Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER Stochastic Processes and their Applications 113 (2004) 143–172

stochastic processes and their applications

www.elsevier.com/locate/spa

Stochastic volatility and fractional Brownian motion

A. Gloter[a], M. Hoffmann[b,*]

[a] *GRAPE, CNRS UMR 5113 and Université Bordeaux IV, av. Léon Duguit, 33608 Pessac Cédex, France*

[b] *Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS UMR 8095, Cité Descartes, 5 Boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2, France*

Received 17 July 2002; received in revised form 18 March 2004; accepted 19 March 2004

Abstract

We observe (Y_t) at times i/n, $i=0,\dots,n$, in the parametric stochastic volatility model

$$\mathrm{d}Y_t = \Phi(\theta, W_t^H)\,\mathrm{d}W_t,$$

where (W_t) is a Brownian motion, independent of the fractional Brownian motion (W_t^H) with Hurst parameter $H \geqslant \frac{1}{2}$. The sample size n increases not because of a longer observation period, but rather, because of more frequent observations.

We prove that the unusual rate $n^{-1/(4H+2)}$ is asymptotically optimal for estimating the one-dimensional parameter θ, and we construct a contrast estimator based on an approximation of a suitably normalized quadratic variation that achieves the optimal rate.

MSC: 62G99; 62F99; 62M99

Keywords: Stochastic volatility models; Discrete samplings; High-frequency data; Fractional Brownian motion; Contrast estimators

1. Introduction

1.1. Statistical model

We are interested in statistical inference in a stochastic volatility model driven by a fractional Brownian motion. For $i=0,\dots,n$ we observe at times i/n a one-dimensional

* Corresponding author. Tel.: +33-01-60-95-75-24; fax: +33-01-60-95-75-45.

E-mail addresses: gloter@montesquieu.u-bordeaux.fr (A. Gloter), hoffmann@math.univ-mlv.fr (M. Hoffmann).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2004.03.008

stochastic process (Y_t) of the form

$$Y_t = y_0 + \int_0^t \sigma_s \, \mathrm{d}W_s, \tag{1}$$

where (W_t) is a standard Brownian motion and (σ_t) is unknown and *stochastic*, in the sense that

$$\sigma_t = \Phi(\theta, W_t^H), \tag{2}$$

where (W_t^H) is a fractional Brownian motion (fBM for short), independent of (W_t), with Hurst parameter $H \geqslant \frac{1}{2}$. The function Φ is known up to a parameter θ lying in a compact interval $\Theta = [\theta_-, \theta_+] \subset \mathbb{R}$. The law of the data $(Y_{i/n},\ i = 0, \dots, n)$ is denoted by $\mathbb{P}_\theta^n$.

1.2. Goal

Recovering θ from data $(Y_{i/n})$ is the objective of the paper. A rate $v_n \to 0$ is said to be *achievable* if there exists an estimator $\hat{\theta}_n$ such that the normalized error

$$\{v_n^{-1}(\hat{\theta}_n - \theta)\}_{n \geqslant 1} \tag{3}$$

is bounded in $\mathbb{P}_\theta^n$-probability, uniformly over open sets in Θ. The rate v_n is a *lower rate of convergence* if, for all open set $U \subset \Theta$, there exists $C > 0$ such that

$$\liminf_{n \to \infty} \inf_F \sup_{\theta \in U} \mathbb{P}_\theta^n \{v_n^{-1} |F - \theta| \geqslant C\} > 0, \tag{4}$$

where the infimum is taken over all estimators.

We prove in the paper that the unusual rate $v_n(H) := n^{-1/(4H+2)}$ is optimal for the sequence of experiments $(\mathbb{P}_\theta^n, \theta \in \Theta)$. This means that (3) and (4) agree, with $v_n = v_n(H)$. We exhibit an optimal estimator based on approximations of integrated functions of fractional Brownian motion. These approximations may have some interest for their own.

1.3. Motivation

1.3.1. Stochastic volatility and fractional Brownian motion

Model (1), (2) is primarily motivated by financial mathematics and econometrics (Hull and White, 1988; Melino and Turnbull, 1990; Heston, 1993). The Brownian dynamics in the stochastic volatility has progressively been replaced by more elaborate Gaussian processes, having in particular the property of long-range dependence (Comte and Renault, 1996, 1998). This is the case for fBM when $H \geqslant \frac{1}{2}$. See also Breidt et al. (1998). More generally, there is a growing interest for using fBM in finance: Dai and Heyde (1996), Salopek (1998), Sottinen (2001) among many others.

We show in the paper that, in the high-frequency data framework—see below—the introduction of a Hurst parameter $H \geqslant \frac{1}{2}$ strongly influences the statistical nature of the stochastic volatility model, even in the simplest form (1), (2).

1.3.2. High-frequency data

There are several ways to discretize (Y_t), whether we allow the experiment duration to grow with n or not. As soon as we want to relax any assumption about ergodicity—for modelling reason, say—and/or if we think of statistical experiments living over a fixed time period, we are led to the high-frequency data framework, where the time between two data is of order $1/n$. For related estimation problems in stochastic volatility models with a different asymptotic framework, we refer to Genon-Catalot et al. (2000a,b).

1.3.3. Maximum likelihood estimator

Conditional on $(W_t^H) = (\omega_t)$, the random variables $Y_{i/n} - Y_{(i-1)/n}$ are independent, centred Gaussian, with variance

$$v_{i,n}(\theta,\omega) = \int_{(i-1)/n}^{i/n} \Phi(\theta,\omega_s)^2\,\mathrm{d}s.$$

However, since (W_t^H) is not observed, there is no simple formula for the likelihood function

$$\theta \rightsquigarrow \mathbb{E}\left\{\prod_{i=0}^{n} p\left(Y_{i/n} - Y_{(i-1)/n}, v_{i,n}(\theta, W_t^H)\right)\right\} \tag{5}$$

with $p(x,a) = (2\pi a)^{-1/2}\exp(-(x-a)^2/2)$. A consequence is that the maximum likelihood estimator is not tractable. Formula (5) can be misleading: a naive guess that the conditional Fisher information (which is of order n) and the true (unconditional) information are of the same order is wrong!

This is why we take a different route for constructing an estimator and proving a lower bound in this model.

1.4. Results

We say that a real-valued function belongs to $\mathscr{C}^l_{\mathscr{P}}$ if it is $\mathscr{C}^l$ and dominated, along with its mixed derivatives up to order l, by a polynomial.

Assumption A. $(x,\theta) \rightsquigarrow \Phi(\theta,x)$ belongs to $\mathscr{C}^3_{\mathscr{P}}$.

Assumption B. For all $\theta\in\Theta$, $x \rightsquigarrow \Phi(\theta,x)^2$ is one-to-one from $\mathbb{R}$ to $\mathscr{X}$; $(x,\theta) \rightsquigarrow b(\theta,x) := \partial_x\Phi^2(\theta,(\Phi^2)^{-1}(\theta,x))$ is $\mathscr{C}^3$ and for all $x\in\mathscr{X}$ and $\theta\in\Theta$: $b(\theta,x) > 0$.

Assumption C. For all $x\in\mathscr{X}$ and $\theta\in\Theta$: $\partial_\theta b(\theta,x)/b(\theta,x) > 0$.

Theorem 1. *Let $H\in(\frac{1}{2},1)$. Grant Assumptions* A–C. *Then the rate $v_n(H) := n^{-1/(4H+2)}$ is achievable. Moreover, the estimator $\hat\theta_n$, explicitly constructed in Section* 3.2 *and given in* (12)–(15) *achieves the rate $v_n(H)$.*

The accuracy $v_n(H)$ is slower by a polynomial order than the usual $n^{-1/2}$ of regular parametric models. Our next result shows that, under a further restriction on the nondegeneracy of the model, this result is indeed optimal.

Assumption D. $\inf_{x\in K,\theta\in\Theta}\Phi(\theta,x)>0$ for some compact $K\subset\mathbb{R}$. Moreover, $\inf_{x\in\mathbb{R},\theta\in\Theta}\partial_x\Phi(\theta,x)>0$.

Theorem 2. *Let* $H\in[\frac{1}{2},1)$. *Grant Assumptions* A–D. *Then the rate* $v_n(H):=n^{-1/(4H+2)}$ *is a lower rate of convergence.*

Several remarks are in order:

(a) Assumption A stands for the regularity of the model. It is not minimal, but ensures that standard computations are straightforward.
(b) Assumption B ensures the invertibility of the function $\Phi(\theta,\cdot)$, so that the paths of (σ_t) roughly behave like that of (W_t^H). Indeed, the resulting function $b(\theta,\cdot)$ may be understood as a diffusion coefficient, see Section 3.
(c) Assumption C is a standard identifiability condition, as soon as we understand $b(\theta,\cdot)$ as a kind of diffusion coefficient.
(d) Theorems 1 and 2 extend former results of Gloter (2000) and Hoffmann (2002) obtained for the case $H=\frac{1}{2}$. However, the cryptic rate $n^{-1/4}$ obtained for the Brownian case can now be linked to the smoothness H of fractional Brownian motion via the formula $n^{-1/(4H+2)}$.
(e) Extension to other loss functions and/or to multidimensional parameter space Θ is presumably possible, at a further technical cost.

1.5. Organization of the paper

In Section 2, we present an approximation result for integrated quadratic functionals of Brownian motion (Theorem 3) that is the cornerstone for our construction of an estimator, explicitly given in Section 3. Sections 4, 5 and 6 successively prove Theorems 3 and 1, and should be read in that order. Section 6 proves Theorem 2. Appendix A contains auxiliary useful technical results.

The notation c will be extensively used to denote a constant that may vary from line to line. The constant c will always be continuous in its arguments (if any), so it will be implicit that upper bounds involving c are uniform in parameters that are omitted in the notation but that live on compact sets.

2. Some approximations

For a continuous process (X_t) with $t\in[0,1]$, we define, for $N\geqslant 1$

$$\bar{X}_i=\bar{X}_{i,N}:=N\int_{i/N}^{(i+1)/N}X_s\,\mathrm{d}s,\quad i=0,\dots,N-1. \tag{6}$$

We base approximation schemes on *generalized differences* (see Istas and Lang, 1997). Let $a=(a_0,\dots,a_p)\in\mathbb{R}^{p+1}$ be such that for some positive integer $m(a)$

$$\text{for } k=0,\dots,m(a)-1:\quad \sum_{i=0}^{p}a_i i^k=0 \quad\text{and}\quad \sum_{i=0}^{p}a_i i^{m(a)}\neq 0. \tag{7}$$

We define, for $i = 0, \dots, N - p - 1$, the generalized difference

$$\Delta_a \bar{X} := \sum_{j=0}^{p} a_j \bar{X}_{i+j}.$$

The integer $m(a)$ is called the order of the difference. For instance, the usual difference $a = (-1, 1)$ is of order 1, and $\tilde{a} = (1, -2, 1)$ is of order 2. Following Istas and Lang (1997), we choose a difference a with order greater or equal than 2.

2.1. An asymptotic result

2.1.1. Model

Let us be given on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ a fractional Brownian motion (W_t^H) with Hurst parameter $H \in (0, 1)$, that is, a centred Gaussian process with covariance

$$\mathbb{E}\{W_t^H W_s^H\} = \frac{c_H}{2}(t^{2H} + s^{2H} - |t - s|^{2H}), \tag{8}$$

where $c_H = \Gamma(2 - 2H)\cos(\pi H)/\pi H(1 - 2H)$. ($\Gamma$ is the Euler function.) The process (W_t^H) has stationary increments, that easily give its covariance structure: for all h, t and t' in $\mathbb{R}$, we have

$$\mathbb{E}\{(W_{t+h} - W_t)(W_{t'+h} - W_{t'})\}$$
$$= \frac{c_H}{2}(|t' + h - t|^{2H} + |t + h - t'|^{2H} - 2|t' - t|^{2H}). \tag{9}$$

We further restrict the parameter t to the time interval $[0, 1]$. We are interested in the asymptotic behaviour of functionals of the type

$$\bar{V}_N(h) := N^{-1} \sum_{i=0}^{N-p-1} \frac{(\Delta_a \bar{X}_i)^2}{N^{-2H}} h(\bar{X}_i) \tag{10}$$

for (X_t) of the form $X_t = f(W_t^H)$, with smooth enough f and test function h. For $h = 1$, $\bar{V}_N(h)$ appears as a kind of renormalized quadratic variation, based on discrete local averages.

2.1.2. Convergence result

Theorem 3. *Let* $H \in (\frac{1}{2}, 1)$ *and*

$$V(h) := \kappa^2 \int_0^1 f'(W_s^H)^2 h(X_s)\, \mathrm{d}s,$$

where $\kappa = \kappa_{a,H}$ *is an explicit constant given in* (20). *Assume that* $f \in \mathscr{C}_{\mathscr{P}}^2$ *and* $h \in \mathscr{C}_{\mathscr{P}}^1$. *The sequence of random variables*

$$\{N^{1/2}[\bar{V}_N(h) - V(h)]\}_{N \geqslant 1}$$

is bounded in $L^1(\mathbb{P})$.

Remark. For appropriate coefficients, extension to fractional Brownian diffusions of the form

$$\mathrm{d}X_t = b(X_t)\, \mathrm{d}W_t^H + \mu(X_t)\, \mathrm{d}t, \quad X_0 = x_0 \tag{11}$$

is then possible, defining f as an inverse of $x \rightsquigarrow \int_{x_0}^{x} b(u)^{-1}\,\mathrm{d}u$ and removing the drift by a change of probability, but this lies beyond the scope of the paper. Recommended references for stochastic calculus with fBM are Decreusefond and Üstünel (1999), Carmona et al. (2003) and the references therein.

3. Construction of an estimator

Using Section 2, we are able to construct a contrast estimator that achieves the optimal rate in model (1), (2). Recall that we observe $Y_{i/n}$ for $i=0,\dots,n$ with

$$Y_t = y_0 + \int_0^t \Phi(\theta, W_s^H)\,\mathrm{d}W_s.$$

Consistently with the previous section, we define

$$X_t := \sigma_t^2 = \Phi(\theta, W_t^H)^2.$$

By Assumptions A and B, the process (X_t) solves the stochastic differential equation (11) with $b = b(\theta,\cdot)$ as defined in Assumption B and $\mu = 0$, but we will not exploit this fact further on.

3.1. Preliminaries

Pick an integer $N = N_n < n$, to be specified later, and define $j_i := [in/N_n]$ for $i = 0,\dots,N_n-1$. Set $m = m_n := [n/N_n] \geqslant 1$, where $[x]$ stands for the integer part of x. Clearly, $j_{i+1} - j_i \in \{m_n, m_n+1\}$.

For $i = 0,\dots,N_n-1$, define

$$\hat{X}_i = \hat{X}_{i,n} := N_n \sum_{k=0}^{j_{i+1}-j_i-1} \left(Y_{(j_i+k+1)/n} - Y_{(j_i+k)/n}\right)^2.$$

The $\hat{X}_i$'s may be seen as approximation variables for $\bar{X}_i$ based on data $Y_{i/n}$, $i=0,\dots,n$.

It is tempting to introduce an approximate quadratic variation like (10) when $\Delta_a \bar{X}_i$ is replaced by $\Delta_a \bar{X}_i$. However, such a quantity presents a bias toward its expected limit, and we need to introduce an appropriate correction term. We define the unbiased quadratic variation

$$\hat{V}_n(h) := N_n^{-1} \sum_{i=0}^{N_n-p-1} \left\{ \frac{(\Delta_a \hat{X}_i)^2}{N_n^{-2H}} - \frac{2\|a\|_2^2}{N_n^{-2H} m_n} \hat{X}_i^2 \right\} h(\hat{X}_i), \tag{12}$$

where $\|a\|_2$ is the Euclidean norm of the vector a that specifies the generalized difference.

3.2. The estimator

We introduce a contrast function based on the approximation of X_{i/N_n} by $\hat{X}_i$. But it may well happen that $\hat{X}_i \notin \mathscr{X}$ although X takes values in $\mathscr{X}$. We therefore need a smooth continuation of $b(\theta,\cdot)$ outside $\mathscr{X}$. Pick a compact interval $\mathscr{X}_0$ such that

$$\Phi^2(\Theta,0) \subset \dot{\mathscr{X}}_0 \subset \mathscr{X}_0 \subset \mathscr{X}$$

and let ψ be a $\mathscr{C}^\infty$ function on $\mathbb{R}$ such that $\psi(x) > 0$ for $x \in \dot{\mathscr{X}}_0$ and $\psi(x) = 0$ for $x \notin \mathscr{X}_0$. We define a family of contrast functions

$$(C_n(\eta), \eta \in \Theta) = (C_{n,N_n}(\eta), \eta \in \Theta)$$

by setting

$$C_n(\eta) := \hat{V}_n\left(\frac{\psi(\cdot)}{b(\eta,\cdot)^2}\right) + N_n^{-1} \sum_{i=0}^{N_n-p-1} \kappa^2 \log b(\eta,\hat{X}_i)^2 \cdot \psi(\hat{X}_i), \tag{13}$$

where $\kappa = \kappa_{a,H}$ is the constant given in (20). Eventually, our estimator of θ is

$$\hat{\theta}_n := \operatorname*{arginf}_{\eta\in\Theta} C_n(\eta), \tag{14}$$

specified by the tuning parameter

$$N_n := [n^{1/(2H+1)}]. \tag{15}$$

4. Proof of Theorem 3

4.1. The case of fBM

In this part, we are interested in integrated functionals of the trajectory of (W_t^H) by means of local averages of the form

$$\bar{W}_i^H = \bar{W}_{i,N}^H := N \int_{i/N}^{(i+1)/N} W_s^H \, \mathrm{d}s, \quad i = 0,\ldots,N-1 \tag{16}$$

for a given integer $N \geqslant 1$. The link with the statistical model will be made later by taking $N = N_n$, where N_n is the tuning parameter of the estimator $\hat{\theta}_n$ of Section 3. Given a generalized difference

$$\Delta_a \bar{W}_i^H := \sum_{j=0}^{p} a_j \bar{W}_{i+j}^H, \quad i = 0,\ldots,N-p-1,$$

we note that (7) implies

$$\sum_{0\leqslant k,l\leqslant p} a_k a_l \, |k-l|^{2q} = 0 \quad \text{for } q = 0,\ldots,m(a)-1 \tag{17}$$

and, for any odd q, by symmetry

$$\sum_{0\leqslant k,l\leqslant p} a_k a_l (k-l)^q = 0.$$

Before stating the main result of the section, we compute an expression for the covariance structure of the process

$$(\Delta_a \bar{W}_i^H)_{i=0,\ldots,N-p-1}.$$

Let us first define $v_H(0)=0$ and for $j \geqslant 1$

$$v_H(j) := \tilde{c}_H \left\{ (j+1)^{2H+2} - 2j^{2H+2} + (j-1)^{2H+2} - 2 \right\}, \tag{18}$$

where $\tilde{c}_H = c_H/(2H+1)(2H+2)$. Then for $j \leqslant -1$, we let $v_H(j) = v_H(-j)$. Set also, for $j \geqslant 0$

$$\rho_{a,H}(j) := -\frac{1}{2} \sum_{0 \leqslant k,l \leqslant p} a_k a_l v_H(k+j-l) \tag{19}$$

and for $j \leqslant 0$, $\rho_{a,H}(j) := \rho_{a,H}(-j)$. Finally, define

$$\kappa_{a,H} := \rho_{a,H}(0)^{1/2}. \tag{20}$$

Lemma 1. *We have*

$$\mathbb{E}\{\Delta_a \bar{W}_i^H \Delta_a \bar{W}_j^H\} = N^{-2H} \rho_{a,H}(i-j) \quad \textit{for } 0 \leqslant i,j \leqslant N-p-1.$$

In particular, $\kappa_{a,H}^2 = N^{2H}\mathbb{E}\{(\Delta_a \bar{W}_i^H)^2\} > 0$. The proof is given in Section A.2.

4.2. Quadratic variation for integrated fBM

For a real function h, let us define

$$\bar{Q}_N(h) := N^{-1} \sum_{i=0}^{N-p-1} \frac{(\Delta_a \bar{W}_i^H)^2}{N^{-2H}} h(\bar{W}_i^H), \tag{21}$$

$$Q(h) := \kappa_{a,H}^2 \int_0^1 h(W_s^H)\,\mathrm{d}s. \tag{22}$$

Proposition 1. *For $h \in \mathscr{C}_{\mathscr{P}}^1$, we have*

$$\mathbb{E}\{|\bar{Q}_N(h) - Q(h)|\} \leqslant cN^{-1/2}. \tag{23}$$

By looking at Lemma 5, we can anticipate and see that it is sufficient to establish the bound

$$\mathbb{E}\{|S_N(h)|\} \leqslant cN^{1/2}, \tag{24}$$

$$S_N(h) := \sum_{i=0}^{N-p-1} \left\{ \frac{(\Delta_a \bar{W}_i^H)^2}{N^{-2H}} - \kappa_{a,H}^2 \right\} h(W_{i/N}^H). \tag{25}$$

The main difficulty comes from the fact that stochastic calculus does not yield simple expansions for a quantity like $(\Delta_a \bar{W}_i^H)^2 h(W_{i/N}^H)$ for $H \neq \frac{1}{2}$. We need to develop specific tools in a first part.

4.2.1. A deterministic intermediate result

First, we treat the simpler case of $t \rightsquigarrow h(W_t^H)$ being replaced by a deterministic function ξ:

Lemma 2. *Let* $\xi : [0,1] \to \mathbb{R}$ *be a bounded function. Define*

$$\Sigma_N(\xi) = \sum_{i=0}^{N-p-1} \left\{ \frac{(\Delta_a \bar{W}_i^H)^2}{N^{-2H}} - \kappa_{a,H}^2 \right\} \xi_{i/N}. \tag{26}$$

Then

$$\mathbb{E}\{\Sigma_N(\xi)^2\} \leqslant c\|\xi\|_\infty^2 N.$$

Proof. The expectation $\mathbb{E}\{\Sigma_N(\xi)^2\}$ is equal to

$$\sum_{0 \leqslant i,j \leqslant N-p-1} \xi_{i/N}\xi_{j/N}\mathbb{E}\left\{ \left[\frac{(\Delta_a \bar{W}_i^H)^2}{N^{-2H}} - \kappa_{a,H}^2 \right] \left[\frac{(\Delta_a \bar{W}_j^H)^2}{N^{-2H}} - \kappa_{a,H}^2 \right] \right\}.$$

For two centred normal variables Z, Z' with same variance v and covariance k, we have $\mathbb{E}\{(Z^2 - v)(Z'^2 - v)\} = 2k^2$. By Lemma 1, we get

$$\mathbb{E}\{\Sigma_N(\xi)^2\} = \sum_{0 \leqslant i,j \leqslant N-p-1} \xi_{i/N}\xi_{j/N} 2\rho_{a,H}(j-i)^2 \leqslant 2\|\xi\|_\infty^2 N \sum_{j=0}^{N} \rho_{a,H}(j)^2.$$

Therefore, the remainder of the proof amounts to show that the series $\sum_{j=0}^{\infty} \rho_{a,H}(j)^2$ converges. Let $j \geqslant p+2$. Using that $\sum_{0 \leqslant k,l \leqslant p} a_k a_l = 0$ together with (18) we can rewrite (19) as

$$\rho_{a,H}(j) = \sum_{0 \leqslant k,l \leqslant p} a_k a_l \sum_{m=0}^{2} \tilde{a}_m w(k + j - l + m - 1)$$

with $\tilde{a} = (\tilde{a}_0, \tilde{a}_1, \tilde{a}_2) = (1, -2, 1)$ and $w(x) = -c_H[2(2H+1)(2H+2)]^{-1}|x|^{2H+2}$. Since $j \geqslant p+2$, we avoid the singularity of w at zero, and by a Taylor expansion of order 5 around $k + j - l - 1$, we have

$$w(k + j - l + m - 1) = \sum_{r=0}^{4} \frac{w^{(r)}}{r!}(k + j - l - 1)m^r + m^5 \int_0^1 \frac{(1-\eta)^4}{4!} w^{(5)}((k + j - l - 1) + m\eta)\,\mathrm{d}\eta.$$

Using that $\tilde{a}$ is a generalized difference of order 2—recall (7)—we deduce

$$\rho_{a,H}(j) = \sum_{0 \leqslant k,l \leqslant p} a_k a_l \sum_{m=0}^{2} \tilde{a}_m \sum_{r=2}^{4} \frac{w^{(r)}}{r!}(k + j - l - 1)m^r + \varepsilon_j, \tag{27}$$

where

$$\varepsilon_j = \sum_{0\leqslant k,l\leqslant p} a_k a_l \sum_{m=0}^{2} \tilde{a}_m m^5 \int_0^1 \frac{(1-\eta)^4}{4!} w^{(5)}((k+j-l-1)+m\eta)\,\mathrm{d}\eta.$$

The expression of $w^{(5)}$ clearly implies that $|\varepsilon_j| \leqslant c(j-p-1)^{2H-3}$. Hence, $\sum_{j=p+2}^{\infty} |\varepsilon_j|^2 < \infty$. Now, by (27), it is sufficient to establish the convergence of the squared sum of

$$\rho_{a,H}^{(r)}(j) := \sum_{0\leqslant k,l\leqslant p} a_k a_l w^{(r)}(k-l+j-1) \quad \text{for } r = 2,3,4.$$

We fix $r\in\{2,3,4\}$. Writing again a Taylor expansion of order 3 around $(j-1)$ for $w^{(r)}(k-l+j-1)$, we get

$$w^{(r)}(k-l+j-1) = \sum_{u=0}^{2} \frac{w^{(r+u)}}{u!}(j-1)(k-l)^u$$
$$+(k-l)^3 \int_0^1 \frac{(1-\eta)^2}{2} w^{(r+3)}(j-1+(k-l)\eta)\,\mathrm{d}\eta.$$

By (17) and the choice $m(a) \geqslant 2$,

$$\rho_{a,H}^{(r)}(j) = \sum_{0\leqslant k,l\leqslant p} a_k a_l (k-l)^3 \int_0^1 \frac{(1-\eta)^2}{2} w^{(r+3)}(j-1+(k-l)\eta)\,\mathrm{d}\eta.$$

But $r+3 \geqslant 5$, so $|\rho_{a,H}^{(r)}(j)| \leqslant c(j-1-p)^{2H-3}$ and $\sum_{j=p+2}^{\infty} \rho_{a,H}^{(r)}(j)^2 < \infty$ for all $H \in (\frac{1}{2}, 1)$. □

4.2.2. *Expansion in the Schauder basis*

The proof of (24) relies on the fact that $S_n(h)$ is linear with respect to $t \rightsquigarrow h(W_t^H)$. In order to exploit this linearity, we expand (W_t^H) in a wavelet basis with random coefficients, following the approach of Ciesielski et al. (1993), see Section A.1.2. We then prove (24) by applying a slight modification of the previous lemma to the low-frequency part of the expansion.

Lemma 3. *Assume that* $\xi : [0,1] \to \mathbb{R}$ *is bounded and vanishes outside the interval* $[k2^{-j_0}, k'2^{-j_0}] \subset [0,1]$, *where* $k, k', j_0 \geqslant 1$, $k \neq k'$. *Then, there exists* $c > 0$ *such that for any* N, j_1 *with* $2^{j_1} \leqslant N < 2^{j_1+1}$ *and* $j_1 \geqslant j_0$, *we have*

$$\mathbb{E}\{\Sigma_N(\xi)^2\} \leqslant c\|\xi\|_\infty^2[(k'-k)2^{j_1-j_0}].$$

Proof. Let $\mathscr{I}_N := \{i : k2^{j_1-j_0} \leqslant i \leqslant (k'2^{j_1+1-j_0}) \wedge (N-p-1)\}$. Since $\xi(t) = 0$ for t outside $[k2^{-j_0}, k'2^{-j_0}]$ and $2^{j_1} \leqslant N < 2^{j_1+1}$, we have

$$\Sigma_N(\xi) = \sum_{i\in\mathscr{I}_N} \left[\frac{(\Delta_a \bar{W}_i^H)^2}{N^{-2H}} - \kappa_{a,H}^2\right] \xi_{i/N}.$$

But the stationarity of $(\Delta_a \bar{W}_i^H)_i$—see Lemma 1—implies that $\Sigma_N(\xi)$ has the same law as

$$\sum_{i\in\mathscr{J}_N}\left[\frac{(\Delta_a\bar{W}_i^H)^2}{N^{-2H}}-\kappa_{a,H}^2\right]\xi_{k2^{j_1-j_0}/N+i/N},$$

where $\mathscr{J}_N := \{i:\ 0\leqslant i\leqslant (k'2^{j_1+1-j_0}-k2^{j_1-j_0})\wedge(N-p-1-k2^{j_1-j_0})\}$. Therefore, an application of Lemma 2 gives

$$\begin{aligned}\mathbb{E}\{\Sigma_N(\xi)^2\} &\leqslant c\|\xi\|_\infty^2\,(k'2^{j_1+1-j_0}-k2^{j_1-j_0}+p-1)\\ &\leqslant 4(p-1)c\|\xi\|_\infty^2(k'-k)2^{j_1-j_0}.\end{aligned}$$

Whence the lemma, up to a modification of c. □

Lemma 4. *Let $h\in\mathscr{C}^1_{\mathscr{P}}$. Denote by c_0, c_1, and $c_{j,k}$ the (random) coefficients of the expansion in the Schauder basis of $t\rightsquigarrow h(W_t^H)$, see Section* A.1.2. *There exists $c>0$ such that*

$$\mathbb{E}\{c_0^2+c_1^2\}\leqslant c,\quad \mathbb{E}\{c_{j,k}^2\}\leqslant c2^{-j(1+2H)}. \tag{28}$$

Proof. The first part of (28) is obvious. For the second part, we just write a Taylor expansion for (61) and obtain, thanks to the assumption $|h'(x)|\leqslant K(1+|x|^K)$ for some $K>0$ that

$$\begin{aligned}|c_{j,k}| &\leqslant 2^{-j/2+1}K(1+\sup_{s\in[0,1]}|W_s^H|^K)\\ &\quad\cdot(|W^H_{(2k-1)/2^{j+1}}-W^H_{2k/2^{j+1}}|+|W^H_{(2k-1)/2^{j+1}}-W^H_{(2k-2)/2^{j+1}}|).\end{aligned}$$

Existence of moments of any order for the supremum of fBM over the time interval $[0,1]$ (e.g. Ciesielski et al., 1993) and (9) yield the second part of (28). □

4.2.3. Proof of Proposition 1

We are now ready to prove (24). The pointwise representation

$$h(W_t^H)=c_0\phi_0(t)+c_1\phi_1(t)+\sum_{j=0}^{\infty}\sum_{k=1}^{2^j}c_{j,k}\phi_{j,k}(t) \tag{29}$$

yields

$$S_N(h)=c_0\Sigma_N(\phi_0)+c_1\Sigma_N(\phi_1)+\sum_{j=0}^{\infty}S_{N,j},\quad S_{N,j}:=\sum_{k=1}^{2^j}c_{j,k}\Sigma_N(\phi_{j,k}).$$

By Lemma 2, for $i=0,1$:

$$\mathbb{E}\{|\Sigma_N(\phi_i)|^2\}\leqslant c\|\phi_i\|_\infty^2 N\leqslant cN.$$

By Cauchy–Schwarz inequality and (28)

$$\mathbb{E}\{|c_i\Sigma_N(\phi_i)|\}\leqslant cN^{1/2}.$$

It remains to show that

$$\mathbb{E}\left\{\sum_{j=0}^{\infty}|S_{N,j}|\right\}=\sum_{j=0}^{\infty}\mathbb{E}\{|S_{N,j}|\}\leqslant cN^{1/2}. \tag{30}$$

Fix $N\geqslant 1$ and let j_1 be the unique integer such that $2^{j_1}\leqslant N<2^{j_1+1}$. Consider first the case $j\leqslant j_1$. Using Lemma 4, we get

$$\mathbb{E}\{|S_{N,j}|\}\leqslant c(K)\sum_{k=1}^{2^j}2^{-(j/2)(1+2H)}\mathbb{E}\{\Sigma_N(\phi_{j,k})^2\}^{1/2}.$$

Using Lemma 3 with an appropriate bound on $\|\phi_{j,k}\|_\infty$ and on the support of $\phi_{j,k}$, we obtain

$$\mathbb{E}\{|S_{n,j}|\}\leqslant c\sum_{k=1}^{2^j}2^{-(j/2)(1+2H)}2^{j/2}2^{(j_1-j)/2}=c2^{j_1/2}2^{(1/2-H)j}.$$

Hence, using $H>\frac{1}{2}$, the sum $\sum_{j=0}^{j_1}\mathbb{E}\{|S_{N,j}|\}$ is less than

$$c(K)2^{j_1/2}\sum_{j=0}^{j_1}2^{(1/2-H)j}\leqslant c(K)2^{j_1/2}\leqslant c(K)N^{1/2}. \tag{31}$$

Consider next the case $j>j_1$. The function $\phi_{j,k}$ is supported on $[(k-1)/2^j,k/2^j]$, hence $\Sigma_N(\phi_{j,k})=0$ unless there exists some $i\in\{0,\dots,N-p-1\}$ satisfying $i/N\in[(k-1)/2^j,k/2^j]$. Since the length of $[(k-1)/2^j,k/2^j]$ is 2^{-j}, which is less than $2^{-j_1-1}<N^{-1}$, there is at most N values of $k\in\{1,\dots,2^j\}$ such that $\Sigma_N(\phi_{j,k})\neq 0$. Moreover, for such a k (with $\Sigma_N(\phi_{j,k})\neq 0$), the sum defining $\Sigma_N(\phi_{j,k})$—see (26)—reduces to a single term, therefore

$$\mathbb{E}\{\Sigma_N(\phi_{j,k})^2\}\leqslant c\|\phi_{j,k}\|_\infty^2\leqslant c2^j.$$

We derive

$$\begin{aligned}\mathbb{E}\{|S_{N,j}|\}&\leqslant\sum_{\substack{1\leqslant k\leqslant 2^j\\ \Sigma_N(\phi_{j,k})\neq 0}}\mathbb{E}\{c_{j,k}^2\}^{1/2}\mathbb{E}\{\Sigma_N(\phi_{j,k})^2\}^{1/2}\\&\leqslant c\sum_{\substack{1\leqslant k\leqslant 2^j\\ \Sigma_N(\phi_{j,k})\neq 0}}2^{-j(1/2+H)}2^{j/2}\leqslant N2^{-jH}.\end{aligned}$$

Summing these inequalities for $j\geqslant j_1+1$ implies, using again $H\geqslant\frac{1}{2}$, that the sum $\sum_{j\geqslant j_1+1}\mathbb{E}\{|S_{N,j}|\}$ is less than

$$cN2^{-j_1H}\leqslant cN2^{-j_1/2}\leqslant cN^{1/2}. \tag{32}$$

Putting together (31) and (32), we obtain (30) whence (24). The proof of Proposition 1 is complete.

We end this section by gathering previously used estimates that will prove useful in the following.

Lemma 5. *Let $h \in \mathscr{C}^1_{\mathscr{P}}(K)$. The following quantities are bounded by $cN^{-1/2}$ in $L^1(\mathbb{P})$:*

(i) $N^{-1+2H} \sum_{i=0}^{N-p-1} \left(\Delta_a \bar{W}_i^H\right)^2 \left[h(\bar{W}_i^H) - h(W_{i/N}^H)\right]$,

(ii) $N^{-1} \sum_{i=0}^{N-p-1} \left[h(\bar{W}_i^H) - h(W_{i/N}^H)\right]$,

(iii) $N^{-1} \sum_{i=0}^{N-p-1} h(W_{i/N}^H) - \int_0^1 h(W_s)\,\mathrm{d}s$.

Proof. Using the regularity of h, the two first properties follow from $H \geqslant \frac{1}{2}$ and the boundedness of the kth moment of $\bar{W}_i^H - W_{i/N}^H$. The third property easily follows from (9). □

4.3. Proof of Theorem 3

We first establish the following two expansions:

$$\Delta_a \bar{X}_i = f'(\bar{W}_i^H)\Delta_a \bar{W}_i^H + e_{i,N}, \tag{33}$$

where, for $k \geqslant 1$

$$\mathbb{E}\{|e_{i,N}|^k\} \leqslant cN^{-2HK} \tag{34}$$

and

$$h(\bar{X}_i) = h(f(\bar{W}_i^H)) + e'_{i,N}$$

with, for $k \geqslant 1$

$$\mathbb{E}\{|e'_{i,N}|^k\} \leqslant cN^{-HK}. \tag{35}$$

For (34), using that $m(a) \geqslant 1$, we write

$$\Delta_a \bar{X}_i = \sum_{j=0}^{p} a_j N \int_{(i+j)/N}^{(i+j+1)/N} \left\{f(W_s^H) - f(\bar{W}_i^H)\right\} \mathrm{d}s.$$

A second-order expansion for f yields

$$\Delta_a \bar{X}_i = \sum_{j=0}^{p} a_j N \int_{(i+j)/N}^{(i+j+1)/N} (W_s^H - \bar{W}_i^H)\,\mathrm{d}s\, f'(\bar{W}_i^H) + e_{i,N}, \tag{36}$$

where the remainder term $e_{i,N}$ fullfills (34). Using again that $m(a) \geqslant 1$, Eq. (36) reduces to (33). The second expansion (35) follows easily from the regularity of f and h. Plugging (34) and (35) in (10)—recall (21)—gives

$$\bar{V}_N(h) = Q_N((f')^2 \times h \circ f) + e_N^{(1)} + e_N^{(2)},$$

where

$$e_N^{(1)} = N^{-1}N^{2H}\sum_{i=0}^{N-p-1}(e_{i,N}^2 + 2\Delta_a \bar{W}_i^H f'(\bar{W}_i^H)e_{i,n})h(\bar{X}_i),$$

$$e_N^{(2)} = N^{-1}N^{2H}\sum_{i=0}^{N-p-1}(\Delta_a \bar{W}_i^H)^2 e'_{i,N}.$$

Simple computations based on Eqs. (34) and (35) show that the expectation of $|e_N^{(1)}|$ and $|e_N^{(2)}|$ are less than $cN^{-H} \leqslant cN^{-1/2}$. We conclude by applying Proposition 1.

5. Proof of Theorem 1

5.1. Preliminaries

Keeping up with the notation of Section 2, we set

$$X_t := \sigma_t^2 = \Phi(\theta, W_t^H)^2,$$

$$V_\theta(h) := \kappa^2 \int_0^1 [\partial_x \Phi^2(\theta, W_s^H)]^2 h(X_s)\,\mathrm{d}s = \kappa^2 \int_0^1 b(\theta, X_s)^2 h(X_s)\,\mathrm{d}s.$$

Although we basically work with the law $\mathbb{P}_\theta^n$ of the data $Y_{i/n}$, it may well happen that we need to consider functionals of the whole trajectory of (X_t) in the limit. Therefore, we denote by $\mathbb{P}_\theta$ the joint law of $(Y_{i/n}, X_t,\ i = 0,\dots,n,\ t \in [0,1])$ on an appropriate probability space.

5.1.1. A first approximation result

Proposition 2. *Consider a parametrization $(h(\eta,\cdot), \eta \in \Theta)$ such that h is in $\mathscr{C}_{\mathscr{P}}^1$ as a function of two variables. Then*

$$\lim_{n\to\infty} \sup_{\theta\in\Theta} \mathbb{E}_\theta \left\{ \sup_{\eta\in\Theta} |\hat{V}_n(h(\eta,\cdot)) - V_\theta(h(\eta,\cdot))| \right\} = 0.$$

Proof. We first compute the difference between $\hat{X}_i$ and $\bar{X}_i$ for large enough m_n. Since the expression of conditional law of Y given X is fairly simple, we introduce the σ-field $\mathscr{G}_X = \sigma(W_s^H, 0 \leqslant s \leqslant 1)$. We rely on the following technical lemma

Lemma 6. *We have $\hat{X}_i = \bar{X}_i + E_{i,n}$, where the error terms $E_{i,n}$ satisfy:*

(i) *Conditional on $\mathscr{G}_X$, the random variables $(E_{i,n})_{i=0,\dots,N-p-1}$ are independent under $\mathbb{P}_\theta$.*

(ii)
$$|\mathbb{E}_\theta[E_{i,n}|\mathscr{G}_X]| \leqslant cm_n^{-1}\|X\|_\infty, \quad |\mathbb{E}_\theta\{E_{i,n}\}| \leqslant cm_n^{-1}, \tag{37}$$

$$\mathbb{E}_\theta[E_{i,n}^2|\mathscr{G}_X] \leqslant cm_n^{-1}\|X\|_\infty^2, \quad \mathbb{E}_\theta\{E_{i,n}^2\} \leqslant cm_n^{-1}, \tag{38}$$

$$\mathbb{E}_\theta[E_{i,n}^4|\mathscr{G}_X] \leqslant cm_n^{-2}\|X\|_\infty^4, \quad \mathbb{E}_\theta\{E_{i,n}^4\} \leqslant cm_n^{-2}. \tag{39}$$

(iii) $\mathbb{E}_\theta[E_{i,n}^2|\mathscr{G}_X] = 2m_n^{-1}X_{i/N_n}^2 + \alpha_{i,n}$,

where for all $l \geqslant 1$

$$\mathbb{E}_\theta\{|\alpha_{i,n}|^l\} \leqslant c\{m_n^{-l}N_n^{-Hl} + m_n^{-2l}\}.$$

(iv) *For all* $l \geqslant 1$: $\mathbb{E}_\theta\{|\hat{X}_i|^l\} \leqslant c$.

Proof. By Itô's formula

$$\hat{X}_i = \tilde{E}_{i,n} + N_n \int_{j_i/n}^{j_{i+1}/n} X_s \, \mathrm{d}s$$

with $\tilde{E}_{i,n} = N_n \sum_{k=0}^{j_{i+1}-j_i} \int_{(j_i+k)/n}^{(j_i+k+1)/n} (Y_s - Y_{(j_i+k)/n})\sigma_s \, \mathrm{d}W_s$, thus

$$E_{i,n} = \tilde{E}_{i,n} + N_n \int_{j_i/n}^{i/N_n} X_s \, \mathrm{d}s + N_n \int_{(i+1)/N_n}^{j_{i+1}/n} X_s \, \mathrm{d}s.$$

We readily derive (i). The sum of the two integrals above is the conditional expectation of $E_{i,n}$ and is clearly bounded by $2m_n^{-1}\|X\|_\infty$ whence the first part of (37). The second part follows from the regularity of Φ and the integrability of $\sup_{t\in[0,1]}|W_t^H|$. Eqs. (38), (39) and (iii) are obtained by direct computation on $\tilde{E}_{i,n}$, the remainder terms proving negligible; the last point (iv) is straightforward. □

Following Lemma 6, we write $\Delta_a\hat{X}_i = \Delta_a\bar{X}_i + \Delta_a E_{i,n}$. We plan to use the decomposition

$$\hat{V}_n(h) = \bar{V}_{N_n}(h) + \sum_{r=1}^{4} B_n^{(r)}(h)$$

with

$$B_n^{(1)}(h) = N_n^{-1+2H} \sum_{i=0}^{N_n-p-1} \left\{(\Delta_a E_{i,n})^2 - 2\|a\|_2^2 m_n^{-1} X_{i/N_n}^2\right\} h(\bar{X}_i),$$

$$B_n^{(2)}(h) = N_n^{-1+2H} \sum_{i=0}^{N_n-p-1} \left\{(\Delta_a \hat{X}_i)^2 - 2\|a\|_2^2 m_n^{-1} \hat{X}_i^2\right\} \left\{h(\hat{X}_i) - h(\bar{X}_i)\right\},$$

$$B_n^{(3)}(h) = N_n^{-1+2H} m_n^{-1} \sum_{i=0}^{N_n-p-1} 2\|a\|_2^2 \left\{X_{i/N_n}^2 - \hat{X}_i^2\right\},$$

$$B_n^{(4)}(h) = N_n^{-1+2H} \sum_{i=0}^{N_n-p-1} 2(\Delta_a E_{i,n})(\Delta_a \bar{X}_i) h(\bar{X}_i).$$

By Theorem 3, if we can prove show that for $r = 1, \dots, 4$

$$\mathbb{E}\{|B_n^{(r)}(h)|\} \leqslant cv(N_n, m_n) := (N_n^{-1/2} + m_n^{-1/2})(1 + m_n^{-1} N_n^{2H}),$$

which is of order $N_n^{-1/2}$, we will obtain

$$\forall \eta \in \Theta, \quad \sup_{\theta \in \Theta} \mathbb{E}_\theta\{|\hat{V}_n(h(\eta, \cdot)) - V_\theta(h(\eta, \cdot))|\} \leqslant cN^{-1/2}. \tag{40}$$

For $B_n^{(1)}(h)$, we set

$$b_i = N_n^{2H} \left\{ (\Delta_a E_{i,n})^2 - 2\|a\|_2^2 m_n^{-1} X_{i/N_n}^2 \right\} h(\bar{X}_i).$$

Remark that by Lemma 6(i), conditional on $\mathscr{G}_X$, b_i and b_j are independent if $|i - j| \geqslant p + 1$. Thus the conditional expectation of $B_n^{(1)}(h)^2$ given $\mathscr{G}_X$ is equal to

$$N_n^{-2} \sum_{\substack{0 \leqslant i,j \leqslant N_n - p - 1 \\ |i-j| \leqslant p}} \mathbb{E}_\theta[b_i b_j | \mathscr{G}_X] + N_n^{-2} \sum_{\substack{0 \leqslant i,j \leqslant N_n - p - 1 \\ |i-j| \geqslant p+1}} \mathbb{E}_\theta[b_i | \mathscr{G}_X] \mathbb{E}_\theta[b_j | \mathscr{G}_X]. \tag{41}$$

Using Cauchy–Schwarz and (39), we have that $\mathbb{E}\{|b_i b_j|\} \leqslant cN_n^{4H} m_n^{-2}$. So we may bound the expectation of the first sum in (41) by $cpN_n N_n^{4H} m_n^{-2} \leqslant cN_n^2 v(N_n, m_n)^2$.

By Lemma 6(i)–(iii), the conditional expectation of b_i given $\mathscr{G}_X$ is equal to

$$\left\{ 2\sum_{l=0}^{p} a_l^2 m_n^{-1} X_{(i+l)/N_n}^2 - 2\|a\|_2^2 m_n^{-1} X_{i/N_n}^2 + \sum_{l=0}^{p} a_l^2 \alpha_{i+l,n} + \beta_{i,n} \right\} N_n^{2H} h(\bar{X}_i),$$

where $|\beta_{i,n}| \leqslant cm_n^{-2} \|X\|_\infty^2$. It follows that

$$\mathbb{E}_\theta\{\mathbb{E}_\theta[b_i | \mathscr{G}_X]^2\} \leqslant c\{m_n^{-2} N_n^{2H} + m_n^{-4} N_n^{4H}\}.$$

Therefore, the expectation of the second sum appearing in (41) is bounded by $cN_n^2 \cdot \{m_n^{-2} N_n^{2H} + m_n^{-4} N_n^{4H}\} \leqslant cN_n^2 v(N_n, m_n)^2$. We conclude that

$$\mathbb{E}_\theta\{B_n^{(1)}(h)^2\} \leqslant cv(N_n, m_n)^2.$$

For the remaining terms, by similar computation based on Lemma 6(ii), (iii) and (vi), we obtain, for $r = 2, 3, 4$

$$\mathbb{E}_\theta\{|B_n^{(r)}(h)|\} \leqslant cv(N_n, m_n).$$

Therefore, we have (40) and only the uniformity in Θ within the expectation requires a proof. One can easily check that the two following conditions are valid and that they imply the result since Θ is convex and compact:

$$\limsup_n \sup_{\theta \in \Theta} \left(\mathbb{E}_\theta \left\{ \sup_{\eta \in \Theta} |\partial_\eta \hat{V}_n(h(\eta, \cdot))| \right\} \vee \mathbb{E}_\theta \left\{ \sup_{\eta \in \Theta} |\partial_\eta V_\theta(h(\eta, \cdot))| \right\} \right) < \infty. \qquad \square$$

5.1.2. Consistency

We heavily rely on Proposition 2, Lemmas 5 and 6 to derive the following estimates: First, if h is in $\mathscr{C}^1_{\mathscr{P}}$,

$$\limsup_n \sup_{\theta\in\Theta} \mathbb{E}_\theta\{N_n^{1/2}|\hat{V}_n(h) - V_\theta(h)|\} < \infty. \tag{42}$$

Next, set $I_n(h) := N_n^{-1}\sum_{i=0}^{N_n-p-1} h(\hat{X}_i)$ and $I(h) := \int_0^1 h(X_s)\,\mathrm{d}s$. We have

$$\sup_n \sup_{\theta\in\Theta} \mathbb{E}_\theta\{N_n^{1/2}|I_n(\log b(\eta,\cdot)^2\cdot\psi(\cdot)) - I(\log b(\eta,\cdot)^2\cdot\psi(\cdot))|\} < \infty \tag{43}$$

and

$$\lim_{n\to\infty} \sup_{\theta\in\Theta} \mathbb{E}_\theta\{\sup_{\eta\in\Theta}|I_n(\log b(\eta,\cdot)^2\cdot\psi(\cdot)) - I(\log b(\eta,\cdot)^2\cdot\psi(\cdot))|\} = 0. \tag{44}$$

These three properties imply the convergence in $\mathbb{P}_\theta$-probability, uniformly in $\theta\in\Theta$ of the function $C_n(\eta)$—recall (13)—to the limit

$$C_\theta(\eta) := \kappa^2\int_0^1 \left(\frac{b(\theta,X_s)^2}{b(\eta,X_s)^2} + \log b(\eta,X_s)^2\right)\psi(X_s)\,\mathrm{d}s. \tag{45}$$

The proof of the consistency of $\hat{\theta}_n$, that is

$$\forall\epsilon > 0,\quad \lim_{n\to\infty}\sup_{\theta\in\Theta}\mathbb{P}^n_\theta\{|\hat{\theta}_n - \theta| \geqslant \epsilon\} = 0 \tag{46}$$

follows from (45) in a classical way. We shall not pursue it here.

5.2. Proof of Theorem 1, completion

Note first that $N_n^{1/2}$ is of order $n^{1/(4H+2)}$, i.e. the rate of convergence. Let U be an open set in Θ. Since the distance d_U between U and Θ^c is positive, by (46), we have

$$\lim_{n\to\infty}\sup_{\theta\in U}\mathbb{P}^n_\theta\{|\hat{\theta}_n - \theta| > \tfrac{1}{2}d_U\} = 0.$$

Next, we use that for $\lambda > 0$, the set $\{N_n^{1/2}|\hat{\theta}_n - \theta| \geqslant \lambda\}$ is included in

$$\left(\{N_n^{1/2}|\hat{\theta}_n - \theta| \geqslant \lambda\}\cap\{\hat{\theta}_n\in(\theta_-,\theta_+)\}\right)\cup\{|\hat{\theta}_n - \theta| > \tfrac{1}{2}d_U\}.$$

This implies that $\limsup_{n\to\infty}\sup_{\theta\in U}\mathbb{P}^n_\theta\{N_n^{1/2}|\hat{\theta}_n - \theta| \geqslant \lambda\}$ is less than

$$\limsup_{n\to\infty}\sup_{\theta\in U}\mathbb{P}^n_\theta(\{N_n^{1/2}|\hat{\theta}_n - \theta| \geqslant \lambda\}\cap\{\hat{\theta}_n\in(\theta_-,\theta_+)\}).$$

Now, on the set $\{\hat{\theta}_n\in(\theta_-,\theta_+)\}$, by a first-order expansion around $\hat{\theta}_n$, we have for $\theta\in\Theta$

$$G_n(\theta)\left[N_n^{1/2}(\hat{\theta}_n - \theta)\right] = L_n(\theta)$$

with

$$L_n(\theta) := -N_n^{1/2}\dot{C}_n(\theta), \;\; G_n(\theta) := \int_0^1 \ddot{C}_n(\theta + (\hat{\theta}_n - \theta)u)\,\mathrm{d}u.$$

Therefore, we have, for all $c > 0$,

$$\mathbb{P}_\theta^n(\{N_n^{1/2}|\hat{\theta}_n - \theta| \geqslant \lambda\} \cap \{\hat{\theta}_n \in (\theta_-, \theta_+)\}) \leqslant \mathbb{P}_\theta^n\left\{|L_n(\theta)| \geqslant \frac{\lambda}{c}\right\} + \mathbb{P}_\theta^n\{G_n(\theta) \leqslant c\}.$$

The remainder of the proof amounts to show that for all $c > 0$:

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} \sup_{\theta\in U} \mathbb{P}_\theta^n\left\{|L_n(\theta)| \geqslant \frac{\lambda}{c}\right\} = 0 \tag{47}$$

and that for all $\epsilon > 0$, there exists c such that

$$\limsup_{n\to\infty} \sup_{\theta\in U} \mathbb{P}_\theta\{G_n(\theta) \leqslant c\} \leqslant \epsilon. \tag{48}$$

Straightforward computation show that

$$L_n(\theta) = -N_n^{1/2}\hat{V}_n\left[-2\,\frac{\dot{b}(\theta,\cdot)}{b^3(\theta,\cdot)}\psi(\cdot)\right] - N_n^{1/2}\kappa^2 I_n\left[2\,\frac{\dot{b}(\theta,\cdot)}{b(\theta,\cdot)}\,\psi(\cdot)\right].$$

Using (42) and (43),

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} \sup_{\theta\in U} \mathbb{P}_\theta^n\left\{N_n^{1/2}|N_n^{-1/2}L_n(\theta) - L(\theta)| \geqslant \frac{\lambda}{c}\right\} = 0,$$

where

$$L(\theta) := V_\theta\left[2\,\frac{\dot{b}(\theta,\cdot)}{b^3(\theta,\cdot)}\,\psi(\cdot)\right] - \kappa^2 I\left[2\,\frac{\dot{b}(\theta,\cdot)}{b(\theta,\cdot)}\,\psi(\cdot)\right] = 0.$$

(For notational convenience, we further abbreviate $\partial_\eta b(\eta,\cdot)$ by $\dot{b}(\eta,\cdot)$.) Therefore only (47) needs to be proved. For (48), we compute the second derivative of $\eta \rightsquigarrow \mathscr{C}_\theta(\eta)$:

$$\ddot{\mathscr{C}}_\theta(\eta) = V_\theta\left[\left(6\,\frac{\dot{b}^2(\eta,\cdot)}{b^4(\eta,\cdot)} - 2\,\frac{\ddot{b}(\eta,\cdot)}{b^3(\eta,\cdot)}\right)\psi(\cdot)\right] + \kappa^2 I\left[\left(-2\,\frac{\dot{b}^2(\eta,\cdot)}{b^2(\eta,\cdot)} + 2\,\frac{\ddot{b}(\eta,\cdot)}{b(\eta,\cdot)}\right)\psi(\cdot)\right].$$

In particular, for $\eta = \theta$, this quantity reduces to

$$\ddot{\mathscr{C}}_\theta(\theta) = 4\kappa^2 \int_0^1 \frac{\dot{b}^2}{b^2}(\theta, X_s)\psi(X_s)\,\mathrm{d}s.$$

By Assumption C, $\dot{b}^2/b^2$ is bounded below on $\Theta \times \mathscr{X}_0$. From this, using $\phi^2(\Theta, 0) \subset \dot{\mathscr{X}}_0$ on which ψ is strictly positive, it can be seen that there exists $c > 0$ such that

$$\sup_{\theta\in U} \mathbb{P}_\theta\{\ddot{\mathscr{C}}_\theta(\theta) \leqslant 3c\} \leqslant \epsilon.$$

Noting that $\ddot{\mathscr{C}}_\theta(\cdot)$ is Lipschitz continuous with a nonrandom Lipschitz constant independent of θ, there exists $\alpha > 0$ such that

$$\sup_{\theta \in U} \mathbb{P}_\theta \left\{ \inf_{|\eta-\theta| \leqslant \alpha} \ddot{\mathscr{C}}_\theta(\eta) \leqslant 2c \right\} \leqslant \epsilon. \tag{49}$$

Now, due to the expression of $G_n(\theta)$, we have that $\mathbb{P}_\theta^n\{G_n(\theta) \leqslant c\}$ is less than

$$\mathbb{P}_\theta \left\{ \inf_{|\eta-\theta| \leqslant \alpha} \ddot{\mathscr{C}}_\theta(\eta) \leqslant 2c \right\} + \mathbb{P}_\theta^n\{|\hat{\theta}_n - \theta| \geqslant \alpha\} + \mathbb{P}_\theta \left\{ \sup_{\eta \in \Theta} |\ddot{C}_n(\eta) - \ddot{\mathscr{C}}_\theta(\eta)| \geqslant c \right\}.$$

Using (46), (42), (44) and (49) we conclude that

$$\limsup_{n\to\infty} \sup_{\theta\in\Theta} \mathbb{P}_\theta\{G_n(\theta) \leqslant c\} \leqslant \epsilon.$$

6. Proof of Theorem 2

6.1. Preliminaries

For $H \in [\frac{1}{2}, 1)$, consider the kernel

$$K_H(t,s) = \frac{(t-s)^{H-1/2}}{\Gamma(H+\frac{1}{2})}\, {}_2F_1\big(\tfrac{1}{2}-H, H-\tfrac{1}{2}, H+\tfrac{1}{2}, 1-t/s\big)\, 1_{[0,t)}(s),$$

where

$${}_2F_1(a,b,c,z) := \sum_{k=0}^{\infty} \frac{(a)_k(b)_k}{(c)_k} \frac{z^k}{k!}$$

denotes the Gauss hypergeometric function and $(a)_k := \Gamma(a+k)/\Gamma(a)$ is the Pochhammer symbol (see Decreusefond and Üstünel, 1999). Given a standard Brownian motion (B_t), we define a Gaussian process (W_t^H) for $t \in [0,1]$ by setting

$$W_t^H = \int_0^t K_H(t,s)\,\mathrm{d}B_s. \tag{50}$$

Putting $R_H(s,t) := \int_0^1 K_H(s,r)K_H(t,r)\,\mathrm{d}r$, we have $R_H(s,t) = \mathbb{E}\{W_t^H W_s^H\}$ and it can be checked that (W_t^H) has covariance (8).

We denote by $\mathbf{W}^H$ the law of $(W_t^H)_{t\in[0,1]}$ under which the canonical process on the Wiener space $\mathscr{C}_0$ is a fractional Brownian motion with Hurst parameter H. The Cameron–Martin space of $\mathbf{W}^H$ is defined as

$$\mathscr{H}_H := \{f \in \mathscr{C}_0 : f = K_H g,\ g \in L^2([0,1])\},$$

where $K_H f(t) := \int_{[0,1]} K_H(t,s) f(s)\,\mathrm{d}s$. The space $\mathscr{H}_H$ is equipped with the norm

$$\|f\|_{\mathscr{H}_H} := \|g\|_2 \text{ if } f = K_H g.$$

6.1.1. fBM and function spaces

We shall need Besov spaces on the interval $B^H_{2,p}([0,1])$, for $p \in \{2,\infty\}$, via their characterization in terms of wavelet sequences on the interval, see Section A.3.1.

Lemma 7. *For all* $H \in [\frac{1}{2}, 1)$

$$\|f\|_{\mathscr{H}_H} \leqslant c\|f\|_{B^{H+1/2}_{2,2}} \quad \text{and} \quad \mathbb{E}\left\{\|W^H\|^2_{B^H_{2,\infty}}\right\} \leqslant c.$$

Proof. The second estimate is part of Theorem IV.3 in Ciesielski et al. (1993). The first estimate follows from classical fractional calculus (recommended reference is Samko et al. (1993, p. 186). It is based on the representation

$$K_H^{-1} = (x^{H-1/2}\mathscr{D}^{H-1/2}x^{1/2-H}\mathscr{I}^{H-1/2})\mathscr{D}^{H+1/2}, \tag{51}$$

where $\mathscr{D}^\alpha$, $\alpha > 0$ is the inverse of the integral operator $\mathscr{I}^\alpha f(t) := \int_0^t f(s)(t-s)^{\alpha-1}\,\mathrm{d}t$ and x^α, $\alpha \in \mathbb{R}$ is the multiplication $[x^\alpha f](t) := t^\alpha f(t)$. By Lemma 3.2 of Samko et al. (1993), since for $H \in [\frac{1}{2}, 1)$, we have $\frac{1}{2} - H > -1$, the operator $x^{1/2-H}$ and $\mathscr{I}^{H-1/2}$ intertwine

$$x^{1/2-H}\mathscr{I}^{H-1/2}f(t) = \mathscr{I}^{H-1/2}x^{1/2-H}[f + Af](t), \quad t \in [0,1] \tag{52}$$

for a certain operator A bounded in $L^2([0,1])$. It follows that if f is such that $\mathscr{D}^{H+1/2} f \in L^2([0,1])$, we have, by (51) and (52):

$$K_H^{-1}f(t) = \mathscr{D}^{H+1/2}f(t) + A\mathscr{D}^{H+1/2}f(t)$$

We derive $\|K_H^{-1}f\|_2 \leqslant c\|\mathscr{D}^{H+1/2}f\|_2 \leqslant c\|f\|_{B^{H+1/2}_{2,2}}$. □

6.1.2. Approximation results for statistical distances

We denote by $\mathbb{D}(\mathbb{P},\mathbb{Q})$ the Kullback–Leibler divergence $\int \log \mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}\,\mathrm{d}\mathbb{P} \leqslant +\infty$ of two probability measures $\mathbb{P}$ and $\mathbb{Q}$.

Lemma 8. *Let* $T(\omega) := \omega + \rho(\omega)$ *be such that the process* $t \rightsquigarrow \rho[\omega](t)$ *is adapted w.r.t. the canonical filtration and differentiable. Assume moreover that* $\mathbf{W}^H \sim T\mathbf{W}^H$ ($\sim$ *denotes equivalence between measures.*) *Then*

$$\mathbb{D}(\mathbf{W}^H, T\mathbf{W}^H) = \tfrac{1}{2}\,\mathbb{E}_{\mathbf{W}^H}\left\{\|\rho(\omega)\|^2_{\mathscr{H}_H}\right\}.$$

Proof. By Girsanov theorem (see for instance Decreusefond and Üstünel, 1999, Theorems 4.8 and 4.9), the density $\mathrm{d}T\mathbf{W}^H/\mathrm{d}\mathbf{W}^H$ can be represented as

$$\frac{\mathrm{d}T\mathbf{W}^H}{\mathrm{d}\mathbf{W}^H} = \exp\left\{-\int_0^1 K_H^{-1}\rho[\omega](s)\,\mathrm{d}B_s - \frac{1}{2}\int_0^1 (K_H^{-1}\rho[\omega](s))^2\mathrm{d}\,s\right\}, \tag{53}$$

where B is a $(\mathbf{W}^H, \mathscr{F})$ standard Brownian motion. Since

$$\mathbb{E}_{\mathbf{W}^H}\left\{-\int_0^1 K_H^{-1}\rho[\omega](s)\,\mathrm{d}B_s\right\} = 0,$$

the fact that $\mathbb{D}(\mathbf{W}^H, T\mathbf{W}^H) = -\mathbb{E}_{\mathbf{W}^H}\left(\log \mathrm{d}T\mathbf{W}^H / \mathrm{d}\mathbf{W}^H\right)$, we obtain by (53) that

$$\mathbb{D}(\mathbf{W}^H, T\mathbf{W}^H) = \frac{1}{2}\,\mathbb{E}_{\mathbf{W}^H}\left\{\int_0^1 (K_H^{-1}\rho[\omega](s))^2\,\mathrm{d}s\right\} = \frac{1}{2}\mathbb{E}_{\mathbf{W}^H}\{\|\rho(\omega)\|^2_{\mathscr{H}_H}\},$$

which is the desired result. □

Lemma 9. *For all* $H \in [\frac{1}{2}, 1)$, *there exists a linear approximation operator* $P_N : \mathscr{C}_0 \to \mathscr{C}_0$ *that satisfies the following properties*:

(i) (*Causality*) *The random process* $t \rightsquigarrow P_N[\omega](t)$ *is adapted.*
(ii) (*Jackson inequality*)

$$\|P_N(\omega) - \omega\|_2 \leqslant c\|\omega\|_{B^H_{2,\infty}} N^{-H}.$$

(iii) (*Stability in* $B^{H+1/2}_{2,2}$) *For all* $N \geqslant 1$, $P_N(\omega) \in \mathscr{H}_H$ *and*

$$\|P_N(\omega)\|_{B^{H+1/2}_{2,2}} \leqslant c\sqrt{N}\|\omega\|_{B^H_{2,\infty}}.$$

For $f \in \mathscr{C}_0$ and $y_0 \in \mathbb{R}$, let $\mathbb{Q}^n_{f,y_0}$ denote the law of the n-dimensional random vector $X^n = (X_{i/n}, i = 1, \ldots, n)$, with

$$X_t = y_0 + \int_0^t f(s)\,\mathrm{d}W_s \tag{54}$$

and where W is the Brownian motion defined in the previous section. The following estimate can be found in Hoffmann (2002).

Lemma 10. *Let* f_0, $f_1 \in \mathscr{C}_0$ *s.t.* $\inf_t f_0(t) \geqslant \mu > 0$, $\|f_1^2 - f_0^2\|_\infty \leqslant \frac{1}{5}\mu^2$ *and for some* $L > 0$

$$\|f_1^2 - f_0^2\|_2 \leqslant \frac{\mu^2 L}{\sqrt{n}}.$$

Let $\lambda > 0$. *For* $n \geqslant 1$, *we have* $\mathbb{Q}^n_{f_1,y_0}\{\mathrm{d}\mathbb{Q}^n_{f_0,y_0}/\mathrm{d}\mathbb{Q}^n_{f_1,y_0} \geqslant e^{-\lambda}\} \geqslant 1 - (L/\lambda)(L/3 + \sqrt{3}/2)$.

Let $T : E \to E$ be measurable map on a probability space $(E, \mathscr{E}, \mu)$ such that $T^{-1}(E) = E$. Let F_1 and F_2 be two positive functions on E.

Lemma 11. *Assume that* $F_1 + F_2 \circ T \geqslant \gamma > 0$ *on* $\mathscr{A} \in \mathscr{E}$, *that* $T\mu \sim \mu$ *and* $\mathbb{D}(\mu, T\mu) \leqslant \beta$. *Then*

$$\int_E (F_1 + F_2)\,\mathrm{d}\mu \geqslant \gamma\left(\sqrt{\mu(\mathscr{A})} - \left(\frac{1}{2}\beta\right)^{1/4}\right)^2.$$

Proof. First, since $T^{-1}(E) = E$, the integral $\int_E (F_1 + F_2)\,\mathrm{d}\mu$ is equal to

$$\int_E \left[F_1 + F_2 \circ T \cdot \frac{\mathrm{d}\mu}{\mathrm{d}T\mu}(T(\cdot))\right]\,\mathrm{d}\mu \geqslant \lambda \int_{\mathscr{A}} [F_1 + F_2 \circ T]\,1_{(\mathrm{d}\mu/\mathrm{d}T\mu)(T(\cdot)) \geqslant \lambda}\,\mathrm{d}\mu$$

for all $\lambda \in [0,1)$. Next, using that the sum of F_1 and $F_2 \circ T$ is bounded below by γ on $\mathscr{A}$ and denoting by $\mu_{\mathscr{A}}$ the probability measure $\mu(\mathscr{A})^{-1} \int_{\mathscr{A}} \cdot \, \mathrm{d}\mu$, the RHS of the last inequality is bounded below by

$$\gamma\lambda\mu(\mathscr{A}) \left[1 - \mu_{\mathscr{A}} \left\{ \frac{\mathrm{d}\mu}{\mathrm{d}T\mu}(T(\cdot)) < \lambda \right\} \right].$$

By Chebyshev's inequality, we further bound this last term from below by

$$\gamma\lambda \left[\mu(\mathscr{A}) - (1-\lambda)^{-1} \int_E \left| \frac{\mathrm{d}\mu}{\mathrm{d}T\mu} - 1 \right| \mathrm{d}T\mu \right] = \gamma\lambda \left[\mu(\mathscr{A}) - \frac{\|\mu - T\mu\|_{TV}}{1-\lambda} \right],$$

where $\|\cdot\|_{TV}$ denotes the variational distance. The conclusion follows by applying the transport inequality $\|\mu - T\mu\|_{TV}^2 \leqslant \frac{1}{2}\,\mathbb{D}(\mu, T\mu) \leqslant \frac{1}{2}\beta$ and by taking the supremum in $\lambda \in [0,1)$. □

6.2. Proof of Theorem 2

Let $\varphi_n := n^{-1/(4H+2)}$. Pick an open set U in Θ and let

$$\theta_0 \in U, \quad \theta_n := \theta_0 + \alpha\varphi_n \in U$$

for $n \geqslant 1$ and a sufficiently small constant $\alpha > 0$. Using $\max' \geqslant \sum/2$, for any estimator F and any $C > 0$, we bound the maximal risk

$$\sup_{\theta \in U} \mathbb{P}_\theta^n\{\varphi_n^{-1}|F - \theta| \geqslant C\}$$

from below by

$$\frac{1}{2}\left(\mathbb{P}_{\theta_0}^n \left\{ \varphi_n^{-1}|F-\theta_0| \geqslant C \right\} + \mathbb{P}_{\theta_n}^n \left\{ \varphi_n^{-1}|F-\theta_n| \geqslant C \right\} \right)$$

$$= \int_{\mathscr{C}_0} \mathbf{W}^H(\mathrm{d}\omega)(\mathbb{P}_{\theta_0,\omega}^n\{\varphi_n^{-1}|F-\theta_0| \geqslant C\} + \mathbb{P}_{\theta_n,\omega}^n\{\varphi_n^{-1}|F-\theta_n| \geqslant C\}), \quad (55)$$

where $\mathbb{P}_{\theta,\omega}^n(\cdot)$ denotes the law of $(Y_{i/n},\ i=1,\dots,n)$, conditional on $W^H = \omega$. We plan to use Lemma 11 to bound (55) from below.

6.2.1. Control of the conditional perturbation

In order to apply Lemma 11, we first need a lower bound for the term under the integral in (55) when we formally replace $\mathbb{P}_{\theta_n,\omega}^n$ by $\mathbb{P}_{\theta_n,T_n(\omega)}^n$ for an appropriate perturbation T_n.

By a change of probability argument and the triangle inequality, it is classical to show that, for large enough n and for all $\lambda > 0$

$$\mathbb{P}_{\theta_{0,\omega}}^n \left\{ \varphi_n^{-1}|F-\theta_0| \geqslant C \right\} + \mathbb{P}_{\theta_n,T_n(\omega)}^n \left\{ \varphi_n^{-1}|F-\theta_n| \geqslant C \right\}$$

$$\geqslant \mathrm{e}^{-\lambda}\mathbb{P}_{\theta_{0,\omega}}^n \left\{ \frac{\mathrm{d}\mathbb{P}_{\theta_n,T(\omega)}^n}{\mathrm{d}\mathbb{P}_{\theta_{0,\omega}}^n} \geqslant \mathrm{e}^{-\lambda} \right\}.$$

We first plan to apply Lemma 10 with

$$\mathbb{P}_{\theta,\omega}^n = \mathbb{Q}_{\Phi(\theta,\omega),y_0}^n.$$

Since Θ is compact, the term $|\Phi(\theta_0,\omega_t)^2-\Phi(\theta_n,T_n(\omega)_t)^2|$ is less than

$$|\Phi(\theta_0,\omega_t)-\Phi(\theta_n,T_n(\omega)_t)|c(\|\omega\|_\infty,\|T_n(\omega)\|_\infty). \tag{56}$$

We therefore need to bound (56) in L^2 norm. By Taylor's approximation, the fundamental term

$$\Phi(\theta_n,T_n(\omega)_t)-\Phi(\theta_0,\omega_t)$$

can be written as

$$\begin{aligned}&(\theta_n-\theta_0)\partial_\theta\Phi(\theta_0,\omega_t)+(T_n(\omega)_t-\omega_t)\partial_x\Phi(\theta_0,\omega_t)\\&\quad+(\theta_n-\theta_0)(T_n(\omega)_t-\omega_t)\partial^2_{\theta,x}\Phi(\theta_0,\omega_t)+\frac{1}{2}(\theta_n-\theta_0)^2\partial^2_\theta\Phi(\theta_0,\omega_t)\\&\quad+\frac{1}{2}(T_n(\omega)_t-\omega_t)^2\partial^2_x\Phi(\theta_0,\omega_t)+r_n[\omega](t),\end{aligned} \tag{57}$$

where the remainder term $r_n[\omega](t)$ satisfies

$$\max\{(\theta_n-\theta_0),\|T_n(\omega)-\omega\|_\infty\}^{-3}\|r_n[\omega]\|_\infty\leqslant c(\|\omega\|_\infty). \tag{58}$$

However, in order to apply Lemma 10, we need that first- and second-order terms cancel in the above expansion. Indeed, $\theta_n-\theta_0=\alpha\varphi_n$ is $O(n^{-1/(4H+2)})$ with $H\in[\frac{1}{2},1)$, so the rate $n^{-1/2}$ required in the assumptions of Lemma 10 is obtained for third-order terms only. Define

$$G(x):=\frac{\partial_\theta\Phi(\theta_0,x)}{\partial_x\Phi(\theta_0,x)}$$

and

$$\delta(x):=\partial_x\Phi(\theta_0,x)^{-1}[G(x)(2\partial^2_{\theta,x}\Phi(\theta_0,x)-G(x)\partial^2_x\Phi(\theta_0,x))-\partial^2_\theta\Phi(\theta_0,x))].$$

Straightforward computations show that

$$T_n^{(0)}(\omega)_t:=\omega_t-\alpha\varphi_nG(\omega_t)+\frac{\alpha^2\varphi_n^2}{2}\delta(\omega_t)$$

satisfies

$$\Phi(\theta_0,\omega_t)-\Phi(\theta_n,T_n(\omega)_t)=\tilde{r}_n[\omega](t), \tag{59}$$

up to a remainder term $\tilde{r}_n[\omega](t)$ which also has property (58). However, the transform $T_n^{(0)}$ is such that $\mathbf{W}^H$ and $T_n^{(0)}\mathbf{W}^H$ are mutually orthogonal, therefore $\mathbb{D}(\mathbf{W}^H,T_n^{(0)}\mathbf{W}^H)=\infty$ and we will not be able to apply Lemma 1 in a second part. Therefore, we consider instead a low frequency approximation of the difference $T_n^{(0)}(\omega)-\omega$ which still encompasses the approximation requirements of Lemma 10. Set $N_n:=n^{1/(2H+1)}$ and let

$$\omega\rightsquigarrow(t\rightsquigarrow P_{N_n}[\omega](t))$$

be a nonanticipative approximation operator satisfying the properties of Lemma 9. Eventually, we set

$$T_n(\omega):=\omega-\alpha\varphi_nP_{N_n}[G(\omega)]+\frac{\alpha^2\varphi_n^2}{2}P_{N_n}[\delta(\omega)]. \tag{60}$$

By Assumptions A and B, the mapping $G(\cdot)$ and $\delta(\cdot)$ are regular enough so that the following approximation properties hold:

$$\|G(\omega)\|_\infty \vee \|\delta(\omega)\|_\infty \leqslant \mathscr{P}_1(\|\omega\|_\infty),$$

$$\|G(\omega)\|_{B^H_{2,\infty}} \vee \|\delta(\omega)\|_{B^H_{2,\infty}} \leqslant \mathscr{P}_2(\|\omega\|_\infty, \|\omega\|_{B^H_{2,\infty}}),$$

where $\mathscr{P}_i$, $i=1,2$, are polynomials.

We see that $\|T_n^{(0)}(\omega) - T_n(\omega)\|_2$, is less than a constant times

$$[\|G(\omega) - P_{N_n}[G(\omega)]\|_2 \varphi_n + \|\delta(\omega) - P_{N_n}[\delta(\omega)]\|_2 \varphi_n^2],$$

which, in turn is less than

$$c[\|G(\omega)\|_{B^H_{2,\infty}} + \|\delta(\omega)\|_{B^H_{2,\infty}}]\varphi_n N_n^{-H} = c[\|G(\omega)\|_{B^H_{2,\infty}} + \|\delta(\omega)\|_{B^H_{2,\infty}}] \cdot n^{-1/2},$$

by Lemma 9 together with the fact that both $\delta(\omega)$ and $G(\omega)$ belong to the space $B^H_{2,\infty}$. It follows that

$$\|\Phi(\theta_0,\omega) - \Phi(\theta_n, T_n[\omega])\|_2 \leqslant c[\|G(\omega)\|_{B^H_{2,\infty}} + \|\delta(\omega)\|_{B^H_{2,\infty}}]n^{-1/2} + \|r_n[\omega]\|_\infty,$$

since $T_n^{(0)}$ solves (59) up to a remainder term of the right order. Back to (56), we obtain

$$\begin{aligned}\|\Phi(\theta_0,\,\omega)^2 - \Phi(\theta_n, T_n[\omega])^2\|_2 &\leqslant [cn^{-1/2} + \|r_n[\omega]\|_\infty]c(\|\omega\|_\infty, \|\omega\|_{B^H_{2,\infty}})\\ &\leqslant c_\star(\|\omega\|_\infty, \|\omega\|_{B^H_{2,\infty}}) \cdot n^{-1/2},\end{aligned}$$

where we used (58) in the last inequality. The constant $c_\star$ can be assumed to be increasing in its arguments. By Assumption D, there exists $m_1 > 0$ such that $\Phi(\theta_n, T_n[\omega]_t)$ is greater than

$$c_{\star\star}(\|\omega\|_\infty) > 0 \quad \text{on } \|\omega\|_\infty \leqslant m_1,$$

and, with no loss of generality, we may assume that $c_{\star\star}$ is decreasing. (if $\Phi(\theta,x)$ is positive; otherwise, we have an analogous inequality with <0 in place of >0 and we apply the same subsequent arguments with obvious changes.)

Let $m_2 > 0$. Define

$$\mathscr{A}(m_1,m_2) := \{\omega \in \mathscr{C}_0 : \|\omega\|_\infty \leqslant m_1 \text{ and } \|\omega\|_{B^H_{2,\infty}} \leqslant m_2\}.$$

We are now ready to apply Lemma 10, for $\omega \in \mathscr{A}(m_1,m_2)$. We take

$$f_0 := \Phi(\theta_n, T_n[\omega]), \quad f_1 := \Phi(\theta_0,\omega),$$

$$\mu := c_{\star\star}(m_1), \quad L := \mu^{-2} c_\star(m_1,m_2) = c_{\star\star}(m_1)^{-2} c_\star(m_1,m_2)$$

and we check that, for large enough n, we have that $\|f_0^2 - f_1^2\|_\infty \leqslant \frac{1}{5}\mu^2$. We derive

$$\begin{aligned}&\mathbb{P}^n_{\theta_0,\omega}\{\varphi_n^{-1}|F-\theta_0| \geqslant C\} + \mathbb{P}^n_{\theta_n,T_n(\omega)}\{\varphi_n^{-1}|F-\theta_n| \geqslant C\}\\ &\quad\geqslant \sup_{\lambda>0} \mathrm{e}^{-\lambda}\left(1-\lambda^{-1}L)(L/3+\sqrt{3}/2)\right) =: \kappa(m_1,m_2) \quad \text{say}.\end{aligned}$$

6.2.2. Control of the separation rate between $\mathbf{W}^H$ and $T_n\mathbf{W}^H$

A direct application of Lemma 8 states

$$\mathbb{D}(\mathbf{W}^H, T_n\mathbf{W}^H) = \tfrac{1}{2}\mathbb{E}_{\mathbf{W}^H}\left\{\|\omega - T_n(\omega)\|^2_{\mathscr{H}^H}\right\}.$$

Using Lemma 7, we have that $\mathbb{D}(\mathbf{W}^H, T_n\mathbf{W}^H)$ is less than a constant times

$$\varphi_n^2(\mathbb{E}_{\mathbf{W}^H}\{\|P_{N_n}(G(\omega))\|^2_{B^{H+1/2}_{2,2}}\} + \mathbb{E}_{\mathbf{W}^H}\{\|P_{N_n}(\delta(\omega))\|^2_{B^{H+1/2}_{2,2}}\}).$$

Applying Lemma 9(iii) this last quantity is less than

$$\varphi_n^2 N_n(\mathbb{E}_{\mathbf{W}^H}\{\|G(\omega)\|^2_{B^H_{2,\infty}}\} + \mathbb{E}_{\mathbf{W}^H}\{\|\delta(\omega)\|^2_{B^H_{2,\infty}}\}).$$

Recalling that $\varphi_n^2 N_n$ remains bounded, we eventually obtain the existence of a constant c_{Kullback} such that

$$\mathbb{D}(\mathbf{W}^H, T_n\mathbf{W}^H) \leqslant c_{\text{Kullback}}.$$

Remark. Note that we can change φ_n into $\sqrt{\tau}\varphi_n$ for any dilation factor $\tau > 0$ without affecting the order of the previous bounds, so c_{Kullback} can be changed into τc_{Kullback} for any $\tau > 0$.

6.2.3. Completion of proof

We are now ready to apply Lemma 11. We take

$$(E, \mu) := (\mathscr{C}_0, \mathbf{W}^H), \quad T := T_n,$$

$$F_1(\omega) := \mathbb{P}^n_{\theta_0,\omega}\{\varphi_n^{-1}|F - \theta_0| \geqslant C\}, \quad F_2(\omega) := \mathbb{P}^n_{\theta_n,\omega}\{\varphi_n^{-1}|F - \theta_n| \geqslant C\},$$

$$\mathscr{A} := \mathscr{A}(m_1, m_2), \quad \gamma := \kappa(m_1, m_2), \quad \beta := c_{\text{Kullback}}.$$

For all $m_1 > 0$ we have $\mathbf{W}^H\{\|\omega\|_\infty \leqslant m_1\} > 0$. So, by Lemma 7, for large enough m_2, we have $\mathbf{W}^H\{\mathscr{A}(m_1, m_2)\} > 0$. Eventually

$$\int_{\mathscr{C}_0} \mathbf{W}^H(\mathrm{d}\omega)(\mathbb{P}^n_{\theta_0,\omega}\{\varphi_n^{-2}(F - \theta_0)^2 \geqslant C\} + \mathbb{P}^n_{\theta_n,\omega}\{\varphi_n^{-2}(F - \theta_n)^2 \geqslant C\})$$

is bigger than

$$\kappa(m_1, m_2)[\mathbf{W}^H(\mathscr{A}(m_1, m_2))^{1/2} - (\tfrac{1}{2}\beta)^{1/4}]^2$$

and the conclusion follows by taking β sufficiently small, a choice that is always possible by the last remark of Section 6.2.2.

Appendix A.

A.3. Wavelets and function spaces

For $f \in \mathscr{C}_0$, we set, as usual.

A.3.1. Biorthogonal wavelets and Besov spaces

A (compactly supported) wavelet basis consists of a (compactly supported) scaling function $\psi_{-1,0}$ and (compactly supported) wavelet functions $\psi_\lambda, \tilde{\psi}_\lambda$ that satisfy a biorthogonality relation

$$\langle \psi_\lambda, \tilde{\psi}_{\lambda'} \rangle = \delta_{\lambda,\lambda'}.$$

The duality bracket $\langle \cdot, \cdot \rangle$ reduces to the usual inner product as soon as ψ_λ and $\tilde{\psi}_{\lambda'}$ are regular enough. The index λ concatenates the scale and space parameters j and k, thus $\psi_\lambda = \psi_{jk} = 2^{j/2}\psi(2^j \cdot -k)$ is obtained from a single function ψ. Likewise for $\tilde{\psi}_\lambda$. The wavelet decomposition of a function f takes then the form

$$f = \sum_{j=-1}^{\infty} \sum_{|\lambda|=j} f_\lambda \psi_\lambda,$$

where $f_\lambda = \langle f, \tilde{\psi}_\lambda \rangle$ and $f_{-1} = \langle f, \psi_{-1,0} \rangle$. For a sequence $(a_n)_{n \geqslant 0}$ of real numbers, we set as usual

$$\|(a_n)_{n\geqslant 0}\|_{\ell_\infty} := \sup_{n \geqslant 0} |a_n| \quad \text{and}$$

$$\|(a_n)_{n\geqslant 0}\|_{\ell_p} := \left(\sum_{n \geqslant 0} |a_n|^p \right)^{1/p} \quad \text{for } 1 \leqslant p < \infty.$$

Definition A.1. For a properly chosen wavelet basis (see Cohen, 2000) and $s > 0$ and $q \in \{2, \infty\}$, the function $f \in \mathscr{C}_0$ belongs to the Besov space $B^s_{2,q}([0,1])$ if

$$\|f\|_{B^s_{2,q}} := \|(2^{s|\lambda|} \|(f_\lambda)_{|\lambda|=j}\|_{\ell_2})_{j \geqslant 0}\|_{\ell_q} < \infty.$$

The Besov space $B^s_{2,2}([0,1])$ is the usual Sobolev space for non integer s.

A.1.2. The Schauder basis

Let $\psi_0 = 1_{[0,1]}$ and for $j \geqslant 0, 1 \leqslant k \leqslant 2^j$:

$$\psi_{j,k} = 2^{j/2}(1_{[(2k-2)/2^{j+1},(2k-1)/2^{j+1}]} - 1_{[(2k-1)/2^{j+1},2k/2^{j+1}]}),$$

i.e. the Haar basis. The Schauder basis is the collection of functions

$$\phi_0(t) = 1, \quad \phi_1(t) = t, \quad \phi_{j,k}(t) = 2^j \int_0^t \psi_{j,k}(s)\,\mathrm{d}s \quad \text{for } j \geqslant 0,\ 1 \leqslant k \leqslant 2^j.$$

Remark that $\phi_{j,k}$ has support in $[(k-1)/2^j, k/2^j]$ and that $\|\phi_{j,k}\|_\infty \leqslant 2^{j/2-1}$. We recall that any continuous function f on $[0,1]$, may be decomposed, with convergence in $\|\cdot\|_\infty$ norm:

$$f = f_0\phi_0 + f_1\phi_1 + \sum_{j=0}^{\infty}\sum_{k=1}^{2^j} f_{j,k}\phi_{j,k}$$

with explicit coefficients, $f_0 = f(0)$, $f_1 = f(1)$ and

$$f_{j,k} = 2^{-j/2+1}\left\{ f\left(\frac{2k-1}{2^{j+1}}\right) - \frac{1}{2}\left(f\left(\frac{2k}{2^{j+1}}\right) + f\left(\frac{2k-2}{2^{j+1}}\right)\right)\right\}. \tag{A.1}$$

The Schauder basis is compatible with the formalism of Section A.3.1 is we set

$$\psi_\lambda := \phi_{jk} \quad \text{and} \quad \tilde{\psi}_\lambda := 2^{-j/2+1}\left\{\delta_{(2k-1)/2^{j+1}} - \tfrac{1}{2}\left(\delta_{2k/2^{j+1}} + \delta_{(2k-2)/2^{j+1}}\right)\right\},$$

where δ_x denotes the Dirac mass at point x.

A.2. Proof of Lemma 1

We have

$$\begin{aligned}
&\mathbb{E}\{\Delta_a \bar{W}_i^H \Delta_a \bar{W}_j^H\}\\
&\quad = \mathbb{E}\left\{\sum_{k=0}^{p} a_k \left(\bar{W}_{i+k}^H \sum_{l=0}^{p} a_l \bar{W}_{j+l}^H\right)\right\}\\
&\quad = \dot{\mathbb{E}}\left\{\sum_{k=0}^{p} a_k \left[\bar{W}_{i+k}^H \sum_{l=0}^{p} a_l (\bar{W}_{j+l}^H - \bar{W}_{i+k}^H)\right]\right\} \quad \left(\text{using that } \sum_{l=0}^{p} a_l = 0\right)\\
&\quad = \mathbb{E}\left\{\sum_{l=0}^{p} a_l \left[\bar{W}_{j+l}^H \sum_{k=0}^{p} a_k (\bar{W}_{i+k}^H - \bar{W}_{j+l}^H)\right]\right\} \quad (\text{by symmetry})\\
&\quad = -\frac{1}{2}\mathbb{E}\left\{\sum_{0\leqslant k,l\leqslant p} a_k a_l (\bar{W}_{i+k}^H - \bar{W}_{j+l}^H)^2\right\}.
\end{aligned}$$

By (19), it suffices to show that for $i, i+k \in \{1,\dots,N-p-1\}$

$$\mathbb{E}\{(\bar{W}_{i+k}^H - \bar{W}_i^H)^2\} = N^{-2H} v_H(k).$$

It is sufficient to consider the case $k \geqslant 1$:

$$\begin{aligned}
&\mathbb{E}\{(\bar{W}_{i+k}^H - \bar{W}_i^H)^2\}\\
&\quad = N^2 \mathbb{E}\left\{\left(\int_{i/N}^{(i+1)/N} (W_{s+k/N}^H - W_s^H)\,\mathrm{d}s\right)^2\right\}\\
&\quad = 2N^2 \int_{i/N}^{(i+1)/N} \int_{i/N}^{s} \mathbb{E}\{(W_{s+k/N}^H - W_s^H)(W_{s'+k/N}^H - W_{s'}^H)\}\,\mathrm{d}s'\,\mathrm{d}s.
\end{aligned}$$

We now replace the expectation in the integral above using (9), and then easily derive —recall (18)—that

$$\mathbb{E}\{(\bar{W}^H_{i+k} - \bar{W}^H_i)^2\} = N^{-2H} v_H(k).$$

It remains to prove that $\kappa^2_{a,H} > 0$. Assume on the contrary, that $\Delta_a \bar{W}^H_i = 0$ for all i. Thus, $(\Delta_a \bar{W}^H_i)_i$ solves a linear equation of order p, thus $\bar{W}^H_i$ is a function of $(\bar{W}^H_j)_{0\leqslant j\leqslant p}$. Choosing $i = i_N$ such that i_N/N converges to an arbitrary $t_0 \in (\frac{1}{2}, 1]$, and since $p/N \leqslant \frac{1}{2}$ for N big enough, we derive that $W^H_{t_0}$ is measurable with respect to $(W^H_t)_{t\in[0,1/2]}$, a contradiction.

A.3. Proof of Lemma 9

Let Φ be $\mathscr{C}^\infty$ with compact support in $[0,1]$ and such that $\int \phi(u)\,\mathrm{d}u = 1$. For $N \geqslant 1$, let j_N be an integer such that $N \leqslant 2^{-j_N} \leqslant 2N$. We set

$$\Phi_N(t) := 2^{j_N} \Phi(2^{-j_N} t)$$

and

$$P_N[\omega](t) := \int_0^1 \omega_s \Phi_N(t-s)\,\mathrm{d}s.$$

We have property (i) by construction. Likewise, property (ii) follows from classical direct estimates by convolution kernels, see e.g. Cohen (2000). It remains to prove the stability property (iii). Expand ω in a wavelet basis $\omega = \sum_j \omega_j$, where $\omega_j = \sum_k \langle \omega, \psi_{jk}\rangle \psi_{jk}$ is the decomposition of ω at scale j. We plan to use the following low–high frequency decomposition

$$P_N[\omega](t) = G_{j_N} + g_{j_N} = \sum_{j\leqslant j_N} \Phi_N * \omega_j + \sum_{j>j_N} \Phi_N * \omega_j.$$

Let $\mathscr{F}$ denote the Fourier transform. Since $\xi \rightsquigarrow (\mathscr{F}\Phi_N)(\xi) = (\mathscr{F}\Phi)(2^{-j_N}\xi)$ is uniformly bounded in N, the Besov (or Sobolev) norm $B^{H+1/2}_{2,2}$ of $\Phi_N * \sum_{j\leqslant j_N} \omega_j$ is less than a constant times the Besov norm of $\sum_{j\leqslant j_N} \omega_j$. If the wavelet basis is sufficiently regular, we further have

$$\left\| \sum_{j\leqslant j_N} \omega_j \right\|_{B^{H+1/2}_{2,2}} \leqslant c \sum_{j\leqslant j_N} 2^{j(H+1/2)} \sum_k \langle \omega, \psi_{jk}\rangle^2,$$

by the characterization of Besov norms in terms of wavelet sequences. Applying Bernstein inequality

$$\sum_k \langle \omega, \psi_{jk}\rangle^2 \leqslant c 2^{-2jH} \|\omega\|_{B^H_{2,\infty}},$$

we obtain that $\|G_{j_N}\|^2_{B^{H+1/2}_{2,2}}$ is less than

$$c\,\|\omega\|^2_{B^H_{2,\infty}} \sum_{j\leqslant j_N} 2^{2j(H+1/2)} \sum_k 2^{-2jH} \leqslant c\,\|\omega\|^2_{B^H_{2,\infty}}\, 2^{j_N} \leqslant c\,\|\omega\|^2_{B^H_{2,\infty}}\, N.$$

We now turn to the high-frequency part. Using again that $(\mathscr{F}\Phi_N)$ is uniformly bounded and Bernstein inequality, we have

$$||g_{j_N}||_2^2 \leqslant c(\Phi) \left\| \sum_{j>j_0} \omega_j \right\|_2^2 \leqslant c2^{-2jH} ||\omega||^2_{B^H_{2,\infty}} .$$

We also have

$$||g_{j_N}||^2_{B^{H+1/2}_{2,2}} \leqslant c \left(||g_{j_N}||_2 + \int |\xi|^{2H+1} |\mathscr{F} g_{j_N}(\xi)|^2 \, \mathrm{d}\xi \right),$$

it remains to bound the last term in the RHS above, which is equal to

$$\int |\xi|^{2H+1} \, |\mathscr{F}\Phi_N(\xi)|^2 \, |\mathscr{F} \left(\sum_{j>j_N} \omega_j \right) (\xi)|^2 \, \mathrm{d}\xi. \tag{A.2}$$

But

$$\begin{aligned} |\xi|^{2H+1} |\mathscr{F}\Phi_N(\xi)|^2 &= |\xi|^{2H+1} |\mathscr{F}\Phi(2^{-j_N}\xi)|^2 \\ &= 2^{j_N(2H+1)} |2^{-j_N}\xi|^{2H+1} |\mathscr{F}\Phi(2^{-j_N}\xi)|^2 \leqslant c(\Phi) 2^{j_N(2H+1)}, \end{aligned}$$

where $c(\Phi) = \sup_\xi |\xi|^{2H+1} |\mathscr{F}\Phi(\xi)|^2$. So the integral in (A.2) is bounded by a constant times

$$2^{j_N(2H+1)} \int \left| \mathscr{F} \left(\sum_{j>j_N} \omega_j \right) (\xi) \right|^2 \mathrm{d}\xi = 2^{j_N(2H+1)} \left\| \sum_{j>j_N} \omega_j \right\|_2^2 .$$

Using the wavelet expansion and Bernstein inequality, this last term is less than

$$2^{j_N(2H+1)} \sum_{j>j_N} \sum_k \langle \omega, \psi_{jk} \rangle^2 \leqslant c 2^{j_N(2H+1)} \sum_{j>j_N} 2^{-2jH} \|\omega\|^2_{B^H_{2,\infty}}$$

and is thus of order $2^{j_N} \|\omega\|^2_{B^H_{2,\infty}}$. The conclusion follows.

References

Breidt, F.J., Crato, N., de Lima, P., 1998. The detection and estimation of long memory in stochastic volatility. J. Econom. 83, 325–348.

Carmona, P., Coutin, L., Montseny, G., 2003. Stochastic integration with respect to fractional Brownian motion. Ann. Inst. H. Poincaré Probab. Statist. 39, 27–68.

Ciesielski, Z., Kerkyacharian, G., Roynette, B., 1993. Quelques espaces fonctionnels associés à des processus gaussiens. Stud. Math. 107, 172–204.

Cohen, A., 2000. Wavelet Methods in Numerical Analysis. Handbook of Numerical Analysis. Vol. VII, North-Holland, Amsterdam, pp. 417–711.

Comte, F., Renault, E., 1996. Long memory continuous time models. J. Econom. 73, 101–150.

Comte, F., Renault, E., 1998. Long memory in continuous-time stochastic volatility models. Math. Finance 8, 291–323.

Dai, W., Heyde, C.C., 1996. Ito formula with respect to fractional Brownian motion and its applications. J. Appl. Math. Stochastic Anal. 914, 439–448.

Decreusefond, L., "Ustünel, A.S., 1999. Stochastic analysis of the fractional Brownian motion. Potential Anal. 10, 177–214.

Genon-Catalot, V., Jeantheau, T., Laredo, C., 2000a. Parameter estimation for discretely observed stochastic volatility models. Bernoulli 5, 855–872.

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

Gloter, A., 2000. Estimation du coefficient de diffusion de la volatilité d'un modl̀e à volatilité stochastique. C.R. Acad. Sci. Paris 330 (Sér. I), 243–248.

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.

Hoffmann, M., 2002. Rate of convergence for parametric estimation in a stochastic volatility model. Stochastic Proc. Appl. 97, 147–170.

Hull, J., White, A., 1988. An analysis of the bias in option pricing caused by a stochastic volatility. Adv. Futures Options Res. 3, 29–61.

Istas, J., Lang, G., 1997. Quadratic variations and estimation of the local Hölder index of a Gaussian process. Ann. Inst. H. Poincaré 33, 407–436.

Melino, A., Turnbull, S.M., 1990. Pricing foreign currency options with stochastic volatility. J. Econom. 45, 239–265.

Salopek, D.M., 1998. Tolerance to arbitrage. Stochastic Proc. Appl. 76, 217–230.

Samko, S.G., Kilbas, A.A., Marichev, O.I., 1993. Fractional Integrals and Derivatives. Gordon and Breach Science, London.

Sottinen, T., 2001. Fractional Brownian motion, random walks and binary market models. Finance Stochastics 5, 343–355.