Inference for Stochastic Volatility Models Driven by Lévy Processes

By MATTHEW P. S. GANDER and DAVID A. STEPHENS

Department of Mathematics, Imperial College London, SW7 2AZ, London, UK

d.stephens@imperial.ac.uk

Summary

We extend the currently most popular models for the volatility of financial time series, Ornstein-Uhlenbeck stochastic processes, to more general non Ornstein-Uhlenbeck models. In particular, we investigate means of making the correlation structure in the volatility process more flexible. For one model, we implement a method for introducing quasi long-memory into the volatility model. We demonstrate that the models can be fitted to real share price returns data, and that results indicate that for the series we study, the long-memory aspect of the model is not supported.

Some key words: Volatility; Long-memory; Fractional Ornstein-Uhlenbeck Process; Power decay process

1. Introduction

This paper describes the development of a class of continuous time stochastic volatility (SV) models for use in statistical finance. Some of these models are generalisations of the SV models of Barndorff-Nielsen & Shephard (2001a) and have a more flexible correlation structure than the exponential decay of the BNS-SV models. The models we consider are motivated by the recent work of Wolpert & Taqqu (2005), and are novel in the context of SV modelling. In this paper, we utilize Markov chain Monte Carlo (MCMC) algorithms to perform inference for the models, applied to simulated and real share data. Our MCMC algorithms exploit a series representation of the latent volatility process that is a Lévy process without a Gaussian component. We describe the adaptation of the MCMC algorithms to this more general inferential setting.

1·1 *Volatility Modelling*

Most option pricing in finance is based on the standard Black-Scholes model (Black & Scholes (1973)). It is well-known that this model does not fit some observed properties of financial data and many generalisations have been proposed. SV models are one generalisation, where the volatility is allowed to vary over time. For a review of recent SV models see Carr et al. (2003) and Schoutens (2003).

The most popular continuous time SV model was proposed in Barndorff-Nielsen & Shephard (2001a) (referred to as the BNS model), where the volatility follows an Ornstein-Uhlenbeck (OU) equation, with increments driven by a background driving Lévy process (BDLP). The volatility process is stationary, with jumps in volatility caused by jumps in the Lévy process. Between jumps, the volatility decays exponentially at a rate determined by one of the parameters in the OU equation. For the BNS model, the marginal distribution of the volatility is completely specified by the type of BDLP which drives the OU equation, and is independent of the rate of exponential decay of the volatility process.

1·2 *Extending the Ornstein-Uhlenbeck Model*

In the BNS-SV model the unknown volatility exhibits short-range dependence, or short-memory. It has been noted that the returns series can exhibit long-range dependence, or long-memory. One approach to long-memory modelling in this context has been to use process superposition: Roberts et al. (2004) and Griffin & Steel (2003) attempt to induce long-memory using superposition of volatility processes each with its own BDLP and correlation parameter - this result is theoretically justified by a theorem in Barndorff-Nielsen & Shephard (2001a). This model is deficient in a number of ways. First, although a superposition of finite numbers of volatility processes allows different BDLPs to describe short-range and longer-range dependencies in the volatility, the resulting process is still short-memory. Secondly, inference for the set of λs is not straightforward. Thus, superposition, although theoretically appealing, is of limited practical use if a real attempt at long-memory modelling is to be made.

1·3 *Contribution of this paper*

The stochastic processes utilized in this paper were introduced in Wolpert & Taqqu (2005), and are novel in the modelling of volatility. The purpose of this paper is to introduce a new class of stochastic volatility models and some of its properties, along with methods to simulate from them. In this paper, we use the series representation of the stochastic processes given in Barndorff-Nielsen & Shephard (2000) to facilitate computational inference. A useful property of one of the proposed models is that it does not use superposition to produce long-memory. Specifically on the long-memory issue, we introduce a volatility process whose correlation structure decays asymptotically like $t^{-\lambda}$, where $\lambda > 1$ is a parameter of the model, such that in the limit as $\lambda \to 1$, long-memory structure is recovered.

2. Volatility Process representations

2·1 *The Barndorff-Nielsen and Shephard Model*

The returns of financial series are often rescaled so that they are of a reasonable size and so it is attractive for volatility to have a *self-decomposable* distribution, as the marginal distribution is altered in a predictable way by rescaling. Wolfe (1982) proved that $\sigma^2(t)$ has a *self-decomposable* distribution if and only if it can be written as

$$\sigma^2(t) = \int_{-\infty}^{0} \exp(s)\, dz(\lambda t + s),$$

where λ is any positive constant and $z(t)$ is a homogeneous Lévy process (see for example Bertoin (1994) and Sato (1999)), referred to as the background driving Lévy process (BDLP). It follows that

$$\sigma^2(t) = \sigma^2(0)\, e^{-\lambda t} + e^{-\lambda t} \int_0^t e^{\lambda s} dz(\lambda s) \tag{2.1}$$

or, equivalently,

$$d\sigma^2(t) = -\lambda \sigma^2(t) + dz(\lambda t). \tag{2.2}$$

This SV model is described in Barndorff-Nielsen & Shephard (2001a).

The BDLP is constant apart from where it has positive jumps. Thus $\sigma^2(t)$ jumps when the BDLP jumps and decays exponentially in-between jumps (where $dz(\lambda t) = 0$).

Therefore, λ controls both the rate at which jumps occur in $\sigma^2(t)$ and the rate at which the volatility decays in-between jumps.

For any self-decomposable distribution, there is a unique BDLP, $z(t)$, that will generate the required marginal for the volatility in equation (2.2). The relationship between the marginal and the BDLP is given by the Lévy-Khintchine Formula (see, for example, Bertoin (1994)), and an important component of this representation is the Lévy measure, denoted here $u(x)$. The Lévy measure plays an integral part in the simulation procedures described below.

If a self-decomposable marginal distribution for $\sigma^2(t)$ is chosen, with Lévy measure $u(x)$, and if $z(1)$ has Lévy measure $w(x)$, Barndorff-Nielsen & Shephard (2000) have shown that if $\sigma^2(t)$ follows the OU equation (2.2), then

$$w(x) = -u(x) - x\frac{du(x)}{dx} \tag{2.3}$$

and, if the infinitely divisible marginal distribution for $\sigma^2(t)$ is chosen, the BDLP is specified by equation (2.3). Following Barndorff-Nielsen & Shephard (2001a), define the Tail Mass function as

$$W_p^+(x) = \int_x^\infty w(y)\,dy = xu(x)$$

and the Inverse Tail Mass function as

$$W_p^{-1}(x) = \inf\left[y > 0 : W_p^+(y) \leq x\right], \tag{2.4}$$

where p are the parameters specifying the exact marginal distribution of $\sigma^2(t)$. These are both monotonic decreasing functions.

We now study two specific examples.

- **The Generalized Inverse Gaussian Marginal** The Generalized Inverse Gaussian (GIG) model is considered extensively by Barndorff-Nielsen & Shephard (2001a), and inference for this general model is described in Gander & Stephens (2004). If $X \sim GIG(\gamma, \nu, \alpha)$, for $\gamma \in \mathbb{R}$ and $\nu, \alpha > 0$, the marginal density is

$$f_X(x) = \frac{(\alpha/\nu)^\gamma}{2K_\gamma(\nu\alpha)} x^{\gamma-1} \exp\left\{-\frac{1}{2}\left(\nu^2 x^{-1} + \alpha^2 x\right)\right\}, \quad \text{for} \quad x > 0,$$

where K_ν is a modified Bessel function of the third kind. The Lévy measure of X is then

$$u(x) = \frac{1}{x}\left\{\frac{1}{2}\int_0^\infty \exp\left(-\frac{x\xi}{2\nu^2}\right) g_\gamma(\xi)\, d\xi + \max(0,\gamma)\right\} \exp\left(-\frac{\alpha^2 x}{2}\right), \qquad (2.5)$$

where

$$g_\gamma(x) = \frac{2}{x\pi^2}\left\{J^2_{|\gamma|}\left(\sqrt{x}\right) + N^2_{|\gamma|}\left(\sqrt{x}\right)\right\}^{-1}$$

and $J_{|\nu|}$ and $N_{|\nu|}$ are Bessel functions of the first and second kind respectively. Special cases of this distribution include the Gamma, Inverse Gamma, Inverse Gaussian and Positive Hyperbolic. It is thus a flexible model that can be used for modelling real volatility processes. For example, in Gander & Stephens (2004), evidence is presented to suggest that the Inverse Gamma is preferable (in terms of out of sample prediction) for modelling the volatility of stocks on the New York Stock Exchange (NYSE).

- **The Tempered Stable Marginal** Another process considered in Gander & Stephens (2004) is the Tempered Stable process $TS(\kappa,\nu,\alpha)$. If $X \sim TS(\kappa,\nu,\alpha)$, for $0 < \kappa < 1$ and $\nu, \alpha > 0$, the marginal density is

$$f_X(x) = e^{\nu\alpha} f_{Y|\kappa,\nu}(x) \exp\left(-\frac{\alpha^{1/\kappa}}{2}x\right), \quad \text{for} \quad x > 0,$$

where, for $x > 0$,

$$f_{Y|\kappa,\nu}(x) = \frac{\nu^{-1/\kappa}}{2\pi}\sum_{j=1}^{\infty}\frac{(-1)^{j-1}}{j!}\sin(j\kappa\pi)\,\Gamma(j\kappa+1)\,2^{jk+1}\left(x\nu^{-1/\kappa}\right)^{-j\kappa-1}$$

is the density function of the positive κ-stable law (see Feller (1971) and Barndorff-Nielsen & Shephard (2001b)); if $\kappa = 0.5$ the *Inverse Gaussian* distribution is recovered. The Lévy measure of X is then

$$u(x) = Ax^{-B-1}e^{-Cx}, \qquad (2.6)$$

where $A = \nu\kappa 2^\kappa/\Gamma(1-\kappa)$, $B = \kappa$ and $C = \alpha^{1/\kappa}/2$. For this Lévy measure the Inverse Tail Mass function is

$$W^{-1}_{\kappa,\nu,\alpha}(x) = \left(\frac{A}{x}\right)^{1/B} \exp\left[-L_W\left(\frac{C}{B}\left(\frac{A}{x}\right)^{1/B}\right)\right],$$

where L_W is the Lambert-W function which satisfies

$$L_W(x)\exp[L_W(x)] = x \tag{2.7}$$

and is a standard function available numerically. For further details on L_W, see Jeffrey et al. (1996).

The last result needed, before we can sample from σ_i^2, is how to sample from stochastic integrals with respect to the BDLP, of the form given in equation (2.1).

2·2 *Integrated Volatility*

The integrated volatility process, $\{\sigma_t^{2*}\}$, related to $\{\sigma_t^2\}$ is defined as

$$\sigma^{2*}(t) = \int_0^t \sigma^2(u)\,du.$$

This is an important quantity for pricing European options and, for the BNS-SV model, it can be shown that

$$\sigma^{2*}(t) = \frac{1}{\lambda}\left\{z(\lambda t) - \sigma^2(t) + \sigma^2(0)\right\}.$$

This relatively simple form for the integrated volatility is an attractive feature of the model. The discretely observed or actual volatility is

$$\sigma_i^2 = \sigma^{2*}(i\Delta) - \sigma^{2*}((i-1)\Delta). \tag{2.8}$$

Barndorff-Nielsen & Shephard (2001a) have shown that

$$\operatorname{corr}\left\{\sigma_i^2, \sigma_{i+s}^2\right\} = d(\lambda\Delta)\,e^{-\lambda\Delta(s-1)},$$

where $d(\lambda\Delta)$ is independent of s and $0 < d(\lambda\Delta) < 1$. The log of the underlying asset, $x(t)$, satisfies

$$dx(t) = \left\{\mu - \frac{\sigma^2(t)}{2}\right\}dt + \sigma(t)\,dW(t)$$

and, if inference about μ and σ_i^2 is required, the likelihood for $y_1, \ldots, y_T$ is given by noting that

$$y_i \sim N\left(\left(\mu - \frac{\sigma_i^2}{2}\right)\Delta, \sigma_i^2\Delta\right).$$

2·3 *The Griffin and Steel Representation*

Griffin & Steel (2003) have shown that the discretely observed volatility can be written as

$$\sigma_i^2 = \frac{1}{\lambda}\left\{\eta_{i,2} - \eta_{i,1} + \left(1 - e^{-\lambda\Delta}\right)\sigma^2\left(\left(i-1\right)\Delta\right)\right\}, \tag{2.9}$$

where

$$\left\{\begin{array}{c} \sigma^2\left(i\Delta\right) \\ z\left(\lambda i\Delta\right) \end{array}\right\} = \left\{\begin{array}{c} e^{-\lambda\Delta}\sigma^2\left(\left(i-1\right)\Delta\right) \\ z\left(\lambda\left(i-1\right)\Delta\right) \end{array}\right\} + \eta_i$$

and

$$\eta_i = \left\{\begin{array}{c} e^{-\lambda\Delta}\int_0^{\Delta} e^{\lambda t} dz\left(\lambda t\right) \\ \int_0^{\Delta} dz\left(\lambda t\right) \end{array}\right\} = \left\{\begin{array}{c} e^{-\lambda\Delta}\int_0^{\lambda\Delta} e^{t} dz\left(t\right) \\ \int_0^{\lambda\Delta} dz\left(t\right) \end{array}\right\} \tag{2.10}$$

is a vector of random jumps, equal to a stochastic integral with respect to the BDLP, $z\left(t\right)$.

Barndorff-Nielsen & Shephard (2000) proved that if $f\left(s\right) \geq 0$ for $0 < s < \Delta$ and, if $f\left(s\right)$ is integrable with respect to $dz\left(s\right)$, then

$$\int_0^{\Delta} f\left(s\right) dz\left(s\right) \overset{\mathfrak{L}}{=} \sum_{i=1}^{\infty} W_p^{-1}\left(a_i/\Delta\right) f\left(\Delta e_i\right), \tag{2.11}$$

where $W_p^{-1}\left(\right)$ is the Inverse Tail Mass function as defined in equation (2.4), a_i are the arrival times of a Poisson process of intensity 1 and e_i are independent standard uniform variates (also independent of a_i). Note that $W_p^{-1}\left(a_i/\Delta\right) \geq 0$ is a decreasing function and that, if it is non-zero for large a_i, the integral can be approximated by truncating the infinite series at some point.

Assume that η_i is truncated by discarding all Poisson points which are greater than a_c (so the same truncation scheme is used for each element of the random shock vector). Let n_i be the number of Poisson points which are less than a_c for the i^{th} entry of the random shock vector (i.e. the number of Poisson points which contribute to η_i). The approximation to equation (2.10) is then

$$\eta_i \overset{\mathfrak{L}}{=} \left\{\begin{array}{c} e^{-\lambda\Delta}\sum_{j=1}^{n_i} W_p^{-1}\left(\frac{a_{i,j}}{\lambda\Delta}\right) e^{\lambda\Delta r_{i,j}} \\ \\ \sum_{j=1}^{n_i} W_p^{-1}\left(\frac{a_{i,j}}{\lambda\Delta}\right) \end{array}\right\}, \tag{2.12}$$

where $a_{i,j}$ and $r_{i,j}$ are Poisson points and uniforms as described previously. We let A and R be the matrices of the $a_{i,j}$ and $r_{i,j}$.

The method to sample from σ_i^2 is as follows: we select a self-decomposable distribution for $\sigma^2(t)$, and find the Lévy measure of this distribution and then the Lévy measure of $z(1)$, using equation (2.3). We then truncate the Poisson point process at a_c and use equations (2.4), (2.9) and (2.12) to generate σ_i^2. Details on how a_c might be chosen are given in Gander & Stephens (2004).

3. Ornstein-Uhlenbeck processes: Alternative representation

The OU process defined in equation (2.2) has solution

$$\begin{aligned}\sigma^2(t) &= \int_{-\infty}^{t} f_2(\lambda,t,s)\,dz(\lambda s) && (3.1)\\ &= \int_0^{\infty} f_1(\lambda,t,s)\,dz(\lambda s)+\int_0^{t} f_2(\lambda,t,s)\,dz(\lambda s) && (3.2)\\ &= e^{-\lambda t}\sigma^2(0)+e^{-\lambda t}\int_0^{t} e^{\lambda s}dz(\lambda s),\end{aligned}$$

where the two Lévy processes of equation (3.2) are independent copies of each other (i.e. series representations for the stochastic integrals use independent realisations from the same Lévy process) and

$$f_1(\lambda,t,s)=e^{-\lambda(t+s)} \qquad f_2(\lambda,t,s)=e^{-\lambda(t-s)}.$$

The process has correlation structure $corr\left(\sigma^2(t),\sigma^2(t+j)\right)=\exp(-\lambda j)$ and $\sigma^2(t)$ is stationary and positive for a wide range of functions f_1 and f_2. Barndorff-Nielsen & Shephard (2001a) mention using models with more general functions f_1 and f_2 and decide to concentrate on OU models, where f_1 and f_2 are as described above.

The timing of the BDLP, $dz(\lambda s)$ in equation (2.2) is chosen so that λ does not influence the marginal distribution of $\sigma^2(t)$. Rather than equation (3.2), consider an amended version

$$\sigma^2(t)=\int_0^{\infty} f_1(\lambda,t,s)\,dz(s)+\int_0^{t} f_2(\lambda,t,s)\,dz(s)=I_{1,t}+I_{2,t}. \qquad (3.3)$$

This is the same representation as used in Wolpert & Taqqu (2005). Unlike equation (3.2), the timing (and the rate of jumps of the Lévy process) now does depend on λ,

unlike the representations used for BNS-SV models, where the rate of jumps of the Lévy process is not influenced by λ.

Simulation from the OU process is relatively straightforward because the time dependent term of f_1 and f_2 can be removed from the stochastic integrals of equation (3.2) and this allows $I_{1,t}$ to be written in terms of the volatility at time zero, $\sigma^2(0)$, in equation (3.3). If the OU equation is generalized, so integrands are not of the form $f_1(t,s) = g_{1,1}(t)\, g_{1,2}(s)$, then $\sigma^2(t)$ can no longer be expressed in terms of $\sigma^2(0)$. For general f_1 and f_2, it is also not possible to separate the t and s terms in $I_{2,t}$. This makes simulating from such models more complicated than the original BNS-SV OU models.

4. Simulation from the generalized model

We now describe how to sample from the models of the previous section. Consider the approximation for $I_{1,t}$,

$$I_{1,t} \approx \int_0^d f_1(\lambda, t, s)\, dz(s),$$

where d is large enough so the approximation is sufficiently accurate. The Barndorff-Nielsen & Shephard (2000) series representation is then

$$\int_0^d f_1(\lambda, t, s)\, dz(s) \overset{\mathfrak{L}}{=} \sum_{j=0}^{n_{1,j}} W_p^{-1}(a_j/d)\, f_1(\lambda, t, r_j), \tag{4.1}$$

where $da_{1,c}$ is the value at which the Poisson point process (order statistics of uniform random variables) is truncated, $n_{1,j} \sim Po(da_{1,c})$, a_j are the order statistics of $n_{1,j}$ $U(0, da_{1,c})$ random variables, $r_j \overset{iid}{\sim} U(0,d)$, all variables are independent and $W_p^{-1}()$ is the Inverse Tail Mass function as defined previously. For every t, the same Poisson points, a_j, and uniforms, r_j, are used and this induces the correlation in $I_{1,t}$, so $I_{1,t} = e^{-\lambda t}\sigma^2(0)$ for the OU case. This allows us to sample from $I_{1,t}$.

For the finite integral, the situation is more complex. Previously a series representation was used, based on independent Poisson point processes and uniforms and this was possible because the volatility could be written in terms of the previous volatility and a stochastic integral (independent of previous stochastic integrals). Further, the stochastic integrals were unaltered by t. For more general functions than $f_2(\lambda, t, s) = e^{-\lambda(t-s)}$, it

is not possible to write the volatility in terms of previous volatilities, though we are able to express the integral as a summation of integrals on disjoint domains and then use independent series representations for these integrals. The second integral at time $t-1$ is

$$I_{2,t-1} = \int_0^{t-1} f_2(\lambda, t-1, s)\, dz(s)$$

and now consider $I_{2,t}|I_{2,t-1}$

$$I_{2,t} = \int_0^{t-1} f_2(\lambda, t, s)\, dz(s) + \int_{t-1}^{t} f_2(\lambda, t, s)\, dz(s). \tag{4.2}$$

The domains of these two integrals are disjoint and so any realisations from these integrals use independent series representations. Equation (4.2) can be rewritten as

$$I_{2,t} = \sum_{j=0}^{t-2} \int_j^{j+1} f_2(\lambda, t, s)\, dz(s) + \int_{t-1}^{t} f_2(\lambda, t, s)\, dz(s)$$

but the integrals of the summation are also disjoint, so by the independent increments assumption,

$$\begin{aligned} I_{2,t} &\overset{\mathfrak{L}}{=} \sum_{j=0}^{t-2} \int_0^1 f_2(\lambda, t, s+j)\, dz(s) + \int_0^1 f_2(\lambda, t, s+t-1)\, dz(s) \\ &= \sum_{j=0}^{t-1} \int_0^1 f_2(\lambda, t, s+j)\, dz(s), \end{aligned}$$

where integral terms in the sum are all with respect to independent realisations of the BDLP (as they represent partitions of the integrals in equation (4.2)). This gives t disjoint independent integrals and these can be simulated using the series representation derived in Barndorff-Nielsen & Shephard (2000) and given in equation (2.11). If the series are again truncated by discarding all Poisson points which are greater than $a_{2,c}$, then the series representation is

$$I_{2,t} \overset{\mathfrak{L}}{=} \sum_{j=0}^{t-1} \sum_{i=0}^{n_{2,j}} W_p^{-1}(a_{2,j,i})\, f_2(\lambda, t, r_{2,j,i} + j), \tag{4.3}$$

where $n_{2,j} \sim Po(a_{2,c})$, $a_{2,j}$ are the order statistics of $n_{2,j}$ $U(0, a_{2,c})$ random variables, $r_{2,j,i} \overset{iid}{\sim} U(0,1)$, all variables are independent of each other and $W_p^{-1}()$ is the Inverse Tail Mass function as defined previously. For the OU process, simulating using these series representations gives the properties of $\sigma^2(t)$ that were discussed in Gander & Stephens

(2004). This is illustrated in Figure 1). We are now able to simulate from processes of the form of equation (3.3) for general f_1 and f_2. Note that simulating from the instantaneous volatility using the series representation of equation (4.3) is an order t^2 algorithm, unlike the series representation that was used for the OU process in previous inference schemes (Roberts et al. (2004), Griffin & Steel (2003), Gander & Stephens (2004)) which are of order t. We now consider which forms of these functions should be examined.

5. General SV models driven by Lévy processes

We retain the assumption that equation (2.1) is driven by the homogeneous BDLP, $z(t)$, where $z(1)$ has Lévy measure

$$w(x) = -u(x) - xu(x),$$

where $u(x)$ is the Lévy measure of the marginal distribution of the BNS-SV model with the same BDLP. We will focus on marginal distributions on the positive real line, so z is a subordinator. The Lévy-Khintchine formula for $z(1)$ is

$$\log \mathrm{E}\left\{e^{i\theta z(1)}\right\} = \int_{-\infty}^{\infty} \left(e^{i\theta x} - 1\right) w(x)\, dx,$$

and $w(x)$ is zero for $x \leq 0$. We shall develop models for the SV process using the approach introduced in Wolpert & Taqqu (2005) in a different modelling context. Full details can be found in that paper; here we give brief details of the representation, concentrating on moment-properties of the resulting stochastic processes. In section 6 we give details of the MCMC algorithms used to make inferences from simulated and real data.

5·1 *Utilizing the Wolpert and Taqqu Representation*

Consider a moving average process, $\{X_t\}$, with representation

$$X_t = \int_0^t f(t,s)\, dz(s),$$

where $z(s)$ is a Lévy process. The function $G(s)$ is *non-anticipating* with respect to $dz(s)$ if $G(s)$ cannot be used to predict future movement in $dz(s)$. The process

$$X_t = \int_{t_0}^{t_n} G(s)\, dz(s)$$

is then also non-anticipating. Consider non-anticipating moving average models for $\sigma^2(t)$ of the form

$$\sigma^2(t) = \int_{-\infty}^{t} h_1(t-s)\, dz(s), \tag{5.1}$$

where the Lévy measure of $z(1)$ is $w(x)$ and $h_1(t-s) \geq 0$ for $s < t$ (so $\sigma^2(t)$ has only positive jumps). Ignoring the timing of the BDLP, this is a generalisation of the solution given in equation (3.1). Therefore

$$\sigma^2(t) = \int_{-\infty}^{\infty} h(t-s)\, dz(s), \tag{5.2}$$

where $h(x) = h_1(x)$ if $x \geq 0$ and zero otherwise. For models of the form of equation (5.2), the negative of the characteristic exponent is

$$\begin{aligned}
\log \mathrm{E}\left\{e^{i\theta\sigma^2(t)}\right\} &= \log \mathrm{E}\left\{\exp\left(i\theta \int_{-\infty}^{\infty} h(t-s)\, dz(s)\right)\right\} \\
&= \log \mathrm{E}\left\{\exp\left(i\theta \sum_{j=-\infty}^{\infty} \int_{j\Delta}^{(j+1)\Delta} h(t-s)\, dz(s)\right)\right\}
\end{aligned}$$

and as $\sigma^2(t)$ is non-anticipative,

$$\begin{aligned}
\log \mathrm{E}\left\{e^{i\theta\sigma^2(t)}\right\} &= \log \mathrm{E}\left\{\exp\left(i\theta \sum_{j=-\infty}^{\infty} h(t-j\Delta)\left(z\left((j+1)\Delta\right) - z(j\Delta)\right)\right)\right\} \\
&= \log \mathrm{E}\left\{\prod_{j=-\infty}^{\infty} \exp\left(i\theta h(t-j\Delta)\, z_j(\Delta)\right)\right\},
\end{aligned}$$

where $z_j(t)$ are independent and identical homogeneous Lévy processes with Lévy measure $w(x)$. Then

$$\log \mathrm{E}\left\{e^{i\theta\sigma^2(t)}\right\} = \sum_{j=-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{\exp\left(i\theta h(t-j\Delta)\, x\right) - 1\right\} w(x)\, dx$$

and letting $\Delta \to 0$ this gives

$$\log \mathrm{E}\left\{e^{i\theta\sigma^2(t)}\right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{\exp\left(i\theta h(t-s)\, x\right) - 1\right\} w(x)\, dx ds$$

and

$$\log \mathrm{E}\left\{e^{i\theta\sigma^2(t)}\right\} = \int_{-\infty}^{\infty} \int_{0}^{\infty} \left(e^{i\theta x h(s)} - 1\right) ds w(x)\, dx, \tag{5.3}$$

The variance and covariance of the process can be calculated by considering the joint characteristic function in the usual way. We have

$$\text{Cov}\left\{\sigma^2(t), \sigma^2(0)\right\} = \tau^2 \int_0^\infty h_1(t+s)\, h_1(s)\, ds$$

where τ^2 is the marginal variance induced by the model, and the process variance is given by

$$\text{Var}\left\{\sigma^2(t)\right\} = \tau^2 \int_0^\infty h_1^2(s)\, ds, \tag{5.4}$$

which we require to be finite. The correlation at lag t is

$$\rho(t) = \frac{\int_0^\infty h_1(|t|+s)\, h_1(s)\, ds}{\int_0^\infty h_1^2(s)\, ds}. \tag{5.5}$$

By picking suitable functions for $h_1(x)$, we are able to generate from a wide range of distributions and correlation structures which have Lévy measure and correlation structure specified by equations (5.3) and (5.5) respectively. In general, the discretely observed volatility in equation (2.8) is not readily available and this makes it difficult - although not impossible - to fit SV models of this form using the discretely observed volatility. Three examples of the flexibility of models of this form are now given, before fractional Ornstein-Uhlenbeck processes are introduced.

5·2 *Ornstein-Uhlenbeck process*

The marginal distribution of the BNS-SV OU volatility models is unaltered by the λ parameter. Rather than using this representation of the OU process, consider instead using $h_1(t-s) = \sqrt{2\lambda}e^{-\lambda(t-s)}$ in equation (5.1). This is the OU process used in Wolpert & Taqqu (2005) where λ influences the marginal distribution of $\sigma^2(t)$. The correlation specified by equation (5.5) is $\rho(t) = e^{-\lambda t}$ as for the BNS-SV OU models. The relationship between the marginal and λ is now given, when the BDLPs of Barndorff-Nielsen & Shephard (2001a) drive the OU process.

Substituting $r = xh(s)$ in equation (5.3) implies the negative of the characteristic exponent is

$$\frac{1}{\lambda}\int_0^\infty \int_0^{x\sqrt{2\lambda}} \left\{e^{itr} - 1\right\} r^{-1} w(x)\, dr dx$$

and exchanging the order of integration, this is

$$\frac{1}{\lambda}\int_0^\infty \int_{r/\sqrt{2\lambda}}^\infty \left\{e^{itr}-1\right\} r^{-1} w\left(u\right) dxdr$$

and so $\sigma^2\left(t\right)$ has Lévy measure

$$\frac{1}{\lambda} r^{-1} \int_{r/\sqrt{2\lambda}}^\infty w\left(x\right) dx = \frac{1}{\lambda} r^{-1} \left[-xu\left(x\right)\right]_{r/\sqrt{2\lambda}}^\infty = \frac{1}{\lambda\sqrt{2\lambda}} u\left(r/\sqrt{2\lambda}\right).$$

When the BDLP that gives a $GIG\left(\gamma,\nu,\alpha\right)$ marginal for the BNS-SV OU model is used to drive the OU equation, using equation (2.5), the Lévy measure of $\sigma^2\left(t\right)$ is

$$\frac{1}{\lambda} r^{-1} \left[\left\{\frac{1}{2}\int_0^\infty \exp\left(-\frac{x\xi}{2\nu^2\sqrt{2\lambda}}\right) g_\gamma\left(\xi\right) d\xi + \max\left(0,\gamma\right)\right\} \exp\left(-\frac{\alpha^2 x}{2\sqrt{2\lambda}}\right)\right].$$

In general, it is not possible to write the distribution of $\sigma^2\left(t\right)$ in terms of a GIG distribution because of the complex nature of the integrand. However, to illustrate the use of these models, we use the BDLP which gives a $Ga\left(\nu,\alpha\right)$ $\left(GIG\left(\nu,0,\sqrt{2\alpha}\right)\right)$ distribution for the BNS-SV model, so the integral is zero. Then

$$\sigma^2\left(t\right) \sim Ga\left(\frac{\nu}{\lambda}, \frac{\alpha}{\sqrt{2\lambda}}\right).$$

This marginal distribution is verified in Figure 1, which also confirms that the correlation structure is $e^{-\lambda t}$, as given by equation (5.5). The simulation results of Figure 1 are as the theory suggests. This demonstrates the correct implementation of the series representation of Section 3 for the OU process.

When the BDLP, which gives a $TS\left(\kappa,\nu,\alpha\right)$ marginal for the BNS-SV OU model, is used to drive equation (5.7), using equation (5.6) and equation (2.6), the Lévy measure of $\sigma^2\left(t\right)$ is

$$A' r^{-B'-1} e^{-C' x},$$

where

$$A' = \frac{A}{\lambda\left(2\lambda\right)^{\kappa/2}}, \quad B' = B \quad \text{and} \quad C' = \frac{C}{\sqrt{2\lambda}}$$

and A, B and C are as defined under equation (2.6). From this it can be shown that

$$\sigma^2\left(t\right) \sim TS\left(\kappa, \nu\lambda^{-1}\left(2\lambda\right)^{-\kappa/2}, \alpha\left(2\lambda\right)^{-\kappa/2}\right)$$

and so the *IG-OU* BDLP generates $\sigma^2\left(t\right) \sim IG$ (with different parameters).

5·3 *Power Decay process*

We now consider a SV model whose correlation decays asymptotically like a power, and thus that has the potential to exhibit long-memory behaviour. Let

$$h_1(t-s) = \frac{1}{(\alpha + \beta|t-s|)^{\lambda}} \tag{5.6}$$

in equation (5.1). We consider the case $\lambda > 1$ only, as this ensures that the correlation is monotonically decreasing in $|t-s|$. We will focus on the case $\alpha = 1$, as other α values do not offer a richer correlation structure, as their effect only rescales the β parameter. Substituting $r = xh(s)$ in equation (5.3) implies the negative of the characteristic exponent is

$$\frac{1}{\beta\lambda}\int_{-\infty}^{\infty}\int_{0}^{1}\left(e^{itr}-1\right)x^{1/\lambda}r^{-(1+1/\lambda)}drw(x)\,dx.$$

Therefore the Lévy measure of $\sigma^2(t)$ is

$$\frac{1}{\beta\lambda}r^{-(1+1/\lambda)}\int_{0}^{1}x^{1/\lambda}w(x)\,dx.$$

Both the Lévy measure and correlation structure of $\sigma^2(t)$ can be expressed in terms of standard numerical functions, though these expressions are complex. For this reason we focus on the cases $\lambda = 1.5$ and $\lambda = 2$, which will be used for simulation purposes later. In the case of the *Ga(ν, α_2)-OU* BDLP, the Lévy measure is

$$\frac{3\nu}{5\alpha_2^{1/3}}e^{-\alpha_2/2}W_M\left(\frac{1}{3},\frac{5}{6},\alpha_2\right)r^{-5/3} \qquad \lambda = 1.5$$

$$\frac{\nu}{2\sqrt{\alpha_2}}\left\{\sqrt{\pi}\mathrm{erf}\left(\sqrt{\alpha_2}\right)-2\sqrt{\alpha_2}e^{-\alpha_2}\right\}r^{-3/2} \qquad \lambda = 2,$$

where erf is the error function, $W_M(\mu,\nu,z)$ is the Whittaker-M special function

$$W_M(a,b,z) = e^{-z/2}z^{b+1/2}\Phi\left(\frac{1}{2}+b-a,1+2b,z\right)$$

and Φ is the confluent hypergeometric function

$$\Phi(a,b,z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_{0}^{1}e^{zt}t^{a-1}(1-t)^{b-a-1}dt.$$

Both functions are available numerically. These models are limiting cases of the $TS\left(1/\lambda, \nu_2, \alpha_3\right)$ distribution as $\alpha_3 \to 0$ (where ν_2 is determined by ν, α_2 and λ). The correlation is specified by equation (5.5) and is

$$\frac{4}{\beta^2}\frac{\left(2+\beta t-2\sqrt{1+\beta t}\right)}{t^2\sqrt{1+\beta t}} \qquad \lambda = 1.5$$

$$\frac{3}{\beta^3 t^3\left(1+\beta t\right)}\left\{2\left(1+\beta t\right)\log\left(\frac{1}{1+\beta t}\right)+\beta t\left(2+\beta t\right)\right\} \qquad \lambda = 2.$$

The asymptotic decay of the correlation is proportional to

$$\lim_{t\to\infty}\int_0^\infty \frac{1}{\left\{\left(1+\beta s\right)\left(1+\beta\left(t-s\right)\right)\right\}^\lambda}ds,$$

truncating the integral at some large K ($\ll t$) gives the correlation proportional to

$$\lim_{t\to\infty}\int_0^K \frac{1}{\left\{\left(1+\beta s\right)\left(\beta t\right)\right\}^\lambda}ds$$

and so the asymptotic decay in the correlation is $t^{-\lambda}$. In the limit as $\lambda \to 1$, the asymptotic decay in the correlation tends to t^{-1} so the model approaches a long-memory model. For $\lambda = 1.5$ the correlation decays asymptotically like $t^{-3/2}$, and for $\lambda = 2$ it decays like t^{-2}, and so on. This gives a slower decay than the BNS-SV OU models. Figure 2 are ACF plots for simulations of size $50,000$ from this process, when $\beta = 0.1$ and for $\lambda = 1.5$ and $\lambda = 2$, using the $Ga\left(1,1\right) - OU$ BDLP and demonstrate the correct decay of the correlation of the volatility process.

As λ increases the asymptotic decay of the process increases. The dashed line shows the theoretical correlation and suggests that the series representation of Section 3 has been implemented correctly. The β parameter can be used to further control the correlation structure. Figure 2 are the same ACF plots as Figure 3 but for $\beta = 1$.

For $\beta = 1$, the initial decay in the correlation is faster than when $\beta = 0.1$. Models with $h_1\left(t-s\right)$ given by equation (5.6) can control the initial decay of the volatility through the β parameter and the asymptotic decay by the λ parameter.

5·4 *Fractional Ornstein-Uhlenbeck process*

Rather than using the Ornstein-Uhlenbeck process, we consider the OU process with solution

$$\sigma^2(1,t) = \sqrt{2\lambda}\int_{-\infty}^{t} e^{-\lambda(t-s)}dz(s), \tag{5.7}$$

as was used in Section 5·2.

The Riemann-Liouville operator of fractional integration of a function, $f(s)$, is defined by

$${}_aD^{-n}f(s) = \frac{1}{\Gamma(n)}\int_a^t (t-s)^{n-1} f(s)\,ds, \tag{5.8}$$

where D^{-n} is the n-fold integral (see Anh & McVinish (2003)). Define $\sigma^2(\kappa,t)$ as

$$\sigma^2(\kappa,t) = \int_{-\infty}^{t} \lambda e^{-\lambda(t-s)}\sigma^2(\kappa-1,t)(s)\,ds, \tag{5.9}$$

for $\kappa \neq 1$ (the $\kappa = 1$ case has been covered in Section 5·2). It can be shown that

$$\sigma^2(\kappa,t) = \sqrt{2\lambda}\int_{-\infty}^{t} \frac{\lambda^{\kappa-1}}{\Gamma(\kappa)}(t-s)^{\kappa-1} e^{-\lambda(t-s)}dz(s)$$

(see Wolpert & Taqqu (2005)). The process $\sigma^2(\kappa,t)$ is therefore called the *fractional Ornstein-Uhlenbeck Lévy* (or *fOUL*) process, as it is of the form of equation (5.8) with $n = \kappa$ and

$$f(s) = \sqrt{2\lambda}\lambda^{\kappa-1}e^{-\lambda(t-s)}.$$

Equations (5.7) and (5.9) are equivalent to those used in Wolpert & Taqqu (2005) to define fOUL processes. Again, unlike the BNS-SV OU models, the marginal distribution of the volatility is influenced by λ for these OU processes. We shall not change the timing of the BDLP to avoid this, as although this is possible with the OU solution, for the fOUL solution, κ also alters the marginal and it is difficult to manipulate equation (5.9) to ensure this is not the case (and therefore we will be unable to make the marginal independent of both λ and κ). Wolpert & Taqqu (2005) use fOUL processes to model the Telecom process and are interested in the covariance function of these processes, whilst we are usually concerned with the correlation function. The fOUL process is a special case of equation (5.2), when

$$h(x) = \sqrt{2\lambda}\frac{\lambda^{\kappa-1}}{\Gamma(\kappa)}x^{\kappa-1}e^{-\lambda x} \quad x \geq 0$$

and zero otherwise. The variance of the process is given by equation (5.4) and we restrict our attention to the finite variance processes, where $\kappa > 1/2$. Equation (5.5) gives the correlation function

$$\rho^{\kappa}(t) = \frac{2}{\Gamma(\kappa - 1/2)} \left(\frac{\lambda t}{2}\right)^{\kappa - 1/2} K_{\kappa - 1/2}(\lambda |t|).$$

For small lags, the correlation decays like a power for the fOUL process (unlike the OU process). Although the fOUL process has a more flexible correlation structure than the OU, both processes decay exponentially for large lags and do not have long-memory. The decay of the correlation for small lags is controlled by κ and the decay for large lags is determined by λ. Figure 4 demonstrates how the correlation structure varies with κ for constant $\lambda = 0.1$.

Unfortunately, using a similar method to that used in Section 5·2, to calculate the marginal distribution of the volatility, it is difficult to derive the Lévy measure or distribution of $\sigma^2(\kappa, t)$ for fOUL processes in general for the commonly used BDLPs.

As before, consider the homogeneous BDLP, $z(t)$, with Lévy measure

$$w(x) = -u(x) - xu(x),$$

where $u(x)$ is the Lévy measure of the marginal distribution of the BNS-SV model with the same BDLP. From equation (5.7), the negative of the characteristic exponent of $\sigma^2(\kappa, t)$ is

$$\log \mathrm{E}\left\{e^{it\sigma^2(\kappa,t)}\right\} = \int_0^{\infty}\int_0^{\infty} \left(e^{ituh(s)} - 1\right) w(x)\, ds dx,$$

where $h(s)$ is given by equation (5.9). Previously, we performed the substitution $r = uh(s)$ and more care must be taken to perform this substitution for general $h(s)$ as, for some κ values, the substitution is not one-to-one on the domain of integration. The difficulties for the case $1/2 < \kappa < 1$ are discussed below.

The substitution $r = uh(s)$ is one-to-one on $\mathbb{R}^+$ for $1/2 < \kappa < 1$ (unlike when $\kappa > 1$). However, this is still more complex than for the OU process, as

$$dr = uh'(s)\, ds$$

and

$$\frac{1}{uh'(s)} = \frac{1}{uh(s)} \frac{s}{(\kappa - 1 - \lambda s)} = r^{-1} \frac{h^{-1}(r/u)}{(\kappa - 1 - \lambda h^{-1}(r/u))},$$

where

$$h^{-1}(s) = x^{1/(1-\kappa)} \exp\left\{-L_W\left(\frac{\lambda x^{1/(1-\kappa)}}{1-\kappa}\right)\right\}.$$

and L_W is the Lambert-W function of equation (2.7). Then

$$\log \mathrm{E}\left\{e^{it\sigma^2(\kappa,t)}\right\} = \int_0^\infty \int_0^\infty \left\{e^{itr} - 1\right\} r^{-1} \frac{h^{-1}(r/u)}{(\lambda h^{-1}(r/u) + 1 - \kappa)} dr w(u)\, du$$

and $\sigma^2(\kappa, t)$ has Lévy measure

$$r^{-1} \int_0^\infty \frac{h^{-1}(r/u)}{(\lambda h^{-1}(r/u) + 1 - \kappa)} w(u)\, du.$$

Due to the complex nature of h^{-1}, it is not possible to simplify this further, even when we have a $Ga - OU$ BDLP. This demonstrates that both λ and κ specify the exact form of the marginal distribution of $\sigma^2(\kappa, t)$. Even for specific κ values (such as $\kappa = 3/4$) it is not possible to evaluate this integral analytically.

We will concentrate on simulating from such models with the $Ga - OU$ BDLP, as this gives finite summations in equations (4.1) and (4.3) and the Inverse Tail Mass function is available directly. Using the numerical methods, along with the series representation of Section 3, allows us to simulate from integrals with respect to any of the BDLPs used in Gander & Stephens (2004).

Even though we were unable to derive the marginal distribution of the volatility for this BDLP, empirical results suggest the volatility might be *Gamma* distributed when the $Ga - OU$ BDLP is used to drive the fOUL process.

6. Inference using MCMC

Inference for continuous-time stochastic volatility models has been attempted using MCMC. For example, MCMC algorithms to estimate the parameters of the BNS-SV OU models have been described in Roberts et al. (2004), Griffin & Steel (2003) and Gander & Stephens (2004). The algorithm in the final reference simulates from the stochastic integral of equation (2.10) using the series representation of equation (2.12). For a data set of size N this is an $O(N)$ algorithm. For the models described in this paper, in general, we need to sample from stochastic integrals of the form of equation (3.3) and can use a similar series representation, given in equations (4.1) and (4.3), which gives an $O\left(N^2\right)$ algorithm to sample from the instantaneous volatilities, $\sigma^2(0\Delta), \ldots, \sigma^2(N\Delta)$.

6·1 *Likelihood construction and Approximation*

The likelihood for BNS-SV OU models in Gander & Stephens (2004) is specified by the discretely observed volatility, σ_i^2, defined in equation (2.8) as

$$\sigma_i^2 = \int_{(i-1)\Delta}^{i\Delta} \sigma^2(u)\, du,$$

which has a simple form for the BNS-OU-SV models. In general, for the models of this paper, this discretely observed volatility is

$$\int_{(i-1)\Delta}^{i\Delta} \int_0^{\infty} f_1(\lambda, u, s)\, dz(s)\, du + \int_{(i-1)\Delta}^{i\Delta} \int_0^{t} f_2(\lambda, u, s)\, dz(s)\, du.$$

As the time dependent term of $\sigma^2(t)$ cannot be separated from the stochastic integral term, this cannot be simplified to a single integral, as was the case for the BNS-SV OU models. The series representations for the models of this paper are slower to implement because of the $O(t)$ series representation of equation (4.3). The double integrals for the discretely observed volatility are very intensive to compute, and are not currently feasible as part of an MCMC algorithm. Fortunately, however, the instantaneous and discretely observed volatilities have similar properties and so we can fit the models of this paper using the same likelihood as before but with the approximation

$$\sigma_i^2 = \int_{(i-1)\Delta}^{i\Delta} \sigma^2(u)\, du \approx \sigma^2((i-1)\Delta)\,\Delta.$$

Alternatively, we could use the approximation

$$\int_{(i-1)\Delta}^{i\Delta} \sigma^2(u)\, du \approx \sigma^2(i\Delta)\,\Delta,$$

or use the trapezium rule to make the approximation

$$\int_{(i-1)\Delta}^{i\Delta} \sigma^2(u)\, du \approx \left\{ \frac{\sigma^2(i\Delta) - \sigma^2((i-1)\Delta)}{2} \right\} \Delta.$$

For simulation purposes, each approximation gives similar results. For this reason we use the approximation $\sigma_i^2 \approx \sigma^2(i\Delta)\,\Delta$, as the correlation structure of $\sigma^2(i\Delta)$ is already known and gives a simple correlation structure for σ_i^2.

To test the correct fit of the models to observed data, we examine the observed and theoretical correlation structure of the square of the log returns, given the estimated

model parameters. To estimate corr $\left\{y_i^2, y_{i+s}^2\right\}$ for $s > 0$, we make the approximation $y_i \sim N\left(0, \sigma^2\left(i\Delta\right)\right)$, then $\mathrm{E}\left\{y_i^2\right\} = \mathrm{E}\left\{\sigma^2\left(i\Delta\right)\right\}$ and so

$$\begin{aligned}\mathrm{Cov}\left\{y_i^2, y_{i+s}^2\right\} &= E\left\{\left(\sigma^2\left(i\Delta\right)X_i^2 - \overline{\sigma^2\left(i\Delta\right)}\right)\left(\sigma^2\left(\left(i+s\right)\Delta\right)X_{i+1}^2 - \overline{\sigma^2\left(\left(i+s\right)\Delta\right)}\right)\right\} \\ &= \rho\left(s\right)\mathrm{Var}\left\{\sigma^2\left(i\Delta\right)\right\},\end{aligned}$$

where $X_i \overset{iid}{\sim} N\left(0,1\right)$ is independent of the volatility process and $\rho\left(t\right)$ is as given in equation (5.5). We also have

$$\mathrm{Var}\left\{y_i^2\right\} = \mathrm{Var}\left\{y_{i+s}^2\right\} = \mathrm{Var}\left\{\sigma^2\left(i\Delta\right)\right\}\mathrm{Var}\left\{X_i^2\right\} = 2\mathrm{Var}\left\{\sigma^2\left(i\Delta\right)\right\}$$

and so corr $\left\{y_i^2, y_{i+s}^2\right\} = \rho\left(s\right)/2$.

The MCMC algorithm is the same as previously, using the new stochastic integrals and series representations. Due to the intensive series representation, efficient coding is very important, so that the models run in sensible time. Note that when the b^{th} row of Poisson points and uniforms are updated, as

$$I_{2,t} \overset{\mathfrak{L}}{=} \sum_{j=0}^{t-1}\sum_{i=0}^{n_{2,j}} W_p^{-1}\left(a_{2,j,i}\right) f_2\left(\lambda, t, r_{2,j,i} + j\right),$$

we have that

$$I_{2,t}^{'}|I_{2,t} \overset{\mathfrak{L}}{=} I_{2,t} - \sum_{i=0}^{n_{2,j}} W_p^{-1}\left(a_{2,b,i}\right) f_2\left(\lambda, t, r_{2,b,i} + b\right) + \sum_{i=0}^{n_{2,j}} W_p^{-1}\left(a_{2,b,i}^{'}\right) f_2\left(\lambda, t, r_{2,b,i}^{'} + b\right),$$

which does not require $O\left(t^2\right)$ operations. To further improve the speed, the values $W_p^{-1}\left(a_{2,b,i}\right)$ and $f_2\left(\lambda, t, r_{2,b,i} + b\right)$ can be stored, to avoid repeat calculations.

6·2 *Analysis of Simulated Data*

We will focus on the $Ga - OU$ BDLP to facilitate algorithm run time, though any of the previous BDLPs could be used. As an approximate comparison, the MCMC algorithm for the *Ga-OU* BDLP for the non-OU models requires similar amounts of CPU time as the MCMC algorithm for the most general OU-driven SV model, the BNS-OU-SV model with *Generalized Inverse Gaussian* marginal. In the simulation study, we studied inference for data sets of size $N = 1000$. Our objective is to show that MCMC inference is feasible, and to illustrate its use for training and real data. We concentrate on the

Power Decay and fOUL processes. For the Power Decay process, we use a $Ga\,(1, 0.1)$ prior for $\lambda + 1$ and a $Ga\,(1, 0.5)$ prior for β. For the fOUL process, we use a $Ga\,(1, 0.5)$ prior for λ and a $Ga\,(1, 2/3)$ prior for $\kappa + \frac{1}{2}$.

Posterior distributions for λ and β for the Power Decay process on training data are shown in Figure 5, where $100,000$ iterations were taken (and thinned by recording every 10^{th} value) after a burn-in of $10,000$. The posterior supports the true values from which the data were generated for the Power Decay process. Correct inferences for the simulated fOUL process are obtained as for the Power Decay process and so details are omitted here for brevity.

From these results, it is evident that the MCMC algorithm is working correctly, and we now attempt to fit the models to the S&P 500 data set.

6·3 *Analysis of the S&P 500 Data*

We now study a real share index returns series, namely the S&P 500 index taken from 29th November 1999 to 1st December 2003, a total of $N = 1000$ observations, similar to the data set analyzed in Griffin & Steel (2003). Posterior histograms of the parameters determining the correlation of the square of the log returns are given in Figures 6 and 7. $10,000$ iterations were taken, thinning by recording every 10^{th} value, after a burn-in of $10,000$. The posterior for λ supports small λ values and fits a volatility process which decays asymptotically at a rate between t^{-1} and t^{-3}. The posterior is not concentrated at $\lambda = 1$, so the volatility process does not have long-memory. These graphs are ACF plots of the square of the log returns and the theoretical distribution of the square of the log returns of the fitted processes for one set of parameters, taken after MCMC had converged. Figure 8 demonstrates the Power Decay and fOUL models accurately fitting the correlation structure.

Thus the Power Decay and fOUL volatility models adequately capture the behaviour of the autocorrelation observed in the S&P 500 data. Analysis (not reported here) using a BNS-OU-SV model revealed that the exponential decay in correlation characteristic of the standard OU-SV model cannot capture the observed behaviour of the autocorrelation, confirming the analysis of Griffin & Steel (2003).

7. Discussion

Wolpert & Taqqu (2005) consider a class of stochastic processes driven by Lévy processes. We recall these models and suggest they could be used for stochastic volatility models because of their rich correlation structure. We describe how to simulate from such models and some of the properties of them. Although the models of this paper can have a more flexible correlation structure than the BNS-SV models, they are less tractable because the stochastic integrands have a more complex form. For example, for the BNS-SV model, the relationship between the BDLP and marginal distribution of the volatility is simple, whilst for the models of this paper it is often not available analytically. The models of this paper require a more complex series representation and this makes simulation slower. Therefore MCMC inference for these models, using the series representation of this paper and a similar algorithm to that of Gander & Stephens (2004) is more involved, although still feasible, for the $Ga - OU$ BDLP.

ACKNOWLEDGEMENT: The authors are very grateful for the helpful comments of Professor Robert Wolpert on the models used.

References

Anh, V. V. & McVinish, R. (2003). Fractional diffrerential equations driven by Lévy noise. *Journal of Applied Mathematics and Stochastic Analysis* **16**, 97–119.

Barndorff-Nielsen, O. E. & Shephard, N. (2000). Modelling by Lévy processes for financial econometrics. In *Lévy Processes - Theory and Applications*, O. E. Barndorff-Nielsen, T. Mikosch & S. Resnick, eds. Boston: Birkhäuser.

Barndorff-Nielsen, O. E. & Shephard, N. (2001a). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B* **63**, 167–241.

Barndorff-Nielsen, O. E. & Shephard, N. (2001b). Normal modified stable processes. *Theory of Probability and Mathematical Statistics* **65**, 1–19.

BERTOIN, J. (1994). *Lévy Processes*. Chapman and Hall.

BLACK, F. & SCHOLES, M. S. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* **81**, 637–654.

CARR, P., GEMAN, H., MADAN, D. P. & YOR, M. (2003). Stochastic volatility for Lévy processes. *Mathematical Finance* **13**, 345–382.

FELLER, W. (1971). *An Introduction to Probability Theory and its Applications*. Wiley.

GANDER, M. P. S. & STEPHENS, D. A. (2004). Stochastic volatility modelling with general marginal distributions: Inference, prediction and model selection for option pricing. Tech. rep., Imperial College London. Submitted.

GRIFFIN, J. E. & STEEL, M. F. J. (2003). Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility. Tech. rep., Department of Statistics, University of Warwick.

JEFFREY, G. H., HARE, D. J. & CORLESS, D. E. G. (1996). Unwinding the branches of the Lambert W function. *The Mathematical Scientist* **21**, 1–7.

ROBERTS, G., PAPASPILIOPOULOS, O. & DELLAPORTAS, P. (2004). Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **66**, 369–393.

SATO, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge university Press.

SCHOUTENS, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives*. Wiley.

WOLFE, S. J. (1982). On a continuous analogue of the stochastic difference equation $x_n = \rho x_{n-1} + b_n$. *Stochastic Processes and their applications* **12**, 301–312.

WOLPERT, R. L. & TAQQU, M. S. (2005). Fractional Ornstein-Uhlenbeck Lévy processes and the telecom process: Upstairs and downstairs. *Signal Processing* To Appear.

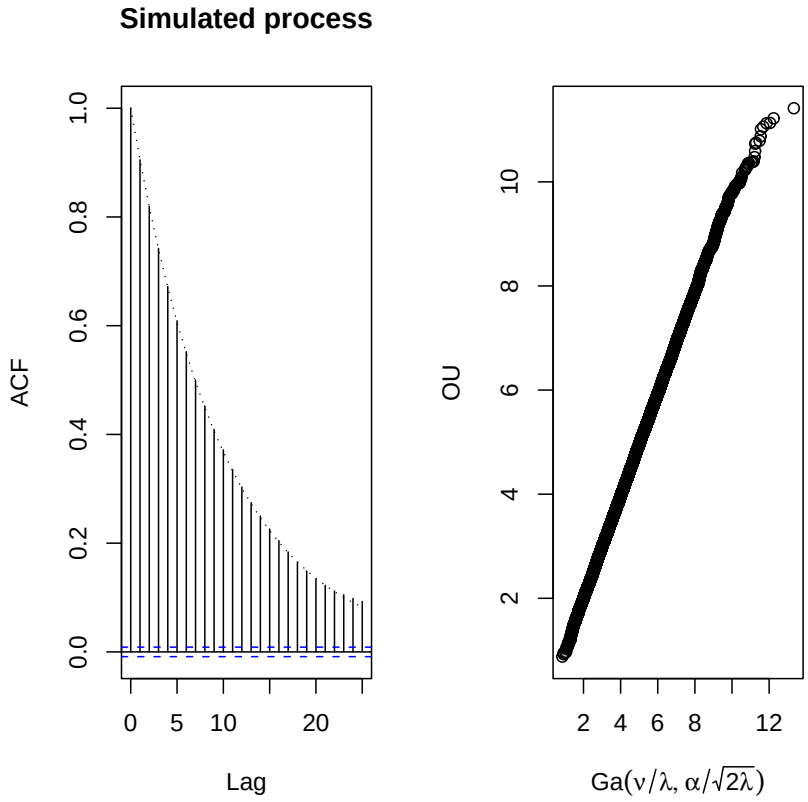

Figure 1: ACF of the OU process of Wolpert & Taqqu (2005) for $\lambda = 0.1$ for a $Ga(1,1)-$*OU* BDLP using the series representation of Section 3.

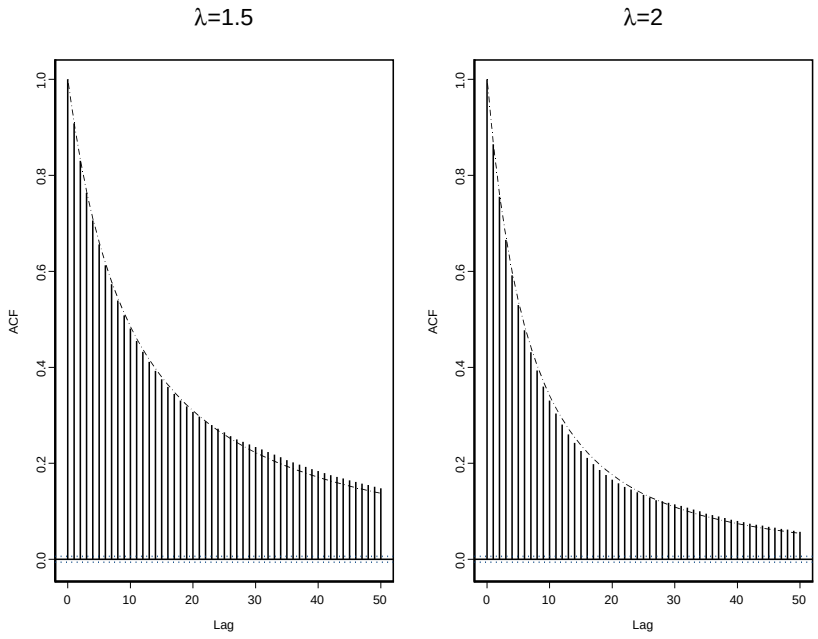

Figure 2: ACF of the Power Decay volatility process for $\beta = 0.1$, $\lambda = 1.5$ and $\lambda = 2$.

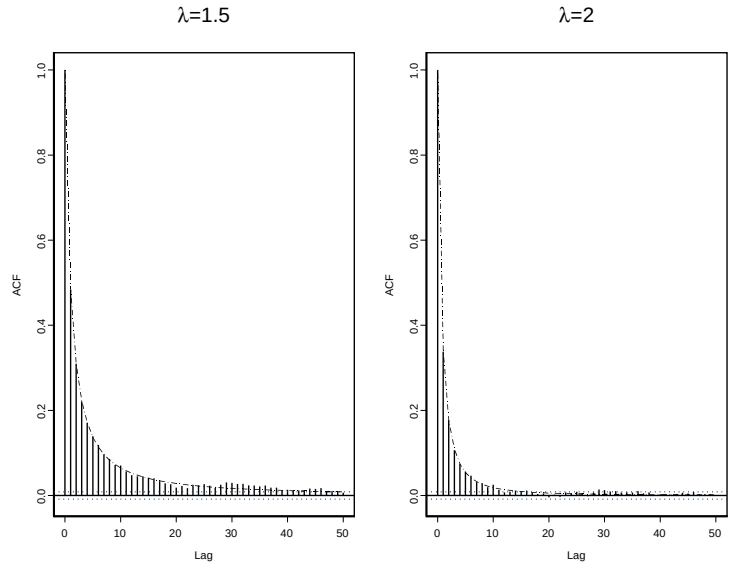

Figure 3: ACF of the Power Decay volatility process for $\beta = 1$, $\lambda = 1.5$ and $\lambda = 2$.

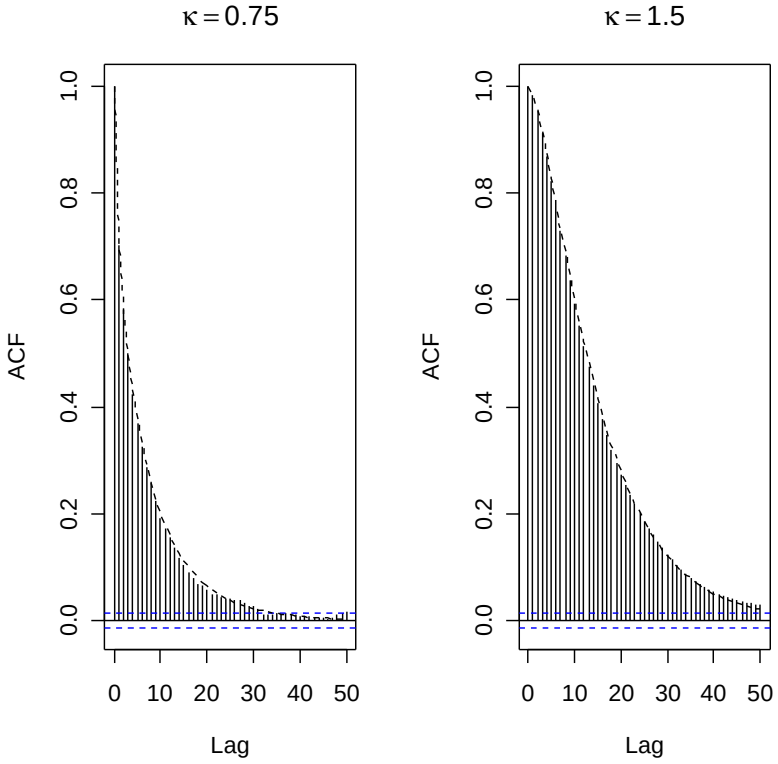

Figure 4: ACF of the fOUL process for $\lambda = 0.1$, $\kappa = 0.75$ and $\kappa = 1.5$.

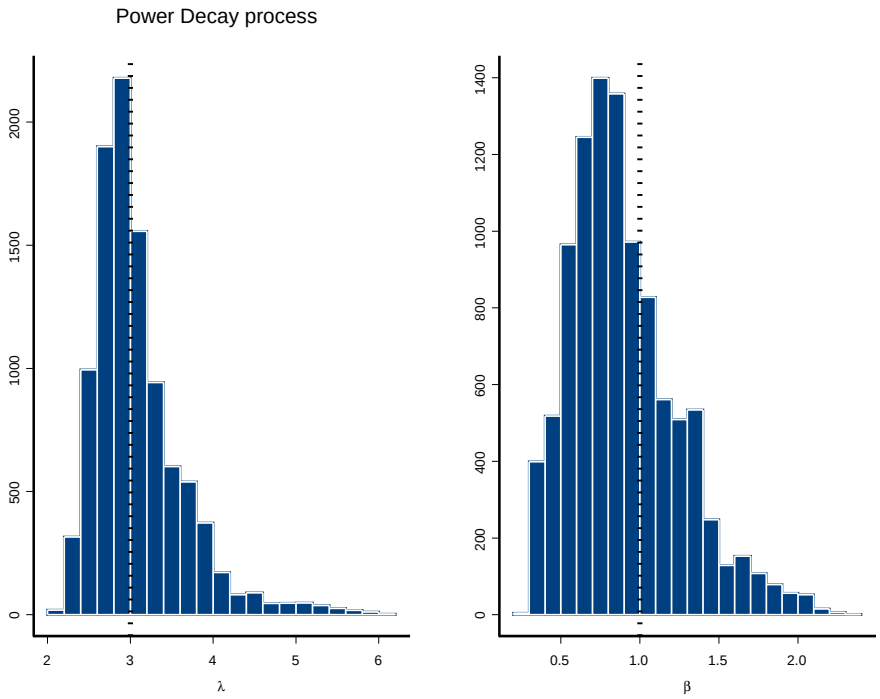

Figure 5: Histograms of the posterior distribution of λ and β for training data. True values indicated by vertical dotted lines.

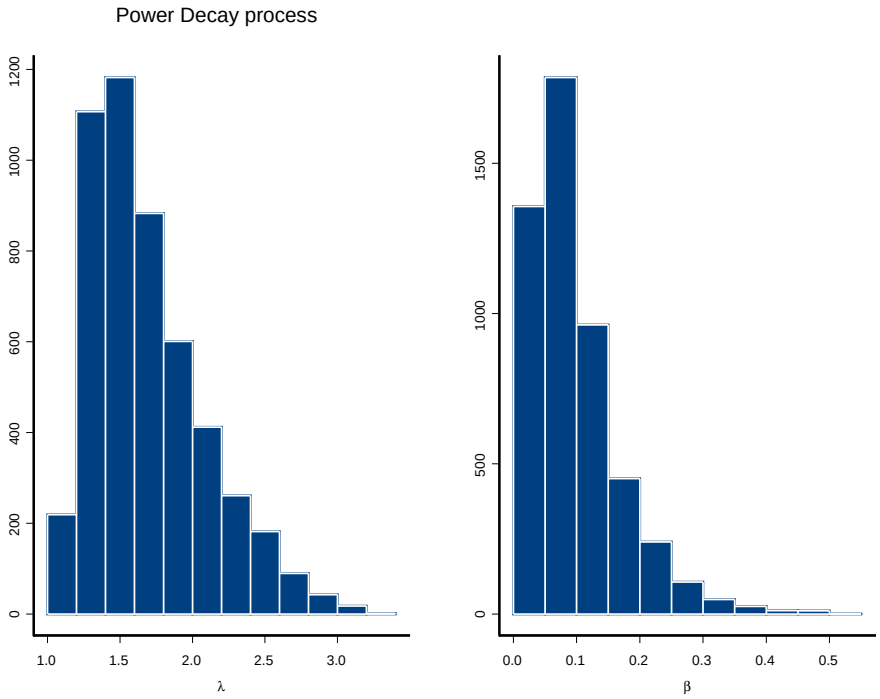

Figure 6: Power Decay process: Histograms of λ and β for S&P 500 data.

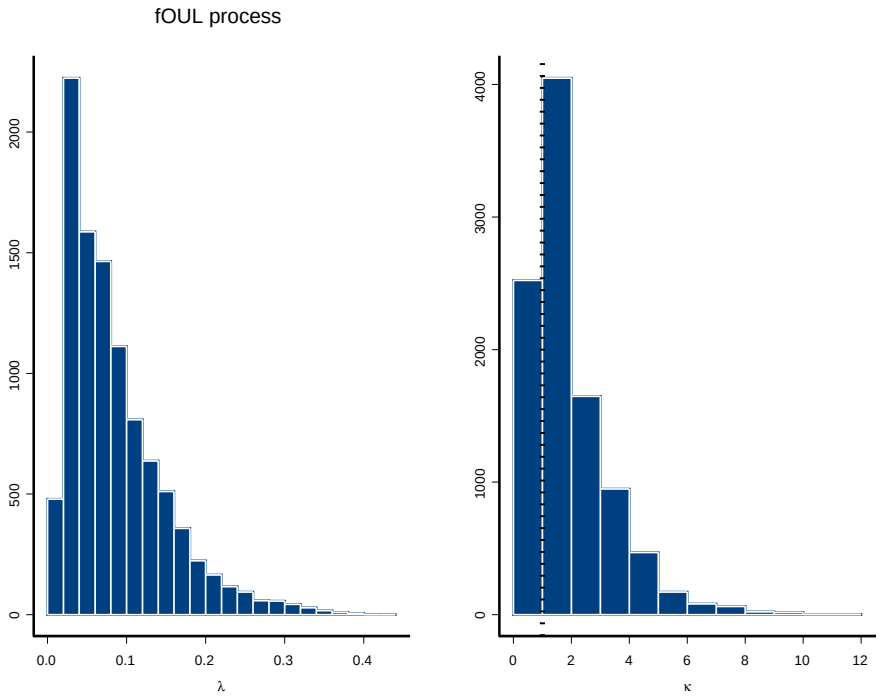

Figure 7: fOUL process: Histograms of λ and κ for S&P 500 data.

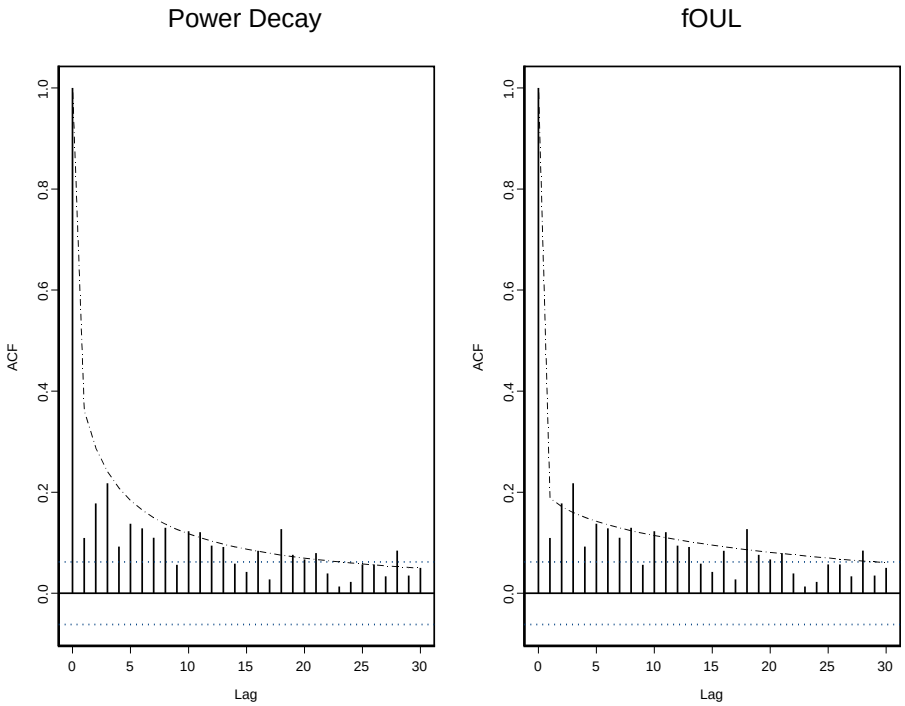

Figure 8: ACF of the square of the log returns of S&P 500 data and theoretical ACF of the fitted Power Decay and fOUL processes.