastic differential equations without drift. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **68**, 287–314.
Feller, W. (1971) *An Introduction to Probability Theory and its Applications, Vol. II.* New York: Wiley.
Genon-Catalot, V., Jeantheau, T. and Larédo, C. (2000) Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli*, **6**, 1051–1079.
Karlin, S. and Taylor, M. (1981) *A Second Course in Stochastic Processes*. Orlando, FL: Academic Press.
Larsen, K.S. and Sørensen, M. (2003) A diffusion model for exchange rates in a target zone. Preprint no. 6, Department of Applied Mathematics and Statistics, University of Copenhagen.
Madan, D.B. and Seneta, E. (1990) The variance gamma (V.G.) model for share market returns. *J. Business*, **63**, 511–524.
Madan, D.B. and Yor, M. (2002) Making Markov martingales meet marginals: with explicit constructions. *Bernoulli*, **8**, 509–536.
Mikkelsen, H.E. (1988) *Turbulence in the Wake Zone of a Vegetated 2-Dimensional Dune*, Geoskrifter 30. Geology Institute, University of Aarhus.
Mikkelsen, H.E. (1989) *Wind Flow and Sediment Transport over a Low Coastal Dune*, Geoskrifter 32. Geology Institute, University of Aarhus.
Pedersen, A.R. (2000) Estimating the nitrous oxide emission rate from the soil surface by means of a diffusion model. *Scand. J. Statist.*, **27**, 385–404.
Skorokhod, A.V. (1989) *Asymptotic Methods in the Theory of Stochastic Differential Equation*. Providence, RI: Americal Mathematical Society.
Steutel, F.W. (1983) Infinite divisibility. In N.L. Johnson, S. Kotz and C.B. Read (eds), *Encyclopedia of Statistical Sciences, Vol. 4*, pp. 114–116. New York: Wiley.

Received April 2003 and revised July 2004