Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 118 (2008) 1434–1462
stochastic processes and their applications
www.elsevier.com/locate/spa

Estimation of the volatility persistence in a discretely observed diffusion model

Mathieu Rosenbaum*

Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées, UMR CNRS 8050, France
CREST-ENSAE, Timbre J120, 3 avenue Pierre Larousse, 92245 Malakoff Cedex, France

Received 16 February 2006; received in revised form 6 June 2007; accepted 16 September 2007
Available online 22 September 2007

Abstract

We consider the stochastic volatility model

$$\mathrm{d}Y_t = \sigma_t \,\mathrm{d}B_t,$$

with B a Brownian motion and σ of the form

$$\sigma_t = \Phi\left(\int_0^t a(t,u)\mathrm{d}W_u^H + f(t)\xi_0\right),$$

where W^H is a fractional Brownian motion, independent of the driving Brownian motion B, with Hurst parameter $H \geq 1/2$. This model allows for persistence in the volatility σ. The parameter of interest is H. The functions Φ, a and f are treated as nuisance parameters and ξ_0 is a random initial condition. For a fixed objective time T, we construct from discrete data $Y_{i/n}, i = 0, \ldots, nT$, a wavelet based estimator of H, inspired by adaptive estimation of quadratic functionals. We show that the accuracy of our estimator is $n^{-1/(4H+2)}$ and that this rate is optimal in a minimax sense.

MSC: 60G18; 60G99; 60H05; 60H40; 60G15; 62F12; 62F99; 62M09; 62P05

Keywords: Stochastic volatility models; Discrete sampling; High frequency data; Fractional Brownian motion; Scaling exponent; Adaptive estimation of quadratic functionals; Wavelet methods

* Corresponding address: CREST-ENSAE, Timbre J120, 3 avenue Pierre Larousse, 92245 Malakoff Cedex, France.
E-mail address: rosenbaum@ensae.fr.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2007.09.004

1. Introduction

1.1. Stochastic volatility and volatility persistence

Since the celebrated model of Black and Scholes, the behavior of financial assets is often modelled by processes of type

$$\mathrm{d}S_t = \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}B_t,$$

where S is the price (or the log-price) of the asset, B a Brownian motion and μ a drift process. The volatility coefficient σ represents the fluctuations of S and plays a crucial role in trading, option pricing and hedging. It is well known that stochastic volatility models, where the volatility is a random process, provide a way to deal with the endemic time-varying volatility and to reproduce various stylized facts observed on the markets; see Shephard [34], Barndorff-Nielsen, Nicolato and Shephard [3]. Among these stylized facts, there are many arguments about volatility persistence. This presence of memory in the volatility has in particular consequences for option pricing; see Ohanissian, Russel and Tsay [32], Taylor [35], Comte, Coutin and Renault [11]. Hence, continuous time dynamics have been introduced to capture this phenomenon; see Comte and Renault [12], Comte, Coutin and Renault [11] or Barndorff-Nielsen and Shephard [4]. However, in statistical finance, the question of the volatility persistence has been mostly treated with discrete time models; see among others Breidt, Crato and De Lima [6], Harvey [18], Andersen and Bollerslev [1], Robinson [33], Hurvich and Soulier [22], Teyssière [36]. Concurrently, statistical methods for detecting this volatility persistence have been specifically developed for these models; see Hurvich, Moulines and Soulier [20], Deo, Hurvich and Lu [14], Hurvich and Ray [21], Lee [26], Jensen [24]. In this paper, our objective is to build, for continuous time models, a statistical program allowing us to recover information about the volatility persistence.

1.2. A diffusion model with fractional stochastic volatility

We consider a class of diffusion models whose volatility is a non-linear transformation of a stochastic integral with respect to fractional Brownian motion. Recall that a fractional Brownian motion $(W_t^H, t \in \mathbb{R})$, with Hurst parameter $H \in (0, 1]$, is a self-similar centered Gaussian process with covariance function

$$\mathbb{E}[W_t^H W_s^H] = \frac{1}{2}(|s|^{2H} + |t|^{2H} - |t-s|^{2H}).$$

If $H > 1/2$, the use of fractional Brownian motion (fbm for short) is a way to allow for persistence. Indeed, its increments are then stationary, positively correlated and the value of the Hurst parameter quantifies the presence of so-called long memory between them; see Mandelbrot and Van Ness [27], Taqqu [15]. We define on a rich enough probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a Brownian motion B, a fractional Brownian motion W^H, independent of B, with unknown Hurst parameter $H \in (1/2, 1)$, and a random variable ξ_0, measurable with respect to the sigma algebra generated by $(W_t^H, t \le 0)$. We fix an objective time $T > 0$ and we consider the one-dimensional stochastic process Y defined by

$$Y_t = y_0 + \int_0^t \sigma_s \,\mathrm{d}B_s, \quad y_0 \in \mathbb{R},\ t \in [0, T], \tag{1}$$

where y_0 is deterministic and σ is another one-dimensional stochastic process of the form

$$\sigma_t = \Phi\left(\int_0^t a(t,u)\mathrm{d}W_u^H + f(t)\xi_0\right). \tag{2}$$

The functions Φ, a and f are deterministic and unknown. Since we only consider continuously differentiable integrands, the stochastic integral with respect to the fractional Brownian motion W^H with $H \in (1/2, 1)$ is simply defined as a pathwise Riemann–Stieltjes integral. In particular, this definition gives that for a continuously differentiable real function g,

$$\int_0^t g(u)\mathrm{d}W_u^H = -\int_0^t g'(u)W_u^H\mathrm{d}u + g(t)W_t^H.$$

This framework is an extension of the long memory stochastic volatility model introduced in mathematical finance by Comte and Renault [12]. We retrieve the volatility function used by Comte and Renault in [12] taking

$$\Phi(x) = \exp(x), \qquad a(t,u) = \gamma\exp(-k[t-u]), \qquad \xi_0 = 0,$$

where k and γ are positive constant parameters. Its stationary version, that is the exponential of a long memory fractional Ornstein–Uhlenbeck process, is obtained taking the same specification for Φ and a and

$$f(t) = \gamma\exp(-kt), \qquad \xi_0 = \int_{-\infty}^0 \exp(ku)\mathrm{d}W_u^H;$$

see Cheridito et al. [7] for details. For the preceding specification of the volatility process, Comte and Renault have shown in [12] that $\mathrm{Cov}[\sigma_{t+h}, \sigma_t]$ is of order $|h|^{-(1-2d)}$ as h tends to infinity, with $d = H - 1/2$. Hence, not only the logarithm of the volatility but also the volatility process itself entails long memory with long memory parameter $d = H - 1/2$.[1]

Remark also that in the limit case $H = 1/2$, under smoothness assumptions on Φ, letting

$$a = 1, \qquad f = 0, \qquad g = (\Phi^2)' \circ \Phi^{-1} \qquad \text{and} \qquad h = (\Phi^2)'' \circ \Phi^{-1},$$

we equivalently have

$$\mathrm{d}\sigma_t^2 = h(\sigma_t^2)\mathrm{d}t + g(\sigma_t^2)\mathrm{d}W_t.$$

Thus, we (partially) retrieve the standard stochastic volatility diffusion framework; see for example Hull and White [19], Melino and Turnbull [28] or Musiela and Rutkowski [29] for a more exhaustive study.

For $I \subseteq \mathbb{R}$, we denote by $\mathcal{C}^k(I)$ the set of all deterministic k-times-differentiable functions from I to $\mathbb{R}$. The assumptions on a, Φ, f and ξ_0 in model (1) and (2) are the following.

Assumption A. (i) For all $t \in [0,T]$, $u \to a(t,u) \in \mathcal{C}^2([0,T])$, with bounded derivative uniformly in t.

(ii) For all $u \in [0,T]$, $t \to a(t,u) \in \mathcal{C}^2([0,T])$, with bounded derivatives uniformly in u.

[1] Note that if $\{X_t, t \in \mathbb{Z}\}$ is a stationary long memory Gaussian process, the statement that $\Phi(X_t)$ is also a long memory process with the same memory parameter is not always true. This is the case provided the linear term in the Hermite expansion of $\Phi(X_t)$ does not vanish. In every instance, a non-linear transform $\Phi(X_t)$ cannot "increase" the memory of X_t.

(iii) $t \to f(t) \in \mathcal{C}^2([0, T])$.
(iv) For $p > 0$, $\mathbb{E}[e^{p|\xi_0|}] < \infty$.
(v) There exist $0 \leq \beta_1 < \beta_2 \leq T$ such that $\inf_{u \in [\beta_1, \beta_2]} a^2(u, u) > 0$.

Assumption B. (i) $x \to \Phi(x) \in \mathcal{C}^2(\mathbb{R})$.
(ii) For some $c_1 > 0$, $c_2 > 0$ and $\gamma \geq 0$, $|(\Phi^2)'(x)| \geq c_1|x|^\gamma \mathbb{1}_{|x| \in [0,1]} + c_2 \mathbb{1}_{|x|>1}$.
(iii) For some $c_3 > 0$, $|(\Phi^2)''(x)| \leq c_3 e^{|x|}$.

1.3. Statistical model and results

We consider model (1) and (2). For technical convenience (see Section 2.1.2), we take $T \geq 3$. We observe the diffusion at the sampling frequency n. This means that we observe

$$Y^n = \{Y_{i/n}, i = 0, \ldots, nT\}.$$

For simplicity, we assume throughout the paper $n = 2^N$. We study the problem of the inference of H based on Y^n.

A rate $v_n \to 0$ is said to be achievable over $\mathcal{H} \subset (1/2, 1)$ if there exists an estimator $\widehat{H}_n = \widehat{H}_n(Y^n)$ such that the normalized error

$$\{v_n^{-1}(\widehat{H}_n - H)\}_{n \geq 1} \tag{3}$$

is bounded in probability, uniformly over $\mathcal{H}$. The rate v_n is moreover a lower rate of convergence on $\mathcal{H}$ if there exists $C > 0$ such that

$$\liminf_{n \to \infty} \inf_F \sup_{H \in \mathcal{H}} \mathbb{P}[v_n^{-1}|F - H| \geq C] > 0, \tag{4}$$

where the infimum is taken over all estimators $F = F(Y^n)$. We prove in this paper that the rate $v_n(H) = n^{-1/(4H+2)}$ is optimal in a minimax sense. This means that (3) and (4) agree with $v_n = v_n(H)$. We also exhibit an optimal estimator based on the behavior of the wavelet coefficients of the process σ^2.

Theorem 1. *Under Assumptions A and B, the rate $v_n(H) = n^{-1/(4H+2)}$ is achievable over every compact set $\mathcal{H} \subset (1/2, 1)$. Moreover, the estimator $\widehat{H}_n$ explicitly constructed in Section 2.2 achieves the rate $v_n(H)$.*

Our next result shows that, under an additional restriction on the non-degeneracy of the model and on the initial condition, this result is indeed optimal.

Assumption C. The variable ξ_0 is deterministic. Moreover, for some $c_4 > 0$, $c_5 > 0$, $c_4 \neq c_5$ and $c_6 > 0$, we have $c_4 \leq |\Phi(x)| \leq c_5$ and $|\Phi'(x)| \leq c_6$.

Theorem 2. *Under Assumptions A–C, the rate $v_n(H) = n^{-1/(4H+2)}$ is a lower rate of convergence over every compact set $\mathcal{H} \subset (1/2, 1)$ with non-empty interior.*

1.4. Discussion

• Contrary to other works, ours does not consider intrinsically discrete data, but considers discretely observed data from an underlying continuous time process. Thus, as the objective time T is fixed, the dynamic between two data depends on the sampling frequency. This approach substantially differs from those based on ergodic properties. In our context, the available information quantity does not increase because of a longer observation period but because of a higher sampling frequency. The estimation rates are naturally different according to the approaches. Compare our accuracy with the rate $n^{-(2/5-\varepsilon)}$ obtained by Hurvich, Moulines and Soulier in an ergodic context; see [20].

• Through this model, we aim at showing that we can recover the smoothness of the volatility from historical data. The following proposition, whose proof is given in Appendix, shows that the Hurst parameter can be interpreted as a regularity parameter thanks to Besov smoothness spaces (see the Appendix).

Proposition 1 (*Smoothness of the Volatility Process*). *For large enough T, under Assumptions A and B, in model (1) and (2),*

(i) *Almost surely, the trajectory of $t \to \sigma_t^2$ belongs to the Besov space $\mathcal{B}^H_{2,\infty}([0,T])$ but, for all $q<\infty$, a.s. it does not belong to $\mathcal{B}^H_{2,q}([0,T])$.*

(ii) *For all $s<H$, almost surely, the trajectory of $t \to \sigma_t^2$ belongs to the Besov space $\mathcal{B}^s_{\infty,\infty}([0,T])$ but, if moreover there exists $c>0$ such that $|(\Phi^2)'(x)|>c$, then a.s. it does not belong to $\mathcal{B}^H_{\infty,\infty}([0,T])$.*

• With the point of view of the estimation of the local Hölder index of a process (in our case, this is equal to the parameter H), Theorem 1 remains true in a slightly more general setting. Consider the model

$$Y_t = y_0 + \int_0^t \sigma_s \, \mathrm{d}B_s, \qquad y_0 \in \mathbb{R},\ t \in [0,T], \tag{5}$$

with $\sigma_t = \Phi(Z_t)$. Here Φ verifies Assumption B and $(Z_t, t \in [0,T])$ is a continuous time process such that for all $(s,t) \in [0,1]^2$, $s \le t$,

$$Z_t - Z_s = a(s)(Z'_t - Z'_s) + (t-s)f(s) + h(t,s) + [v(t)-v(s)]\xi_0$$

where

- $a \in \mathcal{C}^1([0,T])$ and there exist $0 \le \beta_1 < \beta_2 \le T$ such that $\inf_{u\in[\beta_1,\beta_2]} a^2(u) > 0$ and $\mathbb{P}[\forall u \in [\beta_1,\beta_2], Z_u^2 = 0] = 0$.
- $(Z'_t, t \in [0,T])$ is a centered Gaussian process, independent of B, such that $Z'_0 = 0$ and for all $t \ge 0$ and $h > 0$,
$$\mathbb{E}[(Z'_{t+h} - Z'_t)^2] = \mathbb{E}[Z'^2_h] \text{ and } \mathbb{E}[Z'^2_h] = h^{2H}(1 + g(h)h^{1/2}),$$
with $H \in (1/2, 1)$ and $g \in \mathcal{C}^4([0,T])$.
- $f : [0,T] \to \mathbb{R}$ is a random function such that for $p > 0$,
$$\sup_{s\in[0,T]} \mathbb{E}[|f(s)|^p] < \infty.$$
- $h : [0,T]^2 \to \mathbb{R}$ is a random function such that for $p > 0$ and $(t,s) \in [0,T]^2$,
$$\mathbb{E}[|h(t,s)|^p] \le c_p (t-s)^{3p/2} \quad \text{and} \quad \mathbb{E}\left[\sup_{t\in[0,T]} \mathrm{e}^{p|h(t,0)|}\right] < \infty.$$
- $v \in \mathcal{C}^2([0,T])$.
- ξ_0 is a random variable, independent of B, such that for $p > 0$, $\mathbb{E}[\mathrm{e}^{p|\xi_0|}] < \infty$.

This general setting includes various Gaussian processes with stationary increments and local Hölder index equal to H; see for example Istas and Lang [23]. The following proposition enables us to work in the general setting of model (5) for the proof Theorem 1.

Theorem 3 (*General Formalism for Theorem* 1).
(i) *The formalism of model* (5) *includes model* (1) *and* (2).
(ii) *In model* (5), *Theorem* 1 *holds for the estimation of the parameter* H.

Hence, we only prove Theorems 2 and 3.
• The accuracy $v_n(H)$ is slower by a polynomial order than the usual $n^{-1/2}$ of regular parametric models. This rate of convergence seems to be characteristic of high frequency parametric inference from noisy data in the presence of fractional Brownian motion. Indeed, this rate is also found by Gloter and Hoffmann [16] in the high frequency inference of the finite dimensional parameter θ in the model

$$\mathrm{d}Y_t = \sigma_t \,\mathrm{d}B_t, \quad \sigma_t = \Phi(\theta, W_t^H) \tag{6}$$

and in the high frequency estimation of the Hurst parameter in the model

$$Y_i^n = \sigma W_{i/n}^H + a(W_{i/n}^H)\xi_i^n, \tag{7}$$

where a is an unknown function and ξ_i^n a centered white noise; see [17]. In a sense, our approach is a generalization of both (6) and (7) for the estimation of the parameter H.
• In practice, a usual way to estimate the regularity, or the long memory parameter, of the volatility of an asset is to build a volatility proxy[2] from the prices, and then to use classical method for regularity estimation or long memory detection. Although linked with the preceding practice, the method we give in this paper is mathematically rigorous, and in some sense optimal. The optimal rates of convergence are quite slow, but not catastrophic. Hence, our result shows that getting accurate enough information about the smoothness of the volatility process is possible, but compulsorily requires a large amount of data. This is not surprising. Indeed, the volatility is not observed and any pointwise approximation of it is very noisy. For an illustration of this, see the numerical results in the Appendix.

1.5. Organization of the paper

In Section 2, we present our estimation method for the volatility Hurst parameter. Section 3 states the main propositions which lead to Theorems 3 and 2. The proof of Theorem 3(i) is given in Section 4. We prove in Sections 5–7 the results stated in Section 3 concerning the upper bound whereas Theorem 3(ii) is proved in Section 8. We end with the proof of Theorem 2 in Section 9. The proof of Proposition 1 and one numerical illustration are given in the Appendix.

2. Estimation strategy

2.1. Estimation of the Hurst parameter: Preliminaries

2.1.1. Estimation of H from direct observation of a fractional Brownian motion

Imagine that we observe high frequency data

$$\{\sigma W_{i/n}^H, i = 0, \ldots, n\},$$

[2] Such proxies are often based on the absolute or quadratic variation of the log prices, with sampling period higher than 10 min to avoid microstructure noise effects.

where σ is an unknown constant and W_t^H a fractional Brownian motion. Then, we can recover the Hurst parameter at the parametric rate $n^{-1/2}$. Indeed, we can use as follows local properties of the trajectory of the fractional Brownian motion; see Istas and Lang [23], and see also Berzin and Leon [5], Lang and Roueff [25]. Let $s = (s_0, \dots, s_p) \in \mathbb{R}^{p+1}$ be such that

$$\text{for } k = 0, \dots, p-1 : \sum_{i=0}^{p} s_i i^k = 0 \quad \text{and} \quad \sum_{i=0}^{p} s_i i^p \neq 0.$$

The integer $p = m(s)$ is called the order of the difference. For instance, the usual difference $s = (-1, 1)$ is of order 1 and $s = (1, -2, 1)$ is of order 2. For such a sequence s and $i = 0, \dots, n - m(s) - 1$, we define for a function $f : [0, 1] \to \mathbb{R}$, the generalized difference

$$\Delta_{i,n} f = \sum_{j=0}^{m(s)} s_j f\left(\frac{i+j}{n}\right).$$

Consider

$$V_n(H) = \sum_{i=0}^{n-m(s)-1} (\Delta_{i,n} W^H)^2.$$

Istas and Lang [23] show that for $m(s) > 1$, there exists a constant $L_{s,H} > 0$ such that[3]

$$n^{2H-1} V_n(H) = L_{s,H} + \frac{1}{\sqrt{n}} \xi_n,$$

with ξ_n bounded in probability. Then, an estimator achieving the rate $n^{-1/2}$ is for example[4]

$$\widehat{H} = \frac{1}{2}\left(1 + \log_2 \frac{V_{\lfloor n/2 \rfloor}(H)}{V_n(H)}\right).$$

Note that beyond fractional Brownian motion, the problem of estimating the local Hölder index of a process has been heavily studied in the Gaussian context; see in particular Istas and Lang [23] and Lang and Roueff [25].

2.1.2. *Estimation of H from noisy observation of a fractional Brownian motion*

Consider now model (7). Recovering the Hurst parameter in this context of noisy data is more difficult. Indeed, Gloter and Hoffmann show in [17] that the statistical structure of model (7) is significantly modified by the noise. They prove that the rate $n^{-1/(4H+2)}$ is optimal for estimating H in the minimax sense of (3) and (4). Their estimation strategy is based on the behavior of the energy levels of the fractional Brownian motion that reflects the Besov properties of the trajectories. Pick a mother wavelet ψ with 2 vanishing moments. Hence, the wavelet support has a minimal length of 3; see Daubechies [13]. For j and k positive integers, let

$$\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k),\ d_{jk} = \int \psi_{jk} W_s^H \mathrm{d}s \quad \text{and} \quad Q_j = \sum_k d_{jk}^2.$$

[3] The condition $m(s) > 1$ is necessary for $H > \frac{3}{4}$; if $H \leq \frac{3}{4}$, one can take $m(s) = 1$.

[4] Note that if σ is known, an estimator with accuracy $n^{-1/2}(\log n)^{-1}$ can be built; see Cœurjolly [9].

The sequence of energy levels $(Q_j, j \geq 0)$ has the following scaling property[5]:

$$\frac{Q_{j+1}}{Q_j} = 2^{-2H} + o(1) \quad \text{as } j \to +\infty. \tag{8}$$

Gloter and Hoffmann [17] construct estimators $\widehat{d^2_{jk}}$ of the d^2_{jk} up to a maximal resolution level $J_n = \lfloor \frac{1}{2} \log_2(n) \rfloor$. Setting

$$\widehat{Q}_j = \sum_k \widehat{d^2_{jk}}, \tag{9}$$

one obtains a sequence of estimators:

$$\widehat{H}_{j,n} = -\frac{1}{2} \log_2 \frac{\widehat{Q}_{j+1,n}}{\widehat{Q}_{j,n}}, \quad j = 1, \ldots, J_n. \tag{10}$$

The final estimator is $\widehat{H}_{J_n^*,n}$ where the optimal resolution level J_n^* is defined following the rules of adaptive estimation of quadratic functionals,

$$J_n^* = \max\left\{ j = 1, \ldots, J_n, \widehat{Q}_{j,n} \geq \frac{2^j}{n} \right\}. \tag{11}$$

We adapt the preceding strategy in this paper.

2.2. *Construction of an estimator*

We build in this section our estimator in the general setting of model (5).

2.2.1. *An Euler scheme-type transformation*

By an Euler scheme-type transformation, we boil the problem down to a regression model. Indeed, we have

$$z_i^n = n(Y_{(i+1)/n} - Y_{i/n})^2 = \sigma^2_{i/n} + \xi_i^n, \tag{12}$$

with

$$\xi_i^n = n\left[\int_{\frac{i}{n}}^{\frac{i+1}{n}} (\sigma_t^2 - \sigma^2_{i/n}) \mathrm{d}t + \left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma_t \, \mathrm{d}B_t \right)^2 - \int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma_t^2 \mathrm{d}t \right].$$

Conditional on the fbm W^H and up to negligible terms, the ξ_i^n are martingale increments with variance of order 1.

2.2.2. *Estimation of the energy levels*

Let ψ be a mother wavelet with two vanishing moments and support $[0, T]$. Let

$$d_{jk} = \int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \sigma_t^2 \psi_{jk}(t) \mathrm{d}t \quad \text{and} \quad Q_j = \sum_k d^2_{jk}.$$

[5] For the moment, we do not specify the meaning of $o(\cdot)$.

By proving a scaling-type property on the energy levels analogous to (8), we can follow the strategy of Section 2.1.2. The main difficulty lies here in the non-linearity introduced by the function Φ^2. We now present the estimation of the energy levels. To get rid of boundary effects, without any loss of generality in our asymptotic framework, we do not take into account the wavelets ψ_{jk} whose support is not totally included in $[0, T]$. We have

$$d_{jk} = \sum_{l=0}^{T2^{N-j}-1} \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t^2 \psi_{jk}(t)\mathrm{d}t.$$

A first natural estimator of d_{jk} is

$$\widetilde{d_{jk}} = \sum_{l=0}^{T2^{N-j}-1} z^n_{k2^{N-j}+l} \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t.$$

Let

$$M_{k,l,t} = \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{t} \sigma_u \mathrm{d}B_u\right)^2 - \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{t} \sigma_u^2 \mathrm{d}u.$$

From (12), we have the following decomposition:

$$\widetilde{d_{jk}} - d_{jk} = b_{jk} + e_{jk} + f_{jk},$$

with

$$b_{jk} = \sum_{l=0}^{T2^{N-j}-1} \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)(\sigma^2_{k2^{-j}+l2^{-N}} - \sigma_t^2)\mathrm{d}t,$$

$$e_{jk} = n \sum_{l=0}^{T2^{N-j}-1} \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\, M_{k,l,\frac{k}{2^j}+\frac{l+1}{2^N}},$$

$$f_{jk} = n \sum_{l=0}^{T2^{N-j}-1} \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} (\sigma_t^2 - \sigma^2_{k2^{-j}+l2^{-N}})\mathrm{d}t.$$

In order to estimate d_{jk}^2 accurately enough, we cannot use $\widetilde{d_{jk}}^2$ because the remaining term e_{jk}^2 has to be compensated. The other terms are negligible.

Conditional on W^H, $(M_{k,l,t}, t \geq 0)$ is a continuous local martingale. Let $\tilde{\mathbb{E}}$ denote the expectation conditional on the path of the volatility. Then, by the independence of the Brownian increments,

$$\tilde{\mathbb{E}}[e_{jk}^2] = n^2 \sum_{l=0}^{T2^{N-j}-1} \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\right)^2 \tilde{\mathbb{E}}[M^2_{k,l,\frac{k}{2^j}+\frac{l+1}{2^N}}].$$

Let

$$N_{k,l,t} = \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{t} \sigma_u \mathrm{d}B_u.$$

By Ito's formula,

$$M_{k,l,t} = 2\int_0^t \sigma_u N_u \mathbb{1}_{\{u \geq \frac{k}{2^j} + \frac{l}{2^N}\}} \mathrm{d}B_u.$$

Let

$$a_{j,k,l}^2 = \tilde{\mathbb{E}}[M^2_{k,l,\frac{k}{2^j}+\frac{l+1}{2^N}}] = 2\left(\tilde{\mathbb{E}}\left[(Y_{k2^{-j}+(l+1)2^{-N}} - Y_{k2^{-j}+l2^{-N}})^2\right]\right)^2.$$

We need to compensate $a_{j,k,l}^2$, so we estimate it by

$$\widehat{a_{j,k,l}^2} = \left(\frac{\sqrt{2}}{h(n)}\sum_{p=0}^{h(n)}(Y_{k2^{-j}+(l+1+p)2^{-N}} - Y_{k2^{-j}+(l+p)2^{-N}})^2\right)^2,$$

where $h(n) = \lfloor n^{1/2} \rfloor$. Let

$$\nu_{jk} = n^2 \sum_{l=0}^{T2^{N-j}-1}\left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\right)^2 a_{j,k,l}^2,$$

$$\bar{\nu}_{jk} = n^2 \sum_{l=0}^{T2^{N-j}-1}\left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\right)^2 \widehat{a_{j,k,l}^2}.$$

Finally we define

$$\widehat{d_{jk}^2} = \widetilde{d_{jk}}^2 - \bar{\nu}_{jk} \quad \text{and} \quad \widehat{Q}_j = \sum_k \widehat{d_{jk}^2}.$$

We thus obtain our estimator $\widehat{H}_{J_n^*,n}$ of H with the same specifications as in (10) and (11).

3. The behavior of the energy levels

We present here the steps that enable us to prove Theorem 3(ii) and Theorem 2.

3.1. Upper bound

We work in the general setting of model (5). Let

$$d_{jk} = \int \sigma_t^2 \psi_{jk}(t)\mathrm{d}t \quad \text{and} \quad Q_j = \sum_k d_{jk}^2.$$

We write c for a constant depending on Φ, a, f, v, H, ψ and continuous in its arguments.

Proposition 2 (*Limit of the Energy Levels*). *In model (5), there exists a constant $c(\psi) > 0$, depending on ψ and H, continuous in its arguments, and where $c > 0$ such that*

$$\mathbb{E}\left[\left|2^{2jH}Q_j - c(\psi)\int_0^T a(u)^2\{(\Phi^2)'(Z_u)\}^2\mathrm{d}u\right|\right] \leq c2^{-j/2}.$$

More precisely, Proposition 2 enables us to obtain the following result.

Proposition 3 (*Scaling Property*). *In model (5), we have*

(i) *for all* $\varepsilon > 0$, *there exist* j_0 *and* $r > 0$ *such that for all* $j \geq j_0$,

$$\mathbb{P}[2^{2jH} Q_j \geq r] \geq 1 - \varepsilon,$$

(ii) *for all* $\varepsilon > 0$, *there exist* j_0 *and* $M > 0$ *such that for all* $j \geq j_0$,

$$\mathbb{P}\left[2^{j/2} \sup_{l \geq j} \left| \frac{Q_{l+1}}{Q_l} - 2^{-2H} \right| \geq M \right] \leq \varepsilon.$$

Finally, we have the following result for the estimator.

Proposition 4 (*Deviation of the Estimator*). *Let* $j_n(H) = \lfloor (2H+1)^{-1} \log_2(n) \rfloor$ *and* $\mathcal{H}$ *be a compact set included in* $(1/2, 1)$. *In model (5), for all* $H \in \mathcal{H}$, $J_n \geq j_n(H)$ *and for any* $L > 0$, *the sequence*

$$\left\{ n 2^{j_n(H)/2} \sup_{J_n \geq j \geq j_n(H) - L} 2^{-j} |\widehat{Q}_{j,n} - Q_j| \right\}$$

is bounded in probability, uniformly over $\mathcal{H}$.

We then prove in Section 8 that Propositions 3 and 4 together imply Theorem 3(i).

3.2. Lower bound

For the lower bound, we work in model (1) and (2). Let $\mathbb{P}_f^n$ denote the law of the data $Y^n = \{Y_{i/n}, i = 0, \ldots, nT\}$ conditional on $W^H = f$. The key point of the lower bound is the following.

Proposition 5 (*Distance in Total Variation*). *Under Assumptions A–C, there exists* $c > 0$ *such that*

$$\|\mathbb{P}_f^n - \mathbb{P}_g^n\|_{TV}^2 \leq cn \|f - g\|_2^2,$$

where $\|\cdot\|_{TV}$ *denotes the distance in total variation and* $\|\cdot\|_2$ *the usual* L^2 *norm of functions on* $[0, T]$ *with respect to the Lebesgue measure.*

Proposition 5 together with Proposition 5 of Gloter and Hoffmann [17] implies the lower bound.

4. Proof of Theorem 3(i)

We show here than we can prove Theorem 1 under the general formalism of model (5).

4.1. Notation

In all the proofs, we repeatedly use the notation c for constants depending on H, ψ and the functions defined in model (1) and (2) or model (5), continuous in their arguments, and that may vary from line to line. We write the symbol $=$ also for almost sure equality and for a function g, we set $\|g\|_\infty = \sup_t |g(t)|$. Finally, $\partial_i^j f(u,t)$ denotes the j-th derivative of f with respect its i-th variable.

4.2. Proof of Theorem 3(i)

For $s \le t$

$$\int_0^t a(t,u)\mathrm{d}W_u^H - \int_0^s a(s,u)\mathrm{d}W_u^H$$

is equal to $a(s,s)(W_t^H - W_s^H) + (t-s)f(s) + h(t,s)$ with

$$f(s) = \int_0^s \partial_1 a(s,u)\mathrm{d}W_u^H,$$

and

$$h(t,s) = \frac{(t-s)^2}{2}\int_0^s \partial_1^2 a(\theta_1[t,s],u)\mathrm{d}W_u^H + R(t,s),$$

where

$$\begin{aligned} R(t,s) &= \int_s^t [a(t,u)-a(s,u)]\mathrm{d}W_u^H + \int_s^t [a(s,u)-a(s,s)]\mathrm{d}W_u^H \\ &= (t-s)\int_s^t \partial_1 a(\theta_2[t,s],u)\mathrm{d}W_u^H + \int_s^t (u-s)\partial_2 a(s,\theta_3[u,s])\mathrm{d}W_u^H. \end{aligned}$$

Here θ_1, θ_2 and θ_3 are deterministic functions with values in $[0,T]$. Using Assumption A, we get that all the preceding integrands are deterministic, continuously differentiable with respect to the variable u and uniformly bounded with respect to all variables. Hence, in our case, the Riemann–Stieltjes integral with respect to the fbm coincides almost surely with the Wiener integral with respect to the fbm. Consequently, for $f \in \mathcal{C}^1([0,T])$ and $g \in \mathcal{C}^1([0,T])$,

$$\mathbb{E}\left[\int_0^T f(u)\mathrm{d}W_u^H \int_0^T g(u)\mathrm{d}W_u^H\right] = H(2H-1)\int_0^T\int_0^T f(s)g(t)|s-t|^{2H-2}\mathrm{d}s\mathrm{d}t; \quad (13)$$

see for example Norros et al. [30]. Hence, using the fact that the preceding stochastic integrals are Gaussian variables together with Assumption A, we easily get that for $(t,s) \in [0,T]^2$, $s \le t$ and $p > 0$

$$\mathbb{E}\left[\left|\frac{(t-s)^2}{2}\int_0^s \partial_1^2 a(\theta_1[t,s],u)\mathrm{d}W_u^H + R(t,s)\right|^p\right] \le c_p(t-s)^{p(1+H)}.$$

For $p > 0$ and $t > 0$, let

$$V_t = \int_0^t [a(t,u)-a(0,0)]\,\mathrm{d}W_u^H, \quad \tilde{V} = \sup_{t\in[0,T]} |pV_t| \text{ and } \nu = \sup_{t\in[0,T]} \mathbb{E}[(pV_t)^2].$$

We now prove that $\mathbb{E}[\mathrm{e}^{\tilde{V}}] < \infty$. The process $(pV_t, t \ge 0)$ is a Gaussian process starting from 0 with continuous trajectories, so we can use Dudley's entropy bound. There exists a universal constant c such that

$$\mathbb{E}[\tilde{V}] \le c\int_0^{d(0,T)} \sqrt{\log N(T,d,\varepsilon)}\mathrm{d}\varepsilon,$$

where $d^2(s,t) = p^2\mathbb{E}[|V_t - V_s|^2]$ and $N(T,d,\varepsilon)$ is the minimal number of balls of radius ε needed to recover $[0,T]$. Since

$$\mathbb{E}[|V_t - V_s|^2] \leq c|t-s|^{2H},$$

we easily obtain that $N(T, d, \varepsilon)$ is less than $cT\varepsilon^{-1/H}$. Hence, we get $\mathbb{E}[\tilde{V}] < \infty$. We now use Borell's inequality: For $\lambda > \mathbb{E}[\tilde{V}]$,

$$\mathbb{P}[\tilde{V} \geq \lambda] \leq 2\mathrm{e}^{-\frac{1}{2}(\lambda - \mathbb{E}[\tilde{V}])^2/\nu}.$$

As

$$\mathbb{E}[\mathrm{e}^{\tilde{V}}] = \int_0^{+\infty} \mathbb{P}[\mathrm{e}^{\tilde{V}} \geq \lambda]\mathrm{d}\lambda,$$

we get

$$\mathbb{E}[\mathrm{e}^{\tilde{V}}] \leq c + 2\int_{\mathrm{e}^{\mathbb{E}[\tilde{V}]}}^{+\infty} \mathrm{e}^{-\frac{1}{2}(\log\lambda - \mathbb{E}[\tilde{V}])^2/\nu}\mathrm{d}\lambda \leq c + 2\int_0^{+\infty} \mathrm{e}^{\mathbb{E}[\tilde{V}]+u-u^2/2\nu}\mathrm{d}u.$$

Finally, suppose that on $[\beta_1, \beta_2]$,

$$t \to \int_0^t a(t,u)\mathrm{d}W_u^H + f(t)\xi_0$$

is equal to zero with positive probability. This implies that $t \to \int_0^t a(t,u)\mathrm{d}W_u^H$ belongs to $\mathcal{C}^1([\beta_1, \beta_2])$ with positive probability. This is absurd; see the proof of Proposition 1 in the Appendix.

5. Proof of Proposition 2

From now, and until the end of the proof of Theorem 3(ii), we work in model (5).

5.1. Technical lemmas

We establish here several useful lemmas. We apply here ideas of Gloter and Hoffmann [16], initially developed for generalized differences. We first prove two lemmas on the expectation and covariance of the wavelet coefficients for the stochastic integral. Let

$$\beta_{jk} = \int_0^T Z_t\psi_{jk}(t)\mathrm{d}t, \quad \beta'_{jk} = \int_0^T Z'_t\psi_{jk}(t)\mathrm{d}t, \quad F(t) = \int_0^t \psi(u)\mathrm{d}u.$$

We have the following lemma.

Lemma 1. *For all positive integers j, k,*

$$\beta_{jk} = a(k2^{-j})\beta'_{jk} + 2^{-2j}R_{jk},$$

and

$$\mathbb{E}[\beta_{jk}^{\prime 2}] = 2^{-j(1+2H)}\{c(\psi) + 2^{-j/2}R'_{jk}\},$$

with

$$c(\psi) = \mathbb{E}\left[\left(\int_0^T F(t)\mathrm{d}W_t^H\right)^2\right] > 0,$$

where $(W_t^H, t \geq 0)$ is a fractional Brownian motion and $\mathbb{E}[|R_{jk}|^p + |R'_{jk}|^p] \leq c_p$ for $p > 0$.

Proof. The coefficient β_{jk} is equal to

$$2^{-j/2}\int_0^T \psi(v)Z_{(k+v)2^{-j}}\mathrm{d}v.$$

The two vanishing moments of the wavelet easily give the first assertion of the lemma.

Using that

$$2Z'_{(k+u)2^{-j}}Z'_{(k+v)2^{-j}} = Z'^2_{(k+u)2^{-j}} + Z'^2_{(k+v)2^{-j}} - (Z'_{(k+u)2^{-j}} - Z'_{(k+v)2^{-j}})^2, \tag{14}$$

together with the vanishing moments of the wavelet, we get that $\mathbb{E}[\beta'^2_{jk}]$ is equal to

$$-2^{-(j+1)}\int_0^T\int_0^T \psi(u)\psi(v)2^{-j2H}|u-v|^{2H}\mathrm{d}u\mathrm{d}v$$
$$-2^{-(j+1)}\int_0^T\int_0^T \psi(u)\psi(v)2^{-j(2H+1/2)}|u-v|^{2H+1/2}g(|u-v|2^{-j}).$$

Hence,

$$\mathbb{E}[\beta'^2_{jk}] = -2^{-(j+1)}\int_0^T\int_0^T \psi(u)\psi(v)2^{-j2H}|u-v|^{2H}\mathrm{d}u\mathrm{d}v + 2^{-j(3/2+2H)}R'_{jk},$$

with $|R'_{jk}| \le c$. Then we easily show that

$$-\int_0^T\int_0^T \psi(u)\psi(v)|u-v|^{2H}\mathrm{d}u\mathrm{d}v$$
$$= 2H(2H-1)\int_0^T\int_0^T F(u)F(v)|u-v|^{2H-2}\mathrm{d}u\mathrm{d}v.$$

We conclude using (13). $\square$

Lemma 2 (*Decorrelation of the Wavelet Coefficients*). *There exists c such that, for all j, k, k',*

$$|\mathbb{E}[\beta'_{jk}\beta'_{jk'}]| \le 2^{-j(1+2H)}c(1+|k-k'|)^{2H-4}.$$

Proof. For $k \ge k'+T+1$, let $m_{k,k',u,v} = 2^{-j}(k-k'+u-v)$. We have

$$\begin{aligned}\mathbb{E}[\beta'_{jk}\beta'_{jk'}] &= 2^{-j}\int_0^T\int_0^T \psi(u)\psi(v)\mathbb{E}[Z'_{(k+u)2^{-j}}Z'_{(k'+v)2^{-j}}]\mathrm{d}u\mathrm{d}v\\ &= -2^{-(j+1)}\int_0^T\int_0^T \psi(u)\psi(v)\mathbb{E}[Z'^2_{(k-k'+u-v)2^{-j}}]\mathrm{d}u\mathrm{d}v\\ &= -2^{-(j+1)}\int_0^T\int_0^T \psi(u)\psi(v)|m_{k,k',u,v}|^{2H}\\ &\quad -2^{-(j+1)}\int_0^T\int_0^T \psi(u)\psi(v)g(m_{k,k',u,v})|m_{k,k',u,v}|^{2H+1/2}\mathrm{d}u\mathrm{d}v.\end{aligned}$$

For the first term, we make a fourth-order Taylor expansion of $x \to (x+u-v)^{2H}$ around point $k-k'$. Thanks to the two vanishing moments of the wavelet, we get that the first term is less than $c2^{-j(1+2H)}(|k-k'|-T)^{2H-4}$. For the second term, we first make a fourth-order Taylor expansion of g around point $(k-k')2^{-j}$. The result follows thanks to expansions of $x \to (x+u-v)^{2H+1/2}$ up to order 4, 3, 2 and 1. $\square$

Lemma 3. *Let $\xi : [0, T] \to \mathbb{R}$ be a deterministic bounded function. Define*

$$\Sigma_j(\xi) = 2^j \sum_{k=0}^{T(2^j-1)} \{2^{j2H}\beta_{jk}^2 - c(\psi)2^{-j}a(k2^{-j})^2\}\xi_{k2^{-j}}. \tag{15}$$

Then,

$$\mathbb{E}[\Sigma_j(\xi)^2] \le c\|\xi\|_\infty^2 2^j.$$

Proof. We have

$$\begin{aligned}\Sigma_j(\xi) &= 2^j \sum_{k=0}^{T(2^j-1)} 2^{j2H}a(k2^{-j})^2(\beta_{jk}^{\prime 2} - \mathbb{E}[\beta_{jk}^{\prime 2}])\xi_{k2^{-j}} \\ &\quad + 2^j \sum_{k=0}^{T(2^j-1)} 2^{j2H}(2^{-4j}R_{jk} + 2^{-2j}2^{-j(1/2+H)}R'_{jk} + 2^{-j(3/2-2H)}R''_{jk})\xi_{k2^{-j}},\end{aligned}$$

with $\mathbb{E}[|R_{jk}|^p + |R'_{jk}|^p + |R''_{jk}|^p] \le c$, for $p > 0$. Hence $\mathbb{E}[\Sigma_j(\xi)^2]$ is less than

$$c2^j\|\xi\|_\infty^2 + c2^{2j}\mathbb{E}\left[\sum_{k,k'=0}^{T(2^j-1)} 2^{j4H}\{\beta_{jk}^{\prime 2} - \mathbb{E}[\beta_{jk}^{\prime 2}]\}\{\beta_{jk'}^{\prime 2} - \mathbb{E}[\beta_{jk'}^{\prime 2}]\}\xi_{k2^{-j}}\xi_{k'2^{-j}}\right].$$

Let $Y_k = \beta_{jk}^{\prime 2}/\mathbb{E}[\beta_{jk}^{\prime 2}] - 1$. The preceding inequality can be written as

$$\mathbb{E}[\Sigma_j(\xi)^2] \le c2^j\|\xi\|_\infty^2 + c2^{2j}2^{j4H} \sum_{k,k'=0}^{T(2^j-1)} \mathbb{E}[Y_kY_{k'}]\mathbb{E}[\beta_{jk}^{\prime 2}]\mathbb{E}[\beta_{jk'}^{\prime 2}]\xi_{k2^{-j}}\xi_{k'2^{-j}}.$$

We now apply Mehler's formula and we get

$$\begin{aligned}\mathbb{E}[\Sigma_j(\xi)^2] &\le c2^j\|\xi\|_\infty^2 + 2^{2j}2^{j4H}\|\xi\|_\infty^2 2\sum_{k,k'=0}^{T(2^j-1)} \mathrm{Cov}(\beta'_{jk}, \beta'_{jk'})^2 \\ &\le c2^j\|\xi\|_\infty^2 + c2^{2j}2^{j4H}\|\xi\|_\infty^2 \sum_{k,k'=0}^{T(2^j-1)} 2^{-2j(1+2H)}(1+|k-k'|)^{4(H-2)} \\ &\le c2^j\|\xi\|_\infty^2 + c\|\xi\|_\infty^2 \sum_{k=0}^{T(2^j-1)}\sum_{i=0}^{+\infty}(1+i)^{4(H-2)} \le c2^j\|\xi\|_\infty^2. \quad \square\end{aligned}$$

Lemma 4. *Assume that $\xi : [0, T] \to \mathbb{R}$ is bounded and vanishes outside the interval $[k2^{-j_0}, k'2^{-j_0}] \subset [0, T]$ for some $k, k', j_0 \ge 1$, $k \ne k'$. Then, there exists $c > 0$ such that for $j \ge j_0$,*

$$\mathbb{E}[\Sigma_j(\xi)^2] \le c\|\xi\|_\infty^2|k' - k|2^{j-j_0}.$$

Proof. As $\xi_{z2^{-j}}$ is different from zero only if $k2^{-j_0} \leq z2^{-j} \leq k'2^{-j_0}$, there are less than $|k-k'|2^{j-j_0}+1$ admissible values for z. Hence we easily get that $\mathbb{E}[\Sigma_j(\xi)^2]$ is less than

$$c|k'-k|2^{j-j_0}|k'-k|2^{-j_0}\|\xi\|_\infty^2 + c\sum_{z,z'}^{T(2^j-1)} |\xi_{z2^{-j}}||\xi_{z'2^{-j}}|(1+|z-z'|)^{4(H-2)}.$$

By similar computations on the series as in the proof of Lemma 3, we get

$$\mathbb{E}[\Sigma_j(\xi)^2] \leq c|k'-k|2^{j-j_0}\|\xi\|_\infty^2 + c\|\xi\|_\infty \sum_z |\xi_{z2^{-j}}|.$$

The result follows. □

We now decompose the function $t \to (\{\Phi^2\}')^2(Z_t)$ in a wavelet basis with support $[0,T]$. Thus, we use the same wavelet as before but in another context. We have the following lemma.

Lemma 5 (*Decomposition in a Wavelet Basis*). *Let $\Gamma = (\{\Phi^2\}')^2$. Let ϕ be the scaling function associated with ψ. We write $\phi_{0k}(t) = \phi(t-k)$,*

$$c_k = \int \Gamma(Z_t)\phi_{0k}(t)\mathrm{d}t \quad \text{and} \quad c_{jk} = \int \Gamma(Z_t)\psi_{jk}(t)\mathrm{d}t.$$

Then,

$$\Gamma(Z_t) = \sum_{k=0}^{r} c_k\phi_{0k}(t) + \sum_{j=0}^{+\infty}\sum_{k=0}^{T(2^j-1)} c_{jk}\psi_{jk}(t),$$

where r is a constant value depending on T and with

$$\mathbb{E}[c_0+\cdots+c_r] \leq c, \quad \mathbb{E}[c_{jk}^2] \leq c2^{-j(1+2H)}.$$

Proof. We have

$$\begin{aligned} c_{jk} &= 2^{-j/2}\int_0^T \psi(u)\Gamma(Z_{2^{-j}(k+u)})\mathrm{d}u \\ &= 2^{-j/2}\int_0^T \psi(u)[\Gamma(Z_{2^{-j}(k+u)}) - \Gamma(Z_{2^{-j}k})]\mathrm{d}u \\ &= 2^{-j/2}\int_0^T \psi(u)[Z_{2^{-j}(k+u)} - Z_{2^{-j}k}]\Gamma'(\eta)\mathrm{d}u, \end{aligned}$$

with η a random value between $Z_{2^{-j}k}$ and $Z_{2^{-j}(k+u)}$. By the continuity of the sample path of $t \to Z_t$, we know there exists a random point θ between $k2^{-j}$ and $(k+u)2^{-j}$ such that $\eta = Z_\theta$. Thus, we have

$$c_{jk}^2 \leq c2^{-j}\int_0^T \psi^2(u)[Z_{2^{-j}(k+u)} - Z_{2^{-j}k}]^2\{\Gamma'(r[\theta])\}^2\mathrm{d}u,$$

with

$$r(\theta) = a(0)Z'_\theta + \theta f(0) + h(\theta,0) + [v(\theta)-v(0)]\xi_0.$$

As Z' is a Gaussian process, we easily get

$$\mathbb{E}[(Z_{2^{-j}(k+u)} - Z_{2^{-j}k})^4] \leq c2^{-j4H}.$$

Using Assumption B, we have

$$\mathbb{E}[c_{jk}^2] \leq c2^{-j(1+2H)} \int_0^T \psi(u)^2 \{\mathbb{E}[\mathrm{e}^{c|r(\theta)|}]\}^{1/2} \mathrm{d}u.$$

Let $p > 0$ and $\tilde{Z} = \sup_{t\in[0,T]} |pZ_t'|$. By the same arguments as in the proof of Theorem 3(i), we prove that $\mathbb{E}[\mathrm{e}^{\tilde{Z}}] < \infty$. Hence, using the hypothesis of model (5) and Cauchy–Schwarz inequality, we get $\mathbb{E}[c_{jk}^2] \leq c2^{-j(1+2H)}$. By a Taylor expansion, we get $\mathbb{E}[c_0 + \cdots + c_r] \leq c$. □

Lemma 6. *Let Γ be as in Lemma 5. We have*

$$\mathbb{E}\left[\left|2^j \sum_k \{2^{j2H} \beta_{jk}^2 - c(\psi)2^{-j} a(k2^{-j})^2\} \Gamma(Z_{k2^{-j}})\right|\right] \leq c2^{j/2}.$$

Proof. We know from Lemma 5 that

$$\Gamma(Z_t) = \sum_{k=0}^{r} c_k \phi_{0k}(t) + \sum_{j=0}^{+\infty} \sum_{k=0}^{T(2^j-1)} c_{jk} \psi_{jk}(t).$$

Let

$$S_j(\Gamma) = 2^j \sum_k \{2^{j2H} \beta_{jk}^2 - c(\psi)2^{-j} a(k2^{-j})^2\} \Gamma(Z_t).$$

We can write

$$S_j(\Gamma) = \sum_{k=0}^{r} c_k \Sigma_j(\phi_{0k}) + \sum_{j_1=0}^{+\infty} S_{j,j_1},$$

with

$$S_{j,j_1} = \sum_{k=0}^{T(2^j-1)} c_{j_1k} \Sigma_j(\psi_{j_1k}).$$

For $k = 0$ to r, $\mathbb{E}[|c_i \Sigma_j(\phi_{0i})|] \leq c2^{j/2}$, by Lemma 3. Now we prove that

$$\mathbb{E}\left[\sum_{j_1=0}^{+\infty} |S_{j,j_1}|\right] \leq c2^{j/2}.$$

If $j_1 \leq j$, by Lemma 4,

$$\mathbb{E}[|S_{j,j_1}|] \leq c \sum_{k=0}^{T(2^{j_1}-1)} 2^{-j_1(1+2H)/2} (\mathbb{E}[\Sigma_j(\psi_{j_1k})^2])^{1/2} \leq c2^{j_1(1/2-H)} 2^{j/2}.$$

Because $H > 1/2$, we have

$$\sum_{j_1=0}^{j} \mathbb{E}[|S_{j,j_1}|] \leq c2^{j/2}.$$

If $j < j_1$, ψ_{j_1k} has support $[k2^{-j_1}, (k+T)2^{-j_1}]$, so $\Sigma_j(\psi_{j_1k}) = 0$ unless there exists $i \in [0, T(2^j - 1)]$ such that $i2^{-j} \in [k2^{-j_1}, (k+T)2^{-j_1}]$, that is

$$k2^{j-j_1} \leq i \leq (k+T)2^{j-j_1}.$$

Thus, there are less than $c2^j$ possible values for i and moreover, for such i, the sum defining $\Sigma_j(\psi_{j_1k})$ is reduced to one single term, so, combining this result with Lemma 1, we get

$$\mathbb{E}[\Sigma_j(\psi_{j_1k})^2] \le c\|\psi_{j_1k}\|_\infty^2 \le c2^{j_1}$$

and

$$\mathbb{E}[|S_{j,j_1}|] \le c\sum_{k=0}^{T(2^j-1)} 2^{-j_1(1+2H)/2}2^{j_1/2} \le c2^j 2^{-j_1H}.$$

Finally

$$\sum_{j_1=0}^{+\infty} \mathbb{E}[|S_{j,j_1}|] = \sum_{j_1=0}^{j} \mathbb{E}[|S_{j,j_1}|] + \sum_{j_1=j+1}^{+\infty} \mathbb{E}[|S_{j,j_1}|] \le c2^{j/2}. \quad \square$$

Lemma 7 (*Riemann's Approximation*). *Let $H(x,t) = a(x)^2\Gamma(Z_t)$. We have*

$$\mathbb{E}\left[\left|\int_0^T H(t,t)\mathrm{d}t - 2^{-j}\sum_{k=1}^{2^jT} H(k2^{-j},k2^{-j})\right|\right] \le c2^{-j/2}.$$

Proof. We easily get that

$$\mathbb{E}\left[\left|\int_0^T H(t,t)\mathrm{d}t - 2^{-j}\sum_{k=1}^{2^jT} H(k2^{-j},k2^{-j})\right|\right]$$

is smaller than

$$\sum_{k=1}^{2^jT}\int_{(k-1)2^{-j}}^{k2^{-j}} a(t)^2\mathbb{E}[|(Z_t - Z_{k2^{-j}})\Gamma'(Z_{\theta[t,k2^{-j}]})|] + |a(k2^{-j})^2 - a(t)^2|\mathbb{E}[|\Gamma(Z_{k2^{-j}})|]\mathrm{d}t,$$

with $\theta[t,k2^{-j}]$ a random value between t and $k2^{-j}$. The same arguments as in the proof of Lemma 5 give that it is less than $c(2^{-j}+2^{-jH})$. $\square$

5.2. Proof of Proposition 2

Let

$$\tilde{\beta}_{jk} = \int \psi_{jk}(t)\Phi^2(Z_t)\mathrm{d}t.$$

Using the first vanishing moment of ψ, we have

$$\tilde{\beta}_{jk} = (\Phi^2)'(Z_{k2^{-j}})\beta_{jk} + 2^{-j/2}\int \psi(t)(Z_{(k+u)2^{-j}} - Z_{k2^{-j}})^2(\Phi^2)''(Z_{\theta[k2^{-j},(k+1)2^{-j}]}).$$

with $\theta[k2^{-j},(k+1)2^{-j}]$ a random value between $k2^{-j}$ and $(k+1)2^{-j}$. So, $\tilde{\beta}_{jk}^2$ is equal to

$$\Gamma(Z_{k2^{-j}})\beta_{jk}^2 + 2^{-j}2^{-4jH}X_{jk} + 2^{-j(1/2+H)}2^{-j/2}2^{-2jH}Y_{jk},$$

with $\mathbb{E}[|X_{jk}|^p + |Y_{jk}|^p] \leq c_p$, for $p > 0$. Hence

$$\sum_k \{2^{j2H}\tilde{\beta}_{jk}^2 - c(\psi)2^{-j}a(k2^{-j})^2\Gamma(Z_{k2^{-j}})\}$$
$$= \sum_k \{2^{j2H}\beta_{jk}^2 - c(\psi)2^{-j}a(k2^{-j})^2\}\Gamma(Z_{k2^{-j}}) + 2^{-j/2}W_{jk},$$

with $\mathbb{E}[|W_{jk}|^p] \leq c_p$, for $p > 0$. We finally get Proposition 2 by Lemmas 6 and 7.

6. Proof of Proposition 3

We begin by the proof of (i). With the notation of model (5), there exists $\eta > 0$ such that

$$c(\psi)\int_0^T a^2(u)\{(\Phi^2)'(Z_u)\}^2 \mathrm{d}u \geq \eta \int_{\beta_1}^{\beta_2} \{(\Phi^2)'(Z_u)\}^2 \mathrm{d}u.$$

Let

$$\zeta = \eta \int_{\beta_1}^{\beta_2} \{(\Phi^2)'(Z_u)\}^2 \mathrm{d}u.$$

Suppose there exists $\varepsilon > 0$ such that for all $r > 0$, $\mathbb{P}[\zeta \leq r] \geq \varepsilon$. Since $\zeta \geq 0$, $\mathbb{P}[\zeta = 0] \geq \varepsilon$. By Assumption B, this implies $Z_u = 0$ on $[\beta_1, \beta_2]$ with positive probability, which is absurd by the assumptions on model (5). Then, for $\varepsilon > 0$, there exists $r > 0$ such that

$$\mathbb{P}[\zeta \geq 2r] \geq 1 - \varepsilon.$$

By Markov's inequality, we have

$$\mathbb{P}[2^{2jH}Q_j \notin [\zeta - r, \zeta + r]] = \mathbb{P}[|2^{2jH}Q_j - \zeta| > r] \leq c\frac{2^{-j/2}}{r}.$$

Thus,

$$\sum_{j\geq 0} \sup_H \mathbb{P}[2^{2jH}Q_j \notin [\zeta - r, \zeta + r]] < +\infty.$$

Then, by the Borel–Cantelli lemma, for large enough j a.s.

$$2^{2jH}Q_j \geq \zeta - r.$$

We now prove (ii). Let $\varepsilon > 0$, r and j_0 be associated by Proposition 3(i) and $j \geq j_0$. We have

$$\mathbb{P}\left[2^{j/2}\sup_{l\geq j}\left|\frac{Q_{l+1}}{Q_l} - 2^{-2H}\right| \geq M\right] = \mathbb{P}\left[\sup_{l\geq j}|Q_{l+1} - 2^{-2H}Q_l| \geq MQ_l 2^{-j/2}\right]$$
$$\leq \varepsilon + \mathbb{P}\left[\sup_{l\geq j}|Q_{l+1} - 2^{-2H}Q_l| \geq M2^{-j/2}2^{-2lH}r\right]$$
$$\leq \varepsilon + \sum_{l\geq j\geq j_0} \mathbb{E}[|Q_{l+1} - 2^{-2H}Q_l|]2^{2lH}2^{j/2}(Mr)^{-1}.$$

Let

$$L = c(\psi)\int_0^T a^2(u,u)\{(\Phi^2)'[Z_u]\}^2 \mathrm{d}u.$$

The quantity $\mathbb{E}[|Q_{l+1} - 2^{-2H} Q_l|]$ is equal to

$$\mathbb{E}[|Q_{l+1} - 2^{-2(l+1)H} L + 2^{-2(l+1)H} L - 2^{-2H} Q_l|] \leq c 2^{-l(2H+1/2)}.$$

Eventually,

$$\mathbb{P}\left[2^{j/2} \sup_{l \geq j} \left|\frac{Q_{l+1}}{Q_l} - 2^{-2H}\right| \geq M\right] \leq \varepsilon + c \sum_{l \geq j \geq j_0} 2^{-l/2} 2^{j/2} (Mr)^{-1}.$$

For large enough M, this can be made arbitrarily small.

7. Proof of Proposition 4

With the notation of Section 2.2.2, we have

$$\begin{aligned}\widehat{Q_j} - Q_j = {} & \sum_k b_{jk}^2 + \sum_k f_{jk}^2 + \sum_k b_{jk} f_{jk} + \sum_k d_{jk} b_{jk} + \sum_k d_{jk} f_{jk} \\ & + \sum_k (e_{jk}^2 - \bar{\nu}_{jk}) + \sum_k e_{jk} f_{jk} + \sum_k b_{jk} e_{jk} + \sum_k d_{jk} e_{jk} + \sum_k \nu_{jk} - \bar{\nu}_{jk}.\end{aligned}$$

Following Gloter and Hoffmann [17], it is enough to prove

$$\sup_{J_n \geq j \geq j_n(H) - L} \sup_{H \in [H_-, H_+]} 2^{-j/2} \mathbb{E}[|\widehat{Q}_{j,n} - Q_j|] \leq c n^{-1}.$$

Now we bound the 10 terms one by one.

• Term 1: Let $V_{tl} = \sigma_t^2 - \sigma_{k2^{-j}+l2^{-N}}^2$. We have

$$\begin{aligned}\mathbb{E}[b_{jk}^2] &= \sum_{l=0}^{T(2^{N-j}-1)} \sum_{l'=0}^{T(2^{N-j}-1)} \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \int_{\frac{k}{2^j}+\frac{l'}{2^N}}^{\frac{k}{2^j}+\frac{l'+1}{2^N}} \psi_{jk}(t) \psi_{jk}(t') \mathbb{E}[V_{tl} V_{tl'}] \mathrm{d}t \mathrm{d}t' \\ &\leq c 2^j \sum_l \sum_{l'} \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \int_{\frac{k}{2^j}+\frac{l'}{2^N}}^{\frac{k}{2^j}+\frac{l'+1}{2^N}} \left(\mathbb{E}[V_{tl}^2] \mathbb{E}[V_{tl'}^2]\right)^{1/2} \mathrm{d}t \mathrm{d}t'.\end{aligned}$$

Moreover, for $t \in [k2^{-j} + l2^{-N}, k2^{-j} + (l+1)2^{-N}]$,

$$V_{tl} = (Z_t - Z_{k2^{-j}+l2^{-N}}) \Phi^{2'} \{Z_{k2^{-j}+(l+v)2^{-N}}\}$$

with $v \in [0, 1]$. By Assumption B and the same arguments as previously, $\mathbb{E}[V_{tl}^2] \leq c 2^{-2NH}$. Hence $\mathbb{E}[b_{jk}^2] \leq c 2^{-j} n^{-1}$.

• Term 2 and term 3 follow easily with the same order.

• Term 4: As in Lemma 5, we easily prove that $\mathbb{E}[d_{jk}^2] \leq c 2^{-j(1+2H)}$ and then, because $j \geq \frac{1}{2H+1} \log_2(n)$, $\mathbb{E}[|d_{jk} b_{jk}|] \leq c 2^{-j/2} n^{-1}$.

• Term 5 follows as term 4 with the same order.

• Term 6: We argue first conditionally on the path of the volatility. We write $\tilde{\mathbb{E}}$ for the expectation conditional on the path of the volatility. Because of the independence of the Brownian increments and because the variables are centered, we have

$$\tilde{\mathbb{E}}\left[\left(\sum_k e_{jk}^2 - \nu_{jk}\right)^2\right] = \sum_k \tilde{\mathbb{E}}[(e_{jk}^2 - \nu_{jk})^2] \leq c \sum_k \tilde{\mathbb{E}}[e_{jk}^4 + \nu_{jk}^2].$$

Let

$$M_l = \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t \, \mathrm{d}B_t\right)^2 - \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t^2 \mathrm{d}t.$$

Because the variables $M_l, l = 0, \dots, T(2^{N-j}-1)$, are centered and independent, we get that $\tilde{\mathbb{E}}[e_{jk}^4]$ is equal to

$$\sum_{l=0}^{T(2^{N-j}-1)} \sum_{l'=0}^{T(2^{N-j}-1)} n^4 \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t \int_{\frac{k}{2^j}+\frac{l'}{2^N}}^{\frac{k}{2^j}+\frac{l'+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\right)^2 \tilde{\mathbb{E}}[M_l^2 M_{l'}^2].$$

Indeed the products of terms of power 3 with terms of power 1 are equal to zero. But, we have the following equality in law:

$$M_l^2 \overset{\mathcal{L}}{=} \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t^2 \mathrm{d}t\right)^2 (Z^2-1)^2,$$

with Z a standard Gaussian variable. Hence,

$$\tilde{\mathbb{E}}[M_l^4] = c\left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t^2 \mathrm{d}t\right)^4.$$

Now, we have

$$\mathbb{E}\left[\left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t^2 \mathrm{d}t\right)^4\right] \le \iiiint \left(\mathbb{E}[\sigma_{t_1}^8]\mathbb{E}[\sigma_{t_2}^8]\mathbb{E}[\sigma_{t_3}^8]\mathbb{E}[\sigma_{t_4}^8]\right)^{1/4} \mathrm{d}t_1\mathrm{d}t_2\mathrm{d}t_3\mathrm{d}t_4.$$

Moreover, there exists $\theta \in [0, T]$ such that

$$\sigma_t^2 = \Phi^2(Z_t) = \Phi^2(0) + \Phi^{2'}(Z_\theta)Z_t.$$

This leads to $\mathbb{E}[\sigma_t^8] \le c$. Hence $\mathbb{E}[e_{jk}^4] \le cn^{-2}$. We have

$$\mathbb{E}[\nu_{jk}^2] = 4n^4 \sum_{l=0}^{T(2^{N-j}-1)} \sum_{l'=0}^{T(2^{N-j}-1)} \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t \int_{\frac{k}{2^j}+\frac{l'}{2^N}}^{\frac{k}{2^j}+\frac{l'+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\right)^2$$
$$\times \mathbb{E}\left[\left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t^2 \mathrm{d}t\right)^2 \left(\int_{\frac{k}{2^j}+\frac{l'}{2^N}}^{\frac{k}{2^j}+\frac{l'+1}{2^N}} \sigma_t^2 \mathrm{d}t\right)^2\right].$$

In the same way as for $\mathbb{E}[e_{jk}^4]$, we get $\mathbb{E}[\nu_{jk}^2] \le cn^{-2}$.

- Term 7: In the preceding proof, we have shown $\mathbb{E}[e_{jk}^2] \le cn^{-1}$ and so we obtain $\mathbb{E}[|f_{jk}e_{jk}|] \le c2^{-j/2}n^{-1}$.
- Term 8 follows exactly as term 7.

• Term 9: We argue first conditionally on the path of the volatility. Because of the independence of the Brownian increments and because the variables are centered, we have

$$\tilde{\mathbb{E}}\left[\left(\sum_k e_{jk} d_{jk}\right)^2\right] = \sum_k d_{jk}^2 \tilde{\mathbb{E}}[e_{jk}^2].$$

Again because of the independence of the Brownian increments and because the variables are centered, we have

$$\begin{aligned} d_{jk}^2 \tilde{\mathbb{E}}[e_{jk}^2] &= c \int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \psi_{jk}(t_1)\sigma_{t_1}^2 \mathrm{d}t_1 \int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \psi_{jk}(t_2)\sigma_{t_2}^2 \mathrm{d}t_2 \\ &\quad \times \sum_l n^2 \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t_3)\mathrm{d}t_3\right)^2 \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_{t_3}^2 \mathrm{d}t_3\right)^2. \end{aligned}$$

So, we get

$$\begin{aligned} &\mathbb{E}[d_{jk}^2 e_{jk}^2] = cn^2 \sum_l \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t_3)\mathrm{d}t_3\right)^2 \\ &\times \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \mathbb{E}\left[\int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \psi_{jk}(t_1)\sigma_{t_1}^2 \mathrm{d}t_1 \int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \psi_{jk}(t_2)\sigma_{t_2}^2\sigma_{t_3}^2\sigma_{t_4}^2 \mathrm{d}t_2\right] \mathrm{d}t_3 \mathrm{d}t_4. \end{aligned}$$

Because of the vanishing moment of the wavelet, we have

$$\begin{aligned} &\mathbb{E}\left[\int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \psi_{jk}(t_1)\sigma_{t_1}^2 \mathrm{d}t_1 \int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \psi_{jk}(t_2)\sigma_{t_2}^2\sigma_{t_3}^2\sigma_{t_4}^2 \mathrm{d}t_2\right] \\ &= \mathbb{E}\left[\int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \psi_{jk}(t_1) \int_{\frac{k}{2^j}}^{\frac{k+T}{2^j}} \psi_{jk}(t_2) V_{t_1 0} V_{t_2 0} \sigma_{t_3}^2\sigma_{t_4}^2 \mathrm{d}t_2 \mathrm{d}t_1\right] \\ &\leq c 2^j 2^{-2j} \left(\mathbb{E}[V_{t_1 0}^4]\mathbb{E}[V_{t_2 0}^4]\right)^{1/4} \leq c 2^{-j} 2^{-j2H}. \end{aligned}$$

Consequently, $\mathbb{E}[d_{jk}^2 e_{jk}^2] \leq cn^{-1}2^{-3j}$, but, as $j \geq \frac{\log_2 n}{3}$, $\mathbb{E}[d_{jk}^2 e_{jk}^2] \leq cn^{-2}$.

• Term 10: Let $X = \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t \, \mathrm{d}B_t\right)^2$ and $X_i = \left(\int_{\frac{k}{2^j}+\frac{l+i}{2^N}}^{\frac{k}{2^j}+\frac{l+i+1}{2^N}} \sigma_t \, \mathrm{d}B_t\right)^2$. Then,

$$\begin{aligned} \nu_{jk} &= 2 \sum_{l=0}^{2^{N-j}-1} n^2 \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\right)^2 \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_u^2 \mathrm{d}u\right)^2, \\ \bar{\nu}_{jk} &= 2 \sum_{l=0}^{2^{N-j}-1} n^2 \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\right)^2 \left(\frac{1}{h}\sum_{i=0}^{h} X_i\right)^2, \end{aligned}$$

where $h = h(n) = \lfloor n^{1/2} \rfloor$. The term $\nu_{jk} - \bar{\nu}_{jk}$ is equal to

$$2\sum_{l=0}^{2^{N-j}-1} n^2 \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \psi_{jk}(t)\mathrm{d}t\right)^2 \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_u^2\mathrm{d}u + \frac{1}{h}\sum_{i=0}^{h} X_i\right)$$
$$\times \left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_u^2\mathrm{d}u - \frac{1}{h}\sum_{i=0}^{h} X_i\right).$$

We argue first conditionally on the path of the volatility. We have

$$\tilde{\mathbb{E}}\left[\left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_u^2\mathrm{d}u - \frac{1}{h}\sum_{i=0}^{h} X_i\right)^2\right]$$
$$\leq c\tilde{\mathbb{E}}\left[\left(\frac{1}{h}\sum_{i=0}^{h}\{X_i - \tilde{\mathbb{E}}[X_i]\}\right)^2\right] + c\tilde{\mathbb{E}}\left[\left(\frac{1}{h}\sum_{i=0}^{h}\tilde{\mathbb{E}}[X_i] - \tilde{\mathbb{E}}[X]\right)^2\right],$$

with the following equality in law:

$$X_i - \tilde{\mathbb{E}}[X_i] \overset{\mathcal{L}}{=} \left(\int_{\frac{k}{2^j}+\frac{l+i}{2^N}}^{\frac{k}{2^j}+\frac{l+i+1}{2^N}} \sigma_t^2\mathrm{d}t\right)(Z^2-1),$$

with Z a standard Gaussian variable. Now,

$$\mathbb{E}\left[\left(\int_{\frac{k}{2^j}+\frac{l+i}{2^N}}^{\frac{k}{2^j}+\frac{l+i+1}{2^N}} \sigma_t^2\mathrm{d}t\right)^2\right] \leq c2^{-2N}.$$

Then, by independence of the Brownian increments and because the variables are centered,

$$\tilde{\mathbb{E}}\left[\left(\frac{1}{h}\sum_{i=0}^{h}\{X_i - \tilde{\mathbb{E}}[X_i]\}\right)^2\right] = \frac{1}{h^2}\sum_{i=0}^{h}\tilde{\mathbb{E}}[(X_i - \tilde{\mathbb{E}}X_i)^2] \leq \frac{c}{h}2^{-2N}.$$

For the other term, $\mathbb{E}[(\frac{1}{h}\sum_{i=0}^{h}\tilde{\mathbb{E}}X_i - \tilde{\mathbb{E}}X)^2]$ is equal to

$$\mathbb{E}\left[\left(\frac{1}{h}\sum_{i=0}^{h}\int_{\frac{k}{2^j}+\frac{l+i}{2^N}}^{\frac{k}{2^j}+\frac{l+i+1}{2^N}} \sigma_t^2\mathrm{d}t - \int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_t^2\mathrm{d}t\right)^2\right]$$
$$= \frac{1}{h^2}\sum_{i=0}^{h}\sum_{g=0}^{h}\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}}\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \mathbb{E}[(\sigma_{u+i2^{-N}}^2 - \sigma_u^2)(\sigma_{v+g2^{-N}}^2 - \sigma_v^2)]\mathrm{d}u\mathrm{d}v$$
$$\leq \frac{c}{h^2}\sum_{i=0}^{h}\sum_{g=0}^{h}\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}}\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} (ig)^H 2^{-2NH}\mathrm{d}u\mathrm{d}v \leq c2^{-2N(1+H)}h^{2H}.$$

Eventually,

$$\mathbb{E}\left[\left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_u^2\mathrm{d}u - \frac{1}{h}\sum_{i=0}^{h} X_i\right)^2\right] \leq c\frac{n^{-2}}{\sqrt{n}}.$$

We easily check that the term

$$\mathbb{E}\left[\left(\int_{\frac{k}{2^j}+\frac{l}{2^N}}^{\frac{k}{2^j}+\frac{l+1}{2^N}} \sigma_u^2 \mathrm{d}u + \frac{1}{h}\sum_{i=0}^{h} X_i\right)^2\right]$$

is less than cn^{-2} and finally $\mathbb{E}[|\nu_{jk} - \bar{\nu}_{jk}|] \leq cn^{-1}2^{-j/2}$, because $j \leq \frac{\log_2 n}{2}$.

8. Proof of Theorem 3(ii)

We now prove that Propositions 3 and 4 together imply Theorem 3(ii). Following Lemma 1 of Gloter and Hoffmann [17], we easily obtain that for all positive ε, there exist n_0 and $M > 0$, such that for all $n \geq n_0$,

$$\mathbb{P}[n^{1/(4H+2)}|\widehat{H}_n - H| \geq M] \leq \varepsilon. \tag{16}$$

With no loss of generality, we may demand $\widehat{H} \leq C$, with $C > 2$ a constant value, by considering $\widetilde{H} = \widehat{H}\mathbb{1}_{|\widehat{H}|\leq C}$. Let $\varepsilon > 0$, n_0, M be associated by (16). For $n \geq n_0$, if $(C-1)n^{1/(4H+2)} > M$, we have

$$\mathbb{P}[\widehat{H}_n \geq C] \leq \mathbb{P}[n^{1/(4H+2)}|\widehat{H}_n - H| \geq (C-1)n^{1/(4H+2)}] \leq \varepsilon.$$

Let $n_0^* \geq n_0$ be such that $(C-1)n_0^* \geq M$. For all $n \leq n_0^*$,

$$n^{1/(4H+2)}|\widetilde{H}_n - H| \leq (C+1)(n_0^*)^{1/(4H+2)}.$$

Let $M_1 = \max\{M, (C+1)(n_0^*)^{1/(4H+2)}\}$. For all n,

$$\mathbb{P}[n^{1/(4H+2)}|\widetilde{H}_n - H| \geq M_1] \leq \varepsilon.$$

9. Proof of Theorem 2

9.1. *Proof of Proposition 5*

We observe

$$\left\{Y_{i/n} = y_0 + \int_0^{i/n} \Phi\left(\int_0^s a(s,u)\mathrm{d}W_u^H + f(s)\xi_0\right)\mathrm{d}B_s, i = 1,\ldots,nT\right\}.$$

Without loss of generality, we set here $\xi_0 = 0$. Consider the equivalent sample

$$\{Z_{i/n} = Y_{i/n} - Y_{(i-1)/n}, i = 1,\ldots,nT\}.$$

Conditional on $W^H = f$, $Z_{i/n}$ is a centered Gaussian variable with variance γ_i where

$$\gamma_i = \int_{(i-1)/n}^{i/n} \Phi^2\left(\int_0^s a(s,u)\mathrm{d}f_u\right)\mathrm{d}s.$$

Moreover, conditional on W^H, the observations are independent. We define by $K(\mu,\nu) = \int(\log\frac{\mathrm{d}\mu}{\mathrm{d}\nu})\mathrm{d}\mu \leq +\infty$ the Kullback–Leibler divergence between two probability measures

μ and ν. We recall the classical Pinsker's inequality $\|\mu-\nu\|_{TV} \leq \sqrt{2}K(\mu,\nu)^{1/2}$. Let $\mathbb{P}^n_f$ be the law of the sample conditional on $W^H = f$; let

$$\beta_i = \int_{(i-1)/n}^{i/n} \Phi^2\left(\int_0^s a(s,u)\mathrm{d}g_u\right)\mathrm{d}s.$$

We have

$$\|\mathbb{P}^n_f - \mathbb{P}^n_g\|_{TV} \leq \sqrt{2}K(\mathbb{P}^n_f, \mathbb{P}^n_g)^{1/2}.$$

By classical computations, we get

$$K(\mathbb{P}^n_f, \mathbb{P}^n_g) = \frac{1}{2}\sum_{i=1}^{nT}\left(-\log\frac{\gamma_i}{\beta_i} - 1 + \frac{\gamma_i}{\beta_i}\right).$$

By Assumption C, we have $(c_4/c_5)^2 \leq \gamma_i/\beta_i \leq (c_5/c_4)^2$. Let $a = (c_4/c_5)^2$, $b = (c_5/c_4)^2$ and $c \geq 1/2$. Consider

$$z(x) = \log x - 1 + 1/x - c(x-1)^2, \quad x \in [a,b].$$

We have $z(a) = \log a - 1 + 1/a - c(a-1)^2$, so, if $c \geq \frac{\log a - 1 + 1/a}{(a-1)^2}$, we have $z(a) \leq 0$. Take

$$c = c^* = \max\left(\frac{1}{2}, \frac{\log a - 1 + 1/a}{(a-1)^2}\right).$$

Hence z is non-positive on $[a,b]$; consequently, $K(\mathbb{P}^n_f, \mathbb{P}^n_g)$ is less than

$$\begin{aligned}
K(\mathbb{P}^n_f, \mathbb{P}^n_g) &\leq c\sum_{i=1}^{nT}\left(\frac{\beta_i}{\gamma_i} - 1\right)^2 \\
&\leq cn^2\sum_{i=1}^{nT}\left(\int_{(i-1)/n}^{i/n}\left|\Phi\left(\int_0^s a(s,u)\mathrm{d}f_u\right) - \Phi\left(\int_0^s a(s,u)\mathrm{d}g_u\right)\right|\mathrm{d}s\right)^2 \\
&\leq cn\int_0^T\left|\int_0^s a(s,u)\mathrm{d}f_u - \int_0^s a(s,u)\mathrm{d}g_u\right|^2\mathrm{d}s \\
&\leq cn\int_0^T\left|a(s,s)[f(s)-g(s)] + \int_0^s \partial_2 a(s,u)[g(u)-f(u)]\mathrm{d}u\right|^2\mathrm{d}s \\
&\leq cn\|f-g\|_2^2.
\end{aligned}$$

Acknowledgements

I am grateful to Marc Hoffmann, François Simenhaus and Vincent Vargas for their helpful remarks. I also thank an anonymous referee and the associate editor whose comments have substantially improved a former version of this paper.

Appendix

A.1. *Proof of Proposition 1*

The link between Besov spaces and Gaussian processes has been heavily studied; see in particular Ciesielski, Kerkyacharian and Roynette [8] and Nualart and Ouknine [31]. We give here some simple proofs for our case. Let (ϕ, ψ) be a well chosen wavelet basis. For f a real function on $\mathbb{R}$, we set

$$\alpha_{0k} = \int f(x)\phi_{0k}(x)\mathrm{d}x, \qquad \beta_{jk} = \int f(x)\psi_{jk}(x)\mathrm{d}x.$$

Recall that in terms of wavelet coefficients, the Besov space $\mathcal{B}^s_{p,q}(\mathbb{R})$, with $s \in [0, 1]$, $1 \leq p, q < \infty$, is the space of all functions f such that the following norm is finite:

$$\|f\|_{\mathcal{B}^s_{p,q}} = \|\alpha_{0.}\|_{l_p} + \left[\sum_j \left(2^{j(s-1/p+1/2)}\|\beta_{j.}\|_{l_p}\right)^q\right]^{1/q},$$

where

$$\|\beta_{j.}\|_{l_p} = \left(\sum_k |\beta_{jk}|^p\right)^{1/p}.$$

If p or q is equal to ∞, then the corresponding norm in p or q is replaced by the sup norm. For details, we refer the reader to Cohen [10]. Here we say that an $f \in \mathcal{B}^s_{p,q}([0, T])$ if there exists g such that $g \in \mathcal{B}^s_{p,q}(\mathbb{R})$ and the restriction of g to $[0, T]$ is equal to f.

First, we show that the trajectory of $t \to \sigma_t^2$ belongs a.s. to $\mathcal{B}^H_{2,\infty}([0, T])$. It is enough to prove that $\sup_j 2^{2jH} Q_j < \infty$. We know that for all positive ε, there exist j_0 and $M > 0$ such that

$$\mathbb{P}\left[2^{j/2} \sup_{l \geq j \geq j_0} \left|\frac{Q_{l+1}}{Q_l} - 2^{-2H}\right| \geq M\right] \leq \varepsilon.$$

This implies that

$$\mathbb{P}\left[\exists j_0, \exists M, 2^{j/2} \sup_{l \geq j \geq j_0} \left|\frac{Q_{l+1}}{Q_l} - 2^{-2H}\right| \leq M\right] = 1.$$

Let $u_j = 2^{2jH} Q_j$. For such j_0, for all $j \geq j_0$, $|u_{j+1}/u_j| \leq 1 + M2^{-j/2}$. Thus, $\log u_{j+1} - \log u_j \leq \log(1 + M2^{-j/2}) \leq M2^{-j/2}$ and $\log u_n \leq c$. Hence the trajectory belongs a.s. to $\mathcal{B}^H_{2,\infty}([0, T])$. Nevertheless, it does not belong to $\mathcal{B}^H_{2,q}([0, T])$, $q < \infty$. As a matter of fact, for all ε positive, there exist j_0 and $r > 0$ such that for all $j \geq j_0$, $\mathbb{P}[2^{2jH} Q_j \geq r] \geq 1 - \varepsilon$. So, almost surely,

$$\sum_{j=0}^{+\infty} (2^{2jH} Q_j)^q = +\infty.$$

The fact that for $s < H$, the trajectory belongs almost surely to $\mathcal{B}^s_{\infty,\infty}([0, T])$ is clear by Kolmogorov's criterion and preceding calculations on the expectations. We now prove that it does not belong to $\mathcal{B}^H_{\infty,\infty}([0, T])$. Suppose that almost surely, there exists c such that for all $(s, t) \in [\beta_1, \beta_2]$,

$$\left|\Phi^2\left(\int_0^t a(t, u)\mathrm{d}W_u^H\right) - \Phi^2\left(\int_0^s a(s, u)\mathrm{d}W_u^H\right) + [f(t) - f(s)]\xi_0\right| \leq c|t - s|^H.$$

Because there exists $c > 0$ such that for all x, $|(\Phi^2)'(x)| > c$, this implies

$$\left| \int_0^t a(t,u) \mathrm{d}W_u^H - \int_0^s a(s,u) \mathrm{d}W_u^H + [f(t) - f(s)]\xi_0 \right| \le c|t-s|^H.$$

Ito's formula gives

$$|W_t^H a(t,t) - W_s^H a(s,s) + R(t,s)| \le c|t-s|^H,$$

with

$$R(t,s) = [f(t) - f(s)]\xi_0 - \int_s^t \partial_2 a(t,u) W_u^H \mathrm{d}u - \int_0^s \partial_2 [a(t,u) - a(s,u)] W_u^H \mathrm{d}u.$$

For fixed $\varepsilon > 0$ and $|t-s|$ small enough,

$$(t-s)^{1-H} \left| \frac{R(t,s)}{t-s} \right| \le \varepsilon$$

and consequently,

$$\left| \frac{(W_t^H - W_s^H) a(t,t)}{(t-s)^H} - \frac{W_s^H [a(s,s) - a(t,t)]}{(t-s)^H} \right| \le c + \varepsilon.$$

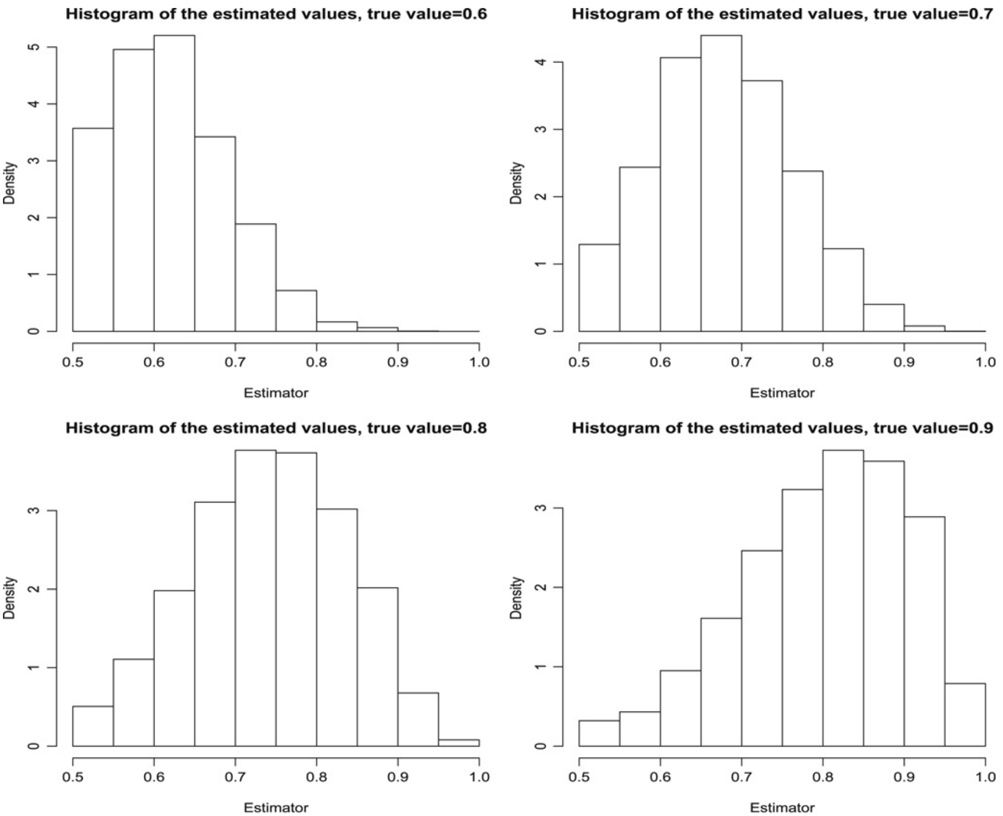
Eventually, we get for $|t-s|$ small enough,

$$\left| \frac{W_t^H - W_s^H}{(t-s)^H} \right| \le \frac{c + 2\varepsilon}{\inf\limits_{x \in [\beta_1, \beta_2]} a(x,x)},$$

which is absurd as a.s. the fbm is H Hölderian on no interval; see Arcones [2].

A.2. One numerical illustration

As explained in Section 1.4, getting accurate estimations with a small number of data is hopeless. Nevertheless, the estimation rates remain polynomial and so estimation procedures are conceivable as soon as we get a "reasonably big" number of data. For example, financial data are available in large amounts. Moreover, the number of data can be increased using aggregation techniques, between assets. Note that the fact that T is fixed and that we consider a "high frequency" asymptotic does not mean we can only consider $T = 1$ day. The value of T might be in the order of magnitude of years. Nevertheless, the sampling period has to be such that the discretized process lives at the "diffusive scale" (that is in general a sampling period bigger than 10 min), in order to avoid microstructure noise effects. We present here some numerical results in the model where $\sigma_t = \mathrm{e}^{W_t^H}$. For $H = 0.6, 0.7, 0.8, 0.9$, 10 000 simulations are done with a frequency $n = 2^{16}$. We compute the estimator for each simulated sample path. We set the estimator to 0.5 if the estimated value is smaller than 0.5 and to 1 if it is bigger than 1. We obtain the following histograms.



We see that the distributions are shifted to the right when going from 0.6 to 0.9.

References

[1] T.G. Andersen, T. Bollerslev, Intraday periodicity and volatility persistence in financial markets, Internat. Econom. Rev. 39 (1997) 885–905.

[2] M.A. Arcones, On the law of the iterated logarithm for Gaussian processes, J. Theoret. Probab. 8 (1995) 877–903.

[3] O.E. Barndorff-Nielsen, E. Nicolato, N. Shephard, Some recent advances in stochastic volatility modelling, Quant. Finance 2 (2002) 11–23.

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

[5] C. Berzin, J.R. León, Convergence in fractional models and applications, Electron. J. Probab. 10 (2005) 326–370.

[6] F. Breidt, N. Crato, P. De Lima, The detection and estimation of long-memory in stochastic volatility, J. Econometrics 83 (1998) 325–348.

[7] P. Cheridito, H. Kawaguchi, M. Maejima, Fractional Ornstein–Uhlenbeck processes, Electron. J. Probab. 8 (2003) 1–14.

[8] Z. Ciesielski, G. Kerkyacharian, B. Roynette, Quelques espaces fonctionnels associés à des processus gaussiens, Studia Math. 107 (1993) 172–204.

[9] J.F. Cœurjolly, Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths, Stat. Inference Stoch. Process. 4 (2) (2001) 199–227.

[10] A. Cohen, Wavelet methods in numerical analysis, in: P.G. Ciarlet, J.L. Lions (Eds.), Handbook of Numerical Analysis, vol. VII, Elsevier Science, 1999.

[11] F. Comte, L. Coutin, E. Renault, Affine fractional stochastic volatility models with application to option pricing, 2003, Preprint, University of Montreal.

[12] F. Comte, E. Renault, Long-memory in continuous-time stochastic volatility models, Math. Finance 8 (1998) 291–323.
[13] I. Daubechies, Orthonormal bases of compactly supported wavelets, Comm. Pure Appl. Math. 41 (1988) 909–996.
[14] R. Deo, C. Hurvich, Y. Lu, On the log-periodogram regression estimator of the memory parameter in long memory stochastic volatility model, Econometric Theory 17 (4) (2004) 686–710.
[15] P. Doukhan, G. Oppenheim, M. Taqqu (Eds.), Long-Range Dependence: Theory and Applications, Birkhäuser, Boston, 2003.
[16] A. Gloter, M. Hoffmann, Stochastic volatility and fractional Brownian motion, Stochastic Process. Appl. 113 (2004) 143–172.
[17] A. Gloter, M. Hoffmann, Estimation of the Hurst parameter from discrete noisy data, The Annals of Statistics (2007) (in press).
[18] A.C. Harvey, Long-memory in stochastic volatility, in: J. Knight, S. Satchell (Eds.), Forecasting Volatility in Financial Markets, Butterworth-Heineman, Oxford, 1998, pp. 307–320.
[19] J. Hull, A. White, An analysis of the bias in option pricing caused by a stochastic volatility, Adv. Futures Options Res. 3 (1988) 29–61.
[20] C. Hurvich, E. Moulines, P. Soulier, Estimating long-memory in stochastic volatility, Econometrica 73 (4) (2005) 1283–1328.
[21] C. Hurvich, K. Ray, The local whittle estimator of long-memory stochastic volatility, J. Financial Econometrics 1 (3) (2003) 445–470.
[22] C. Hurvich, P. Soulier, Testing for long-memory in volatility, Econometric Theory 18 (6) (2002) 1291–1308.
[23] J. Istas, G. Lang, Quadratic variations and estimation of the local Hölder index of a Gaussian process, Ann. Inst. H. Poincaré 33 (4) (1997) 407–436.
[24] M. Jensen, Semiparametric bayesian inference of long-memory stochastic volatility models, J. Time Ser. Anal. 25 (6) (2004) 895–922.
[25] G. Lang, F. Roueff, Semi-parametric estimation of the Hölder exponents of stationary Gaussian processes with minimax rates, Stat. Inference Stoch. Process. 4 (3) (2001) 283–306.
[26] J. Lee, Wavelet transform for estimating the memory parameter in long memory stochastic volatility model, 2004. Preprint, National University of Singapore.
[27] B. Mandelbrot, J. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev. 10 (1968) 422–437.
[28] A. Melino, S.M. Turnbull, Pricing foreign currency options with stochastic volatility, J. Econometrics 45 (1990) 239–265.
[29] M. Musiela, M. Rutkowski, Martingale Methods in Financial Modelling, Springer, 2005.
[30] I. Norros, E. Valkeila, Y. Virtamo, An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motions, Bernoulli 5 (4) (1999) 571–587.
[31] D. Nualart, Y. Ouknine, Besov regularity of stochastic integrals with respect to the fractional Brownian motion with parameter $H > 1/2$, J. Theoret. Probab. 16 (2003) 451–470.
[32] A. Ohanissian, J. Russel, R. Tsay, True or spurious long memory in volatility: Does it matter for option pricing? 2004. Preprint.
[33] P.M. Robinson, The memory of stochastic volatility models, J. Econometrics 101 (2001) 195–218.
[34] N. Shephard, Stochastic volatility. in: New Palgrave Dic. of Econ., 2nd ed., 2006.
[35] S. Taylor, Consequences for option pricing of a long-memory in volatility, 2000. Preprint.
[36] G. Teyssière, Interaction models for common long-range in asset price volatilities, in: G. Rangarajan, M. Ding (Eds.), Processes with Long Range Correlations: Theory and Applications, in: Lecture Notes in Physics, vol. 621, Springer-Verlag, 2003, pp. 251–269.