Chin. Ann. Math.
24B:2(2003),227-238.

PARAMETER ESTIMATION FOR A DISCRETELY OBSERVED STOCHASTIC VOLATILITY MODEL WITH JUMPS IN THE VOLATILITY***

JIANG Wenjiang* J. PEDERSEN**

Abstract

In this paper a stochastic volatility model is considered. That is, a log price process Y which is given in terms of a volatility process V is studied. The latter is defined such that the log price possesses some of the properties empirically observed by Barndorff-Nielsen & Jiang[6]. In the model there are two sets of unknown parameters, one set corresponding to the marginal distribution of V and one to autocorrelation of V. Based on discrete time observations of the log price the authors discuss how to estimate the parameters appearing in the marginal distribution and find the asymptotic properties.

Keywords Stochastic volatility models, NIG distributions, Central limit theorems, Law of large numbers, Levy processes, Ornstein-Uhlenbeck processes
2000 MR Subject Classification 60F15, 60G10, 60J75, 62P05
Chinese Library Classification O211.4, O211.6, O211.62, O213
Document Code A
Article ID 0252-9599(2003)02-0001-12

§1. Introduction

Many recent contributions in the field of finance are devoted to modelling of stock prices. The seminal Black and Scholes model[9] assumes that the stock price is a geometric Brownian motion, which in particular implies that the volatility is constant. However, recent empirical evidence has shown that the latter is unrealistic, see e.g. Bollerslev et al. (1992) or Taylor and Xu (1994, 1995). As an alternative, some authors have proposed models that include stochastic volatility in such a way that the pair consisting of the stock and the volatility constitutes a two-dimensional diffusion. Examples are the models of Hull and White[14], Wiggins[18], Scott[16], Chesney and Scott[10], Stein and Stein[17] and Heston[15]. One essential feature shared by all these models is that the volatility is unobservable when only discrete time observations are available, and clearly this raises some difficulties when trying to estimate the unknown parameters.

Genon-Catalot et al.[11] have proposed a general set-up for estimating the parameters of stochastic volatility models. They consider a two-dimensional diffusion $(Y(t), V(t))$, where

Manuscript received October 10, 2001.

*School of Mathematical Science, Yunnan Normal University, Kunming 650092, China.
E-mail: wenjiang@vip.sina.com

Institute Of Mathematical Sciences, University of Aarhus, Aarhus, Denmark. **E-mail: jan@imf.au.dk

***Project supported by the Yunnan Provincial Natural Science Foundation of China (No.00A0006R).

$Y(t)$ denotes the log price of a stock at time t, while $V(t)$ is the unobserved volatility. Moreover, $(Y(t), V(t))$ is given by

$$dY(t) = V(t)^{\frac{1}{2}} dB_t, \tag{1.1}$$

$$dV(t) = b(\theta, V(t))dt + a(\theta, V(t))dW_t, \tag{1.2}$$

where B and W are independent standard Brownian motions and θ is the unknown parameter. They consider the case when discrete time observations are available at time $0, \triangle_n, 2\triangle_n, \cdots, n\triangle_n$, and let $\triangle_n \to 0$ and $T_n = n\triangle_n \to \infty$ as $n \to \infty$. In this setting they are able to produce estimates of θ that are either consistent or even asymptotically normal. However, it is of fundamental importance that, in a loose sense, the unknown parameter only show up in the stationary distribution of $V(t)$ (see model 2 of their paper for more information on this) in particular, some questions related to the dependence structure of Y can not be addressed in their setting.

Recent empirical investigations by Barndorff-Nielsen and Jiang[6] show that by defining a stochastic volatility model where the volatility process has autocorrelation function given by

$$r(u) = w\rho_0^u + (1 - w)\rho_1^u, \tag{1.3}$$

a very good fit to real data is obtained. Moreover, the same paper also concludes that the NIG distribution often provides an excellent fit to the increments in log prices, measured on a daily basis, for instance. To incorporate these findings we will in the present study modify the approach taken by Genon-Catalot et al.[11]. More precisely, as above we let the log price be specified by Equation (1.1), but instead of (1.2) we take V to be the sum of two independent Inverse Gaussian Ornstein Uhlenbeck processes (short: IG OU processes), as defined by Barndorff-Nielsen[4]. This in particular implies that V is a stationary process with an inverse Gaussian law, $\mathrm{IG}(\alpha, \delta)$, as marginal distribution. Moreover, the autocorrelation of V then is described by (1.3). It also follows easily that, when measured over small intervals, the increments in the log price, suitably normed, are almost defined according to the Normal Inverse Gaussian distribution $\mathrm{NIG}(\alpha, 0, 0, \delta)$.

In other words, there are generally five unknown parameters in the model, two of which (α and δ) appear in the stationary distribution of V and three (ω, ρ_0 and ρ_1) in the autocorrelation. The problem we address is how to estimate α and δ based on discrete time observations of Y. Normally one would rely on the maximum likelihood estimate, but, as pointed out by Genon-Catalot et al.[11], the likelihood function has a somewhat inconvenient structure, and therefore one has to use another procedure. We follow the approach of Genon-Catalot et al.[11]. Essentially we obtain the same kind of results as they do, though the arguments have to be modified due to the fact that our V is a process with jumps whereas Genon-Catalot et al.[11] (1997) use a diffusion. Estimation of the parameters (ω, ρ_0, ρ_1) will be considered in a future paper.

The paper is organized in the following way. Below we first describe our model and the asymptotic framework in more detail. Section 2 is devoted to the parameters α and δ. As in [11] the building blocks are a Law of Large Numbers and a Central Limit Theorem, and these are also established in Section 2.

1.1. The Model

We now describe our model in further detail. First, as above let $B = (B(t))_{t\geq 0}$ denote a standard Wiener process and define the log price Y by (1.1). Second, based on the empirical findings previously mentioned we aim at defining V as a stationary process such that (i) the increments in Y, at least over small intervals, are nearly defined according to a normal inverse Gaussian distribution, and (ii) the autocorrelation of V is given by (1.3).

We show below that (i) and (ii) follow if V appears as the sum of two independent IG OU processes as we now describe.

Condition A *Let $\alpha, \delta_0, \delta_1, \lambda_0$ and λ_1 be positive parameters. Throughout the paper we let $V = (V(t))_{t\geq 0}$ be given by $V(t) = V_0(t) + V_1(0)$, where V_0 and V_1 are two independent* IG OU *processes, such that for $j = 0, 1$ the autocorrelation of V_j is $t \to \exp(-\lambda_j t)$ and the* *marginal distribution is* $\mathrm{IG}(\delta_j, \alpha)$.

Recall from [6] that this determines the distribution of the entire process V_j uniquely. By construction V_j can be written as

$$dV_j(t) = -\lambda_j V_j(t)dt + dZ_j(t), \tag{1.4}$$

where $V_j(0) \sim \mathrm{IG}(\delta_j, \alpha)$ is independent of the process Z_j, which is the so-called background driving Lévy process. The density of $V_j(0)$ (and $V_j(t)$ because V is stationary) hence is

$$(2\pi)^{-\frac{1}{2}}\delta e^{\delta\alpha}x^{-\frac{3}{2}}\exp\Big[-\frac{1}{2}(\delta^2x^{-1}+\alpha^2x)\Big], \quad x>0.$$

For future reference we note that Z_j is a strictly increasing Lévy process satisfying $Z_j(t) = \sum\limits_{0<s\le t} \triangle Z_j(s)$. The Lévy measure of Z_j is

$$F_j(dx) = \frac{1}{\lambda_j}(2\pi)^{-1/2}\delta_j x^{-3/2}e^{-\alpha^2x/2}dx, \quad x>0, \tag{1.5}$$

as can be seen from [6]. For more information on IG OU processes we also refer to [6]. The processes mentioned above are defined on a common probability space $(\Omega, \mathcal{F}, P)$ carrying a filtration $(\mathcal{F}_t)_{t\ge0}$ satisfying the usual conditions.

Recall that the convolution of $\mathrm{IG}(\delta_0, \alpha)$ and $\mathrm{IG}(\delta_1, \alpha)$ gives $\mathrm{IG}(\delta, \alpha)$, where $\delta = \delta_0 + \delta_1$ (see [6]). By defining V according to Condition A it therefore follows that V is a stationary process with $V(t)$ distributed according to the $\mathrm{IG}(\delta, \alpha)$. Moreover, letting $w = \delta_0/\delta$, $\rho_0 = e^{-\lambda_0}$, $\rho_1 = e^{-\lambda_1}$, it is seen by direct calculation that, indeed, the autocorrelation of V is as stated in (1.3).

Once again we emphasize that one of the main differences between our model and the model of Genon-Catalot et al.[11] is that we have jumps in our V process. We are further able to control the marginal distribution of V and the autocorrelation in a convenient way.

Having described our setting we note that there are five unknown parameters: α and δ corresponding to the stationary distribution of V and ω, ρ_0 and ρ_1 for the autocorrelation. When estimating these parameters we assume that discrete time observations of Y are available at time $t_0^n = 0, t_1^n, t_2^n, \cdots, t_n^n = T_n$, where we put $T_n = n\triangle_n$ and let $t_i^n = i\triangle_n$. We further let $\triangle_n \to 0$ and $T_n \to \infty$ as $n \to \infty$.

Often it is convenient to consider the variables defined by

$$Y_i^n := \frac{Y(t_i^n) - Y(t_{i-1}^n)}{\triangle_n^{1/2}}, \quad i = 1, \cdots, n$$

with a similar definition of B_i^n. The point is that B_i^n follows a standard normal distribution no matter what the values of i and n are, and $B_1^n \cdots, B_n^n$ are independent. Moreover, the sequence $Y_1^n, \cdots, Y_n^n$ is stationary and conditional on the processes V_0 and V_1 it follows that $Y_1^n, \cdots, Y_n^n$ are independent of

$$Y_i^n \sim N\Big(0, \frac{1}{\triangle_n}\int_{t_{i-1}^n}^{t_i^n} V(u)du\Big).$$

In particular, as $n \to \infty$ it is seen that Y_i^n converges to a variable Z, with $Z = \varepsilon\sqrt{\eta}$, where ε is standard normal and η is $\mathrm{IG}(\delta, \alpha)$, and ϵ and η are independent. That is, Z follows a $\mathrm{NIG}(\alpha, 0, 0, \delta)$. (Recall that $\delta = \delta_0 + \delta_1$ and the parameters are defined in Condition A). For future reference we state this result as a lemma.

Lemma 1.1 *Let* $\triangle_n \to 0$ *as* $n \to \infty$. *Then it holds that* $Y_1^n \xrightarrow{D} Z$, *where* $Z \sim \mathrm{NIG}(\alpha, 0, 0, \delta)$.

Despite its simplicity Lemma 1.1 is in fact one of the cornerstones when estimating α and δ, essentially because it allows us to pretend that the Y_i^n's are drawings from the $\mathrm{NIG}(\alpha, 0, 0, \delta)$.

§2. Estimation of the Marginal Distribution of V

This section concerns α and δ. Using Lemma 1.1 estimation is performed pretending that the Y_i^n's $(i=1,\cdots,n)$ are a sample of the NIG$(\alpha,0,0,\delta)$ and, as usual, the asymptotic behaviour is deduced via (a uniform version of) the Law of Large Numbers and a Central Limit Theorem. The appropriate versions are stated in Subsection 2.1, while the proofs are to be found in Subsection 2.2. The behaviour of the estimates is the topic of Subsection 2.3.

2.1. The Law of Large Numbers and the Central Limit Theorem

We first state the Law of Large numbers.

Proposition 2.1. *Let $f:\mathbf{R}\to\mathbf{R}$ be C^1 and satisfy $|f'(x)|\le K(1+|x|^r),\forall x$, for some constants $K,r>0$. Let also $\triangle_n\to 0$ and $T_n\to\infty$ as $n\to\infty$. Then it holds that*

$$\frac{1}{n}\sum_{i=1}^{n} f(Y_i^n)\to E[f(Z)]$$

in probability as $n\to\infty$, where $Z\sim$ NIG$(\alpha,0,0,\delta)$.

Remark 2.1. Note that by Lemma 1.1 we have that the Y_i^n's "almost" are distributed according to the NIG$(\alpha,0,0,\delta)$. In view of this the above result is not surprising.

In fact, we need to strengthen the statement in Proposition 2.1 slightly; more precisely, consider the following uniform version of the Law of Large Numbers.

Proposition 2.2. *Let $l\ge 1$ and $K\subseteq\mathbf{R}^l$ be a compact set. Also let $f:\mathbf{R}\times\mathbf{R}^l\to\mathbf{R}$ be C^1 and denote the partial derivatives by $f_i, i=1,\cdots,l+1$. Finally assume that there are constants $K,r>0$ such that for $i=1,\cdots,l+1$ and $(x,\tau)\in\mathbf{R}\times\mathbf{R}^l$ it holds that $|f_i(x;\tau)|\le K(1+|x|^r+||\tau||^r)$, where $||\cdot||$ denotes the usual norm on $\mathbf{R}^l$, and let $\triangle_n\to 0$ and $T_n\to\infty$ as $n\to\infty$. Then*

$$\sup_{\tau\in K}\Big|\frac{1}{n}\sum_{i=1}^{n} f(Y_i^n;\tau)-E[f(Z;\tau)]\Big|\to 0$$

in probability as $n\to\infty$, where Z is specified in Proposition 2.1.

We finally need the Central Limit Theorem, but first we introduce more notation. Let $l\ge 1$ and consider a mapping $f=(f_1,\cdots,f_l):\mathbf{R}\to\mathbf{R}^l$ satisfying the following two conditions:

(i) First, f is C^1 and there are constants $C,r>0$ such that for $j=1,\cdots,l$ the following inequality applies,

$$|f_j'(x)|\le C(1+|x|^r),\quad \forall x\in\mathbf{R}.$$

(ii) Also, $E[f_j(Z)]=0$ for $j=1,\cdots,l$, where Z is defined in Proposition 2.1.

Denote the class of mappings satisfying (i) and (ii) by $\mathcal{H}^l$.

For $f\in\mathcal{H}^l$ we define $h_f:\mathbf{R}_+^2\to\mathbf{R}^l$ by $h_f(w)=E[f(||w||X)]$, where X is standard normal and $||w||$ denotes the norm of $w=(w_0,w_1)$ given by $\sqrt{w_0+w_1}$. Then h_f satisfies a growth condition and is C^1. In the following we use that the process $(W(t))_{t\ge0}=(V_0(t),V_1(t))_{t\ge0}$ (the components are described in Condition A, see Section 1.1) is a Markov process, and we denote the corresponding transition function by $(T_t)_{t\ge0}$. Hence, for $f\in\mathcal{H}^l$ we have

$$T_t h_f(w)=E[h_f(W(t))|W(0)=w].$$

For $f\in\mathcal{H}^l$ we also define g_f by

$$g_f(w)=-\int_0^\infty T_t h_f(w)dt.$$

It follows from Remark 2.3.1 of [8], or by direct calculation, using that V_j is an OU process, that g_f is well-defined, that $g_f(W(t))$ is square integrable and that

$$\lim_{s\downarrow0}\frac{T_s g_f-g_f}{s}=\lim_{s\downarrow0}\frac{\int_0^s T_t h_f dt}{s}=h_f \tag{2.1}$$

pointwise. Whence g_f is in the domain of the generator of $(T_t)_{t\geq 0}$. Moreover, by defining

$$\Sigma_f = -E[g_f^T(W(0))h_f(W(0)))] - E[h_f^T(W(0))g_f(W(0)))] \tag{2.2}$$

where superscript T denotes transposition, we get a positive semidefinite $l \times l$ matrix, as we shall see in the proof of Proposition 2.3.

Proposition 2.3. (a) *Let $f \in \mathcal{H}^l$ and let $\triangle_n \to 0, T_n \to \infty$ and $\triangle_n T_n \to 0$ as $n \to \infty$. Then we have*

$$\frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} f(Y_i^n)\triangle_n \xrightarrow{D} N(0, \Sigma_f).$$

(b) *If moreover the l functions $w \to h_{f,1}(w), \cdots, w \to h_{f,l}(w)$ are linearly independent, then it holds that Σ_f is positive definite. This condition is satisfied if each f_j is symmetric and further the l functions $x \to f_1(x), \cdots, x \to f_l(x)$ are linearly independent.*

Results similar to Propositions 2.1 and 2.3 are also given in [12]. However, as they do not use quite the same setting we can not use their results directly. Specifically, they obtain their analogue to Proposition 2.3 (a) by using a CLT for diffusions, while we have to rely on a CLT stated for more general Markov processes. Besides, they give the asymptotic variance in somewhat more closed form. In the present setting (2.1) is essentially the best representation of the variance we can get, and the property of Σ_f stated in Proposition 2.3 (b) has to be proved by means of martingale techniques. An analogue to Proposition 2.3 (b) is also obtained by [8] in a somewhat different context, but only when h_f is bounded.

2.2. The Proofs

We first state a couple of preliminary results.

Lemma 2.1. *For $j = 0, 1$ and $p, T > 0$ we have that $\sup_{t\leq T} V_j(t)$ is in $L^p(P)$. Moreover, it holds that*

$$\limsup_{t\to 0} \frac{1}{t} E\Big[\sup_{u\in[0,t]} |V(u) - V(0)|\Big]$$

is finite.

Proof. As V_j is an OU process and Z_j is increasing (see Subsection 1.1) we have the following estimate for $t \leq T$

$$V_j(t) = V_j(0)e^{-\lambda_j t} + \int_0^t e^{-\lambda_j(t-s)} dZ_j(s) \leq V_j(0) + Z_j(T).$$

Clearly, both $V_j(0)$ and $Z_j(T)$ have finite moments. This gives the first statement.

To prove the second statement we note that it clearly suffices to argue

$$\limsup_{t\to 0} \frac{1}{t} E\Big[\sup_{u\in[0,t]} |V_j(u) - V_j(0)|\Big] < \infty$$

for $j = 0, 1$. But from the above inequalities we have for $u \leq t$

$$|V_j(u) - V_j(0)| \leq V_j(0)(1 - e^{-\lambda_j t}) + Z_j(t)$$

and using that, $E[Z_j(t)] = C_j t, \forall t \geq 0$, for some constant C_j, it follows

$$\frac{1}{t} E\Big[\sup_{u\in[0,t]} |V_j(u) - V_j(0)|\Big] \leq E[V_j(0)]\frac{(1 - e^{-\lambda_j t})}{t} + C_j.$$

The right-hand side converges to $E[V_j(0)]\lambda_j + C_j$ as $t \to 0$, which gives the result.

Lemma 2.2. *let f and $\triangle_n$ be as in Proposition 2.1. Then it holds that $f(Y_1^n)$ has the expansion*

$$f(Y_1^n) = \{f(Y_1^n) - f(V(0)^{\frac{1}{2}}B_1^n)\} + f(V(0)^{\frac{1}{2}}B_1^n),$$

where the first term, divided by $\triangle_n^{\frac{1}{2}}$, is bounded in $L^1(P)$.

Proof. We first have $f(Y_1^n) - f(V(0)^{\frac{1}{2}} B_1^n) = f'(Z_n)\{Y_1^n - V(0)^{\frac{1}{2}} B_1^n\}$, where Z_n is on the line between Y_1^n and $V(0)^{\frac{1}{2}} B_1^n$. And note that the number $E[f^{'}(Z_n)^2]^{\frac{1}{2}}$ is bounded by a constant C, not depending on n, therefore

$$\begin{aligned}E[|f(Y_1^n) - f(V(0)^{\frac{1}{2}} B_1^n)|] &\le C\{E[(Y_1^n - V(0)^{\frac{1}{2}} B_1^n)^2]\}^{\frac{1}{2}} \\ &= C\Big\{E\Big[\triangle_n^{-1}\int_0^{\triangle_n}(V^{\frac{1}{2}}(0) - V^{\frac{1}{2}}(u))^2 du\Big]\Big\}^{\frac{1}{2}} \\ &\le C\Big\{E\Big[\sup_{0\le u\le\triangle_n}|V(0) - V(u)|\Big]\Big\}^{\frac{1}{2}}.\end{aligned}$$

By applying to Lemma 2.1, this gives the result.

Proof of Proposition 2.1. Consider the following expansion

$$\frac{1}{n}\sum_{i=1}^n f(Y_i^n) = \frac{1}{n}\sum_{i=1}^n f(Y_i^n) - f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n) + \frac{1}{n}\sum_{i=1}^n f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n) \tag{2.3}$$

and use that by stationarity and the previous lemma

$$\begin{aligned}&E\Big[\Big|\frac{1}{n}\sum_{i=1}^n f(Y_i^n) - f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n)\Big|\Big] \le n^{-1}\sum_{i=1}^n E\Big[\Big|f(Y_i^n) - f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n)\Big|\Big] \\ &= E\Big[\Big|f(Y_1^n) - f(V^{\frac{1}{2}}(0)B_1^n)\Big|\Big] \to 0.\end{aligned}$$

This means we just have to consider the second term on the right-hand side of (2.3). Define $q_f(v)$ by $q_f(v) = E[f(v^{\frac{1}{2}}X)]$, where X follows a standard normal, and note that

$$\frac{1}{n}\sum_{i=1}^n f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n) = \frac{1}{n}\sum_{i=1}^n q_f(V(t_{i-1}^n)) + \frac{1}{n}\sum_{i=1}^n [f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n) - q_f(V(t_{i-1}^n))]. \tag{2.4}$$

It is obvious that the second term on the right-hands side of (2.4) vanishes in $L^2(P)$ as $n \to \infty$. Also, it is easily seen that q_f is continuous and satisfies a growth condition, therefore by the previous lemma we have

$$\begin{aligned}&E\Big[\Big|\frac{1}{T_n}\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n} q_f(V(t_{i-1}^n)) - q_f(V(u))du\Big|\Big] \\ &\le E\Big[\sup_{u\in[0,t_1^n]}\Big|q_f(V(0)) - q_f(V(u))\Big|\Big] \to 0.\end{aligned}$$

Finally, as V is ergodic, the Law of Large Numbers gives

$$\frac{1}{T_n}\int_0^{T_n} q_f(V(u))du \to E[q_f(V(0))] = E[f(Z)]$$

as $n \to \infty$ a.s., and this concludes the proof.

Proof of Proposition 2.2. Note there are constants $C, p > 0$ such that $|f(x;\tau)| \le C(1 + |x|^p), \forall x \in \mathbf{R}, \forall \tau \in K$. Put $\gamma = \sup_n E[\{C(1 + |Y_1^n|^p)\}^2]^{\frac{1}{2}}$, which is a finite quantity. We must argue that for any $\epsilon > 0$ it holds

$$\limsup_{n\to\infty} E\Big[\sup_{\tau\in K}\Big|\frac{1}{n}\sum_{i=1}^n f(Y_i^n;\tau) - E[f(Z;\tau)]\Big| \wedge 1\Big] < \epsilon.$$

Choose $L > 0$ such that for $n \in N$ it holds that l $P(|Y_1^n| > L)^{\frac{1}{2}} \le \frac{\frac{\epsilon}{4}}{2\gamma}$. This can be done due to the fact that Y_1^n converges by Lemma 1.1. Also choose an integer q, points $\tau_1, \cdots, \tau_q \in K$ and positive numbers $r_1, \cdots, r_q$ such that (i) $K \subseteq \bigcup_{j=1}^q B(\tau_j; r_j)$, where $B(\tau_j; r_j)$ denotes the open ball in $\mathbf{R}^l$ with center τ_j and radius r_j, (ii) $E[\sup_{\tau\in B(\tau_j;r_j)} |f(Z;\tau_j)-f(Z;\tau)|] \le \frac{\epsilon}{4}$ and (iii) $\sup_{x:|x|\le L, \tau\in B(\tau_j;r_j)} |f(x;\tau_j)- f(x;\tau)| \le \frac{\epsilon}{4}$. Then

$$\begin{aligned}
&\sup_{\tau\in K}\Big|\frac{1}{n}\sum_{i=1}^{n} f(Y_i^n;\tau) - E[f(Z;\tau)]\Big| \\
\le &\sup_{j=1,\cdots,q}\sup_{\tau\in B(\tau_j;r_j)}\Big|\frac{1}{n}\sum_{i=1}^{n} f(Y_i^n;\tau) - E[f(Z;\tau)]\Big| \\
\le &\sup_{j=1,\cdots,q}\sup_{\tau\in B(\tau_j;r_j)}\Big|\frac{1}{n}\sum_{i=1}^{n} [f(Y_i^n;\tau_j) - f(Y_i^n;\tau)]\Big| \\
&+\sup_{j=1,\cdots,q}\Big|\frac{1}{n}\sum_{i=1}^{n} f(Y_i^n;\tau_j) - E[f(Z;\tau_j)]\Big| \\
&+\sup_{j=1,\cdots,q}\sup_{\tau\in B(\tau_j;r_j)}\Big|E[f(Z;\tau) - f(Z;\tau_j)]\Big|.
\end{aligned}$$

The third term is by (ii) less than $\frac{\epsilon}{4}$, the mean of the last term is dominated by $2\gamma P(|Y_i^n| > L)^{\frac{1}{2}} \le \frac{\epsilon}{4}$. Using (iii), we can easily see that the first term is dominated by

$$\frac{\epsilon}{4} + \frac{2C}{n}\sum_{i=1}^{n}(1+|Y_i^n|^p)1_{\{|Y_i^n|>L\}}.$$

Therefore

$$\begin{aligned}
&E\Big[\sup_{\tau\in K}\Big|\frac{1}{n}\sum_{i=1}^{n} f(Y_i^n;\tau) - E[f(Z;\tau)]\Big|\wedge 1\Big] \\
\le &\sup_{j=1,\cdots,q} E\Big[\Big|\frac{1}{n}\sum_{i=1}^{n} f(Y_i^n;\tau_j) - E[f(Z;\tau_j)]\Big|\wedge 1\Big] + \frac{3}{4}\epsilon,
\end{aligned}$$

and by choosing n large enough the left-hand side is less than ϵ, as is seen from Proposition 2.1.

Before proving the Central Limit Theorem Proposition 2.3 we state another useful result.

Proposition 2.4. *Let $f \in \mathcal{H}^l$ and $T_n \to \infty$ as $n\to\infty$. It then holds*

$$\frac{1}{\sqrt{T_n}}\int_0^{T_n} h_f(W(u))du \xrightarrow{D} N(0,\Sigma_f) \quad as \quad n\to\infty.$$

Moreover, Σ_f is positive definite if the functions $w\to h_{f,1}(w),\cdots,w\to h_{f,l}(w)$ are linearly independent.

Proof. Let $l = 1$. The general case follows from this using the usual Cramer-Wold device and the fact that if $f\in\mathcal{H}^l$ and $a\in\mathbf{R}^l$ is constant then $a\cdot f\in\mathcal{H}^1$ and $a\cdot h_f = h_{a\cdot f}$, $a\cdot g_f = g_{a\cdot f}$. The first part of the statement is a consequence of Bhattacharya[8, Theorem 2.1]. However, in order to prove the second part we have to repeat some of the arguments. First, it is known that the relation

$$T_t g_f = g_f + \int_0^t T_s h_f ds$$

implies that the process

$$M(t) = g_f(W(t)) - g_f(W(0)) - \int_0^t h_f(W(s))ds$$

is a martingale. Also, by direct calculation it is seen that $\mathrm{var}(M(t)) = t\Sigma_f$. Clearly, stationarity of W implies that the first part of the result follows by proving $\frac{1}{\sqrt{t}}M(t) \xrightarrow{D} N(0,\Sigma_f)$ as $t \to \infty$. As M has square integrable increments and is defined in terms of a stationary Markov process, this is however an easy consequence of the Central Limit Theorem for martingales (see e.g. [13]).

To prove the second part let us assume $\Sigma_f = 0$. Then $M(t)$ is almost surely constant. By right-continuity it therefore follows that M is indistinguishable from zero, and hence $(g_f(W(t)))_{t\geq 0}$ is a continuous process. However, W has jumps of any positive size and therefore g_f is constant (see Lemma 2.7 below for details). Using the relation (2.1) it is seen that h_f is zero.

The following lemma clarifies one point used above.

Lemma 2.3. *Let the setting be as in Proposition* 2.3 (a) *and let* $l = 1$.

(a) *Let* $w = (w_0, w_1) \in \mathbf{R}_+^2$ *and* ξ, ρ *and* r *be positive numbers such that* ξ *and* ρ *are greater that* $3r$. *Then it holds that*

$$P(\exists t : V_0(t-) \in]w_0 - r, w_0 + r[\quad and \quad V_0(t) \in]w_0 + \xi - r, w_0 + \xi + r[$$
$$and \quad V_1(t-), V_1(t) \in]w_1 - r, w_1 + r[\Big) > 0$$
$$P(\exists t : V_1(t-) \in]w_1 - r, w_1 + r[\quad and \quad V_1(t) \in]w_1 + \rho - r, w_1 + \rho + r[$$
$$and \quad V_0(t-), V_0(t) \in]w_0 - r, w_0 + r[) > 0.$$

(b) *If* $\Sigma_f = 0$ *then* g_f *is constant.*

Remark 2.2. The first statement in (a) states that with positive probability V_0 jumps from a neighbourhood of w_0 to a neighbourhood of $w_0 + \xi$, while V_1 stays near w_1.

Proof. (a) Note that, because the marginal distributions of the V_j processes are independent IG laws, the variable $W(0)$ will be in the set $]w_0 - r, w_0 + r[\times]w_1 - r, w_1 + r[$ with positive probability. Therefore it suffices to argue that the two probabilities are both positive conditional on $W(0) = \bar{w}$, where $\bar{w}$ is an arbitrary point of $]w_0 - r, w_0 + r[\times]w_1 - r, w_1 + r[$. We shall for simplicity only consider the case $\bar{w} = w$. Let P_w denote the measure corresponding to the conditional distribution given $W(0) = w$. Then using that V_j appears as an OU Process driven by the process Z_j, and that Z_j is increasing, we obtain under P_w

$$e^{-\lambda_j t} w_j \leq V_j(t) \leq e^{-\lambda_j t} w_j + Z_j(t), \quad j = 0, 1. \tag{2.5}$$

We give the argument for the first probability only, the second being similar. Take $t_0 > 0$ satisfying

$$e^{-\lambda_j t_0} w_j > w_j - r, \quad j = 0, 1 \tag{2.6}$$

and let us prove

$$P_w(V_1(t) \in]w_1 - r, w_1 + r[, \forall t \leq t_0 \quad and$$
$$\exists t \leq t_0 : V_0(t-) \in]w_0 - r, w_0 + r[, \quad V_0(t) \in]w_0 - r + \xi, w_0 + r + \xi[\Big) > 0.$$

Now, by independence the above probability equals

$$P_{w_1}\Big(V_1(t) \in]w_1 - r, w_1 + r[, \forall t \leq t_0\Big)$$
$$\times P_{w_0}\Big(\exists t \leq t_0 : V_0(t-) \in]w_0 - r, w_0 + r[, \quad V_0(t) \in]w_0 - r + \xi, w_0 + r + \xi[\Big)$$

and by the representation of the Lévy measure of Z_j in (1.5), it follows easily that with positive probability we have $Z_1(t_0) < r$, and hence by (2.5) and (2.6) the first term is positive. Let $\tau = \inf\{t > 0 : \triangle Z_0(t) \geq r\}$. Using (2.5) and the fact that $\triangle Z_0 = \triangle V_0$ it suffices to argue

$$P_{w_0}\Big(\tau \leq t_0 \quad and \quad \triangle Z_0(\tau) \in]w_0 - r + \xi - V_0(\tau-), w_0 + r + \xi - V_0(\tau-)[and \quad Z_0(\tau-) < r\Big) > 0.$$

Let $\mathcal{F}^0_{\tau-}$ be the pre σ–field generated by τ in the natural filtration of Z_0. As the set $A = \{\tau \le t_0, Z_0(\tau-) < r\}$ has positive probability under P_{ω_0} and is an element of $\mathcal{F}^0_{\tau-}$, it suffices to argue that conditional on $\mathcal{F}^0_{\tau-}$ the event

$$\{\triangle Z_0(\tau) \in]w_0 - r + \xi - V_0(\tau-), w_0 + r + \xi - V_0(\tau-)[\}$$

has positive probability on A. Now, $\triangle Z_0(\tau)$ is independent of $\mathcal{F}^0_{\tau-}$ and is distributed according to $\frac{F^0(d\kappa)}{F^0([r,\infty[)}$ for $\kappa > r$, where F^0 denotes the Lévy measure for Z_0. On A the interval $]w_0 - r + \xi - V_0(\tau-), w_0 + r + \xi - V_0(\tau-)[$ is contained in $]r, \infty[$ (because $\xi > 3r$ by assumption). Further, F^0 has positive density with respect to the Lebesgue measure on $\mathbf{R}_+$ (see (1.5)). By combining these facts the result follows.

Concerning (b) this is an easy consequence of (a), continuity of g_f and the fact that $(g_f(W(t)))_{t\ge 0}$ is a continuous process.

Proof of Proposition 2.3. (a) We only consider the case $l = 1$. Use the expansion

$$\frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} f(Y_i^n)\triangle_n = \frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} \big\{ f(Y_i^n) - f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n) \big\}\triangle_n + \frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n)\triangle_n \tag{2.7}$$

and noting that the mean of the first term is dominated by

$$T_n^{\frac{1}{2}} \triangle_n^{\frac{1}{2}} \frac{E[|f(Y_1^n) - f(V^{\frac{1}{2}}(0)B_1^n)|]}{\triangle_n^{\frac{1}{2}}}$$

and that the second is bounded by Lemma 2.2., we have that the first term on the right-hand side of (2.7) converges to zero in $L^1(P)$. Moreover,

$$\begin{aligned} &\frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n)\triangle_n \\ &= \frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} \big\{ f(V^{\frac{1}{2}}(t_{i-1}^n)B_i^n) - h_f(W(t_{i-1}^n)) \big\}\triangle_n + \frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} h_f(W(t_{i-1}^n))\triangle_n \end{aligned}$$

and as in the proof of Proposition 2.1 it is seen that the first term vanishes in $L^2(P)$ as $n \to \infty$. By the preceding proposition we conclude the proof once we have shown that

$$\begin{aligned} &\frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} h_f(W(t_{i-1}^n))\triangle_n - \frac{1}{\sqrt{T_n}} \int_0^{T_n} h_f(W(u))du \\ &= \frac{1}{\sqrt{T_n}} \sum_{i=1}^{n} \int_{t_{i-1}^n}^{t_i^n} h_f(W(u)) - h_f(W(t_{i-1}^n))du \end{aligned}$$

vanishes in $L^1(P)$ as $n \to \infty$. But the mean of this is less than

$$T_n^{\frac{1}{2}} E\Big[\sup_{0 \le u \le \triangle_n} \big| h_f(W(u)) - h_f(W(0)) \big| \Big].$$

It follows that the last term is dominated by a constant times

$$T_n^{\frac{1}{2}} \Big\{ E\Big[\sup_{0 \le u \le \triangle_n} |V(u) - V(0)| \Big] \Big\}^{\frac{1}{2}},$$

and by Lemma 2.1 convergence to zero holds.

To prove (b) the first part follows from Lemma 2.3 while for the second we first state a fact as follows:

If $f \in \mathcal{H}^1$ is symmetric and $Ef(\tilde{X}) = 0$ for any $\tilde{X}$ which follows a normal distribution with zero mean, then f must be zero.

To prove this, denote $g(\sigma) = \int_{-\infty}^{+\infty} f(x) e^{\frac{-x^2}{2}\sigma^2} dx$. Then it follows that $g(\sigma) = 0, \quad \forall \sigma > 0$, and also $g'(\sigma) = 0, \quad \forall \sigma > 0$. Repeating this leads to

$$\int_{-\infty}^{+\infty} f(x) x^{2n} e^{\frac{-x^2}{2}} dx = 0, \quad \forall n \in N,$$

and by symmetry of f, we have

$$\int_{-\infty}^{+\infty} f(x) x^{2n+1} e^{\frac{-x^2}{2}} dx = 0, \quad \forall n \in N.$$

Thus

$$\int_{-\infty}^{+\infty} f^2(x) e^{\frac{-x^2}{2}} dx = 0.$$

Finally we get that f must be zero due to continuity.

Now, we use this to deduce the second part of (b). If Σ_f is not positive definite then by the first part of (b) there is a nonzero vector $a \in \mathbf{R}^l$, such that $a \cdot h_f = 0$. This then leads to $E[a \cdot f(||w||X) = 0]$ for any $w \in \mathbf{R}_+^2$, where $X \sim N(0,1)$, and by the above we get $a \cdot f = 0$.

2.3. Parameter Estimation

Recall from Lemma 1.1 that $Y_1^n \xrightarrow{D} \mathrm{NIG}(\alpha, 0, 0, \delta)$ as $n \to \infty$. That is, for $\triangle_n$ small we have that $Y_1^n, \cdots, Y_n^n$ is a stationary sequence such that Y_i^n almost follows a NIG distribution. Motivated by this we now find an estimate $(\hat{\alpha}_n, \hat{\delta}_n)$ of the true parameter values (α, δ) as a local maximum for

$$(\alpha^0, \delta^0) \to \sum_{i=1}^{n} l(Y_i^n; (\alpha^0, \delta^0)), \tag{2.8}$$

where $l(y; (\alpha^0, \delta^0))$ denotes the log density of the $\mathrm{NIG}(\alpha^0, 0, 0, \delta^0)$ in the point y. In other words we estimate the parameters by maximizing the likelihood function corresponding to iid drawings of a NIG distribution. As we shall see the estimates are in fact asymptotically normal, though the observations are not independent. The asymptotic variance however differs somewhat from the iid case.

From [4] we have that the log density of a $\mathrm{NIG}(\alpha^0, 0, 0, \delta^0)$ is

$$\begin{aligned} l(y; (\alpha^0, \delta^0)) &= \delta^0 \alpha^0 - \log \pi - \log \alpha^0 + \log \delta^0 - \frac{1}{2} \log({\delta^0}^2 + y^2) \\ &\quad + \log K_1(\alpha^0 \cdot \sqrt{{\delta^0}^2 + y^2}), \end{aligned}$$

where $q(y) = \sqrt{1 + y^2}$ and K_1 is the modified Bessel function of the third order and index 1. Denote also the j'th derivative of $\log K_1$ by D_j. To state the asymptotic properties of the estimates we need the first and the second derivative of l with respect to the parameters, so let $U = (U_1, U_2)$ be the two-dimensional vector of the partial derivatives (subscript 1 corresponds to α). Then we have (with $D_1 = D_1(\alpha^0 \cdot \sqrt{{\delta^0}^2 + y^2})$ and $D_2 = D_2(\alpha^0 \cdot \sqrt{{\delta^0}^2 + y^2}))$

$$\begin{aligned} U_1(y; (\alpha^0, \delta^0)) &= -\frac{1}{\alpha^0} + \delta^0 + D_1 \cdot \sqrt{{\delta^0}^2 + y^2}, \\ U_2(y; (\alpha^0, \delta^0)) &= \frac{1}{\delta^0} + \alpha^0 - \frac{\delta^0}{{\delta^0}^2 + y^2} + D_1 \cdot \frac{\alpha^0 \delta^0}{\sqrt{{\delta^0}^2 + y^2}} \end{aligned}$$

and with $j = (j_{ij})_{i,j=1}^2$ denoting the 2 by 2 matrix consisting of minus the second derivatives we get

$$j_{11}(y;(\alpha^0,\delta^0)) = \frac{1}{{\alpha^0}^2} + D_2\cdot({\delta^0}^2 + y^2),$$

$$j_{12}(y;(\alpha^0,\delta^0)) = 1 + D_2\cdot\delta^0\alpha^0 + D_1\cdot\frac{\delta^0}{\sqrt{{\delta^0}^2+y^2}},$$

$$j_{21}(y;(\alpha^0,\delta^0)) = j_{12}(y;(\alpha^0,\delta^0)),$$

$$j_{22}(y;(\alpha^0,\delta^0)) = -\frac{1}{{\delta^0}^2} - \frac{y^2-{\delta^0}^{2}{}^{2}}{{\delta^0}^2+y^2} + D_2\cdot\frac{{\alpha^0}^2{\delta^0}^2}{\sqrt{{\delta^0}^2+y^2}} + D_1\cdot\frac{\alpha^0 y^2}{({\delta^0}^2+y^2)^{\frac{3}{2}}}.$$

By combining this with subsection 2.1 we arrive at the following.

Proposition 2.5. *Let $\triangle_n \to 0$, $T_n \to \infty$ and $\triangle_n T_n \to 0$ as $n \to \infty$.* (a) *First, we have*

$$\frac{1}{\sqrt{T_n}}\sum_{i=1}^n U\big(Y_i^n;(\alpha,\delta)\big)\triangle_n \xrightarrow{D} N(0,\Sigma_{(\alpha,\delta)}), \quad n\to\infty,$$

where $\Sigma_{(\alpha,\delta)}$ denotes the matrix Σ_f in (2.2) *with $f(\cdot) = U(\cdot;(\alpha,\delta))$. Moreover, $\Sigma_{(\alpha,\delta)}$ is positive definite.*

(b) *If $r>0$ is positive such that the open ball, $B((\alpha,\delta);r)$, with radius r and center (α,δ) is contained in $\mathbf{R}_+^2$, then it holds that*

$$\sup_{(\alpha^0,\delta^0)\in B((\alpha,\delta);r)}\Big|\frac{1}{n}\sum_{i=1}^n j(Y_i^n;(\alpha^0,\delta^0)) - E[j(Z;(\alpha^0,\delta^0))]\Big| \to 0$$

in probability as $n\to\infty$, where $Z \sim$ NIG$(\alpha,0,0,\delta)$. Moreover, the matrix $E[j(Z;(\alpha,\delta))]$ is positive definite.

Remark 2.3. Let $i(\alpha,\delta) = E[j(Z;(\alpha,\delta))]$. Then $i(\alpha,\delta)$ corresponds in fact to the expected information for iid drawings of the NIG$(\alpha,0,0,\delta)$. This immediately gives that $i(\alpha,\delta)$ is positive definite.

Proof. We just have to verify the conditions in Propositions 2.2 and 2.3. But for every $\epsilon>0$ it holds that $\sup_{y>\epsilon}|D_j(y)|$ is bounded for $j=1,2,3$ as can be seen from [1]. This gives the various growth conditions, which in turn implies $U(\cdot;(\alpha,\delta))\in\mathcal{H}^2$. As $y\to U_1(y;(\alpha,\delta))$ and $y\to U_2(y;(\alpha,\delta))$ are symmetric and linearly independent it also follows that the asymptotic variance in (a) is positive definite.

Having established the CLT and the Law of Large Numbers, the asymptotic behaviour of the estimate $(\hat\alpha_n,\hat\delta_n)$ is easily derived. More precisely, using the classical Taylor expansion argument (see e.g. [7]) it follows that, indeed, there is local maximum $(\hat\alpha_n,\hat\delta_n)$ for (2.8) such that

$$\sqrt{T_n}\big((\hat\alpha_n,\hat\delta_n)-(\alpha,\delta)\big)i(\alpha,\delta)\Sigma_{(\alpha,\delta)}^{-\frac{1}{2}} \xrightarrow{D} N(0,I_2), \tag{2.9}$$

where I_2 denotes the 2×2 identity matrix. This result holds if $\Delta_n\to0$, $T_n\to\infty$ and $\Delta_n T_n\to0$ as $n\to\infty$.

Remark 2.4. (1) Note that in the case of iid observations from NIG$(\alpha,0,0,\delta)$, one gets asymptotic normality of the estimator as in (2.9). However, the asymptotic variance is different as one has to replace $\Sigma_{(\alpha,\delta)}$ by $i(\alpha,\delta)$ and T_n by n.

(2) We have chosen to estimate α and δ by using the likelihood function corresponding to iid observations. Of course, by choosing a contrast function, which does not necessarily correspond to this likelihood function, and finding a minimum contrast estimate it is still

possible to derive consistency or asymptotic normality in terms of the results of Subsection 2.1.

(3) In this paper we have used that the volatility process V appears as the sum of two Ornstein-Uhlenbeck processes with inverse Gaussian marginals. Clearly, the same kind of results apply by considering other Ornstein Uhlenbeck processes.

References

[1] Abramowitz, M. & Stegun, I. A. paperHandbook of mathematical functions with formulas, Graphs and Mathematical Tables. Dover. 1972.

[2] Barndorff-Nielsen, O. E., Normal inverse Gaussian processes and the modelling of stock returns. Research Report 300. Dept. Theor. Statistics, Aarhus University, 1995.

[3] Barndorff-Nielsen, O. E., Normal inverse Gaussian distributions and stochastic Volatility Modelling, *Scand. J. Statist.,* **24,** (1997), 1–13.

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

[5] Barndorff-Nielsen, O. E., Jensen, J. L. & Sørensen, M., Some stationry process in discrete and continuous type, Research Report 241. Dept. Theor. Statistics, Aarhus University, 1995.

[6] Barndorff-Nielsen, O. E. & Jiang, W., An initial analysis of some German stock price series, CAF working paper, No. 15, 1998.

[7] Barndorff-Nielsen, O. E. & Sørensen, M., Asymptotic likelihood theory for stochastic processes, A review, *Internat. Statist. Rev.* **62**, (1994), 133–165.

[8] Bhattacharya, R. N., On the functional central limit theorem and the law of the iterated logarithm for Markov processes, *Z. Wahrscheinlichkeitstheorie verw, Gebiete,* **60**(1982), 185–201.

[9] Black, F. & Scholes, M., The pricing of options and corproate liabilities, *Journal of Political Economy,* **81**(1973), 673–659.

[10] Chesney, M. & Scott, L., Pricing European currency options: A comparison of the modified Black-Scholes model and a random variance model, *Journal of Financial and Quantitative Analysis,* **24**:3(1989), 267–289.

[11] Genon-Catalot, V., Jeantheau, T. & Larédo, C., Parameter estimation for discretely observed stochastic volatility models, *Bernoulli*, (1997).

[12] Genon-Catalot, V., Jeantheau, T. & Larédo, C., Limit theorems for discretely observed stochastic volatility models, *to appear in Bernouli*, (1996).

[13] Helland, I. S., Central Limit Theorems for Martingales with Discrete or Continuous Time, *Scand. J. Statist.* **9**(1982), 79–94.

[14] Hull, J. & White, A., The pricing of options on assets with stochstic volatilities, *Journal of Finance,* **42**(1987), 281–300.

[15] Heston, S. L., A closed-form solution for options with stochastic volatility with applications to Bond and Currency options, *The review of Financial studies,* **6:**2(1993), 327–343.

[16] Scott, L. O., Option pricing when the variance changes randomly : Theory, estimation and an application, *Journal of Financial and Quantitative Analysis,* **22:**4(1987), 419–438.

[17] Stein, E. M. & Stein, J. C., Stock prices distributions with stochastic volatility: An analytic approach, *Journal of Financial Studies,* **4:**4(1991), 727–752.

[18] Wiggins, J. B., Option values under stochastic volatilities. Theory and empirical evidence, *Journal of Financial Economics,* **19**(1987), 351–372.