Bernoulli **9**(3), 2003, 451–465

Nonparametric volatility density estimation

BERT VAN ES[1], PETER SPREIJ[1] and HARRY VAN ZANTEN[2]

[1]*Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands. E-mail:* *vanes@science.uva.nl; **spreij@science.uva.nl

[2]*Division of Mathematics and Computer Science, Faculty of Sciences, Free University Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands. E-mail: harry@cs.vu.nl*

We consider a continuous-time stochastic volatility model. The model contains a stationary volatility process, the density of which, at a fixed instant in time, we aim to estimate. We assume that we observe the process at discrete instants in time. The sampling times will be equidistant with vanishing distance. A Fourier-type deconvolution kernel density estimator based on the logarithm of the squared processes is proposed to estimate the volatility density. An expansion of the bias and a bound on the variance are derived.

Keywords: deconvolution; density estimation; kernel estimator; mixing; stochastic volatility models

1. Introduction

Let S denote the log-price process of some stock on a financial market. It is often assumed that S can be modelled as the solution of a stochastic differential equation or, more generally, as an Itô diffusion process. So we assume that we can write

$$\mathrm{d}S_t = b_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t, \qquad S_0 = 0, \tag{1}$$

or, in integral form,

$$S_t = \int_0^t b_s\,\mathrm{d}s + \int_0^t \sigma_s\,\mathrm{d}W_s, \tag{2}$$

where W is a standard Brownian motion and the processes b and σ are assumed to satisfy certain regularity conditions (see Karatzas and Shreve 1991) for the integrals in (2) to be well defined. In the financial context, the process σ is called a volatility process.

In this paper we model σ as a strictly stationary positive process satisfying a mixing condition, for example an ergodic diffusion on $[0, \infty)$, and we make the assumption that σ is independent of W. We will assume that the one-dimensional marginal distribution of σ has a density with respect to the Lebesgue measure on $(0, \infty)$. This is typically the case in virtually all stochastic volatility models that are proposed in the literature, where the evolution of σ is modelled by a stochastic differential equation, mostly in terms of σ^2 or $\log\sigma^2$; see, for example, Wiggins (1987) and Heston (1993).

For stochastic differential equations of the type

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + a(X_t)\mathrm{d}B_t,$$

1350–7265

with B_t representing Brownian motion, the invariant density is, up to a multiplicative constant, equal to

$$x \mapsto \frac{1}{a^2(x)} \exp\left(2\int_{x_0}^{x} \frac{b(y)}{a^2(y)}\,\mathrm{d}y\right), \tag{3}$$

where x_0 is an arbitrary element of the state space (l, r); see e.g. Gihman and Skorohod (1972) or Skorokhod (1989). From formula (3) one sees that the invariant distribution of the volatility process (take, for instance, X equal to σ^2 or $\log\sigma^2$) may take on many different forms, as is the case for the various models that have been proposed in the literature. This observation supports our point of view that nonparametric procedures are by all means sensible tools for obtaining insight in the behaviour of the volatility.

In the present paper we propose a nonparametric estimator for the volatility density. Using ideas from deconvolution theory, we will propose a procedure for the estimation of the marginal density at a fixed point. We will assume that we observe the log-asset price S at time instants $0, \Delta, 2\Delta, \ldots, n\Delta$, where the time gap satisfies $\Delta = \Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$. To assess the quality of our procedure, we will study how the bias and variance of the estimator behave under these assumptions.

The remainder of the paper is organized as follows. In the next section, we give the heuristic arguments that motivate the definition of our estimator. In Section 3 the main result concerning the asymptotic behaviour of the estimator is presented and discussed. The proof of the main theorem is given in the last two sections.

2. Construction of the estimator

To motivate the construction of the estimator, we first consider (1) without the drift term, so we have

$$\mathrm{d}S_t = \sigma_t\,\mathrm{d}W_t, \qquad S_0 = 0.$$

It is assumed that we observe the process S at the discrete time instants $0, \Delta, 2\Delta, \ldots, n\Delta$. For $i = 1, 2, \ldots$ we work, as in Genon-Catalot *et al*. (1998; 1999), with the normalized increments

$$X_i^\Delta = \frac{1}{\sqrt{\Delta}}(S_{i\Delta} - S_{(i-1)\Delta}).$$

For small Δ, we have the rough approximation

$$X_i^\Delta = \frac{1}{\sqrt{\Delta}}\int_{(i-1)\Delta}^{i\Delta} \sigma_t\,\mathrm{d}W_t \approx \sigma_{(i-1)\Delta}\frac{1}{\sqrt{\Delta}}(W_{i\Delta} - W_{(i-1)\Delta}) = \sigma_{(i-1)\Delta}Z_i^\Delta,$$

where, for $i = 1, 2, \ldots,$ we define

$$Z_i^\Delta = \frac{1}{\sqrt{\Delta}}(W_{i\Delta} - W_{(i-1)\Delta}).$$

By the independence and stationarity of Brownian increments, the sequence $Z_1^\Delta, Z_2^\Delta, \ldots$ is an independent and identically distributed (i.i.d.) sequence of standard normal random variables. Moreover, the sequence is independent of the process σ by assumption.

Taking the logarithm of the square of X_i^Δ, we obtain

$$\log((X_i^\Delta)^2) \approx \log(\sigma^2_{(i-1)\Delta}) + \log((Z_i^\Delta)^2), \tag{4}$$

where the terms in the sum are independent. Assuming that the approximation is sufficiently accurate, we can use this approximate convolution structure to estimate the unknown density f of $\log(\sigma^2_{i\Delta})$ from the observed $\log((X_i^\Delta)^2)$.

Before we can define the estimator, we need some more notation. Observe that the density of the 'noise' $\log(Z_i^\Delta)^2$, denoted by k, is given by

$$k(x) = \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{x/2}\, \mathrm{e}^{-\mathrm{e}^x/2}. \tag{5}$$

The characteristic function of the density k is denoted by ϕ_k.

We will use a function w satisfying the following condition. For examples of such kernels, see Wand (1998).

Condition 2.1. *Let w be a real symmetric function with real-valued symmetric characteristic function ϕ_w with support $[-1, 1]$. Assume, further, that*

(i) $\int_{-\infty}^{\infty} |w(u)|\mathrm{d}u < \infty,\ \int_{-\infty}^{\infty} w(u)\mathrm{d}u = 1,\ \int_{-\infty}^{\infty} u^2|w(u)|\mathrm{d}u < \infty;$
(ii) $\phi_w(1-t) = At^\alpha + o(t^\alpha)$ *as* $t \downarrow 0$, *for some* $\alpha > 0$.

Following a well-known approach in statistical deconvolution theory, we use a *deconvolution kernel density estimator*; see, for example, Section 6.2.4 of Wand and Jones (1995). Having the characteristic functions ϕ_k and ϕ_w at our disposal, choosing a positive *bandwidth* h, we introduce the kernel function

$$v_h(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\phi_w(s)}{\phi_k(s/h)}\, \mathrm{e}^{-\mathrm{i}sx}\, \mathrm{d}s \tag{6}$$

and the density estimator

$$f_{nh}(x) = \frac{1}{nh}\sum_{j=1}^{n} v_h\left(\frac{x - \log((X_j^\Delta)^2)}{h}\right). \tag{7}$$

One can easily verify that the function v_h is real-valued, as is the estimator f_{nh}.

3. Results

To derive the asymptotic behaviour of the estimator, we need a mixing condition on the process σ. For the sake of clarity, we recall the basic definitions. For a certain process X let $\mathcal{F}_a^b$ be the σ-algebra of events generated by the random variables X_t, $a \leqslant t \leqslant b$. The mixing coefficient $\alpha(t)$ is defined by

$$\alpha(t) = \sup_{A\in\mathcal{F}^0_{-\infty}, B\in\mathcal{F}^\infty_t} |P(A \cap B) - P(A)P(B)|. \tag{8}$$

The process X is called *strongly mixing* if $\alpha(t) \to 0$ as $t \to \infty$.

As we mentioned in Section 1, it is common practice to model the volatility process $V = \sigma^2$ as the stationary, ergodic solution of a stochastic differential equation of the form

$$\mathrm{d}V_t = b(V_t)\mathrm{d}t + a(V_t)\mathrm{d}B_t.$$

The mixing condition that we use in Theorem 3.1 below is satisfied in this set-up. See, for instance, Corollary 2.1 of Genon-Catalot *et al*. (2000).

It is easily verified for such processes that $\mathrm{E}|V_t - V_0| = O(t^{1/2})$, provided that $b \in L_1(\mu)$ and $a \in L_2(\mu)$, where μ is the invariant probability measure. Indeed, we have $\mathrm{E}|V_t - V_0| \leqslant \mathrm{E}\int_0^t |b(V_s)|\mathrm{d}S + (\mathrm{E}\int_0^t a^2(V_s)\mathrm{d}S)^{1/2} = t\|b\|_{L_1(\mu)} + \sqrt{t}\|a\|_{L_2(\mu)}$. Although we will not assume explicitly that σ^2 solves an SDE, the above observation motivates the following condition:

Condition 3.1. *We have* $\mathrm{E}|\sigma_t^2 - \sigma_0^2| = O(t^{1/2})$ *for* $t \to 0$.

The following theorem describes the asymptotic behaviour of our estimator f_{nh}. Note that it also covers the case where there is a drift b_t present in (1). The condition on the drift is boundedness of $\mathrm{E}\,b_t^2$. This condition is typically satisfied in realistic models for the log-returns of a stock, since b_t is the local rate of return and this will itself mostly be bounded.

Theorem 3.1. *Assume that* $\mathrm{E}b_t^2$ *is bounded. Let the process* σ *be strongly mixing with coefficient* $\alpha(t)$ *satisfying, for some* $0 < q < 1$,

$$\int_0^\infty \alpha(t)^q\,\mathrm{d}t < \infty,$$

and suppose that Condition 3.1 holds. Let the kernel function w satisfy Condition 2.1 and let the density f *of* $\log\sigma_t^2$ *be continuous, twice continuously differentiable with a bounded second derivative. Also assume that the density of* σ_t^2 *is bounded in a neighbourhood of zero. Suppose that* $\Delta = n^{-\delta}$ *for given* $0 < \delta < 1$ *and choose* $h = \gamma\pi/\log n$, *where* $\gamma > 4/\delta$. *Then the bias of the estimator* (7) *satisfies*

$$\mathrm{E}\,f_{nh}(x) - f(x) = \tfrac{1}{2}h^2 f''(x)\int u^2 w(u)\mathrm{d}u + o(h^2). \tag{9}$$

Moreover, the variance of the estimator satisfies

$$\operatorname{var} f_{nh}(x) = O\left(\frac{1}{n}\,h^{2\alpha}\,\mathrm{e}^{\pi/h}\right) + O\left(\frac{1}{nh^{1+q}\Delta}\right). \tag{10}$$

The proof of the theorem is deferred to the next section. We conclude the present section with a number of comments on the result.

Remark 3.1. The expectation of the deconvolution estimator is equal to the expectation of an ordinary kernel density estimator, as becomes clear from the proof of Lemma 4.1.

It is well known that the variance of kernel-type deconvolution estimators heavily depends on the rate of decay to zero of $|\phi_k(t)|$ as $|t| \to \infty$. The faster the decay the larger the asymptotic variance. The smoother is k, in other words, the harder is the estimation problem. This follows, for instance, for i.i.d. observations from results in Fan (1991) and for stationary observations from the work of Masry (1993).

The rate of decay of $|\phi_k(t)|$ for the density (5) is given by Lemma 5.1 below, which states that $|\phi_k(t)| \sim \sqrt{2}\, e^{-\pi|t|-2}$ as $|t| \to \infty$. This shows that k is supersmooth; see Fan (1991). By the similarity of the tail of this characteristic function to the tail of a Cauchy characteristic function we can expect the same order of the mean squared error as in Cauchy deconvolution problems, where it decreases logarithmically in n; see Fan (1991) for results on i.i.d. observations. Note that this rate, however slow, is faster than the one for normal deconvolution. Fan (1991) also shows that we cannot expect anything better.

Remark 3.2. The choices $\Delta = n^{-\delta}$, with $0 < \delta < 1$, and $h = \gamma\pi/\log n$, with $\gamma > 4/\delta$, render a variance that is of order $n^{-1+1/\gamma}(1/\log n)^{2\alpha}$ for the first term of (10) and $n^{-1+\delta}(\log n)^{1+q}$ for the second term. Since by assumption $\gamma > 4/\delta$, we have $1/\gamma < \delta/4 < \delta$ so the second term dominates the first term. The order of the variance is thus $n^{-1+\delta}(\log n)^{1+q}$. Of course, the order of the bias is logarithmic, hence the bias dominates the variance and the mean squared error of $f_{nh}(x)$ is also logarithmic.

Remark 3.3. Better bounds on the asymptotic variance can be obtained under stronger mixing conditions. Consider, for instance, *uniform mixing*. In this case the mixing coefficient $\phi(t)$ is defined for $t > 0$ as

$$\phi(t) = \sup_{A \in \mathcal{F}_{-\infty}^{0}, B \in \mathcal{F}_{t}^{\infty}} |P(A|B) - P(A)|, \tag{11}$$

and a process is called *uniform mixing* if $\phi(t) \to 0$ for $t \to \infty$. Obviously, uniform mixing implies strong mixing. As a matter of fact, one has the relation

$$\alpha(t) \leqslant \tfrac{1}{2}\phi(t).$$

See Doukhan (1994) for this inequality and many other mixing properties. If σ is uniform mixing with coefficient ϕ satisfying $\int_0^\infty \phi(t)^{1/2}\, dt < \infty$, then the variance bound is given by

$$\operatorname{var} f_{nh}(x) = O\left(\frac{1}{n}\, h^{2\alpha}\, e^{\pi/h}\right) + O\left(\frac{1}{nh\Delta}\right). \tag{12}$$

The proof of this bound runs similarly to the strong-mixing bound. The essential difference is that in (44) below we use Theorem 17.2.3 of Ibragimov and Linnik (1971) with $\tau = 0$ instead of Deo's (1973) lemma, as in the proof of Theorem 2 in Masry (1983).

Remark 3.4. Smoothness conditions on the density at each time of the solution of a stochastic differential equation are guaranteed under Hörmander's condition; see Theorem 2.3.3 in Nualart (1995). Recall also relation (3), which can be used to relate the smoothness of the invariant density to the smoothness of the drift and diffusion coefficients.

4. Proof of Theorem 3.1

We give the proof under the additional assumption that $b_t = 0$. The general case is an easy consequence. Let $\mathcal{F}_\sigma$ denote the sigma field generated by the process σ and let $\tilde{f}_{nh}$ denote the estimator based on the approximating random variables $\sigma_{(j-1)\Delta} Z_j^\Delta$, written as $\tilde{X}_j$, that is,

$$\tilde{f}_{nh}(x) = \frac{1}{nh} \sum_{j=1}^{n} v_h \left(\frac{x - \log(\tilde{X}_j^2)}{h} \right). \tag{13}$$

The proof of (9) follows from the following two lemmas, whose proofs are given in the next section. The first one deals with the expectation of $\tilde{f}_{nh}$.

Lemma 4.1. *We have*

$$\mathrm{E}\tilde{f}_{nh}(x) = \frac{1}{h} \int_{-\infty}^{\infty} w\left(\frac{x-u}{h} \right) f(u) \mathrm{d}u. \tag{14}$$

The second lemma estimates the expected difference between f_{nh} and $\tilde{f}_{nh}$. The bound is in terms of the functions

$$\gamma_0(h) = \frac{1}{2\pi} \int_{-1}^{1} \left| \frac{\phi_w(s)}{\phi_k(s/h)} \right| \mathrm{d}s \tag{15}$$

and

$$\gamma_1(h, x) = \mathrm{e}^{\pi/2h} + \frac{1}{h} \exp\left(\frac{\pi}{2} \frac{1 + \pi/|x|}{h} \right) \log \frac{1 + \pi/|x|}{h}. \tag{16}$$

Lemma 4.2. *For $h \to 0$ and ε small enough, we have*

$$|\mathrm{E} f_{nh}(x) - \mathrm{E}\tilde{f}_{nh}(x)| = O\left(\frac{1}{h^2} \gamma_0(h) \frac{\Delta^{1/4}}{\varepsilon} + \frac{1}{h} \gamma_0(h) \frac{\Delta^{1/2}}{\varepsilon^2} + \gamma_1(h, |\log 2\varepsilon|/h) \frac{\varepsilon}{|\log 2\varepsilon|} \right).$$

Notice that equality (14) is the same as for ordinary kernel estimators; see, for instance, Wand and Jones (1995). Statement (9) of the theorem then follows by combining standard arguments of kernel density estimation and Lemma 4.2. We will show that the bound in Lemma 4.2 is essentially a negative power of n, whereas h^2 is of logarithmic order. Recall that we have assumed $\delta > 4/\gamma$. It follows that $1/2\gamma < \delta/4 - 1/2\gamma$, so we can pick a $\beta \in (1/2\gamma, \delta/4 - 1/2\gamma)$ and take $\varepsilon = n^{-\beta}$. Up to factors that are logarithmic in n, the order of $|\mathrm{E}\, f_{nh}(x) - \mathrm{E}\, \tilde{f}_{nh}(x)|$ is then

$$n^{1/2\gamma - \delta/4 + \beta} + n^{1/2\gamma + 2\beta - \delta/2} + n^{1/2\gamma - \beta}, \tag{17}$$

which is negligible with respect to $h^2 = \gamma^2 \pi^2 / (\log n)^2$ for the chosen values of the parameters.

To prove the bound (10) we use the two lemmas below, which again are proved in the next section. First consider the variance of $\tilde{f}_{nh}(x)$.

Lemma 4.3. *We have, for $h \to 0$,*

$$\operatorname{var} \tilde{f}_{nh}(x) = O\left(\frac{1}{n} h^{2\alpha} \mathrm{e}^{\pi/h}\right) + O\left(\frac{1}{nh^{1+q}\Delta}\right). \tag{18}$$

The next lemma estimates $\operatorname{var}(f_{nh}(x) - \tilde{f}_{nh}(x))$.

Lemma 4.4. *We have, for $h \to 0$ and $\varepsilon > 0$ small enough,*

$$\operatorname{Var}(f_{nh}(x) - \tilde{f}_{nh}(x)) = O\left(\frac{1}{nh^4} \gamma_0(h)^2 \frac{\Delta^{1/2}}{\varepsilon^2} + \frac{1}{n} \gamma_1(h, |\log 2\varepsilon|/h)^2 \frac{\varepsilon}{|\log 2\varepsilon|^2}\right)$$
$$+ \frac{1}{nh^2\Delta} O\left(\frac{\Delta^{(1-q)/2}}{h^2\varepsilon^2} + \varepsilon^{1-q}\right). \tag{19}$$

The proof of (10) is finished as soon as we show that the estimate in Lemma 4.4 is of lower order than that in Lemma 4.3. Up to terms that are logarithmic in n, the bound in Lemma 4.3 is of order $n^{\delta-1}$. Choosing $\varepsilon = n^{-\beta}$ again, up to logarithmic factors, the order of $\operatorname{var}(f_{nh}(x) - \tilde{f}_{nh}(x))$ is

$$n^{-1+1/\gamma-\delta/2+2\beta} + n^{-1+1/\gamma-\beta} + n^{-1+2\beta+\delta(1+q)/2} + n^{-1+\delta-\beta(1-q)}. \tag{20}$$

Recall our assumption that $\delta\gamma > 4$. If we pick β less than $\frac{1}{4}\delta(1-q)$, then all these terms are indeed of lower order than $n^{\delta-1}$.

5. Technical lemmas

5.1. Analytic properties

We need expansions and order estimates for the functions ϕ_k, the kernel v_h as defined in (6), the function γ_0 given as defined in (15) and γ_1 as defined in (16). These are given in the lemmas of this subsection.

Lemma 5.1. *For $|t| \to \infty$ we have*

$$|\phi_k(t)| = \sqrt{2}\, \mathrm{e}^{-\pi|t|/2}\left(1 + O\left(\frac{1}{|t|}\right)\right).$$

Proof. The characteristic function of k is given by

$$\phi_k(t) = \frac{1}{\sqrt{\pi}} 2^{\mathrm{i}t}\, \Gamma(\tfrac{1}{2} + \mathrm{i}t). \tag{21}$$

The result follows by applying the Stirling formula for the complex gamma function; see Chapter 6 in Abramowitz and Stegun (1964). □

Lemma 5.2. *We have the following order estimate for the L^2 norm of v_h. For $h \to 0$,*

$$\|v_h\|_2 = O(h^{1/2+\alpha}\,\mathrm{e}^{\pi/2h}). \tag{22}$$

Proof. By Parseval's identity,

$$\|v_h\|_2^2 = \frac{1}{2\pi}\int_{-1}^{1}\left|\frac{\phi_w(s)}{\phi_k(s/h)}\right|^2 \mathrm{d}s.$$

The integral on the right-hand side is bounded by

$$\frac{1}{2}\int_{-1}^{1}|\phi_w(s)|^2\mathrm{e}^{\pi|s/h|}\mathrm{d}s + \int_{-1}^{1}|\phi_w(s)|^2\left|\frac{1}{|\phi_k(s/h)|^2} - \tfrac{1}{2}\mathrm{e}^{\pi|s/h|}\right|\mathrm{d}s \tag{23}$$

The first term in (23) can be rewritten as

$$\mathrm{e}^{\pi/h}h^{1+2\alpha}\int_0^{1/h}\left|\frac{\phi_w(1-hv)}{(hv)^\alpha}\right|^2 v^{2\alpha}\,\mathrm{e}^{-\pi v}\mathrm{d}v \sim \mathrm{e}^{\pi/h}h^{1+2\alpha}A^2\int_0^\infty v^{2\alpha}\,\mathrm{e}^{-\pi v}\,\mathrm{d}v,$$

by the dominated convergence theorem. The second term in (23) is

$$2h^{1+2\alpha}\,\mathrm{e}^{\pi/h}\int_0^{1/h}\left|\frac{|\phi_w(1-hv)|}{(hv)^\alpha}\right|^2\left|\frac{2\,\mathrm{e}^{-\pi(1/h-v)}}{|\phi_k(1/h-v)|^2} - 1\right|v^{2\alpha}\,\mathrm{e}^{-\pi v}\,\mathrm{d}v,$$

which is of order $O(h^{1+2\alpha}\,\mathrm{e}^{\pi/h})$ by the dominated convergence theorem. We have used the fact that both the functions $\phi_w(1-u)/u^\alpha$ and $|(2\exp(-\pi u)/|\phi_k(u)|^2) - 1|$ are bounded and that the second function is of order $O(1/u)$ as u tends to infinity. This shows that the second term in (23) is negligible with respect to the first. □

Lemma 5.3. *For $h \to 0$, we have*

$$\gamma_0(h) = O\big(h^{1+\alpha}\,\mathrm{e}^{\pi/2h}\big). \tag{24}$$

Proof. The proof is similar to that of Lemma 5.2. □

Lemma 5.4. *The functions v_h are bounded and Lipschitz. More precisely, for all x, we have $|v_h(x)| \leqslant \gamma_0(h)$, and for all x and u,*

$$|v_h(x+u) - v_h(x)| \leqslant \gamma_0(h)|u|. \tag{25}$$

Proof. The bound for $|v_h(x)|$ is obvious. To prove (25) write

$$|v_h(x+u) - v_h(x)| \leqslant \frac{1}{2\pi}\int_{-1}^{1}\left|\frac{\phi_w(s)}{\phi_k(s/h)}\right||\mathrm{e}^{-\mathrm{i}su} - 1|\mathrm{d}s \leqslant \gamma_0(h)|u|.$$

□

Lemma 5.5. *For $x \to \infty$, we have the following estimate on the behaviour of v_h. For some positive constant D,*

$$|v_h(x)| \leqslant \frac{D}{|x|}\gamma_1(h,\, x) \qquad \textit{as } |x| \to \infty, \tag{26}$$

and

$$\gamma_1(h,\, x) = O\left(\frac{|\log h|}{h}\, \mathrm{e}^{\pi(1+\pi/|x|)/2h}\right) \qquad \textit{as } h \to 0. \tag{27}$$

Proof. By a bound in the proof of the Riemann–Lebesgue lemma in Hewitt and Stromberg (1965, p. 402) we have, with $y = \pi/x$,

$$\begin{aligned}
|v_h(x)| &= \frac{1}{2\pi}\left|\int_{-\infty}^{\infty} \frac{\phi_w(s)}{\phi_k(s/h)}\, \mathrm{e}^{-\mathrm{i}sx}\mathrm{d}s\right| \\
&\leqslant \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\frac{\phi_w(s)}{\phi_k(s/h)} - \frac{\phi_w(s+y)}{\phi_k((s+y)/h)}\right|\mathrm{d}s \\
&\leqslant \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\frac{\phi_w(s)-\phi_w(s+y)}{\phi_k(s/h)}\right|\mathrm{d}s \\
&\quad + \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_w(s+y)\left|\frac{1}{\phi_k(s/h)} - \frac{1}{\phi_k((s+y)/h)}\right|\mathrm{d}s.
\end{aligned} \tag{28}$$

First we need a bound on the first integral in (28). Since it follows from Condition 2.1 that ϕ_w is Lipschitz (the proof is similar to that of (25)), with Lipschitz constant C_1, say, we have

$$\begin{aligned}
\int_{-\infty}^{\infty}\left|\frac{\phi_w(s)-\phi_w(s+y)}{\phi_k(s/h)}\right|\mathrm{d}s &\leqslant C_1\int_{-1}^{1}\frac{1}{|\phi_k(s/h)|}\,\mathrm{d}s\,|y| \\
&\leqslant 2C_1\frac{1}{|\phi_k(1/h)|}|y| \sim \frac{C_1}{\sqrt{2}}\,\mathrm{e}^{\pi/2h}\,|y|,
\end{aligned}$$

by Lemma 5.1. To bound the second integral in (28) we need an estimate on the behaviour of $|\phi_k'|/|\phi_k|^2$. Recall expression (21) for ϕ_k. Hence, with $\Psi = \Gamma'/\Gamma$ the digamma function,

$$\begin{aligned}
|\phi_k'(t)| &= \frac{1}{\sqrt{\pi}}\left|\mathrm{i}\log 2\,\mathrm{e}^{\mathrm{i}t\log 2}\Gamma\left(\frac{1}{2}+\mathrm{i}t\right) + \mathrm{i}\mathrm{e}^{\mathrm{i}t\log 2}\Gamma'\left(\frac{1}{2}+\mathrm{i}t\right)\right| \\
&\leqslant \frac{1}{\sqrt{\pi}}\left(\log 2\left|\Gamma\left(\frac{1}{2}+\mathrm{i}t\right)\right| + \left|\Gamma'\left(\frac{1}{2}+\mathrm{i}t\right)\right|\right)
\end{aligned}$$

and, as $|t| \to \infty$,

$$\left|\frac{\phi_k'(t)}{\phi_k(t)^2}\right| \leqslant \sqrt{\pi}\frac{1}{|\Gamma(\frac{1}{2}+\mathrm{i}t)|}\left(\log 2 + \left|\Psi\left(\frac{1}{2}+\mathrm{i}t\right)\right|\right) \leqslant 4\sqrt{\pi}\log(|t|)\mathrm{e}^{\pi|t|/2}, \tag{29}$$

by Lemma 5.1 and by the expansion $|\Psi(z)| \sim \log z$ for $z \to \infty$, $|\mathrm{Arg}\, z| < \pi$; see Chapter 6 of Abramowitz and Stegun (1964). We now return to the second integral in (28) and write

$$\int_{-\infty}^{\infty} \phi_w(s+y)\left|\frac{1}{\phi_k(s/h)} - \frac{1}{\phi_k((s+y)/h)}\right| \mathrm{d}s$$

$$= \int_{-1}^{1} \phi_w(s)\left|\frac{1}{\phi_k((s-y)/h)} - \frac{1}{\phi_k(s/h)}\right| \mathrm{d}s$$

$$\leqslant \frac{2}{h} \sup_{(-1-|y|)/h \leqslant s \leqslant (1+|y|)/h} \left|\frac{\phi_k'(s)}{\phi_k(s)^2}\right| |y|$$

$$\leqslant \frac{2}{h} \sup_{(-1-|y|)/h \leqslant s \leqslant (1+|y|)/h} 4\sqrt{\pi}\log(|s|)\mathrm{e}^{\pi|s|/2}|y|$$

$$= \frac{8}{h}\sqrt{\pi}\log((1+|y|)/h)\mathrm{e}^{\pi(1+|y|)/2h}|y|$$

in view of (29). This completes the proof. □

5.2. Proofs of Lemmas 4.1–4.4

Recall that $\mathcal{F}_\sigma$ is the σ-algebra generated by the process σ.

Proof of Lemma 4.1. Write

$$\mathrm{E}(\tilde{f}_{nh}(x)|\mathcal{F}_\sigma) = \frac{1}{nh}\sum_{t=1}^{n} \mathrm{E}\left(v_h\left(\frac{x - \log\sigma^2_{(t-1)\Delta} - \log(Z_t^\Delta)^2}{h}\right)|\mathcal{F}_\sigma\right)$$

$$= \frac{1}{nh}\sum_{t=1}^{n}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\phi_w(s)}{\phi_k(s/h)}\mathrm{E}\left(\mathrm{e}^{-\mathrm{i}s(x-\log\sigma^2_{(t-1)\Delta}-\log(Z_t^\Delta)^2)/h}|\mathcal{F}_\sigma\right)\mathrm{d}s$$

$$= \frac{1}{nh}\sum_{t=1}^{n}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\phi_w(s)}{\phi_k(s/h)}\,\mathrm{e}^{-\mathrm{i}sx/h}\,\mathrm{e}^{\mathrm{i}s\log\sigma^2_{(t-1)\Delta}/h}\phi_k(s/h)\mathrm{d}s$$

$$= \frac{1}{nh}\sum_{t=1}^{n}\frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_w(s)\mathrm{e}^{-\mathrm{i}s(x-\log\sigma^2_{(t-1)\Delta})/h}\,\mathrm{d}s$$

$$= \frac{1}{nh}\sum_{t=1}^{n} w\left(\frac{x-\log\sigma^2_{(t-1)\Delta}}{h}\right).$$

By taking expectations the result follows. □

For the proof of Lemma 4.2 we need a few properties of the process σ, valid under Condition 3.1. Since $(x-y)^2 \leqslant |x^2-y^2|$ for $x, y \geqslant 0$, we have that $\mathrm{E}(\sigma_t-\sigma_0)^2 = O(t^{1/2})$ for $t \to 0$. Consequently, there exists a constant $C > 0$ such that

$$\mathrm{E}(X_1^\Delta - \sigma_0 Z_1^\Delta)^2 \leqslant C\Delta^{1/2} \qquad \text{for } \Delta \to 0, \tag{30}$$

since $\mathrm{E}(X_1^\Delta - \sigma_0 Z_1^\Delta)^2 = \frac{1}{\Delta}\mathrm{E}\int_0^\Delta (\sigma_t - \sigma_0)^2 \mathrm{d}t$. Moreover, Condition 3.1 implies that

$$\mathrm{E}\left|\frac{1}{\Delta}\int_0^\Delta \sigma_t^2\,\mathrm{d}t - \sigma_0^2\right| = O(\Delta^{1/2}) \qquad \text{for } \Delta \to 0. \tag{31}$$

Proof of Lemma 4.2. Writing

$$W_j = v_h\left(\frac{x - \log((X_j^\Delta)^2)}{h}\right) - v_h\left(\frac{x - \log(\tilde{X}_j^2)}{h}\right),$$

so that $f_{nh}(x) - \tilde{f}_{nh}(x) = (1/nh)\sum_{j=1}^n W_j$, we have

$$\begin{aligned}|\mathrm{E}f_{nh}(x) - \mathrm{E}\,\tilde{f}_{nh}(x)| &\leqslant \frac{1}{h}\,\mathrm{E}\,|W_j| \\ &= \frac{1}{h}\mathrm{E}|W_j|I_{[|X_1^\Delta|\geqslant\varepsilon \text{ and } |\tilde{X}_1|\geqslant\varepsilon]} \\ &\quad + \frac{1}{h}\mathrm{E}|W_j|I_{[|X_1^\Delta|\leqslant\varepsilon \text{ or } |\tilde{X}_1|\leqslant\varepsilon]}I_{[|X_1^\Delta - \tilde{X}_1|\geqslant\varepsilon]} \\ &\quad + \frac{1}{h}\mathrm{E}|W_j|I_{[|X_1^\Delta|\leqslant\varepsilon \text{ or } |\tilde{X}_1|\leqslant\varepsilon]}I_{[|X_1^\Delta - \tilde{X}_1|<\varepsilon]}.\end{aligned} \tag{32}$$

By Lemma 5.4 and (30) the first term on the right in (32) can be bounded by

$$\begin{aligned}\frac{2}{h^2}\gamma_0(h)\mathrm{E}|\log(X_1^\Delta) - \log(\tilde{X}_1)|I_{[|X_1^\Delta|\geqslant\varepsilon \text{ and } |\tilde{X}_1|\geqslant\varepsilon]} &\leqslant \frac{2}{h^2}\frac{1}{\varepsilon}\gamma_0(h)\mathrm{E}|X_1^\Delta - \tilde{X}_1| \\ &\leqslant \frac{2}{h^2}\,\gamma_0(h)\sqrt{C}\frac{\Delta^{1/4}}{\varepsilon}.\end{aligned}$$

In the same way, the second term can be bounded by

$$\frac{2}{h}\gamma_0(h)P(|X_1^\Delta - \tilde{X}_1| \geqslant \varepsilon) \leqslant \frac{2}{h}\gamma_0(h)C\frac{\Delta^{1/2}}{\varepsilon^2}.$$

Since the absolute value of both arguments of v_h below is eventually larger than $|\log 2\varepsilon|/h$, by Lemma 5.5 the third term on the right in (32) can be bounded by

$$\frac{1}{h}\gamma_1(h, |\log 2\varepsilon|/h)\frac{1}{(|\log 2\varepsilon|/h)}P(|\tilde{X}_1| \leqslant 2\varepsilon) \leqslant C_2\,\gamma_1(h, |\log 2\varepsilon|/h)\frac{\varepsilon}{|\log 2\varepsilon|},$$

for some constant C_2. Here we have used the fact that the density of $\tilde{X}_1$ is bounded, which follows from the assumption that σ_0^2 has a bounded density in a neighbourhood of zero. □

Proof of Lemma 4.3. Consider the decomposition

$$\operatorname{var}(\tilde{f}_{nh}(x)) = \operatorname{var}(\mathrm{E}(\tilde{f}_{nh}(x)|\mathcal{F}_\sigma)) + \mathrm{E}(\operatorname{var}(\tilde{f}_{nh}(x)|\mathcal{F}_\sigma)). \tag{33}$$

By the proof of Lemma 4.1 the conditional expectation $\mathrm{E}(\tilde{f}_{nh}(x)|\mathcal{F}_\sigma)$ is equal to a kernel estimator of the density of $\log\sigma_t^2$. By Theorem 3 of Masry (1983), we can bound its variance by

$$\frac{20(1+o(1))}{nh^{1+q}\Delta} f(x)^{1-q}\left(\int_{-\infty}^{\infty}|w(u)|^{2/(1-q)}\mathrm{d}u\right)^{1-q}\int_0^\infty \alpha(\tau)^q \mathrm{d}\tau = O\left(\frac{1}{nh^{1+q}\Delta}\right).$$

Given the process σ, the random variables $\log\tilde{X}_t^2$ are independent, so we can bound the second term in (33) by

$$\frac{1}{n^2h^2}\sum_{t=1}^n \operatorname{var}\left(v_h\left(\frac{x-\log\tilde{X}_t^2}{h}\right)\right) \leqslant \frac{1}{nh^2}\,\mathrm{E}\left(v_h\left(\frac{x-\log\tilde{X}_1^2}{h}\right)\right)^2 \leqslant \frac{1}{nh^2}\gamma_0(h)^2,$$

by Lemma 5.4. The result follows by an application of Lemma 5.3. □

Proof of Lemma 4.4. Note that for different i, j, conditional on the process σ, the pairs X_i^Δ, $\tilde{X}_i$ and X_j^Δ, $\tilde{X}_j$ are independent. Hence the conditional covariances of functions of these pairs vanish.

With W_j as in the proof of Lemma 4.2, we have

$$\begin{aligned}&\operatorname{var}(f_{nh}(x)-\tilde{f}_{nh}(x))\\ &\quad = \frac{1}{nh^2}\operatorname{var} W_1 + \frac{1}{n^2h^2}\sum_{i\neq j}\operatorname{cov}(\mathrm{E}(W_i|\mathcal{F}_\sigma), \mathrm{E}(W_j|\mathcal{F}_\sigma)).\end{aligned} \tag{34}$$

Let us first derive a bound on $\operatorname{var} W_1$. We have $\operatorname{var} W_1 \leqslant \mathrm{E}\, W_1^2$, which can be split into three terms

$$\begin{aligned}&\frac{1}{h}\mathrm{E} W_j^2 I_{[|X_1^\Delta|\geqslant\varepsilon \text{ and } |\tilde{X}_1|\geqslant\varepsilon]}\\ &\quad+\frac{1}{h}\mathrm{E} W_j^2 I_{[|X_1^\Delta|\leqslant\varepsilon \text{ or } |\tilde{X}_1|\leqslant\varepsilon]} I_{[|X_1^\Delta-\tilde{X}_1|\geqslant\varepsilon]}\\ &\quad+\frac{1}{h}\mathrm{E} W_j^2 I_{[|X_1^\Delta|\leqslant\varepsilon \text{ or } |\tilde{X}_1|\leqslant\varepsilon]} I_{[|X_1^\Delta-\tilde{X}_1|<\varepsilon]}.\end{aligned} \tag{35}$$

By (30) and Lemma 5.4 the first term in (35) can be bounded by $(2/h^2)\gamma_0(h)^2 C\Delta^{1/2}/\varepsilon^2$. Again by (30) and Lemma 5.4 the second term in (35) can be bounded by

$$4\gamma_0(h)^2 P(|X_1^\Delta - \tilde{X}_1| \geqslant \varepsilon) \leqslant 4\gamma_0(h)^2 C\frac{\Delta^{1/2}}{\varepsilon^2}.$$

Since the absolute value of both arguments of v_h below is eventually larger than $|\log 2\varepsilon|/h$, by Lemma 5.5, the third term in (35) can be bounded by

$$\frac{\gamma_1(h, \log 2\varepsilon|/h)^2}{(|\log 2\varepsilon|/h)^2} P(|\tilde{X}_1| \leqslant 2\varepsilon) \leqslant C_2\, h^2\gamma_1(h, |\log 2\varepsilon|/h)^2 \frac{\varepsilon}{|\log 2\varepsilon|^2},$$

for some constant C_2, where we again use the fact, as in the proof of Lemma 4.2, that the density of $\tilde{X}_1$ is bounded. We obtain

$$\mathrm{E}W_1^2 = O\left(\frac{1}{h^2}\gamma_0(h)^2 C\frac{\Delta^{1/2}}{\varepsilon^2} + h^2\gamma_1(h, |\log 2\varepsilon|/h)^2\frac{\varepsilon}{|\log 2\varepsilon|^2}\right), \tag{36}$$

which gives the order bound for the first term on the right in (19).

Next, we concentrate on the sum of covariances in (34). Define

$$\overline{\sigma}_i = \frac{1}{\Delta}\int_{(i-1)\Delta}^{i\Delta}\sigma_t^2 \mathrm{d}t. \tag{37}$$

Note that given $\mathcal{F}_\sigma$, X_i^Δ is $N(0, \overline{\sigma}_i)$ distributed and $\tilde{X}_i$ is $N(0, \sigma^2_{(i-1)\Delta})$. As in the proof of Lemma 4.1, it follows that

$$\mathrm{E}(W_i|\mathcal{F}_\sigma) = w\left(\frac{x - \log\overline{\sigma}_i}{h}\right) - w\left(\frac{x - \log\sigma^2_{(i-1)\Delta}}{h}\right).$$

We follow the line of argument in the proof of Theorem 3 in Masry (1983). The stationarity of W_j implies that the conditional expectations $\tilde{W}_j := \mathrm{E}(W_j|\mathcal{F}_\sigma)$ are also stationary. Hence we have

$$\sum_{i\neq j}\mathrm{cov}(\tilde{W}_i, \tilde{W}_j) = 2\sum_{k=1}^{n-1}(n-k)\mathrm{cov}(\tilde{W}_0, \tilde{W}_k).$$

Now note that the process $\tilde{W}_j$ is strongly mixing with a mixing coefficient $\tilde{\alpha}(k) \leqslant \alpha((k-1)\Delta)$, $k = 1, 2, \ldots,$ where α is the coefficient of the process σ. By a lemma of Deo (1973) for strongly mixing processes it follows that, for all $\tau > 0$,

$$|\mathrm{cov}(\tilde{W}_0, \tilde{W}_k)| \leqslant 10\alpha((k-1)\Delta)^{\tau/(2+\tau)}\left(\mathrm{E}|\tilde{W}_1|^{2+\tau}\right)^{2/(2+\tau)}. \tag{38}$$

By the monotonicity of the mixing coefficient α we obtain

$$\begin{aligned}&\left|\frac{1}{n^2h^2}\sum_{i\neq j}\mathrm{cov}(\tilde{W}_i, \tilde{W}_j)\right| \\ &\leqslant \frac{10}{nh^2}\left(\mathrm{E}|\tilde{W}_1|^{2+\tau}\right)^{2/(2+\tau)}\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\alpha((k-1)\Delta)^{\tau/(2+\tau)} \\ &\leqslant \frac{10}{nh^2}\left(\alpha(0)^{\tau/(2+\tau)} + \frac{1}{\Delta}\int_0^\infty \alpha(t)^{\tau/(2+\tau)}\,\mathrm{d}t\right)\left(\mathrm{E}\,|\tilde{W}_1|^{2+\tau}\right)^{2/(2+\tau)}.\end{aligned}$$

Next we derive a bound on $\mathrm{E}|\tilde{W}_1|^{2+\tau}$. Fix $\kappa \in (0, 1]$. We have

$$
\begin{aligned}
\mathrm{E}|\tilde{W}_1|^{2+\tau} &= \mathrm{E}\left|w\left(\frac{x-\log(\bar{\sigma}_1)}{h}\right) - w\left(\frac{x-\log(\sigma_0^2)}{h}\right)\right|^{2+\tau} I_{[\bar{\sigma}_1 \geqslant \varepsilon \text{ and } \sigma_0^2 \geqslant \varepsilon]} \\
&+ \mathrm{E}\left|w\left(\frac{x-\log(\bar{\sigma}_1)}{h}\right) - w\left(\frac{x-\log(\sigma_0^2)}{h}\right)\right|^{2+\tau} I_{[\bar{\sigma}_1 \leqslant \varepsilon \text{ or } \sigma_0^2 \leqslant \varepsilon]} I_{[|\bar{\sigma}_1^\kappa - \sigma_0^{2\kappa}| \geqslant \varepsilon]} \\
&+ \mathrm{E}\left|w\left(\frac{x-\log(\bar{\sigma}_1)}{h}\right) - w\left(\frac{x-\log(\sigma_0^2)}{h}\right)\right|^{2+\tau} I_{[\bar{\sigma}_1 \leqslant \varepsilon \text{ or } \sigma_0^2 \leqslant \varepsilon]} I_{[|\bar{\sigma}_1^\kappa - \sigma_0^{2\kappa}| < \varepsilon]}. \qquad (39)
\end{aligned}
$$

Note that by Condition 2.1 and Fourier inversion w is Lipschitz with constant $1/\pi$ and bounded by $1/\pi$. Hence the first term on the right in (39) can be bounded by $\mathrm{E}|\bar{\sigma}_1^\kappa - \sigma_0^{2\kappa}|^{2+\tau}/(\kappa\varepsilon h)^{2+\tau}$. The second term can be bounded by

$$
P(|\bar{\sigma}_1^\kappa - \sigma_0^{2\kappa}| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^{2+\tau}} \mathrm{E}\,|\bar{\sigma}_1^\kappa - \sigma_0^{2\kappa}|^{2+\tau}.
$$

Likewise, the third term can be bounded by

$$
\mathrm{E}\left|w\left(\frac{x-\log(\bar{\sigma}_1)}{h}\right) - w\left(\frac{x-\log(\sigma_0^2)}{h}\right)\right|^{2+\tau} I_{[\bar{\sigma}_1 \leqslant \varepsilon(1+\varepsilon^{1-\kappa})^{1/\kappa} \text{ and } \sigma_0^2 \leqslant \varepsilon(1+\varepsilon^{1-\kappa})^{1/\kappa}]},
$$

which is bounded by $P(\sigma_0^2 \leqslant 2\varepsilon) = O(\varepsilon)$ since σ_0^2 was assumed to have a bounded density in a neighbourhood of zero.

With $\tau = 2q/(1-q)$ and $\kappa = 1/(2+\tau) = (1-q)/2$ we have, with an application of the basic inequality $|u^\kappa - v^\kappa| \leqslant |u-v|^\kappa$ for $u, v \geqslant 0$ and $\kappa \in (0, 1]$ in the second equality below and from Condition 3.1 and its consequence (31) in the fourth equality,

$$
\begin{aligned}
&\left|\frac{1}{n^2 h^2}\sum_{i \neq j} \mathrm{cov}(\tilde{W}_1, \tilde{W}_j)\right| \\
&= \frac{1}{nh^2\Delta} O\left(\frac{1}{h^{2+\tau}} \frac{1}{\varepsilon^{2+\tau}} \mathrm{E}\,|\bar{\sigma}_1^\kappa - \sigma_0^{2\kappa}|^{2+\tau} + \varepsilon\right)^{2/(2+\tau)} \\
&= \frac{1}{nh^2\Delta} O\left(\frac{1}{h^{2+\tau}} \frac{1}{\varepsilon^{2+\tau}} \mathrm{E}\,|\bar{\sigma}_1 - \sigma_0^2|^{\kappa(2+\tau)} + \varepsilon\right)^{2/(2+\tau)} \\
&= \frac{1}{nh^2\Delta} O\left(\frac{(\mathrm{E}\,|\bar{\sigma}_1 - \sigma_0^2|)^{2/(2+\tau)}}{h^2\varepsilon^2} + \varepsilon^{2/(2+\tau)}\right), \\
&= \frac{1}{nh^2\Delta} O\left(\frac{\Delta^{1/(2+\tau)}}{h^2\varepsilon^2} + \varepsilon^{2/(2+\tau)}\right) \\
&= \frac{1}{nh^2\Delta} O\left(\frac{\Delta^{(1-q)/2}}{h^2\varepsilon^2} + \varepsilon^{1-q}\right),
\end{aligned}
$$

which gives the order bound for the second term on the right in (19). □

References

Abramowitz, M. and Stegun, I. (1964) *Handbook of Mathematical Functions (9th edn).* Dover, New York: Dover.

Deo, C.M. (1973) A note on empirical processes for strong mixing processes. *Ann. Probab.*, **1**, 870–875.

Doukhan, P. (1994) *Mixing, Properties and Examples*, Lecture Notes in Statist. 85. New York: Springer-Verlag.

Fan, J. (1991) On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.*, **19**, 1257–1272.

Genon-Catalot, V., Jeantheau, T. and Larédo, C. (1998) Limit theorems for discretely observed stochastic volatility models. *Bernoulli*, **4**, 283–303.

Genon-Catalot, V., Jeantheau, T. and Larédo, C. (1999) Parameter estimation for discretely observed stochastic volatility models. *Bernoulli*, **5**, 855–872.

Genon-Catalot, V., Jeantheau, T. and Larédo, C. (2000) Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli*, **6**, 1051–1079.

Gihman, I.I. and Skorohod A.V. (1972) *Stochastic Differential Equations*. Berlin: Springer-Verlag.

Heston, S.L. (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Stud.*, **6**, 327–343.

Hewitt, E. and Stromberg K. (1965) *Real and Abstract Analysis*. New York: Springer-Verlag.

Ibragimov, I.A. and Linnik (1971) *Independent and Stationary Sequences of Random Variables*. Groningen: Wolters-Noordhoff.

Karatzas, I. and Shreve, S.E. (1991) *Brownian Motion and Stochastic Calculus*. New York, Springer-Verlag.

Masry, E. (1983) Probability density estimation from sampled data. *IEEE Trans. Inform. Theory*, **29**, 696–709.

Masry, E. (1993) Strong consistency and rates for deconvolution of multivariate densities of stationary processes. *Stochastic Process. Appl.*, **47**, 53–74.

Nualart, D. (1995) *The Malliavin Calculus and Related Topics*. New York: Springer-Verlag.

Skorokhod, A.V. (1989) *Asymptotic Methods in the Theory of Stochastic Differential Equations*. Providence, RI: American Mathematical Society.

Wand, M.P. (1998) Finite sample performance of deconvolving kernel density estimators, *Statist. Probab. Lett.*, **37**, 131–139.

Wand, M.P. and Jones, M.C. (1995) *Kernel Smoothing*. London: Chapman & Hall.

Wiggins, J.B. (1987) Option valuation under stochastic volatility, *J. Financial Econom.*, **19**, 351–372.

Received August 2001 and revised March 2002